Interface Focus

rsfs.royalsocietypublishing.org

Research

Cite this article: Gallaher J, Anderson ARA. 2013 Evolution of intratumoral phenotypic heterogeneity: the role of trait inheritance. Interface Focus 3: 20130016.
http://dx.doi.org/10.1098/rsfs.2013.0016

One contribution of 11 to a Theme Issue 'Integrated cancer biology models'.

Subject Areas:
biomathematics, computational biology, biocomplexity

Keywords:
heterogeneity, evolution, phenotypic inheritance

Author for correspondence:
Jill Gallaher
e-mail: jill.gallaher@moffitt.org

Evolution of intratumoral phenotypic heterogeneity: the role of trait inheritance

Jill Gallaher and Alexander R. A. Anderson

Department of Mathematical Oncology, H. Lee Moffitt Cancer Center, Tampa, FL 33612, USA

A tumour is a heterogeneous population of cells that competes for limited resources. In the clinic, we typically probe the tumour by biopsy, and then characterize it by the dominant genetic clone. But genotypes are only the first link in the chain of hierarchical events that leads to a specific cell phenotype. The relationship between genotype and phenotype is not simple, and the so-called genotype to phenotype map is poorly understood. Many genotypes can produce the same phenotype, so genetic heterogeneity may not translate directly to phenotypic heterogeneity. We therefore choose to focus on the functional endpoint, the phenotype as defined by a collection of cellular traits (e.g. proliferative and migratory ability). Here, we will examine how phenotypic heterogeneity evolves in space and time and how the way in which phenotypes are inherited will drive this evolution. A tumour can be thought of as an ecosystem, which critically means that we cannot just consider it as a collection of mutated cells but more as a complex system of many interacting cellular and microenvironmental elements. At its simplest, a growing tumour with increased proliferation capacity must compete for space as a limited resource. Hypercellularity leads to a contact-inhibited core with a competitive proliferating rim. Evolution and selection occurs, and an individual cell's capacity to survive and propagate is determined by its combination of traits and interaction with the environment. With heterogeneity in phenotypes, the clone that will dominate is not always obvious as there are both local interactions and global pressures. Several combinations of phenotypes can coexist, changing the fitness of the whole. To understand some aspects of heterogeneity in a growing tumour, we build an off-lattice agent-based model consisting of individual cells with assigned trait values for proliferation and migration rates. We represent heterogeneity in these traits with frequency distributions and combinations of traits with density maps. How the distributions change over time is dependent on how traits are passed on to progeny cells, which is our main enquiry. We bypass the translation of genetics to behaviour by focusing on the functional end result of inheritance of the phenotype combined with the environmental influence of limited space.

1. Introduction

Tumours are phenotypically and genotypically heterogeneous. Understanding even the most superficial definitions of a tumour's complexity may seem an overwhelming and daunting task, yet characterizing this heterogeneity has begun to gain popularity in recent years [1–5]. The pursuit of embracing the complexity of the myriad of possible interactions between individual parts is gaining ground in computational and mathematical oncology, and we are learning how to represent the tumour as the heterogeneous system that it is.

With heterogeneity comes a possibility for competition and selection. A growing tumour is limited for space and resources, so not all individual cells will continue propagating over time. The cellular population may adapt for enhanced overall fitness, but selection can also be used to our advantage when it comes to treatment if we can better understand how the interacting parts combine to form the emergent whole.

The actual function of each cell is not just determined by genotype [6]. The same genotype may yield different phenotypes, and the same phenotype may result from multiple genotypes [3,6]. Add to that complexity the epigenetic

Royal Society Publishing
Informing the science of the future

modulations that affect the cell's response [7], the microenvironment [8,9], differential protein expression and biochemical noise [10–13] and the problem of mapping genotype to phenotype appears to be impossible [14]. To reduce this inherent complexity and because the cell's function is what selection will act on, we focus on the phenotype.

Here, we will present a method that will characterize the key elements of phenotypic heterogeneity in a growing tumour and simulate situations in which the cells that make up this tumour will compete with each other. We investigate, using the traits of proliferation and migration, (i) how trait inheritance matters, (ii) how spatial competition affects trait selection, and (iii) how constraints on trait combinations affect population growth.

With heterogeneity in the phenotypes of a collection of cells, it matters how a cell passes on these traits. The role of phenotypic inheritance in tumour progression under environmental stresses has been considered before, but remains an open question [7,9,15–20]. In the pioneering work by Anderson and colleagues, the evolution of phenotypic heterogeneity of five different traits was investigated. They compared a linear progression of mutation and a more random acquisition of traits (representing phenotypic plasticity) in a growing tumour in a range of different microenvironments [8,21]. We continue this line of thought by exploring trait inheritance in a much simpler model with only two traits, namely proliferation and migration, under the stress of limited space.

Space is the only limited resource we consider here. We assume that with hypercellularity the cells are subject to contact inhibition, which induces quiescence. We look at two different configurations that lead to different degrees of competition: a dispersion of cells and a cluster of cells. The dispersion of cells begin thinly distributed over a large area with plenty of space to form their own colonies. This may be analogous to a cell suspension *in vitro* or an *in vivo* situation in which competition is less pronounced (e.g. circulating tumour cells). The cluster of cells start tightly packed together so that as it grows only the cells on the rim of the mass will be in the proliferating state. Contact inhibition will play a bigger role in providing a selective pressure for evolution of the population in the latter configuration. In a model by Lee *et al.* [22], the growth of populations in similar spatial configurations was investigated and the importance of contact inhibition was highlighted, but this model did not consider heterogeneity or variations in the mode of inheritance, which will be our primary focus here.

With limited energy and resources, the combinations of traits that allow cell viability will be bounded. Much has been studied on the idea of the proliferation–migration dichotomy (i.e. go-or-grow) [23–25], where there is a trade-off between proliferating and migrating. This is a simplistic view of a complicated set of possibilities, but there are some combinations of traits that a cell will never achieve. Many traits are coupled in such a way that an effect on one is sure to have an impact on another. As more traits are considered, the complexity increases dramatically.

This study also provokes speculation for how treatment might be affected by the heterogeneity within a tumour and the inheritance mode of the cells. To keep things as simple as possible we will not include cell death in this model, but the implications remain that wiping out a particular cell type will leave room and resources for other non-targeted cell types to take over [26]. The selection of phenotypes in a growing population with environmental influences is an important concept that should not be ignored. Most standard of care anti-cancer therapies act on phenotypic traits (e.g. anti-mitotics, anti-proliferatives, radiation, chemotherapies). So we are interested in seeing how the underlying lineage of diversity comes about in the first place.

This simple model provides a brief look at how inheritance of proliferation rates and migration rates affects the overall heterogeneity and fitness of a growing population of cells. We build an off-lattice agent-based model to inspect how the many interacting cells compete for space as they pass on their traits in various ways. In the first simulations, we vary each trait individually to get an idea of how they alone influence the collective behaviour of the population. The subsequent simulations allow combinations of traits with different constraints imposed on the phenotype space. After discussing how the heterogeneity of a population is affected by inheritance, we analyse how the different schemes affect the fitness of the population.

2. The model

To capture the competition between individuals with heterogeneity in proliferation and migration rates, we implement an agent-based model. Each cell is given a set of traits initially, but we also need to keep a memory of the previous traits that are important for some inheritance modes. Each cell at any time is defined by a set of parameters given by

$$[p, m] = [(\rho, \rho_0), (\nu, \nu_0, p_T, \theta)]. \tag{2.1}$$

The phenotype of the cell is represented by a combination of trait states; the proliferation state p and the migration state m. The proliferation parameter indicates how quickly a cell moves through the cell cycle in terms of an intermitotic time (IMT), and the migration parameters indicate how fast a cell is moving (speed), how long it moves in the same direction (persistence time) and in which direction it is moving. For a given cell, we define the proliferation parameters, ρ and ρ_0, as the current and previous proliferation rates, respectively. The current and previous migration speeds are ν and ν_0, respectively, p_T is the persistence time and θ is the angular direction in which the cell is moving.

For the simulation, each cell is given a random starting age in the cell cycle, and for each time step it will follow the decision tree presented in figure 1*a*. The simulation starts at the green shape marked '$t = t + 1$', which signifies the start of a new time step. At each time frame, the cell either divides (yellow loop) if it is ready, goes quiescent if it has too many neighbours (red loop) or migrates (blue loop). If it has enough space, the cell will spend most of its time in the migration loop while ageing and only go through the act of division during a single time step. In the quiescent state, the cell discontinues ageing and moving.

All simulations follow the same rules for proliferation and migration, but vary in the inheritance mode, the spatial layout and the combinations of traits permitted. The cell moves by an angle of direction and a speed instead of residing on an orthogonal lattice; the angular lattice is resolved to $\Delta\theta = 1$ degree, and the updating occurs with a time step of 1 min. All cells are 20 μm in diameter.

2.1. Proliferation and quiescence

The cells are each given an IMT, i.e. the time that it takes to go through division. At each time point each cell checks for close

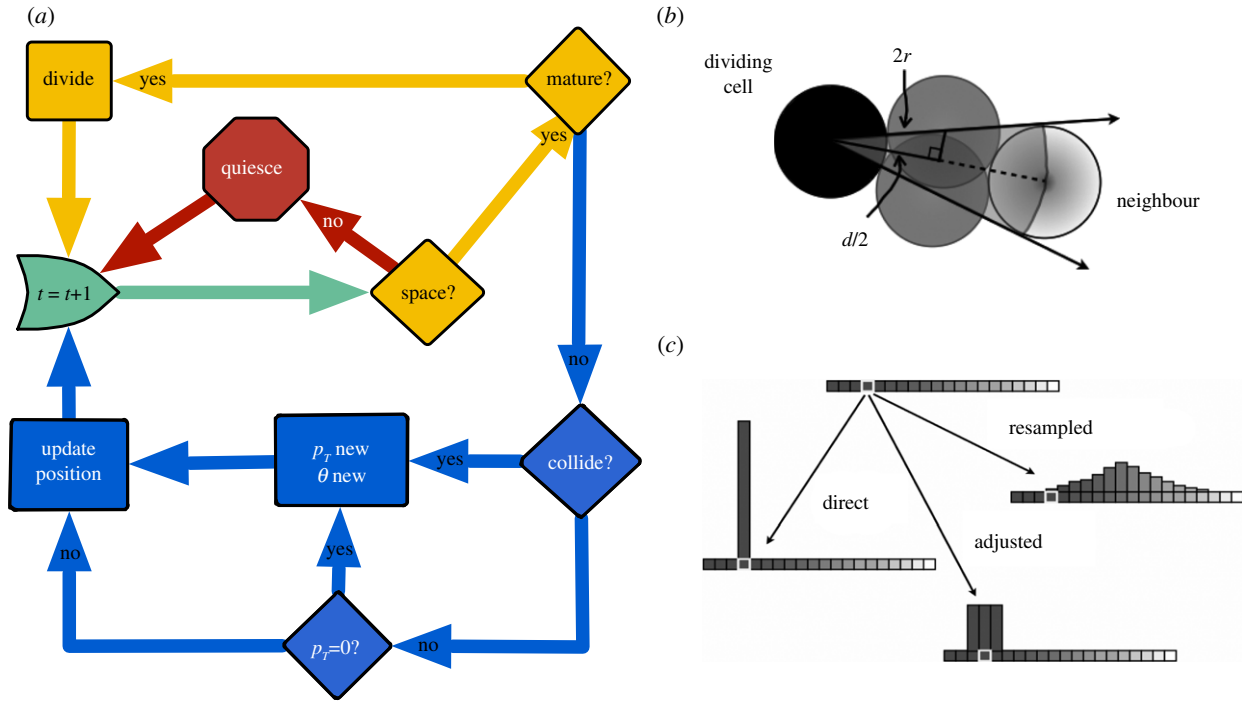

Figure 1. Model details: (*a*) Flow diagram describing decision-making of each cell at each time frame. The time frame starts at the green arrow where it either goes through division (yellow loop), goes into quiescence (red loop) or migrates (blue loop). (*b*) The geometry of angle exclusion that leads to quiescence from contact inhibition from a dividing cell (black) and its neighbour (light grey). The mid-grey 'phantom' cells show the closest possible position of a new cell on each side of the overlap so that there is a block of excluded angles. (*c*) The different modes of inheritance (direct, adjusted and resampled) are shown with the probability of choosing the trait value of the daughter cells given the parent's trait value (white box).

neighbours by searching its surrounding space for other cells positioned at a distance (d) less than four cell radii (r) away (a position close enough for possible overlap with a newly formed daughter cell). For each of these close neighbours, a block of angles $\Delta\theta = \arccos(d/4r)$ to each side of the bisector of the cell is excluded from the bank of allowable angles (figure 1*b*), removing the possibility of overlap upon division. If at any point a cell runs out of available angles, it will go into quiescence, which means here that it stops migrating and stops progressing through the cell cycle. If the quiescent cell has room at a subsequent time to proliferate as neighbours move out of the way, it will pick up where it left off in the cell cycle with its previous migration parameters. The domain boundaries are handled just like neighbouring cells, by simply taking up space and excluding available angles.

When it is time to divide, if there is space, one daughter cell will take up the space previously occupied by the parent, and the second will divide into a neighbouring space. The new angle is chosen randomly from all available angles. Given this new angle θ, the newly positioned cell will lie a cell diameter away in that direction. During the division time step, the cell will not move. Until then, if not quiescent, the cell migrates.

2.2. Migration and cell collisions

Because migration plays such a large role here, we set up an off-lattice model to capture the more subtle nuances in movement. We assume that the cells follow a persistent random walk [17,22,27,28]. In this case, the cell has a 'preferred' or most probable amount of time it will spend continuing in the same direction. We assume that each cell has a predetermined persistence time, p_T, that is randomly drawn from a distribution (normally distributed around 80 min with a standard deviation of 10 min). If the cell has not divided or collided with another cell or the boundary before the persistence time is reached, it will assume a new direction at random and obtain a new persistence time from the distribution.

At each time frame, we loop through all cells currently in the migratory state and identify whether any overlap has occurred with another cell. The time step is sufficiently small that the fastest moving cells will detect an overlap just under 0.5 µm. If two cells are found to overlap, then a collision response begins at the time of first contact within the time frame, then the cells move for the remaining time in the frame as usual with the new trajectory given by the angle of collision. Collisions of one cell with another follow a regular elastic collision response, i.e. the new angle is reflected along the normal. The cells collide with the domain boundaries in the same way. After a collision, a new persistence time is obtained from the persistence distribution for each cell. If the cell divides, each daughter cell obtains a new persistence from the distribution, but they move apart in opposite directions along their normal instead. In the absence of a collision or division, the cell simply continues in the same direction at its given speed.

Variations in the distribution of persistence times and turning angles do result in differences in population-scale dynamics, but in order to keep this manuscript focused they will not be investigated here. We assume that these parameters are less determined by the cell itself and more by the microenvironment. A very directional type of environment (e.g. muscle fibres or blood vessels) might cause a cell to move with long persistences and small changes in turning angles, whereas a more tortuous environment (e.g. grey matter in the brain) might subject a cell to smaller persistences and a more random distribution of turning angles. We, therefore, do not consider these to be appropriate as heritable traits, such as the proliferation rate and the migration speed.

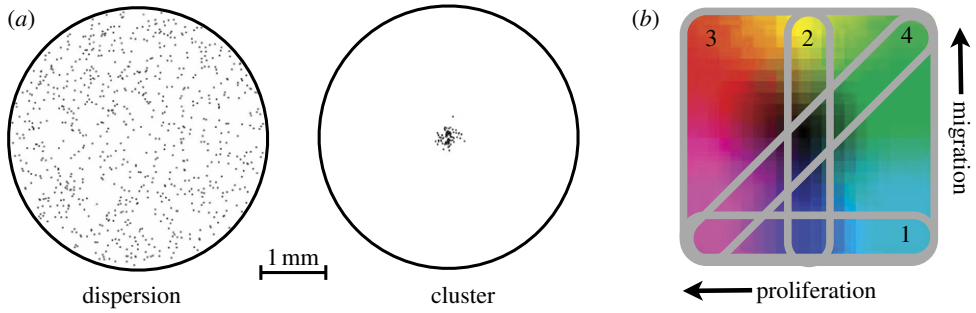

Figure 2. Model details: (*a*) The two initial spatial configurations: dispersion and cluster. (*b*) The different constraints on the phenotype space labelled in order of presentation (1, proliferation only; 2, migration only; 3, no constraints; 4, go-or-grow).

2.3. Inheritance schemes

In this study, we primarily investigate the role of trait inheritance. Three schemes that represent different modes of adaptability will be presented. We refer to the inheritance schemes as: direct, adjusted and resampled. Figure 1*c* relates these various modes of inheritance as a probability that a trait will assume a new value either the same as, close to or completely different from the parent's current value.

With direct inheritance, daughters inherit exactly the same trait value as the parent. We, therefore, define the probability P of the daughter cell inheriting the parent's trait τ_p as simply

$$P(\tau_d) = 1 \quad \text{if} \quad \tau_d = \tau_p. \tag{2.2}$$

With adjusted inheritance, the trait value τ_d of the daughter cells may drift slightly from the parent's trait in either direction by a small value ϵ, where ϵ is the size of one bin in the discretized distribution of allowed values. So, the probability that the daughter cell inherits a trait is

$$P(\tau_d) = \begin{cases} \frac{1}{3} & \text{if } \tau_d = \tau_p - \varepsilon, \\ \frac{1}{3} & \text{if } \tau_d = \tau_p, \\ \frac{1}{3} & \text{if } \tau_d = \tau_p + \varepsilon. \end{cases} \tag{2.3}$$

The values are confined to a range, so surpassing an upper or lower bound for the adjusted inheritance leads to the retention of the previous trait at that bound.

With resampled inheritance, the new trait values for the daughters will have no memory of the parent's trait, but instead will randomly be chosen from a weighted normal distribution:

$$P(\tau_d) \propto e^{(\tau_d - \mu)^2/2\sigma^2}, \tag{2.4}$$

where μ is the mean and σ is the standard deviation. If a chosen value falls outside the allowed range, we simply resample from the weighted distribution. In all of the following simulations, the resampling distribution is the same as the initial distribution of trait values. Using these modes of inheritance, we will investigate their effects on the heterogeneity of a growing cell population over space and time.

3. Results

The three previously introduced inheritance schemes are examined using two different spatial configurations. We refer to these as a cellular dispersion and a cellular cluster. With the cellular dispersion, 800 cells are placed randomly within a circular area 4 mm in diameter, which could represent either an *in vitro* configuration or a disseminated cancer with many seeds spatially spread out. The simulation runs until reaching a population of 25 000 cells, which is sufficient to still discern the colonies just before confluence (approx. 60% full). With the cellular cluster, 80 cells are placed randomly within a much smaller circular area (0.25 mm in diameter) representing a small heterogeneous tumour mass shortly after initiation. From this arrangement, the population grows to 8000 cells before the simulation terminates, which is around the size where vascular recruitment has an effect. The proximity of cells to viable vasculature, which provides nutrients and oxygen, affects both proliferation and migration. To keep the model as simple as possible, we cut off the simulation around the time that the avascular phase ends. These two spatial configurations allow different degrees of competition (figure 2*a*).

Within each section, we examine different constraints on allowable trait combinations. We first examine each trait separately and then in combination with different constraints on allowable values. The heterogeneity of the cellular population will be investigated for variation in proliferation only, variation in migration only, variation in both traits (all combinations allowed) and the go-or-grow constraint (proliferation rate is inversely mapped to migration rate). Figure 2*b* displays these constraints in a two-dimensional proliferation–migration phenotype space that we will use both to show the density of combinations and as a colouring scheme representing phenotypes in the spatial distributions. We conclude with an analysis on how the resulting heterogeneity from these situations contributes to the fitness of the populations as a whole.

3.1. Proliferation: different schemes drive different compositions

Cells are given initial IMTs randomly sampled from a normal distribution with a mean of 18 h and a standard deviation of 4 h. Trait values are confined within a range of 8–28 h, and the width of each bin is 1 h. This is similar to distributions found in real cancer cell lines [17]. In figure 3, we present the results for non-motile cells with diversity only in the proliferation rates.

First and foremost, there is a noticeable difference between the final distributions of traits for the dispersion and the cluster. The dispersion, being spread out spatially, does not provoke much selection, but with the extra competition in the cluster the selection for faster proliferators is much stronger. There are differences in the final distributions of phenotypes across different inheritance schemes. We find that there is selection for faster IMTs if the inheritance is direct or adjusted, whereas no change is observed from the

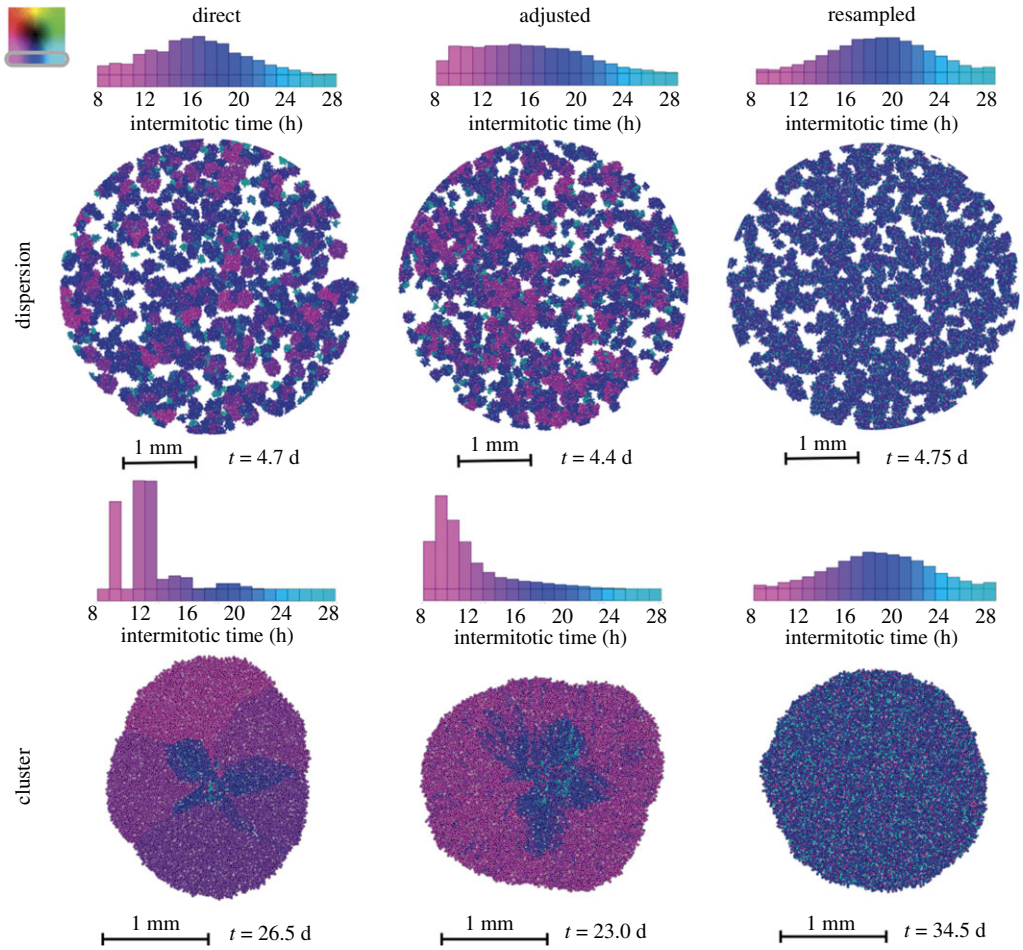

Figure 3. Heterogeneity of IMTs in dispersions upon reaching 25 000 cells (top) and clusters upon reaching 8000 cells (bottom) with different inheriting modes (columns for direct, adjusted and resampled inheritance). The spatial distributions are shown with the time taken to reach the final population recorded below. The histograms above the images show the distribution of IMTs at this final population. The gradient from magenta to blue to cyan represents cells with IMTs going from short to long.

original normal distribution, as expected, with the resampling scheme. We also see that, with direct inheritance, the histogram has more disjointed peaks. With this kind of inheritance, if a clone gets stuck in the bulk and goes quiescent, it cannot further propagate its trait. This is a product of both chance and ability.

The shape and symmetry of the cellular expansion also depends on the inheritance mode. For the dispersions, we find that the colonies with short IMTs are larger in the direct and adjusted inheritance modes, while the slower proliferating colonies are smaller. Inheritance resampling, on the other hand, results in colonies of roughly the same size owing to an equal diversity in each. Just as we refer to colonies in the dispersion, we will use the term 'families' in the cluster to signify a group of cells that are all progeny of an initial seed.

The cellular clusters evolve in space differently because of tougher competition. With direct inheritance, once a family establishes a piece of the proliferating front, it will move outwards as a group in radial bands. But at the interface of these patches of families, competition eventually drives the victory of only one. We can see this progression within the structure where some of the bands consisting of slower IMTs taper off next to a faster growing family. In turn this makes the shape lobular at these interfaces. With adjusted inheritance, initial heterogeneity in space may cause a skewed shape to persist for a while, but the cells eventually drift towards being faster at proliferating. The shape of the cluster might reflect the initial heterogeneity but will eventually round out while wobbling slightly around the lower bound of IMTs. With resampled inheritance, the cluster retains a rounded geometry. It can be thought of as being equally heterogeneous locally and globally with a relatively even radial expansion.

3.2. Migration: movement benefits all

Next, we examine the situation when cells have variation in migration speeds but have the same proliferation rates. We set all of the IMTs to be 18 h long, and initialize the cells with speeds drawn randomly from a normal distribution centred at 12 $\mu m\,h^{-1}$ and with a standard deviation of 5 $\mu m\,h^{-1}$. We bound the available migration speeds within a window from 0 to 25 $\mu m\,h^{-1}$, and the width of each bin is 1 $\mu m\,h^{-1}$. This distribution is similar to the values found in Quaranta *et al.* [17].

Using this method with the cellular dispersion, we find that the trait distributions retain their normal shape with no discernible change (not shown). It appears that once the cells can move around, regardless of speed, the selection pressure from spatial competition (which was already small) is almost negligible. There is a more pronounced difference for the cluster configuration with variable migration. These results are shown in figure 4.

We find that there is more competition for space in the cluster configuration, and the faster movers have a selective

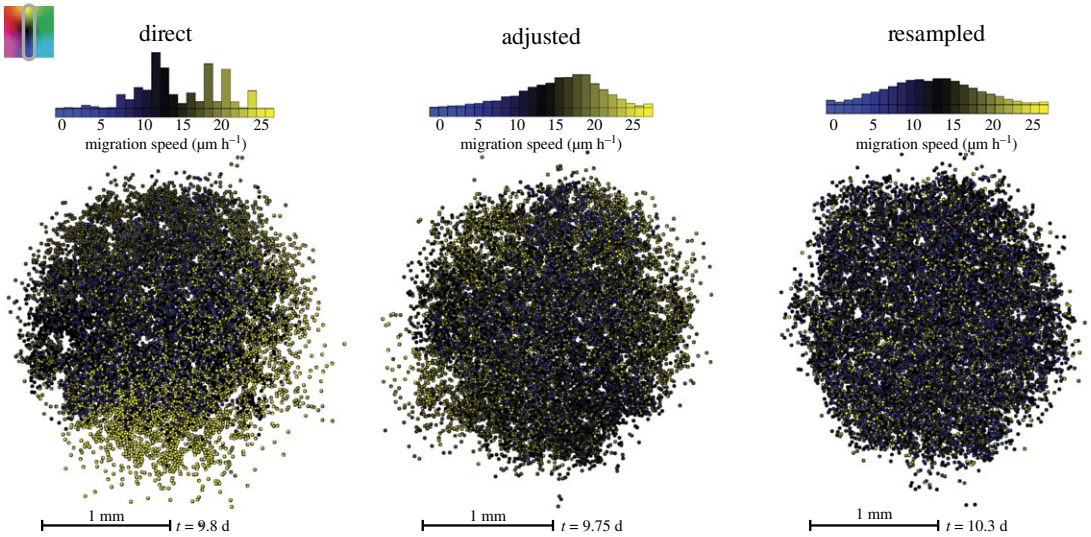

Figure 4. Heterogeneity in migration speeds with different inheriting modes (columns for direct, adjusted and resampled inheritance). The cells all have the same proliferation rate (18 h) but vary in how fast they are moving. The image shows the spatial distribution upon reaching 8000 cells, and the time taken to reach this population is recorded below. The histograms above the images show the distribution of migration speeds at this final population. The gradient from blue to black to yellow represents cells with speeds going from slow to fast.

advantage. However, as they move apart from the mass, it also actually helps the slower movers continue to proliferate by making more space available. So, for direct and adjusted inheritance, we see a shift in the distributions over time weighted towards the faster migrators while still maintaining a spread of values as all cells benefit from the outward expansion. We find that, similar to the effect of inheritance on proliferation rates, direct inheritance results in a distribution that is skewed towards faster migrators but more rugged and irregular where some clones are snuffed out from proliferating by getting caught up in the quiescent bulk. The adjusted inheritance gives a more smooth distribution of values, and for trait resampling, we find again that the normal distribution is maintained.

In general, when cells are allowed to move, the whole structure becomes more jumbled, and the proliferating rim widens. There are subtle shape effects on the cluster at the leading edge. For the direct and adjusted inheritance, we see patches where the cells are more spread out and patches where the cells are more packed together. Looking closer, we find that the more diffuse patches correspond, unsurprisingly, to faster movers, and the more compact edges consist mostly of slower moving cells. This effect is more pronounced in the direct inheritance case, where it is more probable that a cell's neighbours are from the same family and have the same trait values. Resampling migration speeds results in an edge that expands rather evenly but that is still moderately diffuse, and a few stray cells may advance further out because longer persistence times and/or movement directly away from the mass may create slight irregularities.

3.3. Trait combinations: just the sum of the parts?

We now take away all constraints on the phenotypes and allow both traits to vary independently of each other. We initialized the cells by randomly sampling from the same normal distributions for IMTs and migration speeds as in the previous sections. We then ran the same sort of simulations as before. The final distributions of traits are shown in figure 5.

Comparing the histograms from figures 3 and 4, where each trait was varied on its own, with those in figure 5, where both traits vary independently of each other, we note that there is very little difference between the distributions. The dispersions still do not invoke as much selection as the clusters. The direct and adjusted inheritance schemes select for both faster proliferators and faster movers, but the resampled scheme maintains its distribution. The direct inheritance has disjointed peaks as only several clones dominate, and the adjusted inheritance is smooth over more values. Once again, faster proliferators are more strongly selected for than faster movers. As far as the distributions in traits is concerned, the combination of traits appears no different from the sum of the two individual traits.

Because of the greater competition and better image resolution, we show only the spatial results for the cluster configuration, and, though the mass appears somewhat jumbled, we can still pick out some spatial trends. For both the direct and adjusted inheritance, there is some clumping together of similar traits (red and magenta patches), but the clumping is more apparent with direct inheritance. We also note that some slower proliferators are speckled throughout (black and green cells). As in the previous section, it is mainly the fast proliferators that get to the edge, but so do some slow proliferating, fast migrating cells. The shape of the mass with direct inheritance is elongated where the faster proliferators have dominated the region. Where the faster migrators have reached the leading edge, the proliferating rim is more diffuse. The resampled inheritance leads to a mix of cells with normally distributed trait values for both traits. The cells reside in a relatively even spatial distribution throughout the mass and the shape is quite round compared with the others.

The traits do not appear to be correlated. Even though there is a complicated spatial distribution for the traits, it is clear that the combinations of traits that get selected when both traits vary are quite similar to the combination of independent traits that get selected when each trait is

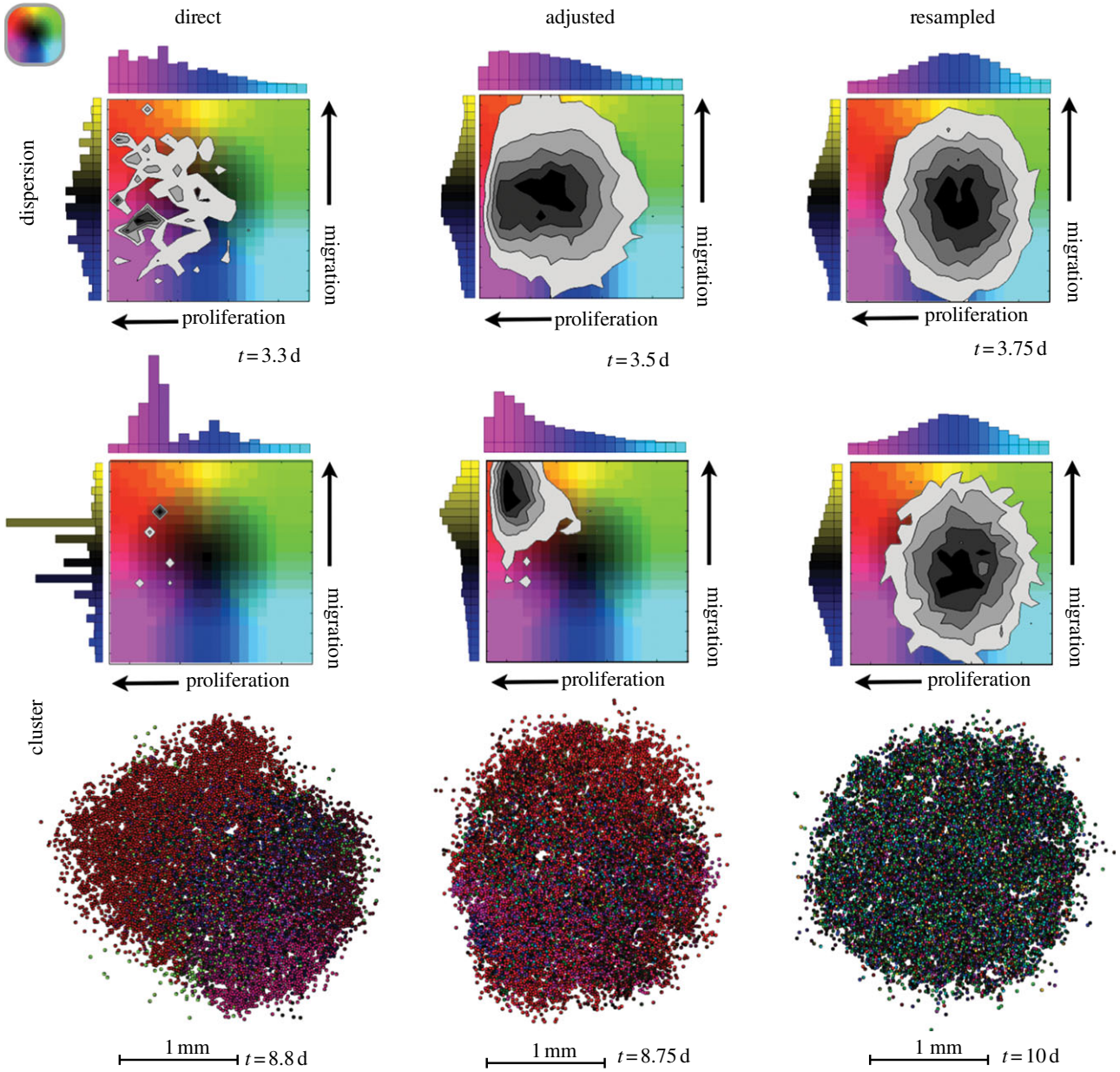

Figure 5. The frequency of occurrence of cells with combinations of traits in the dispersion upon reaching 25 000 cells (top) and the cluster upon reaching 8000 cells (bottom). The time taken to reach these populations is listed. The histograms to the top and side of the phenotype density maps show the distribution of proliferation rates and migration speeds, respectively, with increasing values along the direction of the arrow. The phenotype density map represents the frequency of occurrence of each trait combination (darker greyscale values signify more cells with that combination). The spatial layout shows the cluster of cells coloured according to where they fall in the phenotype space.

varied separately. However, the story changes when we limit the phenotype space by imposing a trade-off between proliferation and migration.

3.4. Go-or-grow: better do both

So far, we have considered the traits of proliferation and migration separately and also together but uncorrelated, but this may not be so. Rapidly advancing through the cell cycle and synthesizing new material for division and actively moving around may cost the cell a large amount of energy. The cell may only have so much energy to expend, so we consider the idea of a trade-off between proliferation and migration, often referred to as go-or-grow. We assume that, if a cell is a faster migrator, it must be a slower proliferator. The migration speed is mapped linearly to the proliferation rate along the diagonal axis of the phenotype space in this manner.

The resulting distributions of traits are shown in figure 6. We see once again that a shift in trait distributions occurs with the direct and adjusted inheritance, whereas trait resampling results in practically no change from the original distribution. For the dispersion, there is a slight shift towards faster proliferating cells, which means that there is a slight shift towards slower migrating cells. This is not surprising with the cells spread out like this, because faster proliferators are more strongly selected for than the faster migrators when the cells already have enough space. But, for the cluster, something interesting happens. It is most apparent in the adjusted inheritance scheme. The same overall shift occurs towards faster proliferators that are slow at moving, but another population also appears to be present. In addition to

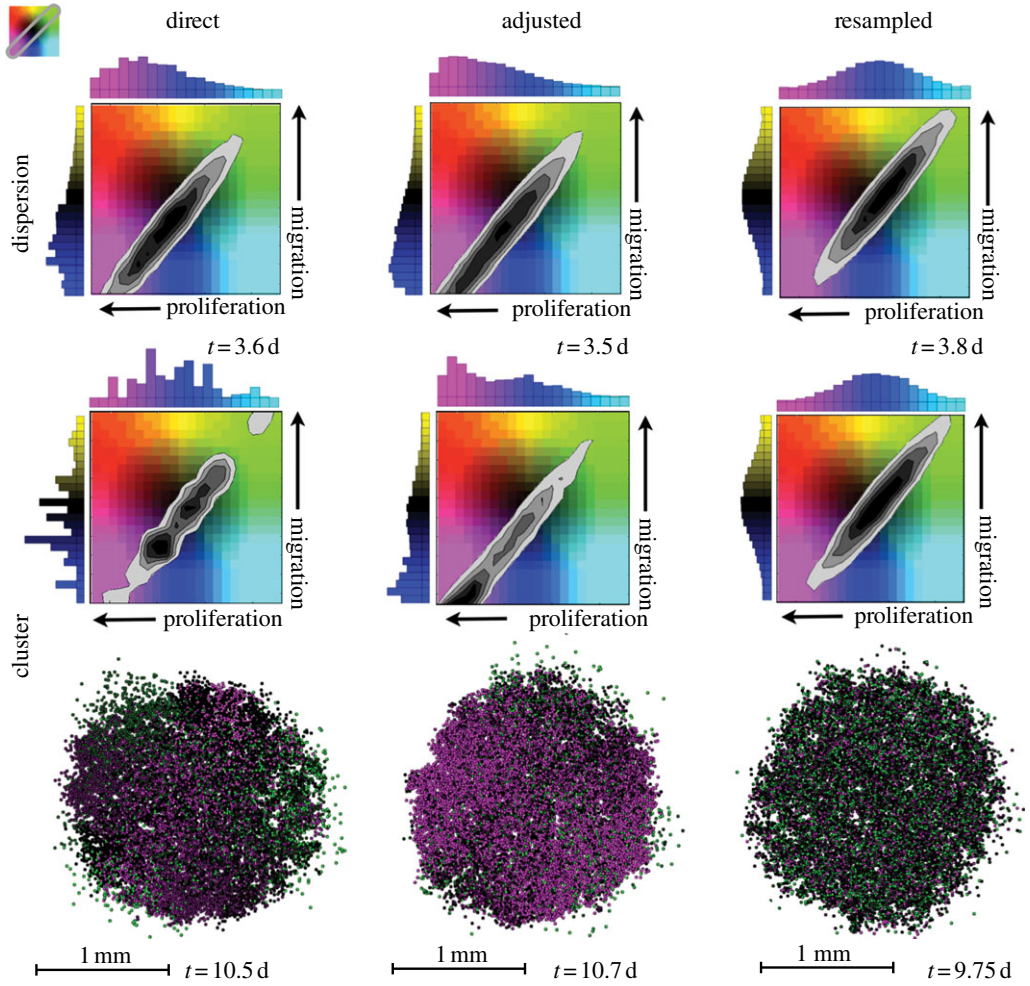

Figure 6. The frequency of occurrence of cells when confined to the diagonal of the phenotype space (go-or-grow) in the dispersion upon reaching 25 000 cells (top) and the cluster upon reaching 8000 cells (bottom). The time taken to reach these populations is listed. The histograms to the top and side of the phenotype density map show the distribution of proliferation rates and migration speeds, respectively, with increasing values along the direction of the arrow. The phenotype density map represents the frequency of occurrence of each trait combination (darker greyscale values signify more cells with that combination). The spatial layout shows the cluster of cells coloured according to where they fall in the phenotype space.

the peak at fast proliferation values, there is another peak in the middle of the distribution where the cells are average at both proliferating and migrating. For direct inheritance, there may be multiple peaks, but with the inherent bumpiness of the distribution it is hard to decipher.

The spatial distributions of the clusters are presented in figure 6 for the three types of inheritance. We find that, in general, the assemblage is more heterogeneously mixed than with the other constraints. With direct and adjusted inheritance, there are some regions with similar neighbours, but there are also regions that are more mixed. The resampled inheritance again maintains its usual diversity.

Further investigation of the adjusted inheritance for the cluster configuration reveals that the weak bimodality is actually quite unstable. The peak may either stay in the middle, move towards faster proliferators or become bimodal with both peaks present to some degree, as previously described and shown in figure 6. The initial distribution has a significant effect on where this final distribution will stabilize. Figure 7 shows the final configurations for simulations starting with all cells either fast at proliferating, fast at migrating or average at both. In these cases, the distribution tends to remain centred around where it began. However, when the trait values begin in the middle of the range, several outcomes are possible: either the peak remains in the middle, another peak forms at fast or slow IMTs, or several peaks arise. The appearance of a multi-modal distribution or relatively stable distributions that have peaks at different values most probably means that many trait values are equally fit.

3.5. The heterogeneity index: quantifying the phenotypic mix

So far, we have shown how different constraints on the phenotype space in heterogeneous tumour growth can affect cell population heterogeneity both temporally and spatially. To understand the trait variation and allotment within a range of values, we have plotted histograms of single traits. To understand the variation and abundance of trait combinations, we have shown phenotype density maps. Now, we want to compare the previous results by describing the heterogeneity numerically. We use the heterogeneity index outlined in Appendix 1 as our metric for comparison.

In the previous sections, we showed typical results for each scenario; however, with the amount of stochasticity involved in these simulations, variation is expected. Here, we compile from 15 different simulations of each scenario the heterogeneity indices for proliferation rates (figure 8)

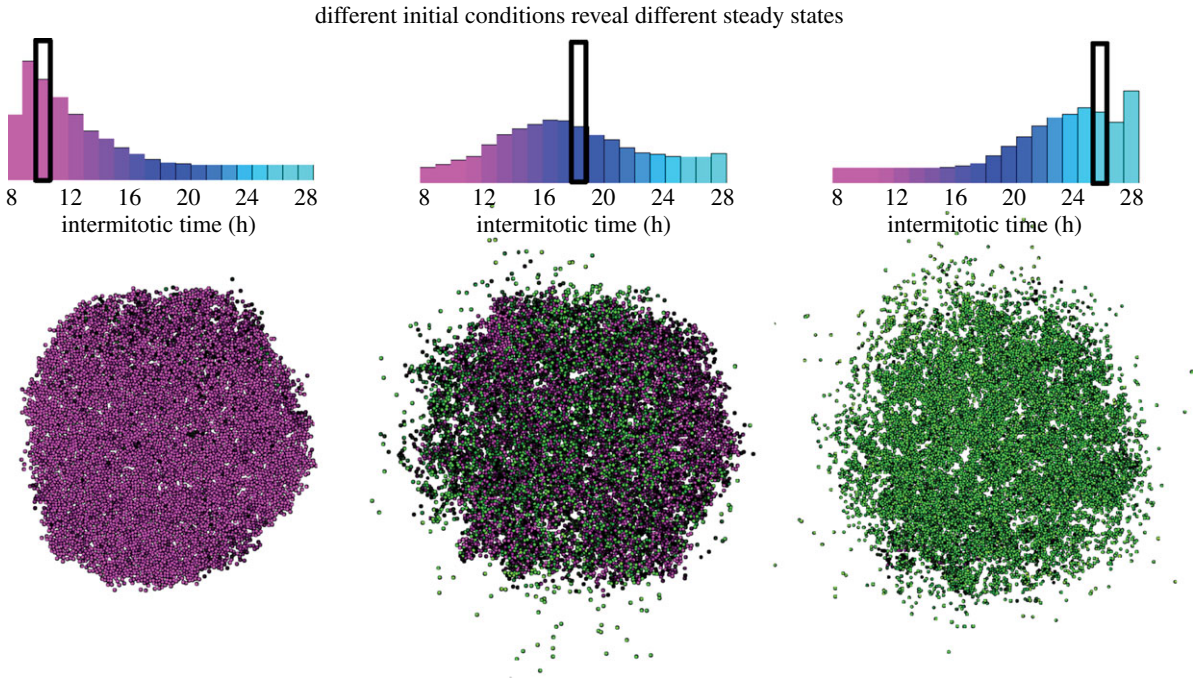

Figure 7. With the adjusted inheritance, the go-or-grow constraint yields different populations of clusters grown to 8000 cells depending on the initial distribution. All initial distributions (black box) are monoclonal with IMTs at 10 h (left), 18 h (middle) and 26 h (right).

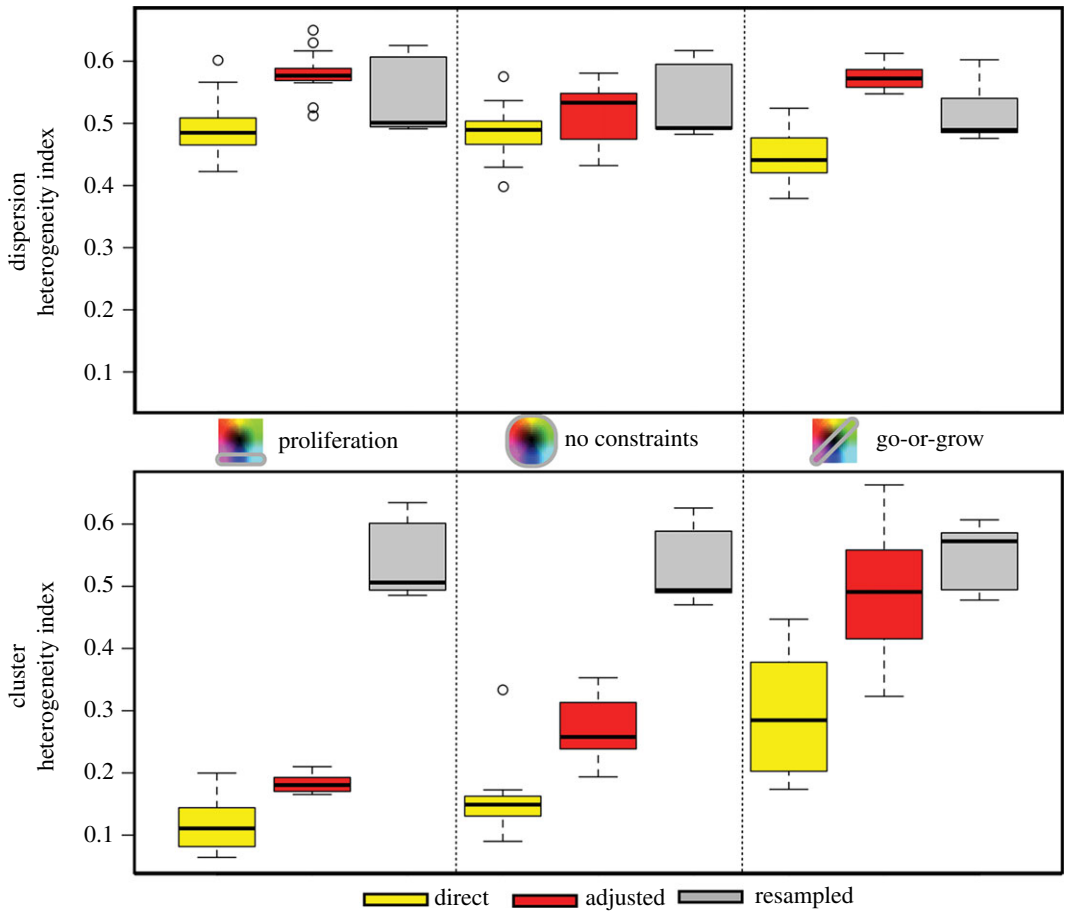

Figure 8. Heterogeneity in proliferation rate as calculated via equation (A 1) for each constraint on phenotype space (columns), for each inheritance type (colours) and for each spatial configuration (top is the dispersion and bottom is the cluster). The columns are for: proliferation only, both traits with no constraints and go-or-grow. From 15 different runs, the range, upper quartile and lower quartile are shown.

and for migration speeds (figure 9). The heterogeneity index H was calculated from the final trait distributions.

The mean heterogeneity index is generally higher in the dispersions where there is less selection than in the clusters, but there are also differences between the different inheritance modes. The direct inheritance scheme consistently yields the smallest mean H across all scenarios. This is to be expected, as no new traits can be gained but only lost. The

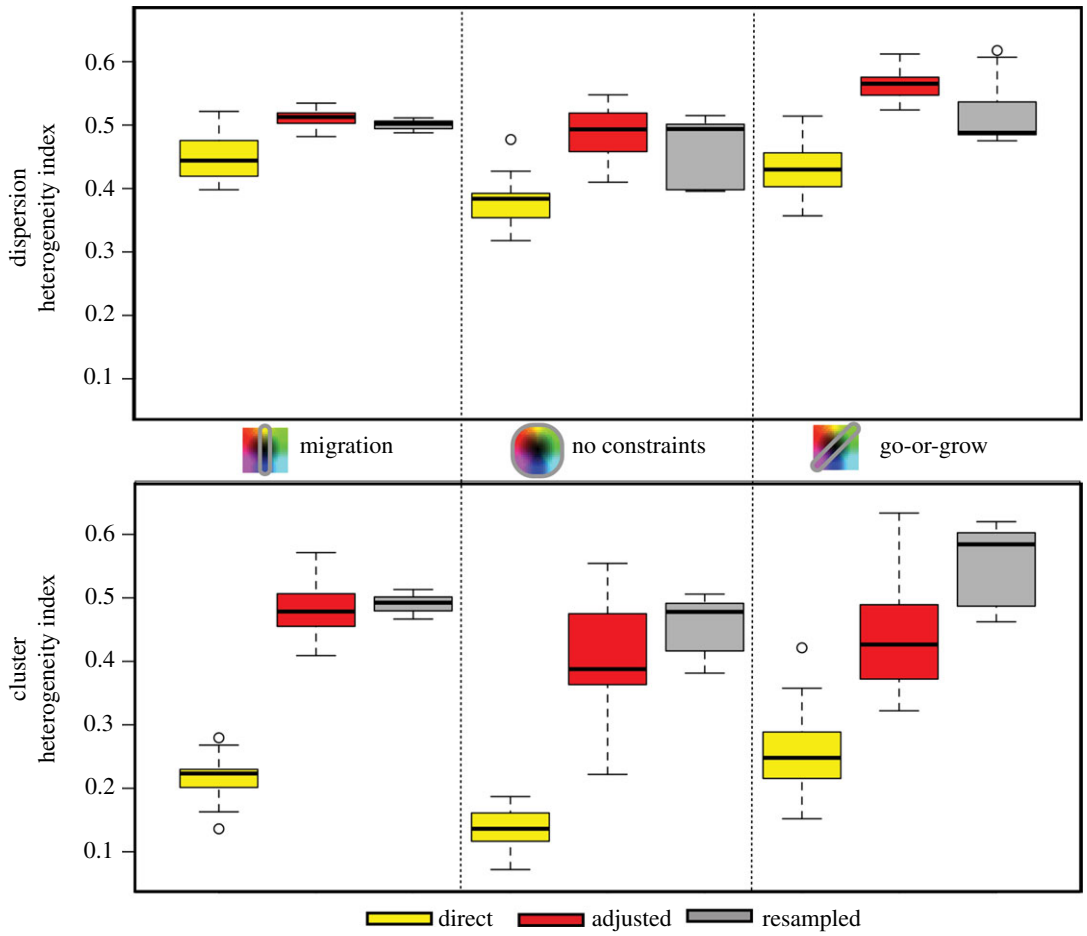

Figure 9. Heterogeneity in migration rate as calculated via equation (A 1) for each constraint on phenotype space (columns), for each inheritance type (colours) and for each spatial configuration (top is the dispersion and bottom is the cluster). The columns are for: migration only, both traits with no constraints and go-or-grow. From 15 different runs, the range, upper quartile and lower quartile are shown.

resampled scheme appears to remain relatively steady over the different constraints, while the mean values of H for the adjusted inheritance scheme jump around to different values.

For the cluster, the mean H always increases from direct to adjusted to resampled inheritance, and the extra competition in the cluster configuration makes the homogeneity from direct inheritance more pronounced. The first two columns reflect that the mean values for both the direct and adjusted modes move towards faster proliferation, so this results in a distribution that is more homogeneous. However, with go-or-grow, there is more heterogeneity, and also more variation in the amount of heterogeneity, because many outcomes are possible. This results in more heterogeneity, and also more variation in the amount of heterogeneity as there are many outcomes possible.

When comparing the heterogeneity indices for migration, similar trends are present, but there are differences. Again, direct inheritance has the lowest mean value of H, but resampling does not always have the highest mean H. With migration, the adjusted scheme will often tend towards a very spread out distribution, which translates to a large H. But the heterogeneity index is largest in the go-or-grow scenario with the resampling scheme. Now, with a good understanding of heterogeneity in this system, we define how it leads to different population fitnesses.

3.6. Fitness: different fates for different constraints

The heterogeneity of a population and its distribution of phenotypes gives some measure of its capacity to survive in different environments, but there are other ways to characterize the population's fitness. The most obvious metric is the overall growth rate of the population, which we will evaluate. We will also look at the proportion of proliferators in the population, which gives a measure of spatial diffusiveness, a degree of competition from neighbours, and an overall capacity for continued proliferation.

The growth rates are quite different for the dispersion and the cluster (figure 10*a*). The dispersion grows exponentially ($N \propto e^{at}$) until reaching confluence, and then it levels off as the space eventually fills to capacity. For the cluster configuration, the cells do not immediately compete with each other as there is some space between them from the random nature of the initial placement. So, after a quick growth spurt when the population is small, the mass gets larger and the fraction of cells at the leading edge that can proliferate reduces. Figure 10*a* shows the cluster growth fit to a power law ($N \propto t^b$).

The proliferating fraction for the dispersion and the cluster have different temporal dynamics (figure 10*b*). For the dispersion, all cells start their own colonies with 100 per cent in the proliferating state. Each colony grows initially like its own separate cluster building up individual quiescent cores, but quickly the space fills up and neighbouring colonies become the competition. The initial proliferating fraction for the cluster is smaller than the dispersion. This fraction shrinks very gradually as the quiescent tumour bulk becomes a larger and larger piece of the whole population compared with the proliferating rim.

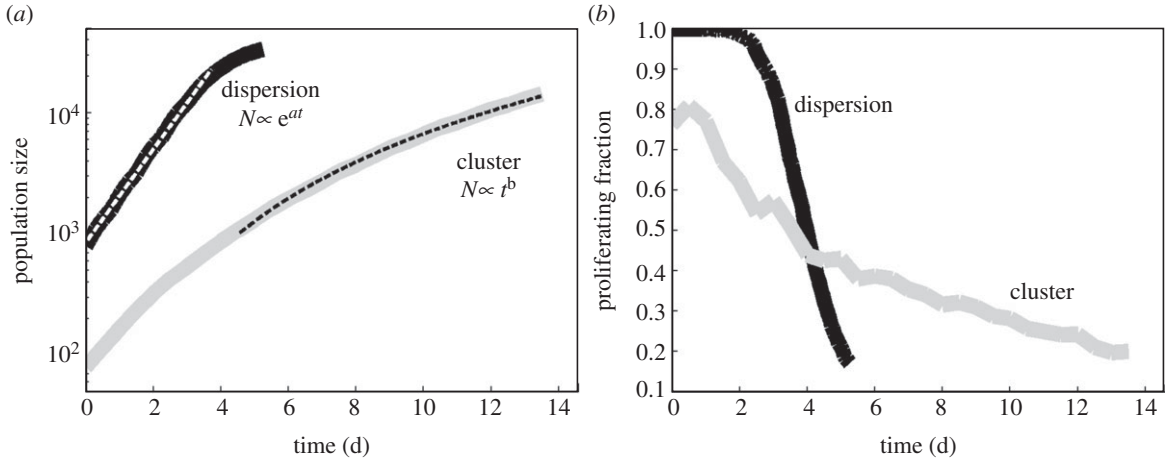

Figure 10. (*a*) Typical growth rates of the dispersion and cluster configurations. The dispersion is fit with an exponential, whereas the cluster is fit to a power law (both with $R^2 > 0.999$). (*b*) The percentage of cells of the whole population that are in the proliferating state. The two spatial configurations clearly lead to different degrees of competition.

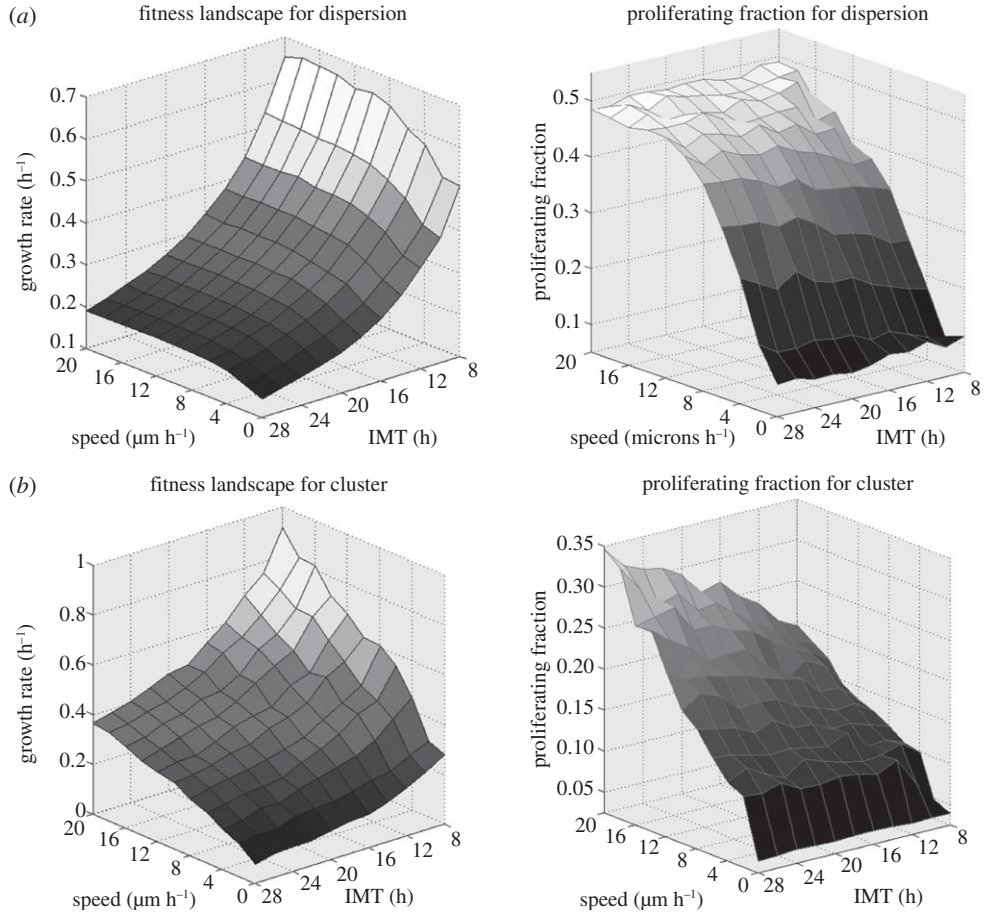

Figure 11. Growth rates (left) and proliferating fractions (right) of monoclonal populations with various combinations of proliferation and migration rates. To quantify the growth rate and proliferating fractions: (i) the dispersion was grown from 800 to 25 000 cells and (ii) the cluster was grown from 80 to 8000 cells. The growth rates are found by dividing the total change in the number of cells by the total change in time. This is then normalized by the initial population. The proliferating fraction is taken at the final time step.

Understanding how the fitness is affected by different inheritance schemes and different constraints on heterogeneity is not straightforward. We ease into this intricacy by first looking at the fitness of trait combinations in homogeneous populations. For simplicity and to be able to compare the dispersion with the cluster, we correlate fitness with the growth rate of the population, which we take as $\Delta N/(N_0 \Delta t)$, where N_0 is the initial population, ΔN is the change in total number of cells and Δt is the total time.

We grow monoclonal populations (all cells have identical trait values which are passed on directly to their progeny) and record the overall growth rate. We take the proliferating fraction at the end of each simulation. The results are shown in figure 11 as fitness landscapes over the two-dimensional space of phenotype combinations.

For the cellular dispersion in figure 11*a*, the growth rate surface plot indicates that being a faster proliferator makes the biggest impact on fitness. Increasing the speed is shown to

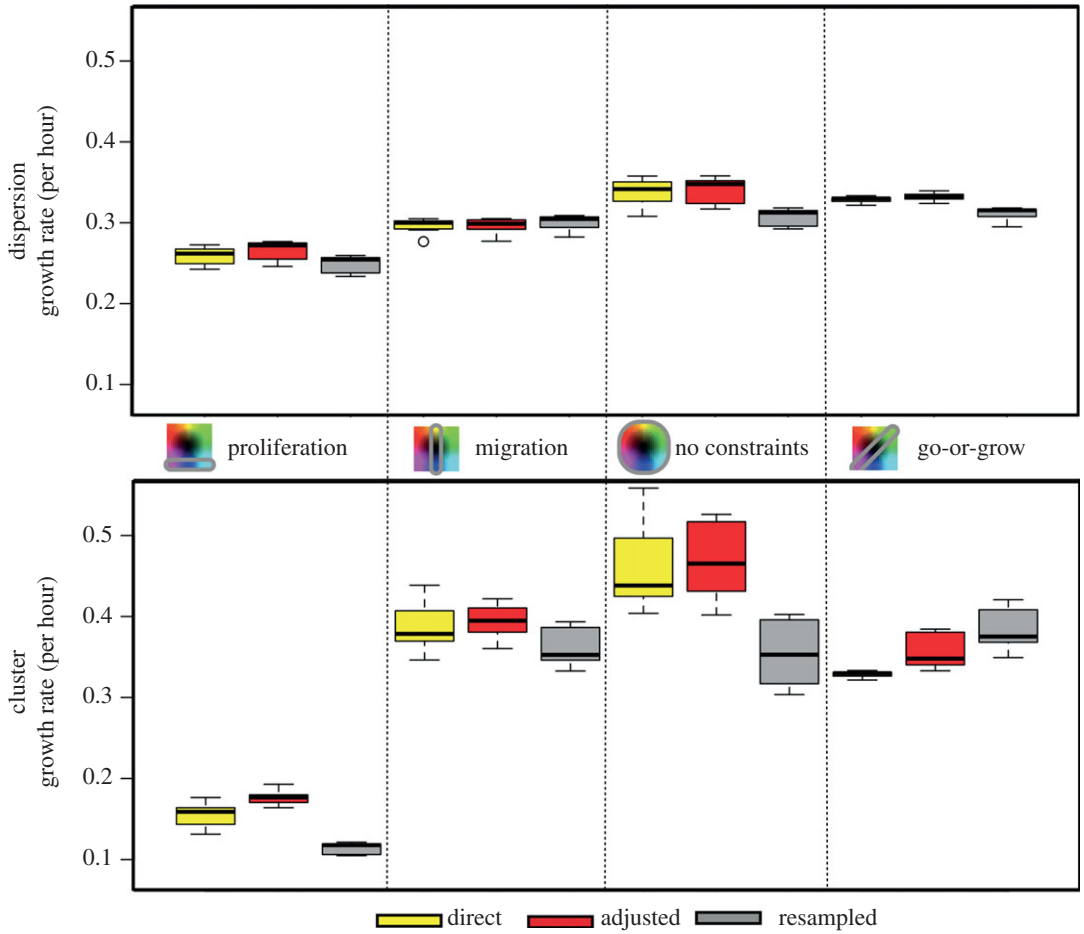

Figure 12. Growth rates for the dispersion (top) and the cluster (bottom) for each constraint on phenotype space (columns) and for each inheritance type (colours). From 15 different runs, the range, upper quartile and lower quartile are shown.

increase the proliferating fraction, but it does little to increase the overall growth. However, there is a noticeable decline in the growth rates at very slow migration speeds. In this configuration, the cells are already proliferating exponentially, so allowing them to spread out hardly makes a difference.

For the cellular cluster in figure 11*b*, the surface plot shows that the most fit combination of traits corresponds to the fast proliferating, fast migrating phenotype and the least fit combination corresponds to the slow proliferating, slow migrating phenotype. This agrees with what we find to be the most fit when there is heterogeneity in the traits—the population trends towards this most fit combination. However, we also see that there are many combinations in between these extremes that share the same values for fitness. From figure 11*b*, we find that, upon the line allowed by the go-or-grow condition, most values are equally fit. The combination with the largest proliferating fraction, however, does not correspond to the fastest growth rate. The largest proliferating fraction occurs when the cells move out fast, but proliferate slowly. This combination creates a very diffuse tumour that invades further before filling in the space left behind.

Analysis on monoclonal populations tells us only so much. It is not always the fastest moving and most proliferative populations that come out on top when many interacting parts make up the whole. Just as certain combinations of traits may work well together, combinations of neighbours may interact differently. Furthermore, the fitness of the population as a whole cannot be determined by summing the fitnesses of the clones that make up the population, as we see next.

We now address the complexity of heterogeneity from the previous sections. We compare the fitness metrics of each of the different inheritance schemes with different constraints and different spatial configurations by compiling results from 15 simulations for each scenario. First, we look at the fitness of the populations, measured by the growth rate (figure 12).

For the dispersion configuration, there is less difference between the three inheritance schemes; however, there is a tendency for the resampled inheritance to yield slower growing populations than the adjusted and direct inheritance schemes. But when migration varies as a single trait, increasing migration speeds (as seen with direct and adjusted inheritance) do little to help when there is already such meagre competition, so there is hardly any advantage to select, and the three inheritance schemes produce similar results.

The growth rates for the cluster configuration are more distinct between inheritance schemes. In the first three columns, the adjusted and direct inheritance schemes yield faster growth rates than the resampled scheme, which lags behind. But in the last column, for go-or-grow, the resampled scheme's mean growth rate is faster. In the first three columns, we know that the distribution tends towards homogeneous populations with faster proliferators and faster migrators. But when these values are no longer allowed and many trait values are equally fit (in go-or-grow), this heterogeneity that previously reduced the fitness of the population is now beneficial. Though the adjusted and direct inheritance schemes are seen to also lead to increases in the heterogeneity indices for both traits in the go-or-grow case, the resampled scheme has more local heterogeneity. Local

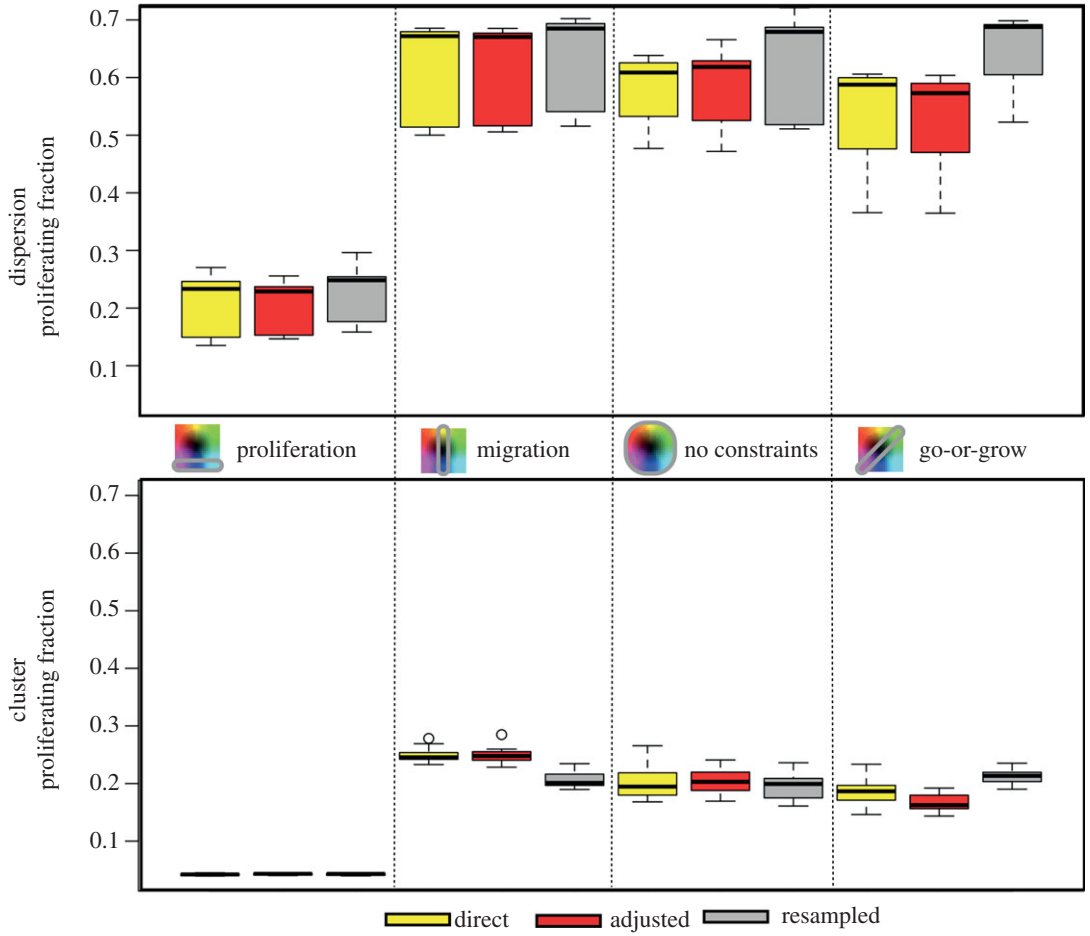

Figure 13. The percentage of the total population of cells that are proliferating for the dispersion (top) and the cluster (bottom). Fifteen different runs make the variance in the plot.

heterogeneity is harder to achieve when the inheritance has some familial memory (direct and adjusted inheritance gives rise to daughter cells exactly like or similar to their parent, respectively). These schemes will never achieve the same degree of local heterogeneity that the resampled scheme is capable of providing when neighbours can easily be different from each other. This is presumably a growth advantage.

Fitness is not just a measure of the growth rate of a population. A population of cells may also be more fit if there is a larger proportion of cells in the proliferating state. So, we examine, for all situations, the percentage of proliferators of the total population at the end of each simulation (figure 13).

We record the proliferating fraction for the dispersion when it is around 60 per cent confluent, whereas the cluster's proliferating fraction is measured when it reaches a size of around 2 mm in diameter. Therefore, the proliferating fraction is always larger in the dispersion than in the cluster because we are comparing the edges of many proliferating colonies with the edge of one big proliferating mass, respectively. Also, when the cells do not migrate (the first column), the mean fraction is always smaller, because there is no diffuse boundary of cells at the edges that increase this proliferation ratio value. Once the cells have some movement, the edges become more diffuse, and this value increases.

For the dispersion, the mean fraction of proliferators for the resampling inheritance mode is larger than the others, and this is more pronounced in the go-or-grow scenario. For the cluster, the resampling scheme actually appears to have mean proliferating fractions that are less than or equal to the other schemes, yet the value once again is largest for the go-or-grow scenario. The drift towards faster migrators with direct or adjusted inheritance could lead to a diffuse boundary, but the local heterogeneity found with the resampled inheritance could also lead to a diffuse edge by avoiding large pockets of closely packed fast proliferators.

We have examined three metrics to numerically evaluate the heterogeneous population growth of this system: the heterogeneity index, the growth rate and the proliferating fraction. Not one of these metrics gives the complete understanding of the dynamics, but each has provided some insight into piecing together how the many interactions between the parts lead to the dynamics of the whole.

4. Discussion

This analysis has brought to light that, with heterogeneity in traits, different modes of inheritance of phenotypes can result in different outcomes for population composition, heterogeneity and fitness. The only environmental constraint considered here is space, and lack of space can drive selection towards populations that are more or less heterogeneous depending on the allowed trait combinations.

An obvious question to ask is, is one of these inheritance schemes more realistic? It is probable that some traits do not change often upon division, as with direct inheritance, and that some adaptability is also allowed, as with the adjusted inheritance. The unchanging diversity of the resampling

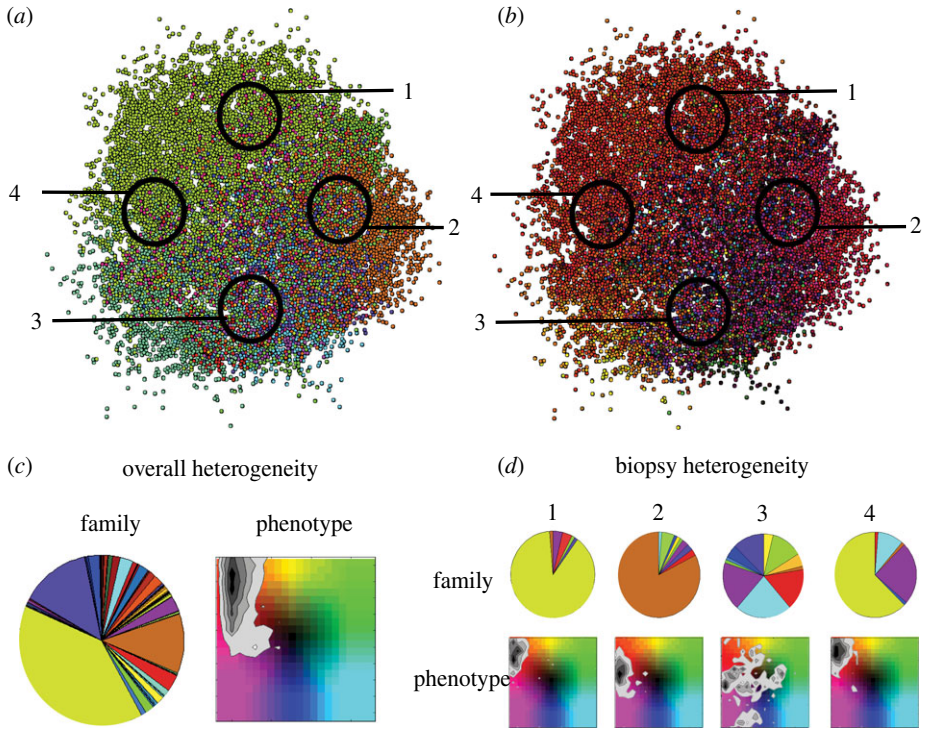

Figure 14. Heterogeneity from (*a*) familial origin and (*b*) phenotype does not always correspond. Site-specific differences are also found. We show heterogeneity (*c*) overall and (*d*) within each biopsy. We see that even if a tumour may have several dominant clones, the phenotypes may converge.

scheme seems quite improbable, yet we saw in the go-or-grow scenario that this local heterogeneity provides an advantage to the growth of the population, and it might also be beneficial when first entering a new microenvironment (e.g. a new metastatic niche) to be maximally heterogeneous to better evaluate which traits work best. With all of the ways that function actually comes about in a cell, it is also not necessary that every trait will be passed on in the same way.

It is clear that, with an inheritance scheme that allows plasticity of a phenotype, the traits of cells in a population may converge to be more homogeneous even though the cells come from different clones of origin. This idea of convergent evolution is not new, but has recently been demonstrated by genetically characterizing biopsies from different sites within a tumour [3,4]. To explicitly confirm this phenomenon in our model, we use the scenario where the trait combinations are unlimited and the inheritance is adjusted (figure 14).

We colour the images in figure 14 according to 'families', or progeny from the same clone of origin (*a*), and position in the two-dimensional phenotype space as before (*b*). We see that each biopsy site results in different distributions of families and different distributions of phenotypes. Further, the originating family and phenotype need not be related. Looking at the overall heterogeneity, it is clear that there are multiple families producing the same phenotypic behaviour (the phenotype density map shows convergence towards faster proliferation and faster migration). If we then look at each biopsy in isolation this pattern is similar for biopsies 1, 2 and 4. In biopsy 3, we see that the families and phenotypes are similarly diverse. This naive example shows how looking at the genotype (or here, the original clonal family) might be misleading, and as these tumours grow the divergence between genotypes will grow while the convergence to phenotypes will also increase. It is worth noting that there is still utility in looking at the spatial distribution of genotypes as those that dominate will be correlated with the fitter phenotypes. However, our example here has only 80 initial 'genotypes' that do not mutate and is certainly a gross underestimation of reality. We therefore believe that it is critical to characterize the phenotypic heterogeneity (at the single cell level) even more than the genotypic heterogeneity as ultimately this will facilitate our understanding of how the tumour grows and might be better treated.

Advances in single cell analysis have allowed us to examine more fully the heterogeneity of genetic and phenotypic states of individual cells [17,29]. But perhaps more important than understanding the variation in single traits is the heterogeneity of certain trait combinations and how they are distributed spatially. For just the two traits we have examined here, there is the potential for one trait to influence the other. The faster a cell divides the faster the whole population can grow, but if it is not moving apart it only keeps a very small proliferating rim so that the growth is limited. There is a balance needed for optimal growth between keeping the cells dividing to grow the population and moving out to keep the cells in the proliferating state. Beyond just looking at growth rates for a measure of fitness, the proportion of the population that is proliferating is significant.

We incorporate competition for space as the only environmental factor (in both the dispersion and the cluster configuration), which provides a very simplistic view of proliferation and invasion of a population. Nevertheless, an important feature of these two very different spatial configurations is that the growth, evolution, heterogeneity and fitness of the populations are dramatically different. We should be careful in interpreting *in vitro* models in which the cells do not experience this competition or make more observations when cells are closer to confluence. We know that *in vitro* models are, nonetheless, models and should be considered

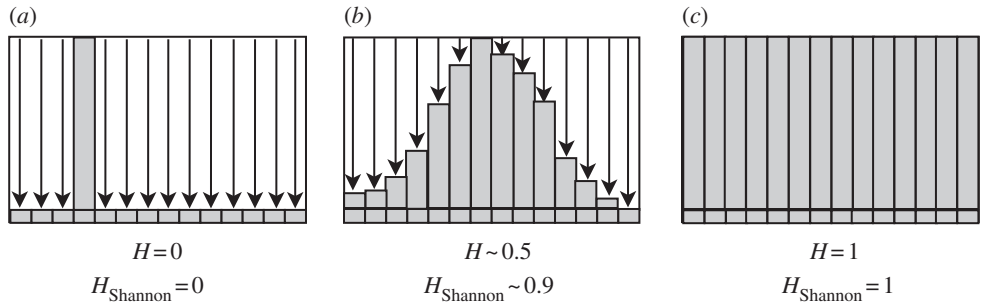

Figure 15. An illustration of the heterogeneity index as calculated via equation (A 1), and with the Shannon index from the frequency distribution of trait values. If all cells in the population have the same trait value, we get a minimum heterogeneity index (*a*). If the population has an equal spread of trait values throughout the range, the heterogeneity index is maximum (*c*). The two values differ in the configuration shown in (*b*), where the Shannon index grows logarithmically compared with the linear growth used here.

as such. Biological models such as tumour spheroids, scratch assays, three-dimensional cultures and the Nest assay [30,31] are better at mimicking spatial competition over standard cell cultures. These models represent the migration much more realistically, but, with so many combinations of traits, there is no way to study them all without the help of mathematical and computational models.

Some scenarios played out here may be more realistic than others though all are quite abstract from the complexity of a real tumour. Though abstract, this method of heterogeneity analysis of trait inheritance could be played out with any combination of traits. We could also consider the evolution of populations with more than two traits as angiogenesis, metabolism, build-up of acidity and other processes change the microenvironment. The complexity with just two traits is already quite extensive, but a multi-dimensional phenotype space is also possible. With each additional trait considered, the combinations of traits grows quickly, but realistic constraints on these combinations may significantly reduce the dimension, as in the go-or-grow scenario.

The spatial and temporal evolution of phenotypic heterogeneity (as defined by trait combinations) in a tumour has significant implications for the treatment of that tumour [32]. With the growing trend of genetic characterization of tumours from a single biopsy, we may only be capturing a skewed subset of the whole heterogeneity. Beyond the potential for misconstruing the type of treatment based on where the biopsy is obtained, when a certain clone or phenotype is targeted, we may be freeing up space and resources for other cell types to take over. If we look at tumour growth and regrowth during and after treatment as an evolutionary process, we may get a better understanding of how to best prevent recurrence. The mode of phenotypic inheritance directly affects tumour growth, and the interplay between the various traits is a source of complexity that deserves much more attention from the biological community as this ultimately may be the place where the best therapeutic strategies for targeting a given tumour are found.

The authors gratefully acknowledge funding from the NCI Integrative Cancer Biology Program grant no. U54 CA113007.

Appendix A. The heterogeneity index

Common formulae for quantifying heterogeneity include those proposed by Shannon [33] and Simpson [34] in the late 1940s. Both of these values grow quickly as sample size increases and slower as the numbers even out. We want to avoid this bias and present a value that represents the occupied proportion of the available space. This point can be made clear with figure 15, which shows a distribution calculated by the heterogeneity index described here, which ranges from 0 to 1 linearly, and also by the normalized Shannon index, which is logarithmic, for comparison.

Basically, we sum up, over all bins, the difference between each bin and the bin with the highest frequency. So, if all traits occupy one bin, we get a maximum value, and if all bins are equally populated we get nil. We then subtract the summation from 1 to get the heterogeneity index. The equation for the heterogeneity index is as follows:

$$H = 1 - \frac{1}{N_{\max}(\tau_{\max} - \tau_{\min} - \Delta\tau)} \sum_j (N_{\max} - N_j)\Delta\tau. \quad \text{(A 1)}$$

This equation defines a simple measure for how disperse the values are within a trait's range $\tau \in (\tau_{\min}, \tau_{\max})$. It ensures that $H = 1$ when the range of trait values are filled equally and $H = 0$ when all cells have the exact same trait value (figure 15). We define $N_{\max}$ as the bin with the highest frequency, N_j as the occupancy of every other bin and $\Delta\tau$ as the bin size. Choosing the range and bin sizes appropriately is an important part of determining this value.

References

1. Merlo LMF, Maley CC. 2010 The role of genetic diversity in cancer. *J. Clin. Invest.* **120**, 401–403. (doi:10.1172/JCI42088)
2. Marusyk A, Almendro V, Polyak K. 2102 Intratumour heterogeneity: a looking glass for cancer? *Nat. Rev. Cancer* **12**, 323–334. (doi:10.1038/nrc3261)
3. Gerlinger M *et al*. 2012 Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366**, 883–892. (doi:10.1056/NEJMoa1113205)
4. Sottoriva A *et al*. 2013 Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc. Natl Acad. Sci. USA* **110**, 4009–4014. (doi:10.1073/pnas.1219747110)
5. Marusyk A, Polyak K. 2013 Cancer cell phenotypes in fifty shades of grey. *Science* **339**, 528–529. (doi:10.1126/science.1234415)
6. Kreso A *et al*. 2013 Variable clonal repopulation dynamics influence chemotherapy response in

colorectal cancer. *Science* **339**, 543–548. (doi:10.1126/science.1227670)
7. Rando OJ, Verstrepen KJ. 2007 Timescales of genetic and epigenetic inheritance. *Cell* **106**, 19 352–19 357.
8. Anderson ARA, Weaver AM, Cummings PT, Quaranta V. 2006 Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. *Cell* **127**, 905–915. (doi:10.1016/j.cell.2006.09.042)
9. Anderson ARA *et al*. 2009 Microenvironmental independence associated with tumor progression. *Cancer Res.* **69**, 8797–8806. (doi:10.1158/0008-5472.CAN-09-0437)
10. Niepel M, Spencer SL, Sorger PK. 2009 Non-genetic cell-to-cell variability and the consequences for pharmacology. *Curr. Opin. Chem. Biol.* **13**, 556–561. (doi:10.1016/j.cbpa.2009.09.015)
11. Brock A, Chang H, Huang S. 2009 Non-genetic heterogeneity—a mutation-independent driving force for the somatic evolution of tumours. *Nat. Rev. Genet.* **10**, 336–342. (doi:10.1038/nrg2556)
12. Spencer SL, Gaudet S, Albeck JG, Burke JM, Sorger PK. 2009 Non-genetic origins of cell-to-cell variability in trail-induced apoptosis. *Nature* **459**, 428–432. (doi:10.1038/nature08012)
13. Ma C *et al*. 2011 A clinical microchip for evaluation of single immune cells reveals high functional heterogeneity in phenotypically similar *t*-cells. *Nat. Med.* **17**, 738–744. (doi:10.1038/nm.2375)
14. Burga A, Lehner B. 2012 Beyond genotype to phenotype: why the phenotype of an individual cannot always be predicted from their genome sequence and the environment that they experience. *FEBS J.* **279**, 3765–3775. (doi:10.1111/j.1742-4658.2012.08810.x)
15. Lachmann M, Jablonka E. 1996 The inheritance of phenotypes: an adaptation to fluctuating environments. *J. Theor. Biol.* **181**, 1–9. (doi:10.1006/jtbi.1996.0109)
16. Staudte RG, Huggins RM, Zhang J, Axelrod DE, Kimmel M. 1997 Estimating clonal heterogeneity and interexperiment variability with the bifurcating autoregressive model for cell lineage data. *Math. Biosci.* **143**, 103–121. (doi:10.1016/S0025-5564(97)00006-0)
17. Quaranta V, Tyson DR, Garbett SP, Weidow B, Harris MP, Georgescu W. 2009 Trait variability of cancer cells quantified by high-content automated microscopy of single cells. *Meth. Enzymol.* **467**, 23–57. (doi:10.1016/S0076-6879(09)67002-6)
18. Feinberg AP, Irizarry RA. 2010 Stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. *Proc. Natl Acad. Sci. USA* **107**, 1757–1764. (doi:10.1073/pnas.0906183107)
19. Sottoriva A, Vermeulen L, Tavaré S. 2011 Modeling evolutionary dynamics of epigenetic mutations in hierarchically organized tumors. *PLoS Comput. Biol.* **7**, 1–11. (doi:10.1371/journal.pcbi.1001132)
20. Guerrero-Bosagna C, Skinner MK. 2012 Environmentally induced epigenetic transgenerational inheritance of phenotype and disease. *Mol. Cell. Endocrinol.* **354**, 3–8. (doi:10.1016/j.mce.2011.10.004)
21. Anderson ARA. 2005 A hybrid mathematical model of solid tumor invasion: the importance of cell adhesion. *Math. Med. Biol.* **22**, 163–186. (doi:10.1093/imammb/dqi005)
22. Lee Y, Kouvroukoglou S, McIntire LV, Zygourakis K. 1995 A cellular automaton model for the proliferation of migrating contact-inhibited cells. *Biophys. J.* **69**, 1284–1298. (doi:10.1016/S0006-3495(95)79996-9)
23. Giese A, Bjerkvig R, Berens M, Westphal M. 2003 Cost of migration: invasion of malignant gliomas and implications for treatment. *J. Clin. Oncol.* **21**, 1624–1636. (doi:10.1200/JCO.2003.05.063)
24. Mansury Y, Diggory M, Deisboeck TS. 2006 Evolutionary game theory in an agent-based brain tumor model: exploring the 'genotype-phenotype' link. *J. Theoret. Biol.* **238**, 146–156. (doi:10.1016/j.jtbi.2005.05.027)
25. Böttget K, Hatzikirou H, Chauviere A, Deutsch A. 2012 Investigation of the migration/proliferation dichotomy and its impact on avascular glioma invasion. *Math. Model. Nat. Phenom.* **7**, 105–135. (doi:10.1051/mmnp/20127106)
26. Greaves M, Maley CC. 2012 Clonal evolution in cancer. *Nature* **481**, 306–313. (doi:10.1038/nature10762)
27. Codling EA, Plank MJ, Benhamou S. 2008 Random walk models in biology. *J. R. Soc. Interface* **5**, 813–834. (doi:10.1098/rsif.2008.0014)
28. Jeon J, Quaranta V, Cummings PT. 2010 An off-lattice hybrid discrete-continuum model of tumor growth and invasion. *Biophys. J.* **98**, 37–47. (doi:10.1016/j.bpj.2009.10.002)
29. Rowat AC, Weitz DA. 2009 Understanding epigenetic regulation: tracking protein levels across multiple generations of cells. *Eur. Phys. J. Spec. Top.* **178**, 71–80. (doi:10.1140/epjst/e2010-01183-5)
30. Kam Y, Karperien A, Weidow B, Estrada L, Anderson ARA, Quaranta V. 2009 Nest expansion assay: a cancer systems biology approach to *in vitro* invasion measurements. *BMC Res. Notes* **2**, 130–139. (doi:10.1186/1756-0500-2-130)
31. Kam Y, Rejniak KA, Anderson ARA. 2011 Cellular modeling of cancer invasion: integration of *in silico* and *in vitro* approaches. *J. Cell Physiol.* **227**, 431–438. (doi:10.1002/jcp.22766)
32. Swanton C. 2012 Intratumor heterogeneity: evolution through space and time. *Cancer Res.* **72**, 4875–4882. (doi:10.1158/0008-5472.CAN-12-2217)
33. Shannon CE. 1948 A mathematical theory of communication. *Bell Sys. Tech. J.* **27**, 379–423.
34. Simpson EH. 1949 Measurement of diversity. *Nature* **163**, 688. (doi:10.1038/163688a0)