Patterns of tumor progression predict small and tissue-specific tumor-originating niches

Thomas Buder[1,2,*], Andreas Deutsch[1], Barbara Klink[3,4], Anja Voss-Böhme[1,2]

1 Center for Information Services and High Performance Computing, Technische Universität Dresden, Nöthnitzer Str. 46, Dresden, 01062, Germany
2 Faculty of Informatics/Mathematics, HTW Dresden - University of Applied Sciences, Friedrich-List-Platz 1, 01069 Dresden, Germany
3 Institute for Clinical Genetics, Faculty of Medicine Carl Gustav Carus, Technische Universität Dresden, Fetscherstr. 74, Dresden, 01307, Germany
4 German Cancer Consortium (DKTK), Dresden, Germany, German Cancer Research Center (DKFZ), Heidelberg, Germany; Center for Molecular Tumor Diagnostics (CMTD), National Center for Tumor Diseases (NCT), Dresden, Germany

* Correspondence to Thomas Buder

HTW Dresden - University of Applied Sciences
Faculty of Informatics/Mathematics
Friedrich-List-Platz 1
D-01069 Dresden, Germany
tel: +49 351 463-31944
mail: thomas.buder@htw-dresden.de

Abstract

Cancer development is a multistep process in which cells increase in malignancy through progressive alterations. The early phase of this process is hardly observable which aggravates an understanding of later tumor development. We shed light on this initial phase with a cell-based stochastic model calibrated with epidemiological data from the tissue scale. Our model allows to estimate the number of tumor cells needed for tumor formation in human tissues based on data on the diagnosed ratios of benign and malignant tumors. We find that the minimal number of cells needed for tumor formation is surprisingly small and largely depends on the tissue type. Our results point towards the existence of tumor-originating niches in which the fate of tumor development is early decided. Our estimate for the human

colon agrees well with the size of the stem cell niche in colonic crypts. Our estimates might help to identify the tumor-originating cell type, e.g. our analysis suggests for glioblastoma that the tumors originate from a cell type competing in a range of 300 - 1900 cells.

Summary

We estimate the number of tumor cells needed for tumor formation in human tissues and propose the existence of small and tissue-specific tumor-originating niches which might help to find tumor-originating cell types, in particular in glioblastoma.

Introduction

Cancer development is a multistep process in which cells acquire a certain number of progressive epigenetic and genetic alterations [1]. This multistep process can be divided into a neutral and a selection phase. In the neutral phase, the epigenetic and genetic alterations do not confer a proliferative fitness advantage to the tumor precursor cells whereas cells gain such an advantage in the selection phase [2,3]. A single genetically altered cell does not necessarily induce tumor formation but is rather exposed to competition with its corresponding wild-type cells [4]. The realization of this competition depends on the tissue and cell type. It can be direct due to replacement of a cell by the offspring of another proliferating cell [4] but also indirect, for example by symmetric and asymmetric division of cells [5]. Importantly, tissues are composed of different types of cells but only those which are capable to give rise to a progeny able to accumulate alterations can be tumor-originating cell types [6]. Tumor-originating cell refers to the wild-type cell of a certain type that acquires the first alteration in the multistep process of cancer development. Within the neutral phase, the progeny of the tumor-originating cell competes with wild-type cells within normal tissue homeostasis. Because this competition is controlled by the original tissue organization, the range of this competition is determined by the tissue structure which provides natural spatial boundaries for the spread of the progeny of the tumor-originating cell [7,8].

In order to induce tumor formation, the progeny of the tumor-originating cell must not go extinct but has to establish within the tissue. This establishment is achieved by clonal expansion to a sufficiently large cell population [9]. For some tissues, there is experimental evidence that this establishment is characterized by an outcompetition of wild-type cells within the homeostatic range of competition. For example, the tumor-originating cell within the human colon has been identified to be almost always a stem cell with a first hit in the APC gene, and a second hit in this gene is sufficient to induce adenoma formation, a benign precursor of malignant adenocarcinoma. These stem cells reside at the bottom of so-called niches within colonic crypts and are capable of self-renewal and multilineage

differentiation [10]. It has been demonstrated that tumor-originating cells neutrally compete with wild-type stem cells for a position within the spatially restricted stem cell niche [5]. Either such an altered stem cell goes extinct due to this competition or eventually replaces all wild-type stem cells within the stem cell niche. This process has been termed monoclonal conversion and represents almost always the first step of tumor formation within the human colon [10]. Hence, the monoclonal conversion of the stem cell niche by the progeny of the tumor-originating cell with loss of the APC gene induces the establishment of an adenoma on the tissue scale. However, in other tissues the early phase of tumor development on the cellular scale is less understood. The main reason is a lack of knowledge regarding the tumor-originating cell type. Similar to the colon, it has been shown that stem cells within the hematopoietic system represent the tumor-originating cell type [11, 12]. In contrast, there is also evidence that non-stem cells can be the tumor-originating cell type, e.g. in oligodendroglioma [13]. Although the lineage in which cancer originates has been revealed for skin, pancreatic, brain and breast tumors, the tumor-originating cell type remains elusive in most cases [14]. Its identification may allow earlier detection of malignancies and may lead to preventive therapies for individuals at high risk of developing cancer [14].

On the tissue scale, one observes different types of tumor progression. Tumors can progress sequentially, i.e. with a clinically detectable benign precursor stage. Alternatively, they can also progress by tunneling without such a prior benign precursor stage. Epidemiological data allow to infer the progression patterns with respect to the ratios of tunneling versus sequential progression of different tumors. Interestingly, these progression patterns differ largely between tissues. Some tumors exhibit predominantly sequential progression, e.g. benign adenoma almost always develop prior to adenocarcinoma in the colon [10]. Similarly, multiple myeloma are in almost all cases preceded by a premalignant state called monoclonal gammopathy of undetermined significance (MGUS) [15]. In contrast, glioblastoma develops in 90% of all cases without evidence of a less malignant precursor lesion (primary glioblastoma) and progresses in 10% of all cases from low-grade tumors (secondary glioblastoma) [16]. In which way these progression patterns on the tissue scale emerge from the multistep process of cancer development on the cellular scale is difficult to infer since the early phase of this multistep process is hardly observable.

In this work, we use observables on the tissue scale to shed light on the early cellular processes of tumor development. We utilize a MORAN model with mutations [17–19] to describe cellular competition between wild-type cells and tumor cells. Benign and malignant tumor subtypes on the tissue scale are represented by two absorbing states within the model. We incorporate epidemiological data on the progression patterns of cancers to calibrate the model. By analyzing the model dynamics with respect to different spatial cell arrangements, we obtain a lower and upper bound for the critical number of tumor cells needed for tumor development on the tissue scale. Interestingly, our estimates are considerably small, tissue-specific and far away from the overall number of cells in a clinically observable tumor. We therefore propose that the fate of tumor development is decided in tissue-specific tumor-originating niches. This proposal is supported by our estimate of the tumor-originating niche size for the human colon which agrees well with the size of the stem cell niche in colonic crypts. In particular, we propose that a tumor-originating niche size of

300 – 1900 cells within the human brain can explain the ratio of primary and secondary glioblastoma. Interestingly, our estimates also agree well with the minimal number of tumor cells needed for tumor formation in mice injection experiments and might allow to infer the tumor-originating cell type.

Materials and methods

State space and representation of benign and malignant tumor subtypes. The multistep process in which cancer cells increase gradually in malignancy differs with respect to the number of steps, e.g. two steps in retinoblastoma [20] compared to seven steps in colon cancer [21]. In our cell-based model, we only regard the last step within the neutral phase and the first step within the selection phase such that we obtain a two-step process. This coarse-grained approach is appropriate for our purpose since we are only interested in modeling tumor progression patterns and not quantities which are largely influenced by the precise number of steps, e.g the time-scale of tumor development or intra-tumor heterogeneity. In the cellular two-step process, genetic or epigenetic alterations can transform *wild-type cells* into *benign tumor cells* which can further progress to *malignant tumor cells*. We assume that the benign progeny of the tumor-originating cell competes with wild-type cells and can clonally expand within normal tissue homeostasis. The parameter N in our model describes the homeostatic range of this competition. We further assume that a benign tumor on the tissue scale will develop if monoclonal conversion of wild-type cells by benign tumor cells within the homeostatic range of competition N is achieved. This assumption is based on experimental observations for example within the colon where mutant cells either go extinct or fixate in the colonic stem cell niche [5]. Moreover, mice injection experiments indicate that a critical number of tumor cells is needed for tumor formation [22–29] which suggests a point of no return on the cellular scale for tumor formation. If the first benign tumor cell progresses to a malignant tumor cell we assume that a malignant tumor will inevitably develop on the tissue scale which reflects a high fitness advantage of malignant tumor cells.

The state space of the underlying stochastic process of the model is $S = \{0, 1, 2, ..., N, E\}$ where states 0 to N represent the occurrence of the respective number of benign tumor cells without the occurrence of malignant tumor cells. State E indicates the presence of a malignant tumor cell. States N and E correspond to emergence of benign and malignant tumor subtypes and therefore to sequential and tunneling tumor progression, see also Figure 1. Both states N and E are absorbing states of the underlying stochastic process, see also Text S1 for details.

Dynamics in the model. In order to describe competition between cells and tumor cell progression, we adopt a MORAN model with mutations. This model class has mostly been investigated from a theoretical point of view [18, 30, 31]. Recently, we applied a MORAN model to evaluate tumor regression in pilocytic astrocytoma [19]. MORAN models are appropriate to describe a population of fixed size N which represents the homeostatic range of competition in our model. The dynamics is as follows. One cell is randomly

chosen to undergo cell death and is replaced by the offspring of another chosen cell. During proliferation, a genetic or epigenetic alteration can lead to tumor cell progression. Wild-type cells can progress to benign tumor cells with probability u and benign tumor cells progress to malignant tumor cells with probability v. We assume that initially all cells are wild-type cells. Hence, the process starts in state 0.

Analysis of the model

Choice of spatial cell arrangement. Theoretical studies demonstrated that the interplay between tissue structure, the population size N and mutation probabilities u and v in MORAN models are crucial for the dynamics of the model [18, 31, 32]. In particular, it has been shown that the absorption probability in state N on regular structures is the highest if all cells can potentially compete with each other and the lowest for a one-dimensional cell arrangement [18]. Since the tumor-originating cell type is unknown for most cancers also the spatial cell arrangement and realization of competition is unknown [4, 33]. Therefore, we consider a space-free and a one-dimensional cell arrangement in order account for this uncertainty by deriving a lower and an upper bound for the absorption probabilities. Figure 2 illustrates the MORAN dynamics on these two structures. For the precise definition of the underlying stochastic processes, see Text S1.

Tumor progression patterns in the model. Three parameter regimes within the model can be distinguished with respect to the tumor progression patterns. Within the *sequential fixation* regime, the benign tumor cell population is primarily able to reach size N before a benign tumor cell progresses to a malignant tumor cell. This regime corresponds to primarily sequential progression on the tissue scale. In the *tunneling* regime [30] a malignant clone will occur before the benign population is able to reach size N which corresponds to primarily tunneling progression in the model. In the *borderline* regime [32] both sequential fixation and tunneling are possible corresponding to both progression types on the tissue scale. An asymptotic classification of the model behavior with respect to these parameter regimes for large N has been theoretically derived in a space-free model [34] and in lattice-like cell arrangements [31]. We showed in a previous work that the exact progression pattern described by the absorption probability in state N in the space-free model solely depends on the so-called risk coefficient $\gamma := N\sqrt{v}$ [19]. Analogously, we derive here that the absorption probability in the one-dimensional model solely depends on a one-dimensional risk coefficient $\gamma_{1D} = N\sqrt[3]{v}$.

Choice of the parameter regime. In order to describe the frequency of sequential tumor formation, we derive the absorption probability of the underlying stochastic process in state N. We assume a parameter regime in which no additional mutations from wild-type cells to benign tumor cells occur as long as benign tumor cells are present in the system. The biological consequence of this assumption is that tumors form from the progeny of a single cell [35]. Formally, in the space-free model

$$Nu \ll 1 \tag{1}$$

and in the $1D$ model

$$Nu \ll \sqrt[3]{v} \tag{2}$$

ensures this model behavior, see [18]. For $u = v = 10^{-6}$, this assumption is valid if $N \ll 10^6$ in the space-free model and $N \ll 10^4$ in the 1D model.

Absorption probabilities. In [19], we have shown that the absorption probability in state N in the space-free model is given by

$$\alpha(\gamma) = \frac{1}{I_0(2\gamma)}, \tag{3}$$

where $\gamma := N\sqrt{v}$ and $I_n, n \in \mathbb{N}_0$, denote the modified BESSEL functions of the first kind, see [36]. We refer to the parameter γ as *space-free risk coefficient* because it determines the fraction of sequential progression.

In this work, we derive the absorption probability in state N of the 1D model. For this purpose, we utilize first step analysis in order to obtain a linear system of equations for the absorption probabilities in state N starting the process with $k, 1 \leq k \leq N$, benign tumor cells. Subsequently, CRAMER'S rule allows to derive the absorption probability starting with one benign tumor cell. Here, it is necessary to calculate two determinants and one of them is approximated by solving a second order difference equation with non-constant coefficients. Finally, we can approximate the absorption probability in state N of the 1D process starting with a single benign tumor cell as

$$\alpha^{\mathrm{1D}}(N, v)$$
$$= \frac{v(1-v)^{N-1}}{\pi \left(J_{N+\frac{2}{v}}\left(\frac{2}{v}\right) \left((2v+1)Y_{1+\frac{2}{v}}\left(\frac{2}{v}\right) - Y_{2+\frac{2}{v}}\left(\frac{2}{v}\right) \right) + Y_{N+\frac{2}{v}}\left(\frac{2}{v}\right) \left(J_{2+\frac{2}{v}}\left(\frac{2}{v}\right) - (2v+1)J_{1+\frac{2}{v}}\left(\frac{2}{v}\right) \right) \right)}, \tag{4}$$

where J denotes a BESSEL function of the first kind and Y a BESSEL function of the second kind, see [36]. A detailed derivation of equation (4) is provided in Text S1.

Numerical analysis of the 1D model indicates that the absorption probability in state N solely depends on the *1D risk coefficient* given by $\gamma_{1D} := N\sqrt[3]{v}$, see Figure S5. A detailed derivation of equation (4) and simulation results of the 1D model showing good accordance to this approximation is provided in Table S2.

Results

The range of competition determines tumor progression patterns. Our analysis allows to determine the progression patterns in both the space-free and the one-dimensional model in dependency of the competition range N. Interestingly, we find that a considerably small value of N corresponds to primarily tunneling progression in both the space-free and one-dimensional model. Moreover, the estimates of the parameter N largely depend on the

considered underlying spatial cell arrangement. In particular, the smaller the number of neighboring cells, the smaller is the estimated competition range. The estimated values for a mutation probability $v = 10^{-6}$ per cell division [37] are summarized in Table 1 and visualized in Figure 3. Note that these conclusions also hold for other values of v although a smaller value of v would increase and a larger value of v would decrease the estimates, see Table S3.

Fate of tumor development is decided in small tissue-specific tumor-originating niches. Our model allows to estimate the range of cellular competition N in different human tissues. For these estimations, we calibrate the space-free and $1D$ model with epidemiological data on the diagnosed fraction of benign and malignant tumor subtypes. We performed an extensive literature research to obtain these data which allow to estimate the risk coefficients $\gamma = N\sqrt{v}$ and $\gamma_{1D} = N\sqrt[3]{v}$ in the following way. We equal the clinically diagnosed fraction of benign tumors p with the absorption probabilities in state N given by equations [3] and [4] in the Material and Methods section. Subsequently, we numerically calculate the risk coefficients by evaluating the inverse of the absorption probability function at the diagnosed fraction of benign tumors p, i.e. $\gamma = \alpha^{-1}(p)$. The resulting estimates of the competition ranges in various tissues are provided in Table 2 and visualized in Figure 3. Our model predicts that the range of competition is considerably small compared to the overall number of cells in a tumor. Note that we do not assume any upper bound for the parameter N in our model. Moreover, although the estimates are considerably small, the range of competition largely depends on the tissue. For example, the estimated competition range within the liver is $383 - 2837$ cells whereas the estimates for the bone marrow are $18 - 31$ cells. Importantly, the estimate of the tumor-originating niche size for the human colon assorts well with the stem cell niche size in colonic crypts of about 40 cells [38] but surely less than 100 cells [39]. Overall, these results can be interpreted as existence of a tissue-specific tumor-originating niche in which the fate of tumor development is decided long before a tumor becomes detectable. Based on our results, we propose that a tumor-originating niche size of $291 - 1928$ cells within the human brain might be responsible for the clinically observed fraction of primary and secondary glioblastoma, see Table 2b).

Predicted tumor-originating niche size agrees well with mice injection experiments. Interestingly, the estimated tumor-originating niche sizes in Table 2 correspond very well to the necessary cell numbers for tumor induction in mice experiments [22–29], see also Figure 3. This observation supports two of our model assumptions. First, the possibility of tumor cell extinction for too small populations in these experiments shows that tumor and wild-type cells compete with each other. Second, the experimental data justify our model assumption that a sufficient number of tumor cells is needed to induce tumor formation on the tissue scale.

Discussion

On the tissue scale, one observes tumor progression types with and without detectable benign precursor stages. Data on the progression patterns with respect to the ratios of these progression types exhibit large differences between tissues. The underlying cellular processes causing these progression patterns are hardly observable and remain unclear. In this work, we shed light on the early phase of tumor development on the cellular scale with the help of a stochastic model. Our model is based on competition between wild-type and tumor cells and assumes that a sufficient amount of tumor cells is needed for tumor formation. We estimate this number by fitting the model to data on the diagnosed ratios of benign and malignant tumor subtypes. Our model predicts that this number is considerably small compared to the overall number of cells in a clinically detectable tumor and largely depends on the tissue which can be interpreted as existence of a tissue-specific tumor-originating niche. Hence, our results suggest that the fate of tumor development is decided long before a tumor becomes detectable.

Interestingly, our estimates of the tumor-originating niche size of about 39 cells for colon cancer agrees well with the number of stem cells found in one colonic crypt [38]. Indeed, it is the current understanding that colon adenomas and carcinomas develop within one colonic crypt with intestine stem cells likely to be the cell type of origin [40]. This demonstrates that our model might be utilized to predict tumor-originating niche sizes, thereby allowing to infer the potential cell type of tumor-origin for cancers from other tissues in which the origin is still elusive, e.g. for glioblastoma [41]. Glioblastoma can be divided into two classes dependent on the progression dynamics. In about 90% of cases, glioblastoma occur *de novo*, i.e. without evidence of a less malignant precursor lesion (primary glioblastoma) whereas 10% develop slowly by progression from low-grade gliomas (secondary glioblastoma). Using this data, our model predicts that the size of the tumor-originating niche from which glioblastoma develop is about 291 – 1928 cells. Neural stem cells (NSCs) of the subependymal zone (SEZ) have been suggested as a potential cell of origin for glioblastoma. Moreover, recent experimental evidence regarding NSCs in the SEZ of the adult brain suggests that the total number and fate of NSCs is regulated by a density-dependent mechanism [42]. Importantly, the finding in [42] that the fate of a NSC, e.g. activation or quiescence, is coupled to its neighbors perfectly fits to our hypothesis of cells competing within a certain range. Interestingly, the authors also suggest that the fate of active NSCs is coupled to the total number of neighboring NSCs in a shared locally restricted area which suggests that this area is a potential candidate for the tumor-initiating niche in the adult brain. It would be interesting to investigate if the range of coupled NSCs fits to our predicted size of the tumor-originating niche for glioblastoma.

From a modeling perspective, our analysis allows to distinguish different model regimes in Moran models not only asymptotically as in [18], but also for finite values of the system size N, see Table S3. and Figure S5. We find that the risk coefficient in both the space-free and the $1D$ model does not only distinguish the different model regimes, but quantifies the proportion of tunneling and sequential progression. Moreover, we show that the corresponding asymptotic results in dependency of the risk coefficient provide very good

approximations even for small population sizes N. This finding obviates the need of ad-hoc rules like $a \ll b$ if $a/b \leq 1/10$ to choose the appropriate parameter regime [43].

The potential existence of tumor-originating niches in which tumor fate is decided at an early stage of the cellular multistep process supports the view that cancer development is an ecological process [44, 45]. Ecology studies the dynamics of communities of species and their interactions and describe the origin of new species. From this point of view, the size of the tumor-originating niche might represent a critical effective population size that has to be reached by the progeny of the tumor-originating cell type in order to establish a tumor on the tissue scale. A deeper understanding of the processes and the origin of the tumor-originating niche contributes to the understanding of the early phase of tumor development. Further modeling and experimental effort is needed to understand this early phase of tumor development on the cellular scale in a better way. In this work, we demonstrated how observable quantities on the tissue scale might be utilized to achieve this goal.

Supplementary Material

S1   Text

The supplementary text contains the precise definitions of the MORAN models and detailed analytical derivations.

S2   Table

Comparison of the analytical approximation given and simulation results from 10000 trajectories of the underlying stochastic process of the one-dimensional model for the probability of absorption in state N. The results are also visualized in Figure S5

S3   Table

This table summarizes the risk coefficient regimes with respect to the different progression patterns of the model. Primarily sequential and primarily tunneling progression patterns refer to a fraction of 99.9% of sequential and tunneling progression, respectively.

S4   Table

This table summarizes regimes for the parameter N with respect to the different progression patterns of the models in dependency of the mutation probability v. Primarily sequential and primarily tunneling progression patterns refer to a fraction of 99.9% of sequential and tunneling progression, respectively.

S5   Figure

We numerically approximated the absorption probability in state N for different values of N and v such that the risk coefficient γ_{1D} is constant. This analysis suggests that the

absorption probability solely depends on the risk coefficient γ_{1D} for approximately $N \geq 40$. The squares indicate the results of simulation studies of the absorption probability in state N and therefore the benign tumor fraction in the model, see also Table S2

Funding

Acknowledgments

Conflict of interest statement. The authors declare that they have no conflict of interest.

References

1. Vogelstein B, Kinzler KW. The multistep nature of cancer. Trends in Genetics. 1993;9(4):138–141.

2. Williams MJ, Werner B, Barnes CP, Graham TA, Sottoriva A. Identification of neutral tumor evolution across cancer types. Nature Genetics. 2016;48(3):238–244.

3. Wu CI, Wang HY, Ling S, Lu X. The ecology and evolution of cancer: the ultra-microevolutionary process. Annual Review of Genetics. 2016;50:347–369.

4. Amoyel M, Bach EA. Cell competition: how to eliminate your neighbours. Development. 2014;141(5):988–1000.

5. Baker AM, Cereser B, Melton S, Fletcher AG, Rodriguez-Justo M, Tadrous PJ, et al. Quantification of crypt and stem cell evolution in the normal and neoplastic human colon. Cell Reports. 2014;8(4):940–947.

6. Lodish H, Berk A, Zipursky S, Matsudaira P, Baltimore D, Darnell J. Tumor cells and the onset of cancer. Molecular Cell Biology. 2000;.

7. LaBarge MA. The difficulty of targeting cancer stem cell niches. Clinical Cancer Research. 2010;16(12):3121–3129.

8. Qin H, Bao D, Tong X, Hu Q, Sun G, Huang X. The role of stem cells in benign tumors. Tumor Biology. 2016;37(12):15349–15357.

9. Kuang Y, Nagy JD, Eikenberry SE. Introduction to mathematical oncology. vol. 59. CRC Press; 2016.

10. Humphries A, Wright NA. Colonic crypt organization and tumorigenesis. Nature Reviews Cancer. 2008;8(6):415.

11. Vishwakarma A, Karp JM. Biology and Engineering of Stem Cell Niches. Academic Press; 2017.

12. Gilbertson RJ, Rich JN. Making a tumour's bed: glioblastoma stem cells and the vascular niche. Nature Reviews Cancer. 2007;7(10):733–736.

13. Persson AI, Petritsch C, Swartling FJ, Itsara M, Sim FJ, Auvergne R, et al. Non-stem cell origin for oligodendroglioma. Cancer Cell. 2010;18(6):669–682.

14. Visvader JE. Cells of origin in cancer. Nature. 2011;469(7330):314–322.

15. Landgren O, Kyle RA, Pfeiffer RM, Katzmann JA, Caporaso NE, Hayes RB, et al. Monoclonal gammopathy of undetermined significance (MGUS) consistently precedes multiple myeloma: a prospective study. Blood. 2009;113(22):5412–5417.

16. Ohgaki H, Kleihues P. The definition of primary and secondary glioblastoma. Clinical Cancer Research. 2013;19(4):764–772.

17. Moran PAP. Random processes in genetics. Math Proc Cambridge Philos Soc. 1958 1;54:60–71. Available from: http://journals.cambridge.org/article_S0305004100033193.

18. Durrett R, Moseley S. Spatial Moran models I. Stochastic tunneling in the neutral case. The Annals of Applied Probability. 2015;25(1):104.

19. Buder T, Deutsch A, Klink B, Voss-Böhme A. Model-based evaluation of spontaneous tumor regression in pilocytic astrocytoma. PLoS computational biology. 2015;11(12):e1004662.

20. Knudson AG. Mutation and cancer: statistical study of retinoblastoma. Proceedings of the National Academy of Sciences. 1971;68(4):820–823.

21. Vogelstein B, Kinzler KW. The genetic basis of human cancer. McGraw-Hill; 2002.

22. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. Proceedings of the National Academy of Sciences. 2003;100(7):3983–3988.

23. Civenni G, Walter A, Kobert N, Mihic-Probst D, Zipser M, Belloni B, et al. Human CD271-positive melanoma stem cells associated with metastasis establish tumor heterogeneity and long-term growth. Cancer Research. 2011;71(8):3098–3109.

24. Huang EH, Hynes MJ, Zhang T, Ginestier C, Dontu G, Appelman H, et al. Aldehyde dehydrogenase 1 is a marker for normal and malignant human colonic stem cells (SC) and tracks SC overpopulation during colon tumorigenesis. Cancer Research. 2009;69(8):3382–3389.

25. O'brien CA, Pollett A, Gallinger S, Dick JE. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. Nature. 2007;445(7123):106.

26. Lau WM, Teng E, Chong HS, Lopez KAP, Tay AYL, Salto-Tellez M, et al. CD44v8-10 is a cancer-specific marker for gastric cancer stem cells. Cancer Research. 2014;74(9):2630–2641.

27. Ginestier C, Hur MH, Charafe-Jauffret E, Monville F, Dutcher J, Brown M, et al. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell Stem Cell. 2007;1(5):555–567.

28. Yang ZF, Ho DW, Ng MN, Lau CK, Yu WC, Ngai P, et al. Significance of CD90+ cancer stem cells in human liver cancer. Cancer Cell. 2008;13(2):153–166.

29. Bao S, Wu Q, McLendon RE, Hao Y, Shi Q, Hjelmeland AB, et al. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. Nature. 2006;444(7120):756.

30. Michor F, Iwasa Y, Nowak MA. Dynamics of cancer progression. Nature Reviews Cancer. 2004;4(3):197–205.

31. Durrett R, Foo J, Leder K. Spatial Moran models, II: cancer initiation in spatially structured tissue. Journal of mathematical biology. 2016;72(5):1369–1400.

32. Durrett R, Schmidt D, Schweinsberg J. A waiting time problem arising from the study of multi-stage carcinogenesis. Ann Appl Probab. 2009 04;19(2):676–718. Available from: http://dx.doi.org/10.1214/08-AAP559.

33. Johnston LA. Competitive interactions between cells: death, growth, and geography. Science. 2009;324(5935):1679–1682.

34. Nowak MA, Komarova NL, Sengupta A, Jallepalli PV, Shih IM, Vogelstein B, et al. The role of chromosomal instability in tumor initiation. Proceedings of the National Academy of Sciences. 2002;99(25):16226–16231.

35. Davis A, Gao R, Navin N. Tumor evolution: Linear, branching, neutral or punctuated? Biochimica et Biophysica Acta (BBA)-Reviews on Cancer. 2017;1867(2):151–161.

36. Abramowitz M. Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables,. Dover Publications, Incorporated; 1974.

37. Drake JW, Charlesworth B, Charlesworth D, Crow JF. Rates of spontaneous mutation. Genetics. 1998;148(4):1667–1686.

38. Kang H, Shibata D. Direct measurements of human colon crypt stem cell niche genetic fidelity: the role of chance in non-Darwinian mutation selection. Frontiers in Oncology. 2013;3.

39. Yang J, Axelrod DE, Komarova NL. Determining the control networks regulating stem cell lineages in colonic crypts. Journal of Theoretical Biology. 2017;429:190–203.

40. Zeuner A, Todaro M, Stassi G, De Maria R. Colorectal cancer stem cells: from the crypt to the clinic. Cell stem cell. 2014;15(6):692–705.

41. Llaguno SRA, Parada LF. Cell of origin of glioma: biological and clinical implications. British journal of cancer. 2016;115(12):1445.

42. Basak O, Krieger TG, Muraro MJ, Wiebrands K, Stange DE, Frias-Aldeguer J, et al. Troy+ brain stem cells cycle through quiescence and regulate their number by sensing niche occupancy. Proceedings of the National Academy of Sciences. 2018;p. 201715911.

43. Durrett R, Schmidt D. Waiting for two mutations: with applications to regulatory sequence evolution and the limits of Darwinian evolution. Genetics. 2008;180(3):1501–1509.

44. Kareva I. What can ecology teach us about cancer? Translational Oncology. 2011;4(5):266–270.

45. Merlo LM, Pepper JW, Reid BJ, Maley CC. Cancer as an evolutionary and ecological process. Nature Reviews Cancer. 2006;6(12):924–935.

46. Vijay A, Elaffandi A, Khalaf H. Hepatocellular adenoma: An update. World Journal of Hepatology. 2015;7(25):2603.

47. Kerlikowske K. Epidemiology of ductal carcinoma in situ. Journal of the National Cancer Institute Monographs. 2010;2010(41):139–141.

48. Cymerman RM, Shao Y, Wang K, Zhang Y, Murzaku EC, Penn LA, et al. De novo vs nevus-associated melanomas: differences in associations with prognostic indicators and survival. JNCI: Journal of the National Cancer Institute. 2016;108(10).

49. Petrovchich I, Ford JM. Genetic predisposition to gastric cancer. In: Seminars in Oncology. vol. 43. Elsevier; 2016. p. 554–559.

50. Larjavaara S, Haapasalo H, Sankila R, Helén P, Auvinen A. Is the incidence of meningiomas underestimated? A regional survey. British Journal of Cancer. 2008;99(1):182.

51. Driman DK, Riddell RH. Flat adenomas and flat carcinomas: do you see what I see? Gastrointestinal Endoscopy. 1994;40(1):106–109.

Tables

progression patterns	$1D$ model	space-free model
primarily sequential	$N \leq 17$	$N \leq 29$
both sequential and tunneling	$17 < N \leq 528$	$29 < N \leq 4530$
primarily tunneling	$N > 528$	$N > 4530$

Table 1: Tumor progression patterns in dependency of the spatial cell arrangement in the model. For these estimates, we have chosen $v = 10^{-6}$. Primarily sequential and primarily tunneling progression patterns refer to a fraction of 99.9% of sequential and tunneling progression, respectively.

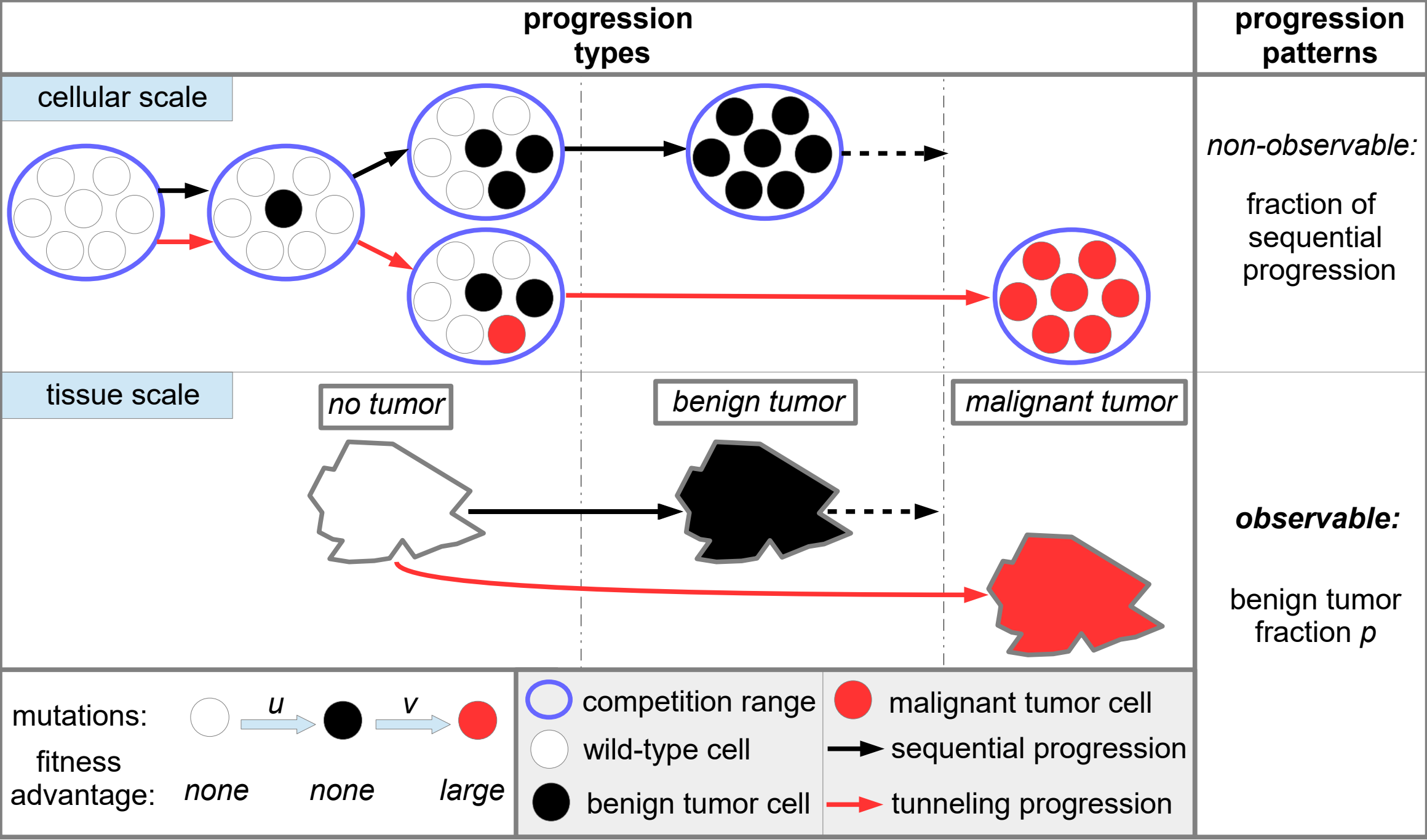|		epidemiological data			model predictions		mice data
	tissue	benign precursor	malignant tumor	benign fraction p	$N_{\text{space-free}}$	N_{1D}	needed cells for tumor formation in mice
a)	liver	hepatocellular adenoma	hepatocellular carcinoma	2% [46]	2837	383	1000 [28]
b)	brain	low-grade astrocytoma	glioblastoma	10% [16]	1928	291	500 [29]
c)	breast	ductal carcinoma in situ	invasive ductal carcinoma	20% [47]	1514	246	200 [22] – 500 [27]
d)	skin	nevus	melanoma	25% [48]	1375	230	1000 [23]
e)	stomach	gastric adenomas	gastric cancer	81% [49]	471	111	200 [26]
f)	meninges	benign meningioma	aggressive meningioma	95% [50]	227	67	NA
g)	colon	colonic adenoma	adenocarcinoma	99% [51]	100	39	25 [24] – 100 [25]
h)	bone marrow	MGUS	myeloma	99.9*% [15]	31	18	NA

Table 2: **Estimation of the homeostatic competition range N in different tissues.**
This table summarizes epidemiological data on benign and malignant tumor subtypes. We calibrate the model such that the absorption probabilities in state N given by equations (3) and (4) are equal to the diagnosed benign tumor fraction p. We assume a mutation probability of $v = 10^{-6}$ throughout these calculations. The last column contains the necessary number of injected cells to obtain a tumor in mice transplantation experiments. The estimates and mice data are also illustrated in Figure 3. * = estimated, MGUS=monoclonal gammopathy of undetermined significance.
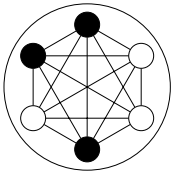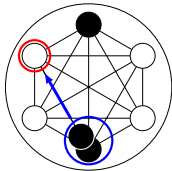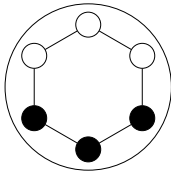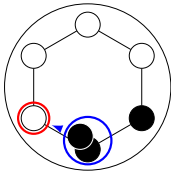
Legends to Figures

Figure 1: **Tumor progression types and patterns in the model.** On the cellular scale, wild-type cells can progress to benign tumor cells during proliferation with mutation probability u. Wild-type cells and benign tumor cells neutrally compete with each other within the homeostatic range of competition which is modeled by MORAN dynamics, see Figure 2. A benign tumor will manifest if all cells within the homeostatic range of competition are converted to benign tumor cells. Furthermore, benign tumor cells can progress to malignant tumor cells during proliferation with probability v. Then, a malignant tumor inevitably develops. These cellular dynamics lead to two distinct progression types at the tissue scale, namely sequential progression and tunneling progression. The benign tumor fraction p determines the progression pattern. We do not model the transition from benign tumors to malignant tumors since we are only interested in the type of progression (dotted arrow).

Figure 2: **Moran dynamics with different spatial cell arrangements.** In the MORAN dynamics, a randomly chosen cell proliferates (blue circle) and replaces a neighboring cell which undergoes cell death (red circle). In **A**, the space-free dynamics is illustrated, i.e. each cell can be replaced by any other cell. In **B**, only neighboring cells can be replaced representing a one-dimensional cell arrangement.

Figure 3: **Estimated tumor-originating niche sizes based on tumor progression patterns.** This plot shows the benign tumor fraction in the space-free (red) and one-dimensional (blue) model as function of the tumor-originating niche size. The blue curve has been numerically evaluated, see equation [4]. The red curve represents the plot of equation (3). The shaded areas illustrate the borderline parameter regimes, i.e. in which both sequential and tunneling progression are possible for the space-free and the $1D$ model, see Table 1. The dots indicate the estimated tumor-originating niche size and the squares represent experimental data for different tissues, see Table 2 for the values.

A

B