Supplemental Information

Supplemental Figure Legends

Supplemental Appendix: Derivation of inequalities for game theoretical model.

Supplemental Table 1: Model Assumptions

Supplemental Figure 1. Proliferation rates under growth factor-rich and –deprived conditions. (A) Experimental proliferation data were collected by plating $3x10^4$ cells/well in 12-well plates in complete growth medium and allowed to adhere for 8 h before counting (0-h) and replenishing with fresh growth media (growth factor-rich) or washing into serum-free DMEM-F12 media (growth factor-deprived). Thereafter, cells were counted with a hemacytometer using trypan blue exclusion to identify live cells, every 12 h, for 84 h. Growth curves for all cell lines represent the mean +/- standard error in that condition (N=3). (B) and (C): Since observed growth data were found to be exponential in nature ($y=a*\exp(g*t)$, where y=cell number, a=scaling constant, g=growth rate, and t=hours; $R^2>0.90$), growth rates (g) were achieved by fitting each proliferation curve with an exponential growth equation . Growth rates obtained from these equations were then used to calculate doubling times for each cell line (doubling time=ln*2/g).

Supplemental Figure 2. Invadopodia assay used to measure cell-associated ECM degradation. Cells were cultured overnight on FITC-fibronectin-coated gelatin films (green) on glass Mat-Tek dishes in assay media with 5% Horse serum and 5 ng/ml EGF before fixation and staining with rhodamine-phalloidin for actin rich invadopodia puncta (red). Representative 40X images are shown. FITC-gelatin degradation is evident as dark areas in the green background (arrowhead) and is frequently associated with active invadopodia protrusions (red puncta, examples pointed out with arrows). Cell area was determined by manually tracing the outline of cells using the F-actin staining to define the footprint of the cell, followed by calculation using the region of interest tool in Metamorph software. Degradation area was determined by performing an inclusive threshold of the FITC channel to include the dark, degraded areas: then, the region of interest tool was used to calculate the thresholded area. Scale bar = 10 µm.

Supplemental Figure 3. Dispase disaggregation assay (DDA) used to measure cell-cell adhesion strength. (A) Diagram depicting the protocol followed: an adherent cell monolayer is grown in fully supplemented growth media in duplicate-60mm dishes until confluency. Confluent cultures were washed with PBS and then incubated in 4.8 U of the ECM-cleaving bacterial metalloproteinase dispase for 30 min at 34°C. Released monolayers were resuspended in PBS (5 ml) and sedimented by low speed centrifugation. The pellet was resuspended in PBS and mechanically disrupted by standardized pipetting (30 times). Cell fragments are counted. The percent cell dissociation under these conditions is the readout of cell-cell junction strength; thus, a higher cell dissociation percentage indicates lower cell-cell adhesion strength. (B) Quantification of percent dissociation for all cell expressed as the percentage of single cells released by mechanical disruption after dispase treatment versus the total number of cells recovered after subsequent treatment of the same samples with trypsin (% dissociation). MCF10A shRNA E-cadherin (shEcad) and shRNA p120 (shp120) are positive controls, with a high level of cell dissociation likely due to loss of E-cadherin (C). All data is reported as mean +/- standard error. (C) Western blot showing the level of expression for E-cadherin, and p120 in the shRNA MCF10A control cells lines.

Supplemental Appendix (for Game Theory):

Given the table in Fig 6A, the fitness of cells that are, relatively, independent from their microenvironment (mEI) and that of cells that rely on their microenvironment (mED) is given by W(mEI) and W (mED). The overall fitness of the mEI population is given by function W (phenotype):

$$W(mEI) = h.$$

Since the fitness of mED cells depends on the proportion p of mEI cells in the tumour population and on the phenotypic interactions defined in the payoff table, the overall fitness W in the population is:

$$W(mED) = p(1-c) + (1-p)(1-nc)$$
$$= pc(n-1) + (1-nc).$$

To calculate the proportion *p* of mEI cells in a biclonal population with both types of cells, we use the Bishop Cannings theorem which states that, in equilibrium, the fitness of the two populations are the same.

$$W(mEI) = W(mED) \Rightarrow$$
$$h = pc(n-1) + (1-nc) \Rightarrow$$
$$pc(n-1) = h + nc - 1 \Rightarrow$$
$$p = \frac{h + nc - 1}{c(n-1)}.$$

Assuming that the cost of relying on the mE is twice as much in environments in which more than one mED cell has to share available resources (n=2) then,

$$p = \frac{2c + h - 1}{c}.$$

The proportion *p'* of mED cells is then p' = 1-p.

.

Although in many situations the tumour evolves to be monoclonal, during the transition from clonal dominance by mED to mEI phenotypes there is a region of coexistence. The

magnitude of this region is dependent on both the mEI fitness and mE parameters and can be found using the expresion for p deduced before. Given that a dominating mEI would lead to p equal to 1, we can deduce the value of c that leads to that situation:

$$p = 1 \Rightarrow c = 1 - h$$

on the other hand a dominating mED population should lead to a p equal to 1 in which case c can be deduced as:

$$p = 0 \Rightarrow 2c + h - 1 = 0 \Rightarrow c = \frac{1-h}{2}$$

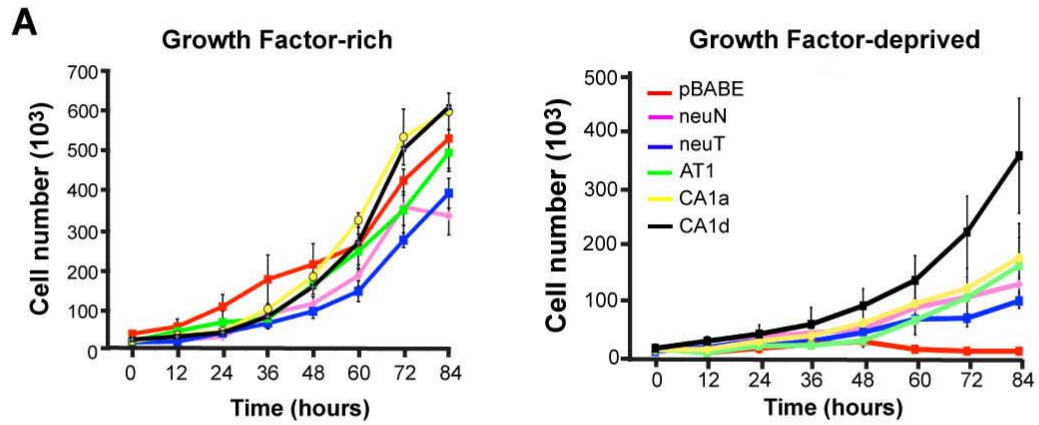
thus, the region of coexistance will be given by the region in which p is smaller than 1 and greater than 0:

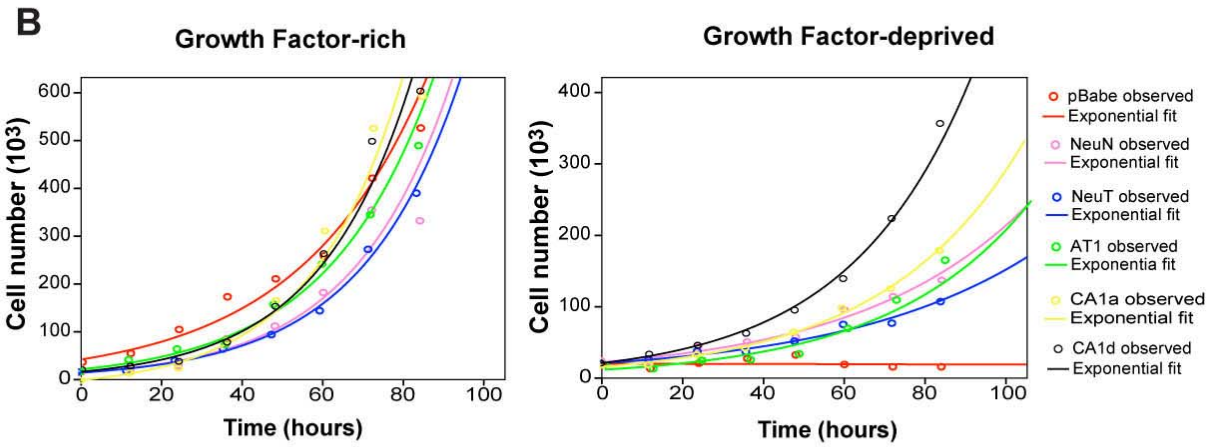
$$\frac{1-h}{2} \leq c \leq 1 - h$$

Supplementary Table 1: Model Details/Assumptions

Reduction of model
a. Removed all dependence on nutrients (such as oxygen)
b. Removed explicit model of matrix degrading enzymes
c. Removed biased migration towards increasing ECM density i.e. haptotaxis
Enhancements
a. Incorporated the physical constraint of the ECM upon both migration and proliferation
b. ECM degradation is modeled as direct contact degradation i.e. where a tumor cell is the ECM is degraded
Tumor Cell Assumptions
a. Migrate only via random motility modulated by both cell-cell adhesion and ECM constraints
b. Proliferate at a constant rate defined by phenotype provided these conditions are satisfied: i. Cell has reached its maturation age ii. Sufficient space around the cell exists for a daughter cell iii. ECM density at the current cell position is lower than the predefined constraint (f_m)
c. Cell-cell adhesion is modeled only from a one sided perspective, the idea behind this is that and individual cell has a given number of adhesion receptors that it can share with neighboring cells, the higher this value the more adhesive the cell is to its neighbors. It is implemented as follows, for a cell with adhesion value A_i: i. Count the number of neighboring (orthogonal) cells, A_n ii. Only if $A_n > A_i$ then the cell can move
ECM Assumptions
a. ECM is a heterogeneous density varying at the same scale as a cell
b. ECM does not move, diffuse or get produced
c. ECM degradation only occurs at grid points where tumor cells are located and at a rate defined in Table 1 for the cell under consideration
d. ECM acts as a physical constraint to both migration and proliferation, such that a certain density of ECM must be achieved (f_m) before a cell can move or proliferate.
Implementation of Migration
a. Cell migration is assumed to be diffusive, as controlled by movement probabilities derived from a discretization of the diffusion equation, see [5] for a formal derivation of this relationship. The probabilities for a cell to remain stationary, or move one grid point left, right, up or down are as follows: $$P_0 = 1 - \frac{4kD_n}{h^2}, P_1 = \frac{kD_n}{h^2}, P_2 = \frac{kD_n}{h^2}, P_3 = \frac{kD_n}{h^2}, P_4 = \frac{kD_n}{h^2}$$
b. Due to the way the movement probabilities are derived they will remain between 0 and 1 provided the space (h) and time (k) steps are chosen such that $4kD_n/h^2<1$, where D_n is the cell diffusion rate.
c. At each time step of the simulation for every cell we calculate the movement probabilities (as above) and then compute movement ranges by summing the probabilities to produce 5 ranges, $R_0 = 0$ to P_0 and $R_i = \sum_{j=0}^{i-1} P_j$ to $\sum_{j=0}^{i} P_j$, where $i = 1$ to 4. We then generate a random number between 0 and 1, and depending on the range which this number falls in, the current individual tumor cell under consideration will remain stationary (R_0) or move left (R_1), right (R_2), down (R_3) or up (R_4) modulated by cell-cell adhesion and ECM constraints
d. Cell movement will rarely be truly diffusive as even when a movement is chosen a cell cannot move for any of the following reasons: i. The grid point the cell wants to move to is already occupied by another cell ii. Its cell-cell adhesion criteria are not met iii. It had not satisfied the ECM constraint

Supplemental Figure 1





C

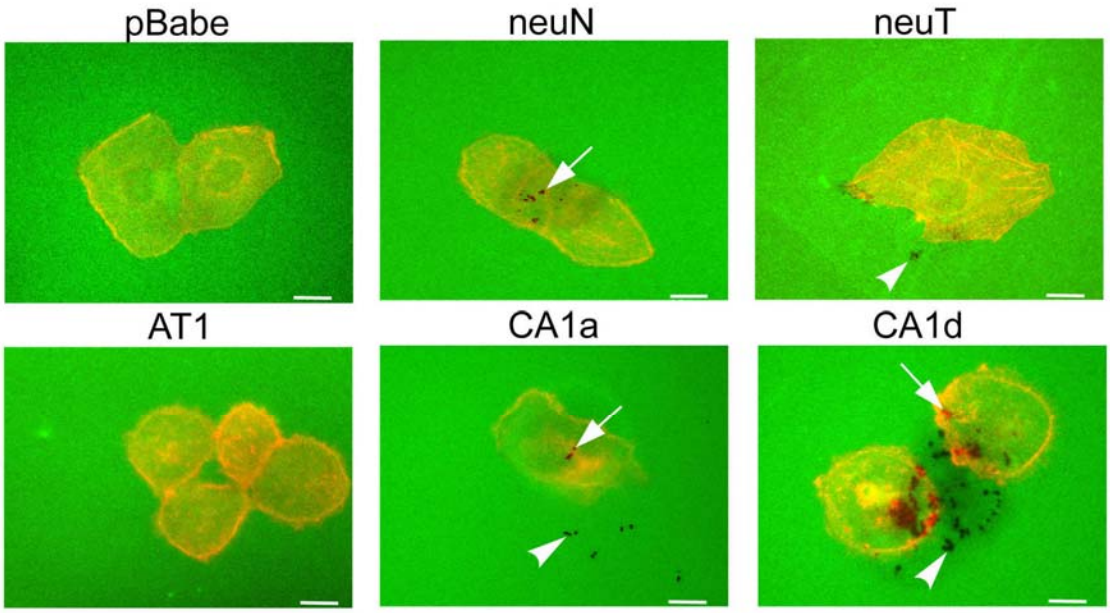
Growth Factor-rich

Cell line	Exponential fits	R^2	Doubling time (hours)
pBabe	$y=42831e^{0.031t}$	0.973	22.36
NeuN	$y=14875e^{0.040t}$	0.964	17.34
NeuT	$y=13568e^{0.040t}$	0.990	17.33
AT1	$y=23037e^{0.037t}$	0.980	18.73
CA1a	$y=18815e^{0.044t}$	0.983	15.75
CA1d	$y=18513e^{0.043t}$	0.984	16.12

Growth Factor-deprived

Cell line	Exponential fits	R^2	Doubling time (hours)
pBabe	$y=20186e^{-0.000t}$	0.003	N/A
NeuN	$y=19460e^{0.023t}$	0.951	30.14
NeuT	$y=17257e^{0.021t}$	0.973	33.01
AT1	$y=12095e^{0.028t}$	0.904	24.76
CA1a	$y=16106e^{0.028t}$	0.988	24.76
CA1d	$y=20676e^{0.024t}$	0.995	21.66

Supplemental Fig 2



Supplemental Fig 3