

## ARTICLE OPEN



# Dedifferentiation-mediated stem cell niche maintenance in early-stage ductal carcinoma in situ progression: insights from a multiscale modeling study

Joseph D. Butner<sup>1</sup>, Prashant Dogra<sup>1,2</sup>, Caroline Chung<sup>3</sup>, Javier Ruiz-Ramírez<sup>1</sup>, Sara Nizzero<sup>1</sup>, Marija Plodinec<sup>4</sup>, Xiaoxian Li<sup>5</sup>, Ping-Ying Pan<sup>6,7</sup>, Shu-hsia Chen<sup>2,6,7</sup>, Vittorio Cristini<sup>1,8,9</sup>, Bulent Ozpolat<sup>10</sup>, George A. Calin<sup>11</sup> and Zihui Wang<sup>1,2,7</sup>✉

© The Author(s) 2022

We present a multiscale agent-based model of ductal carcinoma in situ (DCIS) to study how key phenotypic and signaling pathways are involved in the early stages of disease progression. The model includes a phenotypic hierarchy, and key endocrine and paracrine signaling pathways, and simulates cancer ductal growth in a 3D lattice-free domain. In particular, by considering stochastic cell dedifferentiation plasticity, the model allows for study of how dedifferentiation to a more stem-like phenotype plays key roles in the maintenance of cancer stem cell populations and disease progression. Through extensive parameter perturbation studies, we have quantified and ranked how DCIS is sensitive to perturbations in several key mechanisms that are instrumental to early disease development. Our studies reveal that long-term maintenance of multipotent stem-like cell niches within the tumor are dependent on cell dedifferentiation plasticity, and that disease progression will become arrested due to dilution of the multipotent stem-like population in the absence of dedifferentiation. We have identified dedifferentiation rates necessary to maintain biologically relevant multipotent cell populations, and also explored quantitative relationships between dedifferentiation rates and disease progression rates, which may potentially help to optimize the efficacy of emerging anti-cancer stem cell therapeutics.

*Cell Death and Disease* (2022)13:485; <https://doi.org/10.1038/s41419-022-04939-x>

## INTRODUCTION

Female breast cancer diagnoses in the United States exceeded 250,000 for the first time in 2017 (the latest year data available) [1]; up to 25% of these diagnoses are ductal carcinoma in situ (DCIS) [2], the earliest and non-invasive form of breast cancer, representing a major risk factor for invasive cancer [3]. Women who have breast-conserving surgery (lumpectomy) for DCIS have about a 25% to 30% chance of recurrence, but the mechanistic link between DCIS and invasive breast cancer remains largely unclear. Currently, there are no effective diagnostics that are able to accurately predict which patients will experience disease progression beyond stage zero to invasive cancer, and as a result many patients are often exposed to overtreatment and may even receive surgery and radiation therapy [4]. Even the most aggressive estimates predict that only 1 in 3 DCIS patients will progress to invasive disease [5], further highlighting the need for a better understanding of the biophysical mechanisms involved in DCIS.

DCIS originates from luminal epithelial cells within the mature mammary duct, which undergo unregulated proliferation into the luminal cavity. In the healthy state, bipotent epithelial stem cell niches [6] and a complex interplay of endocrine and paracrine signaling within the mammary gland epithelia are integral to gland homeostasis, where stem cells give rise to daughter cells (daughters) with a proliferative phenotype capable of self-renewal and production of terminally differentiated daughters [7] (Fig. 1). Increasing evidence suggests that tumors originate from cells that gain one or more genetic alterations, resulting in “cells of origin” that initiate cancer and spearhead tumor formation [8]; we refer to these as *tumor initiating cells (TICs)* in this work. These cells of origin must possess the hallmarks of cancer, and the cancer stem cell (CSC) model has emerged to explain these cells of origin [9], which are thought to both maintain a CSC population within the tumor through self-renewal, while also producing daughters through a process of asymmetric division, giving rise to cells that also retain

<sup>1</sup>Mathematics in Medicine Program, Houston Methodist Research Institute, Houston, TX 77030, USA. <sup>2</sup>Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY 10065, USA. <sup>3</sup>Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. <sup>4</sup>Biozentrum and the Swiss Nanoscience Institute, University of Basel, Basel 4056, Switzerland. <sup>5</sup>Department of Pathology & Laboratory Medicine, Emory University School of medicine, Atlanta, GA 30322, USA. <sup>6</sup>Immunotherapy Research Center, Houston Methodist Research Institute, Houston, TX 77030, USA. <sup>7</sup>Neal Cancer Center, Houston Methodist Research Institute, Houston, TX 77030, USA. <sup>8</sup>Department of Imaging Physics, University of Texas MD Anderson Cancer Center, Houston, TX 77230, USA. <sup>9</sup>Physiology, Biophysics, and Systems Biology Program, Graduate School of Medical Sciences, Weill Cornell Medicine, New York, NY 10065, USA. <sup>10</sup>Department of Experimental Therapeutics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. <sup>11</sup>Department of Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA.

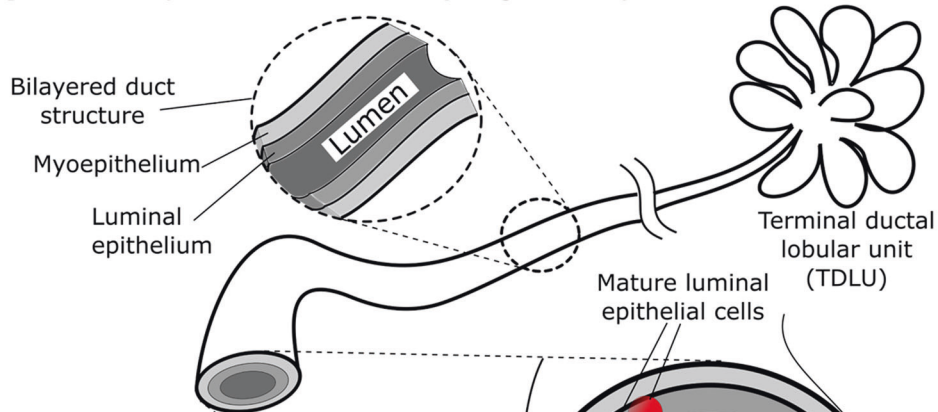
✉email: [zwang@houstonmethodist.org](mailto:zwang@houstonmethodist.org)

Edited by Professor Gerry Melino

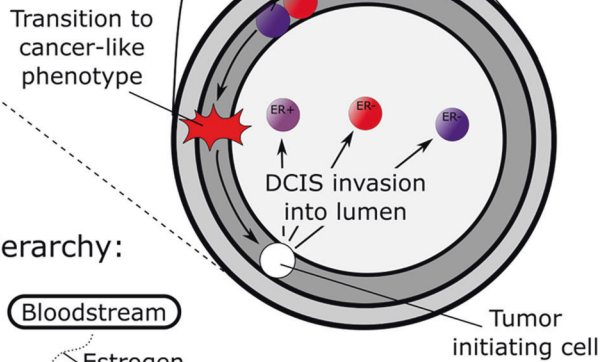
Received: 30 November 2021 Revised: 11 May 2022 Accepted: 11 May 2022

Published online: 21 May 2022

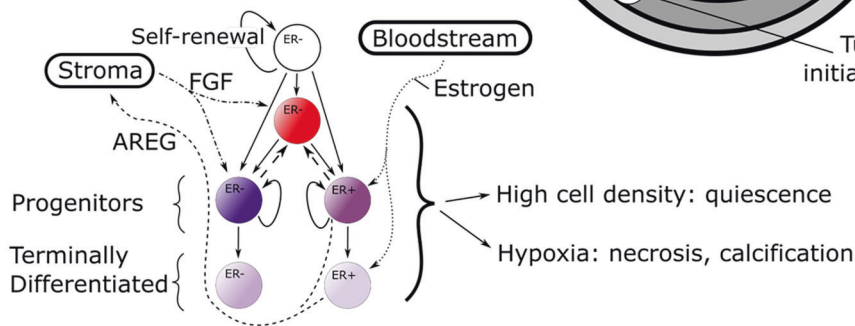
**a) Mammary Gland Structure (Single Duct)**



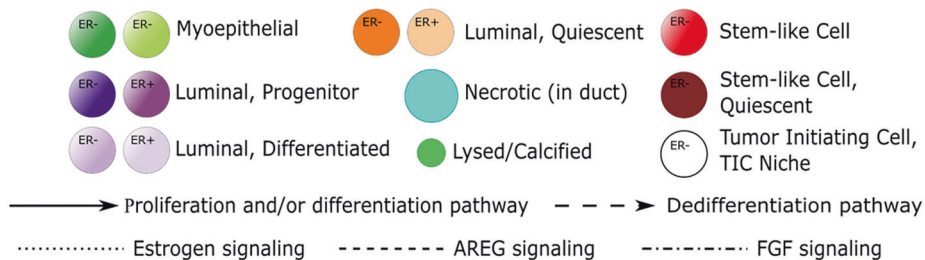
**b) DCIS Initiation**



**c) Modeled DCIS Cell Hierarchy:**



**d) Model Legend:**



**Fig. 1 Model schematic and included biology.** **a** The mammary gland is composed of bilayered epithelial walls surrounding an inner luminal cavity. This tree-like structure branches away from the nipple into the fat pad, and ducts are terminated by ductal lobular units. **b** The duct wall is composed of outer myoepithelial and inner luminal epithelial cell monolayers. In our model, cancer is initiated by transition of a stem cell within the luminal epithelium into a tumor initiating cell (TIC), which is able to proliferate indefinitely, placing its daughters adjacently into the luminal cavity. **c** Mammary epithelial cells are phenotypically distinct cells and are characterized based on their proliferative potential and the presence or absence of estrogen receptor  $\alpha$  (ER+/-, respectively). Proliferation within the DCIS population is regulated by simplified endocrine and paracrine pathways, as shown. **d** Model legend.

self-renewal capacity but with more differentiated phenotypes [7]. Cell proliferation is regulated by a complex interplay of endocrine and paracrine mechanisms, in part by estrogen-induced stimulation of estrogen receptor positive (ER+) mammary epithelial cells to upregulate amphiregulin (AREG) production, which stimulates estrogen receptor negative (ER-) cells via fibroblast growth factor (FGF) [10] (Fig. 1c). These

pathways are central to cell proliferation in mammary gland development, but may become dysregulated, leading to loss of homeostasis and a transition to a cancer phenotype [11, 12].

Increasing evidence has emerged for the existence of cellular plasticity in differentiation within tumors. In this process, cancer cells may experience phenotypic reversal [9], leading to a transition back to a more stem-like state, and resulting in CSCs

**Table 1.** Model parameters perturbed in the global sensitivity analysis.

Parameter	Symbol	Units	Baseline [range]	Distribution	Reference
<b>Discrete (ABM) parameters</b>					
Cell cycle time <sup>*.S</sup>	$\tau_P$	hours	20 [16.5 – 23.5]	Uniform	[29, 30]
Symmetric proliferation probability					
-progenitor (is daughter progenitor or differentiated?) <sup>*.†.S<sup>S</sup></sup>	$\omega_P$	%	75% [50% – 100%]	Uniform	N/A
-stem (is daughter stem?) <sup>*.†.S</sup>	$\omega_{SC}$	%	12% [12% ± 6%]	Lognormal	[46]
-stem (are non-stem daughters ER + or ER-?) <sup>*.S<sup>S</sup></sup>	$\omega_{SC,+/-}$	%	50% [36 – 64%]	Uniform	N/A
Apoptosis probability <sup>*.S</sup>	$\omega_{Apop}$	% $\Delta t^{-1}$	0.705% [0.10% – 1.31%]	Uniform	[47]
Proliferation cycles before differentiation <sup>*.†.S<sup>S</sup></sup>	$P_{max}$	Cell cycles	18 [11 – 25]	Uniform	
Cell density threshold for quiescence <sup>*.†.S<sup>S</sup></sup>	$\theta_Q$	Cells volume <sup>-1</sup>	70% [50 – 90%]	Uniform	N/A
Dedifferentiation probability <sup>†.S</sup>	$\Omega_{DD}$	% cell <sup>-1</sup> cycle <sup>-1</sup>	0.4% [0.1 – 0.7%]	Uniform	[24]
<b>Hybrid (ABM ↔ continuum) parameters</b>					
Cancer cell oxygen metabolism multiplier (cancer cell oxygen metabolism = $\lambda_C \times \lambda_E$ ) <sup>*.†.S</sup>	$\lambda_C$	Mol cell <sup>-1</sup> s <sup>-1</sup>	4.5× healthy cell metabolism [3.0 – 6.0×]	Uniform	[48]
Estrogen metabolism <sup>*.S<sup>S</sup></sup>	$\lambda_E$	Normalized (% baseline)	1.0 [0.9 – 1.1]	Uniform	N/A
FGF uptake/metabolism <sup>*.S<sup>S</sup></sup>	$\lambda_{FGF}$	Normalized (% baseline)	0.72 [0.648 – 0.792]	Uniform	N/A
AREG production <sup>*.S<sup>S</sup></sup>	$\lambda_{AREG}$	Normalized (% baseline)	0.086 [0.0774 – 0.096]	Uniform	N/A
Estrogen proliferation threshold <sup>*.†.S<sup>S</sup></sup>	$\theta_E$	Normalized (% baseline)	0.85 [0.85 ± 0.2]	Lognormal	N/A
FGF proliferation threshold <sup>*.†.S<sup>S</sup></sup>	$\theta_{FGF}$	Normalized (% baseline)	0.7 [0.7 ± 0.15]	Lognormal	N/A

Parameters were tested with either a uniform range [min-max] or a lognormally distributed range [mean ± error factor], with each run under a unique parameter setting determined by LHS. \*Parameters perturbed in the full global sensitivity analysis without dedifferentiation. †Parameters perturbed in the dedifferentiation sensitivity analysis. <sup>S</sup>Parameters quantified from in vivo or in vitro experimental measurements. <sup>S<sup>S</sup></sup>Parameters estimated mathematically or obtained from other modeling work.

that may not have originated from a healthy (non-malignant) stem-like cell [8]. This process is often referred to as dedifferentiation [13], and has been reported in many cancers, including glioblastoma [14], intestinal tumors [15], melanoma [16], and also in breast cancer, where CSCs arise through an epithelial-mesenchymal transition (EMT)-like process facilitated by the ZEB1 promoter [17]. Tumor microenvironmental factors have been shown to play key roles in the dedifferentiation process, including hypoxia [18], Ras-Myc signaling [19], nuclear  $\beta$ -catenin localization and Wnt signaling [20], and TNF- $\alpha$  signaling [16]. These are also thought to play a role in stem cell niche maintenance, where stem cells may also emerge due to dedifferentiation [18, 21]. Specific to breast cancer, dedifferentiation is proposed to be due to ZEB1-based driving of EMT events, which is also required for conversion of non-CSC to CSC and for the maintenance of CSC-like activity [17, 22]. This is also driven by hypoxia [18], oncogenic-induced Yamanaka factors (such as c-Myc or Klf4) [23], or mutated KRAS signaling. This mounting evidence challenges the traditional hierarchy of cellular differentiation, and suggests the need for additional understanding of the role dedifferentiation plasticity plays in cancer [24]. In fact, several groups have made notable strides in quantifying the dynamics of CSC emergence through dedifferentiation events, with a special focus on understanding the prevalence of these events and the conditions under which they may occur [17, 24].

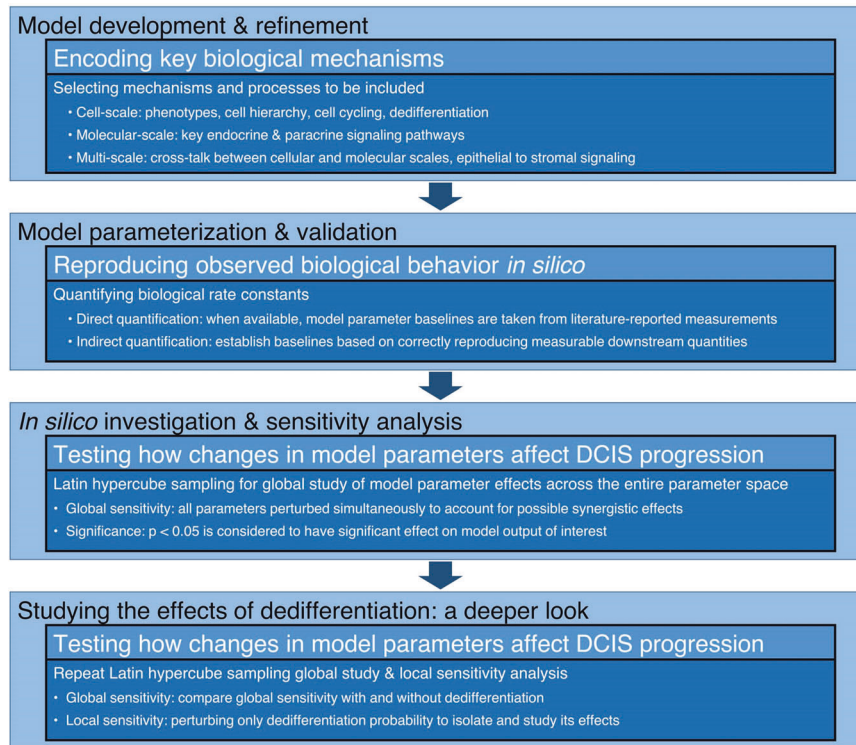
Mathematical modeling has yielded significant insights into DCIS development, and into the roles of cellular hierarchies and dedifferentiation in cancer progression; a detailed review is provided in SI. In the work presented herein, we build on our previous modeling studies of the mammary gland and DCIS

[25–28] to better understand how cellular dedifferentiation and a combination of other tissue, cellular, and molecular scale factors come together to influence the emergence of the early stages of DCIS (see Table 1), as outlined in (Fig. 2).

## METHODS

### Model overview

The basic multiscale agent-based modeling (ABM) framework was published in [27]; interested readers may refer to the SI and other previous publications [25, 28] for descriptions of model design, methods of implementation, and technical details. Briefly, in our model, cells are represented as discrete spherical entities (agents), each with unique phenotypes, positions within the cellular hierarchy, and spatial locations (Fig. 1c, Table 1, Supplementary Table S1), and are free to move lattice-free within the simulated domain (a section of mammary gland duct, Fig. 1c) based on physical interactions with their neighbors and the duct wall. At the molecular scale, continuum molecular profiles (oxygen, estrogen, AREG, and FGF) are described via Fick's law and solved at each time step using finite element methods. These molecular profiles are then mathematically linked to agents at the discrete scale for explicit feedback between scales, thereby implementing key epithelial and stromal signaling pathways involved in cellular proliferation. Cell agents simulate multiple, sequential proliferation cycles, and may undergo phenotypic transitions according to the cell hierarchy shown in Fig. 1c and determined by the reference criteria and probabilities in Table 1, Supplementary Table S1, and may proliferate, grow, terminally differentiate, enter quiescence, become hypoxic or necrotic, or undergo apoptosis, as determined by the conditions at their physical location (approximated as the agent's center of mass) and their phenotype. We have made the simplifying assumption that these state transitions (except for growth, which occurs over multiple steps within the time range specified in Table 1) are completed by the start of the next discrete (ABM) time step. Upon mitosis, daughter phenotypes



**Fig. 2 Study flowchart.** A graphical description of the study is shown. Briefly, after identifying key biological and physical processes and phenotypic and signaling hierarchies to be included in the model, the model was extensively tested to phenomenologically quantify model parameters that were unavailable in the literature, based on successful reproduction of literature-reported measurements. The full parameter space was examined using Latin hypercube sampling (LHS) parameter perturbation, and parameters with significant effects on model outcomes were identified. Perturbation studies were repeated with and without dedifferentiation, and local and global sensitivity analysis was performed.

**Table 2.** Global sensitivity analysis using linear regression without dedifferentiation.

Model output	$R^2$	Significant model parameters ( $p$ -value)
Axial advance rate ( $\mu\text{m}/\text{day}$ )	0.6389	$\omega_p$ ( $<0.001$ ); $\theta_Q$ ( $<0.001$ ); $\theta_E$ ( $<0.001$ ); $\theta_{FGF}$ (0.001); $\omega_{SC}$ (0.003); $\tau_P$ (0.005); $P_{max}$ (0.015)
Proliferation events per day	0.6599	$\theta_Q$ ( $<0.001$ ); $\omega_p$ ( $<0.001$ ); $P_{max}$ (0.001); $\omega_{SC}$ (0.002); $\theta_{FGF}$ (0.015); $\theta_E$ ( $<0.024$ ); $\tau_P$ (0.030)
Axial calcification extent ( $\mu\text{m}$ )	0.7715	$\theta_Q$ ( $<0.001$ ); $\lambda_C$ ( $<0.001$ ); $P_{max}$ ( $<0.001$ ); $\omega_p$ ( $<0.001$ ); $\omega_{SC}$ (0.036)
Total cells calcified	0.7639	$\theta_Q$ ( $<0.001$ ); $\lambda_C$ ( $<0.001$ ); $\omega_p$ ( $<0.001$ ); $P_{max}$ ( $<0.001$ )
Time to leading edge quiescence (days)	0.2377	$\theta_E$ (0.005); $\theta_{FGF}$ (0.006); $\omega_p$ (0.012)
Total cell count at arrest	0.7118	$\theta_Q$ ( $<0.001$ ); $P_{max}$ ( $<0.001$ ); $\omega_p$ ( $<0.001$ ); $\lambda_C$ (0.046);
Total ER + cells at arrest	0.6099	$\theta_E$ ( $<0.001$ ); $P_{max}$ ( $<0.001$ ); $\omega_p$ (0.001); $\theta_Q$ (0.008); $\theta_{FGF}$ (0.009); $\omega_{SC}$ (0.019)
Total ER – cells at arrest	0.6680	$\theta_Q$ ( $<0.001$ ); $\theta_E$ ( $<0.001$ ); $\lambda_C$ (0.016); $P_{max}$ (0.020)

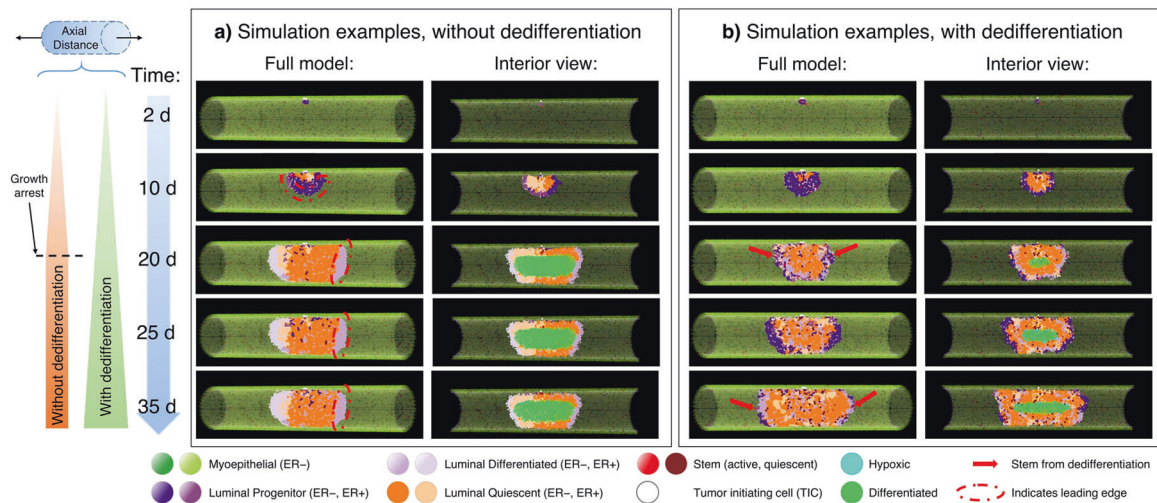
Goodness-of-fit ( $R^2$ ) of the full multivariate model and parameters found to have significant effects on each model output are shown. All parameters in Table 1 marked with an asterisk (\*) were perturbed in a Latin hypercube sampling (LHS) study; parameters not listed here were found to have insignificant effects on model outputs ( $p \geq 0.05$ ). Progenitor symmetric proliferation probability:  $\omega_p$ ; cell density threshold for quiescence:  $\theta_Q$ ; estrogen proliferation threshold:  $\theta_E$ ; FGF proliferation threshold:  $\theta_{FGF}$ ; probability of a stem cell having a stem daughter:  $\omega_{SC}$ ; cell cycle time:  $\tau_P$ ; proliferation cycles before differentiation:  $P_{max}$ ; cancer cell oxygen multiplier  $\lambda_C$ .

are determined stochastically with probabilities described in Table 1. All simulations presented herein were conducted within a simulated section of mature mammary duct represented as a cylinder 1 mm in length with 200  $\mu\text{m}$  inner luminal duct cavity diameter.

### New model components

**Stochasticity in cell cycle time.** In our previous work, we made the simplifying assumption that all cells would experience equal cell cycle time; that is, the minimum time to complete all cell cycle phases before reaching maturity and being able to undergo another mitosis cycle was held constant. While this allowed for focused study of the effects of

variations in cell cycle time on disease progression without the confounding effects of randomness in cell cycle times, it also resulted in an artificial cell cycle synchronicity. In the new version of the model, each daughter cell is assigned a cell cycle time randomized from a bounded lognormal distribution with a mean selected from the range shown in Table 2; means were held constant for each simulation run, but were varied between runs. Baseline mean cell cycle times for mammary epithelial cells were based on cell cycle time of 20 h reported by Shehata et al. [29]. We note that, because we do not explicitly model all phases of the cell cycle for each cell, we imposed a biologically relevant minimum cell cycle time of 16 h based on literature reported values [29, 30] in order to ensure all phases of the cell cycle are complete before another mitosis cycle



**Fig. 3 Representative model simulations.** **a** Representative model states rendered from a simulation without dedifferentiation are shown as full simulated duct (left) and cross-sectional view (right; cross-section taken along  $y = 0$  plane). DCIS is initiated at the top-center of the duct by 5 TIC cells (white) at  $t = 0$  days (d); daughters are placed adjacently into the luminal cavity, and DCIS grows away from site of initiation along the leading edge (examples are indicated by red dashed curves). In the example shown, the leading edge has become terminally differentiated at time  $t = 20$  days (d), and growth arrest is observed as only minimal further expansion of the DCIS mass occurs after this time. Calcification due to hypoxia is observed (green agents) in the cross-sectional view (right column). **b** Representative model images from a simulation with the dedifferentiation pathway activated. At time  $t = 20$  days (d), dedifferentiated cells (lighter purples) are again observed in the leading edge, but in this case new stem cells (red arrows) have emerged in the leading edge; this behavior is again observed in times  $t = 25$  days and  $t = 35$  days.

begins (thereby imposing a lower bound in case a random cell cycle time is selected below this biological threshold).

**Dedifferentiation.** In the work presented here, we have allowed for stochastic dedifferentiation of proliferative phenotype back into the cancer stem cell phenotype. We have also made the simplifying assumptions that quiescent cells do not dedifferentiate, and that terminally differentiated cells do not dedifferentiate back into the proliferative phenotype (we address the implications of this assumption in light of Yamanaka's discoveries that terminally differentiated cells may revert back to pluripotent stem cells [23] in SI). In addition, cells in our model may not dedifferentiate at the same time they undergo mitosis to avoid introducing artificial variations in the probability distributions for daughter phenotypes (Table 1). Lastly, because we only model a single stem phenotype in our model, agents that have a stem phenotype are unable to dedifferentiate. Agents that meet the above criteria are given the opportunity to stochastically dedifferentiate at each time step such that the dedifferentiation rate per cell cycle is as reported in Table 1, and in each simulation run all agents have the same per-cell-cycle dedifferentiation probability (note that this probability may be set to zero). The dedifferentiation pathway may be manually turned on ( $\Omega_{DD} > 0$ ) or off ( $\Omega_{DD} = 0$ ) in our simulations, enabling controlled *in silico* experimentation on the effects of the presence, absence, or stochastic likelihood ( $\Omega_{DD}$ ) of dedifferentiation events.

### Model analysis

We have previously reported extensive local sensitivity analysis of key model parameters without dedifferentiation [27]. Here, we present an expanded analysis covering a greater number of model parameters in a global analysis [31–39], wherein all model parameters of interest were perturbed simultaneously. For each simulation run, model parameter values were determined via Latin hypercube sampling (LHS), based on values and distributions shown in Table 1. All model simulations were run in serial for 96 wall-clock hours, and model states (including all cell locations, phenotypes, cell volumes, and continuum solutions at all nodes) and outputs of interest were recorded or calculated after every 30 minutes of simulated time. In our first global analysis, many model parameters were perturbed simultaneously without dedifferentiation (that is, dedifferentiation probability  $\Omega_{DD} = 0$ ), and multiple regression analysis was conducted to determine parameters that had significant ( $p < 0.05$ ) effects on model outputs of interest. In a second analysis, the dedifferentiation pathway was activated, but parameters tested in the first global analysis that were deemed to be insignificant to all or most of the model outputs of interest

were removed from the perturbation analysis (and held constant at their baseline), allowing us to focus on the key parameters deemed to have the greatest influence on DCIS progression. Finally, a third study was conducted wherein only the dedifferentiation probability was perturbed; these studies are outlined in Fig. 2.

## RESULTS

### DCIS progression is ultimately arrested without stem cell replenishment from dedifferentiation: a global sensitivity analysis

In a first step, a global sensitivity analysis was performed varying all 13 model parameters marked with an asterisk (\*) in Table 1, but without dedifferentiation (that is, dedifferentiation probability  $\Omega_{DD} = 0$ ). It was observed that stem cells experienced exponential dilution over time in the absence of dedifferentiation events (an example is shown in Supplementary Fig. S5), often leading to a lack of stem cells in the leading edge. When combined with the finite limit on mitosis cycles imposed on the progenitor phenotype before terminal differentiation, it was observed that the leading edge would ultimately become terminally differentiated in these cases, at which time DCIS progression was arrested (Figs. 3a and 4a). Times to leading edge arrest ranged between 9.1–90.5 days, with an average of 25.9 days. In this study, we truncated model outputs at the time of growth arrest, and results herein represent the state of the model at that time.

Model outputs examined in this study were (1) DCIS axial advance rate per day, (2) average proliferation events per day, (3) axial extent of calcification and time of growth arrest, (4) total cells calcified at time of growth arrest, (5) time before the DCIS leading edge became quiescent, (6) total cell count at time of growth arrest, and (7) total numbers of ER+/- cells at time of growth arrest. A multivariate linear model was constructed for each output of interest,  $R^2$  assessment of goodness-of-fit was performed, and model parameters that showed significant effects ( $p < 0.05$  was considered significant) on model outputs of interest were identified. We note that because DCIS advances in both directions along the duct axis (away from the location of tumor initiation), total advance per day and axial extent measures

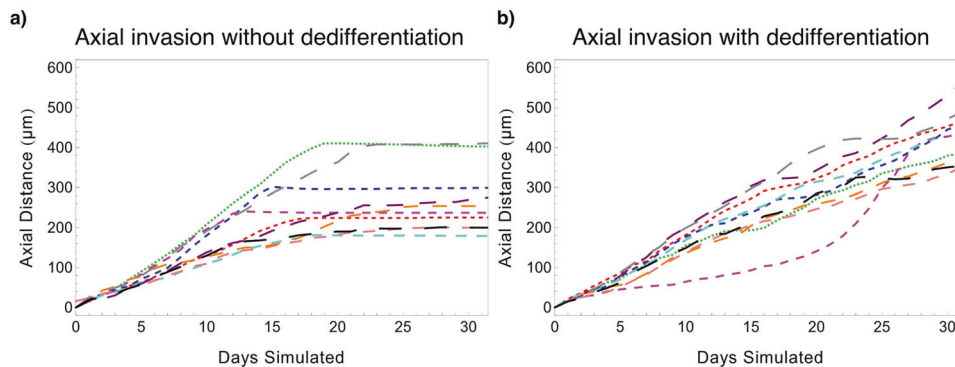
represent the summed change in axial distance in both directions (see Fig. 3).

Multivariate linear regression analysis revealed that only 8 model parameters showed a significant effect on model outputs; these are detailed in Table 2. Of these, progenitor symmetric proliferation probability ( $\omega_p$ ), cell density threshold for quiescence ( $\theta_Q$ ), and proliferation cycles before terminal differentiation in the progenitor phenotype ( $P_{max}$ ) were most often significant (in 7/8 model outputs examined), followed by estrogen proliferation threshold ( $\theta_E$ ) (5/8 model outputs), and then by FGF proliferation threshold ( $\theta_{FGF}$ ), stem cell symmetric proliferation threshold ( $\omega_{SC}$ ) (stem vs. progenitor daughters), and cancer cell oxygen consumption multiplier ( $\lambda_C$ ) (4/8 model outputs). Interestingly, cell cycle time ( $\tau_p$ ) was found to be significant in only 2/8 model outputs examined. The absolute values of the standardized (unitless) effects of each parameter on each examined output are shown in Supplementary Fig. S1.

**Inclusion of dedifferentiation overcomes growth arrest: global sensitivity analysis with active dedifferentiation pathway.** A second global analysis study was performed wherein the dedifferentiation pathway was activated (thus dedifferentiation probability  $\Omega_{DD} > 0$ ) and a reduced parameter set consisting of only those identified to be significant in the previous study plus the dedifferentiation probability were perturbed; these are labeled with a dagger ( $\dagger$ ) in Table 1. Multivariate regression was repeated as described above, and model goodness-of-fit and parameter significance on model outputs of interest were evaluated. We note that cell cycle time was excluded from this study, as it was

only significant in 2/8 outputs in our first study. Growth arrest of the DCIS cell mass due to terminal differentiation in the leading edge was not observed in this analysis study (see Figs. 3b, 4b); in all simulations, axial growth and cell population increase were observed until the end of the 96 wall-clock hour simulation runs. Stochastic dedifferentiation events were able to overcome the leading edge growth arrest behavior that was observed in the first analysis study (see the previous section), and growth arrest was not observed in any simulation run with dedifferentiation in effect (these are further examined later). Accordingly, time to leading edge growth arrest was not included as a model output examined in this second study.

It was observed that quiescence density threshold ( $\theta_Q$ ) and progenitor symmetric proliferation probability ( $\omega_p$ ) were most influential (that is, they exhibit significant influence on the largest number of model outputs examined) on the examined model outputs, with significant effects on 6/7 outputs of interest. Cancer cell oxygen multiplier ( $\lambda_C$ ) showed a significant effect on 5/7 model outputs, followed by estrogen proliferation threshold ( $\theta_E$ ), showing significant effects on 3/7 model outputs, and then by dedifferentiation probability ( $\Omega_{DD}$ ), maximum proliferation cycles before differentiation ( $P_{max}$ ), and FGF proliferation threshold ( $\theta_{FGF}$ ), each showing significant effects on 2/7 model outputs examined; details are provided in Table 3. Note that the probability of a stem cell having a stem daughter ( $\omega_{SC}$ ) was included in this analysis, but was found to be insignificant in all cases. Pareto charts depicting the absolute values of the standardized (unitless) effects of each parameter on each examined output are shown in Supplementary Fig. S2.

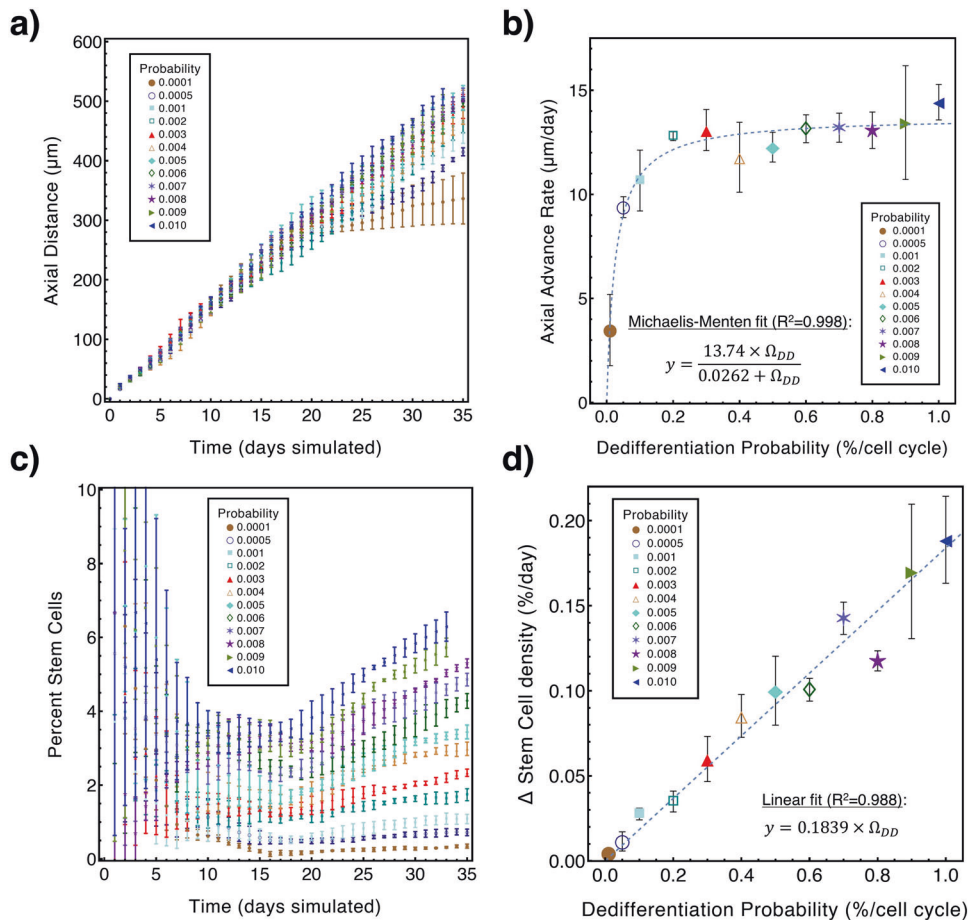


**Fig. 4 Axial invasion over time from selected simulations.** Each simulation run was input with a unique set of parameter values determined by LHS (see Table 1, Supplementary Table S1). **a** Without dedifferentiation, axial invasion is observed to arrest when the stem cell population in the leading edge becomes depleted; as can be seen by curves flattening at higher times. **b** Growth arrest is not seen when the dedifferentiation pathway is active. Only a subset ( $n = 10$ ) of simulations are shown in each panel for ease of visibility; matching colors do not indicate a relationship between curves in each panel.

**Table 3.** Linear regression analysis of global sensitivity analysis with the dedifferentiation pathway enabled.

Model output	$R^2$	Significant model parameters ( $p$ -value)
Axial advance rate ( $\mu\text{m}/\text{day}$ )	0.6940	$\theta_Q$ (<0.001); $\omega_p$ (<0.001)
Proliferation events per day	0.7017	$\theta_Q$ (<0.001); $\omega_p$ (0.001); $\lambda_C$ (0.002)
Axial calcification extent ( $\mu\text{m}$ )	0.7356	$\lambda_C$ (<0.001); $\omega_p$ (<0.001); $\Omega_{DD}$ (0.018)
Total cells calcified	0.9026	$\lambda_C$ (<0.001); $\theta_Q$ (<0.001); $\omega_p$ (0.003); $P_{max}$ (0.020)
Total cell count at end of simulation	0.9012	$\lambda_C$ (<0.001); $\theta_Q$ (<0.001); $\omega_p$ (<0.001); $\theta_{FGF}$ (0.001); $\Omega_{DD}$ (0.004); $\theta_E$ (0.005)
Total ER+ cells at end of simulation	0.8142	$\theta_E$ (<0.001); $\theta_Q$ (0.011); $P_{max}$ (0.028)
Total ER- cells at end of simulation	0.8805	$\theta_E$ (<0.001); $\lambda_C$ (<0.001); $\theta_Q$ (<0.001); $\omega_p$ (<0.001); $\theta_{FGF}$ (0.001)

Goodness-of-fit ( $R^2$ ) of the full multivariate model and parameters found to have significant effects on each model output are shown. All parameters in Table 1 marked with a dagger ( $\dagger$ ) were perturbed in a Latin hypercube sampling study; parameters not listed here were found to not have a significant effect on model outputs ( $p \geq 0.05$ ). Dedifferentiation probability (per cell cycle):  $\Omega_{DD}$ ; progenitor symmetric proliferation probability:  $\omega_p$ ; cell density threshold for quiescence:  $\theta_Q$ ; estrogen proliferation threshold:  $\theta_E$ ; FGF proliferation threshold:  $\theta_{FGF}$ ; proliferation cycles before differentiation:  $P_{max}$ ; cancer cell oxygen multiplier  $\lambda_C$ .



**Fig. 5** Local sensitivity analysis of dedifferentiation effects on axial advance rate and stem cell population. **a** DCIS axial distance per day shows reduced axial advance at lower dedifferentiation probabilities examined after progenitor mitosis cycle limits are first encountered ( $t = 15$  days), while higher dedifferentiation probabilities support consistent advance rates. **b** Average axial advance rate after start of dedifferentiation ( $t \geq 15$  days) was determined via linear regression analysis of data shown in **a**; values are listed in Supplementary Table S2. Michaelis-Menten best fit to data in **b** was found to be axial advance rate =  $(13.74 \times \Omega_{DD}) / (0.0262 + \Omega_{DD})$   $\mu\text{m}/\text{day}$  and is shown (blue, dashed;  $R^2 = 0.998$ ,  $p < 0.001$ ). **c** Density of stem cells in the total DCIS population; note that at early times, stochastic proliferation results in large percentage variations before steady state is achieved at later times. **d** Linear regression analysis to data shown in **c** of stem cell density after differentiation events begin ( $t \geq 15$  days; see Supplementary Table S2) reveals a linear relation between dedifferentiation probability and the average daily percent change ( $\Delta$ ) in stem cell density; linear regression revealed average daily change in stem cell percent density =  $0.189 \times \Omega_{DD}$  (blue, dashed;  $R^2 = 0.988$ ,  $p < 0.001$ ). All simulations were performed in triplicate; data shown as mean  $\pm$  standard deviation (error bars).

*Dedifferentiation overcomes stem cell dilution and terminal differentiation: effects on stem cell population.* We then sought to gain insights into the dynamics of stem cell populations across all global sensitivity analysis simulations, with and without activation of the dedifferentiation pathway. At the time of leading edge arrest in simulations without dedifferentiation, the DCIS stem cell population was found to average 0.55% (range: 0.037–5.5%) without dedifferentiation, well below literature reported values of 4–6% [17, 24]. Conversely, in simulations with the dedifferentiation pathway active, stem cell population was found to have a mean of 3.6% (range: 0.50–9.1%) across all simulations, which is in good agreement with the lower end of the literature-reported range.

The percentage of stem cells resulting from dedifferentiation events was found by multivariate linear regression analysis to be largely dependent on the total cell proliferation rate ( $p < 0.001$ ), with greater proliferation rates (and thus faster DCIS ductal advance rates,  $p < 0.001$ ; and total cell count,  $p < 0.001$ ) significantly associated with lower stem cell density in the final DCIS population. All model outputs examined are plotted against final stem cell count in Supplementary Fig. S3. The relation between stem cell density and progenitor cell phenotypes are also shown in Supplementary Fig. S4a, b; however, because dedifferentiation

occurs stochastically, these panels merely restate the trend observed in the total progenitor cell population. Stem cell density was also observed to be reduced when greater numbers of hypoxia-induced necrosis and subsequent cell calcification were observed ( $p < 0.001$ ; not shown).

*Isolating the effects of dedifferentiation using local sensitivity analysis: a deeper look.* In order to gain further insights into the effects of dedifferentiation in DCIS progression, we performed an additional study where only the dedifferentiation probability ( $\Omega_{DD}$ ) was perturbed while all other parameter values were held at baseline (Table 1, Supplementary Table S1). As an experiment, we expanded the range of dedifferentiation probabilities to 0.01–1.0% in the increments shown in Fig. 5, beyond the more biologically-relevant range listed in Table 1. This range was intentionally imbued with a finer discretization towards the lower end, so that the threshold where dedifferentiation is able to overcome growth arrest or slowing could be identified with higher precision. In this analysis, simulations were not truncated at the time of growth arrest, but instead were truncated at the same time simulated for all tests ( $t = 35$  days, later than growth arrest was observed without dedifferentiation (Fig. 4a)). Additionally, our

analysis of axial advance rates and rate of change of stem cell density (see Fig. 5) was applied only to data collected after our baseline time for dedifferentiation to occur (that is,  $t > 15$  days; Table 2) to focus on steady-state DCIS progression rates. Tests were performed in triplicate for each probability, and model outputs of DCIS axial invasion rate and stem cell percentages were assessed.

DCIS axial advance rate was notably reduced at low dedifferentiation probabilities, but remained consistent at larger dedifferentiation probabilities (Fig. 5a, b). Interestingly, we found that axial advance rate vs. dedifferentiation probability (Fig. 5b) demonstrates a Michaelis-Menten relationship, determined by regression analysis as axial advance rate =  $(13.74 \times \Omega_{DD}) / (0.0262 + \Omega_{DD})$   $\mu\text{m}/\text{day}$  (with  $V_{\text{max}} = 13.74$  and  $K_M = 0.0262$ ). Low dedifferentiation rates resulted in lower stem cell density within the DCIS mass (Fig. 5c), and the rate of stem cell population increase was observed to correlate linearly with dedifferentiation probability (Fig. 5d), quantified by best-fitting linear regression as average daily change in stem cell percent density =  $0.1839 \times \Omega_{DD}$ . A sensitivity score for the results shown in Fig. 5b, d was also determined as we have done previously [27]; this is further discussed in SI and shown in Supplementary Fig. S4.

## DISCUSSION

Our results revealed that, without the dedifferentiation pathway, DCIS behaviors studied in this work (including axial advance rate, proliferation events per day, and axial calcification extent) were most affected by proliferation cycles before differentiation, symmetric probabilities of progenitor cells (progenitor or differentiated daughter), and quiescence thresholds (Table 2, Supplementary Fig. S1). The progenitor population is reduced by proliferation events that give rise to differentiated daughters, terminal differentiation due to hitting the proliferation cycle threshold, and quiescence, which removes proliferative (i.e., progenitor) cells from the population contributing to DCIS growth. Downstream processes that then determine if this population is able to proliferate or not, such as estrogen and FGF thresholds, and cell cycle time before proliferation, showed significant effects on fewer model outputs (a detailed analysis on these effects in our model is presented in [27]). Increased oxygen consumption rates in the DCIS population primarily affected the extent of calcification and total cells lost to hypoxia-induced necrosis, but had minimal effects on DCIS progression rates. This further supports that progression is driven by the leading edge, while hypoxic regions are located at the center of the DCIS mass (this may be observed based on the location of the calcification seen in Fig. 3), and is in good agreement with our previous work [27]. We note that, although our model allows for similar analysis on other potential outputs (e.g., average mitosis cycle length, or detailed analysis of the effects of molecular signaling, which we have previously reported in [27]), we elected to focus on only those that allow for comparison of model output with literature-reported measurements, those pertinent to clinical practice (e.g., mammographic measurement of calcification), or those that are the focus of this study (e.g., growth arrest without dedifferentiation).

Our results reveal that dedifferentiation plasticity is vital for DCIS progression, and that exponential dilution of the bipotent stem-like population ultimately leads to growth arrest as the leading edge of the DCIS mass exhausts its proliferative potential and becomes terminally differentiated (Fig. 4a). We note that the stochasticity of our model allows for stem cell dilution and growth arrest to occur at different times between the two leading edges (e.g., Fig. 3), and axial advance may occur only in one direction for a time after arrest occurs in the opposite leading edge, thereby reducing the total average ductal advance rate and proliferation rate in this case.

Similar to the case without dedifferentiation, progenitor symmetric proliferation probability ( $\omega_p$ ) and quiescence density threshold ( $\theta_Q$ ) were most influential of the examined model outputs when the differentiation pathway was activated (Table 3). Cancer cell oxygen consumption multiplier ( $\lambda_c$ ) showed a significant effect on a larger number of examined model outputs when dedifferentiation was active than in the study without dedifferentiation. This is because at longer simulated times, dedifferentiation overcomes growth arrest in the leading edge, allowing for larger DCIS mass and greater total cell counts, and a larger total DCIS volume results in a similarly larger hypoxic (and thus calcified) volume. Interestingly, the dedifferentiation probability ( $\Omega_{DD}$ ) only showed significant effects on 2/7 total examined model outputs; this is fewer than most parameters examined (Table 3). However, it prevented growth arrest in the leading edge in all simulations performed, thereby demonstrating a necessary role in disease progression. Because lack of dedifferentiation was observed to eventually halt disease progression in all cases examined, this study provides notable evidence that it plays a critical role in the most important clinical aspect: allowing DCIS to overcome growth arrest and facilitating continued disease progression.

The observed trend of lower stem cell density with greater proliferation rates in the global sensitivity analysis (Supplementary Fig. S3) is likely due to exponential expansion in the non-stem (progenitor) cell population, which may double every cell cycle, while dedifferentiation events due to cell plasticity occur at a slower rate. In this study, proliferation rates may be reduced due to unfavorable conditions that are not related to dedifferentiation (such as unsatisfied molecular signaling thresholds), and stem cell density can increase due to continuing dedifferentiation (if  $\Omega_{DD} > 0$ ) even in the absence of proliferative events, counterbalancing this effect. However, in our follow-up study wherein only dedifferentiation probabilities were perturbed while all other variables were held at baseline (and thus confounding factors were removed), dedifferentiation rates demonstrated clear relationships between ductal advance rates (Michaelis-Menten) and total stem cell population (linear) (Fig. 5b, d). These revealed that the maximum theoretical axial advance rate per day under the examined model conditions was close to 14  $\mu\text{m}/\text{day}$  and half-max advance rates occur at a dedifferentiation probability = 0.0262 per cell cycle (Fig. 5b), while the daily percent change in stem cell density was affected by the dedifferentiation probability according to the linear relation  $0.1839 \times \Omega_{DD}$  (Fig. 5d). Reduced stem cell density in the cases of greater cell calcification is likely due to the fact that the center of the DCIS mass (where necrosis leads to hypoxia and calcification) is composed of relatively older cells than at the leading edge. Stochastic dedifferentiation has had more opportunities to occur in these older cells (because they have experienced more cell cycles), and thus stem cell density is likely higher in those regions closer to the site of DCIS initiation. When the region of highest stem cell density is removed from the cell population, the average stem cell density in the remaining population is reduced.

In this study, we implemented dedifferentiation as a purely stochastic event, but it has been reported that dedifferentiation is related to hypoxia in DCIS as well [18, 21]. In upcoming work, we will modify our dedifferentiation mechanism to include a mechanistic hypoxia-based component in order to gain further insights on the mechanistic relationships underlying these phenomena. We are also investigating innate molecular pathways that play natural tumor-suppressing roles in the early stages of DCIS, such as p63 and metalloproteinase-8 [40], which are derived from the mature myoepithelial layer already included in our model. Further, we are collecting in-house data from murine models to identify and calibrate additional key molecular pathways to expand the biological accuracy of the model, and to gain additional insights into molecular effects on the earliest stages of



DCIS. Indeed, it has been reported that dedifferentiation is associated with changes in genetic and molecular expression, and also cell conformation, and it is likely that some of these may be distinct from stemness derived directly from a stem-type mother cell [41, 42]. We have now obtained some DCIS RNAseq data, and spatial single cell sequencing is ongoing in our laboratory. These were not included in the current model, and our study is limited in its ability to shed light onto this interesting problem, but it is our hope that this experimental work will reveal specific pathways for inclusion that will allow us to focus future studies on simulated measures of these quantities. We are currently implementing our own mouse model of dedifferentiation in DCIS (and the related but distinct EMT mechanism) based on previously published mouse models [43–45], which we believe will allow additional validation of the results presented and reveal new molecular pathways to be included for study in future model iterations. Ultimately, we hope that this work will help to reveal new insights into potential clinical strategies to optimize emerging cancer stem cell targeted therapeutics, and to predict the risk of DCIS progression into invasive breast cancer. In combination with our upcoming experimental work, we believe that this may lead to identification of one of more pathways that may be targeted to diminish dedifferentiation events in order to halt disease progression in stage zero, and which could help guide drug design in the future. Moreover, we believe that the underlying software we are developing will lead to a tool that enables the study of the effects and interplay of phenotypic hierarchies and molecular signaling in other solid cancer types, leading to new insights into their developmental biology and strategies for therapy.

### Reporting summary

Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

### DATA AVAILABILITY

All data generated or analyzed during this study are included in this published article and its Supplementary Information files.

### REFERENCES

1. U.S. Cancer Statistics Working Group. United States Cancer Statistics: Leading Cancer Cases and Deaths, All Races/Ethnicities, Male and Female, 2017. In: Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute, 2020.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin.* 2018;68:7–30.
3. Allegra CJ, Aberle DR, Ganschow P, Hahn SM, Lee CN, Millon-Underwood S, et al. NIH state-of-the-science conference statement: diagnosis and management of ductal carcinoma in situ (DCIS). *NIH Consens State—Sci statements.* 2009;26:1–27.
4. Society AC Cancer Facts & Figures 2017. In: Special Section: Rare Cancers in Adults: American Cancer Society, Inc., 2017.
5. Duffy SW, Dibden A, Michalopoulos D, Offman J, Parmar D, Jenkins J, et al. Screen detection of ductal carcinoma in situ and subsequent incidence of invasive interval breast cancers: a retrospective population-based study. *Lancet Oncol.* 2016;17:109–14.
6. Rios AC, Fu NY, Lindeman GJ, Visvader JE. In situ identification of bipotent stem cells in the mammary gland. *Nature.* 2014;506:322–7.
7. Villadsen R, Fridriksson AJ, Rønnow-Jessen L, Gudjonsson T, Rank F, LaBarge MA, et al. Evidence for a stem cell hierarchy in the adult human breast. *J Cell Biol.* 2007;177:87–101.
8. Visvader JE. Cells of origin in cancer. *Nature.* 2011;469:314–22.
9. Visvader JE, Lindeman GJ. Cancer stem cells: current status and evolving complexities. *Cell Stem Cell.* 2012;10:717–28.
10. Gjorevski N, Nelson CM. Integrated morphodynamic signalling of the mammary gland. *Nat Rev Mol Cell Biol.* 2011;12:581–93.
11. André F, Cortés J. Rationale for targeting fibroblast growth factor receptor signaling in breast cancer. *Breast Cancer Res Treat.* 2015;150:1–8.
12. Moerkens M, Zhang Y, Wester L, van de Water B, Meerman JH. Epidermal growth factor receptor signalling in human breast cancer cells operates parallel to estrogen receptor  $\alpha$  signalling and results in tamoxifen insensitive proliferation. *BMC Cancer.* 2014;14:283.
13. Friedmann-Morvinski D, Verma IM. Dedifferentiation and reprogramming: origins of cancer stem cells. *EMBO Rep.* 2014;15:244–53.
14. Friedmann-Morvinski D, Bushong EA, Ke E, Soda Y, Marumoto T, Singer O, et al. Dedifferentiation of neurons and astrocytes by oncogenes can induce gliomas in mice. *Sci (N. Y., N. Y.).* 2012;338:1080–4.
15. Schwitalla S, Fingerle AA, Cammareri P, Nebelsiek T, Göktuna SI, Ziegler PK, et al. Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties. *Cell.* 2013;152:25–38.
16. Landsberg J, Kohlmeyer J, Renn M, Bald T, Rogava M, Cron M, et al. Melanomas resist T-cell therapy through inflammation-induced reversible dedifferentiation. *Nature.* 2012;490:412–6.
17. Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell.* 2008;133:704–15.
18. Helczynska K, Kronblad A, Jögi A, Nilsson E, Beckman S, Landberg G, et al. Hypoxia promotes a dedifferentiated phenotype in ductal breast carcinoma in situ. *Cancer Res.* 2003;63:1441–4.
19. Ischenko I, Zhi J, Moll UM, Nemaierova A, Petrenko O. Direct reprogramming by oncogenic Ras and Myc. *Proc Natl Acad Sci USA.* 2013;110:3937–42.
20. Vermeulen L, De Sousa EMF, van der Heijden M, Cameron K, de Jong JH, Borovski T, et al. Wnt activity defines colon cancer stem cells and is regulated by the microenvironment. *Nat Cell Biol.* 2010;12:468–76.
21. Jögi A, Vaapil M, Johansson M, Pählman S. Cancer cell differentiation heterogeneity and aggressive behavior in solid tumors. *Ups J Med Sci.* 2012;117:217–24.
22. Chaffer CL, Marjanovic ND, Lee T, Bell G, Kleer CG, Reinhardt F, et al. Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity. *Cell.* 2013;154:61–74.
23. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell.* 2006;126:663–76.
24. Chaffer CL, Brueckmann I, Scheel C, Kaestli AJ, Wiggins PA, Rodrigues LO, et al. Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. *Proc Natl Acad Sci USA.* 2011;108:7950–5.
25. Butner JD, Cristini V, Zhihui W. Development of a three dimensional, multiscale agent-based model of ductal carcinoma in situ. *Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Conf.* 2017;2017:86–89.
26. Butner JD, Chuang YL, Simbawa E, Al-Fhaid AS, Mahmoud SR, Cristini V, et al. A hybrid agent-based model of the developing mammary terminal end bud. *J Theor Biol.* 2016;407:259–70.
27. Butner JD, Fuentes D, Ozpolat B, Calin GA, Zhou X, Lowengrub J, et al. A multiscale agent-based model of ductal carcinoma in situ. *IEEE Trans bio-Med Eng.* 2020;67:1450–61.
28. Butner JD, Cristini V, Wang Z. Understanding ductal carcinoma in situ invasion using a multiscale agent-based model(). *Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Conf.* 2018;2018:5846–9.
29. Shehata M, Waterhouse PD, Casey AE, Fang H, Hazelwood L, Khokha R. Proliferative heterogeneity of murine epithelial cells in the adult mammary gland. *Commun Biol.* 2018;1:111.
30. Paine I, Chauviere A, Landua J, Sreekumar A, Cristini V, Rosen J, et al. A geometrically-constrained mathematical model of mammary gland ductal elongation reveals novel cellular dynamics within the terminal end bud. *PLOS Computational Biol.* 2016;12:e1004839–e1004839.
31. Dogra P, Butner JD, Ruiz Ramirez J, Chuang YL, Noureddine A, Jeffrey Brinker C, et al. A mathematical model to predict nanomedicine pharmacokinetics and tumor delivery. *Comput Struct Biotechnol J.* 2020;18:518–31.
32. Dogra P, Ruiz-Ramirez J, Sinha K, Butner JD, Pelaez MJ, Rawat M, et al. Innate immunity plays a key role in controlling viral load in COVID-19: mechanistic insights from a whole-body infection dynamics model. *ACS Pharmacol Transl Sci.* 2021;4:248–65.
33. Wang Z, Birch CM, Sagotsky J, Deisboeck TS. Cross-scale, cross-pathway evaluation using an agent-based non-small cell lung cancer model. *Bioinformatics.* 2009;25:2389–96.
34. Wang Z, Bordas V, Deisboeck TS. Identification of critical molecular components in a multiscale cancer model based on the integration of Monte Carlo, resampling, and ANOVA. *Front Physiol.* 2011;2:35.
35. Wang Z, Deisboeck TS, Cristini V. Development of a sampling-based global sensitivity analysis workflow for multiscale computational cancer models. *IET Syst Biol.* 2014;8:191–7.
36. Brocato TA, Brown-Glaberman U, Wang Z, Selwyn RG, Wilson CM, Wyckoff EF, et al. Predicting breast cancer response to neoadjuvant chemotherapy based on tumor vascular features in needle biopsies. *JCI Insight.* 2019;4:1–12.e126518.

37. Butner JD, Wang Z, Elganainy D, Al Feghali KA, Plodinec M, Calin GA, et al. A mathematical model for the quantification of a patient's sensitivity to checkpoint inhibitors and long-term tumour burden. *Nat Biomed Eng.* 2021;5:297–308.
38. Brocato TA, Coker EN, Durfee PN, Lin YS, Townson J, Wyckoff EF, et al. Understanding the connection between nanoparticle uptake and cancer treatment efficacy using mathematical modeling. *Sci Rep.* 2018;8:7538.
39. Dogra P, Ramirez JR, Butner JD, Pelaez MJ, Chung C, Hooda-Nehra A, et al. Translational modeling identifies synergy between nanoparticle-delivered miRNA-22 and standard-of-care drugs in triple-negative breast cancer. *Pharm Res.* 2022;39:511–28.
40. Sarper M, Allen MD, Gomm J, Haywood L, Decock J, Thirkettle S, et al. Loss of MMP-8 in ductal carcinoma in situ (DCIS)-associated myoepithelial cells contributes to tumour promotion through altered adhesive and proteolytic function. *Breast Cancer Res.* 2017;19:33.
41. Yao Y, Wang C. Dedifferentiation: inspiration for devising engineering strategies for regenerative medicine. *npj Regenerative Med.* 2020;5:14.
42. Schymik B, Buerger H, Krämer A, Voss U, van der Groep P, Meinerz W, et al. Is there 'progression through grade' in ductal invasive breast cancer? *Breast Cancer Res Treat.* 2012;135:693–703.
43. Molyneux G, Geyer FC, Magnay FA, McCarthy A, Kendrick H, Natrajan R, et al. BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. *Cell Stem Cell.* 2010;7:403–17.
44. Guo Q, Li VZ, Nichol JN, Huang F, Yang W, Preston SEJ, et al. MNK1/NODAL signaling promotes invasive progression of breast ductal carcinoma in situ. *Cancer Res.* 2019;79:1646–57.
45. Xu Y, Lee DK, Feng Z, Xu Y, Bu W, Li Y, et al. Breast tumor cell-specific knockout of Twist1 inhibits cancer cell plasticity, dissemination, and lung metastasis in mice. *Proc Natl Acad Sci USA.* 2017;114:11494–9.
46. Giebel B, Zhang T, Beckmann J, Spanholtz J, Wernet P, Ho AD, et al. Primitive human hematopoietic cells give rise to differentially specified daughter cells upon their initial cell division. *Blood.* 2006;107:2146–52.
47. Gandhi A, Holland PA, Knox WF, Potten CS, Bundred NJ. Evidence of significant apoptosis in poorly differentiated ductal carcinoma in situ of the breast. *Br J Cancer.* 1998;78:788–94.
48. Wagner BA, Venkataraman S, Buettner GR. The rate of oxygen utilization by cells. *Free Radic Biol Med.* 2011;51:700–12.

## ACKNOWLEDGEMENTS

This research has been supported in part by the National Science Foundation Grant DMS-1930583 (VC, ZW), the National Institutes of Health (NIH) Grants 1R01CA253865 (VC, BO, ZW), 1U01CA196403 (VC, ZW), 1U01CA213759 (VC, BO, ZW), 1R01CA226537 (VC, ZW), 1R01CA222007 (VC, BO, GAC, ZW), and U54CA210181 (SHC, VC, ZW). Work in GAC's laboratory is also supported by NCI grant 1R01 CA182905-01, NIGMS grant 1R01GM122775-01, DoD Idea Award W81XWH-21-1-0030, a Team DOD grant in Gastric Cancer W81XWH-21-1-0715, a Chronic Lymphocytic Leukemia Moonshot Flagship project, a CLL Global Research Foundation 2019 grant, a CLL Global Research Foundation 2020 grant, the G. Harold & Leila Y. Mathers Foundation, two grants from

Torrey Coast Foundation, an Institutional Research Grant and Development Grant associated with the Brain SPORE 2P50CA127001. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## AUTHOR CONTRIBUTIONS

ZW conceived the study. JDB, ZW conceptualized the idea and designed the model. JDB developed the model. JDB, PD, JRR, SN and MP collected the data. JDB, PD, CC, JRR, SN, MP, XL, PYP, SHC, VC, BO, GC and ZW interpreted the data. JDB, PD, CC, VC, BO, GC and ZW performed the model analysis. JDB, PD, CC, JRR, SN, MP, XL, PYP, SHC, VC, BO, GC and ZW wrote or edited the manuscript.

## COMPETING INTERESTS

GAC is the scientific founder of Ithax Pharmaceuticals. The remaining authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41419-022-04939-x>.

**Correspondence** and requests for materials should be addressed to Zhihui Wang.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022