Genome Medicine

RESEARCH

Open Access

Computational quantification and characterization of independently evolving cellular subpopulations within tumors is critical to inhibit anti-cancer therapy resistance

Heba Alkhatib[1†], Ariel M. Rubinstein[1†], Swetha Vasudevan[1], Efrat Flashner-Abramson[1], Shira Stefansky[1], Sangita Roy Chowdhury[1], Solomon Oguche[1], Tamar Peretz-Yablonsky[2], Avital Granit[2], Zvi Granot[3], Ittai Ben-Porath[3], Kim Sheva[4], Jon Feldman[2], Noa E. Cohen[5], Amichay Meirovitz[4*†] and Nataly Kravchenko-Balasha[1*†]

Abstract

Background: Drug resistance continues to be a major limiting factor across diverse anti-cancer therapies. Contributing to the complexity of this challenge is cancer plasticity, in which one cancer subtype switches to another in response to treatment, for example, triple-negative breast cancer (TNBC) to Her2-positive breast cancer. For optimal treatment outcomes, accurate tumor diagnosis and subsequent therapeutic decisions are vital. This study assessed a novel approach to characterize treatment-induced evolutionary changes of distinct tumor cell subpopulations to identify and therapeutically exploit anticancer drug resistance.

Methods: In this research, an information-theoretic single-cell quantification strategy was developed to provide a high-resolution and individualized assessment of tumor composition for a customized treatment approach.

Briefly, this single-cell quantification strategy computes cell barcodes based on at least 100,000 tumor cells from each experiment and reveals a cell-specific signaling signature (CSSS) composed of a set of ongoing processes in each cell.

Results: Using these CSSS-based barcodes, distinct subpopulations evolving within the tumor in response to an outside influence, like anticancer treatments, were revealed and mapped. Barcodes were further applied to assign targeted drug combinations to each individual tumor to optimize tumor response to therapy.

The strategy was validated using TNBC models and patient-derived tumors known to switch phenotypes in response to radiotherapy (RT).

†Heba Alkhatib and Ariel M. Rubinstein share equal contribution.

†Amichay Meirovitz and Nataly Kravchenko-Balasha are equal contributors.

*Correspondence: amichaym@gmail.com; natalyk@ekmd.huji.ac.il

[1] The institute of Biomedical and Oral Research, The Hebrew University of Jerusalem, 9103401 Jerusalem, Israel
[4] The Legacy Heritage Oncology Center & Dr. Larry Norton Institute, Soroka University Medical Center, Ben Gurion University of the Negev, Faculty of Medicine, 8410101 Beer Sheva, Israel
Full list of author information is available at the end of the article

© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Conclusions: We show that a barcode-guided targeted drug cocktail significantly enhances tumor response to RT and prevents regrowth of once-resistant tumors. The strategy presented herein shows promise in preventing cancer treatment resistance, with significant applicability in clinical use.

Keywords: Cancer resistance, Intra-tumor heterogeneity, Tumor plasticity, Information-theoretic single-cell analysis, Individualized targeted therapy, Radiation oncology, Triple-negative breast cancer

Background

Drug resistance and cancer cell plasticity are principal contributors to therapeutic failure [1, 2]. Discovering a strategy with the ability to transform the potential evolution of certain intra-tumor subpopulations within treated/irradiated tumors into a therapeutic advantage, is an unmet need in cancer research and clinical practice [2].

We propose a novel approach where cancer treatment can be designed based on the changes occurring in patient-specific intra-tumor subpopulations in response to radiotherapy (RT) or cytotoxic treatment. In this study, this approach was assessed using a triple-negative breast cancer (TNBC) model and patient-derived TNBC samples treated with RT.

Recent studies have shown that although being an established and effective anti-cancer treatment, radiation may promote anti-apoptotic and pro-proliferative responses that often result in tumor regrowth [3, 4]. This has initiated numerous attempts to characterize tumor molecular phenotypes expressed in response to RT to identify new potential drug targets and strategies for anti-cancer treatment enhancement [5–10].

TNBC is a clinically unique, aggressive, and highly heterogeneous subtype of breast cancer that does not express estrogen receptors, progesterone receptors, or human epidermal growth factor receptor-2 (Her2), and for which no targeted therapy exists. Chemotherapy (CT) and RT have remained the standardized treatment options for the past 20 years [11, 12]. While TNBC may be sensitive to RT initially, resistance often develops at later stages [12] due to significant intra-tumor cell heterogeneity [13, 14] and potential phenotypic switching due to cellular plasticity (e.g., from $Her2^{-}$ to $Her2^{+}$) [15].

This study proposes an information theoretic approach utilizing thermodynamic-based surprisal analysis (SA) [16] in single cells to elucidate TNBC cellular subpopulations at a high resolution. We quantify evolving subpopulations in response to RT through single-cell barcoding of ongoing processes in TNBC tumor cells. Thermodynamic-based information theory is implemented to identify ongoing processes within each cell. Tumors may be considered as homeostatically disturbed entities that have deviated from a balanced state due to various constraints (e.g., mutational stress and drug treatment) [16]. Each constraint creates a deviation in the expression levels of a subset of proteins in the tumor. In doing so, an ongoing (unbalanced) process in the tumor, consisting of the group of proteins that were altered by the constraint, becomes active. SA identifies the constraints operating in a system as well as the proteins affected by each constraint.

We have previously demonstrated that accurate identification of unbalanced processes in human cells using bulk SA can anticipate the effect of protein inhibitors on protein co-expression network structures [17]. SA of cell-cell signaling in brain tumors has also been shown to predict cellular spatial distributions and the direction of cell-cell movement [18]. Additionally, we have applied SA to large-scale proteomic datasets for the prediction of efficient patient-specific targeted combination therapies [19–21].

In this study, SA is extended to identify cell-specific signaling signatures (CSSSs), consisting of a unique set of ongoing processes that have emerged within the individual cell. Each CSSS is converted into a cell-specific barcode. An intra-tumor subpopulation is then defined to be a group of cells harboring the exact same CSSS-based barcode and these cells are expected to respond similarly to treatment.

The final result of such an analysis is a high-resolution intra-tumoral map of the different subpopulations within a tumor and the CSSS that characterizes every subpopulation. Such a robust and comprehensive map provides guidance on the accurate determination of drug combinations to effectively target evolving subpopulations within the tumor to bring about a potent therapeutic effect.

Methods

Computational methods

Single-cell computational data analysis

Surprisal analysis (SA) was applied on the single-cell level so that each cell could be plotted according to its molecular aberrations and network reorganization (Figs. 1 and 2).

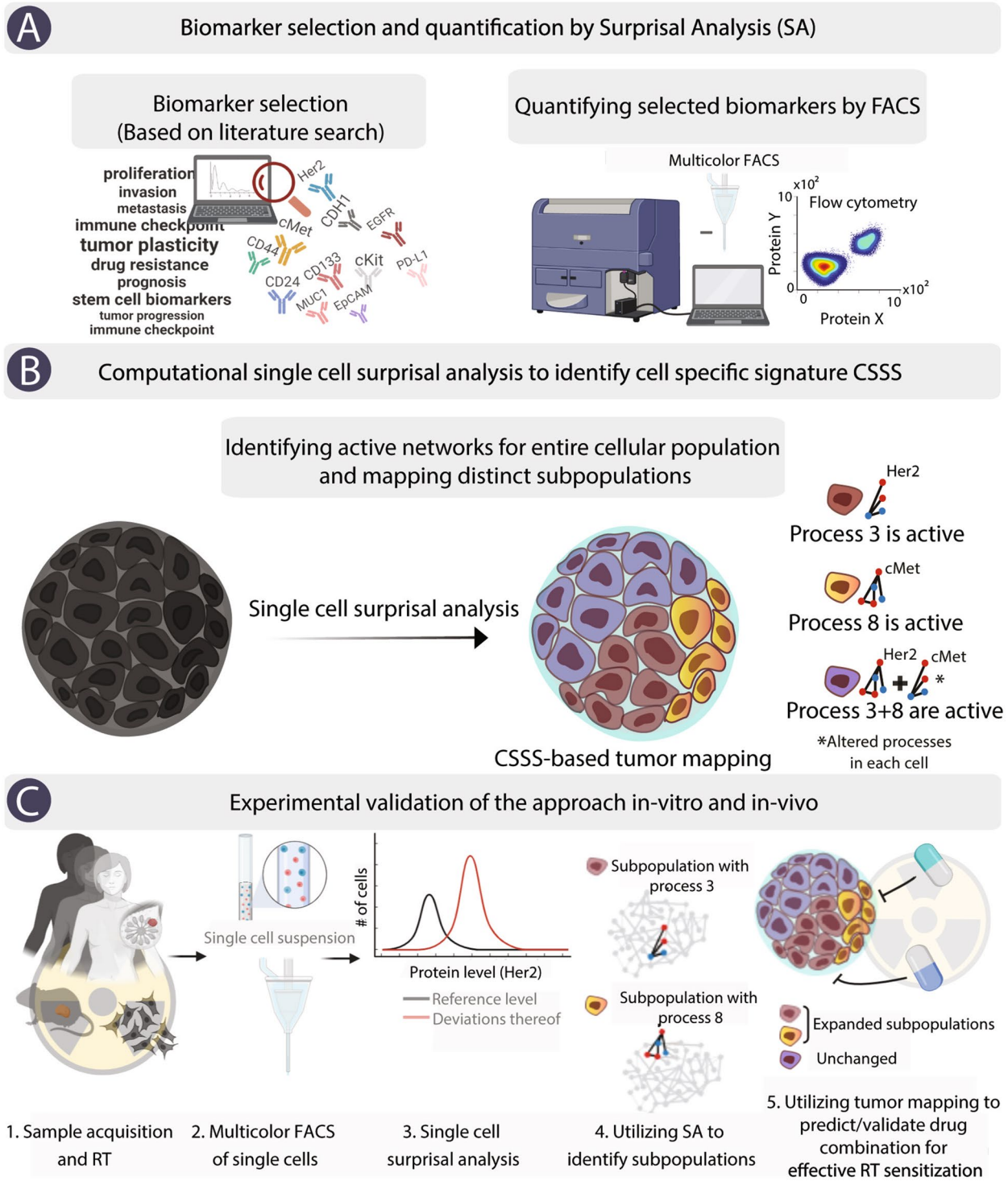

Fig. 1 Study scheme. **A** A literature search was used to compose a list of oncomarkers for single-cell quantification (left panel). Selected cell-surface oncoproteins were quantified using multicolor FACS (right panel). **B** SA was extended to single cells to identify distinct subpopulations based on *sets* of unbalanced processes (cell-specific signaling signatures, CSSS, right panel). **C** To validate the hypothesis that targeting evolving cellular subpopulations in response to RT would enhance TNBC RT sensitization, a series of in vitro and in vivo experiments, using TNBC patient-derived and mice models, were performed

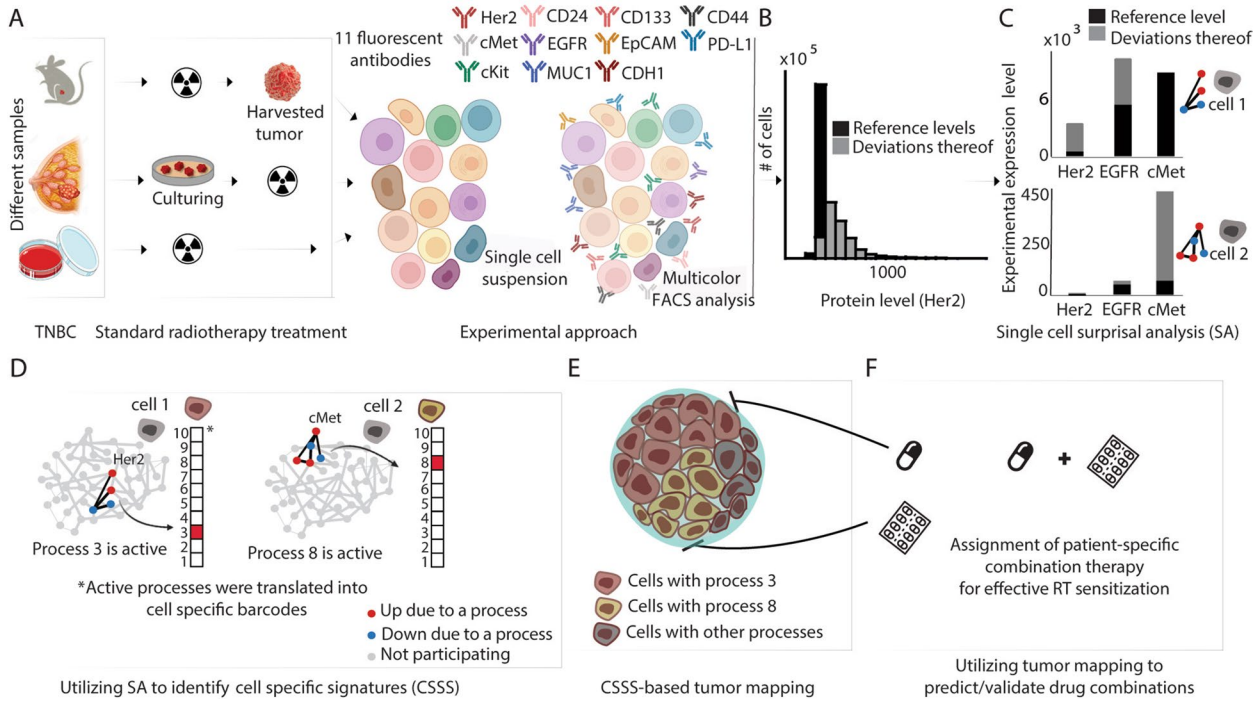

Fig. 2 Schematic of the application of the surprisal analysis algorithm. **A** Preparation of fluorescently-tagged single-cell suspensions from different sample sources (control and post-RT) for multicolor FACS analysis. Each cell was labeled with a mixture of 11 fluorescently tagged antibodies. **B** Surprisal analysis reveals protein expression level distributions at the reference (steady) state and the deviations thereof due to constraints in the system (e.g., irradiation). An example for calculated distribution of the expression levels at the reference state and deviations thereof is presented for Her2, initially quantified by FACS and analyzed by SA, in 4T1 mice model of TNBC. **C** Proteins deviating from the steady state in a coordinated manner are grouped into altered subnetworks referred to as "unbalanced processes." For example, in one 4T1 cell, the levels of Her2 and EGFR deviate significantly (upregulated) from the steady state and in the other cells, cMet levels deviate significantly (upregulated as well). Thus, the two cells are defined by the analysis as possessing different processes. **D** The unbalanced processes in each cell provide a cell-specific signaling signature (CSSS). Each CSSS is schematically transformed into a cell-specific barcode, indicating active and inactive processes. **E** Cells sharing the same barcode are organized into distinct subpopulations. **F** Tumor-specific targeted therapy combinations are tailored against the subpopulations expanding in response to RT

Surprisal analysis

SA is a thermodynamic-based information-theoretic approach [22–24] which has recently been implemented to analyze bulk [16, 19, 20, 25] and single-cell biological data [25, 26]. The analysis is based on the premise that biological systems reach a balanced state when the system is free of constraints [27–29]. However, when under the influence of environmental (e.g., exposure to a drug) and/or genomic constraints (genomic mutations that affect transcript/protein expression and function), the system is prevented from reaching a state of minimal free energy, and instead reaches a state of higher free energy—a constrained state [16].

Expression levels of various macromolecules, e.g., transcripts or proteins are used as an input for SA. Since the varying constraints that act upon living cells ultimately manifest as alterations in the cellular protein/gene expression network, they are viewed as emerging *unbalanced molecular processes* [19, 30]. Recent examples of SA implementation in biology include the characterization of bulk proteomic changes in large datasets, including multiple patient tissues and cancer cell lines, to predict a change in the behavior of systems [17, 18] or to design individualized drug therapies [19–21].

In this study, we analyzed protein expression data obtained from multicolor FACS in which each cell was labeled with a mixture of 11 fluorescently tagged antibodies. Additional file 1 provides the details for the models and single-cell analyses used in the study. It is important to note that this methodology may be applied to any single-cell proteomics data. The data matrix obtained from the flow cytometry analysis (~100,000–500,000 cells), in which columns are protein expression levels and rows are single cells, was used as an input for surprisal analysis. RT treatment imposes a

constraint, but more than one constraint may be identified in the system.

Equation 1 was used to identify different constraints within tumor cells:

$$\underbrace{X_i(cell)}_{\substack{\text{experimental}\\ \text{level of protein } i}} = \underbrace{X_i^o(cell)}_{\substack{\text{level of protein } i\\ \text{in the reference state}}} \exp \underbrace{\left(-\sum_{\alpha=1} G_{i\alpha}\lambda_\alpha(cell)\right)}_{\substack{\text{changes in protein levels}\\ \text{due to the constraints } \alpha = 1, 2, ..}} \quad (1)$$

Here, $X_i^o(cell)$ is the expected expression level of protein i at the reference state in a measured cell. The exponential term in Eq. (1) represents the deviation from the reference value due to the constraints, including those imposed by irradiation. $G_{i\alpha}$ are weights (the degree of participation) of a protein i in the unbalanced processes $\alpha = 1, 2...$ Proteins deviating in a similar manner from the steady state are grouped into unbalanced processes (see examples for 2 cells quantified in the 4T1 TNBC model in Fig. 2C). $\lambda_\alpha(cell)$ is the amplitude of an unbalanced process $\alpha = 1, 2. .$ in each tested cell. For example, Additional file 2: Fig. S2 presents $\lambda_\alpha(cell)$ values for process 8 (the network representing process 8 is shown in Fig. 3F and generated as explained below) in untreated and irradiated cells (6 days post-RT) in the 4T1 in-vitro model. *Negative/positive* amplitude denotes how the cells are correlated with respect to a particular process. The processes are indexed $\alpha = 1, 2, 3...$ so that the significance of the process decreases with an increasing index, i.e., unbalanced process 1 appears in a higher number of cells than unbalanced processes 2, 3, etc. *Several* unbalanced processes may be found in a system, however not all processes are active in all cells (see in the sections below how we define whether a particular process is active in a tested cell). $\sum_{\alpha=1} G_{i\alpha}\lambda_\alpha(cell)$ represents the amount of information we have about each protein i. The partial deviations in the expression level of protein i due to the different constraints add up to the total change in expression level. A protein that is influenced by constraints, i.e., is influenced by one or more unbalanced processes, cannot take on any possible expression level. Its expression level is affected by the expression levels of other proteins in the unbalanced process in the cell.

Calculation of $\lambda_\alpha(cell)$ and $G_{i\alpha}$ We fit the main Eq. (1), to the logarithm of the measured expression level of protein i in each cell using singular value decomposition (SVD). *In practical terms*, a matrix is constructed, containing the natural logarithm of protein expression levels in the different cells [26]. The procedure then utilizes SVD as an intermediate step, which calls for the construction of two square (and symmetric) co-variance matrices. One is smaller with a maximal rank of 11 (as the number of the proteins) and the second is bigger (depends on the number of cells, in this case at least 100,000). These matrices are diagonalized to calculate eigenvectors and eigenvalues. SVD and all other mathematical calculations described here were implemented using Matlab. Codes can be found in Github [31]. Eigenvectors and eigenvalues are further used by SA to calculate the amplitudes of the processes: $\lambda_\alpha(cell)$ for each cell if we use single-cell data, $\lambda_\alpha(k)$ for each sample k if we use bulk data [19, 20, 30], and $G_{i\alpha}$ values, which are weights of the proteins in each process α. A detailed, step-by-step description of the mathematical procedure, namely how the eigenvectors and eigenvalues are used to calculate the amplitudes ($\lambda_\alpha(cell)$ or $\lambda_\alpha(k)$) and $G_{i\alpha}$ values, is described in the Supplementary Information of reference [16]). Any additional information regarding the mathematical procedures/codes can be provided upon request.

The number of calculated constraints is limited by the smaller dimension of the input matrix. In this case, it was limited to 11 (the number of measured proteins) and therefore a maximum of 10 constraints or unbalanced processes (10 constraints plus steady state) could be found. Calculations of the parameters using a smaller matrix (detailed in [16]) allow for the fast and efficient data processing of hundreds of thousands of cells. The number of proteins quantified in each cell can be significantly expanded (hundreds or thousands) without significantly increasing the data processing time.

Generation of protein-protein networks representing unbalanced processes Additional file 2: Fig. S1 complements Fig. 3F and shows additional unbalanced processes active in the 4T1 system. The goal was to generate unbalanced processes composed of proteins with significant $G_{i\alpha}$ values. $G_{i\alpha}$ sign indicates the *correlation or anti-correlation* between proteins in the same process (Additional file 3). Upregulation or downregulation due to process α can be defined further using a product $G_{i\alpha}\lambda_\alpha(cell)$ for each protein in every *cell* in each experimental condition/time point: e.g., proteins with positive $\lambda_\alpha(cell)$ and positive $G_{i\alpha}$ will be upregulated due to process α, since the product $G_{i\alpha}\lambda_\alpha(cell)$, which represents a deviation from the steady state due to process α, will be positive. Protein-protein interactions in each unbalanced process are based on the STRING database [32]. The radius of each circle in the map corresponds to the $G_{i\alpha}$ value (Fig. 3F).

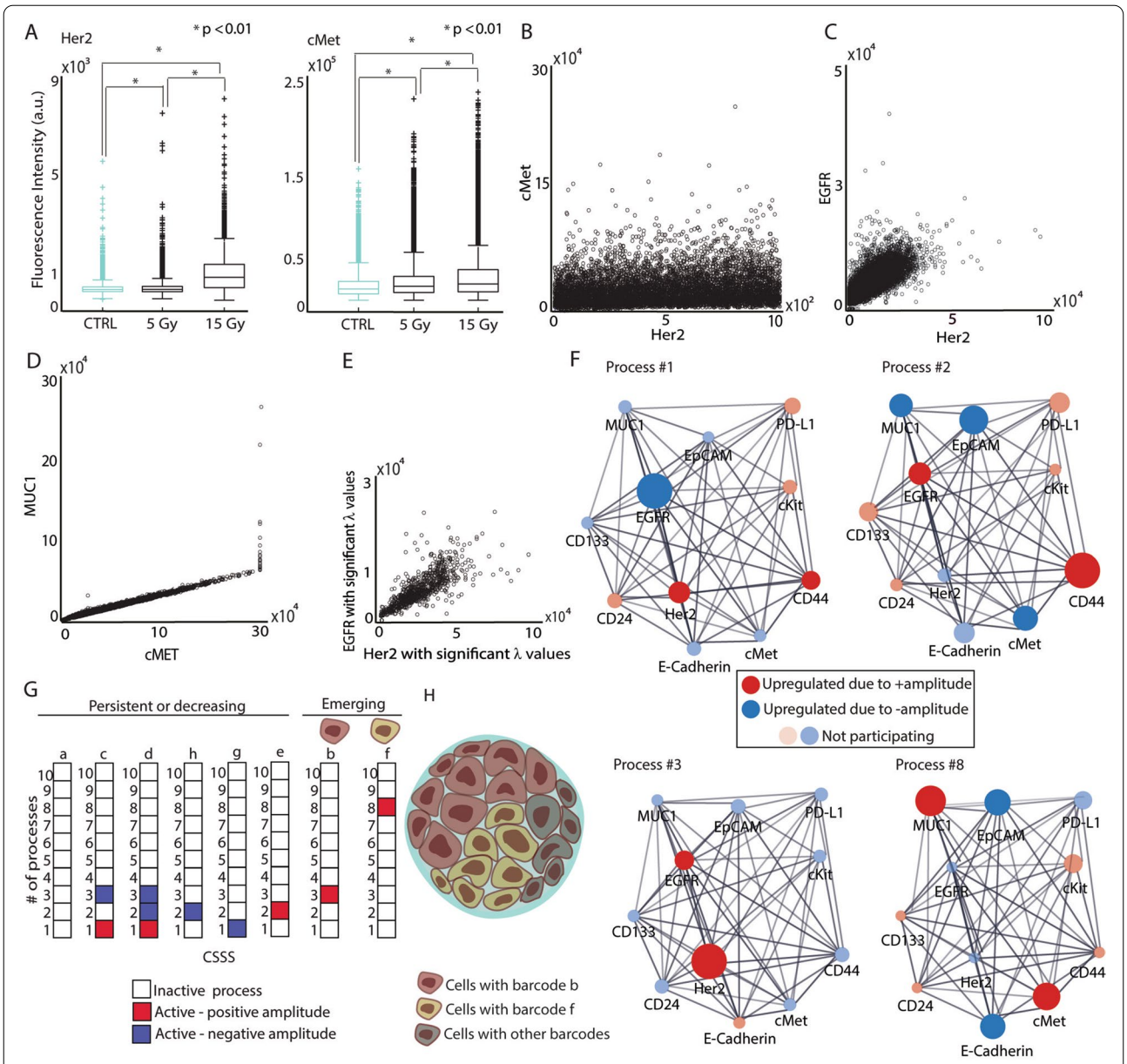

Fig. 3 Resolving expanded 4T1 cellular subpopulations post-RT. **A** FACS expression levels of Her2 and cMet following RT. **B–D** Correlation plots between Her2 and cMet (**B**), Her2 and EGFR (**C**), and MUC1 and cMet (**D**) in irradiated cells. **E** Correlation plot between Her2 and EGFR levels expressed in the cells found to harbor process 3 (only cells with significant amplitudes ($\lambda_3(cell)$ values) were included in this plot, also see the "Methods" section). **F** Protein–protein networks were generated using single-cell SA analysis and STRING to assign the functional connections. To determine the direction of change in every protein (i.e., upregulation or downregulation) a sign of the amplitude in a process *α* in each cell was considered. Four unbalanced subnetworks (processes) out of 10 resolved in 4T1 (Additional file 3) are shown. **G** Each cell was assigned a barcode representing a cell-specific signature (CSSS). The most abundant (>1%) subpopulations are presented. **H** Based on these CSSSs the tumor was divided into distinct subpopulations. Quantification of subpopulations was performed using at least ~30,000 cells from each condition, which were obtained from at least three flasks and from at least three independent experiments for each time point. For **A:** statistically significant differences between control and 5 Gy; control and 15 Gy; and 5 Gy and 15 Gy were determined using a two-tailed Student's *t* test (*$P < 0.01$)

Calculation of cell-specific barcodes based on CSSS It is important to note that not all processes are active in all cells. The term $\lambda_\alpha(cell)$ represents the importance of the unbalanced process *α* in the *cell*. Its sign indicates the correlation or anti-correlation between the same processes in different cells. To further map distinct subpopulations within the entire cellular population, we grouped cells sharing *the same set* of unbalanced processes, indicated by the cell-specific signaling signatures (CSSS), into distinct subpopulations (Figs. 1 and 2). Only unbalanced

processes with significant amplitudes were included in the CSSS of each individual cell as follows:

To determine threshold limits for $\lambda_\alpha(cell)$ values, $\lambda_\alpha(cell)$ values were sorted and plotted as sigmoid plots in each process. Only $\lambda_\alpha(cell)$ values located on the tails of the sorted distributions were considered and used further for the barcode calculations (Additional file 2: Fig. S2).

The combinations of unbalanced processes (CSSS) for each cell were generated using $\lambda_\alpha(cell)$ values that exceeded threshold limits. In this way, CSSSs were assigned to each cell and were converted into cell-specific barcodes for simple representation. Additional file 3-Additional file 8 include the input FACS data and the output parameters obtained using CSSS analysis ($\lambda_\alpha(cell)$ and $G_{i\alpha}$, and barcodes denoting subpopulations) for the major cancer systems, 4T1 and BR45, which have been tested in vitro and in vivo. Based on these cell-specific barcodes, distinct subpopulations were determined in the tumor (Fig. 3G, H). Subpopulations *b* and *f* (Fig. 4A) expanded significantly as detailed in the main text.

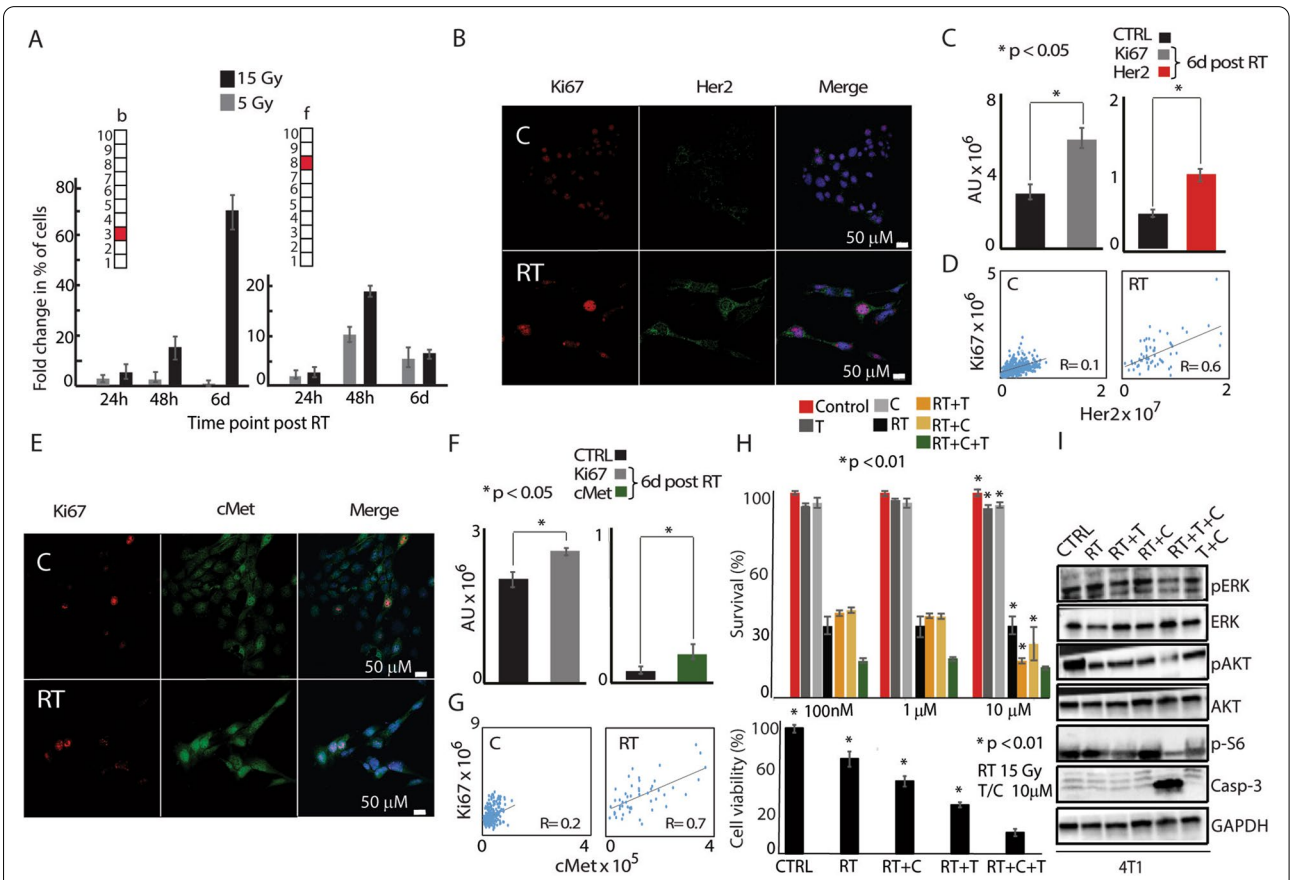

Fig. 4 Two distinct subpopulations expand and show proliferative properties in response to RT. **A** Very small subpopulations (<1%), represented by barcodes b and f, expanded significantly following RT (fold change in % of cells relative to the control of each time point). **B–G** 4T1 cells were irradiated with 15Gy. 6 days post-RT, cells were incubated with antibodies against Ki67, cMet, and Her2 and nuclei were stained with DAPI (fluorogel II). **B**, **E** 40× lens; scale bar represents 50 μm. **C**, **F** Sum intensities of Ki67 (**C**, left panel); Her2 (**C**, right panel); Ki67 (**F**, left panel); and cMet (**F**, right panel) were calculated from 8 to 10 fields using the NIS-Elements software (Nikon). **D**, **G** Correlation plots between **D** Ki67 and Her2 and **G** Ki67 and cMet were generated for each indicated condition to test co-activation represented in **C**, **F**. *R* values indicating the degree of correlation between Ki67 and Her2 (**D**) and Ki67 and cMet (**G**) were calculated before and after RT. **H** Survival rates of 4T1 cells in response to Trastuzumab (T), Crizotinib (C), RT, RT+C, RT+T, and RT+T+C as detected by MB survival assays 6 days post RT (upper panel), and cell viability as measured by the MTT assay (lower panel). Drugs were added from 3 days prior to RT until the end of the experiment. **I** Key downstream to Her2 and cMet signaling proteins are shown following different treatments. The predicted combination induced high levels of cleaved caspase-3 compared to radiation alone, irradiation+T, and irradiation+C. Downregulation of pAKT, pERK, and p-S6 was detected when T+C was applied prior to RT. For **A**: quantification of subpopulations was performed using at least ~30,000 cells from each condition, which were obtained from at least three flasks and from at least three independent experiments for each time point. For **C** and **F**, statistically significant differences between all presented conditions and the cells treated with RT+T+C were determined using a two-tailed Student's *t* test (*$P < 0.05$); for **H**, upper and lower panels *$P < 0.01$

CSSS vs PCA and tSNE Several dimensionality reduction algorithms have been developed to interpret single-cell variations (e.g., variations in protein or gene expression levels), such as clustering-based t-SNE analysis [33] or principal component analysis (PCA) [34–36]. These methods are very useful in the statistical determination of dominant expression patterns but are limited when a more deterministic partitioning of a tumor mass into cellular subpopulations, based on cell-specific sets of altered molecular processes, is required. For example, t-SNE is a non-deterministic method (e.g., different runs with the same hyperparameters may produce different results) and is unable to assign a certain protein to several processes, or to determine which processes are active in every cell (Additional file 2: Fig. S3). Therefore t-SNE will be less efficient when the determination of robust cell-specific signaling signatures is required (e.g., for drug combination design). Similarly, PCA focuses mainly on the most dominant patterns obtained from proteins with the highest variability in the population, rather than on cell-specific sets of altered processes [30, 37]. Additional file 2: Fig. S3 and Fig. S4 show separation of the 4T1 single-cell data, obtained using either t-SNE or PCA analysis (performed using Python), using 2 main principal components. Minority separation between control and RT-treated cells and within RT-treated cells can be observed and therefore CSSS analysis was vital in identifying the two separate subpopulations, b and f, that expanded in response to RT, and were mapped and quantified (Additional file 2: Fig. S3 and Fig. S4, see also Figs. 3-4 and main text for more details).

Experimental methods

Patient-derived tissue used to establish BR45 tumors

Patient-derived tumors were established from a female patient with triple-negative, invasive lobular breast cancer. The tissue was derived from the local chest wall recurrence, s\a mastectomy, chemotherapy, and radiotherapy. When implanted into the NSG mice the tissue formed a tumor, and then was used for the in-vivo experiments as described below.

Cell lines

Murine 4T1 cells were kindly provided by Dr. Zvika Granot. MDA-MB-468 and MDA-MB-231 cells were acquired from ATCC and authenticated by the Genomic Center of the Technion Institute (Haifa). BR45 PDX were obtained from the Oncology Department at Hadassah –Jerusalem Medical Center with prior written informed consent. The BR45-derived and 4T1 cells were maintained and irradiated, after which flow cytometry was performed as indicated in Additional file 2: Supplementary Information file (SI).

Murine models

Syngeneic model 2.0×10^5 4T1 cells were inoculated subcutaneously into 6–7-week-old female BALB/c mice (ENVIGO).

Allogeneic model BR45 tumors were induced in NSG (Jackson Laboratory) female mice either by injecting 4.0×10^6 cells or by transplanting xenografts, orthotopically.

After reaching the initial volume 80–100 mm^3, mice were randomly grouped to approximately 8–10 animals per cage, and treatment was initialized. Tumor sizes were routinely measured with an electronic caliper every two days and their volumes were obtained using the formula V = (W (2) × L)/2. All in vivo experiments were performed with the approval of the Hebrew University of Jerusalem IACUC. See Additional file 2: SI Methods for more details.

In vivo treatments

High dose rate (HDR) brachytherapy (GammaMed™ HDR, Iridium 192) was performed as previously described [38]. 12 Gy was administered on two alternative days. Trastuzumab was purchased from Teva Pharmaceutical Industries Ltd. Crizotinib (cMet inhibitor, #12087-50) and Erlotinib (EGFR inhibitor, #10483-1) were purchased from Cayman Chemical. (See Additional file 2: SI Methods for doses and regimens).

Cohort description for each type of in vivo experiment performed

4T1 In vivo experiment First experiment: 4T1 cells (2×10^5) were subcutaneously injected in female BALB/c mice.

Mice that reached a tumor volume of 80–100 mm^3 were divided into two groups: control and irradiated. The irradiated group was treated by brachytherapy on two alternate days (12 Gy/day). Each group had two exit points: 6 and 12 days. Control samples n=12 (5 mice at day 6 post-RT and 7 mice at day 12 post-RT); RT-treated samples n= 17 (9 mice at day 6 post-RT and 8 mice at day 12 post-RT).

At the end of each time point mice were terminally anesthetized with Ketamine-Xylazine (150 mg/kg/20 mg/kg) IP, after which mice were euthanized by cervical dislocation. An incision was made and the tumor mass as a whole was gently separated from the conjunctive tissue using a sharp blade. Six and 12 days post-RT, whole tumors were collected for FACS analysis after mechanical cell dissociation. CSSS analysis was performed using FACS output.

Second experiment: The setup and procedure of collecting and analyzing the output was exactly as described in the first experiment. When tumors reached volumes of 80–100 mm^3, mice were divided into four groups: each group had two exit points — 6 and 12 days. Mice were treated with 5 mg/kg Trastuzumab (T) and 25 mg/kg Crizotinib (C) starting 3 days prior to brachytherapy until the end of the experiment (day 17). Control: n=12 (7 mice at day 6 post-RT and 5 mice at day 12 post-RT), RT: n= 11 (5 mice at day 6 post-RT and 6 mice at day 12 post-RT), RT+T+C: n= 7 (4 mice at day 6 post-RT and 3 mice at day 12 post-RT), T+C: n= 10 (5 mice at day 6 post-RT and 5 mice at day 12 post-RT), RT+T: n= 12 (6 mice at day 6 post-RT and 6 mice at day 12 post-RT). For more details see Additional file 1.

BR45 PDX in vivo experiments A small portion (~30 mm^3) of BR45 PDX was transplanted orthotopically into each NSG female mouse. When tumor volumes reached 80–100 mm^3, mice were divided into seven groups; each group had two exit points: 6 and 12 days. Mice were treated with 5 mg/kg Trastuzumab (T), 25 mg/kg Crizotinib (C), and 12.5 mg/kg Erlotinib (E) starting 3 days prior to RT until the end of the experiment (day 17). Control: n=6 (3 mice at day 6 post-RT and 3 mice at day 12 post-RT), RT: n= 5 (3 mice at day 6 post-RT and 3 mice at day 12 post-RT), RT+T: n= 6 (3 mice at day 6 post-RT and 3 mice at day 12 post-RT), RT+C: n= 6 (3 mice at day 6 post-RT and 3 mice at day 12 post-RT), RT+T+C: n=6 (3 mice at day 6 post-RT and 3 mice at day 12 post-RT), RT+T+C+E: n= 6 (3 mice at day 6 post-RT and 3 mice at day 12 post-RT), T+C: n= 7 (4 mice at day 6 post-RT and 3 mice at day 12 post-RT). Mice were irradiated by brachytherapy on two alternate days with two doses (12 Gy and 10 Gy). The setup procedure of collecting and analyzing the output data was exactly as described in the 4T1 model. For more details, see Additional file 1.

Flow cytometry Each sample was labeled with an 11 fluorescently tagged antibody (Ab) mixture. In addition when analyzing tumors, an exclusion cocktail including anti-mouse CD45, CD31 and CD140 was used in *in-vivo* experiments to exclude adjacent stromal and immune cells (Additional file 2: Table S8). A LSR-Fortessa Analyzer was utilized to measure all biomarkers simultaneously in each cell. See Additional file 2: SI Methods for more details.

Western blot analysis Cell pellets were lysed with a 20% SDS buffer (targeted drugs were added 1 day prior to RT, which allowed to obtain enough protein concentration for WB). The protein content of each lysate was determined with a Pierce BCA Protein Assay Kit (#23225, ThermoFisher). Equal protein aliquots were subjected to SDS-PAGE (Criterion Stain Free, 4–15% acrylamide, Bio-Rad) under reducing conditions and proteins were transferred to a nitrocellulose membrane. (Millipore). Membranes were blocked with 5% non-fat milk for 1 hour at R.T. and probed with the appropriate antibody (Additional file 2: SI Methods), followed by horseradish peroxidase-conjugated secondary antibody (#123449, Jackson ImmunoResearch) and a chemiluminescent substrate (ECL #170-5061, Bio-Rad).

Survival assay Cells were seeded at 70% confluency and treated as required for different time points. Cells were washed with PBS and fixed with 4% PFA for 10 min at R.T. The fixed cells were stained with methylene blue (MB) for 1 hour at R.T., washed and air dried overnight. The dye was extracted with 0.1M HCl for 1 hour at R.T. Absorbance was read at 630 nm.

MTT assay Cells were seeded and treated as indicated in a 96-well plate for 6 days. Cell viability was determined using an MTT assay kit (#ab211091, Abcam). Equal volumes of MTT solution and culture media were added to each well and incubated for 3 h at 37 °C. MTT solvent was added to each well, and the plate was then covered with aluminum foil and put on the orbital shaker for 15 min. Absorbance was read at 590 nm after 1 h.

Immunofluorescence Cells were grown on coverslips in six-well plates to reach 70% confluency by the next day, then fixed and permeabilized with cold absolute methanol. Afterwards, they were blocked with CAS blocker (cat. no. ZY-008120) and washed 3 times for 5 min with PBS, then stained with the following primary antibodies: Anti-mouse/human Ki-67 (BLG-151202), Rabbit Anti-Met (cMet) Polyclonal Antibody (BS-0668R), Neu (F-11) SC-7301. After washing 3 times with PBS for 5 min, cells were stained with secondary antibodies for 1 h at room temperature in the dark to visualize the aforementioned primary antibodies. The secondary antibodies conjugated to fluorophores were as follows: Goat anti-rat IgG H&L conjugated with Alexa Fluor 647 (1:400) (cat. no. 712605153), Goat anti-mouse IgG (H+L) conjugated with Alexa Fluor 488 (1:150) (cat. no. 115545003), and Goat anti-Rabbit IgG (H+ L) conjugated with Alexa Fluor 488 (1:150) (cat. no. 111545003). All secondary antibodies were purchased from Jackson ImmunoResearch. After washing 3 times with PBS, cell slides were mounted using fluorogel III mixed with DAPI (EMS, cat. no. 17985-01) to stain the nuclei. A spinning disk confocal microscope was used to visualize the expression of biomarkers of interest. The analysis was done using NIS-Elements software (Nikon).

Experimental statistical analysis Significant differences between experimental conditions and experimental reproducibility were determined using the Student's *t*-test (two tails, two samples equal variance); *P* values of ≤ 0.05 were considered statistically significant. All data was represented as the mean ± S.E. (standard error of the mean) if not indicated otherwise. Quantification of subpopulations was performed using a minimum of 30,000 cells from each condition, which were obtained from at least three flasks and from at least three independent experiments for each time point/condition. All experiments were performed minimally in biological triplicate if not indicated otherwise.

Code availability statement All equations and mathematical procedures used in this article are detailed in the "Methods" section and/or referenced. The approach is covered by patent applications "A method for selecting patient specific therapy", PCT/IL2019/050474, and "Methods of Determining Cancer Therapy," PCT/IB2021/056136. Any additional clarification/information regarding mathematical procedures/codes can be provided upon request. The codes for single-cell computational analysis are publicly available from Github [31].

Results

Study overview

To collect high-resolution data regarding the intratumoral composition of TNBC tumors in response to RT, we employed the following computational-experimental strategy: (1) A list of cell-surface oncomarkers for single-cell quantification and analysis (Methods) was determined using a literature search (Fig. 1A, left panel). (2) Single-cell suspensions from multiple sources (e.g., cell lines, tumors from mice, or patient-derived models) were labeled with fluorescently-labeled antibodies targeting selected cell-surface oncoproteins and assayed by multicolor FACS to reveal oncoprotein expression levels in each cell (Fig. 1A, right panel). In every experimental condition, 30,000–50,000 single cells were profiled allowing for the identification of different subpopulations, including very small subpopulations (comprising less than 1% of the total population) that have significantly limited detection rates when using standard pathological tests. (3) SA was extended to single cells ("Methods") to identify *sets* of unbalanced processes (cell-specific signaling signatures (CSSS)) in each cell, thereby identifying distinct cellular subtypes within the tumor (Fig. 1B). (4) We hypothesized that targeting evolving cellular subpopulations in response to RT would enhance TNBC response to RT and inhibit RT resistance development. A series of in vitro and in vivo experiments were performed to validate this hypothesis (Fig. 1C) as detailed in the sections below (Additional file 1 includes the metadata describing all the TNBC models used in the study).

CSSS analysis

We selected 11 cell surface proteins for single-cell quantification using an extensive literature search on relevant oncomarkers for breast cancer [39–45]. The list of oncomarkers included Her2, EGFR, EpCAM, CD44, CD24, PD-L1, cKit, CD133, E-Cadherin, cMet and MUC1.

These oncomarkers are known to be involved in breast cancer proliferation with an aggressive phenotype (EGFR, Her2, MUC1,cMet, and cKit) [46–51], cancer metastasis and invasion (EpCAM, E-Cadherin, CD133, MUC1,cMet and cKit) [50, 52–55], stem cell properties (CD44, CD24, CD133) [54, 56] and immune response (PD-L1) [45], and also represent potential drug targets for breast cancer therapy (Her2, cMet, EGFR, MUC1, cKit, PD-L1).

Protein expression levels of the surface oncomarkers were quantified in single cells (Fig. 2A) using multicolor FACS and analyzed via single-cell SA ("Methods"). Single-cell protein expression levels were used to compute the steady state and deviations thereof due to unbalanced processes operating in the tumor. Proteins that deviated from the steady state (Fig. 2B) in a similar manner were grouped into unbalanced processes (Fig. 2C, "Methods"). Importantly, not all the processes identified by the analysis were active in each cell. Only processes with significant amplitudes were assigned to a cell to identify a *set* of cell-specific processes, which we termed "cell-specific signaling signature" (CSSS) (Fig. 2D). For simplicity of representation, each CSSS was converted into a cell-specific barcode, graphically representing a set of active processes in a cell (Fig. 2D, active processes are labeled in red). Based on matching CSSS we identified distinct subpopulations (Fig. 2E, "Methods").

We suggest that the CSSS is what governs the optimal therapeutic strategy. The in-depth data collected up to this point was utilized to devise a therapeutic strategy that incorporated targeted therapies to aid RT. This was achieved by targeting dominant and RT-resistant subpopulations, to potentially achieve long-term tumor remission (Fig. 2F).

10 unbalanced processes give rise to the expression variations of 11 cell-surface proteins in 4T1 mouse TNBC cells

The first TNBC model used in the study included 4T1 cells, originally derived from a spontaneously arising mammary tumor in BALB/c mice and representing a model for stage IV TNBC [57]. The cells were irradiated at 5 Gy or 15 Gy and grown under standard conditions for 24h, 48h, and 6 days. Overall distribution of Her2

and cMet expression levels (Fig. 3A), as well as additional oncomarkers (Additional file 2: Fig. S5), were altered in response to RT. Two-dimensional correlation plots showed that although certain proteins, such as cMet and Her2, were upregulated in response to RT (Fig. 3A), their expression levels had poor correlation (Fig. 3B). EGFR and Her2 levels, however, demonstrated a strong correlation (Fig. 3C), as did MUC1 and cMet (Fig. 3D).

To examine all possible protein-protein relationships we utilized single-cell surprisal analysis (SA, Methods). Single-cell SA allowed for the identification and mapping of the unbalanced processes operating in the cellular population as a whole as well as in each cell.

The analysis revealed 10 unbalanced processes in the cell population. Four of the processes, all appearing in at least 1% of the treated and/or untreated cells, are shown (Fig. 3F; other processes, quantified by SA, can be found in Additional file 2: Fig. S1). Processes 3 and 8 included co-expressed and induced Her2/EGFR and cMet/MUC1 respectively (Fig. 3F, Additional file 3). Figure 3E exemplifies the procedure of mapping unbalanced processes within single cells – EGFR and Her2 show a higher correlation in the subset of cells which are assigned process 3 compared to all irradiated cells presented in Fig. 3C.

Expansion of Her+ and cMet+ distinct subpopulations are observed in response to RT

Due to the fact that more than one unbalanced process may operate in a single cell, the sets of unbalanced processes (CSSS) in each cell were analyzed to reveal CSSS-based subpopulations. Eight dominant subpopulations, each represented by a unique barcode consisting of processes 1, 2, 3, and 8 (Fig. 3G), were found in the cell population before and/or after RT. Only CSSSs that appeared in at least 1% of the cells were considered.

When we examined the temporal behavior of the dominant subpopulations, we found that the majority of dominant subpopulations did not change or reduced in response to RT (Additional file 3). For example, subpopulation c comprised 9.9% of the cells before RT and decreased to 4.9%, 6 days post-RT. However two subpopulations b and f expanded significantly 6 days post-RT (Fig. 4A). Subpopulation b harbored only process 3, in which Her2 and to a lesser extent EGFR (Fig. 3F), were induced. Strikingly, subpopulation b was induced ~70-fold post-irradiation relative to non-irradiated cells, where an expansion from <1% of the population in untreated cells to ~19–25% of the population 6 days post-RT occurred (Fig. 4A).

Subpopulation f harbored only process 8 (Fig. 3G), with induced cMet/MUC1 and reduced E-Cadherin. Significant induction of subpopulation f was also observed, from undetectable levels to 1.5% 6 days post-RT (Fig. 4A). These results demonstrate an important concept: although cMet and Her2 were both induced in response to RT (Fig. 3A), CSSS-based analysis revealed that these two proteins were expressed in *distinct* cellular subpopulations (processes 3 and 8 did not appear in the same cells; Fig. 3G). The development of such significant, distinct, and well-defined Her2+ and cMet+ subpopulations post-RT suggests that Her2 and cMet signaling may play a significant role in 4T1 cell survival and resistance in response to irradiation.

To characterize proliferative properties of the expanded Her2+ and cMet+ subpopulations in response to RT, we co-stained the 4T1 cell population with anti-Ki67 (proliferative biomarker), anti-cMet, and Her2 antibodies using immunofluorescent assays. Ki67, Her2, and cMet expression increased significantly in the cells surviving RT (Fig. 4B, C, E, and F). Moreover, this result was supported by enhanced coordinated expression of Her2 and Ki67 (Fig. 4D) as well as cMet and Ki67 (Fig. 4G) proteins respectively, as represented by an increased correlation between Her2 and Ki67; and cMet and Ki67 proteins, post-RT. This enhanced correlation in protein expression reveals the increased proliferative properties of Her2 or cMet expressing cells.

Simultaneous inhibition of Her2 and cMet sensitized 4T1 TNBC model to RT treatment

We hypothesized that simultaneous inhibition of both proteins, and thus targeting of both subpopulations, may sensitize 4T1 cells to RT. Her2 and cMet represent good candidates for such a strategy, as they are both druggable oncoproteins against which FDA-approved drugs exist. To validate this hypothesis, we inhibited either each protein alone or in combination, beginning 3 days prior to RT until 6 days post-RT, after which cell survival was measured.

The Her2 inhibitor, Trastuzumab (T), and cMet inhibitor, Crizotinib (C), showed a synergistic effect in sensitizing the cells to RT (Fig. 4H, upper and lower panels). The combination of both drugs with RT increased cell death and depleted downstream signaling to Her2 and cMet, as indicated by the low levels of downstream signaling proteins pERK, pAkt, and pS6 and the enhanced cleavage of the apoptotic marker Casp3 (Fig. 4I).

To further validate our hypothesis, we implanted 4T1 cells into immunocompetent BALB/c mice. The tumors were irradiated using brachytherapy-focused irradiation technology adapted for mice [38] by CT imaging and Monte-Carlo-based dosimetry (Fig. 5A). 4T1 tumors were then excised and single-cell suspensions were analyzed. CSSS-based analysis of the tumors 6 days post-RT, when an initial shrinkage of tumors was observed (Fig. 5B), revealed an expansion of subpopulations b

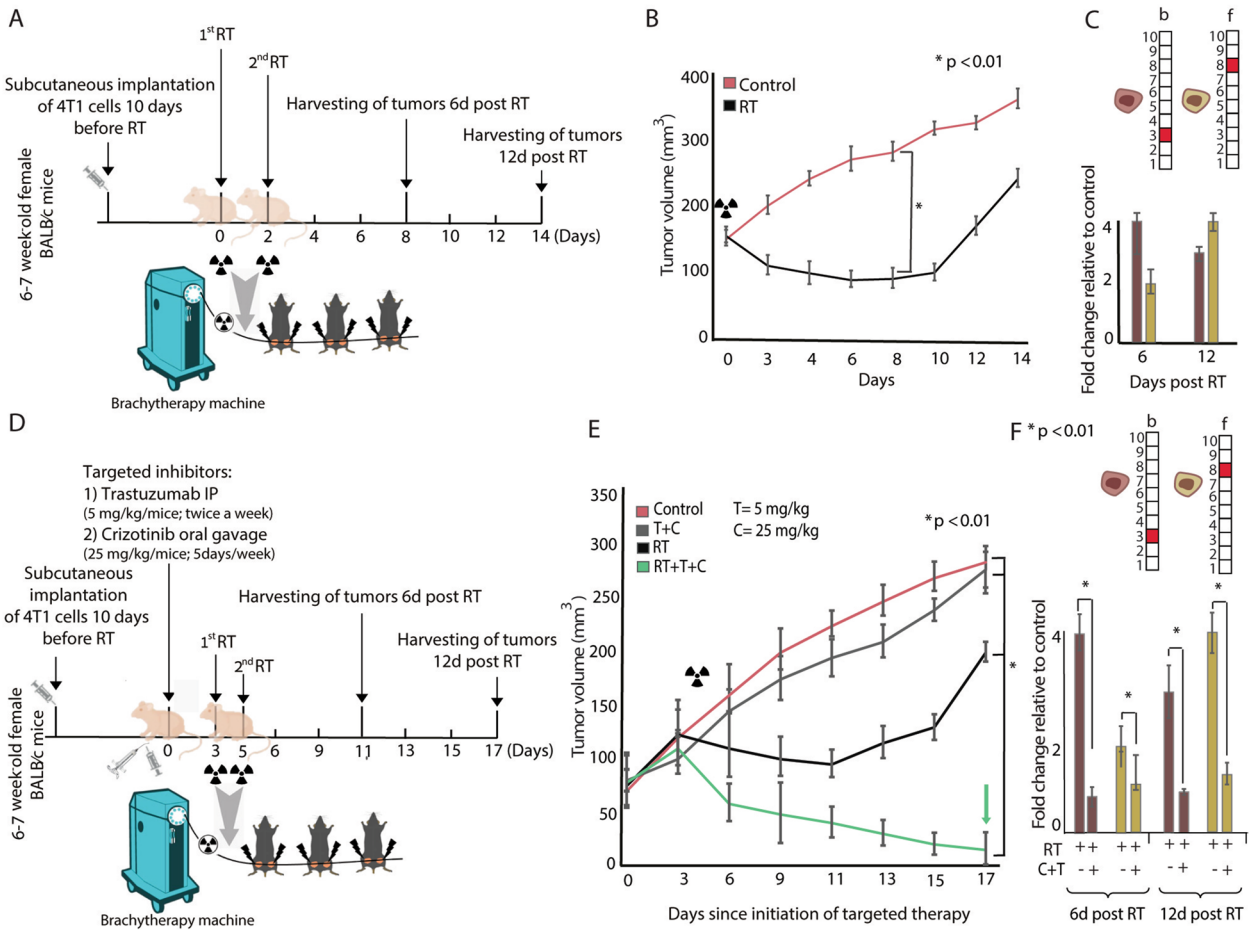

Fig. 5 Inhibition of RT-induced subpopulations sensitized tumor response to RT. **A** 6–7-week-old BALB/c female mice were subcutaneously injected with 4T1 cells. When tumor volumes reached 80–100 mm^3, mice were treated with brachytherapy RT on two alternate days (12 Gy). **B** Tumor volumes in the control group (*red*) versus the RT group (*black*). *$p < 0.01$ and ±SD are shown. **C** Fold change in the abundancy of subpopulations b and f as compared to untreated tumors. A significant expansion due to RT in subpopulation *b* harboring $Her2^+/EGFR^+$ and subpopulation f harboring $cMet^+/MUC1^+$ is detected. **D** Mice were subcutaneously injected with 4T1 cells and treated with RT. Trastuzumab (T), 5 mg/kg, and Crizotinib (C), 25 mg/kg, were administrated IP 2d/week and by gavage 5d/week respectively from d0 (3 days prior to RT) until the end of the experiment (d17). **E** C+T sensitized TNBC to RT in the 4T1 model, ± S.E., and *$p < 0.01$ are shown. **F** In vivo fold changes in subpopulations b and f showed optimal reduction when T and C were used in combination with RT. These results were consistent 6 and 12 days after RT. For **C** and **F**: quantification of subpopulations was performed using at least ~30,000 cells from each condition, which were obtained from at least three independent experiments for each time point, fold change is relative to control of each time point. For **B**, **C** mice used for each condition: control n=12, RT n= 11. On day 6 post-RT, we euthanized 7 control mice and 5 RT mice. On day 12 post-RT the experiment was completed and all remaining mice were euthanized. For **E**, **F** mice for each condition: control n= 12, RT n= 11, RT+T+C n= 7, T+C n= 10, RT+T n=12. On day 6 post-RT, 7 control mice, 5 RT mice, 4 RT+T+C mice, 6 RT+T, and 5 T+C mice were euthanized. On day 12 post-RT, the experiment was completed and all remaining mice were euthanized. For **B**, **E**, and **F**, statistically significant differences compared to cells treated with RT (**B**) and cells treated with RT+T+C (**E**, **F**) were determined using a two-tailed Student's *t* test (*$P < 0.01$)

and f (Fig. 5C). Moreover, 12 days post-RT, when tumor growth resumed (Fig. 5B), these expanded subpopulations were still detected (Fig. 5C, Additional file 4, Additional file 5). Inhibition of both Her2 and cMet proteins (as detailed in Fig. 5D) significantly sensitized the tumors to RT (Fig. 5E). This combined treatment resulted in tumor shrinkage and prevented the development of RT resistance (Fig. 5E). The effect of RT plus the combined targeted therapy was highly synergistic in contrast to the effect of the two targeted drugs without RT, or RT treatment alone. Furthermore, the addition of the targeted drug combination (T+C) prior to RT brought about a significant reduction in the size of subpopulations b and f (Fig. 5F). No other subpopulation expanded following the treatment (Additional file 4, Additional file 5).

Targeting Her2 and cMet to sensitize human cell lines and patient-derived TNBC to RT

To ensure that the expansion of Her2+ and cMet+ cellular subpopulations was not limited to TNBC mouse models, we tested TNBC MDA-MB-231 and MDA-MB-468 human-derived cell lines, and TNBC patient-derived cells (BR45). Inhibition of cell growth, observed in all cell types 6 days post-RT, was followed by significant cell regrowth 14 days post-RT (Fig. 6A). Subpopulations b and f, which expanded 6 days post-RT in all cell types, either maintained their size or expanded following cellular regrowth 14 days post-RT (Fig. 6B). Combined anti-Her2 and anti-cMet pretreatment sensitized all 3 types of human TNBC cells to RT (Fig. 6C). Each drug alone had a significantly smaller effect on cellular survival when compared to the combination of both drugs together with RT (Fig. 6C). Moreover, depletion of the downstream signaling pathways to Her2 and cMet as well as induction of cleaved caspase 3 were observed when the cells were pretreated with anti-Her2 and anti-cMet (Fig. 6D).

Using patient-derived TNBC BR45 cells grown in PDX models, we demonstrated that irradiated BR45 TNBC developed resistance to RT in a short period of time (tumor regrowth was detected 6 days post-RT; Fig. 6E, see black curve). Pretreatment of the mice with each drug alone prior to RT resulted in a small inhibitory effect on tumor growth (Fig. 6E). Pretreatment of the mice with the combination of both drugs, however, showed significant synergistic effects with RT, bringing about significant shrinkage of the tumor and preventing the development of resistance (Fig. 6E, green curve, Additional file 6, Additional file 7 and Additional file 8).

Adding erlotinib (an EGFR inhibitor), which according to our algorithms was not expected to significantly improve the results of the Trastuzumab + Crizotinib + RT treatment (Fig. 6E), did not significantly change the results obtained using Trastuzumab + Crizotinib + RT. Erlotinib monotherapy improved the response of BR45 to RT initially, albeit slightly, most likely due to the participation of EGFR in very small subpopulations (Additional file 7). The tumor, however, regrew after 1 week (Additional file 2: Fig. S6). Subpopulations b and f were reduced when the targeted drug combination (T+C) was applied prior to RT (Fig. 6F, Additional file 2: Fig. S7). These results suggest that CSSS-based single-cell resolution of the plasticity of TNBC in response to RT provides guidance on how effective targeted drug combinations should be designed in order to overcome RT resistance.

Discussion

Integration of computational and biological knowledge into efficient cancer treatment design has become an emerging concept in recent years. Although the induction of tumor cell plasticity [2] in response to anti-cancer therapies has been previously detected [15], a strategy to exploit this plasticity and provide successful treatment is still lacking.

In this study we provided a novel, single-cell framework for the improved resolution of intra-tumor cellular heterogeneity, allowing for the identification of independently evolving subpopulations. High-throughput FACS data was analyzed using single-cell information-theoretic surprisal analysis. The analysis resolved unbalanced protein subnetworks in the tumor [17], which were further attributed to single cells. Each cell was assigned a cell-specific signaling signature (CSSS), composed of a set of altered subnetworks. Cells sharing the same CSSS were considered a subpopulation. This strategy not only resolved overexpressed biomarkers or altered protein-protein correlation networks in response to RT treatment, but also mapped single-cell signaling signatures within the tumor tissue. This information enabled the resolution of distinct cellular subpopulations, information that is critical for accurate treatment design. Our analysis requires only one tissue/sample to elucidate the perturbed networks operating in each tumor, where the large number of single cells analyzed (>100,000/experiment) provides a high resolution of tumor heterogeneity. This is in contrast to bulk analysis which requires large datasets comparing multiple tissues in order to reveal high-resolution altered networks in each patient [19, 20, 30]. Furthermore, CSSS-based analysis efficiently identifies small cellular subpopulations, which are likely to be missed in bulk analyses.

Using the CSSS strategy, we revealed that two distinct cellular subpopulations harboring altered subnetworks with induced Her2 and cMet proteins, respectively, expanded in TNBC tumors in response to RT. Using in vitro and in vivo murine models, human cell lines, and patient-derived TNBC, we showed that efficient sensitization of aggressive TNBC to RT could be achieved only when Her2 and cMet proteins were inhibited simultaneously. Despite the fact that the in-vivo follow-up was only up to three weeks, the results demonstrated a significant synergistic effect in tumor response to RT and combined targeted therapy, compared to RT alone. While RT-treated tumors developed resistance, the tumors pretreated with Her2 and cMet inhibitors exhibited durable remission. In an extended follow-up period, there is a chance that other minor sub-populations may arise, which were not seen during the three weeks. In a clinical setting, a longer patient-specific follow-up might provide additional data for a more accurate treatment plan.

Conclusions

In summary, we reveal a novel approach to resolve in-depth intra-tumor heterogeneity at the single-cell level. This strategy provides an essential step towards the

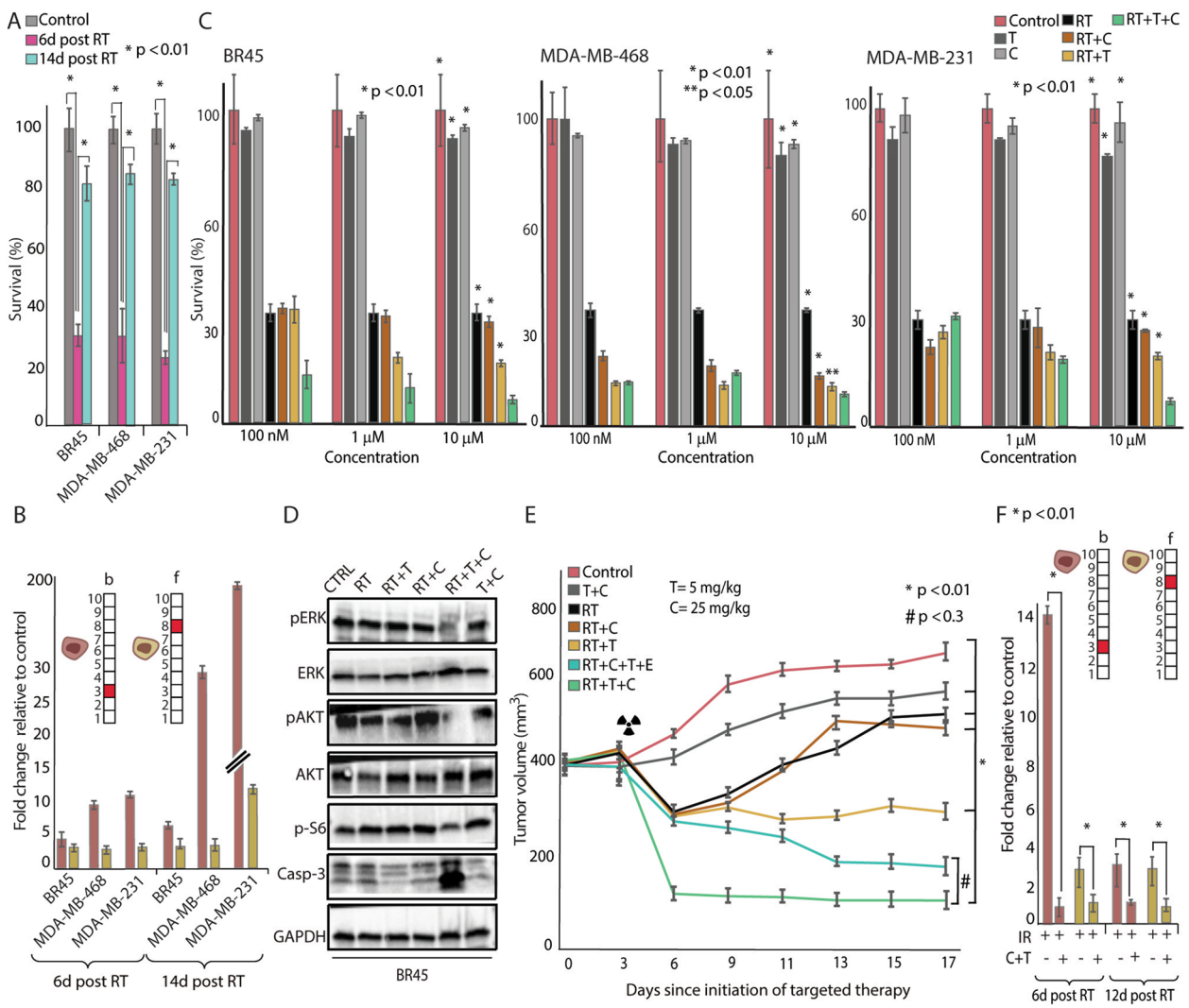

Fig. 6 Inhibition of expanded subpopulations sensitizes human TNBC cell lines and *BR45 PDX to* RT. **A** Survival assays show a ~ 30% cell survival rate 6 days post-RT, with TNBC regrowth to ~80–90% confluency 14 days post-RT. **B** Fold changes in the abundance of subpopulations b and f compared to untreated cells. These subpopulations either remained unchanged or expanded following cellular regrowth; fold change is relative to control of each time point. **C** Survival rates of BR45, MD-468, and MD-231 cells in response to Trastuzumab (T), Crizotinib (C), RT, RT+T, RT+C, and RT+T+C 6 days post-RT. Cellular drug treatment began 3 days prior to RT and was continued until the end of the experiment (d10). **D** Downstream to Her2 and cMet signaling was tested following different treatments. C+T combined with radiation-induced higher levels of cleaved caspase-3 compared to irradiation alone and irradiation with either C or T alone or C+T. C+T administration prior to RT induced the downregulation of pAKT, pERK, and p-S6 levels. **E** C+T sensitized TNBC response to RT in BR45 PDX in vivo. BR45 tissues were transplanted orthotopically into NSG mice treated with brachytherapy on days 3 and 5 with 12 Gy and 10 Gy, respectively. Drugs were administrated from d0 (3 days prior to RT) until the end of the experiment (d17), ± S.E. are shown. **F** In vivo fold changes in the abundance of subpopulations b and f in response to T and C, which is relative to control of each time point. For **A**, **B**, **C**, and **F**, ± S.D. are shown. For **B**, quantification of subpopulations was performed using at least ~30,000 cells from each condition, which were obtained from at least three flasks and from at least three independent experiments for each time point. For **E**, **F**, mice used for each condition: control *n*=6, RT *n*= 5, RT+T *n*=6, RT+C *n*= 6, RT+T+C *n*= 6, RT+T+C+E *n*= 6, and T+C *n*=7. On day 6 post-RT, 3 control mice, 2 RT mice, 3 RT+T mice, 3 RT+C mice, 3 RT+T+C mice, 3 RT+T+C+E mice, and 4 T+C mice were euthanized. On day 12 post-RT the experiment was completed and all remaining mice were euthanized. For **A**, **C**, **E**, and **F** statistically significant differences compared to cells treated with RT (**A**) and to cells treated with RT+T+C (**C**, **E**, **F**) were determined using a two-tailed Student's *t* test (**P* < 0.01; ***P* < 0.05; #*P* <0.3)

accurate design of targeted drug combinations for evolving tumor resistance. We validate this strategy by elucidation and detailed analysis of TNBC plasticity that allows for the sensitization of tumors to RT.

Importantly, this approach allows for the mapping of distinct cellular subpopulations in a single tumor, without the need to be compared to and analyzed relative to other tumors, such as in the case of bulk analyses. This strategy

can be universally applied to any cancer type and any treatment strategy by tailoring the panel of oncomarkers for a particular cancer type, where the computational approach would remain essentially the same. The value of this strategy will increase alongside the continued development of single-cell and mass cytometry techniques, which will allow for the simultaneous detection of dozens [58–60], possibly even hundreds, of signaling proteins in statistically significant numbers (>50,000–1,000,000) of single cells obtained from a single tumor.

Abbreviations

RT: Radiation therapy; CSSS: Cell-specific signaling signature; SA: Surprisal analysis; Her2: Human epidermal growth factor receptor 2; cMet: Cell surface mesenchymal-epithelial transition factor; TNBC: Triple-negative breast cancer; MDA-MB-231: M.D. Anderson-Metastasis Breast cancer-231; MDA-MB-468: M.D. Anderson-Metastasis Breast cancer-468; FACS: Fluorescence-activated single-cell sorting; EGFR: Epidermal growth factor receptor; EpCAM: Epithelial cell adhesion molecule; CD44: Cell surface glycoprotein; CD24: Small cell lung carcinoma cluster 4 antigen; PD-L1: Programmed cell death 1 ligand 1; KIT: Proto-oncogene, receptor tyrosine kinase; CD133: Prominin 1-CD133 antigen; E-Cadherin: Calcium-dependent adhesion protein, epithelial; MUC1: Tumor-associated epithelial mucin; MB: Methylene blue; MTT: 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-2H-tetrazolium bromide, used to assess cell viability as a function of redox potential.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13073-022-01121-y.

Additional file 1. Provides the details for the TNBC models and single cell analyses used in the study.

Additional file 2: Supplementary information file. Includes all the supplementary figures (Fig. S1-Fig. S7; and Table S8) with supplementary methods.

Additional file 3. Data file of in vitro experiments in the 4T1 system. The data includes single cell protein expression levels as measured by FACS, lambda ($\lambda_a(cell)$) values, G (G_{ia}) values and % of subpopulations out of the entire population.

Additional file 4. Data file of in vivo experiments in the 4T1 system 6 days post RT. The data includes single cell protein expression levels as measured by FACS, lambda ($\lambda_a(cell)$) values and G (G_{ia}) values and % of subpopulations out of the entire population.

Additional file 5. Data file of in vivo experiments in the 4T1 system 12 days post RT. The data includes single cell protein expression levels as measured by FACS, lambda ($\lambda_a(cell)$) values and G (G_{ia}) values and % of subpopulations out of the entire population.

Additional file 6. Data file of in vitro experiments in BR45 cells 6 and 14 days post RT. The data includes single cell protein expression levels as measured by FACS, lambda ($\lambda_a(cell)$) values and G (G_{ia}) values and % of subpopulations out of the entire population.

Additional file 7. Data file of in vivo experiments in BR45 PDX mice 6 days post RT. The data includes single cell protein expression levels as measured by FACS, lambda ($\lambda_a(cell)$) values and G (G_{ia}) values and % of subpopulations out of the entire population.

Additional file 8. Data file of in vivo experiments in BR45 PDX mice 12 days post RT. The data includes single cell protein expression levels as measured by FACS, lambda ($\lambda_a(cell)$) values and G (G_{ia}) values and % of subpopulations out of the entire population.

Acknowledgements

We acknowledge Science Training Encouraging Peace (STEP) for partial funding of this research, and Efrat Flashner-Abramson (H.A.'s STEP partner) for helpful input. All the tutors from the Core Research Facility at the Hebrew University of Jerusalem and all the physicists from Focused and Stereotactic Radiotherapy Unit, Dept. of Oncology-Hadassah hospital.

Authors' contributions

H.A., A.M.R., A.M., and N.K.-B. designed research; H.A., A.M.R., A.M., S.V., S.S., K.S., A.M., and N.K.-B. performed research; S.V., E.F.-A., S.S., T.P.Y., A.G., Z.G., I.B.P., S.R.C., and A.M. contributed experimental/analytical tools. H.A., A.M.R., S.V., N.E.C., and N.K.-B. analyzed data; H.A., E.F.A., A.M., K.S., and N.K.-B wrote the paper with contributions from all authors. The authors read and approved the final manuscript.

Funding

This work was supported by the Israel Science Foundation (ISF) (grant number 1961/19) and National Cancer Institute NIH (grant number U01 CA238720).

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files. The codes for single-cell computational analysis are publicly available from Github [31] (https://github.com/cohenoa/Computational-quantification-of-cellular-subpopulations-within-tumors-in-anti-cancer-therapy).

Declarations

Ethics approval and consent to participate

All in-vivo experiments were performed with the approval of the Hebrew University of Jerusalem IACUC, ethics committee-research number: MD-20-16116-5. NIH approval number: OPRR-A01-5011. The study was conducted in accordance with the protocol, good practice guidelines, and in accordance with regulations of the Authority For Biological and Biomedical Models at the Hebrew University of Jerusalem.
All experimentation involving human samples was approved by the Ethics Committee of the Hebrew University-Hadassah Medical School (approval no.: 0346-12-HMO) and was conducted under the provisions of the Declaration of Helsinki. Informed consent was obtained from all study participants.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]The institute of Biomedical and Oral Research, The Hebrew University of Jerusalem, 9103401 Jerusalem, Israel. [2]Sharett Institute of Oncology, Hebrew University-Hadassah Medical Center, 9103401 Jerusalem, Israel. [3]Department of Developmental Biology and Cancer Research, Institute for Medical Research-Israel-Canada, The Hebrew University-Hadassah Medical School, 91120 Jerusalem, Israel. [4]The Legacy Heritage Oncology Center & Dr. Larry Norton Institute, Soroka University Medical Center, Ben Gurion University of the Negev, Faculty of Medicine, 8410101 Beer Sheva, Israel. [5]School of Software Engineering and Computer Science, Azrieli College of Engineering, 9103501 Jerusalem, Israel.

Received: 20 January 2022 Accepted: 28 September 2022
Published online: 20 October 2022

References

1. Clairambault J, Shen S. Cell plasticity in cancer cell populations. F1000Research. 2020;9:1–16.
2. Yuan S, Norgard RJ, Stanger BZ. Cellular plasticity in cancer. Cancer Discovery. 2019;9:837–51.

3. Wang J-S, Wang H-J, Qian H-L. Biological effects of radiation on cancer cells. Mil Med Res. 2018;5:20.
4. Bernier J, Hall EJ, Giaccia A. Radiation oncology: a century of achievements. Nat Rev Cancer. 2004;4:737–47.
5. Arnold KM, Opdenaker LM, Flynn NJ, Appeah DK, Sims-Mourtada J. Radiation induces an inflammatory response that results in STAT3-dependent changes in cellular plasticity and radioresistance of breast cancer stem-like cells. Int J Radiat Biol. 2020;96:434–47.
6. Speers C, et al. Androgen receptor as a mediator and biomarker of radioresistance in triple-negative breast cancer. NPJ Breast Cancer. 2017. https://doi.org/10.1038/s41523-017-0038-2.
7. Jutzy JMS, Lemons JM, Luke JJ, Chmura SJ. The evolution of radiation therapy in metastatic Breast cancer: from local therapy to systemic agent. Int J Breast Cancer. 2018;2018:1–7.
8. Mehta M, et al. HuR silencing elicits oxidative stress and DNA damage and sensitizes human triple-negative breast cancer cells to radiotherapy. Oncotarget. 2016;7:64820–35.
9. Kyndi M, et al. Estrogen receptor, progesterone receptor, HER-2, and response to postmastectomy radiotherapy in high-risk breast cancer: the Danish Breast cancer cooperative group. J Clin Oncol. 2008. https://doi.org/10.1200/JCO.2007.14.5565.
10. Johnson J, et al. Targeting PI3K and AMPKα signaling alone or in combination to enhance Radiosensitivity of triple negative breast cancer. Cells. 2020;9:1253.
11. Abramson VG, Mayer IA. Molecular heterogeneity of triple-negative Breast cancer. Curr Breast Cancer Rep. 2014;6:154–8.
12. He MY, et al. Radiotherapy in triple-negative breast cancer: current situation and upcoming strategies. Crit Rev Oncol Hematol. 2018;131:96–101.
13. Karaayvaz M, et al. Unravelling subclonal heterogeneity and aggressive disease states in TNBC through single-cell RNA-seq. Nat Commun. 2018;9:3588.
14. Lee KL, Kuo YC, Ho YS, Huang YH. Triple-negative breast cancer: current understanding and future therapeutic breakthrough targeting cancer stemness. Cancers. 2019;11:1–36.
15. Castiglioni F, et al. Radiation effects on development of HER2-positive breast carcinomas. Clin Cancer Res. 2007;13:46–51.
16. Remacle F, Kravchenko-Balasha N, Levitzki A, Levine RD. Information-theoretic analysis of phenotype changes in early stages of carcinogenesis. Proc Natl Acad Sci U S A. 2010;107:10324–9.
17. Flashner-Abramson E, Abramson J, White FM, Kravchenko-Balasha N. A thermodynamic-based approach for the resolution and prediction of protein network structures. Chem Phys. 2018;514:20–30.
18. Kravchenko-Balasha N, Shin YS, Sutherland A, Levine RD, Heath JR. Intercellular signaling through secreted proteins induces free-energy gradient-directed cell movement. Proc Natl Acad Sci. 2016;113:5520–5.
19. Flashner-Abramson E, Vasudevan S, Adejumobi IA, Sonnenblick A, Kravchenko-Balasha N. Decoding cancer heterogeneity: studying patient-specific signaling signatures towards personalized cancer therapy. Theranostics. 2019;9:5149–65.
20. Vasudevan S, et al. Overcoming resistance to BRAFV600E inhibition in melanoma by deciphering and targeting personalized protein network alterations. npj Precis Oncol. 2021;5:50.
21. Jubran MR, et al. Overcoming resistance to EGFR monotherapy in HNSCC by identification and inhibition of individualized cancer processes. Theranostics. 2022;12:1204–19.
22. Levine RD, Bernstein RB. Energy disposal and energy consumption in elementary chemical reactions. Information theoretic approach. Acc Chem Res. 1974;7:393–400.
23. Levine RD. An information theoretical approach to inversion problems. J Phys A Math Gen. 1980;13:91.
24. Levine RD. Molecular reaction dynamics. Cambridge: The University Press (The University Press; 2005.
25. Vasudevan S, et al. Drug-induced resistance and phenotypic switch in triple-negative Breast cancer can be controlled via resolution and targeting of individualized signaling signatures. Cancers. 2021;13:5009.
26. Poovathingal SK, Kravchenko-Balasha N, Shin YS, Levine RD, Heath JR. Critical points in tumorigenesis: a carcinogen-initiated phase transition analyzed via single-cell proteomics. Small. 2016;12:1425–31.
27. McMillan WG, Mayer JE. The statistical thermodynamics of multicomponent systems. J Chem Phys. 1945;13:276–305.
28. Mayer JE, Mayer MG. Statistical mechanics: Wiley; 1977.
29. McQuarrie DA. Statistical mechanics. 1st ed: University science books; 2000.
30. Vasudevan S, Flashner-Abramson E, Remacle F, Levine RD, Kravchenko-Balasha N. Personalized disease signatures through information-theoretic compaction of big cancer data. Proc Natl Acad Sci U S A. 2018;115:7694–9.
31. Alkhatib H, Rubinstein AM, Vasudevan S, Flashner-Abramson E, Stefansky S, Chowdhury SR, et al. Computational quantification and characterization of independently evolving cellular subpopulations within tumors is critical to inhibit anti-cancer therapy resistance. Github. https://github.com/cohenoa/Computational-quantification-of-cellular-subpopulations-within-tumors-in-anti-cancer-therapy. 2022.
32. Szklarczyk D, et al. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res. 2011;39:D561–8.
33. Linderman GC, Rachh M, Hoskins JG, Steinerberger S, Kluger Y. Fast interpolation-based t-SNE for improved visualization of single-cell RNA-seq data. Nat Methods. 2019;16:243–5.
34. Kholodenko B, Yaffe MB, Kolch W. Computational approaches for analyzing information flow in biological networks. Sci Signal. 2012;5:re1–re1.
35. Wei W, et al. Single-cell Phosphoproteomics resolves adaptive signaling dynamics and informs targeted combination therapy in Glioblastoma. Cancer Cell. 2016;29:563–73.
36. Amir ED, et al. viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. Nat Biotechnol. 2013;31:545–52.
37. Kravchenko-Balasha N, Johnson H, White FM, Heath JR, Levine RD. A thermodynamic-based interpretation of protein expression heterogeneity in different Glioblastoma Multiforme tumors identifies tumor-specific unbalanced processes. J Phys Chem B. 2016;120:5990–7.
38. Meirovitz A, et al. Role of heparanase in radiation-enhanced invasiveness of pancreatic carcinoma. Cancer Res. 2011;71:2772–80.
39. Fleisher B, Clarke C, Ait-Oudhia S. Current advances in biomarkers for targeted therapy in triple-negative breast cancer. Breast Cancer (Dove Med Press). 2016;8:183–97.
40. Duru N, et al. HER2-associated Radioresistance of Breast cancer stem cells isolated from HER2-negative Breast cancer cells. Clin Cancer Res. 2012;18:6634–47.
41. Soysal SD, et al. EpCAM expression varies significantly and is differentially associated with prognosis in the luminal B HER2+, basal-like, and HER2 intrinsic subtypes of breast cancer. Br J Cancer. 2013;108:1480–7.
42. Nakagawa M, et al. Expression of p53, Ki-67, E-cadherin, N-cadherin and TOP2A in triple-negative breast cancer. Anticancer Res. 2011;31:2389–93.
43. Sahlberg SH, Spiegelberg D, Glimelius B, Stenerlöw B, Nestor M. Evaluation of cancer stem cell markers CD133, CD44, CD24: association with AKT isoforms and radiation resistance in colon cancer cells. PLoS One. 2014;9:e94621.
44. Siroy A, et al. MUC1 is expressed at high frequency in early-stage basal-like triple-negative breast cancer. Hum Pathol. 2013;44:2159–66.
45. Wang Z-Q, et al. PD-L1 and intratumoral immune response in breast cancer. Oncotarget. 2017;8:51641–51.
46. Dieci MV, Miglietta F, Griguolo G, Guarneri V. Biomarkers for HER2-positive metastatic breast cancer: beyond hormone receptors. Cancer Treat Rev. 2020;88:102064.
47. Nguyen TH, Nguyen VH, Nguyen TL, Qiuyin C, Phung TH. Evaluations of biomarker status changes between primary and recurrent tumor tissue samples in Breast cancer patients. Biomed Res Int. 2019;2019:1–7.
48. Rimawi MF, et al. Epidermal growth factor receptor expression in breast cancer association with biologic phenotype and clinical outcomes. Cancer. 2010;116:1234–42.
49. Shams TM, Shams ME. Overexpression of c-KIT (CD117) in triple-negative breast cancer. Egypt J Pathol. 2011. https://doi.org/10.1097/01.xej.0000406601.42226.2d.
50. Jing X, Liang H, Hao C, Yang X, Cui X. Overexpression of MUC1 predicts poor prognosis in patients with breast cancer. Oncol Rep. 2019;41:801–10.
51. Kashiwagi S, et al. Significance of C-MET as a therapeutic target in triple-negative Breast cancer. Ann Oncol. 2012. https://doi.org/10.1016/s0923-7534(20)32770-8.
52. Zeng L, et al. Prognostic value of biomarkers EpCAM and αb-crystallin associated with lymphatic metastasis in breast cancer by iTRAQ analysis. BMC Cancer. 2019;19:1–11.

53. Horne HN, et al. E-cadherin breast tumor expression, risk factors and survival: pooled analysis of 5,933 cases from 12 studies in the Breast cancer association consortium. Sci Rep. 2018;8:1–11.
54. Brugnoli F, Grassilli S, Al-Qassab Y, Capitani S, Bertagnolo V. CD133 in Breast cancer cells: more than a stem cell marker. J Oncol. 2019;2019:1–8.
55. Harrell JC, Shroka TM, Jacobsen BM. Estrogen induces c-kit and an aggressive phenotype in a model of invasive lobular breast cancer. Oncogenesis. 2017;6:396:1–13.
56. Breast L. Tumorigenic and metastatic role of CD44 — /; 2020. p. 1–23.
57. Ferrari-Amorotti G, Chiodoni C, Shen F, et al. Suppression of Invasion and Metastasis of Triple-Negative Breast Cancer Lines by Pharmacological or Genetic Inhibition of Slug Activity. Neoplasia. 2014;16(12):1047–58
58. Spitzer MH, Nolan GP. Leading edge primer mass cytometry: single cells. Many Features. 2016. https://doi.org/10.1016/j.cell.2016.04.019.
59. Lin JR, et al. Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes. Elife. 2018;7:1–46.
60. Labrie, M. et al. Multi-omics analysis of serial samples from metastatic TNBC patients on PARP inhibitor monotherapy provide insight into rational PARP inhibitor therapy combinations: July 29, 2020. https://doi.org/10.1101/2020.07.25.20146431.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions