Cell

Stochastic State Transitions Give Rise to Phenotypic Equilibrium in Populations of Cancer Cells

Piyush B. Gupta,[1,6,*] Christine M. Fillmore,[2] Guozhi Jiang,[1] Sagi D. Shapira,[1] Kai Tao,[3] Charlotte Kuperwasser,[2,3] and Eric S. Lander[1,4,5,*]

[1]Broad Institute, Cambridge, MA 02142, USA
[2]Department of Anatomy and Cellular Biology, Sackler School of Graduate Biomedical Sciences, Tufts University School of Medicine, 136 Harrison Avenue, Boston, MA 02111, USA
[3]Molecular Oncology Research Institute, Tufts Medical Center, Boston, MA 02111, USA
[4]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142, USA
[5]Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA
[6]Present address: Department of Biology, Massachusetts Institute of Technology, and Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA
*Correspondence: pgupta@wi.mit.edu (P.B.G.), lander@broadinstitute.org (E.S.L.)
DOI 10.1016/j.cell.2011.07.026

SUMMARY

Cancer cells within individual tumors often exist in distinct phenotypic states that differ in functional attributes. While cancer cell populations typically display distinctive equilibria in the proportion of cells in various states, the mechanisms by which this occurs are poorly understood. Here, we study the dynamics of phenotypic proportions in human breast cancer cell lines. We show that subpopulations of cells purified for a given phenotypic state return towards equilibrium proportions over time. These observations can be explained by a Markov model in which cells transition stochastically between states. A prediction of this model is that, given certain conditions, any subpopulation of cells will return to equilibrium phenotypic proportions over time. A second prediction is that breast cancer stem-like cells arise de novo from non-stem-like cells. These findings contribute to our understanding of cancer heterogeneity and reveal how stochasticity in single-cell behaviors promotes phenotypic equilibrium in populations of cancer cells.

INTRODUCTION

The regulation of cell-state decisions is critical for the survival of living systems. In unicellular organisms, cell-state changes occur in response to environmental stressors or nutrient changes. Direct sensing of environmental stimuli often involves stochastic cell-fate decisions, modulated by random noise in gene expression (Süel et al., 2006, 2007). Such probabilistic behavior has been shown to be advantageous in certain environmental conditions (Kussell and Leibler, 2005; Thattai and van Oudenaarden, 2004; Wolf et al., 2005). Noisy gene-expression levels can also stochastically influence cell-state decisions in eukaryotes (Di Talia et al., 2007). However, less is known about the role that stochasticity might play in regulating cell-state equilibria in populations of cells.

Cell-state dynamics are of particular significance in tumor pathobiology. Even within individual tumors, cancer cells frequently exist in any of several possible phenotypic states. Cancer cells in distinct phenotypic states often exhibit important differences in functional properties. For example, subpopulations of stem-like cancer cells with increased tumor-seeding ability and drug resistance have been identified in a variety of tumor types (Al-Hajj et al., 2003; Lapidot et al., 1994; Li et al., 2007; Singh et al., 2004; Smalley and Ashworth, 2003; Stingl and Caldas, 2007). The proportion of cancer cells in the various states is related to both tumor type and grade (Chiou et al., 2008; Harris et al., 2008). Additionally, because anticancer therapies preferentially kill specific cancer cell states, treatment can result in selective changes in phenotypic proportions within tumors (Creighton et al., 2009; Gupta et al., 2009; Li et al., 2008; Woodward et al., 2007). Understanding how cancer cell states coexist and evolve within tumors is of fundamental interest and could facilitate the development of more effective therapies.

Phenotypic equilibrium in cell-state proportions is observed in vivo and in cell lines adapted to in vitro culture. Under fixed conditions, both normal and cancerous epithelial lines display stable proportions of cells in stem-like, basal, or luminal states during propagation in culture. The mechanisms that stabilize phenotypic proportions within cellular populations remain unclear. Two general classes of mechanisms can be envisioned: (1) in the absence of interconversion between states, equilibrium proportions could be maintained through intercellular signals that modulate the proliferation rates of distinct states and (2) proliferation rates remain equal, but cancer cells could interconvert between different states in a manner that maintains equilibrium cell-state proportions.

Here, we study the mechanisms that underlie phenotypic diversity in populations of cancer cells. We develop and validate a quantitative Markov model of phenotypic transitions that predicts evolution toward equilibrium proportions in cancer cell populations. The Markov model makes several unanticipated predictions about cell-state transitions and dynamics in cancer cell populations. In addition, the proposed model is useful in characterizing the effects of genetic and chemical perturbations on phenotypic proportions in cancer cell populations.

RESULTS

Characterization of Luminal, Basal, and Stem-like Differentiation States in Breast Cancer Cells

To study cell-state dynamics in cancer cells, we used fluorescence-activated cell sorting (FACS) to isolate three mammary epithelial cell states that have been previously defined and characterized using cell-surface markers: stem-like ($CD44^{hi}CD24^{neg}EpCAM^{lo}$), basal ($CD44^{hi}CD24^{neg}EpCAM^{neg}$), and luminal ($CD44^{lo}CD24^{hi}EpCAM^{hi}$) (Fillmore and Kuperwasser, 2008; Shipitsin et al., 2007). Using this system, we isolated each of these cellular fractions from two human breast cancer lines derived from primary tumors (SUM159 and SUM149), resulting in subpopulations that exhibited significant differences in morphology in culture (Figures 1A and 1B). To confirm that these markers indeed define cells in the expected cell-differentiation states, we collected global gene expression data from sorted $CD44^{hi}CD24^{neg}EpCAM^{lo}$ (stem-like), $CD44^{hi}CD24^{neg}EpCAM^{neg}$ (basal), or $CD44^{lo}CD24^{hi}EpCAM^{hi}$ (luminal) subpopulations of the SUM149 and SUM159 breast cancer lines. We next applied gene set enrichment analysis (GSEA) (Mootha et al., 2003; Subramanian et al., 2005) to determine whether the expression levels of genes previously reported to be associated with stem, basal, and luminal differentiation were specifically enriched in the relevant sorted subpopulations (Assou et al., 2007; Blackman et al., 2005; Böcker et al., 2002; Brambrink et al., 2008; Bussolati et al., 1996; Charafe-Jauffret et al., 2006; Dontu et al., 2003; Drabsch et al., 2007; Dressman et al., 2001; Gill et al., 2001; Gudjonsson et al., 2002; Jones et al., 2004; Kariagina et al., 2007; Liu et al., 2006; Lu et al., 2008; Monaghan et al., 1995; Parisi et al., 2008; Perry et al., 2007; Phinney et al., 2005; Rasbridge et al., 1993; Reis-Filho et al., 2006; Sansone et al., 2007; Shipitsin et al., 2007; Sleeman et al., 2006; Sørlie et al., 2006) (Table S1 available online). We observed that in both the SUM159 and SUM149 lines, the expression of genes associated with luminal differentiation was indeed enriched in the $CD44^{lo}CD24^{hi}EpCAM^{hi}$ subpopulation (Figures 1C ans 1D and Table S1). For example, the luminal markers *CD24*, *claudin 1* (*CLDN1*), and *cytokeratins 6B* and *8* were all specifically upregulated in $CD44^{lo}CD24^{hi}EpCAM^{hi}$ cells (Figure 1C and Table S1). Similarly, we found a significant enrichment in the expression of genes associated with basal differentiation in the $CD44^{hi}CD24^{neg}EpCAM^{neg}$ subpopulation (Figures 1C and 1D and Table S1); these genes included vimentin (*VIM*), *Zeb1* (*TCF8*), membrane metallo-endopeptidase (*CALLA/CD10*), *decorin* (*DCN*), and plasminogen activator inhibitor type 1 (*SERPINE2*). Genes implicated in the self-renewal of embryonic stem cells, including *EZH2*, *KLF5*, *KLF4*, and *SOX9* (Brambrink et al., 2008; Ivanova et al., 2006; Parisi et al., 2008), were significantly enriched in their expression in the $CD44^{hi}CD24^{neg}EpCAM^{lo}$ fraction (Figures 1C and 1D and Table S1).

The expression levels of genes associated with basal differentiation in the $CD44^{hi}CD24^{neg}EpCAM^{neg}$ fraction was further confirmed via quantitative RT-PCR. Consistent with the microarray data, we observed that $CD44^{hi}CD24^{neg}EpCAM^{neg}$ cells from SUM149 and SUM159 populations expressed elevated levels of *Vimentin*, *N-cadherin*, and *Zeb1* compared to unsorted cells (Figure 1E). In contrast, $CD44^{lo}CD24^{hi}EpCAM^{hi}$ cells, which are in a luminal state, exhibited a reduction in the expression levels of these basal genes and elevated levels of the luminal epithelial marker, E-cadherin (Figure 1E). Collectively, these data indicated that, as previously reported, the cell-surface expression of CD24, CD44, and EpCAM does indeed allow for an accurate fractionation of cells in distinct differentiation states.

Cell-State Dynamics in Breast Cancer Populations

We next determined the proportions of the individual cell-states in the SUM159 and SUM149 lines and found that they contained distinct cell-state proportions: SUM159 populations exhibited predominantly basal (B) differentiation with minority subpopulations that were stem-like (S) or luminal (L) (proportion of B, S, and L = 97.3%, 1.9%, and 0.62%, respectively; Figure 2B); in contrast, SUM149 populations exhibited predominantly luminal differentiation with stem-like and basal minority subpopulations (B, S, and L = 3.3%, 3.9%, and 92.8%, respectively; Figure 2B).

Using FACS, we sorted stem-like, basal, or luminal cells from the SUM159 and SUM149 lines (resulting in subpopulations that were at least 96% pure based on reanalysis immediately after sorting). By isolating relatively pure subpopulations of cells in a given differentiation state and allowing them to expand in culture, we could monitor how cell-state dynamics evolve over time (Figure 2A). After 6 days of growth in culture, we assessed the proportions of stem-like, basal, and luminal cells (Figure 2C). For each isolated subpopulation, we observed a rapid progression toward equilibrium proportions (Figure 2C). Two lines of evidence indicated that this progression was not due to differential growth rates of cells in the basal, stem-like, or luminal states but rather to interconversion between states. First, we observed no difference in the proliferation rates of the stem-like, basal, or luminal subpopulations sorted from either SUM159 or SUM149 (Figure S1). Second, given the purity of the sorted populations and the rapid rate of return to equilibrium proportions, some minority subpopulations would need to divide more than three times per day to achieve the observed proportions through differential growth alone. Such a high proliferation rate is implausible because even the most rapidly dividing human cells—embryonic stem cells—require at least 24 hr to complete a proliferation cycle (Cowan et al., 2004).

Quantitative Markov Model of Cell-State Interconversion

Given that interconversion between cell states was occurring, we modeled these observations as a stochastic process in which cancer cells transition randomly between states with each proliferative cycle. We additionally made the assumption that cell transitions follow a Markov process—that is, transition

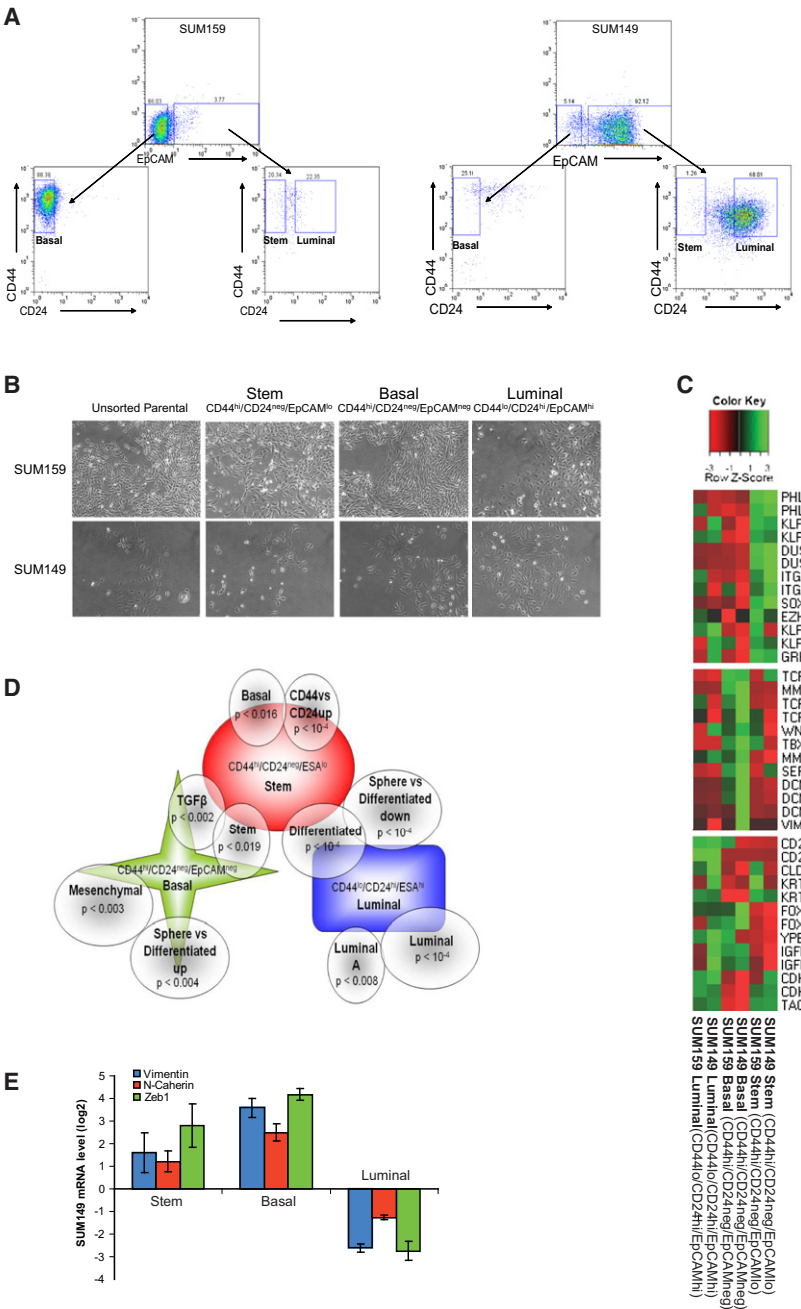

Figure 1. Isolation of Stem-like, Luminal, and Basal Cell-Differentiation States in Human Breast Cancer Lines

(A) Flow cytometry gating strategy for isolating $CD44^{lo}CD24^{hi}EpCAM^{hi}$, $CD44^{hi}CD24^{neg}EpCAM^{neg}$, and $CD44^{hi}CD24^{neg}EpCAM^{lo}$ subpopulations from SUM149 and SUM159 human breast cancer lines.

(B) Morphology of SUM149 and SUM159 colonies cultured immediately after cell sorting. Representative phase-contrast bright-field images of colonies from sorted $CD44^{lo}/CD24^{hi}/EpCAM^{hi}$ cells comprised of epithelial islands exhibiting cobblestone morphology and tight cell-cell junctions, consistent with a luminal cell phenotype. $CD44^{hi}/CD24^{neg}/EpCAM^{neg}$ cells from colonies exhibiting spindloid features reminiscent of basal cell morphology. $CD44^{hi}/CD24^{neg}/EpCAM^{lo}$ colonies are able to exhibit mixed morphologies representing luminal and basal phenotypes, consistent with differentiation of stem-like cells.

(C) Messenger RNA (mRNA) was extracted from the sorted luminal, basal, and stem-like subpopulations and Affymetrix gene chips were then hybridized to pooled mRNA samples (between three to five independent sorts were pooled per marker profile combination and cell line). The heat map displays differentially expressed genes that can distinguish between the three cell states. Genes having multiple instances correspond to distinct probe sets.

(D) Schematic depicting the results of GSEA performed to determine whether genes previously published to be associated with luminal, mesenchymal or stem cell differentiation are upregulated in sorted $CD44^{lo}CD24^{hi}EpCAM^{hi}$ (luminal), $CD44^{hi}CD24^{neg}EpCAM^{neg}$ (basal), and $CD44^{hi}CD24^{neg}EpCAM^{lo}$ (stem-like) subpopulations. The ten gene lists analyzed are described in Table S1 and include basal, CD44vsCD24up, TGFβ, stem, differentiated, sphere versus differentiated down, luminal, luminal A, sphere versus differentiated up, and mesenchymal. The gene lists are shown superimposed upon the cellular subpopulation(s) in which they were enriched, together with an associated p value calculated by permutation testing. Gene lists with corresponding references are described in Table S1.

(E) Quantitative RT-PCR analysis of Vimentin, N-cadherin, E-cadherin, and Zeb1 expression in SUM149 or SUM159 sorted cell populations relative to expression in the parental (unsorted) cell line. Data are represented as the average ΔΔCt ± standard error of the mean (SEM). n = 3. See also Table S1.

probabilities depend only on a cell's current state, not on its prior states (see the Experimental Procedures). Under a Markov model, it is possible to use data from short-term cell culture of isolated subpopulations to infer the probabilities of transition between any two cell states (for example, $P_{B\rightarrow S}$ denotes the probability per generation of a basal cell transitioning into a stem-like state). By revealing whether certain transitions are allowed or forbidden, these inferred probabilities could, in principle, provide significant insights into the cell-state dynamics in cell lines.

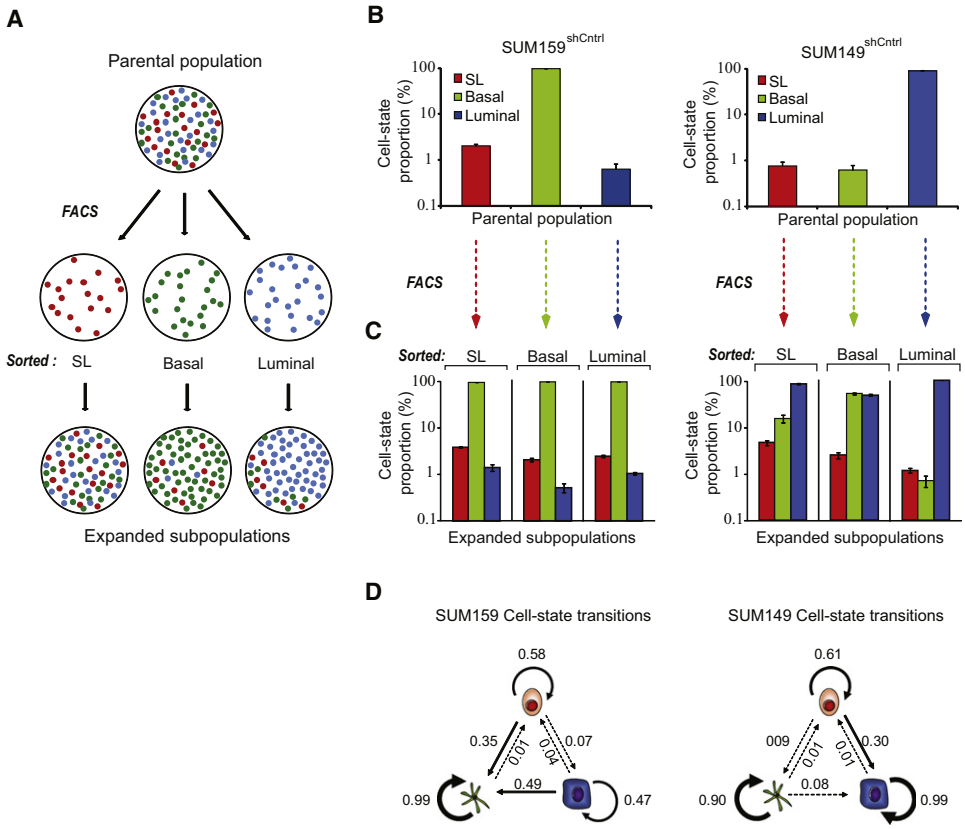

Figure 2. Determination of Breast Cancer Cell-State Transition Probabilities from Population Cell-State Proportions

(A) Schematic of experimental procedure used to determine cell-state transition dynamics.

(B) Proportions of cell-states in parental SUM159^{shCntrl} and SUM149^{shCntrl} breast cancer lines.

(C) Cellular subpopulations in stem-like (SL), basal, or luminal states were isolated by FACS with antibodies directed against the CD44, CD24, and EpCAM cell-surface antigens. Bar charts show the proportion of cells in each cell-differentiation state as assessed by FACS after in vitro culture for 6 days.

(D) Lineage hierarchies for the SUM159^{shCntrl} and SUM149^{shCntrl} lines were calculated from the data in (C). The corresponding cell-state transition probabilities for each cell line are shown. Solid arrows denote transition probabilities greater than 0.1. Dashed arrows denote transition probabilities between 0.01 and 0.1. See also Figure S1.

The transition probabilities for the SUM159 and SUM149 lines inferred from the experiment above (Figure 2D and Table 1) revealed several interesting similarities and distinctions between the cell lines:

(1) For both lines, we found that stem-like cells could either self-renew or transition into either of the basal or luminal states (Figure 2D and Table 1). While the self-renewal probabilities are comparable ($P_{S\rightarrow S} = 0.58, 0.61$), stem-like cells were more likely to transition into a luminal state in the SUM149 line ($P_{S\rightarrow L} = 0.30$, $P_{S\rightarrow B} = 0.09$), but into a basal state in the SUM159 line ($P_{S\rightarrow L} = 0.07$, $P_{S\rightarrow B} = 0.35$).

(2) In both cell lines, basal cells exhibited a high probability of self-renewing divisions ($P_{B\rightarrow B} = 0.99, 0.90$). In contrast, the behavior of luminal cells differed strongly between the cell lines. Luminal cells in the SUM149 line displayed a high probability of self-renewal ($P_{L\rightarrow L} = 0.99$), while luminal cells in the SUM159 line exhibited roughly equal probabilities of either remaining in the luminal state or transitioning to basal state ($P_{L\rightarrow L} = 0.47$; $P_{L\rightarrow B} = 0.49$). These distinctions help explain the higher proportion of basal cells observed at equilibrium in the SUM159 line relative to the SUM149 line.

Validation of Markov Model Predictions In Vitro and In Vivo

The inferred Markov transition probabilities make it possible to quantitatively predict how a population of cells evolves over time given the initial proportion of cell states. Examples are shown in Figure 3 for several initial cell state proportions of SUM159 populations.

The model makes several unanticipated predictions. First, even though a subpopulation of sorted stem-like cells from SUM159 has a low proportion of luminal cells both immediately after sorting and at 6 days postsort (0.6%), the model predicts that the proportion of luminal cells will actually show a sharp transient rise to ~7.3% at 1 day postsort (arrow in Figure 3). We tested this prediction and indeed observed an increase in the proportion of luminal cells (to 6.5%) at 1 day postsort. This unexpected prediction could not have been made without a quantitative model of the underlying cell-state dynamics.

Table 1. Cell-State Transition Probabilities for Control or TBX3-Inhibited Human Breast Cancer Cells

	Control			shTBX3		
Transition Probabilities	S	B	L	S	B	L
SUM159						
Stem	0.58	0.35	0.07	0.50	0.49	0.01
Basal	0.01	0.99	0.00	0.01	0.99	0.00
Luminal	0.04	0.49	0.47	0.06	0.16	0.78
SUM149						
Stem	0.61	0.09	0.30	0.63	0.07	0.30
Basal	0.01	0.90	0.08	0.02	0.86	0.12
Luminal	0.01	0.00	0.99	0.01	0.00	0.99

S, B and L correspond to stem-like, basal, and luminal states, respectively. The rows and columns correspond to initial (pretransition) and final (posttransition) states, respectively. Transition probabilities are shown per cell division.

A second striking prediction of the model is that basal and luminal cells have a non-zero probability of transitioning to a stem-like state—that is, that cancer stem-like cells can arise from non-stem-like cells. This notion stands in stark contrast to the classical understanding of the concept of "stem cells" in normal tissues, which posits the existence of a rigid lineage-hierarchy in which stem cells can give rise to nonstem cells but not vice versa.

To test this prediction, we evaluated the ability of sorted SUM159 subpopulations to seed tumors in mice, either after their initial isolation or after propagation in culture. Consistent with previous observations, the stem-like fraction could efficiently seed tumors, but neither the luminal nor basal fraction was capable of doing so (Table 2). The lack of tumor-seeding potential of luminal and basal subpopulations could be due to either an inherent inability to give rise to cancer stem cells (CSCs) in vivo or simply an inability to survive long enough at the site of implantation to do so. We speculated that admixture with irradiated carrier cells could allow sorted subpopulations to survive longer in vivo following injection, as has been previously shown in the context of hematopoietic reconstitution (Bonnet et al., 1999).

Accordingly, we coinoculated with GFP-labeled irradiated parental SUM159 carrier cells. Under these conditions, all three fractions (stem-like, basal, and luminal) were equally capable of efficiently seeding tumors (Table 2). In contrast, irradiated parental SUM159 populations failed to seed tumors on their own, indicating that irradiation had ablated their tumor-seeding ability. Examination of the tumors arising from basal and luminal subpopulations mixed with irradiated carrier cells revealed the presence of significant numbers of stem-like cells (Table 2). Moreover, the proportions of basal, stem-like, and luminal cells in tumors were comparable irrespective of the sorted subpopulation used to seed the tumor (Table 2). Collectively, these results demonstrated that the luminal and basal fractions can indeed regenerate functional stem-like cells in vivo and suggested that convergence toward equilibrium cell-state proportions could be occurring due to cell-state interconversion within tumors.

The ability of basal and luminal cells to regenerate stem-like cells de novo is consistent with our finding that basal and luminal subpopulations return to equilibrium cell-state proportions over time both in vitro and in vivo. More generally, it is a basic mathematical theorem (Perron-Frobenius [Frobenius, 1912; Perron, 1907]) that any finite Markov process satisfying mild conditions (essentially that there is a path of transitions between any two states) necessarily returns to a unique stationary distribution regardless of the starting state. In the context of cancer populations with nonzero transition probabilities, a Markov process would cause cellular subpopulations to return to equilibrium proportions upon propagation.

Influence of Genetic Perturbation on Cell-State Dynamics

We next examined whether the study of cell-state transition probabilities could help elucidate how genetic perturbations affect cell-state dynamics in cancer populations. We studied the inhibition of the *TBX3* gene, which encodes a transcription factor that is overexpressed in breast cancers (Yarosh et al., 2008) and regulates differentiation in multiple cell types (Fillmore

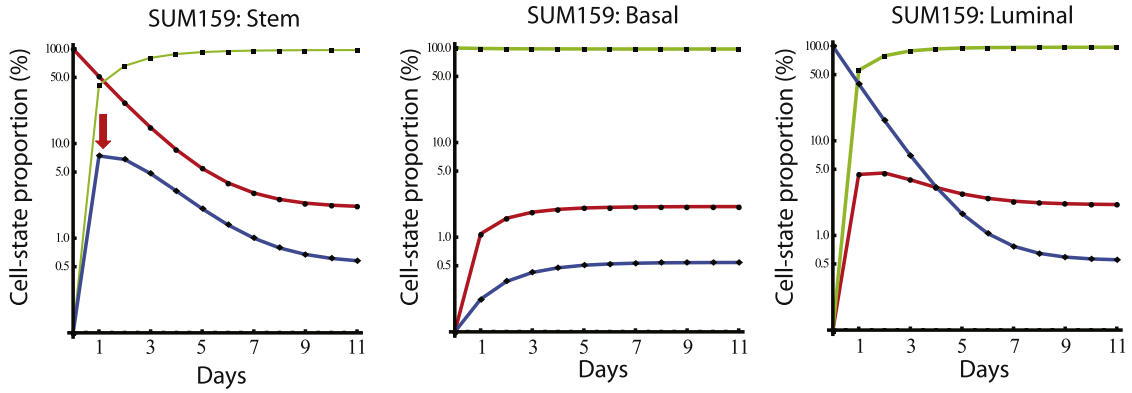

Figure 3. Prediction of Cell-State Dynamics with the Stochastic Cell-State Transition Model

The transition probabilities determined for the SUM159 line were used to compute the expected cell-state proportions over time for isolated subpopulations of stem-like, basal, or luminal cells. For all three isolated subpopulations, an eventual return to equilibrium cell-state proportions is predicted. The model predicts a transient increase in the proportion of SUM159 luminal cells one day after isolation of a stem-like subpopulation (red arrow).

Table 2. Incidence and Phenotype Analyses of Tumors Arising from Sorted SUM159 Subpopulations

SUM159 Subpopulations				Analysis of Formed Tumors				
Basal	Stem-like	Luminal	Tumor Incidence	Viable cells (%)	GFP-neg H2K-neg (%)	Basal (%)	Stem-like (%)	Luminal (%)
Direct Injection								
+	–	–	0/4					
–	+	–	4/4	17.11	49.34	93.38	6.03	0.59
–	–	+	0/4					
With GFP + Irrad. SUM159								
–	–	–	0/4					
+	–	–	4/5	58.1 ± 2.1	56.0 ± 6.0	81.7 ± 5.9	11.4 ± 3.5	6.9 ± 2.4
–	+	–	4/5	53.4 ± 4.7	56.7 ± 7.7	67.2 ± 11.4	24.1 ± 8.0	8.6 ± 3.5
–	–	+	4/5	65.7 ± 6.3	57.7 ± 7.2	82.4 ± 8.1	12.0 ± 6.5	5.6 ± 2.7

Sorted SUM159 $CD44^{lo}CD24^{hi}EpCAM^{hi}$ (luminal) $CD44^{hi}CD24^{neg}EpCAM^{neg}$ (basal), and $CD44^{hi}CD24^{neg}EpCAM^{lo}$ (stem-like) subpopulations (10^4 per injection) were either directly implanted into NOD/SCID mice alone (direct injection) or coinoculated with 10^6 GFP-labeled irradiated SUM159 parental carrier cells prior to implantation in vivo. The fraction of viable cells after tumor dissociation was assessed by propidium iodide staining coupled with flow-cytometry analysis. H2K and GFP were used to exclude contamination of mouse and GFP-positive carrier cells. Luminal, basal, and stem-like proportions within the viable cancer cells of the resulting SUM159 tumors are shown.

et al., 2010; Govoni et al., 2009; Howard and Ashworth, 2006; Ivanova et al., 2006; Lee et al., 2007). To inhibit TBX3 function, we developed and validated short hairpin RNA (shRNA) constructs targeting the *TBX3* gene. We then isolated stem-like, basal, or luminal subpopulations from SUM149 and SUM159 cells in which *TBX3* expression had been inhibited. After 6 days in culture, we determined the cell-state proportions in *TBX3*-inhibited populations and computed the corresponding transition probabilities (Table 1).

The results reveal that in SUM159 cells, *TBX3* inhibition decreased the probability of luminal-to-basal cell state transitions and increased the likelihood of self-renewing divisions for the luminal state ($P_{L\rightarrow B}$ decreases from 0.49 to 0.16, while $P_{L\rightarrow L}$ increases from 0.47 to 0.78; Table 1). In SUM149 cells, one cannot measure whether *TBX3* inhibition causes a decrease in $P_{L\rightarrow B}$ or an increase in $P_{L\rightarrow L}$, because the former is 0.00 and the latter is 0.99 in the absence of inhibition. However, *TBX3* inhibition in the SUM149 line resulted in an increase in the probability of basal-to-luminal transitions ($P_{B\rightarrow L}$ increases from 0.08 to 0.12; Table 1). These findings indicated that *TBX3* promotes basal differentiation by distinct mechanisms, depending on the underlying cell-state transitions present in a given cancer cell population.

Inferring Cell-State Treatment Sensitivities with a Markov Model

Cell-state dynamics can also be applied to study the differential toxicities of collections of potential anticancer drugs on various cell states. Such studies are of particular importance because of the observed resistance of cancer stem-like cells to chemo- and radiation therapies (Bao et al., 2006; Dean et al., 2005; Diehn and Clarke, 2006; Eyler and Rich, 2008; Li et al., 2008; Woodward et al., 2007). This has fueled interest in identifying drugs that selectively target specific cell-differentiation states. One approach to finding such drugs would be to use high-throughput flow cytometry to screen collections of chemicals for their effects on cell-state proportions.

To explore this idea, we treated SUM159 and SUM149 cells with two commonly used breast cancer chemotherapy drugs: paclitaxel and 5-fluorouracil (5-FU) (Figure 4A). For both cell lines, we observed that treatment with either drug resulted in an increase in the relative proportion of stem-like cells (each drug was applied at a concentration that resulted in 90% inhibition of viability). In SUM159 cells, treatment with either paclitaxel or (to a lesser degree) 5-FU also increased the proportion of cells in a luminal state. In contrast, treatment of SUM149 cells with either paclitaxel or 5-FU resulted in expansion of cells in the basal state.

To understand these observations, we expanded our model to allow the three cell states to have different relative viabilities in response to drug treatment. We used the same transition probabilities as previously determined for each cell line, but incorporated drug treatment effects by pre-multiplying the transition matrix by a diagonal matrix of relative viabilities (v_B, v_S, v_L). The relative viabilities represent the survival probability per generation for each cell state, with the values normalized to sum to 1. Because they sum to 1, the values can be represented as points in a ternary plot (that is, an equilateral triangle in which the three coordinates are represented by the distance to each side). We then used Monte Carlo simulation to infer the relative viability matrix most consistent with the cell-state proportions observed following drug treatment. Results of this analysis are shown in the ternary density plots in Figure 4C, with the color of each point representing the normalized difference between the experimentally observed and predicted cell-state proportions. The viability values most consistent with the data are shown as a dark brown region in Figure 4C.

For the SUM159 line, the viability parameters were constrained to a small region and imply that basal cells were more sensitive than the other cell-states to paclitaxel treatment (values tightly centered around $v_B = 0.1 \pm 0.025$, $v_S = 0.4 \pm 0.05$, $v_L = 0.5 \pm 0.05$) (Figure 4C). In contrast, the differential viabilities were not well constrained for 5-FU treatment. For SUM149 cells, the viabilities were tightly constrained and approximately equal for both drugs ($v_B = 0.15 \pm 0.15$, $v_S = 0.68 \pm 0.12$, $v_L = 0.22 \pm 0.05$), and indicate that the stem-like cells were selectively resistant to treatment (Figure 4C).

These results provide insight into the observation that paclitaxel treatment of SUM149 cells leads to an ~5-fold increase

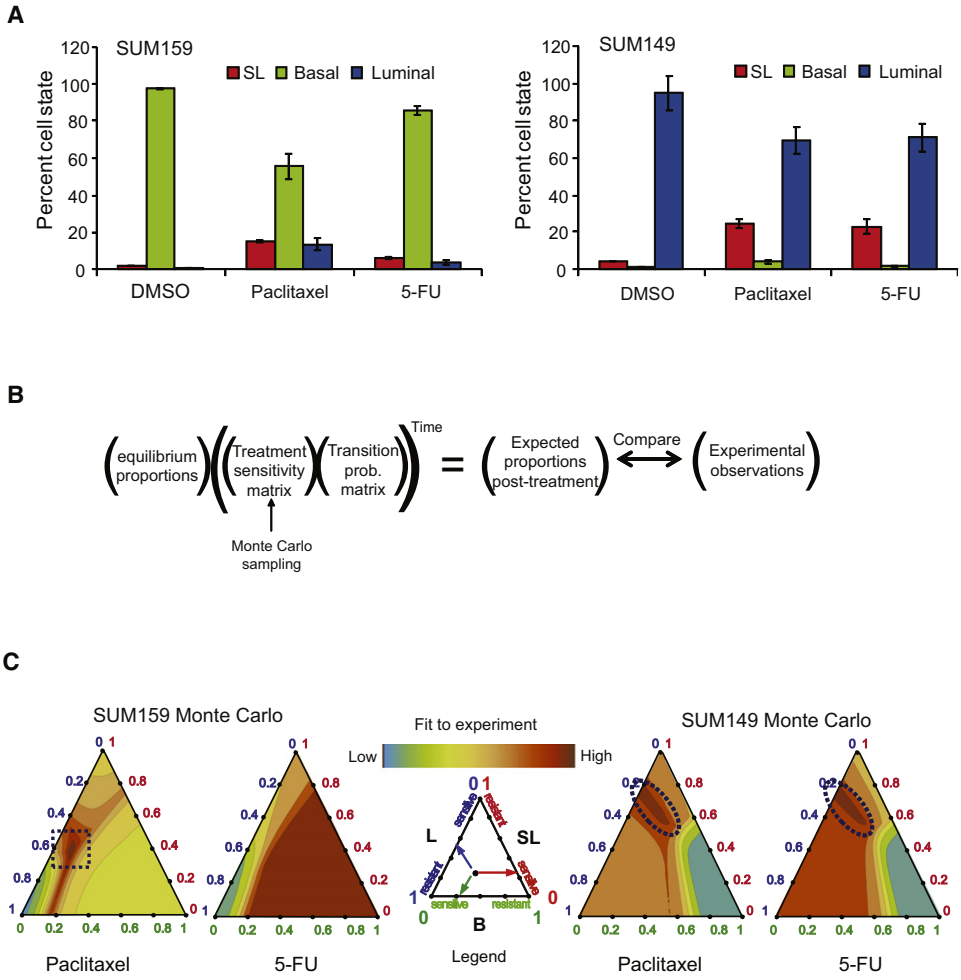

Figure 4. Analysis of Chemical Treatment Sensitivities by Monte Carlo Simulation with the Stochastic Cell-State Transition Model

(A) Cell-state proportions after a 6 day treatment with either paclitaxel or 5-FU are shown for SUM159 and SUM149 populations.

(B) A schematic showing how differential chemical treatment sensitivities can be incorporated into the Markov stochastic model. The matrix of transition probabilities is premultiplied by a differential viability matrix. The diagonal entries of this matrix, (v_B, v_S, and v_L), encode the survival probabilities of basal, stem-like and luminal cells in the presence of chemical treatment. These entries are normalized to sum to 1. Monte Carlo simulation is performed by random sampling of differential viability vectors.

(C) The results of Monte Carlo simulation, shown as ternary density contour plots in which coloration represents the normalized difference (see the Experimental Procedures for the metric used) between experimentally observed and predicted cell-state proportions following chemical treatment. Ten thousand random points within the triangle simplex were sampled, each representing a distinct choice of differential viability vector, (v_B, v_S, v_L), with entries ranging from zero (sensitive) to one (resistant).

in both stem-like and basal cells. In particular, they reveal that the relative increase in the proportion of basal cells does not reflect a selective resistance of basal cells to paclitaxel treatment. Rather, it is an indirect effect arising from the selective resistance of stem-like cells, which then expand and give rise to basal cells. This important inference would not have been possible without taking into account the lineage-structure (i.e., cell-state transition probabilities) of the SUM149 line.

DISCUSSION

While malignant transformation is invariably associated with the perturbation of normal regulatory mechanisms, one of the puzzling features of cancer cell populations is their ability to retain phenotypic equilibrium over extended periods of time. Populations of cancer cells, for example, often harbor subpopulations with specific cell-surface marker profiles that are stably maintained across many cell divisions in culture. The observations here demonstrate that such equilibrium is maintained through the interconversion of cancer cells between states, the latter being defined based on cell-surface marker expression. We have described a Markov model of cell-state dynamics that assumes interconversation rates depend only on a cell's current state and remain constant under fixed microenvironmental conditions.

A specific prediction of this quantitative model is that *any* subpopulation of cancer cells will return to a fixed equilibrium

of cell-state proportions over time, provided that it is possible through one or more interconversions to transition between any two states. The cell-state proportions at this equilibrium can be derived from the transition rates as the second eigenvector of the corresponding transition probability matrix. One corollary of this prediction is that deviations from equilibrium proportions will be corrected even in the absence of any intercellular communication. Therefore, while intercellular signals clearly influence cell-state decisions, they are not necessarily required for stability in phenotypic proportions. Thus, for many types of cancer populations, it is possible that phenotypic equilibrium is maintained in culture in the absence of regulatory feedback mechanisms.

The prediction that cell-state proportions in any subpopulation would return to a stable equilibrium does not depend on the strong form of the Markov property assumed above—namely, that cells lack any memory of their prior states. A more modest form of the Markov condition—that cell-state decisions depend only a fixed depth of prior history—would also be sufficient to promote a stable phenotypic equilibrium, provided there is weak interconnectedness between states as described above.

The prediction of a return to phenotypic equilibrium is consistent with the finding that CSCs give rise to tumors that recapitulate the differentiation-state heterogeneity present in the parental tumors from which they were derived. Our observations with breast cancer lines in vitro indicate that cells in CSC-depleted fractions give rise to cells expressing marker profiles that enrich for CSCs (Chaffer et al., 2011). If a similar interconversion occurs after in vivo injection, a prediction of the model would be that CSC-depleted fractions would give rise to tumors if they survived long enough at the site of implantation to regenerate CSCs. In support of this idea, we found that coimplantation with irradiated carrier cells enabled the luminal, basal and stem-like fractions of the SUM159 breast cancer line to form tumors with high efficiency. Additionally, as predicted by the Markov model, similar cell-state proportions were observed in tumors generated by the luminal, basal and stem-like fractions and in tumors generated from unfractionated populations. While these observations are consistent with the Markov model proposed here, further experimentation will be needed to quantify the rates of interconversion between cancer cell states in vivo.

The experiments described here used sorted subpopulations with relatively large numbers of cells. Accordingly, the transition rates determined represent population averages over numerous cancer cells. Therefore, it is possible that interconversion rates might vary significantly across the individual cancer cells of any particular sorted subpopulation. To determine whether such variability exists, further experimentation will be needed to quantify interconversation rates among the progeny of single-cell clones. For such an experiment, cancer-cell clones would need to be grown under conditions that control for effects on interconversion rates arising from sparsity in culture. Such variability, if observed, could arise from two sources: (1) the cell-surface markers failed to segregate epigenetically uniform subpopulations or (2) there are genetic differences among cancer cells that are in identical epigenetic states.

If observed, variability in interconversion rates across individual cells would not alter the predictive value, at a population level, of the Markov model proposed here. Indeed, the ability of the Markov model to predict the dynamics of cell-surface marker profiles does not depend on a one-to-one correspondence between biological states and the cell-surface marker profiles used to segregate subpopulations. Moreover, a subsequent refinement of the markers allowing for the discrimination of additional states would lead to a natural extension of the model.

In adult tissues, specialized niche cells supply stem and progenitor cells with paracrine signals necessary for their maintenance or expansion. In contrast, due to their acquired genetic lesions, cancer cells can stochastically enter into stem-like states in the absence of normal niche microenvironments. For example, *RB* loss in animal models of osteosarcoma can cause dedifferentiation of bone-committed progenitors, resulting in cancer cell plasticity and the ability to generate tumors of multiple lineages (Calo et al., 2010; Gutierrez et al., 2008). As another example, mutations in the *BRCA1* gene inhibit the ability of breast cancer cells to differentiate into the luminal lineage, resulting in less-differentiated tumors.

While our study was limited to mutated cancer cell populations, there is evidence indicating that ostensibly normal cells that are differentiated can either transdifferentiate or transition into more primitive states. Such differentiation-state plasticity occurs in response to either genetic or microenvironmental perturbations (Deng et al., 2011; Thorel et al., 2010; Zhou et al., 2008) and has also been observed during normal developmental processes (Richard et al., 2011). It will be of significant interest to determine how adult cells acquire plasticity, since modulation of the underlying mechanisms could lead to novel therapies for regenerative medicine.

While our findings indicate that cancer stem-like cells can arise de novo from non-stem-like cells at a low but significant rate, it is unclear whether the presence of such plasticity indicates that the concept of cancer stem cells is fundamentally different from that of normal stem cells. Indeed, the plasticity described above suggests that the lineage-hierarchy associated with normal cells may be more flexible than is commonly depicted (Figure 5, model I versus II).

The de novo generation of CSCs has implications for the effectiveness of anticancer therapies that exclusively target CSCs, because non-CSCs would regenerate CSCs after cessation of therapy and lead to renewed tumor growth. Therefore, in order to be effective, cancer therapies will need to combine agents that are selectively toxic to CSCs with agents that either target the bulk non-CSC populations within tumors or inhibit transitions from non-CSC to CSC states.

The Markov model described here allows for the quantitative inference of transition probabilities between phenotypic states within cancer cell populations. The marker profiles studied here correlate with particular differentiation states associated with mammary epithelial cells. However, the model described here could be applied to subpopulations isolated using any other cell-surface marker profiles. In principle, the model could also be extended to other biological settings in which stochastic state transitions occur, either in normal or diseased contexts. In such applications, it is possible that proliferation or survival rates could differ across cell-states. If this is the case, additional measurements quantifying the viability differences or model-fitting

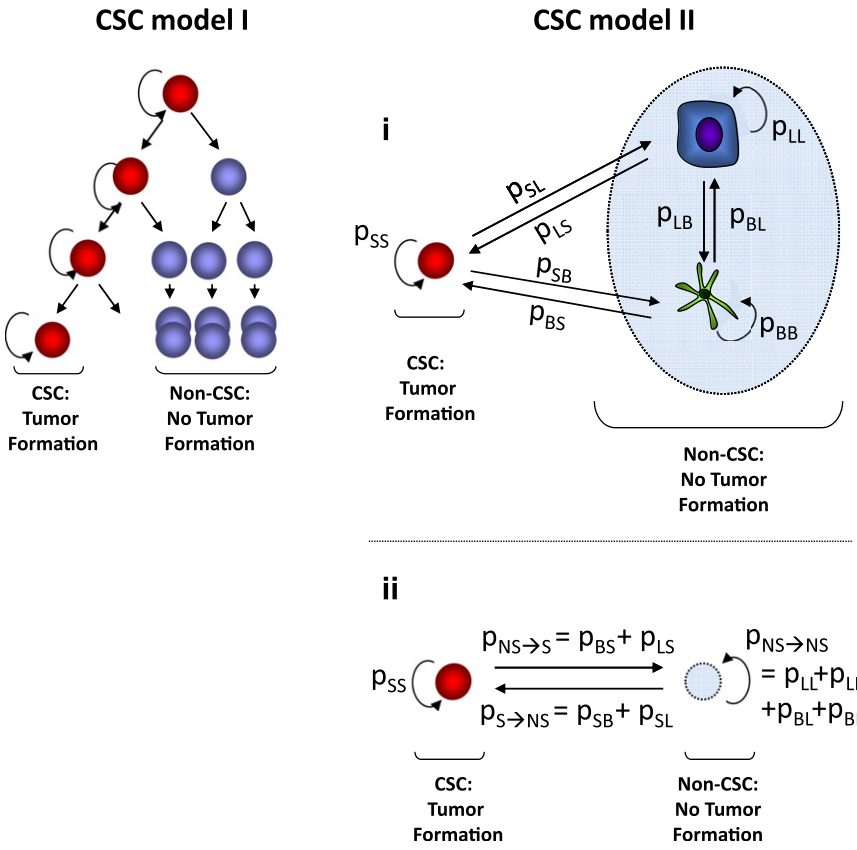

Figure 5. Two Distinct Models of Cancer Cell Populations

In the existing paradigm (model I), CSCs give rise to non-CSCs but not vice versa, resulting in a hierarchical cell-lineage structure reflective of normal tissue biology. We propose an alternative scenario (model II) in which there is bidirectional interconversion between CSC and non-CSC states. The rates of transition between cell states, which vary across distinct cancer cell populations, can be computed with the Markov modeling approach described in the main text and Experimental Procedures.

approaches would be needed in order to extend the model described here.

Interconversion of cells between differentiation states complicates the interpretation of experiments in which cell-state proportions are measured in populations of cells following chemical or genetic perturbation. Our findings indicate that knowledge of inter-conversion rates is needed to accurately characterize the effects of perturbations on cell-state dynamics. The model proposed makes it possible to systematically characterize how cell-state transitions are influenced by experimental perturbations. In addition to being practically useful, the ability to relate cell-state transition probabilities at the single-cell level to the dynamics of population cell-state proportions should prove useful in future theoretical studies of phenotypic complexity.

EXPERIMENTAL PROCEDURES

Description of the Cell-State Transition Markov Model

The Markov model makes several assumptions: (1) cells within a population can exist in any one of **M** possible states and (2) under fixed genetic and environmental conditions, cells transition from one state to another with transition rates per unit time that are constant.

Let $\mathbf{P}_{ij}$ denote the probability that a cell transitions from state ***i*** to state ***j***. Since a given cell must either remain in the same state or enter into one of the other possible **M-1** states, it follows that for all ***i***,

$$\sum_{j=1}^{M} P_{ij} = 1.$$

A population of cells can be described by a single row-vector $\boldsymbol{v}$ of dimension **k** in which the i^{th} entry of the vector is the proportion of cells that are present in the i^{th} phenotypic state. Since all cells must exist in one of the **k** states, it follows that for any $\boldsymbol{v}$,

$$\sum_{i=1}^{M} v_i = 1.$$

Given these assumptions, the dynamic evolution over time of the various cell state proportions is a Markov process. The rates of transition between states can be encoded in a (**M** × **M**) stochastic matrix **P** with row-sums equal to 1 that encodes the transition rates per unit time. Given a vector encoding the initial cell-state proportions at time 0, $\boldsymbol{v}^{(0)}$, and the stochastic matrix **P**, it is possible to calculate cell-state proportions after **T** time units:

$$v^{(T)} = v^{(0)} \cdot (P^T).$$

Computation of the Stochastic Matrix Associated with a Population of Cells

To predict the evolution of phenotypic proportions in a cellular population, the transition probabilities and initial state vector for the population need to be determined. The initial state vector corresponds to the proportions of the various cell states at time 0. To compute the matrix of transition probabilities, **P**, it is necessary to observe the evolution of at least **M** populations for a fixed length of time, **T.** Any set of M populations would suffice to determine P, provided that the initial state vectors corresponding to the distinct populations span $\mathbf{R}^{\mathbf{M}}$. This information is sufficient to determine the stochastic matrix in cases where proliferation and survival rates do not differ between states.

For example, one could use as starting populations the **M** distinct subpopulations for which all cells in a population are present in one of the states $\mathbf{S_1}$, $\mathbf{S_2}$,…, $\mathbf{S_M}$. The observed state-vector at time **T** for population $\mathbf{S_i}$ is given by the i^{th} row of the matrix $\mathbf{P}^{(T)}$ where P represents the transition rates per unit time. Accordingly, the matrix **P** would equal the matrix of observations raised to the power (**1/T**).

Monte Carlo Simulation with Transition Probabilities for Analysis of Treatment Sensitivity

We assumed that compound treatment affects cell viability but not the cell-state transition probabilities for a given population. The differential treatment sensitivities for the k distinct cell states were represented in a diagonal (**M** × **M)** matrix with diagonal elements encoding survival probabilities summing to one. Together, the differential treatment sensitivity and transition probabilities allow for in silico predictions of cell state proportions after treatment. To

determine the differential drug sensitivities that best correspond to experimental observations following treatment, Monte Carlo simulation was performed by randomly sampling treatment sensitivity vectors from the (**M-1**)-simplex. The initial cell-state proportions prior to treatment were assumed to be at equilibrium. For each cell line, the predicted cell-state proportions after 6 days of treatment were determined from the differential treatment sensitivity matrix and the matrix of transition probabilities for the cell line. The normalized distance, ***D***, between the observed ν^i_{obs} and predicted ν^i_{pred} cell-state proportions was calculated with the metric

$$D = \sum_{i=1}^{M} \frac{|\nu^i_{obs} - \nu^i_{pred}|}{\nu^i_{eq}},$$

where ν^i_{eq} are the equilibrium proportions.

Cell Lines and Tissue Culture

SUM cell lines were obtained from Stephen Ethier (Karmanos Institute, MI) and are commercially available (Asterand); SUM149 and SUM159 cells were cultured in Ham's F12 medium with 5% calf serum, insulin (5 μg/ml), and hydrocortisone (1 μg/ml). All cell lines were grown at 37°C in a 5% CO_2 incubator.

Microarray Hybridization and Global Gene Expression Profiling

Total RNA for gene-expression studies was isolated from sorted stem-like ($CD44^{hi}CD24^{neg}EpCAM^{lo}$), basal ($CD44^{hi}CD24^{neg}EpCAM^{neg}$), and luminal ($CD44^{lo}CD24^{hi}EpCAM^{hi}$) subpopulations from SUM149 and SUM159 cell cultures with the RNeasy Mini kit (QIAGEN). Synthesis of complementary DNA from total RNA and hybridization/scanning of microarrays were performed with Affymetrix GeneChip products (HGU133A) as described in the GeneChip manual. Raw data files (.CEL) were converted into probe-set values by RMA normalization using the affy package in R (Gentleman et al., 2004).

After RMA normalization, the limma R package (Smyth, 2004) was used to determine regression coefficients, identifying probes differentially expressed between stem versus basal cells, stem versus luminal and luminal versus basal cells for the SUM149 and SUM159 lines. Hierarchical clustering was performed with the Pearson correlation metric, using probes consistently regulated across both lines. GSEA was performed as described previously (Mootha et al., 2003; Subramanian et al., 2005), using gene lists previously reported to characterize stem, luminal, and basal cells (Table S1). GSEA was on performed for each pair-wise comparison on probe lists preranked based on the sum of the regression coefficients for both lines. Probe sets were collapsed to gene scores based on maximal intensity.

Lentiviral Infection of shRNA Constructs

Bacterial glycerol stocks of MISSION shRNAs were obtained (SIGMA) and plasmid DNA was isolated by mini-prep (QIAGEN). pLKO.1 (0.6 μg) plasmid containing the shRNA constructs targeting Tbx3 or no known target (Scrambled) were transiently cotransfected into S293T cells along with the VSV-G-expressing construct pCMV-VSV-G and the packaging construct pCMV ΔR8.2Δvpr generously provided by Inder Verma (Salk Institute). Viral supernatant was collected and introduced to subconfluent SUM149 or SUM159 cultures. Lentiviral integration was selected with 1 μg/ml puromycin for 7 days, and knockdown efficiency was measured by quantitative RT-PCR. The Tbx3 clone ID is NM_016569.2-443s1c1 5′ATTGATCCATGATC GGCTTGG3′.

Fluorescence-Activated Cell Sorting

Subconfluent cultures were trypsinized, counted, washed with PBS, and stained with antibodies specific for the following human cell-surface markers: EpCAM (ESA)-FITC (clone VU-ID9, AbD Serotec), CD24-PE (clone ML5, BD PharMingen), and CD44-APC (clone G44-26, BD PharMingen). For each staining reaction, cells were incubated with antibody (20 μl antibody per million cells) for 15 min at room temperature. Unbound antibody was washed off, and cells were analyzed on a BD FACSCaliber no more than 1 hr after staining. Isotype controls included mouse IgG_1-FITC, mouse $IgG_{2a\kappa}$-PE, and mouse $IgG_{2b\kappa}$-APC (BD PharMingen). Cell sorting for subsequent culture was performed on a MoFlo cell sorter (Cytomation).

Tumor formation Experiments

All animal procedures were conducted in accordance with a protocol approved by the Tufts University IACUC committee. Nulliparous NOD/SCID female mice aged 6–8 weeks were used for all injection experiments. Human breast cancer cells were suspended in 30 μl phosphate-buffered saline and injected into the fourth inguinal mammary glands of mice. Tumor formation was assessed by palpitation. Tumors were dissociated with Collagenase Type IV and 0.25% Trypsin (Invitrogen). Viable human cells were isolated from dissociated tumors by negative selection of GFP-positive irradiated carrier cells, propidium iodide-positive cells, and mouse pan H2K-MHCI-PE-positive cells. The FACS-sorted viable human cells were then stained for flow cytometry as described above for cell lines.

SUPPLEMENTAL INFORMATION

Supplemental Information includes one figure and one table and can be found with this article online at doi:10.1016/j.cell.2011.07.026.

ACKNOWLEDGMENTS

We thank Aviv Regev, Robert Weinberg, Todd Golub, and Gaorav Gupta for valuable discussions and criticism of the manuscript, Supriya Gupta for assistance with gene-expression profiling, and Tom DiCesare for assistance with graphical design. This work was supported by the Breast Cancer Research Foundation, the Raymond and Beverly Sackler Foundation, and the Broad Institute.

Received: May 2, 2010
Revised: March 15, 2011
Accepted: July 20, 2011
Published: August 18, 2011

REFERENCES

Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J., and Clarke, M.F. (2003). Prospective identification of tumorigenic breast cancer cells. Proc. Natl. Acad. Sci. USA *100*, 3983–3988.

Assou, S., Le Carrour, T., Tondeur, S., Ström, S., Gabelle, A., Marty, S., Nadal, L., Pantesco, V., Réme, T., Hugnot, J.P., et al. (2007). A meta-analysis of human embryonic stem cells transcriptome integrated into a web-based expression atlas. Stem Cells *25*, 961–973.

Bao, S., Wu, Q., McLendon, R.E., Hao, Y., Shi, Q., Hjelmeland, A.B., Dewhirst, M.W., Bigner, D.D., and Rich, J.N. (2006). Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. Nature *444*, 756–760.

Blackman, B., Russell, T., Nordeen, S.K., Medina, D., and Neville, M.C. (2005). Claudin 7 expression and localization in the normal murine mammary gland and murine mammary tumors. Breast Cancer Res. *7*, R248–R255.

Böcker, W., Moll, R., Poremba, C., Holland, R., Van Diest, P.J., Dervan, P., Bürger, H., Wai, D., Ina Diallo, R., Brandt, B., et al. (2002). Common adult stem cells in the human breast give rise to glandular and myoepithelial cell lineages: a new cell biological concept. Lab. Invest. *82*, 737–746.

Bonnet, D., Bhatia, M., Wang, J.C., Kapp, U., and Dick, J.E. (1999). Cytokine treatment or accessory cells are required to initiate engraftment of purified primitive human hematopoietic cells transplanted at limiting doses into NOD/SCID mice. Bone Marrow Transplant. *23*, 203–209.

Brambrink, T., Foreman, R., Welstead, G.G., Lengner, C.J., Wernig, M., Suh, H., and Jaenisch, R. (2008). Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. Cell Stem Cell *2*, 151–159.

Bussolati, G., Cassoni, P., Ghisolfi, G., Negro, F., and Sapino, A. (1996). Immunolocalization and gene expression of oxytocin receptors in carcinomas and non-neoplastic tissues of the breast. Am. J. Pathol. *148*, 1895–1903.

Calo, E., Quintero-Estades, J.A., Danielian, P.S., Nedelcu, S., Berman, S.D., and Lees, J.A. (2010). Rb regulates fate choice and lineage commitment in vivo. Nature *466*, 1110–1114.

Chaffer, C.L., Brueckmann, I., Scheel, C., Kaestli, A.J., Wiggins, P.A., Rodrigues, L.O., Brooks, M., Reinhardt, F., Su, Y., Polyak, K., et al. (2011). Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. Proc. Natl. Acad. Sci. *108*, 7950–7955.

Charafe-Jauffret, E., Ginestier, C., Monville, F., Finetti, P., Adélaïde, J., Cervera, N., Fekairi, S., Xerri, L., Jacquemier, J., Birnbaum, D., and Bertucci, F. (2006). Gene expression profiling of breast cell lines identifies potential new basal markers. Oncogene *25*, 2273–2284.

Chiou, S.H., Yu, C.C., Huang, C.Y., Lin, S.C., Liu, C.J., Tsai, T.H., Chou, S.H., Chien, C.S., Ku, H.H., and Lo, J.F. (2008). Positive correlations of Oct-4 and Nanog in oral cancer stem-like cells and high-grade oral squamous cell carcinoma. Clin. Cancer Res. *14*, 4085–4095.

Cowan, C.A., Klimanskaya, I., McMahon, J., Atienza, J., Witmyer, J., Zucker, J.P., Wang, S., Morton, C.C., McMahon, A.P., Powers, D., and Melton, D.A. (2004). Derivation of embryonic stem-cell lines from human blastocysts. N. Engl. J. Med. *350*, 1353–1356.

Creighton, C.J., Li, X., Landis, M., Dixon, J.M., Neumeister, V.M., Sjolund, A., Rimm, D.L., Wong, H., Rodriguez, A., Herschkowitz, J.I., et al. (2009). Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. Proc. Natl. Acad. Sci. USA *106*, 13820–13825.

Dean, M., Fojo, T., and Bates, S. (2005). Tumour stem cells and drug resistance. Nat. Rev. Cancer *5*, 275–284.

Deng, H., Wang, H.F., Gao, Y.B., Jin, X.L., and Xiao, J.C. (2011). Hepatic progenitor cell represents a transitioning cell population between liver epithelium and stroma. Med. Hypotheses *76*, 809–812.

Di Talia, S., Skotheim, J.M., Bean, J.M., Siggia, E.D., and Cross, F.R. (2007). The effects of molecular noise and size control on variability in the budding yeast cell cycle. Nature *448*, 947–951.

Diehn, M., and Clarke, M.F. (2006). Cancer stem cells and radiotherapy: new insights into tumor radioresistance. J. Natl. Cancer Inst. *98*, 1755–1757.

Dontu, G., Abdallah, W.M., Foley, J.M., Jackson, K.W., Clarke, M.F., Kawamura, M.J., and Wicha, M.S. (2003). In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. Genes Dev. *17*, 1253–1270.

Drabsch, Y., Hugo, H., Zhang, R., Dowhan, D.H., Miao, Y.R., Gewirtz, A.M., Barry, S.C., Ramsay, R.G., and Gonda, T.J. (2007). Mechanism of and requirement for estrogen-regulated MYB expression in estrogen-receptor-positive breast cancer cells. Proc. Natl. Acad. Sci. USA *104*, 13762–13767.

Dressman, M.A., Walz, T.M., Lavedan, C., Barnes, L., Buchholtz, S., Kwon, I., Ellis, M.J., and Polymeropoulos, M.H. (2001). Genes that co-cluster with estrogen receptor alpha in microarray analysis of breast biopsies. Pharmacogenomics J. *1*, 135–141.

Eyler, C.E., and Rich, J.N. (2008). Survival of the fittest: cancer stem cells in therapeutic resistance and angiogenesis. J. Clin. Oncol. *26*, 2839–2845.

Fillmore, C.M., and Kuperwasser, C. (2008). Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy. Breast Cancer Res. *10*, R25.

Fillmore, C.M., Gupta, P.B., Rudnick, J.A., Caballero, S., Keller, P.J., Lander, E.S., and Kuperwasser, C. (2010). Estrogen expands breast cancer stem-like cells through paracrine FGF/Tbx3 signaling. Proc. Natl. Acad. Sci. USA *107*, 21737–21742.

Frobenius, G.F. (1912). Uber Matrizen aus nicht negativen Elementen. Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin, 456–477.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., et al. (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. *5*, R80.

Gill, S., Peston, D., Vonderhaar, B.K., and Shousha, S. (2001). Expression of prolactin receptors in normal, benign, and malignant breast tissue: an immunohistological study. J. Clin. Pathol. *54*, 956–960.

Govoni, K.E., Linares, G.R., Chen, S.T., Pourteymoor, S., and Mohan, S. (2009). T-box 3 negatively regulates osteoblast differentiation by inhibiting expression of osterix and runx2. J. Cell. Biochem. *106*, 482–490.

Gudjonsson, T., Villadsen, R., Nielsen, H.L., Rønnov-Jessen, L., Bissell, M.J., and Petersen, O.W. (2002). Isolation, immortalization, and characterization of a human breast epithelial cell line with stem cell properties. Genes Dev. *16*, 693–706.

Gupta, P.B., Onder, T.T., Jiang, G., Tao, K., Kuperwasser, C., Weinberg, R.A., and Lander, E.S. (2009). Identification of selective inhibitors of cancer stem cells by high-throughput screening. Cell *138*, 645–659.

Gutierrez, G.M., Kong, E., Sabbagh, Y., Brown, N.E., Lee, J.S., Demay, M.B., Thomas, D.M., and Hinds, P.W. (2008). Impaired bone development and increased mesenchymal progenitor cells in calvaria of RB1-/- mice. Proc. Natl. Acad. Sci. USA *105*, 18402–18407.

Harris, M.A., Yang, H., Low, B.E., Mukherjee, J., Guha, A., Bronson, R.T., Shultz, L.D., Israel, M.A., and Yun, K. (2008). Cancer stem cells are enriched in the side population cells in a mouse model of glioma. Cancer Res. *68*, 10051–10059.

Howard, B., and Ashworth, A. (2006). Signalling pathways implicated in early mammary gland morphogenesis and breast cancer. PLoS Genet. *2*, e112.

Ivanova, N., Dobrin, R., Lu, R., Kotenko, I., Levorse, J., DeCoste, C., Schafer, X., Lun, Y., and Lemischka, I.R. (2006). Dissecting self-renewal in stem cells with RNA interference. Nature *442*, 533–538.

Jones, C., Mackay, A., Grigoriadis, A., Cossu, A., Reis-Filho, J.S., Fulford, L., Dexter, T., Davies, S., Bulmer, K., Ford, E., et al. (2004). Expression profiling of purified normal human luminal and myoepithelial breast cells: identification of novel prognostic markers for breast cancer. Cancer Res. *64*, 3037–3045.

Kariagina, A., Aupperlee, M.D., and Haslam, S.Z. (2007). Progesterone receptor isoforms and proliferation in the rat mammary gland during development. Endocrinology *148*, 2723–2736.

Kussell, E., and Leibler, S. (2005). Phenotypic diversity, population growth, and information in fluctuating environments. Science *309*, 2075–2078.

Lapidot, T., Sirard, C., Vormoor, J., Murdoch, B., Hoang, T., Caceres-Cortes, J., Minden, M., Paterson, B., Caligiuri, M.A., and Dick, J.E. (1994). A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. Nature *367*, 645–648.

Lee, H.S., Cho, H.H., Kim, H.K., Bae, Y.C., Baik, H.S., and Jung, J.S. (2007). Tbx3, a transcriptional factor, involves in proliferation and osteogenic differentiation of human adipose stromal cells. Mol. Cell. Biochem. *296*, 129–136.

Li, C., Heidt, D.G., Dalerba, P., Burant, C.F., Zhang, L., Adsay, V., Wicha, M., Clarke, M.F., and Simeone, D.M. (2007). Identification of pancreatic cancer stem cells. Cancer Res. *67*, 1030–1037.

Li, X., Lewis, M.T., Huang, J., Gutierrez, C., Osborne, C.K., Wu, M.F., Hilsenbeck, S.G., Pavlick, A., Zhang, X., Chamness, G.C., et al. (2008). Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. J. Natl. Cancer Inst. *100*, 672–679.

Liu, G., Yuan, X., Zeng, Z., Tunici, P., Ng, H., Abdulkadir, I.R., Lu, L., Irvin, D., Black, K.L., and Yu, J.S. (2006). Analysis of gene expression and chemoresistance of CD133+ cancer stem cells in glioblastoma. Mol. Cancer *5*, 67.

Lu, S., Simin, K., Khan, A., and Mercurio, A.M. (2008). Analysis of integrin beta4 expression in human breast cancer: association with basal-like tumors and prognostic significance. Clin. Cancer Res. *14*, 1050–1058.

Monaghan, P., Clarke, C.L., Perusinghe, N.P., Ormerod, M.G., and O'Hare, M.J. (1995). Epidermal growth factor receptor expression on human breast luminal and basal cells in vitro. Epithelial Cell Biol. *4*, 52–62.

Mootha, V.K., Lindgren, C.M., Eriksson, K.F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E., et al.

(2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat. Genet. *34*, 267–273.

Parisi, S., Passaro, F., Aloia, L., Manabe, I., Nagai, R., Pastore, L., and Russo, T. (2008). Klf5 is involved in self-renewal of mouse embryonic stem cells. J. Cell Sci. *121*, 2629–2634.

Perron, O. (1907). Zur Theorie der Matrizen. Mathematische Annalen *64*, 248–263.

Perry, S.S., Zhao, Y., Nie, L., Cochrane, S.W., Huang, Z., and Sun, X.H. (2007). Id1, but not Id3, directs long-term repopulating hematopoietic stem-cell maintenance. Blood *110*, 2351–2360.

Phinney, D.G., Gray, A.J., Hill, K., and Pandey, A. (2005). Murine mesenchymal and embryonic stem cells express a similar Hox gene profile. Biochem. Biophys. Res. Commun. *338*, 1759–1765.

Rasbridge, S.A., Gillett, C.E., Sampson, S.A., Walsh, F.S., and Millis, R.R. (1993). Epithelial (E-) and placental (P-) cadherin cell adhesion molecule expression in breast carcinoma. J. Pathol. *169*, 245–250.

Reis-Filho, J.S., Milanezi, F., Steele, D., Savage, K., Simpson, P.T., Nesland, J.M., Pereira, E.M., Lakhani, S.R., and Schmitt, F.C. (2006). Metaplastic breast carcinomas are basal-like tumours. Histopathology *49*, 10–21.

Richard, J.P., Zuryn, S., Fischer, N., Pavet, V., Vaucamps, N., and Jarriault, S. (2011). Direct in vivo cellular reprogramming involves transition through discrete, non-pluripotent steps. Development *138*, 1483–1492.

Sansone, P., Storci, G., Giovannini, C., Pandolfi, S., Pianetti, S., Taffurelli, M., Santini, D., Ceccarelli, C., Chieco, P., and Bonafé, M. (2007). p66Shc/Notch-3 interplay controls self-renewal and hypoxia survival in human stem/progenitor cells of the mammary gland expanded in vitro as mammospheres. Stem Cells *25*, 807–815.

Shipitsin, M., Campbell, L.L., Argani, P., Weremowicz, S., Bloushtain-Qimron, N., Yao, J., Nikolskaya, T., Serebryiskaya, T., Beroukhim, R., Hu, M., et al. (2007). Molecular definition of breast tumor heterogeneity. Cancer Cell *11*, 259–273.

Singh, S.K., Hawkins, C., Clarke, I.D., Squire, J.A., Bayani, J., Hide, T., Henkelman, R.M., Cusimano, M.D., and Dirks, P.B. (2004). Identification of human brain tumour initiating cells. Nature *432*, 396–401.

Sleeman, K.E., Kendrick, H., Ashworth, A., Isacke, C.M., and Smalley, M.J. (2006). CD24 staining of mouse mammary gland cells defines luminal epithelial, myoepithelial/basal and non-epithelial cells. Breast Cancer Res. *8*, R7.

Smalley, M., and Ashworth, A. (2003). Stem cells and breast cancer: A field in transit. Nat. Rev. Cancer *3*, 832–844.

Smyth, G.K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat. Appl. Genet. Mol. Biol. *3*, Article3.

Sørlie, T., Wang, Y., Xiao, C., Johnsen, H., Naume, B., Samaha, R.R., and Børresen-Dale, A.L. (2006). Distinct molecular mechanisms underlying clinically relevant subtypes of breast cancer: gene expression analyses across three different platforms. BMC Genomics *7*, 127.

Stingl, J., and Caldas, C. (2007). Molecular heterogeneity of breast carcinomas and the cancer stem cell hypothesis. Nat. Rev. Cancer *7*, 791–799.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550.

Süel, G.M., Garcia-Ojalvo, J., Liberman, L.M., and Elowitz, M.B. (2006). An excitable gene regulatory circuit induces transient cellular differentiation. Nature *440*, 545–550.

Süel, G.M., Kulkarni, R.P., Dworkin, J., Garcia-Ojalvo, J., and Elowitz, M.B. (2007). Tunability and noise dependence in differentiation dynamics. Science *315*, 1716–1719.

Thattai, M., and van Oudenaarden, A. (2004). Stochastic gene expression in fluctuating environments. Genetics *167*, 523–530.

Thorel, F., Népote, V., Avril, I., Kohno, K., Desgraz, R., Chera, S., and Herrera, P.L. (2010). Conversion of adult pancreatic alpha-cells to beta-cells after extreme beta-cell loss. Nature *464*, 1149–1154.

Wolf, D.M., Vazirani, V.V., and Arkin, A.P. (2005). Diversity in times of adversity: probabilistic strategies in microbial survival games. J. Theor. Biol. *234*, 227–253.

Woodward, W.A., Chen, M.S., Behbod, F., Alfaro, M.P., Buchholz, T.A., and Rosen, J.M. (2007). WNT/beta-catenin mediates radiation resistance of mouse mammary progenitor cells. Proc. Natl. Acad. Sci. USA *104*, 618–623.

Yarosh, W., Barrientos, T., Esmailpour, T., Lin, L., Carpenter, P.M., Osann, K., Anton-Culver, H., and Huang, T. (2008). TBX3 is overexpressed in breast cancer and represses p14 ARF by interacting with histone deacetylases. Cancer Res. *68*, 693–699.

Zhou, Q., Brown, J., Kanarek, A., Rajagopal, J., and Melton, D.A. (2008). In vivo reprogramming of adult pancreatic exocrine cells to beta-cells. Nature *455*, 627–632.