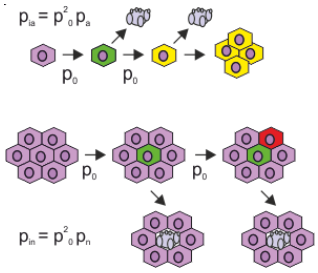
Cooperation of partially-transformed clones: an invisible force behind the early stages of carcinogenesis

Alessandro Esposito

The Medical Research Council Cancer Unit, University of Cambridge, Hills Road, Cambridge CB2 0XZ, UK

Correspondence should be addressed to A.E. (email: ae275@mrc-cu.cam.ac.uk)

Running title: Oncogenic field effect and clonal heterogeneity



Keywords: oncogenesis, non-cell-autonomous, model

Abstract

Most tumours exhibit significant heterogeneity and are best described as communities of cellular populations competing for resources. Growing experimental evidence also suggests, however, that cooperation between cancer clones is important as well for the maintenance of tumour heterogeneity and tumour progression. However, a role for cell communication during the earliest steps in oncogenesis is not well characterised despite its vital importance in normal tissue and clinically manifest tumours. By modelling the interaction between the mutational process and cell-to-cell communication in three-dimensional tissue architecture, we show that non-cell-autonomous mechanisms of carcinogenesis are likely to support and accelerate pre-cancerous clonal expansion through the cooperation of different, non- or partially- transformed mutants. We predict the existence of a 'cell-autonomous time-horizon', a time before which cooperation between cell-to-cell communication and DNA mutations might be one of the most fundamental forces shaping the early stages of oncogenesis. The understanding of this process could shed new light on the mechanisms leading to clinically manifest cancers.

Introduction

The cooperation between tumour cells and its environment and the competition between different tumour clones during carcinogenesis are well-established[1]. Other types of cooperations, for instance, the positive cooperation between tumour clones, or even non-transformed clones, have been increasingly recognised as a possible fundamental driving force in cancer as well[2, 3]. The complexity of all possible clonal interactions, particularly during the late stages of cancer, is, therefore, fostering research aimed to model cancer from an ecological perspective[1, 2, 4]. Competition for resources is one of the driving forces for clonal interaction. However, cell-to-cell communication is an equally fundamental mechanism mediating the interaction of cellular populations through shared diffusible or immobile molecules, such as cytokine or metabolites. After early modelling work on angiogenesis[2], the possibility that partially transformed tumour cells might cooperate was generalised by Axelrod and colleagues[3]. Several recent experimental findings are now supporting the notion that cooperation of clones and polyclonality play an important role in the emergence of cancer.

Glioblastoma multiforme tumours, for instance, exhibit considerable intra-tumoral heterogeneity including the pathogenic expression of an oncogenic truncation of the epidermal growth factor receptor (ΔEGFR) gene and EGFR amplification[5]. The less frequent ΔEGFR clones can support an increased fitness of the more prevalent cells overexpressing EGFR, through secretion of IL6 and LIF and a paracrine effect. Recently, Reeves and colleagues[6] have used multi-colour lineage tracing with a Confetti mouse line together with the topical administration of a carcinogen, to study clonal evolution during early oncogenesis. Interestingly, the authors observed benign papillomas harbouring an HRAS Q61L mutation with streaks of Notch mutant clones. Although these Notch mutants were considered infiltrating clones with no active role in the oncogenic process, Janiszewska and Polyak[7] noted that cooperation between the Notch and HRAS mutants could not be excluded and that streaks of Notch clones are reminiscent of structures found in non-mutualistic colonies of budding yeast. Although unproven, it is conceivable that the less frequent clones can provide, altruistically, a fitness advantage to the HRAS mutant cells similarly to what has been observed for glioblastoma multiforme[5, 8] or for WNT-secreting wild-type HRAS clones supporting HRAS mutants[9]. While facilitating the oncogenic process, a non-mutualistic clone would be then outcompeted by more aggressive clones after a clonal sweep and diversification into multiple intermixed mutants[6] suggestive of mutualistic clonal interactions[7].

However, it is unclear if these observations, often obtained using model systems with carcinogens or established tumour clones, can be recapitulated at the low mutational rates occurring naturally[10]. Furthermore, it is unknown at which stage of carcinogenesis, non-cell-autonomous mechanisms might have a role[9]. As cell-to-cell communication and clonal interaction are often neglected in formal models of carcinogenesis, we propose a model for the interaction between the mutagenic process and cell-to-cell communication within a three-dimensional tissue architecture. We developed the simplest possible models to capture the basic emergent properties of early oncogenesis in the presence of mutations and clonal cell-to-cell communication. We propose that the extremely low mutational frequency encountered in physiological conditions does not render cooperation between mutations in adjacent cells unlikely but – rather the opposite – that synergy between the mutational process and cell-to-cell communication might play a fundamental role in carcinogenesis.

Results

A model for mutationally-driven cooperation in oncogenesis. The question addressed in this work is not *if* cooperation between mutant (partially-transformed) cells can occur, but *how likely* or *when* distinct mutations can occur in different cells cohabiting within the same tissue. Therefore, we develop a simple mathematical model to gain insights into answering these fundamental questions. We consider a low mutational rate ρ_0, constant throughout oncogenesis and equal for each possible oncogenic mutation[11]. With oncogenic mutation, we refer to any mutation that at any given time (not necessarily when it occurs) might contribute to the increased fitness of a clone that will eventually evolve into cancer, either through cell-autonomous or non-cell-autonomous mechanisms.

The probability for a single cell to accrue two specific mutations independently within a given time interval is thus p_0^2 (with $p_0 = \rho_0 t << 1$). The probability that two neighbouring cells exhibit one given mutation each independently is, unsurprisingly, the same. Initially, we assume non-dividing cells in a well-organised tissue that after accumulating these two mutations acquire a fitness advantage. We will refer to these cells as initiated or transformed, but we will use these terms very loosely only to indicate a gain in fitness.

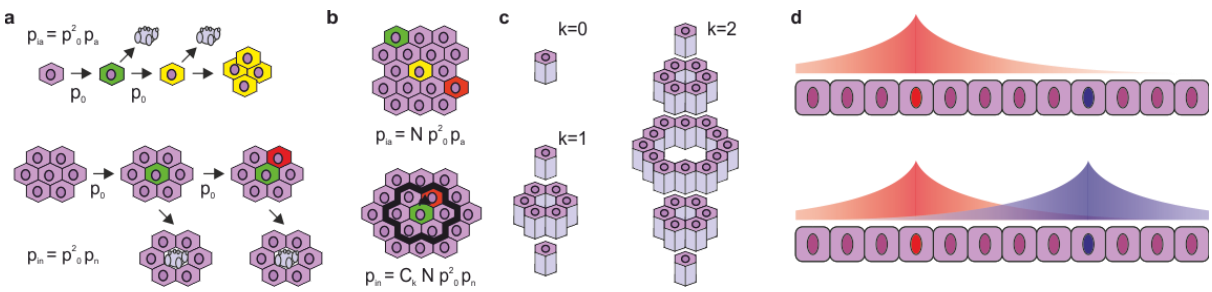


Figure 1 | Tissue organisation and non-cell-autonomous mechanisms. a*) A simple model where cells accumulate two mutations (top) or two mutations occur in different cells within the same neighbourhood (bottom).* ***b****) When the probability to accrue mutations is low, within a tissue of N cells, there will be more opportunities for mutations to co-occur within a given neighbourhood (bottom) rather than within the same cell (top).* ***c****) The neighbourhood of a cell can be described as a problem of geometrical tessellation of space which will depend on tissue organisation, here shown a simple example of hexagonal pillars tesselating space.* ***d****) gradients of shared resources (e.g., growth factors or metabolites) might be then induced by either one or the other cell triggering interactions by juxtacrine or paracrine effects.*

In tissue with N cells, the probability of cell-autonomous initiation of one mutant cell is simply $p_{ia}=Np_0^2p_a$. Similarly, the probability of non-cell-autonomous initiation is $p_{in}=NCp_0^2p_n$. p_a and p_n are defined as the probabilities that one cell harbouring the right pair of mutations – either by itself or within its neighbourhood – survives tumour-suppressive mechanisms (**Fig. 1a**). C is a coordination number, *i.e.* the number of cells within the neighbourhood of a reference cell (**Fig. 1b, c**). Within the validity of common assumptions (*e.g.*, equally probable, spatiotemporally-invariant and independent mutational events), the probability of initiation within a group of N cell is the sum of p_{ia} and p_{in}: $p_i = Np_0^2p_a(1+\Omega_{p_{n0}/p_a})$ with $\Omega = Cp_n/p_{n0}$ and where p_{n0} is the probability that one cell is transformed when directly in contact with another mutant.

p_n (and thus Ω) depends on tissue organization and the type of cell-to-cell cue that contributes to the process of transformation (**Fig. 1d**). With this simple notation, the answer to our central question can be thus separated into the study of tissue organization (the factor Ω) and the magnitude of p_{n0} compared to p_a.

Tissue organisation. To model the organisation of tissue wherein mutated cells are resident, several aspects of tissue organisation have to be considered:

i) the more distant a neighbouring cell is, the lower the probability of cooperative non-cell-autonomous effects should be, *i.e.* p_n shall be a function of distance (*d*);
ii) *C* is the sum of cells in extended neighbourhoods or the sum of C_k, *i.e.*, the number of cells in the k-neighbourhood (at a distance d_k), where k=1 defines cells in contact (*i.e.*, d_1=0);
iii) C_k depends on tissue architecture that we model as a problem of three-dimensional tessellations of space;
iv) tissues are compartmentalised and, therefore, boundaries effects should be considered.

Therefore, in general, the factor Ω can be described as the cumulative effect on the probability of initiation of a reference from each cell within a tissue as consequence of a cell-to-cell communication, $\Omega = p_{n0}^{-1} \sum_k C_k p_{nk}(d_k)$. For convenience, we describe C_k just for two different tissue topologies, a tissue organised in stacked hexagonal pillars or a thin layer of similar hexagonal pillars. In the former case, cells tessellate a three-dimensional space, and we neglect effects at the periphery. In other words, we assume that the number of cells contained within a tissue is larger than the cells at its periphery. **Fig. 1** illustrates the progression of the number of cells included in subsequent neighbourhoods. In **Appendix 1** we demonstrate that $C_k = 6k^2+2$. For a significantly more constrained topology where only three layers exist $C_1 = s_0 + 2$ and $C_{k>1} = s_0(3k-2)$. This description permits us to evaluate analytically the effects of tissue organisation on the probability of cooperation between mutations. In the next section, we provide also numerical examples showing the general validity of this model in the presence of even more stringent topological constrains.

Oncogenic field effect. Without loss of generality, we assume that the interaction between two mutant cells is mediated by a sharing diffusible product[3], for instance a growth factor or a metabolite. Eldar *et al.* have modelled how the concentration of a signalling molecule (a morphogen) secreted by a cell decays in space[12]. Typically, the morphogen concentration is abated by passive diffusion and linear degradation resulting in exponentially decaying concentration gradients. However, ligand-morphogen interactions can induce non-linear mechanisms of morphogen degradation resulting in power-law decays. Therefore, we first analyse the decay of an oncogenic field akin morphogen gradients using power or exponential decays because of their physiological relevance[12-15].

For the case of a power function ($p_n(k) = p_{n0}\, k^{-l}$) and a three-dimensional tissue described by hexagonal pillars ($C_k = 6k^2+2$), the factor Ω can be described analytically with $\Omega(l) = 6\zeta(l-2) + 2\zeta(l)$ (**Appendix 2**). ζ is the Riemmann Zeta function and is finite only for an argument larger than one (here l>3). Therefore, for a large interconnected tissue, oncogenic biochemical gradients induced by a mutant cell must decay very steeply for non-cell-autonomous mechanisms not to dominate. In the limiting case where only the 1-neighbourhood is relevant for transformation ($l \to \infty$), the Riemann Zeta function converges to unity and therefore $\Omega = 8$. This is just the number of cells in direct contact to the reference cell (C_1) showing mathematical consistency and providing a lower boundary to Ω in the case of small effects in a very constrained topology. Conversely, for shallower gradients where the Riemann Zeta function does not converge (l<4), these probabilities will be significantly larger. We obtained these results modelling tissues of non-finite extensions to derive analytical solutions. However, through numerical estimations, it is simple to demonstrate how these observations are generally valid and correct also for small volumes of cells (**Fig. 2a-b**). For example, in a small neighbourhood with a radius of 10 cells, $\Omega \sim 11.5$ (l=4) and $\Omega \sim 340$ (l=1),

values (**Fig. 2a**, solid circles) that reach 12 and 3 10^4, respectively, for a neighbourhood with a radius of 100 cells (**Fig. 2a**, empty circles). Similarly, we demonstrate that for a thin three-layer tissue, $\Omega = 12\zeta(l-1) - 8\zeta(l) + 4/3$ (see **Appendix 2**). This series converge for l>2, it assume a value of 11.7 for l=3, and numerical estimations show that Ω reaches values of ~24 and ~50 for l=2 within a limiting neighbourhood with a 10 or 100 cell radius, respectively (**Fig. 2a**, empty and solid circles). In the limit case where only the first neighbourhood is relevant ($l \to \infty$), Ω ~ 5.3. Therefore, even within this rather constrained topology, Ω obtains rather large values.

Gradients described by power functions are shallower than exponentially decaying gradients at longer distances. Although both gradients are physiologically relevant, power-like functions might overestimate Ω. The formalism for exponentially decaying oncogenic fields is less elegant (see **Appendix 3**); however, it can be readily demonstrated that even for steep gradients decaying of a third at every cell distance (k_c=1), Ω can assume double-digit values (see also **Fig. 2b**).

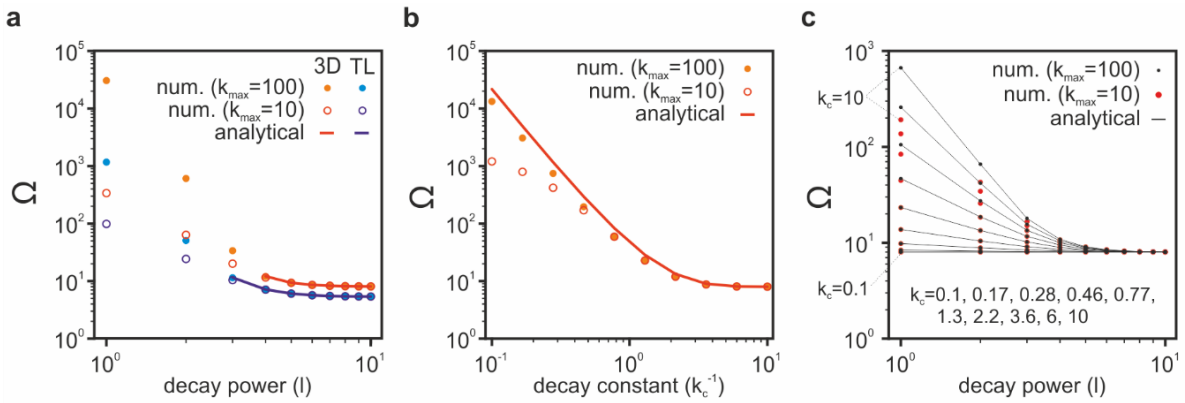


Figure 2. Numerical validation of the oncogenic field solutions. a*) Comparison between numerical and analytical solutions to estimate the value of the oncogenic field factor Ω for a three-dimensional (3D) and three-layered (TL) tissue model. The solid lines represent the analytical solutions within the limits of its convergence (l>2 for TL in blue and l>3 for 3D in red). The numerical estimations, computed over a small 10 (empty circles) and larger 100 (solid circles) cells radius neighbourhood, confirms the observations we reported, more generally, on the analytical solutions, including the necessity of steep power decays and a convergence to Ω~8 (3D) and Ω~5.3 (TL) for these limiting cases.* ***b****) Identical computations described in a) but for the exponential decay model for a three-dimensional tissue. Both the analytical solution and numerical estimates converge to the value of Ω~8 for steep decays.* ***c****) Values for Ω computed for a general case where the oncogenic field decays jointly as the inverse of a power-law and exponentially. The analytical and numerical solutions are shown for the same parameter sweep shown in a) and b), i.e. with the inverse power from 1 to 10 and with a decay constant k_c from 0.1 to 10.*

The analysis of power-law and exponential decays are rather instructive, and they are often used to model morphogen gradients as a solution to the reaction-diffusion equation for one-dimensional problems and for specific three-dimensional architectures. We can also demonstrate (**Appendix 4**) that for $p_{nk} = p_{n0}k^{-l}e^{-(k-1)k_c^{-1}}$, *i.e.*, when the oncogenic gradients jointly decays as an inverse power-law and exponentially, $\Omega = e^{k_c^{-1}}\left[2Li_l\left(e^{-k_c^{-1}}\right) + s_0 Li_{l-2}\left(e^{-k_c^{-1}}\right)\right]$, where *Li* is the polylogarithm. This analytical solution describes the expectation for Ω for an oncogenic field induced by stochastic (mutationally-driven) point-sources of shared resources in an ideal three-dimensional tissue (see **Fig. 2c**) in the presence of degradation. Once again, at the limit for a fast decaying concentration gradient, the value of Ω~8, long-distance interactions (k_c>>0) can drastically increase the magnitude of Ω and with high values found also for small clusters of cells (**Fig. 2c,** blue and red circles).

We can thus infer a general consideration from the mathematical description of the proposed case studies that are aimed to exemplify the possible synergy between the mutational process

and non-cell-autonomous effects. Unsurprisingly, the specific tissue geometries and the properties of concentration gradients results in rather different magnitudes of an 'oncogenic field'. However, either through *juxtacrine* (contact-dependent) or *paracrine* (short- or long-distance) signalling, mutations in tissue neighbourhoods that can cooperate through cell-to-cell communication are likely to have a significant role in oncogenesis, in addition to mutations co-occurring within a cell.

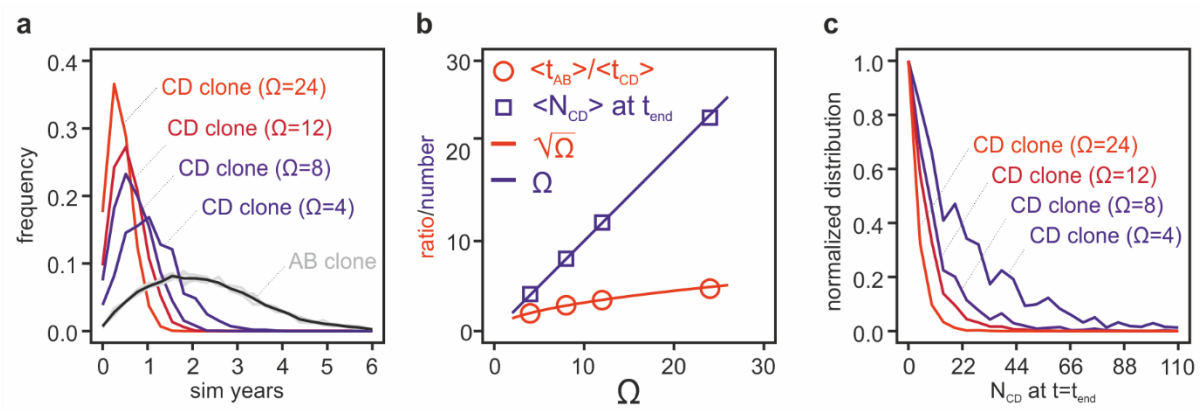


Figure 3. Monte Carlo simulations of the cell-autonomous time-horizon. a*) Probability distribution of the waiting-times for the occurrence of the first two co-occurring mutations (AB-clones, grey and black curves) or for the first cooperating mutations (CD-clones, coloured curves) through non-cell-autonomous mechanisms for Ω values equal to 4, 8, 12 and 24. The grey curves are the four independent simulation runs related to these four conditions and are identical except for noise; the black curve is the average of these four runs and can be used as reference.* ***b****) The average time (<t_{CD}>) at which the first cooperating CD-clone is observed scales with the square root of Ω (red line) compared to the average time (<t_{AB}>) at which the first AB-mutant appears. The average number of cooperating mutations within a neighbourhood at t=t_{end} (<N_{CD}>) scales as Ω (blue curve).* ***c****) Distribution (normalized to maximum for better visualization) of the number of CD-clones at the end of the simulations (t=t_{end}). t_{end} is the time at which at least one AB- and one CD- clone are detected.*

Cell-autonomous time-horizon. So far, we have discussed *if* and *how likely* mutationally-driven non-cell-autonomous mechanisms might be; next, we address the question about *when* these mechanisms are more likely to occur. Indeed, we have shown that non-cell-autonomous mechanisms can increase the probability that mutations contribute to carcinogenesis by a factor Ω. A corollary to this observation is that cooperation between non-transformed cells might contribute to tumour initiation earlier than cell-autonomous mechanisms. For simplicity, we consider only the mutational process and neglect p_a, p_n and p_{n0}. One cell accrues pairs of mutations at the rate ρ_0 but within a neighbourhood cooperating cells accrue mutations at an apparent rate of $\rho_0\sqrt{\Omega}$. We tested this simple mathematical inference with Monte Carlo simulations (see **Fig. 3** and **Methods**). We simulate the independent and stochastic appearance of four types of mutations (A, B, C and D) at a rate of $\rho_0=10^{-6}$ mutations/day on a lattice of $N=10^6$ cells with 2,000 replicates. When one cell accrues mutations A and B, it is flagged as an AB mutant; when a cell becomes a C mutant and in its neighbourhood there is a D mutant, the C mutant will be listed as a CD (cooperating) clone. The average time for a double-mutant cell to appear (<t_{AB}>) is ~29 months (**Fig. 3a**). At the net of noise, the distribution of <t_{AB}> values (**Fig. 3a**, grey curves) depends only on ρ_0 and N but not on Ω and, therefore, we overlay the average of the four distributions of <t_{AB}> values (**Fig. 3a**, black curve) as reference for the cooperating mutants. The distribution of the waiting times for the appearance of CD clones depends on the value of Ω (**Fig. 3a**, coloured curve) and exhibits an average time (<t_{CD}>) of about 14.7, 10.1, 8.5 and 6.2 months for Ω values equal to 4, 8, 12 and 24, respectively – scaling as $\Omega^{-0.5}$. Thus, we can define t_a (<t_{AB}> in our simulations) as the average time for a tissue of N cells to accrue two mutations which is inversely proportional to ρ_0N. By definition, t_a - the time-horizon after which cell-autonomous mechanisms might

dominate - is preceded by a latency period during which single mutations are more likely. However, our model predicts the existence of a significantly long period $t_\Omega = t_a \Omega^{-0.5} \leq t < t_a$ when mutationally-driven cooperation between adjacent cells is more likely than mutationally-driven cell-autonomous mechanisms to occur. For instance, in the limiting case where only the first-neighbourhood significantly contribute to tumour initiation ($\Omega = 8$), during ~65% of the time interval preceding t_a, clonal cooperation is likely to be a fundamental mechanism that synergizes with the mutational process to support partially transformed clones. For Ω values of 4, 12 and 24, this interval will be about 50%, 70% and 80% of t_a. The scaling of t_Ω as $t_a \Omega^{-0.5}$ is shown in **Fig. 3b** from the plot of $<t_{AB}>/<t_{CD}>$ (**Fig. 3b**, red) and the scaling of the number of mutation in a given neighbourhood defined by Ω at t_a is shown is $<N_{CD}>$ (**Fig. 3b**, blue). Because of the stochastic nature of the mutational process, the distribution of waiting times for double mutants (AB and CD) are broad. This heterogeneity result in no CD co-operating clones in about 40%, 20%, 15% and 8% of the simulation runs with Ω = 4, 8, 12 and 24, respectively. The majority of the cases, however, co-operating mutants preceded AB-clones and for larger values of Ω the probability for mutations to appear within a tissue neighbourhood is so high that CD mutants are likely to reoccur multiple times randomly (**Fig. 3c**). Therefore, even assuming a role for synergy between the mutational process and cell-to-cell communications during the earliest steps in oncogenesis ($t < t_a$), chance will determine the first occurrence of co- or cooperating- mutations (as a function of Ω), possibly influencing the evolutionary trajectory of a tumour and contributing to tumour heterogeneity.

Discussion

A role for non-cell-autonomous mechanisms in cancer is well-established, often as a mechanism of interaction between cancer cells and the surrounding tissue[2, 16-19]. The cooperation of non- or partially- transformed clones as a driving force underlying oncogenesis has also been hypothesized[3], and there is nowadays accumulating evidence suggesting that a description of oncogenesis focused exclusively on cell-autonomous mechanisms might under-represent the importance of oncogenic signalling in cancer[5, 9, 19].

Experiments in *Drosophila melanogaster* have also shown that inter-clonal cooperation between mutants harbouring an oncogenic KRAS mutation or inactivation of the tumour suppressor *scrib* can support tumorigenesis mediated by JNK and JAK/STAT signalling[20]. Recently, Marusyk and colleagues (2014) have used a mouse xenograft model to test the effects of clonal heterogeneity demonstrating that clones expressing the chemokines IL11 are capable of stimulating overall tumour growth through a non-cell-autonomous mechanism, while clonal interference maintains genetic intra-tumour heterogeneity[9]. Similarly, Inda and colleagues (2010) have shown how intra-tumour heterogeneity observed in glioblastoma can be maintained through cross-talk between mutants harbouring a ΔEGFR that secrete IL6 and LIF to support fitness in clones with EGFR amplification[5]. Clearly and colleagues (2014) has also shown that WNT-producing HRAS wild-type clones can support tumorigenicity and clonal heterogeneity by cooperating with clones harbouring mutant oncogenic HRAS[21]. These observations support the emerging notion that intra-tumoral heterogeneity is often of polyclonal origin and is an active process supported by non-cell-autonomous mechanisms. However, the role for poly-clonality and clonal cooperation during earlier stages of oncogenesis can be seen in contradiction with the low estimates of mutational rates in cancer[10]. Furthermore, it is unclear if clonal cooperation has a role during early oncogenesis or only at later stages when a heterogeneous tumour is established[21].

Aiming to contribute to filling this gap in knowledge, we developed simple analytical and computational models for mutation-driven oncogenesis in the presence of cell-to-cell communication. Using similar assumptions used to model mutationally-driven oncogenesis, we have studied *if*, *how* and *when* is likely that cell-to-cell communication might cooperate with the mutational process from the perspective of basic principles. Our analysis raises provoking observations on the earliest steps in oncogenesis. We show that irrespectively of the background mutational rate if a set of transforming mutations are sufficiently likely to occur within a single cell in the lifetime of a patient, an equally rare yet oncogenic set of mutations are equally (or more) likely to contribute to tumorigenesis through non-cell-autonomous mechanisms. We have introduced the parameter Ω which capture the impact of tissue organisation and non-cell-autonomous mechanisms on cancer evolution. We modelled non-cell-autonomous mechanisms in analogy to morphogens during embryonic developments. Ω describes the magnitude with which paracrine, juxtacrine and other mechanisms mediated by shared substrates (*e.g.*, growth factors and metabolites) might impact the transformation of a cell or clone. As such, Ω represents an oncogenic field effect, where oncogenic fields have the opposite outcome of morphogens by contributing to the de-regulation of tissue homeostasis. Furthermore, we have identified a stage of oncogenesis during which clonal cooperation might not simply coexist with clonal competition but even dominate before the emergence of clones capable of growing autonomously. With the help of our model, experimentally, the problem is reduced to the measurement of quantities such as p_{n0} and p_a or the abundance of genes that, once mutated, can drive oncogenesis by non-cell-autonomous mechanisms. We argue that the magnitude of the oncogenic field effect (Ω) and the prediction of an autonomous time-horizon suggest a significant role for mutationally-driven and non-cell-autonomous mediated poly-clonal evolution of cancer during, at least, a very early stage of oncogenesis.

The model described here is purposely simple aiming to illustrate the basic principles emerging by the cooperation of the mutational process with non-cell-autonomous mechanisms[3], a phenomenon that, to our knowledge, is often neglected when models of somatic evolution of cancer are studied analytically[22]. For this reason, we did not include the description of more complex and important features of real tissues such as clonal dynamics, tissue homeostasis, tissue mechanics, and other mechanisms for gradient formation of biomolecules. Each of these processes can change considerably the magnitude of the effects we described. The concentration gradients on their own, for instance, can be enhanced by compartmentalisation, abrogated by diffusion into lumens or the vascular system, or affected by systemic alterations of shared resources (*e.g.*, hormones, lipids). If a proliferative tissue is considered, with a fitness advantage for cooperative clones compared to wild-type cells, the presence of these non- or partially- transformed clones could be even more significant, increasing the probability to accrue further mutations at a faster pace and shaping the initial period of oncogenesis.

However, tissues are complex systems and diverse mechanisms of tissue homeostasis in different tissues might conflict with this perspective. In the case of a fast self-renewing tissue like the intestinal epithelium within which the cell-of-origin for common tumours is likely to be a stem cell[23], the highly compartmentalized stem-cell niche might pose an effective barrier to oncogenic field effects. The intestinal epithelium is one of the most proliferative tissues subject to a high mutagenic burden and it has been broadly studied both mathematically[24, 25] and experimentally[23, 26, 27]. A small group of adult stem cells reside within the colonic crypt and maintain the homeostasis of the villi lining the intestine[23]. Within the crypt and the villus, the balance between proliferation and differentiation is maintained by a complex network of signals (*e.g.*, WNT, NOTCH, BMP, and EGF) generated by specialized Paneth cells within the crypt and cells within mesenchyme lining the crypt[27]. Mutations in the WNT (*e.g.*, APC or

CTNNB), EGFR (*e.g.*, KRAS, PIK3C or BRAF) and TGF-β (*e.g.*, SMAD4) signalling pathways gradually render cells independent from niche signals to grow autonomously and promoting cancer. While it is still more likely that two mutations are acquired within adjacent stem cells rather than within one cell, neutral genetic drift or selection fix or purge genetic mutations within the crypt that will be thus, most of the times, monoclonal[23]. However, not all mutations occurs in the stem cells within the crypt[28] and tissue homeostasis is maintained by crypt fusion and fission leading to field cancerization and the possibility that partially- or non- transformed clones might interacts not within a monoclonal crypt but between patches harbouring different mutations[23, 29]. These considerations might hold true also for other highly proliferative tissues, such as the well-described human epidermis and the oesophageal epithelium[30-33]. In these tissues homeostasis is maintained by a balance between the probability for progenitor cells to divide symmetrically or asymmetrically, giving birth to two progenitor cells, two differentiating cells, or – more commonly – one differentiating and one progenitor cell[30, 33]. Alcolea and colleagues, for example, have shown that mutations in the NOTCH pathway reduce the probability for a progenitor cell to generate two differentiating cells and induce wild-type cells to differentiate[30]. In combination with P53 mutations, this cell fate imbalance leads to field cancerization[30]. Considering the experimental observations on later stages of carcinogenesis we have already discussed[5, 6, 8, 9, 20, 21] and our results, it is conceivable that mutations might also cause cell fate imbalance through non-cell-autonomous mechanisms during early oncogenesis. As cell fate determination and the occurrence of mutations are stochastic processes, the role of non-cell-autonomous mechanisms not only might vary across different tissues depending on their organization but also within the same tissue of origin. We described in Fig. 3 that the occurrence of a double-mutant clone that might acquire a fitness advantage autonomously, even if less likely, can still precede the occurrence of two cooperating single-mutants. Similarly, genetic drift and selective pressure could either maintain or collapse one or both of the cooperating populations[9]. Therefore, each tissue and each tumour might be affected differently by non-cell-autonomous mechanisms, mechanisms that could alter the evolutionary trajectory of tumours that later acquire full independence from cooperating clones thus contributing also to tumour heterogeneity. As computational modelling of multicellular tissues can describe complex homotypic and heterotypic interactions, including short- and long- range interactions, and tissue mechanics[34-36], computational models rather than analytical tools might be more appropriate to investigate the possible role for 'oncogenic fields' in complex mutagenic environments.

While the mathematical analysis we presented was not elaborated to capture more complex phenomena occurring during oncogenesis, our model highlights the importance of identifying the genes and the shared resources that can mediate clonal cooperation, such as growth factors (*e.g.*, mitogens, interleukins, *etc.*) or even metabolic by-products that are often at the basis of cooperative behaviour in lower organisms[37-39]. The theory described here was aimed to be simple and, at the best of our knowledge, it is a first attempt to describe the problem with an explicit mathematical model to extend the existing models of oncogenesis[11, 22, 40-43] to non-autonomous mechanisms. Indeed, we emphasise that our model is not in contradiction with the prevailing models of oncogenesis, as it is based on similar assumptions, but it highlights an equally important role for tissue organisation and cell-to-cell communication that cooperate synergistically with a mutationally driven process.

Methods

The mathematical demonstrations of the analytical equations presented in this work are described in **Appendices 1-3**.

The numerical evaluation of these analytical results (**Fig. 2a-c**) was performed with the Matlab script '*analytical_and_numerical_comparisons_v4'* (Mathworks, version 2018) freely available from the GitHub repository *alesposito/CloE-PE*. The numerical estimations simply compare the value obtained from the approximated analytical solutions described in the appendices to direct numerical estimate computed on given neighbourhoods with features described in the main text. For the case of a three-dimensional tissue (**Appendix 2**) the values of the analytical solution $\Omega = s_0\zeta(l-2) + 2\zeta(l)$ were compared to those of the finite series $\Omega = \sum_{k=1}^{kn}(k^2 s_0 + 2)k^{-l}$ for different power functions. Data shown in **Fig. 2a** are computed with *kn* equal to 50 and 100, *l* ranging from 1 to 10, and s_0 = 6. Within the same parameter space, we compare the analytical description of a three-layer thin tissue $\Omega = 12\zeta(l-1) - 8\zeta(l) + 4/3$ to the numerical estimates of $\Omega = 2 + \sum_{k=1}^{kn} s_0(3k-2)k^{-l}$. Similarly, **Fig. 2b** shows a comparison between the analytical description $\Omega = e^{k_c^{-1}}\left[(2+s_0)e^{2k_c^{-1}} + (4+s_0)e^{k_c^{-1}} + 2\right]\left[e^{k_c^{-1}} - 1\right]^{-3}$ of a three-dimensional tissue with an oncogenic field decaying as an exponential function (**Appendix 3**) and the finite series $\Omega = \sum_{k=1}^{kn}(k^2 s_0 + 2)e^{-(k-1)k_c^{-1}}$. Also in this case, the parameters used in the numerical evaluations were *kn* = 50 and 100, and s_0 = 6 with the inverse of the decay constant k_c^{-1} spanning the 0.1 to 10 range. Last, for the distribution jointly decaying as a power and exponential function (**Appendix 4** and **Fig. 2c**), numerical estimates of $\Omega = \sum_{k=1}^{kn}(k^2 s_0 + 2)k^{-l}e^{-(k-1)k_c^{-1}}$ were compared with the analytical solution $\Omega = e^{k_c^{-1}}\left[2Li_l\left(e^{-k_c^{-1}}\right) + s_0 Li_{l-2}\left(e^{-k_c^{-1}}\right)\right]$ on the same parameter space described for the other cases.

Thee Monte Carlo simulations (**Fig. 3**) used to evaluate the relationship between the time-horizon for cell-autonomous mechanisms (t_a) and non-cell-autonomous mechanisms are available as the Matlab script '*polyclonal_mutation_cooccurence_check_v4'* freely available from the GitHub repository *alesposito/CloE-PE*. We simulated a lattice of 10^6 cells with a mutational rate equal to 10^{-6} mutations per cell per simulated day (simday). At each simday and at each node of the lattice, four random numbers (n_m, with m=A, B, C or D) were drawn from uniformly distributed numbers in the [0,1] interval. For any of the indexes A, B, C, or D where n_m was lower than or equal to 10^{-6}, the correspondent cell was switched from non-mutant to mutant. Cells were then allowed to accumulate these four mutations for a maximum of 100,000 simdays. When a cell acquires both A and B mutations, an AB-mutant cell is established and logged. When a D mutation appears in a neighbourhood of a C-mutant, a CD-cooperative clone is logged. CD-cooperative mutants are detected utilizing convolution filters that detects the co-occurrence of a D mutant within the centre of a reference neighbourhood and C-mutant in its immediate vicinity. The publicly available code implements the following neighbourhoods scans equivalent to Ω value of 4, 8 and 12. For Ω = 4, detection in position North (N), East (E), South (S) and West (W); for Ω = 8, as for the previous case but with the addition of NE, NW, SE and SW; for Ω = 12, as for the previous case but with the addition of one non-adjacent cell in N, E, S and W position. As soon as at least one AB- and one CD-cooperating clone occurs, the simulation is interrupted. Simulations are then repeated 2,000 times and the distributions of the appearance of first AB- or CD- clones, and number of CD-clones at the appearance of an AB- clone are generated.

The scripts were run on a Dell Precision 5810 workstation utilizing an Intel Xeon E5-1625 CPU and 64GB RAM. The four Monte Carlo simulations shown in Fig. 3 are computationally intensive and were ran in parallel for about two days.

Acknowledgements

A.E. acknowledges the financial support provided from Medical Research Council program grants (MC_UU_12022/1 and MC_UU_12022/8) awarded to Prof. Ashok Venkitaraman.

Supporting information

Appendix 1 – Description of tissue organization. The 1-neighbourhood of an individual cell will contain 6 adjacent cells within the plane and 2 polar cells, one at the top and one at the bottom of the reference cell. The 2-neighbourhood will contain 12 cells within the plane, 6 cells in the above and bottom layers and two polar cells. The 3-neighbourhood will have 18 cells within the plane, 24, 12, 6 and 2 in the other layers. By induction, we infer that the k-neighbourhood of an individual cell in such tessellation is made of ks_0 cells within the plane, the sum of is_0 for each of the above and below layers, with i from 1 to k-1, and the two polar cells:

$$C_k = ks_0 + 2\sum_{i=1}^{k-1} is_0 + 2$$

To simplify, ks_0 (s_0 here is 6) can be brought within the sum, then allowing to simplify to the correspondent triangular number:

$$C_k = -ks_0 + 2\sum_{i=1}^{k} is_0 + 2 = k^2 s_0 + 2$$

It should be noted that for intercalated layers, the coordination values can be larger.

Another example, this time including a significant constraint in topology, is represented by the same topology, where only three layers are considered.

$$C_1 = s_0 + 2$$

$$C_{k>1} = ks_0 + 2(k-1)s_0 = s_0(3k-2)$$

Appendix 2 – Probability of initiation (power function). Let's assume the probability of tumour initiation is proportional to a concentration gradient, similar to a morphogen, or an oncogenic mitogen/morphogen field decaying as a power function:

$$p_{nk} = \frac{p_{n0}}{k^l}$$

Where p_{n0} indicate the probability of tumour initiation when cells are attached (the 1-neighbourhood). Therefore, in the case of 3D hexagonal tessellation, we can derive the factor Cp_n:

$$\Omega = \sum_{k=1}^{\infty} \frac{k^2 s_0 + 2}{k^l}$$

This sum is carried over an infinite neighbourhood, and the validity of the results will be checked numerically. First, we can expand p_{tn}:

$$\Omega = s_0 \sum_{k=1}^{\infty} k^{2-l} + 2 \sum_{k=1}^{\infty} k^{-l}$$

These series can now be described by Riemann Zeta functions:

$$\Omega = s_0 \zeta(l-2) + 2\zeta(l)$$

Let's now consider the thin 3-layer tissue which tessellation was already discussed. In this case, for one cell:

$$\Omega = s_0 + 2 + \sum_{k=2}^{\infty} \frac{s_0(3k-2)}{k^l} = 2 + \sum_{k=1}^{\infty} \frac{s_0(3k-2)}{k^l}$$

Following the same process described before, we can obtain:

$$\Omega = 2 - 2s_0 \sum_{k=1}^{\infty} \frac{1}{k^l} + 3s_0 \sum_{k=1}^{\infty} \frac{1}{k^{l-1}}$$

And,

$$\Omega = 18\zeta(l-1) - 12\zeta(l) + 2$$

This describe the probability of transformation for a cell in the middle layer. We can approximate the result over the tissue equal to this value by N/3 (middle layer) and with half contribution for the top and bottom layer resulting in

$$\Omega = 12\zeta(l-1) - 8\zeta(l) + 4/3$$

Appendix 3 – Probability of initiation (exponential function). Let's now assume the oncogenic field decays as an exponential function:

$$p_{nk} = p_{n0} e^{-(k-1)k_c^{-1}}$$

Where k_c is a decay constant expressed in terms of k-neighbourhood for simplicity. If two cells are in contact, the probability of initiation will be p_{n0} as per definition of p_{n0}. When cells are at a k_c+1 distance, this probability is 1/e lower, i.e. ~30% lower. In the case of 3D hexagonal tessellation, the factor Cp_n can be now expressed as:

$$\Omega = \sum_{k=1}^{\infty} (k^2 s_0 + 2) e^{-(k-1)k_c^{-1}}$$

Or the sum of the series:

$$\Omega = 2 \sum_{k=1}^{\infty} e^{-(k-1)k_c^{-1}} + s_0 \sum_{k=1}^{\infty} k^2 e^{-(k-1)k_c^{-1}}$$

The first series converges to:

$$\sum_{k=1}^{\infty} e^{-(k-1)k_c^{-1}} = \frac{e^{k_c^{-1}}}{e^{k_c^{-1}} - 1}$$

The second series can be represented as:

$$\sum_{k=1}^{\infty} k^2 e^{-(k-1)k_c^{-1}} = e^{k_c^{-1}} \sum_{k=1}^{\infty} k^2 e^{-k/k_c} = \frac{e^{2k_c^{-1}}\left(e^{k_c^{-1}}+1\right)}{\left[e^{k_c^{-1}}-1\right]^3}$$

Therefore,

$$\Omega = e^{k_c^{-1}} \frac{(2+s_0)e^{2k_c^{-1}} + (4+s_0)e^{k_c^{-1}} + 2}{\left[e^{k_c^{-1}}-1\right]^3}$$

With $s_0 = 6$, once again to confirm mathematical consistency, $\lim_{k_c \to 0} \Omega = 8$, as in the case where only adjacent cells are important. Shallower decays will again increase this value (see Fig. 2).

Appendix 4 – Probability of initiation (generalization). We have characterized the oncogenic field in relation to typical descriptions of morphogenic gradients[15]. While relevant for specific cases, steady-state concentration gradients of shared resources in space, generated by passive diffusion and linear or non-linear degradation, can adopt different shapes. One useful analytical description is represented by concentrations that decay as the product of exponential and power-law functions, for instance as:

$$p_{nk} = \frac{p_{n0}}{k^l} e^{-(k-1)k_c^{-1}}$$

With the same formalism and strategies described in **Appendices 2** and **3** we can show that, for a three-dimensional tissue:

$$\Omega = \sum_{k=1}^{\infty} (k^2 s_0 + 2) k^{-l} e^{-(k-1)k_c^{-1}}$$

This analytical representation of Ω can be expressed as sums of polylogarithm functions:

$$\Omega = e^{k_c^{-1}}\left[2Li_l\left(e^{-k_c^{-1}}\right) + s_0 Li_{l-2}\left(e^{-k_c^{-1}}\right)\right]$$

This representation converges to those shown in **Appendix 2** and **3** in the cases where k_c is very large or where l is very small, respectively, *i.e.* when the power-law or the exponential decay components are negligible. The case l=1 represents an oncogenic field induced by continuous point-sources in an unconstrained three-dimensional space in the presence of linear degradation. In this geometry:

$$\Omega = e^{k_c^{-1}}\left[-2\log\left(1-e^{-k_c^{-1}}\right) + s_0 \frac{e^{-k_c^{-1}}}{\left(1-e^{-k_c^{-1}}\right)^2}\right]$$

References

1. Merlo, L.M.F., Pepper, J.W., Reid, B.J. & Maley, C.C. Cancer as an evolutionary and ecological process. *Nature Reviews Cancer* **6**, 924-935 (2006).
2. Tomlinson, I.P. & Bodmer, W.F. Modelling the consequences of interactions between tumour cells. *Br J Cancer* **75**, 157-160 (1997).
3. Axelrod, R., Axelrod, D.E. & Pienta, K.J. Evolution of cooperation among tumor cells. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 13474-13479 (2006).
4. Gatenby, R.A. & Vincent, T.L. Application of quantitative models from population biology and evolutionary game theory to tumor therapeutic strategies. *Mol Cancer Ther* **2**, 919-927 (2003).
5. Inda, M.D. *et al.* Tumor heterogeneity is an active process maintained by a mutant EGFR-induced cytokine circuit in glioblastoma. *Genes & Development* **24**, 1731-1745 (2010).
6. Reeves, M.Q., Kandyba, E., Harris, S., Del Rosario, R. & Balmain, A. Multicolour lineage tracing reveals clonal dynamics of squamous carcinoma evolution from initiation to metastasis. *Nat Cell Biol* **20**, 699-709 (2018).
7. Janiszewska, M. & Polyak, K. A confetti trail of tumour evolution. *Nat Cell Biol* **20**, 639-641 (2018).
8. Bonavia, R., Inda, M.D., Cavenee, W.K. & Furnari, F.B. Heterogeneity Maintenance in Glioblastoma: A Social Network. *Cancer Research* **71**, 4055-4060 (2011).
9. Marusyk, A. *et al.* Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. *Nature* **514**, 54-58 (2014).
10. Parsons, B.L. Many different tumor types have polyclonal tumor origin: evidence and implications. *Mutat Res* **659**, 232-247 (2008).
11. Beerenwinkel, N. *et al.* Genetic progression and the waiting time to cancer. *PLoS Comput Biol* **3**, e225 (2007).
12. Eldar, A., Rosin, D., Shilo, B.Z. & Barkai, N. Self-enhanced ligand degradation underlies robustness of morphogen gradients. *Dev Cell* **5**, 635-646 (2003).
13. Ibanes, M. & Izpisua Belmonte, J.C. Theoretical and experimental approaches to understand morphogen gradients. *Mol Syst Biol* **4**, 176 (2008).
14. Melen, G.J., Levy, S., Barkai, N. & Shilo, B.Z. Threshold responses to morphogen gradients by zero-order ultrasensitivity. *Mol Syst Biol* **1**, 2005 0028 (2005).
15. Wartlick, O., Kicheva, A. & Gonzalez-Gaitan, M. Morphogen gradient formation. *Cold Spring Harb Perspect Biol* **1**, a001255 (2009).
16. Uhlirova, M., Jasper, H. & Bohmann, D. Non-cell-autonomous induction of tissue overgrowth by JNK/Ras cooperation in a Drosophila tumor model. *Proc Natl Acad Sci U S A* **102**, 13123-13128 (2005).
17. Hanahan, D. & Folkman, J. Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis. *Cell* **86**, 353-364 (1996).
18. Riedel, A., Shorthouse, D., Haas, L., Hall, B.A. & Shields, J. Tumor-induced stromal reprogramming drives lymph node transformation. *Nat Immunol* **17**, 1118-1127 (2016).
19. Tape, C.J. *et al.* Oncogenic KRAS Regulates Tumor Cell Signaling via Stromal Reciprocation. *Cell* **165**, 910-920 (2016).
20. Wu, M., Pastor-Pareja, J.C. & Xu, T. Interaction between Ras(V12) and scribbled clones induces tumour growth and invasion. *Nature* **463**, 545-548 (2010).
21. Cleary, A.S., Leonard, T.L., Gestl, S.A. & Gunther, E.J. Tumour cell heterogeneity maintained by cooperating subclones in Wnt-driven mammary cancers. *Nature* **508**, 113-117 (2014).
22. Beerenwinkel, N., Schwarz, R.F., Gerstung, M. & Markowetz, F. Cancer evolution: mathematical models and computational inference. *Syst Biol* **64**, e1-25 (2015).
23. Nicholson, A.M. *et al.* Fixation and Spread of Somatic Mutations in Adult Human Colonic Epithelium. *Cell Stem Cell* **22**, 909-918 e908 (2018).

24. Zhao, R. & Michor, F. Patterns of proliferative activity in the colonic crypt determine crypt stability and rates of somatic evolution. *PLoS Comput Biol* **9**, e1003082 (2013).
25. Nowak, M.A., Michor, F. & Iwasa, Y. The linear process of somatic evolution. *Proc Natl Acad Sci U S A* **100**, 14966-14969 (2003).
26. Farin, H.F. *et al.* Visualization of a short-range Wnt gradient in the intestinal stem-cell niche. *Nature* **530**, 340-343 (2016).
27. Spit, M., Koo, B.K. & Maurice, M.M. Tales from the crypt: intestinal niche signals in tissue renewal, plasticity and cancer. *Open Biol* **8** (2018).
28. Schwitalla, S. *et al.* Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties. *Cell* **152**, 25-38 (2013).
29. Snippert, H.J., Schepers, A.G., van Es, J.H., Simons, B.D. & Clevers, H. Biased competition between Lgr5 intestinal stem cells driven by oncogenic mutation induces clonal expansion. *EMBO reports* **15**, 62-69 (2014).
30. Alcolea, M.P. *et al.* Differentiation imbalance in single oesophageal progenitor cells causes clonal immortalization and field change. *Nat Cell Biol* **16**, 615-622 (2014).
31. Doupe, D.P. *et al.* A single progenitor population switches behavior to maintain and repair esophageal epithelium. *Science* **337**, 1091-1093 (2012).
32. Mascre, G. *et al.* Distinct contribution of stem and progenitor cells to epidermal maintenance. *Nature* **489**, 257-262 (2012).
33. Klein, A.M., Brash, D.E., Jones, P.H. & Simons, B.D. Stochastic fate of p53-mutant epidermal progenitor cells is tilted toward proliferation by UV B during preneoplasia. *Proc Natl Acad Sci U S A* **107**, 270-275 (2010).
34. Farhadifar, R., Roper, J.C., Aigouy, B., Eaton, S. & Julicher, F. The influence of cell mechanics, cell-cell interactions, and proliferation on epithelial packing. *Curr Biol* **17**, 2095-2104 (2007).
35. Ranft, J. *et al.* Fluidization of tissues by cell division and apoptosis. *Proc Natl Acad Sci U S A* **107**, 20863-20868 (2010).
36. Osborne, J.M., Fletcher, A.G., Pitt-Francis, J.M., Maini, P.K. & Gavaghan, D.J. Comparing individual-based approaches to modelling the self-organization of multicellular tissues. *PLoS Comput Biol* **13**, e1005387 (2017).
37. Ponomarova, O. & Patil, K.R. Metabolic interactions in microbial communities: untangling the Gordian knot. *Curr Opin Microbiol* **27**, 37-44 (2015).
38. D'Souza, G. *et al.* Ecology and evolution of metabolic cross-feeding interactions in bacteria. *Nat Prod Rep* **35**, 455-488 (2018).
39. Momeni, B., Waite, A.J. & Shou, W. Spatial self-organization favors heterotypic cooperation over cheating. *eLife* **2**, e00960 (2013).
40. Nowell, P.C. The clonal evolution of tumor cell populations. *Science* **194**, 23-28 (1976).
41. Knudson, A.G., Jr. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A* **68**, 820-823 (1971).
42. Braakhuis, B.J., Tabor, M.P., Kummer, J.A., Leemans, C.R. & Brakenhoff, R.H. A genetic explanation of Slaughter's concept of field cancerization: evidence and clinical implications. *Cancer Res* **63**, 1727-1730 (2003).
43. Curtius, K., Wright, N.A. & Graham, T.A. An evolutionary perspective on field cancerization. *Nat Rev Cancer* **18**, 19-32 (2018).