

SUPPLEMENTAL INFORMATION

Bartlomiej Waclaw, Ivana Bozic, Meredith E. Pittman, Ralph H. Hruban, Bert Vogelstein, and Martin A. Nowak

1 Computer algorithm

In each time step, a random cell is chosen, its fate (replication or death) decided, and the time variable updated to account for the elapsed time since the last event. More specifically, for Model A,

1. A random tumor cell i is chosen. Let its genotype be g .
2. A nearest neighbor of cell i is picked at random and if it is empty, cell i replicates there and creates a new cell j with probability proportional to the birth rate b_g of cell i divided by the maximal birth rate b_{\max} of all cells present in the tumor. This ensures that the replication probability is always less or equal to one, and that it is proportional to the number of empty neighbors of i . In the case of no replication (site not empty or rejected), proceed to step 5.
3. Each of the two cells i, j receive n_i and n_j new mutations, with n_i, n_j drawn independently from the Poisson probability distribution (Eq. (1) in Methods). If either n_i, n_j is zero, the genotype of the corresponding cell does not change.
4. With probability M , the new cell j migrates and forms a new ball (microlesion) in the proximity of the surface of the ball from which it originates. The 3d position \vec{X}_j of the new ball is determined as $\vec{X}_j = \vec{X}_i + R_i \vec{x}_i / |\vec{x}_i|$, where \vec{X}_i denotes the position of the ball of cell i , R_i is the radius of that ball, and \vec{x}_i is the position of cell i relative to the center of the ball, see Extended Data Fig. 1a.
5. Cell i dies with probability equal to its death rate d_g divided by the maximal death rate d_{\max} of all existing tumor cells. If all cells within a ball die, the ball is removed.
6. Time is updated by a small increment $dt = 1/(b_{\max}N)$, where N is the total number of cells in the tumor. This ensures that b has the correct interpretation of probability per unit time, and the number of cells in an exponentially growing population (for example, for large M) increases as $N(t) \propto \exp[(b - d)t]$.

Therefore, cells replicate only if there is at least one empty site in the neighborhood, but die with rates independent of the number of empty neighbors. If a cell is completely surrounded by other cells, it cannot replicate. However, in simulations presented in the main text we always assume $d > 0$ so that even such “quiescent” cells can eventually resume replication if they survive long enough so that one of their neighbors die first and make space for their progeny.

Note that (i) each daughter cell receives different GAs, and that the average number of new GAs per daughter cell per replication is $\gamma/2$ in our notation, (ii) for small γ (realistic situation, see Methods), the probability of a single, new GA in any of the daughter cells is approximately equal to γ .

The simulation always begins with a single cell (and thus also a single ball of cells) at position $\vec{x} = (0, 0, 0)$. We assume that this cell already has a selective advantage over normal cells. We therefore do not model the process of tumor initiation e.g. by inactivation of tumor suppressor genes. The initial cell begins to replicate as described above, and once two or more balls of cells

are formed as the result of growth and short-range dispersal, the balls start to repel each other mechanically. This reduces the otherwise imminent and physically impossible overlap between the balls of cells. Because calculating forces and solving equations of motion for hundreds of balls every time a new cell is added to the tumor would be computationally too expensive and required a detailed biophysics model of the tissue, the algorithm we use to reduce the overlap is a less-realistic but more efficient “shoving” algorithm similar to the method used in the biofilm simulator iDynamics [64, 65]. The algorithm works by moving apart each pair of overlapping balls along the line joining their centers of mass, see Extended Data Fig. 1b, and is executed only if the radius R of any ball (measured as the distance from the center of the ball to the most distant cell) has increased by 5% or more compared to the last time the algorithm was used. A few iterations are usually necessary to relocate all overlapping balls, but because this is done rarely as compared to replication/death of individual cells, it does not have a huge impact on algorithm’s performance for typical scenarios we simulated (a few hundreds of balls).

For Model B, step (2) of the algorithm is modified as follows: If cell i has no empty neighbors, proceed to step 5. Otherwise, pick up one of them at random and, with probability b_g/b_{\max} , create a new cell j . If no replication, proceed to step 5. Model B differs from Model A in that the replication probability does not depend on the number of empty (normal) cells in the neighborhood.

For Model C, in step (2) cell i always replicates, regardless of whether there are empty sites nearby or not. The site to which the cell replicates is determined as follows. Ten randomly selected directions out of 26 possible directions (vectors (x, y, z) where $x, y, z = 0, -1, +1$ and x, y, z are not simultaneously zero) are scanned until an empty site is reached, and the direction along which the number of occupied sites is the smallest is chosen. The algorithm proceeds to the first site in this direction, and the scanning is repeated. The whole cycle (scanning/moving by one site) is continued until an empty site is reached. Then each cell encountered during the procedure is shifted to the position of the next site in the chain. This eventually creates an empty site near cell i to which it replicates. This algorithm mimics mechanical displacement of cells towards the surface caused by replication inside the tumor. The path of “minimal drag” chosen by the algorithm is self-avoiding, i.e., when choosing the direction for the next step the algorithm checks if a candidate site has been already visited and, if it has, this direction is discarded and another one is chosen at random. The procedure introduces small artifacts related to lattice symmetries, but deviations from the expected spherical shape of the lesion are negligible, c.f. Extended Data Fig. 6c.

For Model D, step (5) is modified as follows: cell i dies with probability equal to its death rate d_g multiplied by the fraction of empty nearby sites and divided by d_{\max} . Thus, both the replication and death rates are proportional to the number of empty neighbors. This causes that cells completely surrounded by other cancer cells not only do not replicate, but also do not die, and a “quiescent core” made of such cells arises inside the tumor.

The computer code (in C++) that simulates the process of tumor growth and treatment discussed in this work is available at <http://www2.ph.ed.ac.uk/~bwaclaw/cancer-code>. Given enough memory ($>10\text{GB}$) and time ($\sim \text{day}$), the code can simulate clinically meaningful (10^9 cells, $>1\text{cm}$) tumors.

2 Model parameters

Growth and death rates prior to treatment. The initial neoplastic/cancer cell has the birth rate $b = \ln 2 \approx 0.69 \text{ days}^{-1}$. This corresponds to a 24h minimum doubling time, though in Model A this is achieved only for cells surrounded by empty space (normal cells), and most cells (surrounded by other tumor cells) will replicate slower on average.

In the case of metastatic (secondary) lesions, the birth and death rates $b_g = b$, $d_g = d$ are constant and do not depend on the genotype g . The death rate d assumes values between zero and b , depending on how aggressive we want the tumor to be. As for primary lesions, in the simulations presented in the main text $b_g = b$ remains the same for all cells prior to treatment and is not affected by mutations. Driver mutations change only the death rate d_g as follows: $d_g = b(1 - s)^{k_g}$, where k_g is the number of driver mutations in genotype g including the initial driver mutation (hence $k_g \geq 1$), and s is the selective advantage of each new driver, similarly to Ref. [8]. In the case of primary lesions, the rate b sets the time scale of the model. There is no independent parameter d for the death rate which is expressed as a fraction of b .

We also consider the case in which $d_g = b(1 - s)$ and $b_g = b(1 + s)^{k_g-1}$, i.e. the initial cell replicates and dies respectively with rates b and $b(1 - s)$, and driver mutations increase the birth rate. As before, the only parameter that sets the time scale here is b , the birth rate of the initial cell, and the death rate d is the function of b and s . We show below that the results for this case are almost identical to the case from the previous paragraph where drivers reduce the rate of death. Consequently, the model behavior is determined by the difference between the birth and death rates, and it does not matter which of the two rates is affected by driver mutations.

For primary tumors we have explored a range of selective advantages $s = 0.5\% \dots 5\%$. Most of the results presented in the main text are for $s = 1\%$. For metastatic lesions, where we do not model driver mutations, we take $d/b = 0.5 \dots 0.9$, with lower death rates corresponding to more aggressive and faster growing tumors. This corresponds to the total selective advantage $0.5 \dots 0.1 = 50\% \dots 10\%$ over normal cells.

Mutation probabilities $\gamma, \gamma_d, \gamma_r$. These are strictly speaking the average numbers of GAs per parental cell per replication, but since these numbers are small we can think about them as the probabilities of all mutations (γ), driver mutations (γ_d), and resistant mutations (γ_r). The mutation probabilities are not known precisely and also vary among different cancers. Here we have used numbers that are representative for many cancers. For the total mutation probability we take $\gamma = 0.02$. This value corresponds to each of $\sim 6 \cdot 10^7$ bp in the diploid exome mutating with probability $\sim 3.3 \cdot 10^{-10}$ per bp per generation. For driver mutations, we take $\gamma_d = 4 \cdot 10^{-5}$ which is about the same as in Ref. [8]. For resistant mutations we assume that $\gamma_r = 10^{-7}$. This is equivalent, for example, to about 100 potential mutations that confer resistance to any drug targeting a specific protein in cancer cells, each of them occurring with probability $\sim 10^{-9}$, which is an upper estimate on the probability of point mutations in somatic cells [66, 67]. The value of γ_r is of the same order of magnitude as estimated in Ref. [68] for targeted EGFR blockade.

Dispersal probability M . Since no reliable *in vivo* data exists for this parameter, in our model we explore a range of values between $M = 10^{-7}$ and $M = 10^{-2}$. Our results for the growth time of a metastatic lesion indicate that $M = 10^{-7} \dots 10^{-6}$ gives a reasonable agreement with clinically observed tumor growth times, and hence we typically use values from this range. Note that M as defined in our model is the probability that a cell not only detaches from the lesion and moves to a nearby place, but that it also anchors there and starts proliferating.

3 Tumor growth rate for the neutral model

Figure 2 from the main paper and Extended Data Figure 2 show that in the absence of driver mutations ($s = 0$) and for $M = 0$ (no dispersal), tumors grow to a roughly spherical shape. The number of cells (and hence also the volume) of the tumor grows approximately as $N \sim T^3$, see Extended Data Fig. 3a,b. This is easy to understand - the inside of the tumor remains in the state of equilibrium between replication and death, and only the surface is able to expand and invade the surrounding tissue as the result of cancer cells' net growth advantage. Because the number of cells in the surface layer of a spherical tumor is proportional to $N^{2/3}$, the rate

of change of the total number of cells $dN(t)/dt$ is $\sim N(t)^{2/3}$ which gives $N(t) = At^3$. The proportionality coefficient A depends on the thickness of the outermost layer of cells that is not at equilibrium and on the net growth rate of cells in this layer, and in general is a complicated function of b and d which we could not determine analytically.

For $M > 0$, however, the model exhibits a transition to exponential growth when the size N becomes much larger than $1/M$, see Fig. 2. This can be explained as follows. Let $N(n, t)$ be the number of balls of cells having n cells at time t . Since in our model the balls do not interact with one another (besides mechanical repulsion) and their growth is unaffected by the presence of other balls, we can write the following Master equation for the rate of change of $N(n, t)$ in time:

$$\frac{dN(n, t)}{dt} = \delta_{n,1} M \sum_{k=1}^{\infty} B(k) N(k, t) + (1 - M) B(n-1) N(n-1, t) - (1 - M) B(n) N(n, t), \quad (1)$$

where $B(n)$ is the rate at which a ball of size n grows to size $n+1$. The first term in Eq. (1) proportional to M accounts for the production of new balls of cells due to dispersal. This happens with the rate equal to the total rate of replication, hence the sum over balls of all sizes $k = 1, \dots, \infty$ weighted by $N(k, t)$, the number of balls of size k . The second term is the gain term due to replication in a ball of size $n-1$ which creates a ball of size n , and the last term is the loss term due to a ball of size n increasing its size by one cell.

If we knew $B(n)$ exactly, we could in principle solve Eq. (1) and find the total size $N(t) \equiv \sum_n n N(n, t)$ as a function of time. Unfortunately, it is not easy to determine $B(n)$ analytically for arbitrary n because of the vast number (and the lack of regularity of) configurations that cells can take in a single ball of size $n \gg 1$. This problem was identified many years ago in the Eden model [62], and it has no known exact solution for large n . However, here we are only interested in the asymptotic behavior of $B(n)$ for large sizes n because we anticipate that only this large- n behavior will affect the tumor growth rate at large times. From our earlier results for $M = 0$ we conclude that $B(n) \sim n^{2/3}$ for $n \rightarrow \infty$ and although finite-size corrections to this formula can be quite strong, we shall neglect them here for the sake of simplicity.

Let us now discuss the large- t solution of Eq. (1). We shall keep a general form of $B(n)$ and replace it by $B(n) \sim n^{2/3}$ in the final step. Equation (1) is a linear differential equation in $N(n, t)$ and therefore its long-time behavior must be dominated by

$$N(n, t) \cong e^{B_{\text{exp}} t} p(n), \quad (2)$$

where B_{exp} plays the role of the largest eigenvalue of the linear operator acting on $N(n, t)$ on the r.h.s. of Eq. (1). Its biological interpretation is that of the exponential growth rate of the whole tumor,

$$N(t) = \sum_n n N(n, t) \cong e^{B_{\text{exp}} t} \sum_n n p(n), \quad (3)$$

and $p(n)$ is now a time-independent, stationary distribution of ball sizes. One could think that we have just shown that $N(t)$ grows exponentially, but this is not true; it remains to be seen that $B_{\text{exp}} = \text{const} > 0$ for $M > 0$, otherwise sub-leading terms could lead to a sub-exponential (i.e. power-law as we have seen above) growth.

To proceed, we insert Eq. (2) into Eq. (1) and obtain the following recursion relation for $p(n)$:

$$B_{\text{exp}} p(n) = \delta_{n,1} M \sum_k p(k) B(k) + (1 - M) B(n-1) p(n-1) - (1 - M) B(n) p(n). \quad (4)$$

For $n = 1$ this reduces to

$$p(1) = M \frac{\sum_k p(k) B(k)}{B_{\text{exp}} + (1 - M) B(1)}, \quad (5)$$

whereas for $n > 1$ we can rewrite Eq. (4) and express $p(n)$ as a function of $p(n - 1)$ as follows

$$p(n) = p(n - 1) \frac{B(n - 1)(1 - M)}{B_{\text{exp}} + B(n)(1 - M)}. \quad (6)$$

Iterating the above equation we obtain $p(n)$ as a product of n fractions as in the above equation, times $p(1)$. If we now use this to expand $\sum_k p(k)B(k)$, we have that

$$\sum_k p(k)B(k) = \left[1 + \sum_{k=2}^{\infty} B(k) \left(\prod_{n=2}^k \frac{B(n - 1)(1 - M)}{B_{\text{exp}} + B(n)(1 - M)} \right) \right] M \frac{\sum_k p(k)B(k)}{B_{\text{exp}} + (1 - M)B(1)}, \quad (7)$$

and after dividing both sides by the common factor $\sum_k p(k)B(k)$ we obtain that M obeys the following equation:

$$\frac{1}{M} = \frac{1}{B_{\text{exp}} + B(1)(1 - M)} \left[1 + \sum_{k=2}^{\infty} B(k) \left(\prod_{n=2}^k \frac{B(n - 1)(1 - M)}{B_{\text{exp}} + B(n)(1 - M)} \right) \right]. \quad (8)$$

The above equation allows us to calculate B_{exp} as a function of the dispersal probability M , albeit in general it must be solved numerically. However, if we now use the asymptotic form $B(n) \sim n^{2/3}$, upon inserting it into Eq. (8) and canceling lower-order terms in M we have

$$B_{\text{exp}} \sim M^{1/3} \quad (9)$$

for $M \ll 1$. This concludes our proof that the rate of exponential growth is non-zero for $M > 0$. Equation (9) also predicts that B_{exp} is proportional to the cubic root of the dispersal probability M . Numerical simulations (results not shown) indicate that, for $M > 10^{-6}$, the actual B_{exp} increase slower with M than $\sim M^{0.33}$, approximately as $M^{0.25...0.3}$. This is caused by deviations from the law $B(n) \sim n^{2/3}$ for balls smaller than $\sim 10^7$ cells, for which $B(n)$ is better approximated by $\sim n^{0.7...0.75}$.

4 Reseeding and long-range migration

In our model each cell that migrates from its original site starts a new ball of cells and there are no interactions (besides mechanical repulsion) between balls of cells. Thus, if we neglect spatial distribution of genetic alterations (GAs) and focus only on the total number of GAs and whole-tumor growth, our algorithm can also model long-range metastasis and reseeding of new lesions. All our predictions except those related to the spatial distribution of GAs remain valid for the reseeding model. For example, we have shown that, with short-range dispersal, tumor growth becomes exponential with rate $\sim M^{1/3}$, where M is the dispersal probability. If we also include the possibility of reseeding (long-range migration) with probability R , the total mass will grow at an increased rate $\sim (M + R)^{1/3}$.

5 Accumulation of driver mutations

Extended Data Figure 5a-c suggests that the number of drivers per cell increases linearly in time. We have made simulations of larger tumors (up to 10^9 cells), and different s , and although the growth rate increases in time, the drivers accumulate approximately at a constant rate (data not shown). This is in contrast to what has been shown in Ref. [8] for exponentially growing tumors without any spatial structure, where the number of drivers increases exponentially in time.

This counter-intuitive outcome is the result of the spatial structure of the tumor and all driver mutations having the same selective advantage s in our model. If s is small, a new driver arising in the background of the previous driver mutations spreads within the population with the same speed as the previous drivers, equal to the speed with which the tumor expands into the normal tissue. This expansion, which is similar to travelling fronts in other spatial models, can be demonstrated mathematically. Our statement remains true for sufficiently large tumors in which most cells are inside the ball and not on the surface. Therefore, apart from first few drivers, the i -th clonal subpopulation follows the same growth law as the $(i - 1)$ -th clone but it starts at a later time t_i , i.e. its linear extension (“diameter”) $l_i(t) = l(t - t_i)$ where the function $l(t)$ is approximately the same for all clones. Moreover, the times t_i at which consecutive driver mutations arise are approximately equally spaced in time. This is because the rate at which new drivers are produced increases very slowly with i . This symmetry upon time translation causes linear accumulation of driver mutations over time. We plan to investigate this effect in a subsequent publication.

6 Alternative simulation method: kinetic Monte Carlo (KMC) algorithm

Although we speak about rates of different processes, our simulation algorithm is not an exact KMC such as e.g. the Gillespie algorithm. The reason we did not use the standard algorithm was that a fully stochastic simulation would be much slower than our algorithm, even when more advanced methods than the Gillespie algorithm are used. As we will now show, our algorithm offers a significant gain in speed at the expense of small inaccuracy as compared to KMC.

We simulated models A and C using a KMC, with cells undergoing independent stochastic birth/death events with rates b and d , and our algorithm, for a number of cases (different d and s). It took 5s to simulate 100 instances of Model A for $N = 10^4$ and no death using our algorithm, whereas it took 17s for a fast KMC (similar to the one used in Ref. [101]) and 3.5h for the Gillespie algorithm. Despite this significant difference in speed, we typically found only small differences between the algorithms (plots not shown), affecting both models A and C in a different way. The biggest difference we found was in $N(t)$ for large t and $d \rightarrow b$.

If Model A is simulated using the KMC algorithm, zero net growth rate does no longer correspond to $d = b$, but to $d \approx 0.897b$. This can be explained as follows: as the difference $b - d$ decreases, the density of cancer cells decreases as well, and the average fraction of empty neighbour sites surrounding a cancer cell increases. However, this fraction never increases beyond 0.897 (as determined numerically) for d approaching b . Since the replication rate in Model A is proportional to the number of empty neighbors, the maximal possible average replication rate is therefore $0.897b$. The net growth rate is thus zero for $d = 0.897b$, as this is when replication and death (occurring with constant rate identical for each cancer cell) balance each other. The reason why the factor 0.897 does not arise in our non-KMC simulation is that division and growth are not entirely independent in this algorithm. If and only if a cell replicates, it attempts to undergo apoptosis with probability d_g/d_{\max} . Killing daughter cells enables the average fraction of empty sites to increase to 1 with decreasing difference $b - d$, and beyond 0.897 from the KMC algorithm. Therefore, the average replication rate tends conveniently to b as $b \rightarrow d$.

To compare both algorithms, we therefore replaced d by $d^* = 0.897d$ in the KMC algorithm so that cells were dying with rate d^* . Consecutively, if s is to be interpreted as the selective advantage, the initial $(b - d)/d$ (the selective advantage of the first driver) used in the KMC algorithm must be multiplied by 0.897 to make the results comparable to our algorithm. In this way, the zero net growth rate corresponds to $d = b$ and the same value of s can be used in both

algorithms, which behave now very similarly (plots not shown); $N(t)$ differs very little by eye between the two models.

For Model C in which cells replicate with constant rate and hence the average size increases exponentially as $N(t) \propto \exp[(b-d)t]$ in the absence of drivers, both the KMC and our algorithm agree exactly. In the presence of drivers, however, there is a small difference between the two algorithms (data not shown). This difference can be decreased by doubling the frequency of driver mutations in the KMC algorithm. This is again related to the fact that replication (during which driver mutations occur) and death are not independent in our algorithm.

7 6-site model

To see how sensitive our model is to how the neighborhood of a cell is defined, we simulated models A and C for the case in which each cell has only 6 neighbors (von Neumann neighborhood) instead of 26 (Moore neighborhood). The results are shown in Extended Data Fig. 9. Red curves are for the 26-sites neighborhood, green curves are for the 6-sites neighbourhood. Panels a,b are for Model A with $d = 0$ and $d = 0.8b$, respectively. Simulation snapshots presented in these panels show that the shapes for 6- and 26-sites neighborhood are indistinguishable by eye for $N = 10^7$ cells, but that the ball of cells expands radially $\sim 2x$ slower for the 6-sites model. This reduction in the growth rate is due to the inability of the 6-site model to replicate to diagonal sites even if they are empty. This can be easily remedied if the replication rate is "corrected" to account for the reduced number of available sites. For example, by letting a cell to attempt one more replication (point 2 of the algorithm in Section 1) if the previous one was unsuccessful due to an occupied site, the growth curve (blue lines in panel b) is almost identical to that in the 26-sites model.

Panel c shows the case for driver mutations when $s = 5\%$. The asymmetry in the shape is caused by non-uniform expansion of clonal sectors with different numbers of driver gene mutations. The difference in the growth rate between the two models is smaller than in panels a,b because the density of occupied sites is smaller due to b being very close to d . Panel d compares the two models in the case of exponential growth and cells pushing away other cells (Model C). This situation is very similar to Ref. [49]. The shape deviates now visibly from spherical symmetry in both cases, but more for the 6-sites model. There is no difference in the growth rate between the 6- and 26-site models because replication is not affected by the number of empty neighbours in Model C.

8 Mechanical repulsion of balls of cells and the off-lattice model

One of our assumptions in the lattice model is the existence of isolated balls of cells. Although this assumption is not needed for the increase in the growth rate and high levels of relatedness postulated in our paper, we have investigated under what conditions balls of cells repel each other. Our goal was only to demonstrate the plausibility of the assumption of repelling balls and hence we considered only a 2d model, as is sometimes done to reduce computational complexity when modelling tissues [58,93]. We also assumed periodic boundary conditions.

Because most cancers are of epithelial origin (carcinomas) [89], we focused on the case in which cancer arises in epithelial linings (e.g., ducts of the breast or pancreas). We modeled an $L \times L$ fragment of such tissue as ring-like regions of proliferating epithelial cells surrounding empty areas ("ducts") and separated by non-proliferating stroma composed of extracellular matrix (ECM) and mesenchymal cells, see Supplementary Video 8.

All cells are represented as very short rods with spherical caps of diameter d . This is similar to the approach taken in Ref. [59] where a cell was represented as a pair of connected, partially overlapping spheres. The stroma is represented by non-proliferating mesenchymal cells interacting through spring-like forces which model the behavior of the elastic medium (ECM) surrounding the cells. The force between two connected mesenchymal cells is given by Hooke's law:

$$F = E_s x(d/2), \quad (10)$$

where x denotes the displacement from the equilibrium configuration which is also the overlap distance between the two cells (positive if they overlap, negative if they are separated), and the dimensional factor $d/2$ ensures that F has proper units (N) if E is expressed in N/m² = Pa. The above equation approximates the stroma as a linear elastic medium; our choice of the multiplicative constant ($d/2$) guarantees that the Young elastic modulus of the medium is equal to E_s . We have confirmed this by simulating compression (with known force) of the modelled tissue and measuring the displacement. Two mesenchymal cells become connected when $x > 0$ for the first time, and either remain connected forever or (in a variant of the model, see below) the connection breaks when $x < -1\mu\text{m}$.

Epithelial (green) and mesenchymal (blue) cells adhere to each other. Mechanical interactions between these cells are predicated upon the JKR model [59, 60]. The contact force F and the contact area a can be calculated by solving a pair of equations:

$$x = \frac{4a^2}{d} - \left(\frac{2\pi a\sigma}{E_{ec}} \right)^{1/2}, \quad (11)$$

$$a^3 = \frac{(3/4)d \left(((3/2)\pi F d\sigma + ((3/4)\pi d\sigma)^2)^{1/2} + F + 3(\pi/4)d\sigma \right)}{4E_{ec}}, \quad (12)$$

where E_{ec} is the Young modulus of cancer, epithelial, and mesenchymal cells (assumed the same for all cells), and σ is the cell-cell adhesion energy. Cancer (red) cells are assumed to have lost adhesion-mediating molecules [29] and hence $\sigma = 0$ if one of the interacting cells is a cancer cell.

Growth is modelled as a linear increase in cell's length (i.e. the separation between the two spherical caps) over time. When the length becomes equal to the diameter d of each cap, the cell splits into two spherical cells with randomly assigned directions in which they will elongate. Growth stops if the cell is surrounded by more than a certain number of cells (5 or more for non-neoplastic and 6 or more for cancer cells). This growth cessation results from various forms of contact inhibition (weaker for cancer cells) and may be related to nutrient depletion or mechanical stress [15, 55, 85], but the model does not require, and is independent of, any specific mechanisms. To simulate the effect of nutrients diffusing into the epithelium from blood vessels embedded in the stroma, epithelial cells replicate only if they are in contact with mesenchymal cells. Similarly, cancer cells replicate only when in contact with either mesenchymal or non-neoplastic epithelial cells. Non-neoplastic epithelial cells are shed (removed) from the upper layer of normal epithelium, while cancer and mesenchymal cells are assumed to be immortal for simplicity. The parameters of the model are listed in a table in Extended Data Figure 7a; their values have been taken from the literature.

To generate the initial configuration of cells, a few mesenchymal cells are distributed in the system and allowed to replicate until they fill the available space. Then all cells inside regularly arranged circles of radius 50 μm ("ducts") are removed, and the cells on the walls of thus formed cavities are transformed into epithelial cells. The tissue is then simulated for some time until it equilibrates, see Supplementary Video 8.

Supplementary Video 9 shows what happens when two cancer cells, presumably invaders from the same primary lesion, begin to proliferate in two nearby ducts. The emerging balls of

cancer cells expand in size but remain detached. Expansion compresses the surrounding tissue and causes nearby ducts to shrink. We think this is a plausible scenario that explains what is seen in naturally occurring human tumors, such as those shown in Fig. 1.

In Supplementary Video 9, growth eventually ceases when the density of cancer cells becomes too high, because the stroma opposes further expansion. If mesenchymal cells are able to replicate in response to the stretching of ECM, growth slows down less (Supplementary Video 10). If ECM can break upon stretching, the balls can expand even more, but they eventually merge (Supplementary Video 11).

We also explored what assumptions are critical for the mechanical repulsion of balls of cells. It seems that as long as cancerous lesions are surrounded by elastic medium that is stiff enough to retain its structure, and is not "fluidized" by cellular movement or turnover, the lesions will repel. This is the case for real, healthy tissues which must remain structured in order to fulfill their function. If a tissue is "fluidized" by replication and death [80], balls of cancer cells will merge (Supplementary Video 12).