A phenotypic switch in the dispersal strategy of breast cancer cells selected for metastatic colonization

Article

Accepted Version

Title

A phenotypic switch in the dispersal strategy of breast cancer cells selected for metastatic colonisation

Authors
George Butler,[1] Shirley Keeton,[1] Louise Johnson[1], Philip Dash[1]*

Affiliations
[1] School of Biological Sciences, University of Reading, UK

*Corresponding author

Email: p.r.dash@reading.ac.uk

Abstract

An important question in cancer evolution concerns which traits make a cell likely to successfully metastasise. Cell motility phenotypes, mediated by cell shape change, are strong candidates. We experimentally evolved breast cancer cells in vitro for metastatic capability, using selective regimes designed to simulate stages of metastasis, then quantified their motility behaviours using computer vision. All evolved lines showed changes to motility phenotypes, and we have identified a previously unknown density-dependent motility phenotype only seen in cells selected for colonisation of decellularized lung tissue. These cells increase their rate of morphological change with an increase in migration speed when local cell density is high. However, when the local cell density is low, we find the opposite relationship: the rate of morphological change decreases with an increase in migration speed. Neither the ancestral population, nor cells selected for their ability to escape or invade extracellular matrix-like environments, display this dynamic behavioural switch. Our results suggest that cells capable of distant site colonisation may be characterised by dynamic morphological phenotypes and the capacity to respond to the local social environment.

Main Text

Introduction

Metastasis is a form of long-range dispersal (1,2) and central to understanding how cancers metastasise is understanding how cells migrate (3,4). During migration, as cancer cells become more invasive and begin to migrate independently, they adopt an altered morphology, typically taking on elongated shapes characteristic of epithelial-mesenchymal transition (EMT) (5,6). This change in cellular morphology is an important marker of migratory state (7,8). Quantitative measures of cell morphology taken from static images have been shown to effectively differentiate between cancer cell lines with high and low metastatic potential (9,10). However, there are important aspects of migratory behaviour linked to metastasis that cannot be measured from static images.

Successful metastasis requires a cell to navigate through a series of sequential steps known as the metastatic cascade. The cascade begins with a cell escaping from the primary tumour before migrating through the extracellular matrix (ECM) towards a nearby blood vessel. The cell must then intravasate into the blood before it is carried around the body. After reaching a distant site the cell then needs to extravasate from the blood and invade the foreign tissue. Finally, the cell must reinitiate aggressive proliferation enabling a secondary tumour to form (11).

In addition to the cellular changes needed for metastatic success, environmental changes are also necessary for a cell to metastasise (12). This is evident at the onset of cellular dispersal where nearby collagen fibres are straightened perpendicular to the tumour boundary (13). The straightened fibres then act as a pathway for future migrants in turn improving their migratory success (14). This dynamic cell-environment interplay continues throughout the metastatic cascade.

To identify the precise changes in cell phenotype that are associated with metastatic success, it is preferable to compare cell lines that differ only in their ability to metastasise. Experimental evolution (15), a powerful approach that has led to major advances in evolutionary biology, is now being applied to cancer evolution and provides the means to generate such cell lines (16,17). Initially identical populations of cancer cells can be selected in replicate for specific capabilities (18). We experimentally evolved populations of cancer cells using selective regimes corresponding to three separate stages of metastasis (19,20): escape from the primary tumour, invasion of foreign tissue, and distant site colonisation.

Distant site colonisation, the rate-limiting step of metastasis (21), requires a cell to migrate through the unpredictable microenvironment of the primary tumour (22) and into the novel environment of the distant metastatic site (11). Success in both stages is achieved, in part, by the cell's capacity to detect and respond to changes in the environment (23-26). Cells selected for distant site colonisation might therefore be expected to be more reactive to environmental changes, and as such display a greater degree of morphological change in response. We would also expect morphological change to be positively correlated with migration speed in successfully metastasizing cells, because a faster-moving cell will experience a greater degree of environmental variation over a given time period, and therefore change its morphology more rapidly in response.

To test these hypotheses, we have combined an experimental evolution framework with video microscopy and novel statistical analysis that quantifies morphological change in individual cells over time. This approach has identified unique cell behavioural phenotypes that may be advantageous for successful metastasis.

Materials and Methods

Evolved population summary

We used experimental evolution methods (15) on an initial population of MDA-MB-231 breast cancer cells (Figure. 1), subjecting them to three separate selective regimes. The experimental selective regimes were designed to be similar to those experienced whilst traversing the metastatic cascade (11). We also froze two biological replicate *ancestor* populations (Figure. 1) at the start of the experiment to act as a control for comparison with our evolved lines.

We selected *escape* populations (Figure. 1) by tightly packing cells into a high density core of collagen and then allowing them to escape outwards into a low density collagen outer ring (27). After 10-14 days the cells that had escaped into the outer collagen ring were recovered from the matrix, expanded and then seeded back into a new collagen escape assay, completing one round of selection. In total, 7 rounds of selection were applied to each of four biological replicate escape populations. The high density collagen core and the low density outer collagen ring were both three-dimensional (3D) culture environments designed to be similar to those experienced during tumour dissemination.

We selected *invasion* populations (Figure. 1) following a similar protocol to the escape populations whereby repeated consecutive rounds of selection were applied. In contrast to the escape assay, however, cells moved from a 2D to 3D environment, similar to the change in environment experienced during the arrest of a cell at a distant site. The cells were seeded around the outside of a Matrigel island - a synthetic basement membrane matrix widely used in cell culture - and left to invade. After 7 days the cells were collected from the Matrigel, expanded and seeded around the outside of another Matrigel island. This process was repeated 15 times for each of the four biological replicate populations over the course of the 6 month experiment.

We selected *colonisation* populations (Figure. 1) by culturing cells on a piece of decellularized rat lung, which acted as a scaffold for growth similar to that experienced by cells colonizing a distant site (27). The protocol involved cells being seeded onto a decellularized scaffold and left to colonize over a 6 month period. Decellularized tissue is generated by removing all cells from a piece of tissue such that only the extracellular matrix is left. At the end of the experiment cells were released from the scaffold, ensuring that the population represented cells from within the tissue core as well as the edges. Again, this selection was applied to four biological replicate populations.

Finally, all twelve experimentally evolved cell populations were frozen and then thawed alongside the ancestor populations prior to experimental analysis. This step ensured that any selective pressure from the freezing-thawing process was constant across all treatments and replicate populations.

Experimental assays

Escape Assay

Initially, MDA-MB-231 cells (LGC) were encapsulated in a 2mg/ml collagen gel (rat-tail collagen type 1, First Link) and set into a 24-well plate which was used as a mould (750,000 cells per gel, Greiner Bio-One). The collagen gels were compressed for 2 minutes as described in (27), then set into a 1mg/ml low density collagen gel (rat tail collagen type 1, First Link). Once set, cell culture medium (Dulbeco's Modified Eagles Medium (DMEM) supplemented with 10% Fetal Bovine Serum (FBS), and Penicillin 100 µg/ml, Streptomycin 100 U/ml (Gibco, Fisher Scientific)) was added over the top. Medium was replaced every 3-4 days. After 10-14 days, the compressed collagen disc was separated from the low density collagen and collagenase type 1 diluted in phospho-buffered saline solution (Gibco, Fisher Scientific) used to retrieve the cells from the collagen matrix, 200 U/ml for compressed collagen and 100 U/ml for low density collagen. Cells in collagenase/PBS were incubated at 37^{o}C in a stirred water-bath at 45 rpm for 30-60 minutes, then washed in Phospo-buffered saline solution (PBS, Gibco Fisher Scientific).

Cells extracted from the compressed collagen were placed in liquid nitrogen storage and those collected from the low density collagen were seeded into 2mg/ml collagen gel with medium over for population expansion. Once expanded, cells were retrieved from collagen using collagenase in PBS then seeded into 2mg/ml collagen for compression or frozen at -80°C and transferred to liquid nitrogen for storage.

Invasion Assay

MDA-MB-231 cells (LGC) were re-suspended in PBS, and seeded around the outside of a 5mg/ml set Matrigel island in a 6-well plate Matrigel (#35623, Corning), was diluted using DMEM without supplements. Cells were seeded in excess at the island margins, with around 40,000 cells seeded in 200µl per experiment for the initial set-up. Cells were left to settle and adhere to the 2D surface for 60 minutes then cell culture medium added over the top (DMEM supplemented with 10% FBS, and penicillin 100 µg/ml, streptomycin 100 U/ml). Medium was changed every 3-4 days and cells were harvested after 7 days. Cells were retrieved from Matrigel using Cell Recovery solution (#354253, Corning) on ice for 45-60 minutes, washed with ice cold PBS then reseeded into Matrigel at 5mg/ml to expand cell numbers. After 7 days the cells were released from Matrigel using cell recovery solution as described above (typically 400,000 – 500,000 per gel), re-suspended in PBS and seeded in excess around the outside of a new Matrigel island (5mg/ml) for the next round of the 2D/3D invasion assay or cells were frozen at -80°C and transferred to liquid nitrogen for storage.

Colonisation Assay

Rat lung was retrieved from 9 week old Wistar rats (Envigo) and flash frozen. It was then thawed and decellularized using repeated rounds of treatment following an adapted version of the protocol published in (28). Briefly: frozen lung was thawed and cut into small pieces of around 100mg, which were then placed into deionized water (ddH_2O), stirred at 60 rpm for 16 hours at 4°C. Lung tissue was treated with 0.02% trypsin/0.05% EDTA for 60 minutes at 37°C at 60 rpm, 3% Triton-X 100/PBS for 70 minutes, 1M sucrose/PBS for 30 minutes, 4% deoxycholate/ddH_2O for 60 minutes, 0.1% peracetic acid in 4% ethanol for 120 minutes, PBS for 5 minutes, and finally twice in ddH_2O for 15 minutes. The tissue was washed thoroughly between each treatment with ddH_2O. De-cellularization was checked between rounds using epifluorescence microscopy and staining with DAPI H1200 Vectashield (Vectorlabs) to identify whether cell nuclei remained within the matrix structure. Decellularized lung tissue was freeze-dried and stored in an airtight container.

Using decellularized lung as a culture matrix: tissue was soaked in 70% ethanol, washed with PBS and then rehydrated in PBS pH 7.2 (Gibco) in a tissue culture incubator for 5 days, then soaked in cell culture medium (DMEM supplemented with 10% FBS and penicillin/streptomycin as described above) for 48 hours. Cells grown in 2D tissue culture flasks were trypsinized, re-suspended in medium then 750,000 cells added in low volume of medium (100-150 µl) over the decellularized lung tissue in a 6-well plate and left to adhere for 2 hours. Medium was then added over the top so that the decellularized lung rafts floated. Rafts were transferred to new wells when the bottom of the well was confluent with shed and adhered cells. To feed the cells growing in/on the raft, ½ of the medium (2ml of 4ml) was aspirated and replaced every 2-4 days. After 140 and 189 days, rafts were retrieved from medium, washed with PBS and cells harvested by incubating in: collagenase I (170 U/ml, Gibco 17018-029), collagenase IV (170 U/ml, Gibco 17104-019), elastase (0.075 U/ml, Sigma E7885) (based on the protocol described in (29)) incubated at 37°C 45rpm in a stirred water-bath, then washed twice with PBS before seeding in 2D tissue culture plates for expansion. Expanded cells were then frozen at -80°C and transferred to liquid nitrogen for storage.

Time-lapse microscopy

Cells were retrieved from liquid nitrogen, cultured in 2D tissue culture flasks (25cm^2 or 75cm^2 Greiner bio-one), trypsinized and seeded into 6-well plates (Greiner bio-one) at 10-15% cell confluence. Time-lapse movies were made for 12 hour periods with images taken at 2 minute intervals, using a Nikon TiE phase contrast microscope with an environmental chamber (37^{o}C) and moveable platform stage. x10 Plan Apo DIC L Lens was used in conjunction with an intermediate magnification changer set to x1.5 to give x15 magnification. NIS Elements software was used for image capture.

Cell Tracking

All cells that were present in each time-lapse video were tracked using the Usiigaci pipeline (30). The neural network was trained on 300 randomly selected images that were manually annotated using ImageJ (31). The manually annotated images were than randomly split so that 80% were used for training and a further 20% were used for testing, 240 images in the training set and 60 in the test set. The 240 training images were then further split for training and validation 90:10 so that 216 images were used for training and 24 for validation. We trained 3 neural networks using the same 240 images however different images were used for the training and validation stage each time. All hyperparameter settings were the same as Usiigaci protocol except the gradient clip norm was increased to 10. We trained the network on all layers over 300 epochs with the learning rate starting at 0.01 and decreasing by an order of magnitude every 100 epochs.

Once the morphologies had been segmented we tracked them through time using the inbuilt semi-automated Usiigaci tracker. After tracking we manually checked the segmented morphologies and corrected any errors. We checked for cases whereby a cell had divided, been mis-identified or incorrectly segmented. Finally we excluded the 30 minutes prior to and after a cell division to remove the rounded morphologies typical of cell division from our analysis.

Quantifying values

All values were quantified using a custom-built pipeline in Python (32) that can be found on GitHub, https://github.com/george-butler/2d_microscopy, any reference to distance refers to the Euclidean distance. The morphology was quantified using the first 20 Zernike moments. Zernike moments capture the information that is encoded in a shape and translate it into a high dimensional vector, in a similar fashion to spatial location being represented by Cartesian coordinates. When taken to a high enough degree, Zernike moments are capable of representing every shape uniquely and are invariant to rotation, scale and translation (33). We followed the methods of (10) to pre-process the morphologies and make them invariant to scale and translation. We determined that 20 Zernike moments were adequate to quantify the morphology of each cell by plotting the mean squared error against the number of moments (34) and finding where the gradient approached 0.

Statistical analysis

All statistical analysis was performed in R (35) and Figures 3-5 were made using GGPlot (36). All code and corresponding data can be found on GitHub, https://github.com/george-butler/2d_microscopy/tree/master/statistical_analysis . A cell needed to appear in at least 30 frames to be included in our analysis and be present for at least 75% of the track. Some cells were not detected in a given frame or had to be removed due to being incorrectly segmented. Throughout our analysis we used linear mixed models to account for the differences between replicate populations within the four treatments (37). The mean rate of morphological change and the mean speed of migration were calculated through the use of an intercept only linear mixed population with independent intercepts for each treatment. The rate of morphological change model is defined below:

Rate of morphological change = $\alpha + \beta_1$ (speed of migration) + β_2 (distance to nearest neighbour) + β_3 (speed of migration) * (distance to nearest neighbour) + (1|well id)

The model was selected through forward selection whereby parameters were only included if they were significant at the 5% level. The marginal R^2 values were calculated using the method detailed by (38).

Results

Quantifying dispersal in evolved populations

To analyse their dispersal behaviour cells were placed onto 2D tissue-culture plates and their migration was recorded over a 12-hour period , with images taken at two-minute intervals. The 2D plastic environment was intentionally chosen as a neutral testing environment and to ensure that the morphology could be clearly seen without the use of fluorescent tags, a factor that might have applied an additional selective pressure (39). The cells were tracked through the use of a semi-automated pipeline, Usiigaci (30), that combined a convolutional neural network with our own manually annotated images to trace the morphology of each cell at every time point (Figure. 2A).

We extracted three quantitative measures per cell per frame of time-lapse video: morphology, speed and the distance to the closest neighbouring cell. Morphology was quantified using Zernike moments. Zernike moments (33) have been used previously to quantify cancer cell morphology in fixed populations (10) and are a method that captures all of the morphological information available rather than needing to make a prior decision about which morphological features might be important i.e. the length of a cell. The rate of morphological change is then measured as the distance between the vector of moments in frame *t* and *t+1* relative to the time between frames (Figure. 2B). Speed of migration was calculated from the change in spatial location between consecutive frames (Figure. 2C). The distance to the closest neighbour cell was calculated as the shortest distance from the edge of the cell contour to another neighbouring cell contour without crossing the body of the cell (Figure. 2D). Finally the average was calculated for each metric over the entire trajectory of the cell, providing a summary of the dispersal phenotype of each cell.

After extracting these three metrics we sought to evaluate whether the rate of morphological change or the speed of migration was significantly different among the four treatments. We used an analysis of variance (ANOVA) to compare the mean rate of morphological change and the mean speed of migration across all populations; differences in wells were accounted for as a random effect. We found that there was significant variation among population in their mean rate of morphological change ($p = 0.0296$, $N = 813$). We then conducted a post-hoc Bonferroni multiple comparison test to identify which populations were different, controlling for any possible between-replicate variation through the use of a random effect. Escape populations had a significantly higher rate of morphological change compared with the invasion populations, ($p = 0.0152$, $N = 813$; Figure. 3). There was no significant difference in the mean speed of migration among the four treatments.

Speed of migration predicts rate of cell-morphological change in evolved populations

Next we investigated how the morphological behaviour of a cell related to its speed and its social environment. We fitted a linear mixed model across our data whereby the rate of morphological change is dependent on the speed of migration, the distance to the nearest neighbouring cell and the interaction of the two, as detailed in our Methods. We set treatment as a fixed effect and

allowed intercepts and slopes to vary between treatments. The significant parameters were then used to fit a reduced model to the ancestor, escape and invasion populations (Figure. 4).

In the ancestor populations neither the speed of migration nor the distance to neighbouring cells significantly affected the rate of morphological change. We proceeded by fitting an intercept only model to our data (Figure. 4). However, the intercept model explained only a small proportion of the variance, (marginal $R^2 = 0$). This might suggest that the rate of morphological change is highly stochastic, or that it depends on factors not included in our model.

In contrast, in both escape and invasion populations, the rate of morphological change is significantly positively correlated with the speed of migration, ($\beta = 0.680$ and 0.319 respectively: Figure. 4). Furthermore, the escape and invasion models both explain a significant proportion of the variation (marginal $R^2 = 0.347$ and 0.099 respectively). To ensure that our results were not affected by a small cluster of potential outliers we repeated the same analysis after having removed influential data points, defined by a Cook's distance > (4 / N) where N is the sample size (40).

The slope of the relationship is steeper for escape than for invasion populations suggesting that selection for escape may favour cells that can change their morphology rapidly when migrating at a high speed. This might be a result of the collagen escape assay being a 3D to 3D environment compared with the 2D to 3D environment of the Matrigel invasion assay. However, this also could be due to the different number of rounds of selection between the two assays, or difference in the strength of selection within each.

Spatial density affects morphological dynamics

The colonisation populations displayed a complex morphological behaviour dependent on the speed of migration, the distance to the nearest neighbouring cell and the interaction of the two: as the distance between neighbouring cells increases, the relationship between the rate of morphological change and the speed of migration becomes negative (Figure. 5A). When close to a neighbouring cell, the rate of morphological change is positively correlated with the speed of migration: a faster speed of migration results in a higher rate of morphological change. However, when the distance between neighbouring cells is large and a cell is isolated, the rate of morphological change is negatively correlated with the speed of migration: a faster speed of migration has a lower rate of morphological change. We repeated the same analysis after the removal of any influential data points and found that the interaction term was still significant in these colonisation populations (Fig. S1). We also found that the colonisation model explained a significant proportion of the variation in the rate of morphological change (marginal $R^2 = 0.236$).

Next we sought to determine whether the switch in morphological behaviour with distance was gradual or sudden. To investigate this hypothesis, we centred the nearest neighbour data at a distance x and then refitted the same morphological change model. After fitting the model, we evaluated whether the speed of migration was significant in the model. If the speed of migration is not significant then we know that at a distance x there is not a significant difference in the rate of morphological change for cells migrating at different speeds. We can then repeat the same method for different values of x to find a range of distances over which the speed of migration is not significant. The smaller the range the more sudden the switch.

We found that for nearest neighbour distances between 57.9µm and 147.2µm the speed of migration is not significant in our model, as seen by the shaded region in Figure. 5B. Therefore, at distances < 57.9µm or > 147.2µm the speed of migration is significantly related to the rate of morphological change. The small range of distance values suggests that the cells have a high degree of sensitivity to the location of neighbouring cells. Interestingly, the range of distance values coincides with values from the literature whereby cells within a tumour core have been

seen to display a correlated mode of migration at spatial distances < 50μm compared with distances greater than 250μm (41).

Discussion

We have conducted novel phenotypic analysis across experimentally evolved populations of MDA-MB-231 breast cancer cells to investigate their behaviour during dispersal. Combining experimental evolution with computer vision we have generated a multidimensional data set that quantifies single cell dispersal dynamics within each population. In turn we have built a continuous data driven morphological model that has uncovered fundamental dispersal behaviour at a cellular level and is capable of distinguishing cells selected for colonisation.

The flow of migratory cells through the microenvironment creates a landscape that is heterogeneous both spatially and temporally (42). This landscape variability might in turn explain the correlation between the rate of morphological change and the speed of migration for both the escape and invasion populations (Figure 4). The collagen escape and Matrigel invasion assays used to select the escape and invasion populations are porous and complex (43) but yet they are also malleable. The malleability of these two environments means that large structural changes can occur and thus migration routes that were previously accessible may become blocked. Therefore, a cell may need to respond to its environment by changing its morphology to ensure that it can continue to migrate and does not become trapped. Likewise, as the speed of migration increases, an increase in the rate of morphological change might be necessary to ensure that the cells aren't temporarily stuck by any potential obstacles. This would also explain why there is no correlation in the ancestor populations where the environment remains constant and there would therefore be no selective advantage to this behaviour.

Distant-site colonisation requires a cell to switch from a mode of long-range dispersal and focus on re-initiating aggressive proliferation; the subsequent increase in local cell density may reduce available space and thus intensify competition. A similar selective pressure can be seen in our experimental assays. In contrast to the ancestor, escape and invasion populations, where cells are periodically moved to a new expansive environment, the colonisation population remain fixed. As such in addition to the structural changes that occurred in the microenvironment there was a high density of cells migrating locally so cells themselves could block potential migration routes, and therefore might explain the significance of the neighbour location in our model. This hypothesis would also explain the interaction that is observed between neighbouring cells. If a cell is migrating at a high speed and is close to other neighbouring cells, then changing its morphology rapidly might be necessary to avoid other cells that are changing location dynamically. However, when isolated the location of neighbouring cells is no longer of concern and thus a reduction in the rate of morphological change might allow a cell to conserve resources.

The significance of the neighbour sensitivity may also suggest that the ability of a cell to sense contact has been re-acquired within the colonisation population. A loss of contact inhibition is seen as one of the earliest developments in cancer progression as it allows aggressive proliferation to ensue, which in turn gives rise to the formation of a primary tumour (44). However, the high degree of neighbour sensitivity seen in Figure. 5 questions whether contact sensing is in fact lost, or instead down-regulated earlier in the metastatic cascade. If true, this could suggest that cells selected for distant-site colonisation are able to vary their own contact sensing ability dependent on the exogenous environmental stresses they encounter.

In summary, we have shown that evaluating cell morphology as a dynamic process provides novel insight into the behaviour of breast cancer cells, and furthers our understanding of the phenotypic route to metastasis. A pivotal next step will be evaluating morphological dynamics within a native 3D environment (45) and in the vicinity of stromal cells such as a fibroblasts which

are known to have a critical role in metastasis (46). The presence of stromal cells might also change the relationship seen within our escape and invasion populations, as cells would then be able to interact via matrix metalloproteinases. Thus, rather than needing to change their morphology quickly to prevent being trapped, they could exploit the matrix metalloproteinases to cut them free, as seen previously during metastatic dispersal (47). It would also be of value to subject multiple starting cell lines to a similar selective regimes, in case the MDA-MB-231 line used here behaves atypically. However, we believe that this work highlights the power of phenotypic analysis in discovering the complex emergent behaviours that would not have been apparent from genetic data .

Acknowledgments

We thank Chris Venditti and Mark Pagel for critical reading of the manuscript.

Author Contributions

G.B, S.K, L.J and P.D conceived and designed the study. S.K evolved the populations and collected the time-lapse data. G.B developed the methodology and performed the formal analysis. L.J and P.D supervised the work. G.B wrote the manuscript. All authors gave final approval for publication and agree to be held accountable for the work performed therein.

References

1. Amend SR, Roy S, Brown JS, Pienta KJ. Ecological paradigms to understand the dynamics of metastasis. Cancer Lett. 2016:237–42.
2. Tissot T, Massol F, Ujvari B, Alix-Panabieres C, Loeuille N, Thomas F. Metastasis and the evolution of dispersal. Proc R Soc B Biol Sci . 2019:20192186.
3. Paul CD, Hung W-C, Wirtz D, Konstantopoulos K. Engineered Models of Confined Cell Migration. Annu Rev Biomed Eng. 2016:159–80. 1
4. Wells A, Grahovac J, Wheeler S, Ma B, Lauffenburger D. Targeting tumor cell motility as a strategy against invasion and metastasis. Trends Pharmacol Sci. 2013283–9
5. Odenwald MA, Prosperi JR, Goss KH. APC/β-catenin-rich complexes at membrane protrusions regulate mammary tumor cell migration and mesenchymal morphology. BMC Cancer 2013:12.
6. Cowden Dahl KD, Dahl R, Kruichak JN, Hudson LG. The epidermal growth factor receptor responsive miR-125a represses mesenchymal morphology in ovarian cancer cells. Neoplasia 2009:1208–15.
7. Prasad A, Alizadeh E. Cell Form and Function: Interpreting and Controlling the Shape of Adherent Cells. Trends Biotechnol. 2018
8. Wu P-H, Gilkes DM, Phillip JM, Narkar A, Cheng TW-T, Marchand J, et al. Single-cell morphology encodes metastatic potential. Sci Adv. 2020
9. Lyons SM, Alizadeh E, Mannheimer J, Schuamberg K, Castle J, Schroder B, et al. Changes in cell shape are correlated with metastatic potential in murine and human osteosarcomas. Biol Open. 2016:289 – 299
10. Alizadeh E, Lyons SM, Castle JM, Prasad A. Measuring systematic changes in invasive cancer cell shape using Zernike moments. Integr Biol. 2016:1183–93.
11. Valastyan S, Weinberg RA. Tumor metastasis: molecular insights and evolving paradigms. Cell . 2011:275–92.
12. Shieh AC. Biomechanical Forces Shape the Tumor Microenvironment. Ann Biomed Eng. 2011:1379–89.

13. Provenzano PP, Inman DR, Eliceiri KW, Knittel JG, Yan L, Rueden CT, et al. Collagen density promotes mammary tumor initiation and progression. BMC Med. 2008.
14. Wershof E, Park D, Jenkins RP, Barry DJ, Sahai E, Bates PA. Matrix feedback enables diverse higher-order patterning of the extracellular matrix. PLOS Comput Biol. 2019.
15. Kawecki TJ, Lenski RE, Ebert D, Hollis B, Olivieri I, Whitlock MC. Experimental evolution. Trends Ecol Evol. 2012:547–60
16. Sprouffske K, Merlo LMF, Gerrish PJ, Maley CC, Sniegowski PD. Cancer in light of experimental evolution. Vol. 22, Current Biology. 2012.
17. Taylor TB, Johnson LJ, Jackson RW, Brockhurst MA, Dash PR. First steps in experimental cancer evolution. Evol Appl. 2013):535–48.
18. Taylor TB, Wass A V, Johnson LJ, Dash P. Resource competition promotes tumour expansion in experimentally evolved cancer. BMC Evol Biol. 2017:268.
19. Chaffer CL, Weinberg RA. A Perspective on Cancer Cell Metastasis. Science. 2011:1559 – 1564.
20. Pantel K, Brakenhoff RH. Dissecting the metastatic cascade. Nat Rev Cancer. 2004:448–56.
21. Massagué J, Obenauf AC. Metastatic colonization by circulating tumour cells. Nature. 2016:298–306. 0
22. Clark AG, Vignjevic DM. Modes of cancer cell invasion and the role of the microenvironment. Curr Opin Cell Biol. 2015:13–22.
23. Peinado H, Lavotshkin S, Lyden D. The secreted factors responsible for pre-metastatic niche formation: Old sayings and new thoughts. Semin Cancer Biol. 2011:139–46.
24. Sceneay J, Smyth MJ, Möller A. The pre-metastatic niche: finding common ground. Cancer Metastasis Rev. 2013:449–64.
25. Psaila B, Kaplan RN, Port ER, Lyden D. Priming the “soil” for breast cancer metastasis: the pre-metastatic niche. Breast Dis. 2015:26:65–74.
26. Costa-Silva B, Aiello NM, Ocean AJ, Singh S, Zhang H, Thakur BK, et al. Pancreatic cancer exosomes initiate pre-metastatic niche formation in the liver. Nat Cell Biol. 2015:816–26.
27. Keeton SJ, Delalande JM, Cranfield M, Burns A, Dash PR. Compressed collagen and decellularized tissue – novel components in a pipeline approach for the study of cancer metastasis. BMC Cancer. 2018:622.
28. Medberry CJ, Crapo PM, Siu BF, Carruthers CA, Wolf MT, Nagarkar SP, et al. Hydrogels derived from central nervous system extracellular matrix. Biomaterials. 2013:1033–40.
29. Quatromoni JG, Singhal S, Bhojnagarwala P, Hancock WW, Albelda SM, Eruslanov E. An optimized disaggregation method for human lung tumors that preserves the phenotype and function of the immune cells. J Leukoc Biol. 2015:201–9.
30. Tsai HF, Gajda J, Sloan TFW, Rares A, Shen AQ. Usiigaci: Instance-aware cell tracking in stain-free phase contrast microscopy enabled by machine learning. SoftwareX. 2019:230–7.
31. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012:676–82.
32. Van Rossum G, Drake FL. Python 3 Reference Manual. Scotts Valley, CA: CreateSpace; 2009.
33. Zernike F. Phase contrast, a new method for the microscopic observation of transparent objects part II. Physica. 1942):974–86.
34. Liao SX. Image Analysis by Moments. The University of Manitoba; 1994.
35. R Core Team. R: A Language and Environment for Statistical Computing. 2017
36. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York2009.
37. Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH, et al. Generalized linear mixed models: a practical guide for ecology and evolution. Trends Ecol Evol. 2009:127–35.
38. Nakagawa S, Schielzeth H. A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods Ecol Evol. 2013:133–42.
39. Liu H-S, Jan M-S, Chou C-K, Chen P-H, Ke N-J. Is Green Fluorescent Protein Toxic to the Living Cells? Biochem Biophys Res Commun. 1999:712–7.

40. Bollen KA, Jackman RW. Regression Diagnostics: An Expository Treatment of Outliers and Influential Cases. Sociol Methods Res. 1985:510–42.
41. Staneva R, El Marjou F, Barbazan J, Krndija D, Richon S, Clark AG, et al. Cancer cells in the tumor core exhibit spatially coordinated migration patterns. J Cell Sci. 2019. t
42. Yuan Y. Spatial Heterogeneity in the Tumor Microenvironment. Cold Spring Harb Perspect Med. 2016.
43. Anguiano M, Castilla C, Maška M, Ederra C, Peláez R, Morales X, et al. Characterization of three-dimensional cancer cell migration in mixed collagen-Matrigel scaffolds using microfluidics and image analysis. PLoS One. 2017.7
44. Pavel M, Renna M, Park SJ, Menzies FM, Ricketts T, Füllgrabe J, et al. Contact inhibition controls cell survival and proliferation via YAP/TAZ-autophagy axis. Nat Commun. 2018:2961
45. Petrie RJ, Yamada KM. At the leading edge of three-dimensional cell migration. J Cell Sci. 2012:5917–26.
46. Malanchi I, Santamaria-Martínez A, Susanto E, Peng H, Lehr H-A, Delaloye J-F, et al. Interactions between cancer stem cells and their niche govern metastatic colonization. Nature. 2012:85–9.
47. Page-McCaw A, Ewald AJ, Werb Z. Matrix metalloproteinases and the regulation of tissue remodelling. Vol. 8, Nature Reviews Molecular Cell Biology. 2007.

Figures

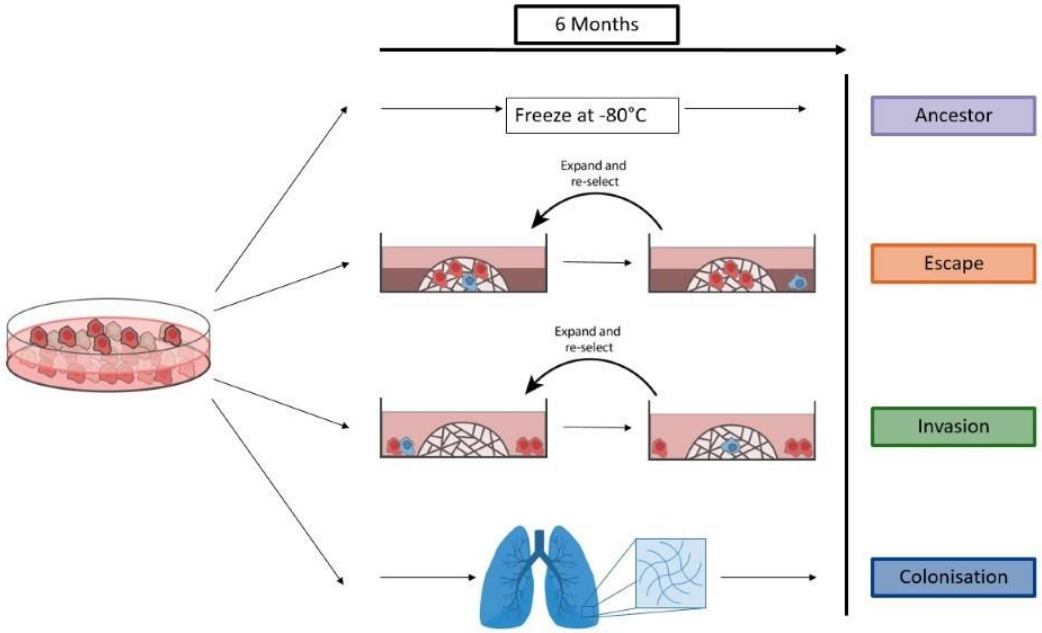


Figure 1. Experimental evolution of cancer cell populations. Ancestor populations were kept frozen throughout. Escape populations were placed in a high density collagen matrix the surrounded by a low density outer collagen ring: after 10-14 days cells that had escaped into the outer ring (shown in blue) were released, expanded and reseeded back into a new high density collagen core; this process was repeated 7 times over the course of 6 months. Invasion populations were seeded around a Matrigel island; after 7 days cells that had invaded the Matrigel (shown in blue) were released, expanded and reseeded around a new Matrigel island this was repeated 15 times over the course of 6 months. Colonisation populations were seeded

onto a piece of decellularized rat lung which acted as a novel scaffold for colonisation and left to establish for 6 months. Four replicate lines were maintained for each treatment.

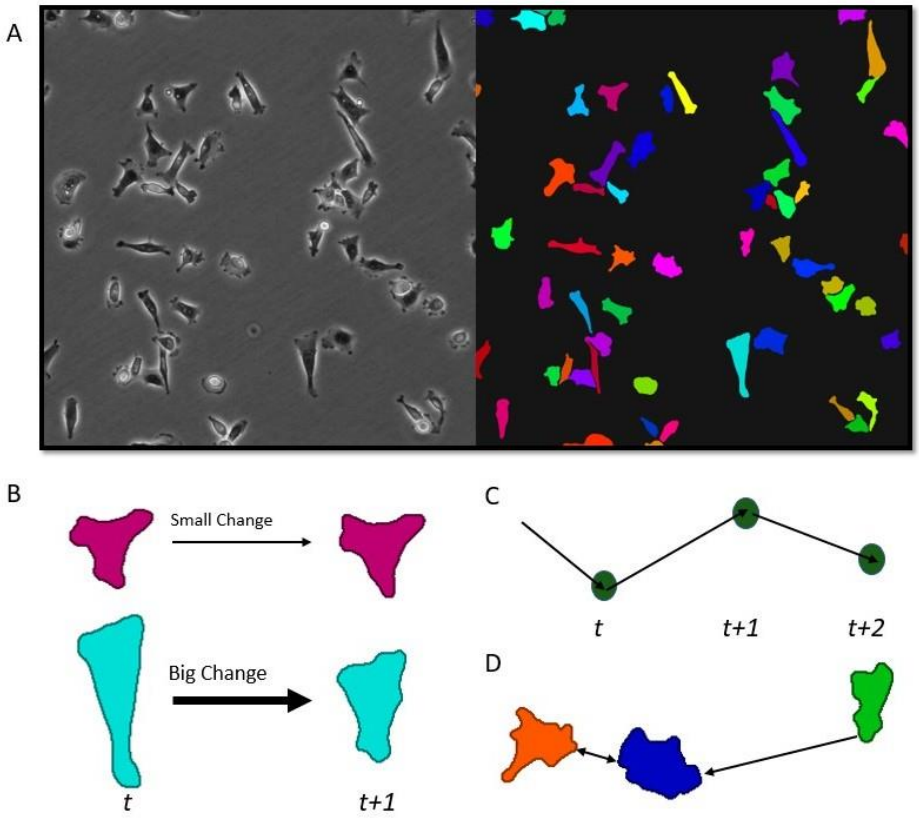
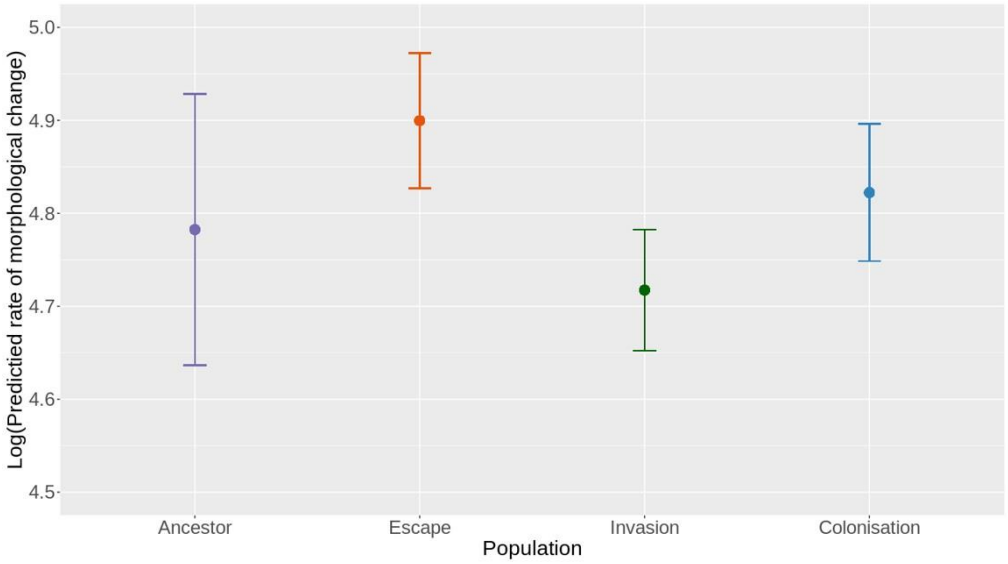


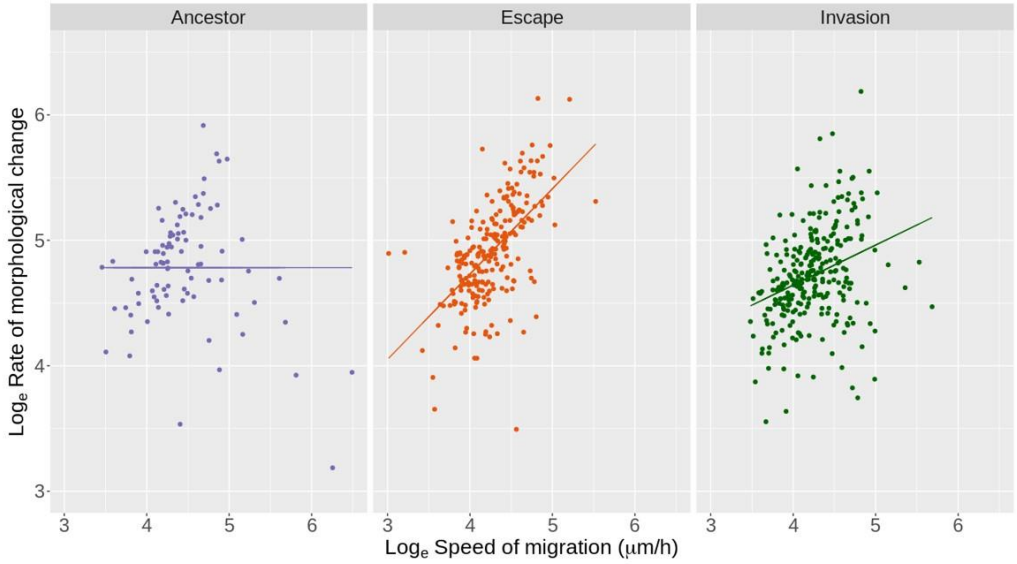
Figure 2. Quantifying dispersal from time-lapse videos. (**A**) Cells were tracked over a 12 hour period with images taken at two minute intervals using phase contrast time-lapse microscopy to generate movies from which morphology could be segmented through the use of a convolutional neural network. (**B**) The rate of morphological change was recorded as the distance between Zernike moments in consecutive frames. (**C**) The speed of migration is calculated as the distance between the spatial location of cells in consecutive frames. (**D**) The distance between neighbouring cells is quantified as the shortest distance between the contour of one cell and the contour of another. The direction of the arrow points from a given cell to the point on the contour of the closest neighbouring cell.



525

526 **Figure 3. Comparing the mean rate of morphological change among the four treatments.** A
527 plot of the natural log-transformed rate of morphological change for each of the four treatments.
528 The centre dot signifies the mean rate of morphological change with errors bars signifying 95%
529 confidence intervals. The escape populations had a significantly faster rate of morphological
530 change compared with the invasion populations, p = 0.0152 (N = 813). The mean, standard error
531 and number of observations for each population can be found in Table S1.

532



533

Figure 4. The rate of morphological change against the speed of migration. The natural log-transformed rate of morphological change plotted against the natural log-transformed speed of migration. The straight lines represent the reduced model for each treatment using only parameters that are significant at the 5% level. The ancestor populations have an intercept-only model fitted (N = 88). The speed of migration is the only significant variable in the escape (N = 230, $p = 1.765 \times 10^{-3}$) and invasion (N = 283, p = 0.018) populations. For both escape and invasion populations the rate of morphological change is positively correlated with the speed of migration, the faster the speed of migration the higher the rate of morphological change.

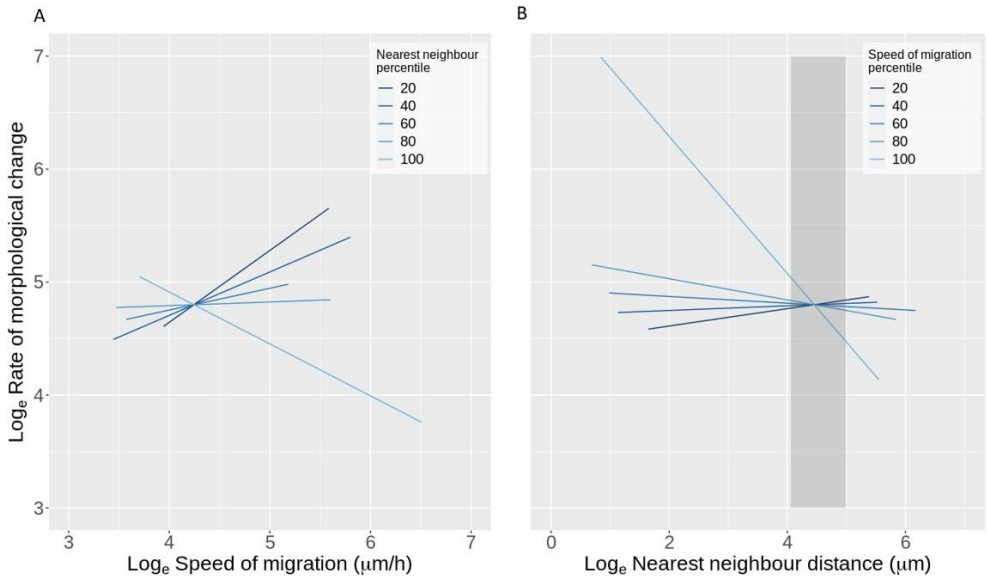


Figure 5. A dynamic switch in the morphological behaviour within cells selected for colonisation. Data points have been removed to highlight the behaviour of the model, the same model with data points can be seen in Fig. S2. The speed of migration ($p = 5.418 \times 10^{-14}$), the distance to the nearest neighbouring cell ($p = 2.207 \times 10^{-10}$) and the interaction of the two ($p = 2.219 \times 10^{-11}$) was significant in the colonisation population (N = 212). (**A**) The predicted natural log-transformed rate of morphological change against the natural log-transformed speed of migration. The shaded lines indicate the natural log transformed nearest neighbour percentile. The lighter the line, the further away from a neighbouring cell with distance values ranging from 2μm - 477μm. (**B**) The predicted natural log-transformed rate of morphological change against the natural log-transformed nearest neighbour distance. The shaded lines indicate the speed of migration percentile. The lighter the line the faster the speed of migration. The shaded region indicates the range of distances over which there is no significant relationship in the rate of morphological change and the speed of migration when the data is centred at these distances, between 57.9μm and 147.2μm.