## Supplemental Material and Methods

### *The SCA Model*

The SCA model is a hybrid cellular automaton (1) with Von Neumann neighborhood on which the tumor is represented by a discrete set of cells on a 2-dimensional lattice Ω of N×N sites with zero-flux boundary conditions. A biological cell is a point in the lattice and can be a normal, a cancer or a necrotic cell. It is 10×10μm in size (2) and possesses the attributes in **Table 1**. The parameter set we apply in this is largely comparable to parameters used in previous computational model to allow for comparison (3, 4). Upon a cell division, an empty place for the daughter cell is created in one of its neighboring sites by shifting outward the surrounding cells.

See **Quick Guide** for the implementation of the processes of metabolism, migration and hierarchical organization

### *3D SCA Model*

We extended the 2D SCA model to at 3-dimensional Von Neumann cellular automaton by converting all the algorithms and the numerical schemes in the model to 3 dimensions. Moreover, because of the large need of memory resources for 3 dimensional cellular automata models and the limited availability of such in all 32-bit computer architectures (4GB), we neglect tumor metabolism in the 3D SCA. This conserves the geometrical and morphological properties that our model yields and makes it possible to simulate large tumor masses (in the order of millions of cells).

### *Phenotypical Evolution*

During tumor progression cancer cells acquire further malignant features (5) such as rapid proliferation speed, hypoxia resistance, migration and loss of cell adhesion. To implement the essential process of clonal evolution in the progress of malignancies in our model system, we assume that at each symmetrical cell division, CSCs have a certain probability $P_{Mut}$ to acquire a *hit* and therefore generate a daughter cell with a different phenotype. Before the start of the experiments we randomly generated a pool of 30 different phenotypes that vary in proliferation speed, consumption rate, random mobility and cell-to-cell adhesion, as shown in **Table S2**. In the SCA model we use a non-supervised scheme in a sense that after a *hit* occurs the daughter cell is randomly assigned a previously generated phenotype from the

pool. This implicates that a cell can acquire both beneficial and disadvantageous new traits, like in reality. For the characteristics of the generated phenotypes that we use in the presented experiments see **Table S3**.

All the randomness in our model is generated using a Mersenne Twister random number generator (6) which allows for fast and high quality random number generation. The model is entirely implemented in C++ (7) as independent open source code for GNU/Linux and it is available under the GPLv3 software license.

***Degree of Tumor Invasiveness***

Tumor invasion is responsible of roughness of the tumor surface, large inhomogeneities and fingering tumor fronts. As a simple measure of invasiveness in 2D, we consider a non-invasive tumor, being spherical, to have the minimum ratio between the perimeter $P$ and the surface $S$. As the irregularity of the tumor borders increases with its invasiveness, such ratio increases as well. By comparing the $P/S$ ratio of a tumor mass with the minimum ratio (the one of a perfect disk), we can measure the level of invasiveness of a solid malignancy. Hence, the invasiveness $IM$ reads:

$$IM = P \Big/ \sqrt{4 \pi S} \qquad \text{(c)}$$

A non-invasive, spherical tumor would have $IM$ close to 1 (the closest to a disk or, in 3 dimensions, to a sphere) while a tumor with a highly irregular border, made by fingers and cluster of invasive cells, would display a high value of $IM$. For quantification of *in vitro* invasion the same measure was applied.

***Degree of Tumor Heterogeneity***

To measure the level of clonal diversity in the tumor population we make use of the *Shannon index* (8). The Shannon index is one of the many diversity indices employed to calculate the diversity of categorical data. It measures the information entropy of a population of items treating species as symbols and relative population sizes as probabilities. The Shannon index is frequently used for measuring biodiversity because it takes into the account the number of species and their evenness. The index increases with additional unique species or with greater species evenness and it is given by equation (d).

$$H = -\sum_{i=1}^{S} p_i \ln(p_i) \qquad \text{(d)}$$

In which $p_i$ represent the relative abundance of the species $i$ over a total of $N$ individuals belonging to $S$ species. Moreover, for any given number of species $S$ there is a maximum possible Shannon index of $H_{max}=\ln(S)$.

***Implementation of a Simulated Therapy***

To simulate a *virtual* tumor treatment we let the malignancy grow up to a target volume $V_T$, then we apply the therapy using three different policies depending on the model of growth (**Figure 5**). For the cancer stem cell model we analyze the response in the case where CSCs are therapy-resistant and in the case where they are therapy-sensitive. In the first scenario the therapy would kill all DCCs, leaving all the CSCs initially present in the tumor to reform the malignancy. In second case instead, we kill the same amount of cells (99.8% of the total volume) but choosing them at random. Then we examine the tumors where the therapy has failed and has spared at least one CSC that regenerates the tumor.

For the classical model, where all cells are tumorigenic, we assume the therapy to have a random effect and to exterminate 99.8% of the tumor population.

To understand which phenotypes present before therapy that are involved in the regrowth of the malignancy we assume the mutation rate $P_{Mut}=0$ after treatment (**Figure 5** and **S8**).

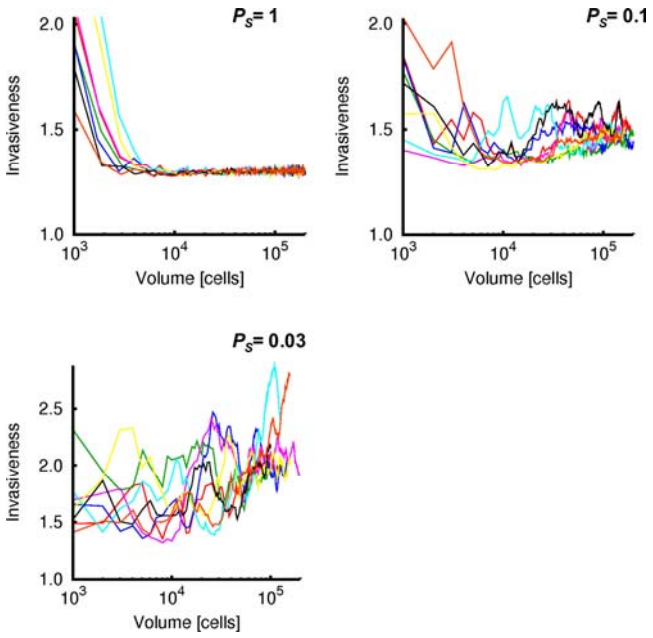
***In vitro experiments***

Cells were cultured according to standard protocol in DMEM, IMDM or MEM medium (Gibco/Invitrogen) supplemented with 10% Fetal Calf Serum (BioWhittaker) and 5 mmol/L L-glutamine (Invitrogen) and were kept in 2%, 5% or 10% $CO_2$ conditions. Limiting dilution assay was performed by FACS deposition of 1, 2, 4, 8, 16, 32, 64, 128, 256 cells in a 96 well plate (Corning). Clonogenicity was calculated using the *limdil* function in the 'statmod' software package (http://bioinf.wehi.edu.au/software/limdil/). To determine the invasiveness cells were plated at clonal density on adherent plates or in Growth Factor Reduced Matrigel (BD Biosciences) and overlaid with medium. After 10 days, surface and perimeter of cellular structures were measured and an Invasion Measure was calculated. Cell lines

that did not form clear visible aggregates of cells but completely dispersed were excluded from analysis as indicated in **Table S1**.

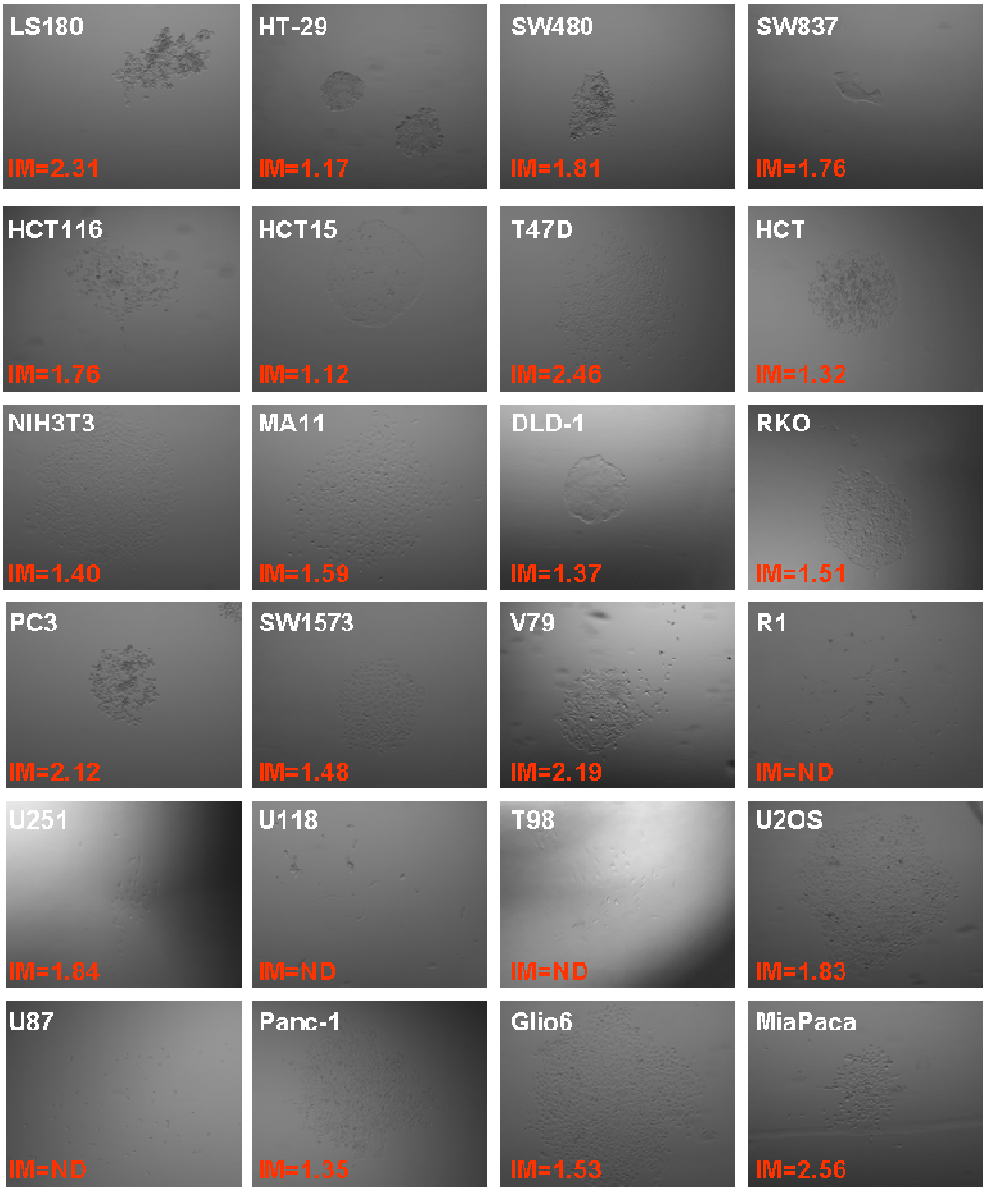
## Supplemental Figures

**Figure S1. Invasive Morphology in the Cancer Stem Cell Model**



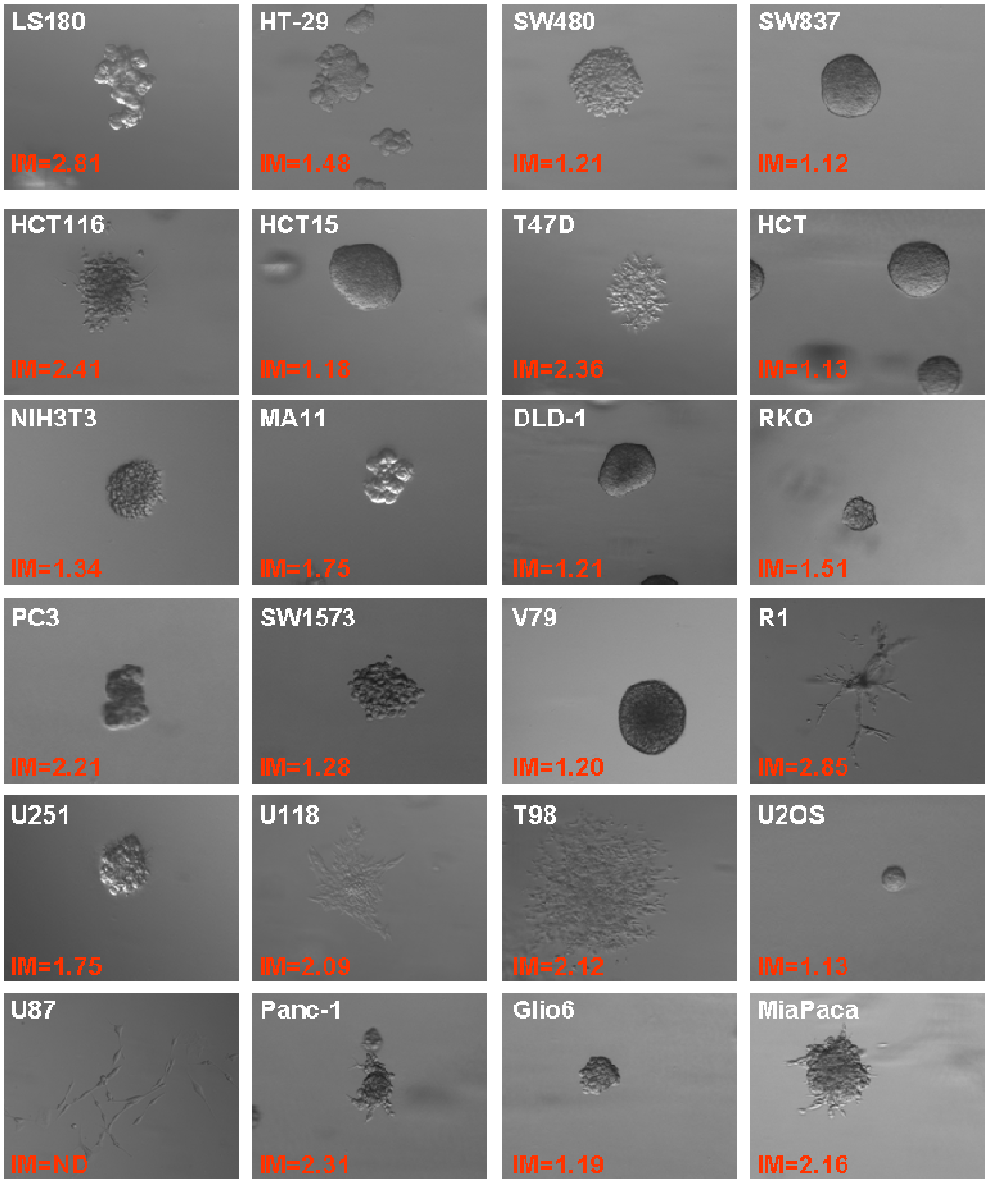

Invasiveness measure as defined in the Material and Methods for individual simulations with the indicted $P_S$ value (8 of the total 16 replicas for every condition). Lines in the plot end when tumor mass reaches the edge of the grid. Evident is a wave-like pattern for invasive behavior. For combined graph (n=16) see **Figure 1C**.

**Figure S2. In vitro validation; adherent cultures**



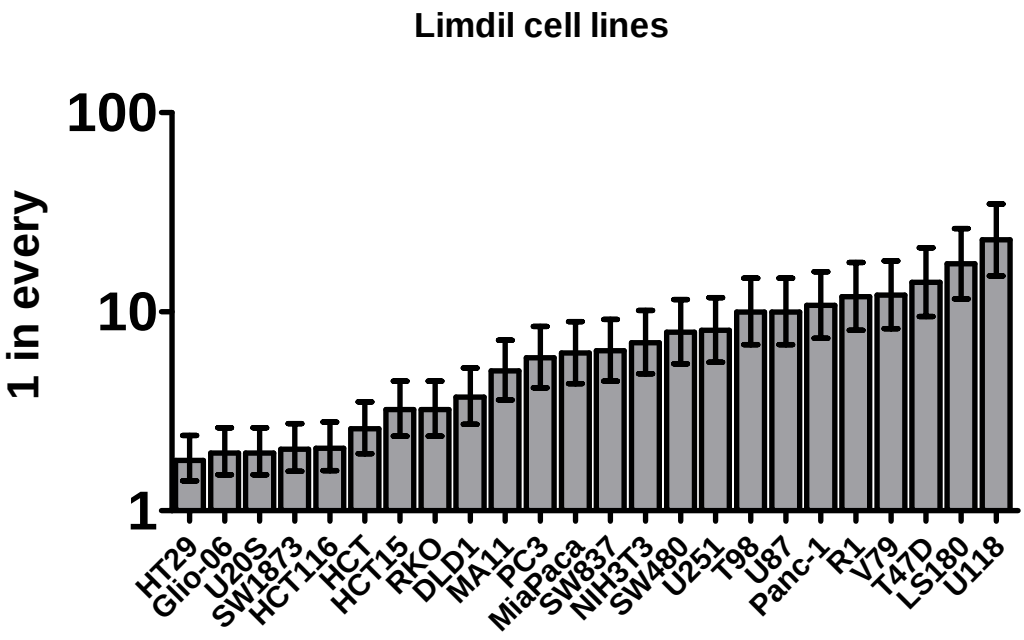

Cells from various cell lines have been plated on adherent plates at clonal density. At day 10 pictures have been taken and invasion/irregularity was quantified as described in material and method section.

**Figure S3. In vitro validation; matrigel cultures**




Cells from various cell lines have been mixed with growth factor reduced matrigel at clonal density and overlaid with medium. At day 10 pictures have been taken and invasion/irregularity was quantified as described in material and method section.

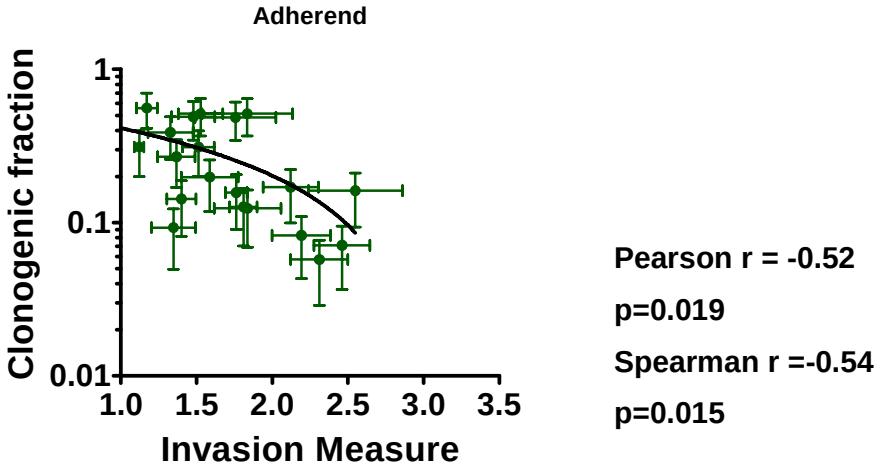
**Figure S4. In vitro validation: limiting dilution assay**



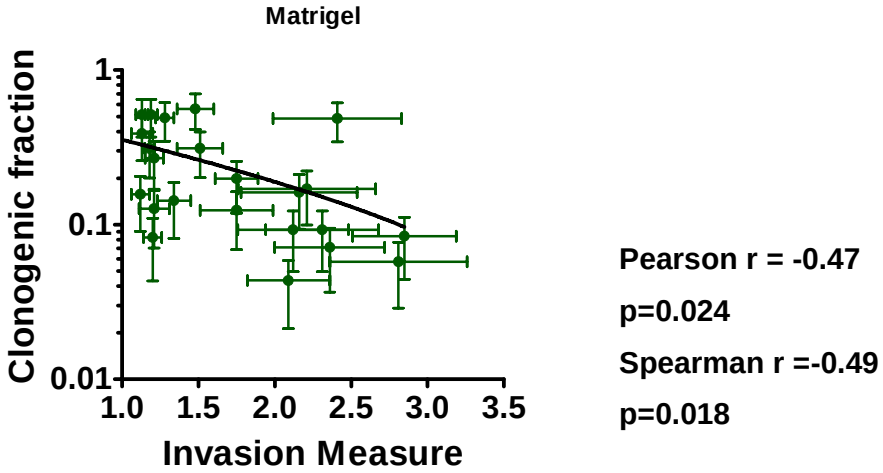

Cells from various cell lines have been deposited by FACS in a limiting dilution fashion. 1, 2, 4, 8, 16, 32, 64, 128, 256 cells have been sorted in a 96 well plate. Limiting dilution calculation has been performed using the *limdil* function of the 'statmod' software package (http://bioinf.wehi.edu.au/software/limdil/). Error bars represent 95% CI.

**Figure S5. In vitro validation; correlation of clonogenicity and Invasion Measure**

**A**

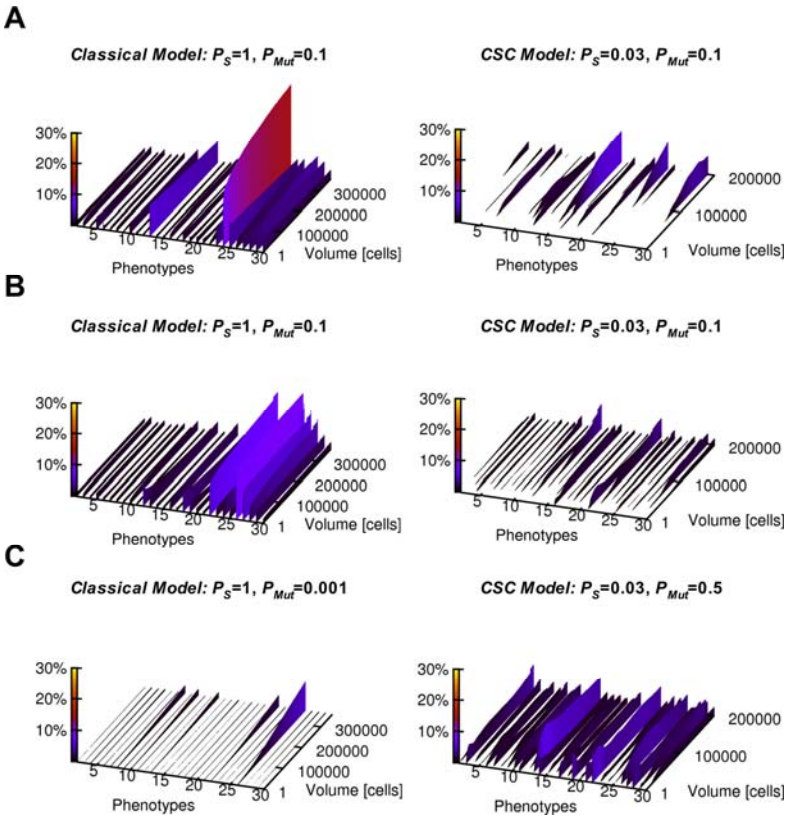



**B**




A significant correlation between clonogenicity and invasion/irregularity measure can be observed both on adherent plates (**A**) and in matrigel (**B**). Every dot Error bars for y-axis represent 95% confidence intervals for clonogenic fraction. Error bars on x-axis represent SEM for invasion measurement of at least 6 aggregates.

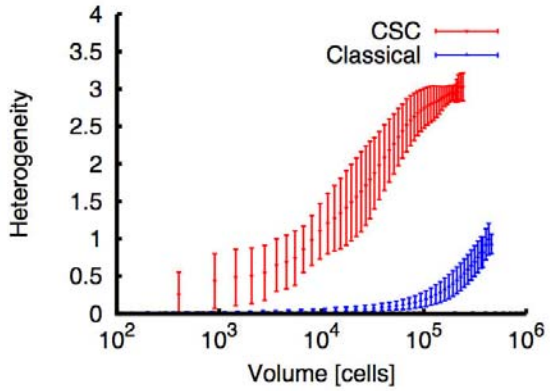
**Figure S6. Tumor Evolution and Phenotypical Selection in a Cancer Stem Cell Context**



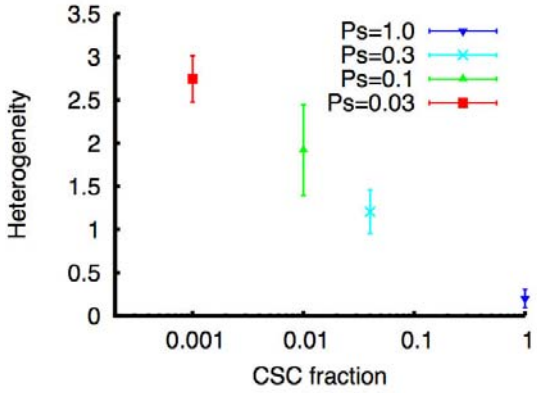

30 phenotypes are randomly generated and described in **Table S3**. (**A**) Diagrams showing a representative example of the phenotype distributions in the classical model compared to the CSC model with equal mutation rates $P_{Mut}$=0.1. (**B**, **C**) In these figures mean results are shown for 8 replicated experiments. (**B**) Diagrams showing the mean of 8 replicas of the process of phenotypical evolution in classical model compared to the CSC model. (**C**) Interestingly, while the pace at which new traits emerge is equal and all other parameters of the system are kept constant, the CSC model stimulates phenotypical heterogeneity. (**A**-**C**) The 3-dimensional plots show the cell fractions present for each newly generated phenotype. The coloration is a heat map linked to the cell fraction. For an individual example of the experiment see **Figure 4**.

**Figure S7. The Cancer Stem Cell Model Promotes Phenotypical Heterogeneity**

**A**



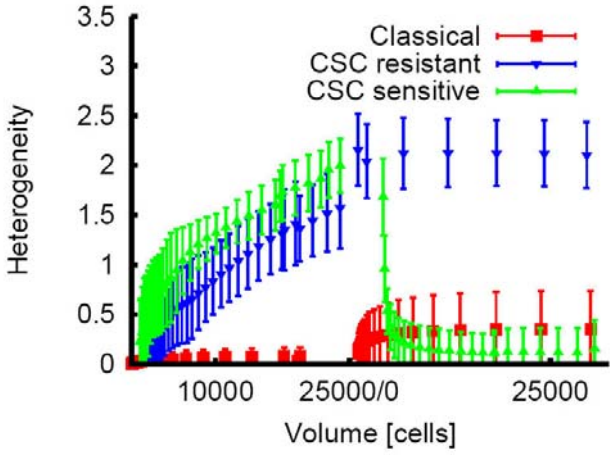

**B**




**(A)** Depicted is the Shannon index for heterogeneity (see **Supplemental Material and Methods** section) for the phenotypical evolution experiments presented in **Figure 4B** and **4C.** A significant increase in heterogeneity occurs in the cancer stem cell model ($P_S$=0.03, $P_{Mut}$=0.5; red) as compared to the classical model ($P_S$=1, $P_{Mut}$=0.001; blue).

**(B)** Moreover the heterogeneity shows a well-defined exponential relationship with the CSC fraction. n=8, error bars represent standard deviation.

**Figure S8. Effects of Therapy in Tumor Heterogeneity**




This plot illustrates the effects of treatment (applied at $V_T$=25000) on the heterogeneity of the models of growth. Interestingly, applying the therapy to the classical model or to the model where CSC are therapy resistant, results in a further increment of heterogeneity. In particular for the case of the CSC resistant settings, the boost in heterogeneity is consistent and maintained until the re-growth of the tumor. In the scenario where CSCs are sensitive to treatment instead, the clonal diversity drops instantly due to the few remaining CSCs. Therapy at intersection of x-axis. In this setting accumulation of phenotypes after therapeutic intervention occurs in prohibited illustrating the heterogeneity in the tumor cell population change as an effect of therapy more clearly. Heterogeneity represenst Shannon index (see **Supplemental Material and Methods** section).

## Supplemental tables

***Table S1. Summary of clonogenicity and invasion meaure for the cell lines used***

| Cell line | Origin | Clonogen fraction | Invasion Measure (adherend) | Invasion Measure (matrigel) |
|---|---|---|---|---|
| LS180 | Colon | 0.06 | 2.31 | 2.81 |
| HT29 | Colon | 0.56 | 1.17 | 1.48 |
| SW480 | Colon | 0.13 | 1.81 | 1.21 |
| SW837 | Colon | 0.16 | 1.76 | 1.12 |
| HCT116 | Colon | 0.49 | 1.76 | 2.41 |
| HCT15 | Colon | 0.31 | 1.12 | 1.18 |
| T47D | Breast | 0.07 | 2.46 | 2.36 |
| HCT | Colon | 0.39 | 1.32 | 1.13 |
| NIH-3T3 | Em. fibrobl. | 0.14 | 1.40 | 1.34 |
| MA11 | Breast | 0.27 | 1.59 | 1.75 |
| DLD1 | Colon | 0.27 | 1.37 | 1.21 |
| RKO | Colon | 0.31 | 1.51 | 1.51 |
| PC3 | Pancreas | 0.17 | 2.12 | 2.21 |
| SW1573 | Lung | 0.49 | 1.48 | 1.28 |
| V79 | Lung fibrobl. | 0.08 | 2.19 | 1.20 |
| R1 | Mouse ES | 0.08 | ND | 2.85 |
| U251 | GBM | 0.12 | 1.84 | 1.75 |
| U118 | GBM | 0.04 | ND | 2.09 |
| T98 | GBM | 0.09 | ND | 2.12 |
| U2OS | GBM | 0.51 | 1.83 | 1.13 |
| U87 | GBM | 0.10 | ND | ND |
| Panc-1 | Pancreas | 0.09 | 1.35 | 2.31 |
| Glio-06 | GBM | 0.51 | 1.53 | 1.19 |
| Mia-Paca | Pancreas | 0.16 | 2.56 | 2.16 |

***Table S2. Description and Range of Phenotypic Traits Subject to Selection***

| Parameter | Symbol | Value Range |
|---|---|---|
| $O_2$ consumption of proliferative cells | $\kappa_p$ | $\kappa_p - 6\ \kappa_p$ |
| $O_2$ consumption of senescent cells | $\kappa_s$ | $\kappa_s - 6\ \kappa_s$ |
| Cell cycle duration | $T$ | 12 - 30 hours |
| Random mobility | $D_n$ | $D_n - 6\ D_n$ |
| Cell-to-cell adhesion | $\alpha$ | 0 - 4 |

This table displays the parameters that were subjected to phenotypical selection in addition to the range. With these criteria we randomly generated a pool of 30 phenotypes that we used during the experiments. This set of phenotypes is displayed in detail in **Table S3**.

***Table S3. Characteristics of Randomly Generated Phenotypes***

| Phenotype ID | Consumption Proliferation | Consumption Quiescence | Cell cycle duration (h) | Adhesion coeffiecient | Mobility |
|---|---|---|---|---|---|
| 1 | 1.00 | 1.00 | 29.60 | 0 | 1.00 |
| 2 | 5.37 | 1.96 | 21.34 | 2 | 3.27 |
| 3 | 1.99 | 1.37 | 18.49 | 1 | 3.26 |
| 4 | 1.87 | 5.62 | 24.21 | 3 | 5.40 |
| 5 | 1.03 | 4.71 | 19.29 | 0 | 2.47 |
| 6 | 1.70 | 1.92 | 27.83 | 0 | 2.87 |
| 7 | 5.15 | 2.72 | 22.60 | 0 | 3.09 |
| 8 | 5.35 | 1.97 | 19.21 | 0 | 1.59 |
| 9 | 5.99 | 5.20 | 27.95 | 0 | 3.91 |
| 10 | 5.38 | 1.22 | 17.08 | 0 | 2.14 |
| 11 | 4.02 | 4.24 | 24.42 | 1 | 5.97 |
| 12 | 3.05 | 2.58 | 25.75 | 1 | 3.55 |
| 13 | 4.14 | 2.66 | 16.58 | 2 | 1.86 |
| 14 | 4.77 | 3.60 | 26.52 | 2 | 3.83 |
| 15 | 4.97 | 2.70 | 23.69 | 1 | 5.94 |
| 16 | 5.91 | 4.54 | 16.79 | 1 | 1.12 |
| 17 | 3.67 | 2.42 | 29.54 | 0 | 4.08 |
| 18 | 4.94 | 3.29 | 19.13 | 2 | 5.89 |
| 19 | 3.72 | 2.90 | 26.25 | 1 | 4.60 |
| 20 | 4.54 | 5.35 | 21.72 | 2 | 3.88 |
| 21 | 4.01 | 5.62 | 25.21 | 0 | 2.04 |
| 22 | 5.12 | 1.13 | 13.19 | 2 | 3.42 |
| 23 | 3.31 | 4.77 | 17.37 | 0 | 1.67 |
| 24 | 5.27 | 4.10 | 12.95 | 2 | 3.45 |
| 25 | 1.33 | 3.44 | 17.63 | 2 | 3.06 |
| 26 | 4.10 | 5.22 | 13.44 | 1 | 5.34 |
| 27 | 2.76 | 4.49 | 15.28 | 3 | 3.25 |
| 28 | 3.93 | 4.30 | 14.48 | 1 | 2.40 |
| 29 | 4.85 | 5.36 | 16.18 | 3 | 2.80 |
| 30 | 4.87 | 4.40 | 17.13 | 1 | 3.62 |

The table displays the random set of 30 phenotypes that we use in the phenotypical selection experiments (**Figure 4**). Red, grey and green indicate the arbitrary malignant potential with respect to that cellular trait.

## Supplemental movie legends

**Movie S1. Tumor growth in SCA model: Classical Model**

(**Figure 1D**) Development of tumor mass in the classical model ($P_S=1$) on a $600^2$ lattice. White: normal tissue, Blue: cells capable of proliferation, Brown: necrotic core.

**Movie S2. Tumor Growth in SCA model: Cancer Stem Cell Model**

(**Figure 1D**) Development of tumor mass in the CSC model ($P_S=0.03$) on a $600^2$ lattice. White: Normal tissue; Red: Cancer stem cells; Dark blue: cells capable of proliferation; Light blue: senescent differentiated cells; Brown: necrotic core.

**Movie S3. 3-Dimensional Tumor Morphology: Classical Model**

(**Figure 3B**) 3D tumor mass simulated with the classical model ($P_S=1$) on a $400^3$ lattice. White: cancer cells, Black: normal tissue.

**Movie S4. 3-Dimensional Tumor Morphology: Cancer Stem Cell Model**

(**Figure 3A**) 3D tumor mass simulated with the CSC model ($P_S=0.03$) on a $400^3$ lattice. White: cancer cells; Black: normal tissue.

**Movie S5. Phenotypical Selection: Classical Model**

(**Figure 4C**) Movie shows tumor development of the classical model ($P_S=1$) with mutation rate 0.001 per cell division. Different colors represent different clones.

**Movie S6. Phenotypical Selection: Cancer Stem Cell Model**

(**Figure 4C**) Movie shows tumor development of the CSC model ($P_S=0.03$) with mutation rate 0.5 per cell division. Different colors represent different clones. Clear spatial heterogeneity can be observed.

## References

1. Sloot PMA, Hoekstra AG. Modeling dynamic systems with cellular automata. In: Fishwick PA, editor. Handbook of dynamic system modeling: Chapman & Hall/CRC; 2007. p. 1-20.
2. Sherwood L. Human physiology: from cells to systems. Belmont: Wadsworth Publishing Company; 1997.
3. Anderson AR. A hybrid mathematical model of solid tumour invasion: the importance of cell adhesion. Math Med Biol 2005; 22: 163-86.
4. Anderson AR, Weaver AM, Cummings PT, Quaranta V. Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. Cell 2006; 127: 905-15.
5. Nowell PC. The clonal evolution of tumor cell populations. Science 1976; 194: 23-8.
6. Matsumoto K, Nishimura T. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. ACM Transactions on Modeling and Computer Simulation (TOMACS) 1998; 8: 3-30.
7. Stroustrup B. The C++ programming language. Special ed. Reading, Mass.: Addison-Wesley; 2000.
8. Shannon CE. A mathematical theory of communication. Bell Systems Technical Journal 1948; 27: 379-423.