

A hybrid modeling environment to describe aggregates of cells heterogeneous for genotype and behavior with possible phenotypic transitions

*Original*

A hybrid modeling environment to describe aggregates of cells heterogeneous for genotype and behavior with possible phenotypic transitions / Chiari, Giulia; Delitala, Marcello Edoardo; Morselli, David; Scianna, Marco. - In: INTERNATIONAL JOURNAL OF NON-LINEAR MECHANICS. - ISSN 0020-7462. - ELETTRONICO. - 144:(2022), p. 104063. [10.1016/j.ijnonlinmec.2022.104063]

*Availability:*

This version is available at: 11583/2964765 since: 2022-05-26T19:07:03Z

*Publisher:*

Elsevier

*Published*

DOI:10.1016/j.ijnonlinmec.2022.104063

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

Elsevier postprint/Author's Accepted Manuscript

© 2022. This manuscript version is made available under the CC-BY-NC-ND 4.0 license  
<http://creativecommons.org/licenses/by-nc-nd/4.0/>. The final authenticated version is available online at:  
<http://dx.doi.org/10.1016/j.ijnonlinmec.2022.104063>

(Article begins on next page)

# A hybrid modeling environment to describe aggregates of cells heterogeneous for genotype and behavior with possible phenotypic transitions

Giulia Chiari<sup>a,b,1</sup>, Marcello Edoardo Delitala<sup>a,2</sup>, David Morselli<sup>a,b,c,3</sup>, Marco Scianna<sup>a,4</sup>

<sup>a</sup>*Department of Mathematical Sciences “G. L. Lagrange”, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy*

<sup>b</sup>*Department of Mathematics “G. Peano”, Università di Torino, Via Carlo Alberto 10, 10124 Torino, Italy*

<sup>c</sup>*Department of Mathematics, School of Science, Computing and Engineering Technologies, Swinburne University of Technology, John St, 3122, Hawthorne, VIC, Australia*

---

## Abstract

Biological systems are typically composed of cells heterogeneous for genotype and phenotype, the latter being time-evolving in response to internal or external stimuli. In order to take these aspects into account, we here propose a modeling framework in which a discrete structuring variable distinguishes cells according to their genotype while a specific mathematical representation (i.e., individual/pointwise vs. collective/density-based) is assigned to each individual on the basis of its phenotypic hallmarks. A coherent procedure is then set to reproduce mechanisms of phenotypic plasticity: based on the definition of a bubble function, which gives the spatial distribution of the mass of a single cell, it possibly accounts the role played by stochasticity and environmental conditions. The proposed modeling environment is then enriched with the inclusion of further cell behavior, such as migratory dynamics and duplication/apoptotic processes, as well as with chemical kinetics. The resulting multiscale hybrid approach is finally applied to the scenario of a heterogeneous tumor aggregate cultured *in vitro*.

*Keywords:* heterogeneous cell population, phenotypic plasticity, hybrid model, discrete vs. continuous mathematical descriptions  
37N25, 45J05, 92C17

---

<sup>1</sup>giulia.chiari@polito.it

<sup>2</sup>marcello.delitala@polito.it

<sup>3</sup>david.morselli@polito.it

<sup>4</sup>marco.scianna@polito.it

## 1. Introduction

We here propose a theoretical/computational approach that allows to characterize cells both at the *genotypic* and at the *phenotypic* level. In particular, a *discrete trait variable* is used to structure a cell population with respect to individual genetic makeup. In other words, each value of this variable is set to correspond to a given sequence of genes. A distinct *mathematical representation*, i.e., pointwise/discrete or density-based/continuous, is instead employed to distinguish cells with respect to their (possibly dichotomic) phenotype/effective behavior, which is established by gene transcription and therefore expression levels. In this respect, the subpopulation of cells with a given phenotype is represented by a set of particles, whereas the remaining group of individuals, characterized by the alternative phenotype, is represented by a continuous density function.

The association between a cell phenotype and a mathematical representation is here determined by reasonable biological arguments. A pointwise description is in fact more appropriate for *specialized/activated/highly metabolic* cells or for cells with *mesenchymal* determinants, i.e., with the ability to undergo individual directional movement in response to environmental cues. On the other hand, a density-based representation, characterized by a lower level of individual detail, is more suitable for *non-specialized/quiescent/poorly metabolic* cell ensembles or for cells with *epithelial* determinants, i.e., which undergo collective dynamics mainly guided by intercellular communication.

In our model, the “discrete vs. continuous” dichotomy is indeed not referred to the spatial scale at which the system is modeled (i.e., “microscopic vs. macroscopic”); rather it is employed to differentiate cell behavior. This aspect distinguishes the mathematical environment proposed here from classical approaches presented in the literature, which typically rely on the idea that a density-based description is a good approximation of the discrete counterpart in the case of systems composed of substantially high amounts of particles with negligible mass. These mathematical frameworks are typically based on mean-field limits [1], heuristic laws of large numbers [2], or coarse-graining procedures [3]. Macroscopic formulations have been also derived by selected lattice-gas cellular automata (LGCA) in [4].

By the definition of a bubble function that represents a plausible spatial distribution of the mass of a single individual, we then implement the passage between the two descriptive instances. This strategy, firstly proposed in [5, 6], allows to model the phenomenon of *phenotypic plasticity*, that is the ability of cells to switch back and forth among multiple phenotypes while maintaining unaltered their genotype [7]. In particular, we here assume that phenotypic conversions are triggered by environmental signals, dependent on cell genetic traits, and affected by randomness. The inclusion of the last aspect represents a significant novelty w.r.t. the previously-cited works [5, 6]. The possibility for cells to have an evolving phenotype has been taken into account in some other approaches. For instance, in individual-based/cellular automata models, each single cell is allowed to vary the label indicating its actual phenotype, as in the

46 case of the well-celebrated Cellular Potts Model, see [8] and reference therein.  
47 Also a hybrid approach allows the description of different phenotypes with differ-  
48 ent discrete populations, as done in [9]; differently from the modeling approach  
49 we present, in this context the discrete setting is used for all cell dynamics while  
50 the continuous description is adopted for microenvironmental dynamics, such as  
51 oxygen or extracellular matrix. Models based on a continuous cell description  
52 (also in the framework of the Theory of Mixtures) instead typically associate  
53 to each subpopulation a distinct density function: phenotypic conversions are  
54 then implemented by mass exchanging terms included in the evolution equa-  
55 tions for cell dynamics, as done for instance in [10, 11]. We refer to [12] for  
56 a comprehensive review of mathematical approaches to model cell plasticity in  
57 the framework of tumor growth. Finally, in approaches dealing with structured  
58 populations, where the trait variable is not referred to the genotype of cells but  
59 rather to their behavioral determinants, and therefore takes values in a con-  
60 tinuous interval, random phenotypic transitions are accounted by including a  
61 diffusion term on the trait domain, see, e.g., [13, 14, 15]. However, in all these  
62 cases phenotypic switches do not imply variations in the mathematical repre-  
63 sentation of cells, which is a novelty introduced in the above-cited works [5, 6]  
64 and here extended by the inclusion of genetic traits and probabilistic aspects.  
65 Our model is finally enriched with cell migratory and growth dynamics, that  
66 are assumed to depend on cell genotype and phenotype.

67 *Applicative potential of the proposed model.* The proposed modeling approach  
68 is indeed able to capture and represent genetic and phenotypic heterogeneity  
69 among a given system of cells, as well as selected mechanisms underlying phe-  
70 notypic plasticity. Its possible applications therefore span a wide spectrum of  
71 phenomena since the evolution of aggregates of cells, from small clusters to large  
72 populations, is typically determined by cooperative dynamics and interactions  
73 between the component individuals differing both at the DNA and at the protein  
74 level.

75 For instance, in most collective cell movement, few specialized individuals,  
76 able to sense environmental chemical signals, typically behave as a pattern-  
77 ing guidance for the rest of the system, which instead passively displaces only  
78 due to adhesion, see [16, 17] and references therein. It is the case of angio-  
79 genic processes, where a small number of endothelial cells forming the walls  
80 of pre-existing vessels acquire a leader/tip phenotype, representing migratory  
81 cues for the neighboring individuals with a follower/stalk behavior [18]. These  
82 mechanisms are triggered by a number of diffusing growth factors (e.g., vascular  
83 endothelial growth factor - VEGF, hepatocyte growth factors - HGF) and  
84 mediated by the well-known Delta-Notch signaling pathways [19, 20].

85 Similarly, during skin repair after injury, the cells located at the front of  
86 the epidermal monolayers that invade the wounded region are characterized by  
87 actin-rich lamellipodia and pseudopodia, that allow active movement, and are  
88 able to synthesize a new basement membrane, whereas individuals at the rear  
89 regions only passively displace dragged by cell-cell adhesive interactions.

90 Cell heterogeneity is also observed in pathological situations, such as in tu-

91 mor growth. For instance, individuals exhibiting different sequences of genes  
 92 and/or phenotypic determinants have been found in several types of disease, in-  
 93 cluding breast cancer [21], colorectal cancer [22], brain cancer [23], and prostate  
 94 cancer [24]. Interestingly, it has been shown that malignant cells within the  
 95 same mass exhibit different behavior in spite of carrying the same genetic al-  
 96 terations [25]. Cancer cells have been also demonstrated to be able to switch  
 97 between alternative phenotypic states either spontaneously or in response to  
 98 ecological inputs. For example, nutrient-deprived malignant individuals activate  
 99 downstream pathways that result in a shift towards a more aggressive behavior.  
 100 These cells in fact lose epithelial characteristics, such as high adhesiveness and  
 101 high duplication capacity, and acquire mesenchymal features, such as enhanced  
 102 motility, which allow them to more effectively invade surrounding tissue. This  
 103 phenomenon, denoted as *epithelial-to-mesenchymal transition (EMT)*, is also  
 104 involved in physiological scenarios, such as morphogenesis and organogenesis.  
 105 The inverse process may occur as well: tumor cell with mesenchymal deter-  
 106 minants can lose their migratory freedom and re-acquire epithelial hallmarks,  
 107 including expression of junctional proteins, when experiencing a sufficient level  
 108 of environmental substrates [26]. Phenotypic differentiation and conversions  
 109 of genetically identical tumor cells have been also shown to (i) facilitate sur-  
 110 vival and adaptation of the entire disease, which can play “hide-and-peek” with  
 111 multiple therapeutic regimes [27, 28], and (ii) fuel subsequent genetic evolution  
 112 [29, 30].

113 *Structure of the article.* The remaining part of the article is organized as fol-  
 114 lows. In Section 2, we introduce the main model ingredients and present a  
 115 sample numerical realization that shows how the procedure for the phenotypic  
 116 switch works. In Section 3, we apply our approach to the representative case  
 117 of a heterogeneous tumor aggregate evolving in an oxygen-deprived domain. In  
 118 Section 4, we finally give some conclusive remarks and hints for further model  
 119 developments.

## 120 2. Proposed approach and representative simulation

121 We are interested in modeling the evolution of an aggregate of cells within  
 122 a closed two-dimensional domain  $D \subset \mathbb{R}^2$ , whose dynamics are studied for the  
 123 period  $T = [0, t_F] \subset \mathbb{R}_0^+$ ,  $t$  being the time variable. The spatial domain  $D$  may  
 124 reproduce, for instance, a planar section of an *in vivo* tissue or the surface of a  
 125 *Petri dish*, usually employed in experimental studies.

126 The cells composing the system are here differentiated according to two  
 127 determinants, as sketched in Fig. 1 (A):

- 128 • their *genotype*, by the use of a *discrete trait variable*  $u$ ;
- 129 • their *phenotype*, by the use of different *mathematical representations*.

130 Our approach is indeed based on the assumption that there is not a deterministic  
 131 and/or necessary relation between the genetic trait of a cell and its phenotype:

132 the latter is in fact determined at the protein level, i.e., by effective gene tran-  
 133 scription and expression levels, which are eventually affected by stochasticity  
 134 and extracellular/environmental stimuli and conditions (by the so-called sur-  
 135 rounding *ecology*).

136 The structuring variable  $u$  is set to assume a given number of values, say  $K$ ,  
 137 i.e.,  $u \in U = \{u_k\}_{k=1}^K$ . In this respect, the generic state  $u_{\hat{k}}$  defines the cell clone  
 138 characterized by the  $\hat{k}$ -th genetic makeup, i.e., by the  $\hat{k}$ -th sequence of genes.

We then consider two alternative cell phenotypes, say “A” and “B”, and associate each of them to a distinct mathematical descriptive instance, as proposed in [5, 6]. In particular, for a given cell variant with genotypic trait  $u_{\hat{k}} \in U$ , the individuals that show phenotype A have a *discrete* representation: they are reproduced with dimensionless points with concentrated unitary mass and identified by their actual position in space (see panel (A) in Fig. 1). Such subgroup of agents can be indeed collected in the following set:

$$\mathbf{X}_{u_{\hat{k}}}^A(t) = \left\{ \mathbf{x}_{1,u_{\hat{k}}}(t), \dots, \mathbf{x}_{N_{u_{\hat{k}}}^A(t),u_{\hat{k}}}(t) \right\}, \quad (1)$$

with  $\mathbf{x}_{i,u_{\hat{k}}}(t) \in D$ , for  $i = 1, \dots, N_{u_{\hat{k}}}^A(t)$ , being  $N_{u_{\hat{k}}}^A(t)$  the number of cells with phenotype A and genotype  $u_{\hat{k}}$  at time  $t$ . The overall amount of individuals with phenotype A within the entire aggregate can be therefore computed as

$$N^A(t) = \sum_{k=1}^K N_{u_k}^A(t). \quad (2)$$

The remaining part of the clone of cells with the  $\hat{k}$ -th genetic trait is instead characterized by phenotype B and *collectively* described by the number density function  $a^B(t, \mathbf{y}, u_{\hat{k}}): T \times D \times U \mapsto \mathbb{R}_0^+$  (as shown in Fig. 1 (A)). The local amount of individuals with phenotype B can be therefore evaluated as

$$\rho^B(t, \mathbf{y}) = \sum_{k=1}^K a^B(t, \mathbf{y}, u_k). \quad (3)$$

In this respect,  $a^B(t, \mathbf{y}, \cdot)$  can be interpreted as the local distribution of cells with phenotype B on the genotype space  $U$ . The overall number of agents with phenotype B actually present within the entire domain  $D$  can be then approximated by integrating  $\rho^B$  along the space variable and rounding down the obtained value:

$$N^B(t) = \left\lfloor \int_D \rho^B(t, \mathbf{y}) \, d\mathbf{y} \right\rfloor. \quad (4)$$

139 The total number of cells composing the aggregate at any given time  $t$  is finally  
 140 equal to  $N(t) = N^A(t) + N^B(t)$ .

141 *Remarks.* For the sake of completeness, we now give some comments on the  
 142 above-proposed modeling framework:

- 143 • the association between the different cell genetic makeups and the corre-  
144 sponding values of the variable  $u$  is arbitrary;
- 145 • the association between a cell phenotype and the corresponding mathe-  
146 matical representation is instead suggested by biological considerations,  
147 as explained in the Introduction of this article;
- 148 • in principle, our approach could include more than two cell phenotypes.  
149 This would require the use of hybrid mathematical representations, i.e.,  
150 able to account for a proper amount of microscopic granularity within a  
151 macroscopic/continuous description of the system of interest, which would  
152 be possible, for instance, by employing tools of Measure Theory [31, 32];
- 153 • the proposed modeling environment is *hybrid* but not, so far, multiscale,  
154 in the sense that different mathematical objects (i.e., material points and  
155 number densities) are used together but represent biological elements at  
156 the same spatial scale, i.e., different types of cells.

*Modeling cell phenotypic plasticity.* In a wide range of biological phenomena, cells are able to change phenotype while maintaining their genetic makeup, i.e., to vary the expression level of one or more of their genes. To reproduce this phenomenon in our modeling framework, we need to set up a procedure to switch between the two cell descriptive instances. It is indeed necessary to define a proper correspondence between the pointwise and the density-based representation of a single cell. In this respect, let us proceed as in [5, 6] and introduce a function  $\varphi_{\mathbf{x}}(\mathbf{y}) : D \times D \mapsto \mathbb{R}_0^+$  such that:

$$\int_D \varphi_{\mathbf{x}}(\mathbf{y}) \, d\mathbf{y} = 1. \quad (5)$$

$\varphi_{\mathbf{x}}$  approximates the spatial distribution of a cell whose center is located in  $\mathbf{x} \in D$ . In principle, there exist several possible options to explicit  $\varphi_{\mathbf{x}}(\mathbf{y})$ . However, in accordance with the already-cited works [5, 6], we hereafter use the following *bubble* function, which assumes a greater amount of cell mass around  $\mathbf{x}$ , as shown in Fig. 1 (B):

$$\varphi_{\mathbf{x}}(\mathbf{y}) = \begin{cases} \frac{4}{\pi r^8} (r^2 - |\mathbf{y} - \mathbf{x}|^2)^3, & \text{if } |\mathbf{y} - \mathbf{x}| \leq r; \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

157 In Eq. (6),  $|\cdot|$  identifies the Euclidean norm while  $r$  is set to approximate a  
158 mean cell radius: hereafter, it will have a value of 15  $\mu\text{m}$ .

Let us now assume that, at a certain time  $t$ , the  $i$ -th cell with phenotype A and genotype  $u_{\hat{k}} \in U$  undergoes a transition to phenotype B. From a biological perspective, this may be the result of environmental stimuli (triggered by chemical signals or by intercellular communication) or of the fact that the individual  $i$  is able to maintain phenotype A only for a limited period of time (e.g., due to high metabolic costs). The proposed A-to-B phenotypic switch can be then

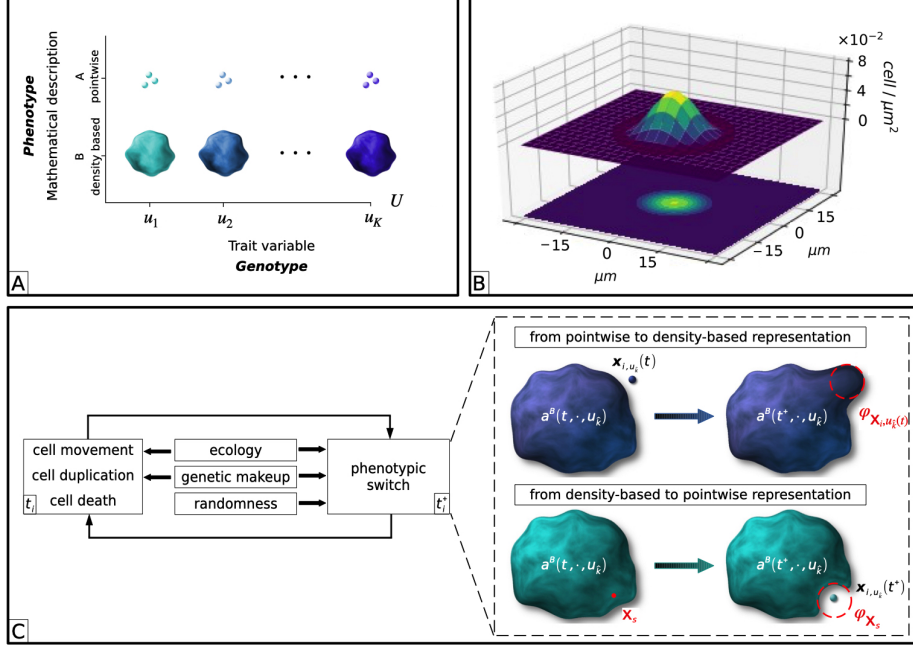


Figure 1: (A) In our modeling environment, each cell is differentiated for genotype, i.e., by the use of a discrete structuring variable  $u \in U$ , and for phenotype, i.e., by the use of a specific mathematical representation. In particular, we only consider two alternative individual phenotypes, which are set to correspond either to a pointwise or to a density-based descriptive instance. (B) For representative purposes, bidimensional and three-dimensional plots of the bubble function centered in  $\mathbf{x} = (0, 0)$ , i.e.,  $\varphi_{(0,0)}$  (cf. Eq. (6)). We recall that the radius  $r$  of the round support of  $\varphi$  is constantly taken equal to  $15 \mu\text{m}$ . (C) We here set that cell dynamics such as growth, migration, and phenotypic switches are affected both by individual genetic trait and by variations in environmental (i.e., ecological) conditions. Stochasticity plays a role as well. In particular, A-to-B phenotypic transition of the generic cell  $i$  with genotype  $u_{\hat{k}}$  is implemented by the removal of the material point located in  $\mathbf{x}_{i, u_{\hat{k}}}$  and the simultaneous addition of the corresponding bubble function  $\varphi_{\mathbf{x}_i, u_{\hat{k}}}$  to the mass distribution  $a^B(\cdot, \cdot, u_{\hat{k}})$ . Conversely, a B-to-A phenotypic switch, stimulated in the domain point  $\mathbf{x}_s$  and involving the cell variant with genotype  $u_{\hat{k}}$ , amounts in the local creation of a new material point  $\mathbf{x}_{N_{u_{\hat{k}}}^A(t)+1, u_{\hat{k}}}$  and in the simultaneous removal of the bubble function  $\varphi_{\mathbf{x}_s}$  to the mass distribution  $a^B(\cdot, \cdot, u_{\hat{k}})$ .

implemented in our modeling framework by removing the material point located in  $\mathbf{x}_{i, u_{\hat{k}}}(t)$  and by simultaneously adding the equivalent mass function  $\varphi_{\mathbf{x}_i, u_{\hat{k}}}(t)$  to the density of the cell variant characterized by the same trait  $u_{\hat{k}}$ , as shown in Fig. 1 (C). In mathematical terms, we indeed get the following relations:

$$\begin{cases} \mathbf{X}_{u_{\hat{k}}}^A(t^+) = \mathbf{X}_{u_{\hat{k}}}^A(t) \setminus \{\mathbf{x}_{i, u_{\hat{k}}}(t)\}; \\ \mathbf{X}_{u_k}^A(t^+) = \mathbf{X}_{u_k}^A(t), \quad \text{for all } k \neq \hat{k}; \\ a^B(t^+, \mathbf{y}, u_{\hat{k}}) = a^B(t, \mathbf{y}, u_{\hat{k}}) + \varphi_{\mathbf{x}_i, u_{\hat{k}}}(t)(\mathbf{y}), \quad \text{for all } \mathbf{y} \in D; \\ a^B(t^+, \mathbf{y}, u_k) = a^B(t, \mathbf{y}, u_k), \quad \text{for all } k \neq \hat{k}; \text{ and } \mathbf{y} \in D. \end{cases} \quad (7)$$

Finally, the remaining particles with phenotype A and genotype  $u_{\hat{k}}$  are renumbered according to the rule

$$\mathbf{x}_{j,u_{\hat{k}}}(t^+) = \begin{cases} \mathbf{x}_{j,u_{\hat{k}}}(t), & \text{if } j < i; \\ \mathbf{x}_{j-1,u_{\hat{k}}}(t), & \text{if } j > i. \end{cases} \quad (8)$$

159 In Eqs. (7) and (8), as well as in the following, the notation  $t^+$  is used to specify  
 160 that, from a *numerical point of view*, phenotypic transitions are not simultane-  
 161 ously implemented with the other processes, e.g., cell movement, duplication,  
 162 death, that occur at the same time instant (see also [5, 6]). The generalization  
 163 of the above procedure to more cells that actually switch from phenotype A to  
 164 phenotype B, possibly with different genotypic traits, is straightforward.

Let us then conversely assume that, at time  $t$ , an environmental stimulus, that is in principle able to trigger a transition from phenotype B to phenotype A in individuals with the generic genotype  $u_{\hat{k}} \in U$ , is active in a given domain location, say  $\mathbf{x}_s \in D$ . Such a switch can occur only if there is a sufficient density of the cell variant of interest to have a localized agent placed in  $\mathbf{x}_s$ . In mathematical terms, this amounts to satisfy the following local constraint:

$$a^B(t, \mathbf{y}, u_{\hat{k}}) \geq \varphi_{\mathbf{x}_s}(\mathbf{y}), \quad \text{for all } \mathbf{y} \in D. \quad (9)$$

In this case, the cell phenotypic transition from B to A (and the corresponding representation switch) results from the removal of  $\varphi_{\mathbf{x}_s}$  from the distribution  $a^B(t, \cdot, u_{\hat{k}})$ , accompanied by the addition of the corresponding new element to the set  $\mathbf{X}_{u_{\hat{k}}}^A$  (see panel (C) in Fig. 1):

$$\begin{cases} \mathbf{X}_{u_{\hat{k}}}^A(t^+) = \mathbf{X}_{u_{\hat{k}}}^A(t) \cup \{\mathbf{x}_{N_{u_{\hat{k}}}^A(t)+1, u_{\hat{k}}}(t) \equiv \mathbf{x}_s\}; \\ \mathbf{X}_{u_k}^A(t^+) = \mathbf{X}_{u_k}^A(t), \quad \text{for all } k \neq \hat{k}; \\ a^B(t^+, \mathbf{y}, u_{\hat{k}}) = a^B(t, \mathbf{y}, u_{\hat{k}}) - \varphi_{\mathbf{x}_s(t)}(\mathbf{y}), \quad \text{for all } \mathbf{y} \in D; \\ a^B(t^+, \mathbf{y}, u_k) = a^B(t, \mathbf{y}, u_k), \quad \text{for all } k \neq \hat{k} \text{ and } \mathbf{y} \in D. \end{cases} \quad (10)$$

165 Furthermore, the following rules are set:

- 166 • in the case of B-to-A phenotypic transitions involving the same cell clone,  
 167 e.g, with genotype  $u_{\hat{k}}$ , and simultaneously stimulated in two distinct  
 168 domain points  $\mathbf{x}_{s1}$  and  $\mathbf{x}_{s2}$  such that  $\varphi_{\mathbf{x}_{s1}}$  and  $\varphi_{\mathbf{x}_{s2}}$  overlap, two al-  
 169 ternative options are accounted: (i) if  $a^B(t, \mathbf{y}, u_{\hat{k}}) \geq \varphi_{\mathbf{x}_{s1}}(\mathbf{y}) + \varphi_{\mathbf{x}_{s2}}(\mathbf{y})$   
 170 for any  $\mathbf{y} \in D$ , then both behavioral switches occur; (ii) if, otherwise,  
 171  $a^B(t, \mathbf{y}, u_{\hat{k}}) \geq \varphi_{\mathbf{x}_{s1}}(\mathbf{y}), \varphi_{\mathbf{x}_{s2}}(\mathbf{y})$  but  $a^B(t, \mathbf{y}, u_{\hat{k}}) < \varphi_{\mathbf{x}_{s1}}(\mathbf{y}) + \varphi_{\mathbf{x}_{s2}}(\mathbf{y})$  for  
 172 at least one domain point, then only one transition takes place, which is  
 173 randomly established. The same rule is extended in the case of more than  
 174 two phenotypic transitions with analogous characteristics;
- 175 • B-to-A phenotypic transitions are not allowed in any domain point effec-  
 176 tively occupied by a pointwise agent (regardless its genotype). Coherently,

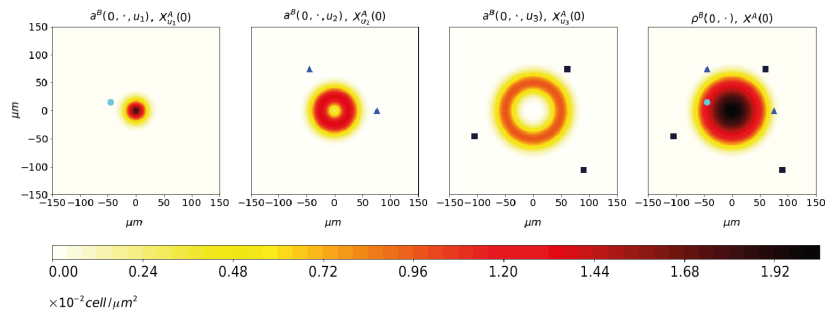


Figure 2: Initial condition of the representative simulation, as specified by Eq. (11). The subpopulation with phenotype B has a radial symmetry: in particular, the cell variant with genotype  $u_1$  is mainly located at the bulk of the cluster, the cell variant with  $u_3$  forms an external ring, whereas the cell variant with  $u_2$  is distributed in the intermediate region. A group of individuals with phenotype A is then dispersed around and within the distribution of cells with phenotype B. In particular, we hereafter use light blue circles to indicate particles with phenotype A and genotype  $u_1$ , blue triangles to indicate particles with phenotype A and genotype  $u_2$ , and dark blue squares to indicate particles with phenotype A and genotype  $u_3$ . Such an initial cell configuration is maintained in the case of the model application proposed in Section 3.

177 only one B-to-A phenotypic switch is allowed (and arbitrarily established)  
 178 at the same time in the same domain point. These constraints are consistent  
 179 with the observation that, in a wide range of phenomena, a cell  
 180 that activates inhibits the surrounding individuals to undergo the same  
 181 process. It is the case, for instance, of the tip cell selection and lateral  
 182 inhibition mechanism controlled by the Delta-Notch pathways during  
 183 physio-pathological angiogenesis;

- 184 • simultaneous B-to-A phenotypic switches occurring at far enough spatial  
 185 regions are instead always permitted.

186 It is instead useful to remark that the above ones are tailored rules and therefore  
 187 can be in principle neglected and/or replaced by other assumptions.

*Sample simulation.* Before including in the proposed modeling framework more realistic biological mechanisms and dynamics, let us propose and comment a representative numerical realization. It deals with a colony of cells which do not grow or move but only undergo arbitrarily selected phenotypic transitions. In more details, in the spatial domain  $D = [-150 \mu\text{m}, 150 \mu\text{m}]^2$ , we place an aggregate whose component individuals can have three different genetic make-ups, i.e.,  $U = \{u_1, u_2, u_3\}$ , while showing the usual dichotomy in the phenotype, i.e., A and B. The initial system configuration is then given by the following

distribution of cells:

$$\begin{cases} \mathbf{X}_{u_1}^A(0) = \{ \mathbf{x}_{1,u_1} = (-45, 15) \}; \\ \mathbf{X}_{u_2}^A(0) = \{ \mathbf{x}_{1,u_2} = (75, 0); \mathbf{x}_{2,u_2} = (-45, 75) \}; \\ \mathbf{X}_{u_3}^A(0) = \{ \mathbf{x}_{1,u_3} = (60, 75); \mathbf{x}_{2,u_3} = (90, -105); \mathbf{x}_{3,u_3} = (-105, -45) \}; \\ a^B(0, \mathbf{y}, u_1) = 3.1 m_\varphi \exp\left(-\frac{|\mathbf{y}|^2}{325}\right); \\ a^B(0, \mathbf{y}, u_2) = 2.4 m_\varphi \exp\left(-\frac{|\mathbf{y} - 25|^2}{325}\right); \\ a^B(0, \mathbf{y}, u_3) = 1.7 m_\varphi \exp\left(-\frac{|\mathbf{y} - 50|^2}{325}\right), \end{cases} \quad (11)$$

for all  $\mathbf{y} \in D$ , being  $m_\varphi = 4/\pi r^8$  the maximum of the *bubble* function (cf. Eq. (6)), see Fig. 2. The overall number of cells at the onset of the simulation, which remains constant in time due to the absence of duplication/death mechanisms, amounts to:

$$\begin{aligned} N(0) &= N^A(0) + N^B(0) \\ &= [\mathbf{X}_{u_1}^A(0)] + [\mathbf{X}_{u_2}^A(0)] + [\mathbf{X}_{u_3}^A(0)] + \left[ \int_D \rho^B(0, \mathbf{y}) d\mathbf{y} \right] \\ &= 6 + \left[ \int_D [a^B(0, \mathbf{y}, u_1) + a^B(0, \mathbf{y}, u_2) + a^B(0, \mathbf{y}, u_3)] d\mathbf{y} \right] = 6 + 188 = 194, \end{aligned} \quad (12)$$

188 where  $[Q]$  indicates the cardinality of a generic set  $Q$ .

189 At a given time  $t_1$ , an external input able to stimulate a switch from phe-  
190 notype B to phenotype A for all cell clones, regardless their genetic trait, ac-  
191 tivates in an arbitrary set of domain points, radially disposed along the main  
192 axes:  $\mathbf{x}_{s1} = (15, 0)$ ,  $\mathbf{x}_{s2} = (50, 0)$ ,  $\mathbf{x}_{s3} = (85, 0)$ ,  $\mathbf{x}_{s4} = (0, 15)$ ,  $\mathbf{x}_{s5} = (0, 50)$ ,  
193  $\mathbf{x}_{s6} = (0, 85)$ ,  $\mathbf{x}_{s7} = (-15, 0)$ ,  $\mathbf{x}_{s8} = (-50, 0)$ ,  $\mathbf{x}_{s9} = (-85, 0)$ ,  $\mathbf{x}_{s10} = (0, -15)$ ,  
194  $\mathbf{x}_{s11} = (0, -50)$ , and  $\mathbf{x}_{s12} = (0, -85)$ , see top panels in Fig. 3. In this respect:

- 195 • no transition takes place in  $\mathbf{x}_{s3}$ ,  $\mathbf{x}_{s6}$ ,  $\mathbf{x}_{s9}$ , and  $\mathbf{x}_{s12}$  due to the lack of  
196 sufficient mass density of any cell genetic variant;
- 197 • in  $\mathbf{x}_{s2}$ ,  $\mathbf{x}_{s5}$ ,  $\mathbf{x}_{s8}$ ,  $\mathbf{x}_{s11}$ , only the subpopulation with genetic trait  $u_3$  is able  
198 to undergoes phenotypic switch, as  $a^B(0, \mathbf{y}, u_3) \geq \varphi_{\mathbf{x}_{s_j}}(\mathbf{y})$  for all  $\mathbf{y} \in D$   
199 and  $s_j \in \{s2, s5, s8, s11\}$ , a condition that instead is not satisfied by the  
200 distributions of the other cell genotypes;
- 201 • in  $\mathbf{x}_{s1}$ ,  $\mathbf{x}_{s4}$ ,  $\mathbf{x}_{s7}$ ,  $\mathbf{x}_{s10}$ , both the cell clone with genotype  $u_1$  and the cell  
202 clone with genotype  $u_2$  have in principle enough mass to undergo a single-  
203 cell switch from phenotype B to phenotype A (i.e.,  $a^B(0, \mathbf{y}, u_1), a^B(0, \mathbf{y}, u_2) \geq$   
204  $\varphi_{\mathbf{x}_{s_j}}(\mathbf{y})$  for all  $\mathbf{y} \in D$  and  $s_j \in \{s1, s4, s7, s10\}$ ). However, as previously  
205 commented, only a single B-to-A phenotypic switch is allowed to occur  
206 at a given time in a given domain location: in this respect, we arbitrarily  
207 establish that in each of the four points, only the genetic variant  $u_2$  is  
208 subjected to phenotypic conversion.

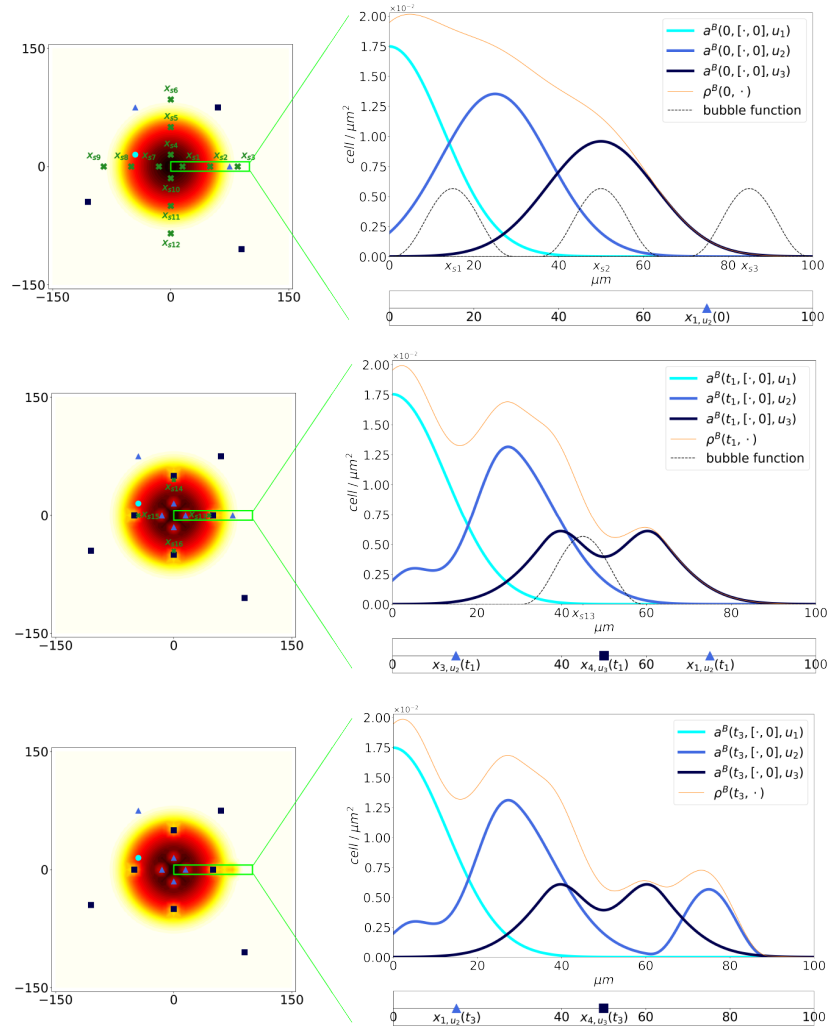


Figure 3: Sample simulation showing how phenotypic switches are implemented in the proposed modeling environment. In the left panels, we represent the evolution of the entire aggregate of cells: in particular, we plot both the overall density of the subpopulation with phenotype B, i.e.,  $\rho^B$  (cf. Eq. (3)), and the set of particles with phenotype A. Within this subgroup, the light blue circles identify cells with genotype  $u_1$ , the blue triangles identify cells with genotype  $u_2$ , and the dark blue squares identify cells with genotype  $u_3$ . The right panels magnify the dynamics of a representative section of the domain.

The above-described dynamics are schematically visualized, in the case of a representative domain section, in the top-right graph of Fig. 3. The updated

system configuration then reads as<sup>5</sup>

$$\left\{ \begin{array}{l} \mathbf{X}_{u_1}^A(t_1) = \mathbf{X}_{u_1}^A(0); \\ \mathbf{X}_{u_2}^A(t_1) = \mathbf{X}_{u_2}^A(0) \cup \{ \mathbf{x}_{3,u_2} \equiv \mathbf{x}_{s1}; \mathbf{x}_{4,u_2} \equiv \mathbf{x}_{s4}; \mathbf{x}_{5,u_2} \equiv \mathbf{x}_{s7}; \mathbf{x}_{6,u_2} \equiv \mathbf{x}_{s10} \}; \\ \mathbf{X}_{u_3}^A(t_1) = \mathbf{X}_{u_3}^A(0) \cup \{ \mathbf{x}_{4,u_3} \equiv \mathbf{x}_{s2}; \mathbf{x}_{5,u_3} \equiv \mathbf{x}_{s5}; \mathbf{x}_{6,u_3} \equiv \mathbf{x}_{s8}; \mathbf{x}_{7,u_3} \equiv \mathbf{x}_{s11} \}; \\ a^B(t_1, \mathbf{y}, u_1) = a^B(0, \mathbf{y}, u_1); \\ a^B(t_1, \mathbf{y}, u_2) = a^B(0, \mathbf{y}, u_2) - \varphi_{\mathbf{x}_{s1}}(\mathbf{y}) - \varphi_{\mathbf{x}_{s4}}(\mathbf{y}) - \varphi_{\mathbf{x}_{s7}}(\mathbf{y}) - \varphi_{\mathbf{x}_{s10}}(\mathbf{y}); \\ a^B(t_1, \mathbf{y}, u_3) = a^B(0, \mathbf{y}, u_3) - \varphi_{\mathbf{x}_{s2}}(\mathbf{y}) - \varphi_{\mathbf{x}_{s5}}(\mathbf{y}) - \varphi_{\mathbf{x}_{s8}}(\mathbf{y}) - \varphi_{\mathbf{x}_{s11}}(\mathbf{y}), \end{array} \right. \quad (13)$$

209 for all  $\mathbf{y} \in D$ . We indeed have that  $N(t_1) = N^A(t_1) + N^B(t_1) = 14 + 180 =$   
 210  $194 = N(0)$ .

211 Successively, at  $t_2$ , an analogous local signal is present in the following set  
 212 of points:  $\mathbf{x}_{s13} = (45, 0)$ ,  $\mathbf{x}_{s14} = (0, 45)$ ,  $\mathbf{x}_{s15} = (-45, 0)$ , and  $\mathbf{x}_{s16} = (0, -45)$ ,  
 213 see the central panels in Fig. 3. In all cases, no phenotypic switch actually  
 214 occurs. In fact, no cell genetic variant has a sufficient amount of mass over the  
 215 support of  $\varphi_{\mathbf{x}_{sj}}$  (with  $j=13, 14, 15, 16$ ) despite the overall mass of individuals  
 216 with phenotype B, measured by  $\rho^B$  would be in principle high enough. In this  
 217 respect, the system does not vary with respect to (13).

We finally set that at time  $t_3$ , the cell  $\mathbf{x}_{1,u_2}$ , located in  $(75, 0)$  from the  
 beginning of the observation time, is triggered to turn back to phenotype B, as  
 shown in the bottom panels of Fig. 3. The pointwise particle is indeed replaced  
 by the corresponding bubble function, that is added to the mass of the proper  
 cell genetic variant, as

$$\left\{ \begin{array}{l} \mathbf{X}_{u_1}^A(t_3) = \mathbf{X}_{u_1}^A(t_2) = \mathbf{X}_{u_1}^A(t_1) = \mathbf{X}_{u_1}^A(0); \\ \mathbf{X}_{u_2}^A(t_3) = \mathbf{X}_{u_2}^A(t_2) \setminus \{ \mathbf{x}_{1,u_2} \} = \mathbf{X}_{u_2}^A(t_1) \setminus \{ \mathbf{x}_{1,u_2} \}; \\ \mathbf{X}_{u_3}^A(t_3) = \mathbf{X}_{u_3}^A(t_2) = \mathbf{X}_{u_3}^A(t_1); \\ a^B(t_3, \mathbf{y}, u_1) = a^B(t_2, \mathbf{y}, u_1) = a^B(t_1, \mathbf{y}, u_1) = a^B(0, \mathbf{y}, u_1); \\ a^B(t_3, \mathbf{y}, u_2) = a^B(t_2, \mathbf{y}, u_2) + \varphi_{\mathbf{x}_{1,u_2}}(\mathbf{y}) = a^B(t_1, \mathbf{y}, u_2) + \varphi_{\mathbf{x}_{1,u_2}}(\mathbf{y}); \\ a^B(t_3, \mathbf{y}, u_3) = a^B(t_2, \mathbf{y}, u_3) = a^B(t_1, \mathbf{y}, u_3), \end{array} \right. \quad (14)$$

218 for all  $\mathbf{y} \in D$ , so that  $N(t_3) = N^A(t_3) + N^B(t_3) = 13 + 181 = 194 = N(0)$ . For  
 219 the sake of reader's convenience, we recall that the element belonging to the set  
 220  $\mathbf{X}_{u_2}^A$  have to be renumbered according to (8).

221 *Remark.* As already commented in the Introduction, and sketched in Fig. 1 (C),  
 222 a cell is stimulated to undergo phenotypic plasticity by environmental signals,  
 223 but the effective transition depends on its genetic makeup and on the intrinsic  
 224 stochasticity of the mechanism. These aspects have not been accounted so far,  
 225 as all the proposed cell phenotypic switches have been set to actually take place

---

<sup>5</sup>*Notation remark:* since in this simulation setting cell dynamics only include phenotypic plasticity, the differentiation between  $t_i$  and  $t_i^+$  (for  $i = 1, 2, 3$ ) is not necessary, and therefore avoided for the sake of simplicity.

226 (provided a sufficient cell mass in the case of B-to-A conversions). Such a model  
227 shortcoming is tackled in the next section, where more realistic rules underlying  
228 variations in cell phenotype will be given.

### 229 3. Model application: early dynamics of an *in vitro* tumor aggregate

230 We then turn to apply the proposed model to one of the scenarios introduced  
231 in Section 1, i.e., the tumor growth. In particular, we hereafter show how our  
232 approach can be used to reproduce selected aspects of the early dynamics of a  
233 malignant aggregate cultured *in vitro*. In the context of our interest, the trait  
234 variable  $u$  is set to assume three values, i.e.,  $U = \{u_1 = 0; u_2 = 0.5; u_3 = 1\}$ ,  
235 each indicating a distinct sequence of genes. In this respect, the higher is the  
236 value of  $u$  the more the corresponding genotype is associated to cells that in  
237 principle have high migratory potential and low proliferation capacity, see Fig. 4  
238 (A). The definition of the structuring variable  $u$  is indeed coherent with the “Go  
239 or Grow” (GoG) assumption, which finds support from both the experimental  
240 [33, 34] and the theoretical literature [35]. Phenotype A, and therefore an  
241 individual pointwise representation, is then assigned to describe tumor cells  
242 with *mesenchymal* determinants (i.e., that show an effectively high invasiveness  
243 and a poor mitotic activity). Phenotype B, as long as a collective density-  
244 based representation, is instead assigned to malignant individuals with *epithelial*  
245 hallmarks (i.e., low migratory ability but high duplication rates). Such modeling  
246 assumptions are sketched in the already-cited panel (A) of Fig. 4.

247 In agreement with the scheme shown in Fig. 1 (C), we then assume that  
248 phenotypic transitions are:

- 249 • stimulated by variations in environmental conditions, in particular in the  
250 availability of oxygen, whose local concentration will be given by the field  
251 variable  $O(t, \mathbf{y}) : T \times D \mapsto \mathbb{R}_0^+$ . In this respect, hypoxia has been widely  
252 shown to boost phenotypic instability, acting as a fuel of selective pres-  
253 sure that stimulates tumor cells to shift towards more aggressive (mes-  
254 enchymal) hallmarks [36]. For instance, tumor cells displaying high levels  
255 of hypoxia-inducible factors, such as HIF-1, have been demonstrated to  
256 overexpress genes relative to the migratory machinery and underexpress  
257 genes related to mitotic processes, see [37] and references therein. In the  
258 case of a sufficient amount of resources, malignant individuals have been  
259 instead shown to maintain or recover a less invasive (epithelial) behavior.  
260 In this respect, cells with low levels of HIF-1 have been shown to transcript  
261 mainly genes implicated in duplication activities [37];
- 262 • affected by the cell genetic makeup: for instance, a variant characterized  
263 by a sequence of genes mainly relative to the migratory machinery more  
264 likely maintains or acquires a mesenchymal behavior (and *vice versa*) [38];
- 265 • subjected to randomness, which is a critical aspect in most biological  
266 phenomena.

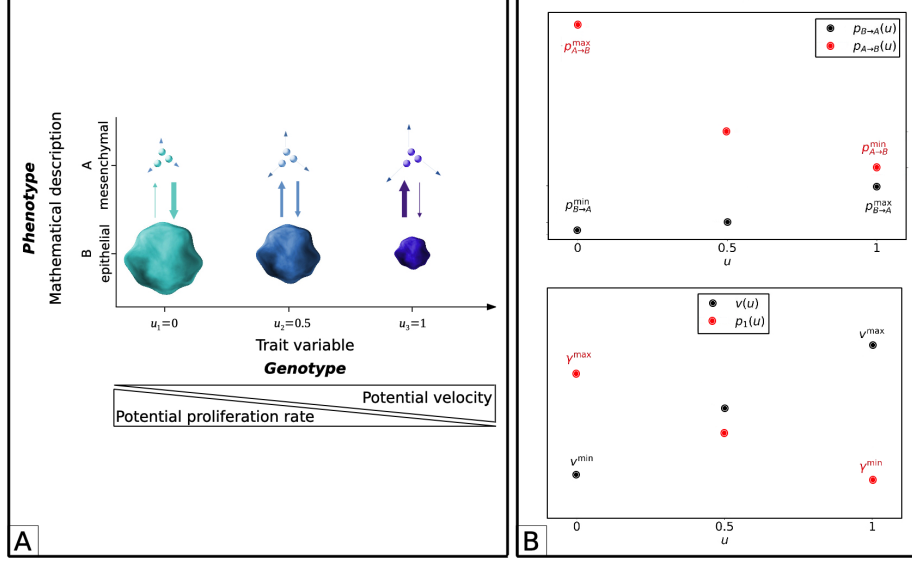


Figure 4: (A) In the proposed model application, the trait value  $u$  is set to qualitatively evaluate the cell motility/proliferation potential. In particular, the higher is the value of  $u$  the more a tumor individual is assigned a sequence of genes that, if expressed, enhances its migratory ability while dropping its duplication capacity. The phenotype A, and therefore the corresponding pointwise representation, is given to malignant cells with mesenchymal characteristics; the phenotype B, and the corresponding density-based representation, is instead assigned to tumor agents with epithelial hallmarks. The thickness of the vertical arrows gives a qualitative indication of the probability that a cell with a given genotype has to undergo one of the two phenotypic transitions. In particular, as also shown in the bottom graph of the panel (B), cells with genotype  $u = u_1 = 0$  more likely acquire (or maintain) an epithelial behavior. In contrast, cells with genotype  $u = u_3 = 1$  more likely acquire (or maintain) mesenchymal hallmarks. (B) Top plot: influence of the genetic trait of a cell on the probability of phenotypic conversions (see Eqs. (17) and (20)). Bottom plot: genotypic-dependent duplication rate of malignant epithelial cells ( $p_1$ , see Eq. (23)) and speed of mesenchymal individuals ( $v$ , see Eq. (27)).

In principle, transition probabilities have to be given as random variables defined on spatio-temporal continuous domains. However, in the perspective of numerical realizations of the proposed model, we here account only for their discretized counterpart. According to these considerations, the probability of a cell  $\mathbf{x}_{i,u_{\hat{k}}}(t)$  with phenotype A and genotype  $u_{\hat{k}} \in U$  to undergo phenotypic transition in an interval of time  $(t - \Delta t, t] \subset T$ , being  $\Delta t$  the size of the time grid (see below), is equal to:

$$P_{A \rightarrow B}(O(t, \mathbf{x}_{i,u_{\hat{k}}}(t)), u_{\hat{k}}) = q_{A \rightarrow B}(O(t, \mathbf{x}_{i,u_{\hat{k}}}(t))) p_{A \rightarrow B}(u_{\hat{k}}). \quad (15)$$

In (15), the first factor evaluates the environmental conditions experienced by the  $i$ -th individual, i.e.,

$$q_{A \rightarrow B}(O(t, \mathbf{x}_{i,u_{\hat{k}}}(t))) = H(O(t, \mathbf{x}_{i,u_{\hat{k}}}(t)) - O_M) \quad (16)$$

being

$$H(O(t, \mathbf{x}_{i, u_{\hat{k}}}(t)) - O_M) = \{1, \text{ if } O(t, \mathbf{x}_{i, u_{\hat{k}}}(t)) \geq O_M; 0, \text{ if } O(t, \mathbf{x}_{i, u_{\hat{k}}}(t)) < O_M\}$$

the Heaviside function and  $O_M$  the amount of molecular substance needed by tumor cells to remain in a normoxic condition, i.e., to avoid hypoxia. With Eq. (16), we are assuming that mesenchymal cells experiencing oxygen deprivation do not undergo phenotypic transitions. The second factor in (15) instead reads as:

$$p_{A \rightarrow B}(u_{\hat{k}}) = (p_{A \rightarrow B}^{\max} - p_{A \rightarrow B}^{\min})(1 - u_{\hat{k}})^2 + p_{A \rightarrow B}^{\min}. \quad (17)$$

267 It indeed sets a quadratic dependence between the genetic makeup of the cell  
 268 and its possibility to switch phenotype. In this respect, in the case of normoxic  
 269 conditions, mesenchymal individuals with genotype  $u_1 = 0$  acquire epithelial  
 270 hallmarks with a probability equal to  $p_{A \rightarrow B}^{\max}$  whereas particles with genotype  
 271  $u_3 = 1$  with a probability equal to  $p_{A \rightarrow B}^{\min}$  where, according to the above-explained  
 272 biological arguments,  $p_{A \rightarrow B}^{\min} < p_{A \rightarrow B}^{\max}$ , see Fig. 4 (B-top plot).

Conversely, considering the same time and space discretization of the previous case, a cell clone with genotype  $u_{\hat{k}}$  and phenotype B, i.e., whose distribution is given by the density  $a^B(t, \cdot, u_{\hat{k}})$ , is set to acquire mesenchymal determinants at a certain point  $\mathbf{x}_s \in D$  of the discretized space and in an interval of time  $(t - \Delta t, t] \subset T$  with a probability equal to

$$P_{B \rightarrow A}(O(t, \mathbf{x}_s), u_{\hat{k}}) = q_{B \rightarrow A}(O(t, \mathbf{x}_s)) p_{B \rightarrow A}(u_{\hat{k}}), \quad (18)$$

where, recalling (16),

$$q_{B \rightarrow A}(O(t, \mathbf{x}_s)) = H(O_M - O(t, \mathbf{x}_s)). \quad (19)$$

The above formula implies that only hypoxic conditions can trigger epithelial-to-mesenchymal transitions, whose probability to effectively occur depends also in this case by the cell genotype:

$$p_{B \rightarrow A}(u_{\hat{k}}) = (p_{B \rightarrow A}^{\max} - p_{B \rightarrow A}^{\min})u_{\hat{k}}^2 + p_{B \rightarrow A}^{\min}, \quad (20)$$

273 where  $p_{B \rightarrow A}^{\max}$  characterizes the cell clone with trait  $u_3 = 1$  and  $p_{B \rightarrow A}^{\min}$  the cell  
 274 variant with  $u_1 = 0$ , being  $p_{B \rightarrow A}^{\max} > p_{B \rightarrow A}^{\min}$ , as plotted in the top graph of Fig.  
 275 4 (B). Obviously, the B-to-A phenotypic transition actually takes place if the  
 276  $u_{\hat{k}}$ -th cell variant has enough mass over the support of  $\varphi_{\mathbf{x}_s}$ .

277 *Remarks.* For the sake of completeness, we now give some comments on the  
 278 above-proposed modeling framework:

- 279 • as we will see in details in the section devoted to the simulation details,  
 280 the sizes of the time and space discretization steps affect the estimate of  
 281 the parameters  $p_{A \rightarrow B}^{\max}$ ,  $p_{A \rightarrow B}^{\min}$ ,  $p_{B \rightarrow A}^{\max}$  and  $p_{B \rightarrow A}^{\min}$ ;
- 282 • phenotypic transitions are actually employed according to the correspond-  
 283 ing procedures explained in the previous section;

- 284 • in the case of simultaneously possible epithelial-to-mesenchymal switches  
285 occurring in the same domain point, it only takes place the one involving  
286 the cell variant with the highest value of  $u$ ;
- 287 • in Eqs. (17) and (20), we have assumed a quadratic relationship between  
288 the value of the structuring variable  $u$  and the transition probabilities.  
289 Different laws may of course be chosen: however, they have to maintain  
290 the same qualitative trends of those proposed here;
- 291 • more sophisticated functions may be set also to describe the influence of  
292 oxygen on phenotypic variations. For instance, the probability of a cell to  
293 acquire mesenchymal determinants may increase upon decrements in the  
294 chemical concentration below the threshold  $O_M$ . One could also consider  
295 two different oxygen thresholds  $O_{M1} < O_{M2}$  such that the phenotypic  
296 switch from A to B occurs for oxygen concentrations above  $O_{M1}$  and the  
297 phenotypic switch from B to A occurs for oxygen concentrations below  
298  $O_{M2}$ .

*Cell dynamics.* Malignant cells with epithelial determinants are here assumed to proliferate and undergo random movement. The evolution of the density of the  $u_{\hat{k}}$ -th variant with phenotype B can be indeed described by means of the following partial differential equation (PDE), whose boundary and initial conditions will be specified later on:

$$\frac{\partial a^B}{\partial t}(t, \mathbf{y}, u_{\hat{k}}) = \underbrace{D_B \Delta a^B(t, \mathbf{y}, u_{\hat{k}})}_{\text{diffusive movement}} + \underbrace{p(u_{\hat{k}}, \rho(t, \mathbf{y})) a^B(t, \mathbf{y}, u_{\hat{k}})}_{\text{proliferation}}, \quad (21)$$

where  $\rho(t, \mathbf{y})$  account for the local tumor mass (see below Eqs. (25) and (26)). The diffusion term at the r.h.s. of Eq. (21), with constant coefficient  $D_B > 0$ , models Brownian cell displacements. The reaction term instead expresses local variations in the mass of the  $u_{\hat{k}}$ -th epithelial cell variant. In particular, they are assumed to depend on (i) individual genetic trait and (ii) physical limitations determined by the available space. In this respect,  $p$  can be factorized as it follows:

$$p(u_{\hat{k}}, \rho(t, \mathbf{y})) = p_1(u_{\hat{k}}) p_2(\rho(t, \mathbf{y})). \quad (22)$$

The duplication law  $p_1$  accounts for the fact that higher proliferation rates characterize cell variants with lower values of the trait variable  $u$  (that, as previously seen, are associated to sequence of genes mainly implicated in the mitotic machinery). In this respect, to avoid overcomplications, we assign to  $p_1$  a linear trend, see Fig. 4 (B-bottom plot):

$$p_1(u_{\hat{k}}) = (\gamma^{\max} - \gamma^{\min})(1 - u_{\hat{k}}) + \gamma^{\min}, \quad (23)$$

being  $\gamma^{\max}$  a maximal duplication rate, characteristic of cells with genotype  $u = u_1 = 0$ , and  $\gamma^{\min}$  the corresponding minimal value, that is instead assigned to individuals with genotype  $u = u_3 = 1$ . The factor  $p_2$  in Eq. (22) instead

models the fact that the mitotic cycle is typically disrupted in overcompressed cells, although abnormal proliferation is a relevant characteristic of malignant masses. This phenomenon can be replicated by setting the following logistic law:

$$p_2(\rho(t, \mathbf{y})) = 1 - \frac{\rho(t, \mathbf{y})}{c}, \quad (24)$$

where  $c > 0$  is a carrying capacity while

$$\rho(t, \mathbf{y}) = \rho^A(t, \mathbf{y}) + \rho^B(t, \mathbf{y}), \quad (25)$$

being  $\rho^B$  defined as in Eq. (3), and

$$\rho^A(t, \mathbf{y}) = \sum_{k=1}^3 \sum_{i=1}^{N_{u_k}^A} \varphi_{\mathbf{x}_{i,u_k}(t)}(\mathbf{y}). \quad (26)$$

299 In Eq. (24), we consider that the available space is reduced by the presence also  
 300 of mesenchymal individuals, whose influence on the overall mass distribution can  
 301 be accounted by the use of the corresponding set of bubble functions, as given  
 302 in Eq. (26). Eq. (21) is then equipped by Neumann homogeneous boundary  
 303 conditions on the spatial domain  $D$ , which are consistent with the fact that cells  
 304 can not physically cross the border of an experimental *Petri dish*.

The dynamics of tumor cells with mesenchymal determinants only include a directional movement towards domain regions with higher oxygen concentrations. In this respect, for the  $i$ -th individual with phenotype A and generic genotype  $u_{\hat{k}}$ , we set:

$$\frac{d\mathbf{x}_{i,u_{\hat{k}}}(t)}{dt} = \frac{\nabla O(t, \mathbf{x}_{i,u_{\hat{k}}}(t))}{|\nabla O(t, \mathbf{x}_{i,u_{\hat{k}}}(t))|} v(u_{\hat{k}}), \quad (27)$$

305 with  $v(u_{\hat{k}}) = (v^{\max} - v^{\min})u_{\hat{k}} + v^{\min}$ , see the bottom graph in Fig. 4 (B).  
 306 In Eq. (27), cell speed and direction of movement are decoupled, given their  
 307 distinct physical meaning. The former depends on the pattern of available  
 308 resources, the latter, quantified by the scalar functions  $v : U \mapsto [v^{\min}, v^{\max}]$ , is  
 309 instead affected by individual genetic makeup. In this respect, recalling that  
 310 higher values of  $u$  imply higher motile potential,  $v^{\max}$  is the speed of cells with  
 311 genotype  $u = u_3 = 1$ , whereas  $v^{\min}$  of cells with genotype  $u = u_1 = 0$ . It  
 312 is finally useful to underline that Eq. (27) is based on the *overdamped force-*  
 313 *velocity assumption*: it establishes that, in extremely viscous regimes such as  
 314 biological environments, the velocity of moving agents and not their acceleration  
 315 is proportional to the sensed forces (see [8] and references therein for a detailed  
 316 comment). When a mesenchymal cancer cells reaches a point of the border of  $D$ ,  
 317 the component of its velocity locally normal to the boundary itself is arbitrarily  
 318 set equal to zero.

319 Summing up, it is possible to conclude that, in this sample model application,  
 320 genetic trait and ecological/environmental conditions not only affect phenotypic

321 transitions of the cancer cells but also their effective growth and migratory  
 322 dynamics, as sketched in panel (C) of Fig. 1.

*Chemical dynamics.* We assume that oxygen diffuses within the domain and is consumed equally by all tumor individuals, regardless their genotype and phenotype. Its kinetics can be therefore described by the following reaction-diffusion (RD) equation:

$$\frac{\partial O}{\partial t}(t, \mathbf{y}) = \underbrace{D_O \Delta O(t, \mathbf{y})}_{\text{diffusion}} - \underbrace{\lambda_O \rho(t, \mathbf{y}) O(t, \mathbf{y})}_{\text{consumption by tumor cells}} - \underbrace{\alpha_O O(t, \mathbf{y})}_{\text{decay}}, \quad (28)$$

323 where  $D_O$ ,  $\lambda_O$ , and  $\alpha_O$  are constant coefficients, that quantify chemical diffusion,  
 324 consumption by malignant cells and natural decay, respectively, being  $\rho$  defined  
 325 as in Eq. (25). Eq. (28) is finally completed with Dirichlet conditions along the  
 326 entire domain boundary  $\partial D$ , i.e.,  $O(t, \partial D) = \bar{O}$ , for all  $t \in T$ : we are indeed  
 327 assuming a continuous and constant chemical supply within our virtual *Petri*  
 328 *dish*. The oxygen initial pattern will be instead specified below. It is useful to  
 329 remark that the inclusion of chemical dynamics gives to our model a *multiscale*  
 330 aspect, as it now deals with elements characteristic of both the cellular and the  
 331 subcellular levels.

332 *Numerical details.* For the spatial domain  $D$ , we have employed a triangular  
 333 mesh with radial symmetry with respect to the center point  $(0, 0)$ . The charac-  
 334 teristic diameter of each grid element has been taken equal to  $\Delta x = 5 \mu\text{m}$ . For  
 335 the time domain  $T$ , we have used an uniform discretization with step equal to  
 336  $\Delta t = 1 \text{ h}$ .

337 Eqs. (21) and (28), describing the dynamics of the continuous population  
 338 and of the oxygen, have been solved employing a time-explicit Euler method  
 339 coupled with a Galerkin finite-element technique. An explicit Euler method has  
 340 been also employed for the system of ODEs describing movement of pointwise  
 341 cells (cf. Eq. (27)). At any discrete time-step, phenotypic switches are im-  
 342 plemented (as explained in Section 2) just *after* the numerical solution of the  
 343 above-cited equation for cell dynamics.

344 Considering B-to-A switches, the following algorithmic rules are implemented  
 345 for each numerical node of the domain:

- 346 (i) the oxygen level is checked: if it is higher than  $O_M$ , then no phenotypic  
 347 transition occurs and we pass to another domain point;
- 348 (ii) otherwise, we check the mass of the cell subpopulation with  $u = u_3 = 1$ :  
 349 if it satisfies condition (9) then a random number from the uniform distri-  
 350 bution between 0 and 1 is drawn. If this number is lower than the value of  
 351 the probability given in (18) and evaluated in the case of our interest, then  
 352 the phenotypic transition occurs and we pass to another domain point (re-  
 353 call that a B-to-A phenotypic transition of a given subpopulation locally  
 354 inhibits analogous processes involving other subpopulations);
- 355 (iii) otherwise, the same evaluations described at point (ii) are performed for  
 356 the other subpopulations in descending order with respect to  $u$  (to be

357 coherent with the fact that cells with higher genotypic traits  $u$  are more  
 358 likely to switch phenotype).

359 We keep into account that, when a B-to-A transition takes place in one point,  
 360 it affects the possibility of transition in neighboring points, as some of the con-  
 361 tinuous mass is removed. Thus, in order to avoid biases in spatial location of  
 362 B-to-A phenotypic switches, at every iteration we randomize the order in which  
 363 the points of the numerical lattice are visited.

364 We then turn on considering possible A-to-B transitions, which take place  
 365 in areas with oxygen concentration above  $O_M$  with probability given by (15)  
 366 (using the same drawing algorithm described above). We finally remark that  
 367 the order in which cells with phenotype A are checked for possible transitions  
 368 does not affect numerical outcomes, since A-to-B transitions are independent of  
 369 each other.

370 All numerical computations have been performed in Fenics, see [39, 40] and  
 371 references therein.

*Parameter estimate.* As previously commented, the probabilities of phenotypic transitions introduced in Eqs. (15) and (18) are the discretized approximations of the corresponding continuous-in-time (and in-space) laws. In more details, the coefficient  $p_{A \rightarrow B}^{\max}$  ( $p_{A \rightarrow B}^{\min}$ , resp.) defines the probability that the  $i$ -th cell with genotype  $u = u_1 = 0$  ( $u = u_3 = 1$ , resp.) undergoes phenotypic transition at a given time step, i.e., in the case of normoxic conditions. The estimation of these values is based on the average time that a cell with mesenchymal characteristics takes to re-acquire epithelial hallmarks; in our model we assume that it ranges from  $T_{A \rightarrow B}^{\min} = 50$  h to  $T_{A \rightarrow B}^{\max} = 200$  h. Such quantities (poorly measured in the empirical literature, see [41] for one of the few contributions in this respect) have been fixed in order to have a reasonable number of phenotypic transitions in the period of observation. By recalling that our model is based on the assumption that cells with lower values of the trait variable more likely undergo A-to-B transitions, we can indeed set

$$p_{A \rightarrow B}^{\max} = \frac{\Delta t}{T_{A \rightarrow B}^{\min}} \quad \text{and} \quad p_{A \rightarrow B}^{\min} = \frac{\Delta t}{T_{A \rightarrow B}^{\max}},$$

so that  $p_{A \rightarrow B}^{\max} = 2 \times 10^{-2}$ ,  $p_{A \rightarrow B}^{\min} = 5 \times 10^{-3}$ . The coefficients  $p_{B \rightarrow A}^{\max, \min}$  instead give the probability that a single-cell-fraction of mass with phenotype B and centered in  $\mathbf{x}_s$  changes phenotype at a given time step when falls in hypoxic conditions. A proper estimate can be obtained by taking into account three aspects: (i) epithelial cells experiencing oxygen deprivation are here assumed to acquire mesenchymal determinants in a time lapse that ranges from  $T_{B \rightarrow A}^{\min} = 8.8$  h to  $T_{B \rightarrow A}^{\max} = 35.4$  h; (ii) in our modeling framework higher values of the genotypic variable imply more possibility to switch towards phenotype A; and (iii) a finer spatial grid requires a smaller transition probability for each node  $\mathbf{x}_s$ , otherwise a higher amount of possible nodes of the domain in principle could allow a higher number of transitions. Taken together, the above considerations lead to

$$p_{B \rightarrow A}^{\max} \propto \Delta t, (T_{B \rightarrow A}^{\min})^{-1}, \Delta x^2 \quad \text{and} \quad p_{B \rightarrow A}^{\min} \propto \Delta t, (T_{B \rightarrow A}^{\max})^{-1}, \Delta x^2.$$

372 In particular, after preliminary simulations, we have fixed  $p_{B \rightarrow A}^{\max} = 4 \times 10^{-3}$ ,  
 373 and  $p_{B \rightarrow A}^{\min} = 10^{-3}$ , which have allowed us to have a reasonable rate of B-to-A  
 374 phenotypic conversions.

375 The diffusion coefficient of epithelial cell movement, i.e.,  $D_B$ , has been taken  
 376 equal to  $1.29 \times 10^3 \mu\text{m}^2/\text{h}$ , as in [42]. The coefficients  $\gamma^{\min}$  and  $\gamma^{\max}$  quantify the  
 377 minimal and maximal mitotic rate of cells with phenotype B, in the case of fully  
 378 available space. The chosen values  $\gamma^{\min} = \ln(2)/48 \text{ h}^{-1}$  and  $\gamma^{\max} = \ln(2)/24$   
 379  $\text{h}^{-1}$  fall within the range quantified for glioblastoma cell lines in either hypoxic or  
 380 normoxic conditions, see again [42]. The carrying capacity  $c$  has been set equal  
 381 to  $1.69 \text{ cell}/\mu\text{m}^2$ , in order to maintain a quasi-monolayered cell configuration,  
 382 in agreement with the bidimensional nature of experimental cultures in a *Petri*  
 383 *dish*.

384 Cells with phenotype A are allowed to freely move within the domain. In this  
 385 respect, the maximal value of their speed  $v^{\max}$ , which characterize mesenchymal  
 386 individuals with trait  $u_3 = 1$  has been fixed to  $10 \mu\text{m}/\text{h}$ , whereas the minimal  
 387 threshold  $v^{\min}$ , which characterizes mesenchymal individuals with trait  $u_1 = 0$ ,  
 388 to  $2.5 \mu\text{m}/\text{h}$ . These parameters have been taken from [43] and assure that the  
 389 modulus of the overall cell velocity substantially falls within the range of the  
 390 corresponding experimental counterparts evaluated for different malignancies.

391 The chemical threshold that leads to hypoxia, i.e.,  $O_M$ , has been set equal to  
 392  $2.56 \times 10^{-15} \mu\text{mol}/\mu\text{m}^2$ , as it is done in [42]. The diffusion coefficient of oxygen  
 393 has been fixed to  $D_O = 3.60 \times 10^6 \mu\text{m}^2/\text{h}$ , and taken again from [42]. The chem-  
 394 ical consumption rate then amounts to  $\lambda_O = 1.67 \times 10^{-10} \mu\text{m}^2/(\text{cell} \cdot \text{h})$ : it has  
 395 been empirically measured taking into account of the proposed computational  
 396 setup, in order to have a realistic time-evolution of the molecular pattern. The  
 397 oxygen decay coefficient has been fixed to  $\alpha_O = 3.60 \times 10^{-4} \text{ h}^{-1}$ , according to  
 398 [44]. The constant production of oxygen at the domain border, i.e.,  $\bar{O}$ , has been  
 399 set equal to  $2.8 \times 10^{-15} \mu\text{mol}/\mu\text{m}^2$ : for the reader's convenience, we remark that  
 400 this value is  $1.1 \times O_M$ . The final observation time  $t_F$  has been instead set equal  
 401 to 35 h.

402 The employed parameter setting is listed in Table 1.

403 *Simulation results.* The spatial domain  $D$ , as well as the initial configuration of  
 404 the cell system, is exactly the same employed in the representative simulation  
 405 given in Section 2, specified by Eqs. (11) and (12), and represented in Fig 2.  
 406 At the onset of the forthcoming numerical realization, we indeed have a tumor  
 407 aggregate with few mesenchymal cells (heterogenous for genotype) dispersed  
 408 within and around a cluster of malignant epithelial individuals. In particular,  
 409 the node of tumor cells with phenotype B has a radial distribution w.r.t. the  
 410 center of the domain, with the bulk mainly constituted by the cell variant with  
 411  $u_1 = 0$  and the external region by the cell variant with  $u_3 = 1$ . The initial  
 412 oxygen concentration is instead given by the stationary solution of Eq. (28),  
 413 evaluated in the absence of cancer cells (i.e., in the case only of chemical diffusion  
 414 and decay): given the low value of the decay rate  $\alpha_O$  (see above and Table  
 415 1), it consists of a spatially quasi-homogeneous pattern with a chemical level  
 416 approximately equal to  $2.8 \times 10^{-15} \mu\text{mol}/\mu\text{m}^2$ . The initial oxygen level indeed

Parameter	Value [Units]	Reference
$r$	15 [ $\mu\text{m}$ ]	[45]
$p_{A \rightarrow B}^{\min}$	$5 \times 10^{-3}$	model estimate
$p_{A \rightarrow B}^{\max}$	$2 \times 10^{-2}$	model estimate
$p_{B \rightarrow A}^{\min}$	$10^{-3}$	model estimate
$p_{B \rightarrow A}^{\max}$	$4 \times 10^{-3}$	model estimate
$D_B$	$1.29 \times 10^3$ [ $\mu\text{m}^2/\text{h}$ ]	[42]
$\gamma_{\min}$	$\ln(2)/48$ [ $\text{h}^{-1}$ ]	[42]
$\gamma_{\max}$	$\ln(2)/24$ [ $\text{h}^{-1}$ ]	[42]
$c$	1.69 [ $\text{cell}/\mu\text{m}^2$ ]	model estimate
$v^{\min}$	2.5 [ $\mu\text{m}/\text{h}$ ]	[43]
$v^{\max}$	10 [ $\mu\text{m}/\text{h}$ ]	[43]
$D_O$	$3.60 \times 10^6$ [ $\mu\text{m}^2/\text{h}$ ]	[42]
$\lambda_O$	$1.67 \times 10^{-10}$ [ $\mu\text{m}^2/(\text{cell} \cdot \text{h})$ ]	model estimate
$\alpha_O$	$3.60 \times 10^{-4}$ [ $\text{h}^{-1}$ ]	[44]
$O_M$	$2.56 \times 10^{-15}$ [ $\mu\text{mol}/\mu\text{m}^2$ ]	[42]

Table 1: Simulation parameter setting.

417 exceeds the hypoxic threshold  $O_M$  in the entire domain.

418 Oxygen consumption then starts to occur at the domain area occupied by  
419 the tumor aggregate, with the extent of local decrements obviously determined  
420 by the density of malignant individuals. The level of chemical at the inner part  
421 of the mass indeed drops to the critical value  $O_M$  and an increasing number of  
422 epithelial tumor cells (characterized by negligible motility) experiences hypoxia.  
423 Some of them are then able to undergo phenotypic transition and to acquire  
424 mesenchymal determinants, see Fig. 5. This group is mainly composed of  
425 individuals with a trait value  $u_3 = 1$ , which is associated to the sequence of  
426 genes that favors (from a probabilistic point of view) such a phenotypic switch.

427 The just-differentiated mesenchymal cells, as long as those already present  
428 at the onset of the simulation, crawl towards oxygenated domain regions: in  
429 particular, each of them moves with a speed dictated by its genetic trait, as  
430 shown by the length of the arrows attached to the particles in Fig. 5. The  
431 remaining fraction of epithelial individuals is instead not able to escape harsh  
432 environmental conditions: in the case of long-term hypoxia (e.g., long-lasting  
433 oxygen deprivation), their fate would be an irreversible necrosis.

434 As the simulation proceeds, the domain region with low chemical level en-  
435 larges: as a result, the above-described cell dynamics take place in more periph-  
436 eral areas of the tumor aggregate and involves an increasing amount of epithelial  
437 mass. In particular, at the end of the observation time (i.e., at  $t = t_F = 35$   
438 hours), the cell configuration consists of a hypoxic cluster of epithelial tissue,  
439 mainly formed by individuals with a trait variable equal to  $u_1 = 0$ . It is sur-  
440 rounded by scattered mesenchymal cells, that have reached the external regions  
441 of the domain, i.e., those with higher oxygen availability. Interestingly, few  
442 of these agents have been able to undergo the inverse transition and reacquire  
443 epithelial hallmarks (see the bottom panels of Fig. 5). During the entire obser-

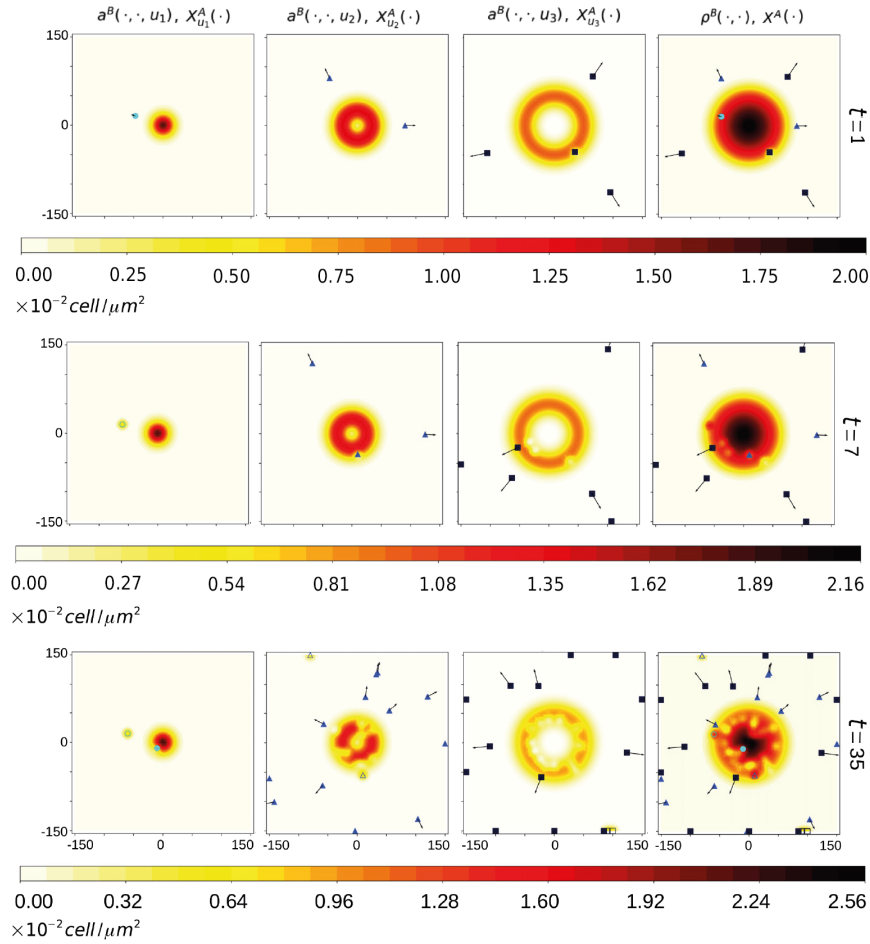


Figure 5: Representative time instants of the evolution of our virtual tumor aggregate. The initial condition of the cell system is exactly the same as in Section 2, see Fig. 2. At the onset of the numerical realization, the oxygen is quasi-homogenously present within the entire domain with a level that is higher than the hypoxic threshold  $O_M$ . Subsequent oxygen consumption results in harsh conditions for malignant epithelial cells: some of them are then able to acquire mesenchymal hallmarks (according to the genotype-dependent probabilistic rule given in (18)) and move towards domain regions with more availability of resources (see top and middle panels, i.e., those relative to  $t = 1$  and  $7$  h). Arrived close to the border of our virtual *Petri dish*, few of them experience normoxia and recover epithelial determinants (see the bottom panels, i.e., those relative to  $t = t_F = 35$  h). We remark that light blue circles identify mesenchymal cells with genotype  $u_1$ , blue triangles identify mesenchymal cells with genotype  $u_2$ , and dark blue squares identify mesenchymal cells with genotype  $u_3$ . The same *empty* geometric labels instead identify mesenchymal cell variants that have undergone the inverse, i.e., A-to-B, phenotypic transition. The arrow attached to each mesenchymal individual identify its velocity: its length is qualitatively proportional to the individual genotype-dependent speed.

444 vation time, the fraction of malignant epithelial mass goes on proliferating (cf.  
445 the variations in the values of the colorbar in Fig. 5).

446 Our numerical results qualitatively agree with a wide range of experimen-  
447 tal evidence, which has shown that malignant cells with different phenotypic  
448 properties occupy tumor regions characterised by different oxygen levels. For  
449 instance, glioblastoma spheroids cultured *in vitro* have the core mainly pop-  
450 ulated by cells with a proliferative activity higher than those located at the  
451 invasive edges [46, 47, 48, 49]. Analogously, mesenchymal cancer stem cells  
452 have been found to be abundant near the tumor-stroma boundary (i.e., at the  
453 external region of the malignant mass) [19]. Similar phenotypic spatial hetero-  
454 geneity has been observed in malignant spheroids of ovarian [50, 51] or breast  
455 [52] carcinomas grown in spinner cultures.

456 Similar growth of tumor masses, i.e., characterized by an inner region of  
457 poorly motile individuals unable to escape nutrient deprivation and by an ex-  
458 ternal possibly scattered ring of aggressive cells, has been also predicted by a  
459 wide spectrum of theoretical models, see the comprehensive books [53, 54] and  
460 the excellent reviews [55, 56, 57, 58, 59].

#### 461 4. Conclusions and future perspectives

462 We have here proposed a modeling framework where cells are distinguished  
463 in terms of genotype by a discrete structuring variable and in terms of phenotype  
464 by the assigned mathematical representation (i.e., pointwise or density-based).  
465 A procedure to consistently switch between the two descriptive instances, which  
466 is based on the definition and the use of a bubble function, has then allowed to  
467 account for phenotypic plasticity.

468 We have then presented a representative simulation to show how pheno-  
469 typic transitions actually take place within our theoretical environment, that  
470 has been finally applied to a more realistic scenario, i.e., the early evolution of a  
471 heterogeneous tumor aggregate hypothetically cultured *in vitro*. In particular,  
472 we have assumed that malignant cells can have one of three distinct genotypes  
473 and one of two alternative, i.e., mesenchymal vs. epithelial, behavior. Pheno-  
474 typic conversions have been set to depend on (i) oxygenation levels, (ii) intrinsic  
475 genotype, and (iii) randomness, which is a novelty of this work w.r.t. [5, 6]. The  
476 resulting numerical realization has captured the realistic emergence of a hypoxic  
477 core within the tumor cluster with the consequent cell tendency to acquire a  
478 more aggressive and invasive (i.e., mesenchymal) phenotype.

479 *Model improvements.* The proposed mathematical environment may be im-  
480 proved at least in two direction.

481 From a strictly modeling perspective, it would be relevant to account for ge-  
482 netic alterations, that may be induced by cell-cell communication and changes  
483 in environmental conditions but that are usually determined by random muta-  
484 tions. This last aspect can be included in the proposed modeling environment  
485 by stochastic variations of the value of the trait variable  $u$  assigned to one or  
486 more pointwise individuals and/or to one or more portions of the cell mass with

487 the density-based representation. Furthermore, one could consider a continuous  
488 trait  $u$  that takes values in a given interval (e.g.,  $[0, 1]$ ). This would amount  
489 in using a structuring variable to represent not only genetic heterogeneity (as  
490 in our model), but also epigenetic heterogeneity: each value of  $u$  in fact would  
491 represent the (normalized) expression of a gene or of a group of genes (or the  
492 level of one or more proteins). In this case, epigenetic variations in the cell pop-  
493 ulation could be accounted by including a diffusion term in the trait domain, as  
494 done in the already cited works [13, 14, 15].

495 From an application perspective, our model could be extended to reproduce  
496 the evolution of a malignant mass *in vivo*, i.e., to shed lights on the effect of  
497 intratumoral heterogeneity and phenotypic plasticity on the invasiveness of the  
498 disease. In this respect, one may include in the picture the presence of both the  
499 preexisting and the tumor-induced vasculature. As a natural extension of our  
500 model assumptions, we would in fact have to take into account that cancer cells  
501 in hypoxic conditions not only shift towards more aggressive phenotypes but also  
502 secrete proangiogenic factors which induce the formation of new blood vessels  
503 departing from existing ones [60]. In addition, our model could be developed to  
504 incorporate a more comprehensive description of the metabolism of the different  
505 cell variants. However, in order to provide consistent results of a such an *in*  
506 *vivo* scenario, model parametrization should be better calibrated, for instance  
507 by focusing on a specific tumor type and using proper sets of existing data.

## 508 Acknowledgments

509 This research was partially supported by the Italian Ministry of Education,  
510 University and Research (MIUR) through the “Dipartimenti di Eccellenza” Pro-  
511 gramme (2018-2022) – Dipartimento di Scienze Matematiche “G. L. Lagrange”,  
512 Politecnico di Torino (CUP: E11G18000350001). All authors are members  
513 of GNFM (Gruppo Nazionale per la Fisica Matematica) of INdAM (Istituto  
514 Nazionale di Alta Matematica).

## 515 References

- 516 [1] J. A. Carrillo, M. Fornasier, G. Toscani, F. Vecil, Particle, kinetic, and  
517 hydrodynamic models of swarming, in: G. Naldi, L. Pareschi, G. Toscani  
518 (Eds.), *Mathematical Modeling of Collective Behavior in Socio-Economic*  
519 *and Life Sciences, Modeling and Simulation in Science, Engineering and*  
520 *Technology*, Birkäuser, 2010, pp. 297–336.
- 521 [2] V. Capasso, D. Morale, Asymptotic behavior of a system of stochastic par-  
522 ticles subject to nonlocal interactions, *Stochastic Analysis and Applications*  
523 27 (3) (2009) 574–603.
- 524 [3] D. Drasdo, Coarse graining in simulated cell populations, *Advanced Com-*  
525 *plex Systems* 8 (2-3) (2005) 319–363.

- 526 [4] K. Bottger, H. Hatzikirou, A. Chauviere, A. Deutsch, Investigation of the  
527 migration/proliferation dichotomy and its impact on avascular glioma in-  
528 vasion, *Mathematical Modelling Natural Phenomena* 7 (1) (2012) 105–135.
- 529 [5] A. Colombi, M. Scianna, L. Preziosi, Coherent modelling switch between  
530 pointwise and distributed representations of cell aggregates, *Journal of*  
531 *mathematical biology* 74 (4) (2017) 783–808.
- 532 [6] M. Scianna, A. Colombi, A coherent modeling procedure to describe cell  
533 activation in biological systems, *Communications in Applied and Industrial*  
534 *Mathematics* 8 (1) (2017) 1–22.
- 535 [7] M. Holzel, A. Bovier, T. Tuting, Plasticity of tumour and immune cells: a  
536 source of heterogeneity and a cause for therapy resistance?, *Nature Reviews*  
537 *Cancer* 13 (5) (2012) 365–376.
- 538 [8] M. Scianna, L. Preziosi, Multiscale developments of cellular potts models,  
539 *Multiscale Modeling & Simulation* 10 (2) (2012) 342–382.
- 540 [9] A. R. A. Anderson, A. M. Weaver, P. T. Cummings, V. Quaranta, Tumor  
541 morphology and phenotypic evolution driven by selective pressure from the  
542 microenvironment, *Cell* 127 (5) (2006) 905–915.
- 543 [10] C. Giverso, T. Lorenzi, L. Preziosi, Effective interface conditions for con-  
544 tinuum mechanical models describing the invasion of multiple cell popula-  
545 tions through thin membranes, *Applied Mathematics Letters* 125 (107708)  
546 (2022) 1–9.
- 547 [11] S. M. Wise, J. S. Lowengrub, H. B. Frieboes, V. Cristini, Three-dimensional  
548 multispecies nonlinear tumor growth— I Model and numerical method,  
549 *Journal of Theoretical Biology* 253 (3) (2008) 524–543.
- 550 [12] H. N. Weerasinghe, P. M. Burrage, K. Burrage, D. V. Nicolau, Mathe-  
551 matical models of cancer cell plasticity, *Journal of Oncology* 2019 (2019)  
552 1–14.
- 553 [13] G. Fiandaca, M. Delitala, T. Lorenzi, A mathematical study of the influence  
554 of hypoxia and acidity on the evolutionary dynamics of cancer, *Bulletin of*  
555 *mathematical biology* 83 (83) (2021) 1–29.
- 556 [14] T. Lorenzi, C. Venkataraman, A. Lorz, M. A. J. Chaplain, The role of  
557 spatial variations of abiotic factors in mediating intratumour phenotypic  
558 heterogeneity, *Journal of theoretical biology* 451 (2018) 101–110.
- 559 [15] A. Lorz, T. Lorenzi, J. Clairambault, A. Escargueil, B. Perthame, Modeling  
560 the effects of space structure and combination therapies on phenotypic  
561 heterogeneity and drug resistance in solid tumors, *Bulletin of mathematical*  
562 *biology* 77 (1) (2015) 1–22.

- 563 [16] O. Ilina, P. Friedl, Mechanisms of collective cell migration at a glance,  
564 *Journal of Cell Science* 122 (18) (2009) 3203–3208.
- 565 [17] A. A. Khalil, P. Friedl, Determinants of leader cells in collective cell migra-  
566 tion, *Integrative Biology* 2 (11-12) (2010) 568–574.
- 567 [18] M. Boareto, M. K. Jolly, A. Goldman, P. M., S. A. Mani, S. Sengupta,  
568 B.-J. Eshel, L. Herbert, N. O. Jose', Notch-jagged signalling can give rise  
569 to clusters of cells exhibiting a hybrid epithelial/mesenchymal phenotype,  
570 *Journal of the Royal Society Interface* 13 (118) (2016) 1–11.
- 571 [19] S. Liu, Y. Cong, D. Wang, Y. Sun, L. Deng, Y. Liu, R. Martin-  
572 Trevino, L. Shang, S. P. McDermott, M. D. Landis, S. Hong, A. Adams,  
573 R. D'Angelo, C. Ginestier, E. Charafe-Jauffret, S. G. Clouthier, D. Birn-  
574 baum, W. S. T., M. Zhan, J. C. Chang, M. S. Wicha, Breast cancer stem  
575 cells transition between epithelial and mesenchymal states reflective of their  
576 normal counterparts, *Stem Cell Reports* 2 (1) (2013) 78–91.
- 577 [20] C. K. Williams, J. L. Li, M. Murga, A. L. Harris, G. Tosato, Upregulation  
578 of the notch ligand delta-like inhibits vegf induced endothelial cell function,  
579 *Blood* 107 (3) (2006) 931–939.
- 580 [21] M. Al-Hajj, M. S. Wicha, A. Benito-Hernandez, S. J. Morrison, M. F.  
581 Clarke, Prospective identification of tumorigenic breast cancer cells, *Pro-  
582 ceedings of the National Academy of Sciences* 100 (7) (2003) 3983–3988.
- 583 [22] C. A. O'Brien, A. Pollett, S. Gallinger, J. E. Dick, A human colon cancer  
584 cell capable of initiating tumour growth in immunodeficient mice, *Nature*  
585 445 (7123) (2007) 106–110.
- 586 [23] S. K. Singh, C. Hawkins, I. D. Clarke, J. A. Squire, J. Bayani, T. Hide,  
587 R. M. Henkelman, M. D. Cusimano, P. B. Dirks, Identification of human  
588 brain tumour initiating cells, *Nature* 432 (7015) (2004) 396–401.
- 589 [24] A. T. Collins, P. A. Berry, C. Hyde, M. J. Stower, N. J. Maitland, Prospec-  
590 tive identification of tumorigenic prostate cancer stem cells, *Cancer Re-  
591 search* 65 (23) (2005) 10946–10951.
- 592 [25] A. Marusyk, V. Almendro, K. Polyak, Intra-tumour heterogeneity: a look-  
593 ing glass for cancer?, *Nature Reviews Cancer* 12 (2012) 323–334.
- 594 [26] M. A. Nieto, R. Y.-J. Huang, R. A. Jackson, J. P. Thiery, Emt: 2016, *Cell*  
595 166 (1) (2016) 21–45.
- 596 [27] A. Roesch, Tumor heterogeneity and plasticity as elusive drivers for resis-  
597 tance to mapk pathway inhibition in melanoma, *Oncogene* 34 (23) (2015)  
598 2951–2957.
- 599 [28] J. Varga, T. De Oliveira, F. R. Greten, The architect who never sleeps:  
600 tumor-induced plasticity, *FEBS Letters* 588 (15) (2014) 2422–2427.

- 601 [29] M. D. Brooks, M. L. Burness, M. S. Wicha, Therapeutic implications of  
602 cellular heterogeneity and plasticity in breast cancer, *Cell Stem Cell* 17 (3)  
603 (2015) 260–271.
- 604 [30] S. M. Mooney, M. K. Jolly, H. Levine, P. Kulkarni, Phenotypic plasticity  
605 in prostate cancer: role of intrinsically disordered proteins, *Asian Journal*  
606 *of Andrology* 18 (5) (2016) 704–710.
- 607 [31] A. Colombi, M. Scianna, A. Tosin, Differentiated cell behaviour: a mul-  
608 tiscale approach using measure theory, *Journal of Mathematical Biology*  
609 71 (5) (2014) 1049–1079.
- 610 [32] A. Colombi, M. Scianna, L. Preziosi, A measure-theoretic model for cell  
611 migration and aggregation, *Mathematical Modelling of Natural Phenomena*  
612 10 (1) (2015) 4–35.
- 613 [33] A. Giese, L. Kluwe, M. E. Berens, M. Westphal, Migration of human glioma  
614 cells on myelin, *Neurosurgery* 38 (4) (1996) 755–764.
- 615 [34] A. Giese, M. A. Loo, N. Tran, D. Haskett, S. W. Coons, M. E. Berens,  
616 Dichotomy of astrocytoma migration and proliferation, *Int. J. Cancer* 67 (2)  
617 (1996) 275–282.
- 618 [35] S. M. Schaller, A. Deutsch, H. Hatzikirou, D. Basanta, ‘Go or grow’: the  
619 key to the emergence of invasion in tumour progression?, *Mathematical*  
620 *Medicine and Biology* 29 (1) (2012) 49–65.
- 621 [36] S.-H. Kao, K.-J. Wu, W.-H. Lee, Hypoxia, epithelial-mesenchymal tran-  
622 sition, and tet-mediated epigenetic changes, *Journal of Clinical Medicine*  
623 5 (2) (2016) 24–38.
- 624 [37] H. N. Barrak, A. K. Maitham, Y. A. Luqmani, Hypoxic environment may  
625 enhance migration/penetration of endocrine resistant mcf7-derived breast  
626 cancer cells through monolayers of other non-invasive cancer cells in vitro,  
627 *Scientific Reports* 10 (1) (2020) 1–14.
- 628 [38] H. L. Rocha, I. Godet, F. Kurtoglu, J. Metzcar, K. Konstantinopoulos,  
629 S. Bhoyar, D. M. Gilkes, P. Macklin, A persistent invasive phenotype in  
630 post-hypoxic tumor cells is revealed by fate mapping and computational  
631 modeling, *Science* 24 (9) (2021) 1–22.
- 632 [39] M. S. Almaes, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. . Logg,  
633 A. Richardson, J. Ring, M. E. Rognes, G. N. Wells, The fenics project  
634 version 1.5, *Archive of Numerical Software* 3 (100) (2015) 9–23.
- 635 [40] A. Logg, K.-A. Mardal, G. N. Wells, *Automated Solution of Differential*  
636 *Equations by the Finite Element Method*, Springer, 2012.

- 637 [41] N. M. Aiello, R. Maddipati, R. J. Norgard, D. Balli, J. Li, S. Yuan, T. Yamazoe, T. Black, A. Sahmoud, E. E. Furth, D. Bar-Sagi, B. Z. Stanger, Emt subtype influences epithelial plasticity and mode of cell migration, *Developmental Cell* 45 (2018) 681–695.
- 638  
639  
640
- 641 [42] A. Martínez-González, G. F. Calvo, L. A. Pérez Romasanta, V. Pérez-García, Hypoxic cell waves around necrotic cores in glioblastoma: a biomathematical model and its therapeutic implications, *Bulletin of Mathematical Biology* 74 (12) (2012) 2875–2896.
- 642  
643  
644
- 645 [43] J. A. Gallaher, J. S. Brown, A. R. A. Anderson, The impact of proliferation-migration tradeoffs on phenotypic evolution in cancer, *Scientific Reports* 9 (2425) (2019) 1–10.
- 646  
647
- 648 [44] P. Cumsille, A. Coronel, C. Conca, C. Quiñinao, C. Escudero, Proposal of a hybrid approach for tumor progression and tumor-induced angiogenesis, *Theoretical Biology and Medical Modelling* 12 (13) (2015) 1–22.
- 649  
650
- 651 [45] B. Alberts, A. Johnson, J. Lewis, *Molecular Biology of the Cell* 4th edition, Garland Science, 2002.
- 652
- 653 [46] R. Abramovitch, G. Meir, M. Neeman, Neovascularization induced growth of implanted c6 glioma multicellular spheroids: magnetic-resonance microimaging, *Cancer Research* 55 (9) (1995) 1956–1962.
- 654  
655
- 656 [47] M. A. A. Castro, F. Klamt, V. A. Grieneisen, I. Grivicich, J. C. F. Moreira, Gompertzian growth pattern correlated with phenotypic organization of colon carcinoma, malignant glioma and non-small cell lung carcinoma cell line, *Cell Proliferation* 36 (2) (2003) 65–73.
- 657  
658  
659
- 660 [48] D. Khaitan, S. Chandna, M. B. Arya, D. B. S., Establishment and characterization of multicellular spheroids from a human glioma cell line: implications for tumor therapy, *J. Transl. Med.* 4 (12) (2006) 1–13.
- 661  
662
- 663 [49] A. M. Stein, T. Demuth, D. Mobley, M. Berens, L. M. Sander, A mathematical model of glioblastoma tumor spheroid invasion in a three-dimensional in vitro experiment, *Biophysical Journal* 92 (1) (2007) 356–365.
- 664  
665
- 666 [50] K. M. Burleson, M. P. Boente, S. E. Parmabuccian, A. P. Skubitz, Disaggregation and invasion of ovarian carcinoma ascites spheroids, *Journal of Translational Medicine* 4 (6) (2006) 1–16.
- 667  
668
- 669 [51] K. Shield, M. L. Ackland, N. Ahmed, G. E. Rice, Multicellular spheroids in ovarian cancer metastases: biology and pathology, *Gynecologic Oncology* 113 (1) (2009) 143–148.
- 670  
671
- 672 [52] R. A. Gatenby, K. Smallbone, P. K. Maini, F. Rose, J. Averill, R. B. Nagle, L. Worrall, R. J. Gillies, Cellular adaptations to hypoxia and acidosis during somatic evolution of breast cancer, *British Journal of Cancer* 97 (5) (2007) 646–653.
- 673  
674  
675

- 676 [53] V. Cristini, J. Lowengrub, Multiscale modeling of cancer: An integrated  
677 experimental and mathematical modeling approach, Cambridge University  
678 Press, 2010.
- 679 [54] L. Preziosi, Cancer modelling and simulation, CRC Press, 2003.
- 680 [55] R. P. Araujo, D. L. S. McElwain, A history of the study of solid tumour  
681 growth: the contribution of mathematical modelling, *Bulletin of Mathe-*  
682 *matical Biology* 66 (2004) 1039–1091.
- 683 [56] N. Bellomo, N. K. Li, P. K. Maini, On the foundations of cancer modelling:  
684 selected topics, speculations, and perspectives, *Mathematical Models and*  
685 *Methods in Applied Sciences* 18 (04) (2008) 593–646.
- 686 [57] H. M. Byrne, T. Alarcon, M. R. Owen, S. D. Webb, P. K. Maini, Mod-  
687 elling aspects of cancer dynamics: a review, *Philosophical Transactions*  
688 *of the Royal Society A: Mathematical, Physical and Engineering Sciences*  
689 364 (1843) (2006) 1563–1578.
- 690 [58] M. A. Chaplain, Avascular growth, angiogenesis and vascular growth in  
691 solid tumours: The mathematical modelling of the stages of tumour devel-  
692 opment, *Mathematical and Computer Modelling* 23 (6) (1996) 47–87.
- 693 [59] V. Quaranta, A. M. Weaver, P. T. Cummings, A. R. A. Anderson, Mathe-  
694 matical modeling of cancer: the future of prognosis and treatment, *Clinica*  
695 *Chimica Acta* 357 (2) (2005) 173–179.
- 696 [60] K. H. Plate, G. Breier, H. A. Weich, W. Risau, Vascular endothelial growth  
697 factor is a potential tumour angiogenesis factor in human gliomas in vivo,  
698 *Nature* 359 (6398) (1992) 845–848.