

WHY INCREASE OF HETEROGENEITY SIGNALS PRE-DETERIORATION DURING TUMOR PROGRESSION: A UNIFIED MATHEMATICAL MODEL*

YUANLING NIU[†], HAO KANG[‡], PINGYANG WANG[§], CHENCHEN GUO[§], FAN NIE[¶],
HONGBIN JI[§], JIARUI WU^{§||#}, AND LUONAN CHEN^{§#}

Abstract. Heterogeneity plays an important role in cancer genesis and progression. In this paper, we theoretically and computationally show that the increase of heterogeneity signals predegeneration during tumor progression by unified random ordinary differential equations (RODEs). Tumor formation results from a systematic change in organisms, and its causality is complex. Heterogeneity may be one of the key factors under some conditions. Specifically, the interactions between cell populations are modeled by random community matrices in normal tissues, benign tumors, and malignant tumors, and then we prove that the RODEs system describing the cell density of normal tissues or benign tumors becomes unstable as heterogeneity of cell populations or species increases. The increase of heterogeneity is an important signal. Heterogeneity can be viewed as a feature to distinguish benign tumors from malignant tumors under certain circumstances. Furthermore, we show that with the increase of heterogeneity, the divergence speed of the RODEs system becomes faster, which implies that malignant tumors with higher heterogeneity may develop faster and have poorer prognoses. Our theoretical findings can explain some noteworthy phenomena in various datasets in tumor patients and our biological experiments in mice from a mathematical viewpoint. Particularly, clinical data and mutation information from the cancer genome atlas and our experiments in mice revealed that tumors with higher heterogeneity usually show shorter survival time, whereas tumors with lower heterogeneity tend to have better prognosis, which also indicates that heterogeneity can be used as a potential biomarker in future clinical diagnosis. Targeting the heterogeneity may be a potential strategy for the cancer treatment.

Key words. random ordinary differential equations, random matrix, stability, heterogeneity, malignancy

MSC codes. 92-10, 92B05, 92C42

DOI. 10.1137/22M1497729

* Received by the editors May 23, 2022; accepted for publication (in revised form) November 3, 2022; published electronically April 28, 2023. Y. Niu, H. Kang and P. Wang contributed equally to this work.

<https://doi.org/10.1137/22M1497729>

Funding: This work was supported by National Natural Science Foundation of China (12071488, 11971488, 12131020, 31930022), National Key Research and Development Program of China (2022YFA1004800), Strategic Priority Research Program of the Chinese Academy of Sciences (XDB38040400, XDB38020000), Special Fund for Science and Technology Innovation Strategy of Guangdong Province (2021B0909050004, 2021B0909060002), Natural Science Foundation of Hunan Province (2020JJ2040), Major Key Project of Peng Cheng Laboratory (PCL2021A12), and JST Moonshot R&D (JPMJMS2021).

[†] School of Mathematics and Statistics, HNP-LAMA, Central South University, Changsha, 410083, People's Republic of China (yuanlingniu@csu.edu.cn).

[‡] Center for Applied Mathematics, Tianjin University, Tianjin, 300072, People's Republic of China (haokang@tju.edu.cn).

[§] Key Laboratory of Systems Biology, Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, Shanghai, 200031, People's Republic of China (wangpingyang6@sibcb.ac.cn, chenchen199111@hotmail.com, hbji@sibcb.ac.cn, wujr@sibs.ac.cn, lnchen@sibcb.ac.cn).

[¶] School of Computer Science and Engineering, Central South University, Changsha, 410083, People's Republic of China (niefan@csu.edu.cn).

^{||} School of Life Science and Technology, ShanghaiTech University, Shanghai, 201210, China.

[#] Key Laboratory of Systems Health Science of Zhejiang Province, School of Life Science, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Hangzhou, 310024, China.

1. Introduction. The global burden of cancer is considerable. The number of new cancer cases will reach 28.4 million worldwide by 2040, assuming that the incidence rate is unchanged [49]. Many scientists and doctors devote themselves to the pathogenesis of cancer. Furthermore, why do some benign neoplasms remain dormant for the lifetime of a patient, while others become malignant? This question remains poorly understood [31].

Some researchers realized that cancer is genetically related. Blokzijl et al. claimed that accumulation of genetic mutation of stem cells in human tissues during life is significantly associated with tumorigenesis [6]. Furthermore, some studies have revealed that the stochastic effects associated with DNA replication greatly affect the incidence rate of human cancers [52]. Cao et al. proposed a model which provides an explanation for differential global DNA methylation levels across cancer types and supports the observation that more malignant cancers tend to exhibit reduced DNA methylation levels [8]. Recently, Rozhok and Degregori claimed that cancers are age-dependent and are driven by age-altered systemic processes regulated above the cell level [40]. Some contexts, including smoking, other carcinogenic exposures, or certain inherited syndromes, lead to the emergence of abnormal cells containing mutations, which changes the tissue microenvironment. Palmer et al. proposed that the decline of the immune system is the major cause of cancer as an age-related disease [33].

Intratumor heterogeneity is widely observed in various types of tumor samples [12]. And it is found that heterogeneity is a distinguishing feature of malignant tissues [53, 57]. Furthermore, Schwarz et al. demonstrated that quantitative measures of intratissue heterogeneity may have predictive value for survival analysis after chemotherapy treatment [43]. In the past few years, several methods for inferring tumor clonal composition have been developed, including PyClone [39], THeTa2 [32], EXPANDS [3], etc.

Research into heterogeneity can provide a better understanding of the causes and progression of diseases. Although several studies have been conducted using specific mathematical models for description of tumor growth kinetics, for instance, exponential models [46], sigmoidal models [5, 22, 47], power law models [16], proliferation-invasion models [50, 58], models integrating angiogenesis [29], and immune system [7], the heterogeneity of tumors is rarely considered. West and Newton considered tumor heterogeneity in their model, but the model links heterogeneity to several different deterministic growth laws [59]. Zhou et al. proposed a mathematical framework of population dynamics which considers the growth rates and intercellular transitions between cancer cell populations [67]. This work can contribute to understanding the dynamics of the heterogeneity of cancer cells. All of the models mentioned above are only applicable to malignant tumors. For adult stem cells, Lei, Levin, and Nie investigated the role of heterogeneity in stem cell regeneration [24]. To our knowledge, however, there are few related studies giving unified models for normal tissues, benign tumors, and malignant tumors.

In this work, we introduce random ordinary differential equations (RODEs), which model the cell density of normal tissues, benign tumors, or malignant tumors, where the interactions between cell populations are described by a community matrix. For simplicity, we mainly focus on short-range interactions, long-range interactions, and self-inhibition of the cell populations for the community matrix. The RODEs model becomes a linear ODEs model when coefficients of the community matrix are determined in the initial choice of interactions. The ODEs can be considered to be obtained from the linearization around a feasible equilibrium of nonlinear equations describing the dynamics of the tissue. The equilibrium point is generally assumed to be zero. If not, we can do a translation transformation and change it to zero. We claim that

the increase of heterogeneity (in terms of cell populations or mutations) is not only a feature of malignant tissues, but also a signal for predeterioration during tumor progression and even a possible inducement of malignant tissues. Our work theoretically shows that the biological system will approach instability as the number of cell populations S increases in the tissue. Normal tissue or a benign tumor are supposed to be stable systems, while a malignant tumor is supposed to be an unstable system. It implies that normal tissues or benign tumors may become malignant tumors with the increase of heterogeneity. Causality of tumors is complex, and one possible factor is heterogeneity. Although we cannot experimentally prove that the increase of heterogeneity drives the cancer progression, it is an important signal for predeterioration or malignancy during tumor progression. Heterogeneity can be viewed as a feature to distinguish benign tumors from malignant tumors under certain circumstances, especially, since one of the tumor hallmarks is genome instability [15], which can result in heterogeneity.

Furthermore, a malignant tumor is increasingly regarded as a “tumor ecosystem.” Tumor cells cooperate with other tumor cells and host cells in their microenvironment, and could evolve to adapt to changing conditions [30, 38]. In view of this, we regard a malignant tumor as an ecosystem in which mutualism and competition coexist and model it by RODEs. We show that the divergence speed of the RODEs system becomes faster with the increase of heterogeneity, which implies that malignant tumors with higher heterogeneity may develop faster and have poorer prognoses. We study heterogeneity from a systematic point of view in this work and point out that heterogeneity may be the key node of tumorigenesis and progression.

Our theoretical findings can explain the phenomenon that predeterioration or malignancy is usually accompanied by the increased heterogeneity in various real datasets in tumor patients and our biological experiments in mice from a mathematical viewpoint. Based on our analysis of the data in [19], it is observed that the heterogeneity level of malignant tumors is higher than that of benign tumors in a study of pheochromocytomas (PCCs) and extraadrenal sympathetic paragangliomas (sPGLs). Our study on tumor drug resistance in the mice model indicated the tumor volume of malignant tumors with higher heterogeneity increases faster during drug treatment. Clinical data and mutation information from the cancer genome atlas (TCGA) reveals that patients with more Single Nucleotide Variants (SNVs) or more clone numbers associate with worse survival expectations. The branched evolution of the passenger mutations is considered the main source of clonal diversity [55]. Therefore, a larger number of SNVs indicates more passenger mutations, which reflects the increase of tissue heterogeneity. Tumors with more SNVs or clones are of higher heterogeneity and are related to shorter survival time, which is consistent with our theoretical findings and experiment results.

The paper is organized as follows. The model and experimental results are presented in section 2, and the discussion is given in section 3.

2. Model.

2.1. Model for normal tissues and benign tumors. Suppose cell populations in normal tissues are distributed regularly. Based on the existence of short-range interactions such as gap junctions in normal tissues and even cancerous cells of early stages [26, 59], we take the spatial structure into consideration in a normal tissue, and assume that a cell population in the tissue is mainly affected by itself and six neighbor cell populations (left, right, up, down, front, and back) in the three-dimensional space.

(In section SM2 of the supplemental material, we analyzed the situation when a cell population is mainly affected by itself and four neighbors (left, right, up, and down) in two-dimensional space.) Other interactions between nonadjacent cell populations including the long-range interactions mediated by neural or endocrine mechanisms are assigned from a distribution of random numbers. Randomness is only for the initial choice of the interactions, which then define a specific system. It is important to note that the actual interactions are not random. We hypothesize that the effect of cell populations on themselves and the interaction between adjacent cell populations play a major role in normal tissues, while long-range interaction is relatively small. For the sake of simplicity, we assume that there are n^3 cell populations in the tissue and they are arranged in a cube of length n as shown in Figure 1 while $n = 3$. Here, we suppose that in the tissue, for simplicity, the effect of a cell population on itself equals a , the effect on its left and right neighbors equals b , the effect on its up and down neighbors equals c , the effect on its front and back neighbors equals d . In fact, our model and results are also applicable to a more general case. It is shown in section 2.1.2 “Stability results in the more general case” that, when there are stronger interactions than the above assumption between some adjacent cell populations and stronger long-range interactions between some nonadjacent cell populations in the tissue, we can obtain similar results. The basic assumptions of the model are shown in Figure 1.

In view of the above assumptions and the similarity between benign tumors and normal tissues [37], we develop the following model for them. Suppose the cell densities in a tissue can be described by

$$(2.1) \quad \frac{dX(t)}{dt} = M_N X(t), \quad X(0) = \eta,$$

where the vector X describes the cell density of different cell populations, $X(0) = \eta$ is the initial value. M_N is a community matrix of size $S \times S$, M_{Nij} describes the effect that cell population j has on i . S is the number of cell populations. Actually, there are many functionally different subpopulations in one cell population. Consequently, S may be very large. In our model, we suppose that M_N is a random matrix, and can be written as $M_N(\omega) = D_N + R_N(\omega)$, $\omega \in \Omega$. ω is usually omitted for simplicity but, in fact, the existence and influence of ω is essential. D_N is a deterministic matrix, which describes the short-range interactions including the interaction between adjacent cell populations and the effect of cell populations on themselves. Usually, $D_{Nii} = a$ is assumed to be negative. R_N is a random matrix which describes the long-range interactions. Let $(\Omega, \mathcal{F}, P)$ be a probability space, where Ω is a sample space, $\mathcal{F}$ is a σ -algebra on Ω , and P is a probability measure. R_{Nij} is a real-valued random variable on Ω . $R_{Nij} \sim N(0, \sigma^2)$. σ is the standard derivation, which can be considered to represent the average strength of long-range interactions. It is relatively small compared with other effects in the tissue and the standard derivation of R_T in the model for malignant tumors. Please see section 2.3 for more details about R_T . Apparently, (2.1) is a RODEs model with random coefficients.

The form of the matrix D_N depends on how the cell populations are indexed in a cube. We give a three-dimensional (3D) lattice model for a tissue with 27 cell populations in Figure 1. On the same plane, cell populations are indexed from left to right and from top to bottom. Cell populations on different planes are indexed

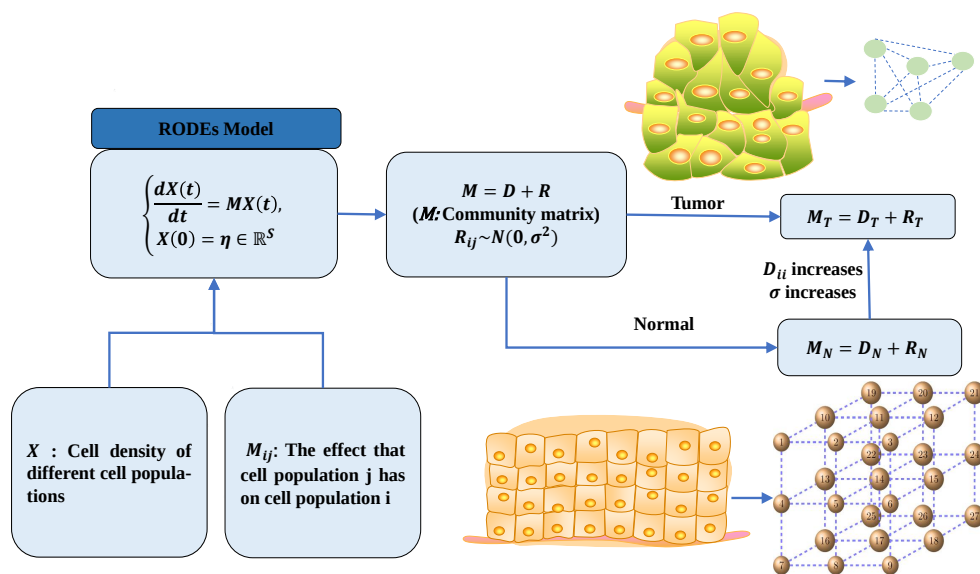


FIG 1. The mathematical description by RODEs of normal tissues and malignant tumors. Because of the different interactions between cell populations as well as the different distribution of cell populations, the community matrices M for normal tissues and malignant tumors are different. For normal tissues, a deterministic matrix D_N describing the short-range interactions including the interaction between adjacent cell populations and the effect of cell populations on themselves is listed separately. Usually, $D_{N,ii}$ is assumed to be negative. Meanwhile, the long-range interactions are described by a random matrix R_N with small standard deviation σ . The community matrix is $M_N = D_N + R_N$. Randomness only enters in the initial choice of $R_{N,ij}$, which then define a specific system. σ is employed here to describe the interaction strength. For simplicity, benign tumors are considered to have the same model with normal tissues because there are many similarities between them [37]. It is assumed that in the process of carcinogenesis, the self-inhibition of the cell populations is weakened [34], and the long-range interaction is enhanced [59]. D_T in the community matrix $M_T = D_T + R_T$ for malignant tumors is a diagonal matrix and represents the effect of cell populations on themselves. R_T here is a random matrix with larger σ representing interactions between adjacent cell populations and long-range interactions. S is the number of variables in the RODEs model, it describes the number of cell populations in the tissue. Larger S implies higher heterogeneity inside the tissue.

from front to back. For the tissue with 27 cell populations as shown in Figure 1, the deterministic part D_N of the community matrix is a symmetric matrix with

$$\begin{aligned}
 \text{diag}(D_N, 0) &= [a, a, a, a, a, a, a, a, a, a, a, a, a, a, a, a, a, a, a, a, a, a, a, a, a, a, a], \\
 \text{diag}(D_N, 1) &= [b, b, 0, b, b, 0, b, b, 0, b, b, 0, b, b, 0, b, b, 0, b, b, 0, b, b, 0, b, b], \\
 \text{diag}(D_N, 3) &= [c, c, c, c, c, c, 0, 0, 0, c, c, c, c, c, c, 0, 0, 0, c, c, c, c, c, c, c, c], \\
 \text{diag}(D_N, 9) &= [d, d, d, d, d, d, d, d, d, d, d, d, d, d, d, d, d, d, d, d, d, d, d, d, d], \\
 (2.2) \quad \text{diag}(D_N, i) &= [0, 0, \dots, 0], \quad i = 2, 4, 8, 10, 11, \dots, 26,
 \end{aligned}$$

where $\text{diag}(D_N, i)$ denotes the collection of elements in the diagonal of a distance of i units with a main diagonal; for example, see (SM2.2) and (SM2.3) in the supplemental material. For a more general case with n^3 cell populations, the deterministic part D_N is a symmetric matrix with:

$$\text{diag}(D_N, 0) = [\underbrace{a, \dots, a}_{n^3}],$$

$$\begin{aligned}
 \text{diag}(D_N, 1) &= \underbrace{[\underbrace{b, \dots, b}_{n-1}, 0, \underbrace{b, \dots, b}_{n-1}, 0, \dots, \underbrace{b, \dots, b}_{n-1}]}_{n^2 \text{ sets of } \underbrace{b, \dots, b}_{n-1}}, \\
 \text{diag}(D_N, n) &= \underbrace{[\underbrace{c, \dots, c}_{n(n-1)}, 0, \dots, 0, \underbrace{c, \dots, c}_{n(n-1)}, 0, \dots, 0, \dots, \underbrace{c, \dots, c}_{n(n-1)}]}_{n \text{ sets of } \underbrace{c, \dots, c}_{n(n-1)}}, \\
 (2.3) \quad \text{diag}(D_N, n^2) &= [\underbrace{d, \dots, d}_{n^2(n-1)}],
 \end{aligned}$$

and other elements are 0.

Remark 1. Noting that in normal tissues and benign tumors, we assume that long-range interactions are relatively small compared with short-range interactions. In other words, the absolute values of corresponding elements in R_N are much smaller than the ones in D_N . Thus, in the positions of M_N representing interactions between adjacent cell populations and the effect of the cell populations on themselves, the values of R_N can be ignored.

2.1.1. Stability results when the community matrix M_N is symmetric.

From (2.2) and (2.3) we find that matrix D_N can be approximated as a symmetric Toeplitz matrix as n increases. A Toeplitz matrix is an $n \times n$ matrix $T_n = [t_{i,j}; i, j = 1, 2, \dots, n]$, where $t_{i,j} = t_{i+1,j+1} = t_{i-j}$, i.e., a matrix of the form

$$(2.4) \quad \begin{pmatrix} t_0 & t_{-1} & t_{-2} & \cdots & \cdots & t_{-(n-1)} \\ t_1 & t_0 & t_{-1} & \cdots & \cdots & \cdots \\ t_2 & t_1 & t_0 & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ t_{n-1} & \cdots & \cdots & \cdots & \cdots & t_0 \end{pmatrix}.$$

The $n \times n$ Toeplitz matrix is associated with a function f in the sense of

$$(2.5) \quad f(\lambda) := \lim_{n \rightarrow \infty} \sum_{k=-n}^n t_k \exp(\mathbf{i}k\lambda),$$

where $\mathbf{i}$ is the imaginary unit. Furthermore, if

$$(2.6) \quad \sum_{k=-\infty}^{\infty} |t_k| < \infty$$

holds, $f \in L^1([0, 2\pi])$, and t_k can be recovered from f ,

$$(2.7) \quad t_k = \frac{1}{2\pi} \int_0^{2\pi} f(\lambda) \exp(-\mathbf{i}k\lambda) d\lambda.$$

Under (2.6), we also have

$$(2.8) \quad \lim_{n \rightarrow \infty} \left| f(\lambda) - \sum_{k=-n}^n t_k \exp(\mathbf{i}k\lambda) \right| = 0.$$

According to [14], we obtain the bounds on the eigenvalues of Hermitian Toeplitz matrices by the following.

LEMMA 2.1 (see [14]). Suppose f is a real-valued function that is associated with the Toeplitz matrix. We define the essential supremum $M_f = \text{esssup } f$ of f as the smallest number α for which $f(x) \leq \alpha$ except on a set of total length of measure 0. Similarly, the essential infimum $m_f = \text{essinf } f$ is defined as the largest value of α for which $f(x) \geq \alpha$ except on a set of total length of measure 0. Let $\tau_{n,k}$ be the eigenvalues of a Toeplitz matrix. If the Toeplitz matrix is Hermitian, then

$$(2.9) \quad m_f \leq \tau_{n,k} \leq M_f.$$

Therefore, we expect that eigenvalues of D_N distribute on a real interval and the interval is independent of the order S of the matrix according to Lemma 2.1, when S is relatively large.

LEMMA 2.2 (see [51]). Let A_n be the $n \times n$ Hermitian random matrix whose upper diagonal entries are independent and identically distributed (i.i.d.) complex random variables with mean 0 and variance 1. Then the empirical spectral distribution of $\frac{1}{\sqrt{n}}A_n$ converges (both in probability and in the almost sure sense) to the semicircle distribution.

According to Lemma 2.2, the eigenvalues of R_N are real and distribute on the interval $[-2\sigma\sqrt{S}, 2\sigma\sqrt{S}]$ when R_N is a real symmetric random matrix, where S is the order of the matrix.

Furthermore, according to Corollary IV.4.9 of the book [48], we have the following.

LEMMA 2.3 (see [48]). Let A denote a Hermitian matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. And $\bar{A} = A + E$ denotes a Hermitian perturbation of A with eigenvalues $\bar{\lambda}_1 \geq \bar{\lambda}_2 \geq \dots \geq \bar{\lambda}_n$. Let the eigenvalues of E be $\epsilon_1 \geq \epsilon_2 \geq \dots \geq \epsilon_n$. Then for $i = 1, \dots, n$, $\bar{\lambda}_i \in [\lambda_i + \epsilon_n, \lambda_i + \epsilon_1]$.

On the basis of Lemmas 2.1, 2.2 and 2.3, we have the following theorem.

THEOREM 2.4. Assume that in the system described by (2.1), a cell population is mainly affected by itself and 6 neighbors. The community matrix can be written as $M_N = D_N + R_N$. The effect of a cell population on the cell population itself equals a , the effect on its left and right neighbors equals b , the effect on its up and down neighbors equals c , the effect on its front and back neighbors equals d . Suppose there are n^3 cell populations in the system and they are arranged in a cube of length $L = n$. R_N is a symmetric random matrix, and $R_{Nij} \sim N(0, \sigma^2)$ ($i \leq j$). It describes the long-range interactions between cell populations. Then the largest eigenvalue of the community matrix M_N belongs to the interval $[2\sigma\sqrt{S} + \lambda_{\min}, 2\sigma\sqrt{S} + \lambda_{\max}]$. Here, $S = n^3$, $\lambda_{\min} = a - 2|b| - 2|c| - 2|d|$, and $\lambda_{\max} = a + 2|b| + 2|c| + 2|d|$. $\lambda_{\max}$ and $\lambda_{\min}$ are the maximum and minimum eigenvalue of D_N , respectively, and they are independent of S for matrices with relatively large order.

Proof. As mentioned previously, matrix D_N can be approximated as a symmetric Toeplitz matrix for large n . The corresponding Toeplitz matrix has 7 nonzero descending diagonals for a 3D system where a cell population is mainly affected by itself and 6 neighbors; please see (2.3) for more details. According to (2.5) and (2.8), for a Toeplitz matrix with 7 such nonzero descending diagonals, we obtain the associated function f in the following:

$$\begin{aligned} f(\lambda) &= \lim_{n \rightarrow \infty} \sum_{k=-n}^n t_k \exp(\mathbf{i}k\lambda) \end{aligned}$$

$$\begin{aligned}
&\approx a \exp(\mathbf{i} \times 0 \times \lambda) + b \exp(\mathbf{i} \times 1 \times \lambda) + b \exp(\mathbf{i} \times (-1) \times \lambda) + c \exp(\mathbf{i} \times n \times \lambda) \\
&\quad + c \exp(\mathbf{i} \times (-n) \times \lambda) + d \exp(\mathbf{i} \times n^2 \times \lambda) + d \exp(\mathbf{i} \times (-n^2) \times \lambda) \\
&= a + b(\cos(\lambda) + \mathbf{i} \sin(\lambda) + \cos(-\lambda) + \mathbf{i} \sin(-\lambda)) + c(\cos(n\lambda) + \mathbf{i} \sin(n\lambda) + \cos(-n\lambda) \\
&\quad + \mathbf{i} \sin(-n\lambda)) + d(\cos(n^2\lambda) + \mathbf{i} \sin(n^2\lambda) + \cos(-n^2\lambda) + \mathbf{i} \sin(-n^2\lambda)) \\
&= a + 2b \cos \lambda + 2c \cos(n\lambda) + 2d \cos(n^2\lambda).
\end{aligned}$$

Noting that $-1 \leq \cos \lambda, \cos(n\lambda), \cos(n^2\lambda) \leq 1$, and $|b|, |c|, |d|$ are nonnegative, we obtain

$$(2.10) \quad a - 2|b| - 2|c| - 2|d| \leq f(\lambda) \leq a + 2|b| + 2|c| + 2|d|.$$

According to Lemma 2.1, the eigenvalues of D_N have lower bound λ_{min} and upper bound λ_{max} , where

$$(2.11) \quad \lambda_{min} = a - 2|b| - 2|c| - 2|d|$$

and

$$(2.12) \quad \lambda_{max} = a + 2|b| + 2|c| + 2|d|.$$

In addition, we obtain from Lemma 2.2 that the eigenvalues of R_N are real and distribute on the interval $[-2\sigma\sqrt{S}, 2\sigma\sqrt{S}]$ when S is relatively large.

It is noted that both D_N and R_N are real symmetric matrices. Thus they are also Hermitian matrices. Consequently, from Lemma 2.3, we obtain that the largest eigenvalue of the community matrix M_N belongs to the interval $[2\sigma\sqrt{S} + \lambda_{min}, 2\sigma\sqrt{S} + \lambda_{max}]$. $\square$

Remark 2. The semicircular law is performed in the limit of large scale matrices, and distributions of eigenvalues are expected to converge in systems of several hundreds (or larger) of species. Furthermore, eigenvalues of Hermitian Toeplitz matrices have bounds that are independent of the order according to Lemma 2.1. Matrix D_N in Theorem 2.4 can be approximated as a symmetric Toeplitz matrix as S increases. Thus, it is expected that their eigenvalues distribute on a real interval and the interval is independent of the order of the matrix, when the order is relatively large. Based on the above-mentioned factors, Theorem 2.4 is performed for relatively large systems.

Let $a = -90, b = -11, c = 12, d = -10$ in the matrix D_N . We can see from Table SM1 and Figure SM1 (in the supplemental material) that when S is large, the eigenvalues almost distribute on the interval $[-156, -24]$ which coincides with our theoretical result $\lambda_{min} = a - 2|b| - 2|c| - 2|d|$ and $\lambda_{max} = a + 2|b| + 2|c| + 2|d|$ as given in Theorem 2.4.

One can see from Theorem 2.4 that the largest eigenvalue of M_N increases as the order S increases. Furthermore, we can obtain the stability condition is

$$(2.13) \quad 2\sigma\sqrt{S} + \lambda_{max} < 0.$$

In addition, it is obtained that the instability condition is

$$(2.14) \quad 2\sigma\sqrt{S} + \lambda_{min} > 0.$$

We illustrate the above theoretical findings in the following. The initial values of the system are set to be $X(0) = [1000, 500, 800, 50, 50, \dots, 50]^T$. Let $a = -90, b =$

$-11, c = 12, d = -10$ in D_N again. According to Table SM1, Figure SM1 (in the supplemental material) and Theorem 2.4, the instability condition for such a system is $2\sigma\sqrt{S} - 156 > 0$, namely, $S > 67600$ when $\sigma = 0.3$.

And the stability condition is $2\sigma\sqrt{S} - 24 < 0$, namely, $S < 1600$ when $\sigma = 0.3$. Then we apply the backward Euler scheme to (2.1) with step size $h = 0.005$ and obtain numerical solutions with $S = 1331$ and $S = 27000$ on Figure 2(c). There are too many components in the solution so we plot $\|X(t)\|$ which is the infinity norm of the solution at time t , i.e., $\|X(t)\| = \max\{|X_i(t)|, i = 1, 2, \dots, S\}$. M_N is random, but given the fact that the simulations are similar for different ω , only one-path simulations are depicted in Figure 2 for different cases. One can see from Figure 2(c) that systems become unstable with increasing S . It should be pointed out that we present sufficient and unnecessary conditions in (2.13) and (2.14), thus the system keeps stable when $S = 8000$ and becomes unstable when $S = 27000$ as shown in Figure 2. In addition, from numerous simulations and Theorem 2.4 we find that, for sufficiently large S , any selection of diagonal elements will only lead to the shift of the eigenvalue distribution, and has no obvious effect on its shape. Namely, λ_{min} and λ_{max} increase as the diagonal elements a of D_N increase. The largest eigenvalue of M_N increases accordingly. Combining this with (2.13) and (2.14), we conclude that the system approaches instability as a and the interaction strength σ increase, which are confirmed in Figures 2(a), (b), respectively.

Remark 3. From (2.11), (2.12), (2.13), and (2.14), one can infer that the system is stable when $2\sigma\sqrt{S} + a + 2|b| + 2|c| + 2|d| < 0$; it is unstable when $2\sigma\sqrt{S} + a - 2|b| - 2|c| - 2|d| > 0$. In other words, the stability range of parameter a is $a < -2\sigma\sqrt{S} - 2|b| - 2|c| - 2|d|$ based on given values of σ and S . When $a > -2\sigma\sqrt{S} + 2|b| + 2|c| + 2|d|$, the system becomes unstable. Letting $\sigma = 0.3$, $S = 8000$, and $a = -90, b = -11, c = 12, d = -10$, we obtain that the system is stable when $a < -120$. As mentioned previously, (2.13) is a sufficient and unnecessary condition to keep the system stable. Therefore, Figure 2(a) shows that the system keeps stable when $a = -90$.

Remark 4. “ $X(t)$ approaches zero” implies that the system returns to the equilibrium after small perturbations and the tissue stays in the normal state while “ $X(t)$ approaches infinity” implies that the system moves away from the equilibrium after small perturbations and the tissue changes to the diseased state from the normal state. Our model is a qualitative model, not a quantitative model. $X(t)$ does not necessarily represent the real cell density of different cell populations. In this work, we intend to provide a theoretical explanation for our hypothesis qualitatively by using the concept of stability in differential equations.

2.1.2. Stability results in the more general case. It is more general when the community matrix $M_N = D_N + R_N$ is asymmetric. According to [51] and [13], we have Conjecture 1 which gives the limiting distribution of eigenvalues for more general community matrices.

CONJECTURE 1. For each S , let D_N be a deterministic $S \times S$ matrix satisfying

$$(2.15) \quad \sup_S \frac{1}{S^2} \|D_N\|_2^2 < \infty.$$

Let $M_N = D_N + R_N$ with R_N being a random matrix whose elements are i.i.d. drawings from a distribution with mean 0 and variance σ^2 . Then the support of $\frac{M_N}{\sigma\sqrt{S}}$ converges towards the given set of $z \in \mathbb{C}$ for which

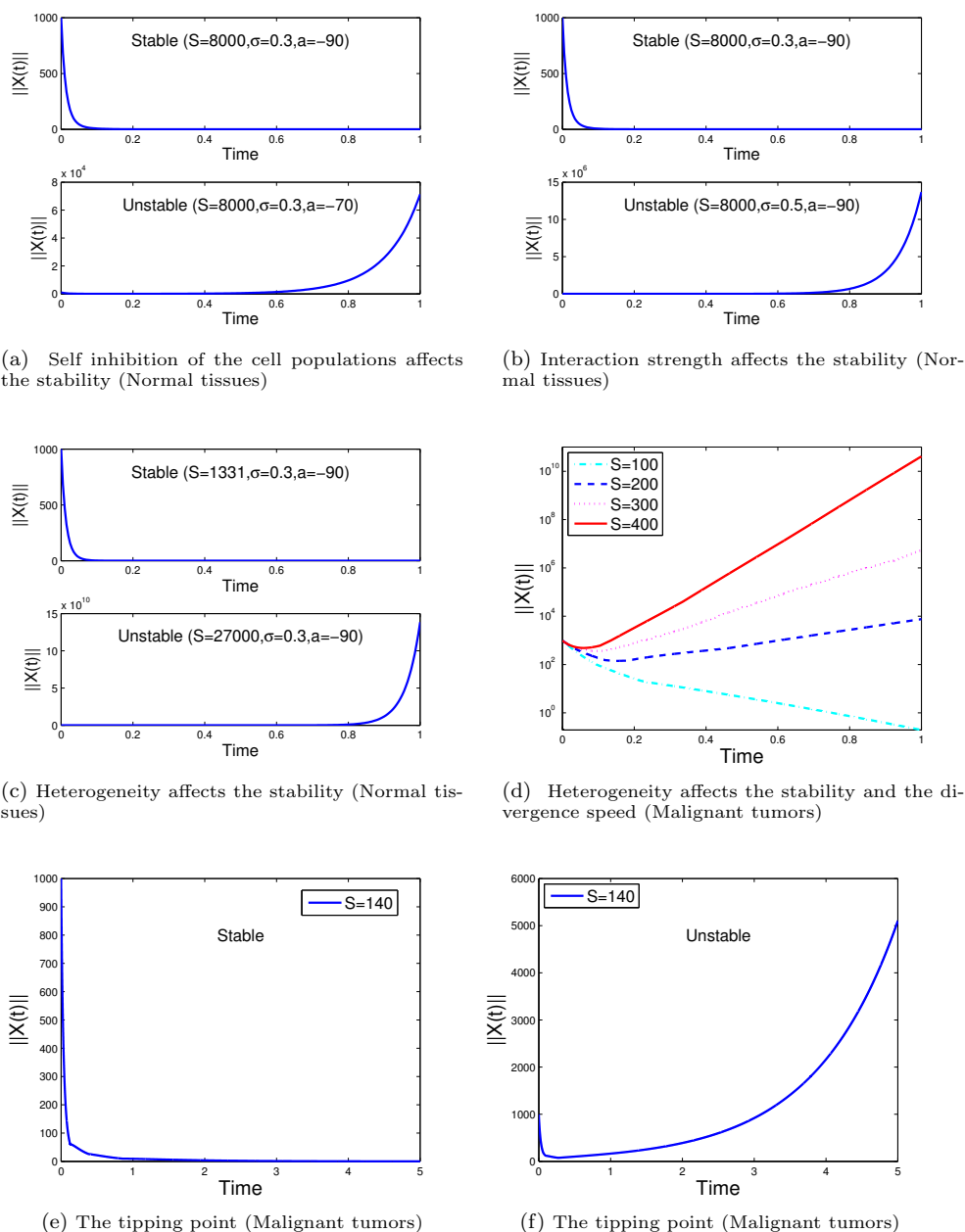


FIG 2. Heterogeneity, the self-inhibition of the cell populations a and interaction strength σ affect the phase transition (from stable state to unstable state, which implies the tissue is in a normal state or in a diseased state, respectively) of the system (RODEs (2.1)) describing normal tissues; see subfigures (a)-(c). Meanwhile, heterogeneity affects the stability and the divergence speed of the system (RODEs (2.18)) describing malignant tumors; see subfigure (d). In the early stages, the tumor is still at the primary site, there is no macrovascular invasion, and the disease is in a relatively stable and controllable state compared with later stages of tumors. The system becomes unstable with higher heterogeneity and the disease enters the deteriorating state. Systems with S such that $\text{Re}(\lambda_{max}^{M_T(\omega)}) \approx 0$ can be regarded facing a “tipping point”—the system is sometimes stable and sometimes unstable; see subfigures (e),(f). Subfigures (a)–(c) give the dynamics of RODEs (2.1) with different a , σ , and S , respectively. Here, $b = -11, c = 12, d = -10$ in the deterministic matrix D_N . (d)–(f) show dynamics of RODEs (2.18) with different S . Here, $a_T = -30, \sigma = 2, C_1 = 0.6$. $\|X(t)\|$ is the infinity norm of the solution.

$$(2.16) \quad \int |z - x|^{-2} d\mu(x) \geq 1,$$

where μ is the limiting distribution of the empirical spectral distribution of $\frac{D_N}{\sigma\sqrt{S}}$. Specially, if the eigenvalues of $\frac{D_N}{\sigma\sqrt{S}}$ are $\lambda_i^{\frac{D_N}{\sigma\sqrt{S}}}$ with multiplicities $\alpha_i, i = 1, \dots, k$ and $\sum_{i=1}^k \alpha_i = S$, then the limit support of the empirical spectral distribution of $\frac{M_N}{\sigma\sqrt{S}}$ is given by

$$(2.17) \quad \int |z - x|^{-2} d\mu(x) = \frac{1}{S} \sum_{i=1}^k \frac{\alpha_i}{|z - \lambda_i^{\frac{D_N}{\sigma\sqrt{S}}}|^2} \geq 1.$$

Suppose that R_N is an $S \times S$ matrix containing values drawn from the normal distribution. According to the circular law [51], its eigenvalues are distributed on a disk whose radius increases in pace with S . Furthermore, the deterministic matrix D_N can be approximated as a Toeplitz matrix when S is relatively large as mentioned previously. Consequently, its eigenvalues have bounds according to Lemma 2.1. In addition, we find from numerous simulations and Conjecture 1 that the distribution of the eigenvalue of M_N is a superposition of those of D_N and R_N , and give examples in Figures SM2 and SM3 in the supplemental material. Denote the eigenvalue of M_N with the largest real part by $\lambda_{max}^{M_N}$, i.e., $Re(\lambda_{max}^{M_N}) = \max_{1 \leq i \leq S} Re(\lambda_i^{M_N})$. Then we conjecture that $Re(\lambda_{max}^{M_N})$ increases as S increases. By numerical simulations we do find that $Re(\lambda_{max}^{M_N})$ increases along with S , which can be seen from Figure 3 and Figures SM2 and SM3 in the supplemental material. Figure SM4 in the supplemental material shows that one can obtain similar results when the effect on the cell populations themselves is different. Please see section SM3 in the supplemental material for more details about the eigenvalue distribution of M_N .

b, c, d in (2.3) are negative which means that the surrounding cell populations restrain the growth of a cell population. They are positive means that the surrounding cell populations promote the growth of the cell population. They are zero when the surrounding cell populations almost do not affect the growth of the cell population.

Figure 3(a) reveals that $Re(\lambda_{max}^{M_N})$ increases as S increases while the surrounding cell populations restrain the growth of the cell population. Next, we let the interaction strength to the surrounding 6 cell populations of some cell populations in the system be 10 times stronger than that in (2.3), and see what will happen. Obviously, we can find similar results; see Figures SM5 and SM6 (in the supplemental material). Moreover, Figure SM6 shows that we can obtain similar results when there are several significant long-range interactions in the tissue.

Let b, c, d in (2.3) be positive which means that the surrounding cell populations promote the growth of the cell population; we obtain that the system approaches instability when S increases in Figure 3(b). Especially, the system cannot be stable while the cell population promotes the growth by itself, that is, a is also positive. The reason is that the trace of matrix M_N is positive since its diagonal elements are positive. If the real parts of all eigenvalues are negative, the trace cannot be positive since it equals the sum of all eigenvalues. Therefore, there is at least one eigenvalue that has positive real part. Accordingly, the system cannot be stable in such a case; see Figure SM7 given in the supplemental material.

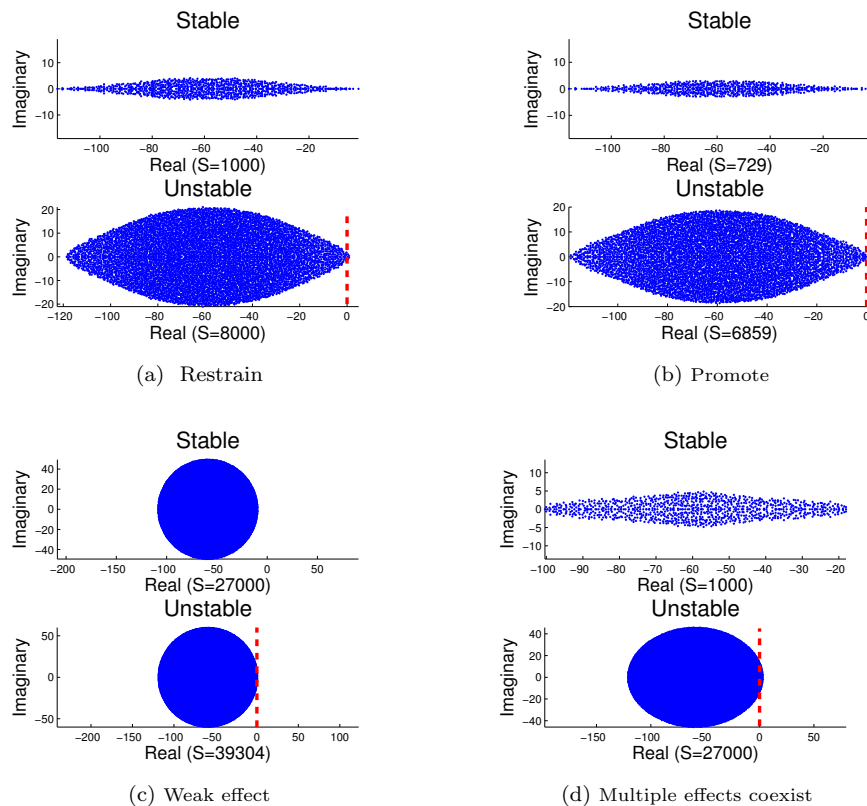


FIG 3. RODEs model (2.1) for normal tissues approaches instability as heterogeneity (described by number of cell populations S) increases no matter what the effect from other cell populations is. Furthermore, different cell population interactions affect the stability region. Keeping other test conditions constant, a system with weak cellular interaction is more stable and will stay stable for longer time as shown in subfigure (c) compared with systems in which cell populations are affected by surrounding neighbors as shown in subfigures (a), (b), and (d). Subfigures (a)–(d) give eigenvalues of community matrices $M_N = D_N + R_N$ with different S . System (2.1) is stable when the real parts of all eigenvalues of M_N are negative. When one or several of the eigenvalues have positive real parts, the system is unstable. R_N is an $S \times S$ matrix containing values drawn from the normal distribution with $\sigma = 0.3$. (a): The parameters in D_N are $a = -59, b = -11, c = -9, d = -10$ which mean that surrounding cell populations restrain the growth of the cell population. (b): The parameters in D_N are $a = -59, b = 11, c = 9, d = 10$, which mean that the surrounding cell populations promote the growth of the cell population. (c): The parameters in D_N are $a = -59, b = 0, c = 0, d = 0$, which mean that surrounding cell populations almost do not affect the growth of the cell population. (d): The parameters in D_N are $a = -59, b = 11, c = 0, d = -10$ which mean that some surrounding neighbors restrain the growth of the cell population, while other surrounding neighbors promote or almost do not affect its growth.

We set b, c, d in (2.3) to be 0, which means that cell populations almost do not affect each other, i.e., there exists a weak effect between adjacent cell populations like the weak long-range interaction in the assumption, and see what will happen. From Figure 3(c) we find that the system will also approach instability while S increases. But such a system is more stable and will stay stable for a longer time compared with systems in which cell populations are affected by surrounding neighbors, while other test conditions (a and σ) are kept constant. In other words, the system will stay stable until S increases to a greater extent compared with Figures 3(a), (b), and (d).

And finally we let some parameters in (2.3) be negative, some be positive, and some take values of 0, which means that some surrounding neighbors restrain the growth of the cell population, while other surrounding neighbors promote or almost do not affect its growth. From Figure 3(d) we find that the system will approach instability while S increases. Figure SM8 (in the supplemental material) shows that we can obtain similar results when there are several significant long-range interactions and the interaction strength to surrounding neighbors of some cell populations in the system are 10 times stronger than that in (2.3).

Remark 5. There are different relationships within a tissue depending on the components of its cells and the strength of their interaction. The degree of cell population to cell population connection is also different depending on the different cell population interaction patterns within different tissues. It can be simulated into two different situations that there is an interaction between adjacent cell populations as shown in Figures 3(a), (b), and (d), or there does not exist any connection between adjacent cell populations (which corresponds to weak interaction between two neighbors) as shown in Figure 3(c). There are complex intercellular interactions between cell populations, which constitute various cellular microenvironments. Important cellular microenvironments in the human body include: tumor microenvironment, immune cell microenvironment, stem cell microenvironment, synaptic microenvironment, and virus-host cell interactions, etc. The specific proportion of cell distribution in the tumor microenvironment [61, 62] is a key factor in determining the direction of progression. Different rates and directions of tumor initiation, progression, and metastasis process, are a result of the continuous crosstalk between different cell individuals or other components in the tissue microenvironment. For example, there are two different types of immune T cells, Treg and T-effector. Treg, as a regulator, regulates the effect of T-cell-related immune response [60]. A reduced T-effector/Treg ratio was thus a strong signature of the lung tumor lesion [23]. Moreover, there are two different types of macrophages named M1 and M2. M1 macrophages are macrophages that produce proinflammatory cytokines, known as classic macrophages which play a role in suppressing cancer cell growth in tumor tissue. And M2 macrophages are also known as alternative activated macrophages. They have a role in helping cancer cells evade immune surveillance in cancer tissue. These interregulatory relationships between immune cells are widespread not only in cancer tissue, but also within normal tissues [41, 42, 60].

From the above analysis, we obtain that the system will approach instability as S increases despite what the effect from other cell populations is. Thus, we conjecture that the system will tend to instability while cell populations become larger and larger in the tissue. Mutation in genes results in emergence of new cell populations. Heterogeneity may increase consequently. Healthy tissue is assumed to be a stable system. Malignancy increases, disease ensues, and even cancer develops while the system is unstable. The increase of heterogeneity is an important signal. Next, we present some noteworthy phenomena involving heterogeneity in various datasets in tumor patients and our biological experiments in mice.

2.2. Some noteworthy phenomena in real data and our biological experiments.

2.2.1. Data of comparing heterogeneity between benign tumors and malignant tumors.

PCCs and extraadrenal sPGLs are catecholamine-producing tumors. To understand the role of heterogeneity in the pathogenesis of these tumors,

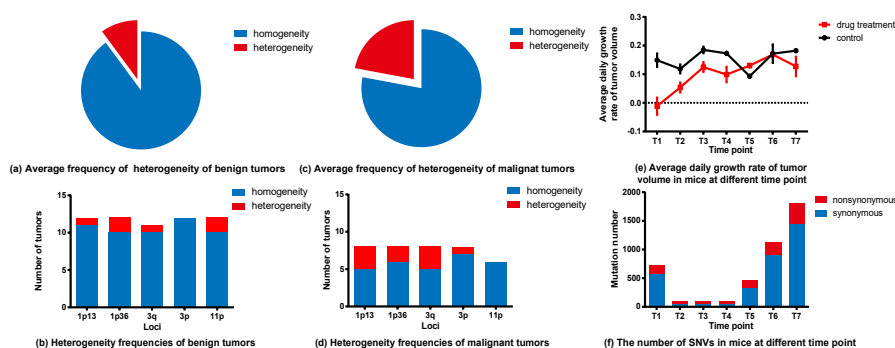


FIG 4. Heterogeneity signals the degree of malignancy of tumors. (a)-(d): Malignant tumors were found to have more heterogeneity (overall 23%) than benign tumors (overall 10%). Here, a tumor was considered to be genetically heterogeneous when LOH was present in at least one sample but not in all the samples of the same tumor or if the samples differed in loss of an upper or lower allele. 12 benign and 8 malignant PCCs and sPGLs for LOH on DNA extracted from four macroscopically separated regions of each tumor and from metastases were investigated [19]. The markers of LOH were selected on chromosomal regions frequently deleted in PCC, containing 1p, 3p, 3q, and 11p. (e) (f): Malignant tumors with higher heterogeneity (described by number of SNVs) showed strong drug resistance. Drug resistance was calculated using average daily growth rates of tumor volumes from the drug treatment group (red line) and control group (black line) (e), and the smaller difference indicated strong drug resistance. The number of SNVs in tumors at different time points (T1, T2, T3, T4, T5, T6, T7) were shown in subfigure (f).

12 benign and 8 malignant PCCs and sPGLs were investigated for loss of heterozygosity (LOH) on DNA extracted from four macroscopically separated regions of each tumor and from metastases [19]. The markers of LOH were selected on chromosomal regions that were frequently deleted in PCC, containing 1p, 3p, 3q, and 11p. A tumor was considered to be genetically heterogeneous when LOH was present in at least one sample but not in all the samples of the same tumor or if the samples differed in the loss of an upper or lower allele. Here we reused the data in [19]. Results in Figures 4(a) and (c) revealed that the average frequency of heterogeneity of benign tumors (10%) was lower than that of malignant tumors (23%). The heterogeneity frequency for chromosome 1p36 and 11p in benign tumors was the highest (17%), and that for 1p13 and 3q in malignant tumors was the highest (38%) as shown in Figures 4(b) and (d). It's concluded that the heterogeneity level of malignant tumors in PCCs and SPGLs is higher than benign tumors.

2.2.2. Data of the daily growth rate and drug resistance of malignant tumors in mice. The NCI-H82 cell line was injected into the mice to establish an implanted tumor mouse model, and the mice were then treated with anticancer drugs (cisplatin and etoposide). The tumors were removed from the mice and transplanted into other mice after drug treatment for 2 to 3 weeks. Tumor volumes were measured for a series of seven transplantations and the tumors at each time point were analyzed with exome sequencing. Mutation calling was performed with VarScan2. Drug resistance was calculated using average daily growth rates of tumor volumes from the drug treatment group and the control group, and the smaller difference indicated increased drug resistance as shown in Figure 4(e). We observed more mutations in tumors at the first time point (T1) than tumors at T2, T3, T4, and T5 in Figure 4(f), which might be due to dramatic DNA damage caused by drug treatment. We found that the number of mutations increased in tumors from T2 to T7. The branched evolution of

the passenger mutations is widely considered to be the main source of clonal diversity [55]. Therefore, a larger number of SNVs indicates more passenger mutations, which reflects the increase of tissue heterogeneity. It can be observed in Figures 4(e) and (f) that tumors with higher heterogeneity are related to higher malignant degree, higher daily growth rate, and higher drug resistance rate.

2.2.3. Overall survival analysis of TCGA data classified by SNV numbers and clone numbers. Data of all 9058 samples were downloaded from TCGA database, including 24 cancer projects as shown in Figure 5(a). Patients with a higher number of SNVs have significantly worse survival expectations from analyzing the data set ($p < 0.01$, Figure 5(b)). Then we analyzed tumor cell fractions after correcting for tumor purity with PyClone [39] to predict the clonal population structure of each sample in the data set; tumor purity was estimated with ABSOLUTE [9]. The proportion of samples decreases gradually with the increase of clone numbers as shown in Figure 5(c). Nearly half of the samples containing two or fewer clones were thought to have lower intratumor heterogeneity (ITH), which was defined as a low ITH group, and the remaining half of the samples were defined as a high ITH group. Notably, patients with lower levels of ITH have greater overall survival than high ITH group as shown in Figure 5(d). To further investigate the differences in ITH of different types of cancer, we divided all samples into 28 clusters referred to a previous study [17]. The first two clusters with the highest and the lowest percentage of high ITH patients, respectively, were selected (Figure 5(e)). We found from Figure 5(f) that those clusters with a higher proportion of high ITH patients have a significantly worse survival expectation. Taken together, these results suggest a clear correlation between ITH level and malignancy degree of tumors in pancancer. Heterogeneity can be considered a biomarker of cancer prognosis or treatment response.

We note that predeterioration or malignancy is usually accompanied by the increased heterogeneity from data of a comparison of heterogeneity between benign tumors and malignant tumors. Our model in section 2.1 can explain this phenomenon from a mathematical viewpoint. In the last two datasets, we find that patients with higher heterogeneity usually have stronger drug resistance and worse prognosis. To study and explain these phenomena, we further give a model for malignant tumors.

2.3. Model for malignant tumors. Malignant tumors contain not only tumor cells, but also other nontumor cells, such as endothelial cells, stromal fibroblasts, macrophages, and lymphocytes [4]. The distribution of cell populations in malignant tumors is disordered. Long-range intercellular interactions significantly affect the growth of malignant tumors [59]. These interactions are not only controlled by soluble factors such as cytokines, growth factors, and neurotransmitters, but also controlled by genetic information exchange through microRNAs. The interaction between malignant tumor cells and their microenvironment is dynamic and bidirectional. Stromal cells produce many cytokines that are conducive to tumor cell survival, proliferation, angiogenesis, etc., and have a dialogue with tumor cells. Tumor associated macrophages (TAMs) and T cells play a double-edged role in the occurrence and development of tumors as stated in Remark 5. Compared with interactions between nontumor cells, the interactions between tumor cells are very different. For example, malignant tumor cells lose contact inhibition, which can prevent unlimited proliferation and growth in normal cells [34]. They can form juxtaposition stimulation with their receptors through membrane binding factors, which in fact form contact stimulation. Besides, malignant tumor cells have stronger self-connection, but the connection with surrounding normal tissue is weakened. The gap junctional

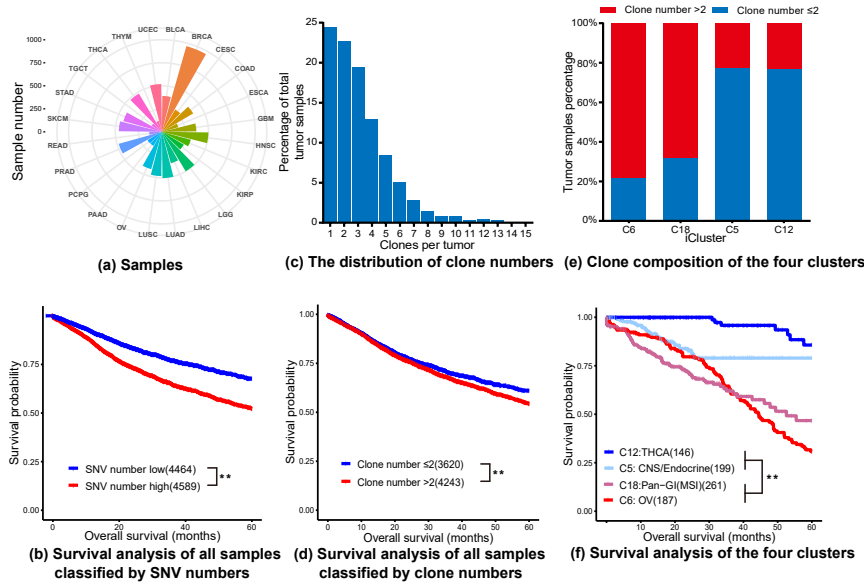


FIG 5. High levels of ITH lead to poor prognosis. (a) (b): Patients with a higher number of SNVs have significantly worse survival expectations. Exome sequencing data of 9058 patients including 24 cancer projects was downloaded from the TCGA database as shown in subfigure (a). Patients with more SNVs can be considered to have higher heterogeneity. (c) (d): Patients with higher levels of ITH predicted by PyClone are predictive of worse overall survival. Exome sequencing data includes 7863 patients. Subfigure (c) gives the clone number distribution predicted by PyClone. Nearly half of the patients have more than two clones, which are defined as a high ITH group. In subfigure (d), red curves are for patients predicted to present more than two clones (high ITH), and blue curves for those predicted to present one or two clones (low ITH). (e) (f): Pancancer molecular subtypes referred to previous studies with a smaller proportion of high ITH patients distribution have better survival expectations. The first two clusters with the highest and the least percentage of high ITH patients were selected, respectively, in subfigure (e). Those clusters with a higher proportion of high ITH patients have a significantly worse survival expectations as shown in subfigure (f).

communication function between adjacent cells of most malignant tumor cells is weak or lost [1]. Furthermore, the intercellular adhesion of a human tumor is weakened. The anchoring dependence of malignant tumor cells was significantly reduced, and some of them could even be suspended in liquid culture medium.

In view of the above characteristics of the malignant tumor, we regard it as an ecosystem in which mutualism and competition coexist, and model it with the RODEs as shown in Figure 1:

$$(2.18) \quad \frac{dX(t)}{dt} = M_T X(t), \quad X(0) = \eta.$$

$M_T = D_T + R_T$ is a community matrix of size $S \times S$, M_{Tij} describes the effect that cell population j has on i . M_{Tij} is a real-valued random variable on Ω . S is the number of cell populations. As mentioned in the model for normal tissues and benign tumors, S may be very large. We suppose that S is large throughout this subsection. D_T is a diagonal matrix and represents the effect of cell populations on themselves. R_T is a random matrix with larger standard derivation and represents the interaction between adjacent cell populations and long-range interactions. It is assumed that in the process of carcinogenesis, the self-inhibition of the cell populations is weakened [34], and the long-range interaction is enhanced [59]. Therefore, compared with D_{Nii}

in the model for normal tissues, D_{Tii} increases here. And the standard derivation of R_{Tij} ($i \neq j$) is larger than that of R_{Nij} in the model for normal tissues.

We generate the random matrix R_T in the following way:

- For each pair of R_{Tij} and R_{Tji} , ($i \neq j$), a random variable θ_1 with uniform distribution is generated. Namely, $\theta_1 \sim U[0, 1]$.
- If $\theta_1 \leq C_1$ (C_1 represents the relevance of the system; it can be set according to the characteristics of tumors), the second random variable θ_2 with uniform distribution is generated, namely, $\theta_2 \sim U[0, 1]$.
- If $\theta_2 \leq C_2$ (C_2 represents the proportion of mutualism in the system, which can be set according to the characteristics of tumors), two independent random variables with half-normal distributions are generated and assigned to R_{Tij} and R_{Tji} . That is, $R_{Tij} \sim |N(0, \sigma^2)|$, $R_{Tji} \sim |N(0, \sigma^2)|$ ($i \neq j$).
- If $\theta_2 > C_2$, two independent random variables with negative half-normal distributions are generated and assigned to R_{Tij} and R_{Tji} . That is, $R_{Tij} \sim -|N(0, \sigma^2)|$, $R_{Tji} \sim -|N(0, \sigma^2)|$ ($i \neq j$).
- If $\theta_1 > C_1$, we assign 0 to both R_{Tij} and R_{Tji} ($i \neq j$).
- All diagonal terms R_{Lii} are set to 0.

Here, σ is the standard deviation of the half-normal distribution which represents the interaction strength. Furthermore, let $D_T = a_T E = \text{diag}(a_T, a_T, \dots, a_T)$, where E is an identity matrix. $a_T = D_{Tii}$ represents the effect of the cell populations on themselves. These parameters in M_T can be set according to the characteristics of tumors. For the convenience of research, we make $C_2 = 0.5$ throughout this paper. It can be seen from [2] that the eigenvalues of M_T are uniformly distributed on an ellipse with $(a_T, 0)$ as the center and $\sigma\sqrt{SC_1}(1 + \frac{2}{\pi})$ as the horizontal half-axis.

2.3.1. The stability of the homogeneous linear ODEs with constant coefficients. Suppose that there are linear homogeneous ODEs with constant coefficients

$$(2.19) \quad \frac{dX(t)}{dt} = AX(t), \quad X(0) = \eta,$$

where A is a real matrix of order $S \times S$. Denote the eigenvalue of A with the largest real part by λ_{max}^A . By the classical theory of ODE systems, we have the following result.

LEMMA 2.5. *Suppose that there are two ODEs systems like (2.19) with matrices A_1 and A_2 , respectively, and each component of their initial values is nonzero. If both A_1 and A_2 have eigenvalues with positive real parts, then both the systems are unstable. Furthermore, if $\text{Re}(\lambda_{max}^{A_1}) > \text{Re}(\lambda_{max}^{A_2}) > 0$, then the system with the matrix A_1 will diverge faster than the system with the matrix A_2 .*

2.3.2. The stability of RODEs model (2.18). Since M_T is a random matrix, and its elements are random variables, we can write it as $M_T(\omega)$, $\omega \in \Omega$. Then the solution of (2.18) can be described as $X(t, \omega)$. From the eigenvalue distribution of the random matrix $M_T(\omega)$, it is easy to obtain the following result.

LEMMA 2.6. *Denote the eigenvalue of the $S \times S$ matrix $M_T(\omega)$ in (2.18) with the largest real part by $\lambda_{max}^{M_T(\omega)}$. Then we have, with relatively large S ,*

$$(2.20) \quad \text{Re}(\lambda_{max}^{M_T(\omega)}) \approx a_T + \sigma\sqrt{SC_1} \left(1 + \frac{2}{\pi}\right),$$

which means that $\text{Re}(\lambda_{max}^{M_T(\omega)})$ is independent of ω .

From (2.20), it can be obtained that $Re(\lambda_{max}^{M_T(\omega)})$ increases with S . The system is stable when $Re(\lambda_{max}^{M_T(\omega)}) < 0$ and it becomes unstable when $Re(\lambda_{max}^{M_T(\omega)}) > 0$. A system with S such that $Re(\lambda_{max}^{M_T(\omega)}) \approx 0$ can be regarded as facing a tipping point, that is, the system is sometimes stable and sometimes unstable. When ω is determined in the initial choice of the coefficients R_T , RODEs (2.18) become ODEs (2.19). Lemma 2.5 holds for (2.18) in such a situation. Combining with Lemma 2.6, we have the following result.

THEOREM 2.7. *Suppose there are two RODEs systems like (2.18) with random matrices $M_{T_1}(\omega)$ and $M_{T_2}(\omega)$, respectively, and each component of their initial values is nonzero. $M_{T_1}(\omega)$ and $M_{T_2}(\omega)$ are generated in the manner given in the beginning of this section, $\omega \in \Omega$. Their orders are S_1 and S_2 , respectively. Parameters a_T , σ , and C_1 are the same in both matrices. Suppose $\lambda_{max}^{M_{T_1}(\omega)}$ and $\lambda_{max}^{M_{T_2}(\omega)}$ are their eigenvalues with the largest real part, respectively. If $S_1 > S_2$ then $Re(\lambda_{max}^{M_{T_1}(\omega)}) > Re(\lambda_{max}^{M_{T_2}(\omega)}) \approx a_T + \sigma\sqrt{S_2 C_1}(1 + \frac{2}{\pi})$. Furthermore, if $Re(\lambda_{max}^{M_{T_2}(\omega)}) > 0$, then both systems are unstable and the system with the random matrix $M_{T_1}(\omega)$ diverges faster than the system with the random matrix $M_{T_2}(\omega)$.*

We apply the Euler scheme to RODEs (2.18) with step size $h = 0.0001$ and obtain numerical solutions with $S = 100$, $S = 140$, $S = 200$, $S = 300$, and $S = 400$, respectively. The initial values for the RODEs are $(1000, 500, 800, 50, 50, \dots, 50)^T$. Numerical experiments given in Figures 2(d)–2(f) further confirm our theoretical results. As mentioned previously, only one-path simulations are depicted in these subfigures for different cases given the fact that the simulations are similar with different ω . In the early stage, the tumor is still at the primary site, there is no macrovascular invasion, and the disease is in a relatively stable and controllable state compared with later stages of tumors. The system becomes unstable with higher heterogeneity such that $Re(\lambda_{max}^{M_T(\omega)}) > 0$ and the disease enters the deteriorating state. A system with S such that $Re(\lambda_{max}^{M_T(\omega)}) \approx 0$ faces a tipping point. In addition, we show that the higher the heterogeneity, the faster the RODEs system diverges. Therefore, we conclude that for mixed systems with mutualism and competition, the more cell populations there are, the easier the system will become unstable, the faster the system will collapse. And for malignant tumors, we deduce that, the higher the heterogeneity, the faster the disease may develop, the worse the prognosis, and the higher the degree of malignancy. In other words, the increase of heterogeneity signals predeterioration during tumor progression. The long-term accumulation of mutations due to various causes and the formation of a tumor microenvironment eventually lead to heterogeneity of the tumor. This can explain the last two phenomena mentioned in section 2.2 from a mathematical viewpoint.

Actually, it can be seen from (2.20) and Theorem 2.7 that the same conclusion holds for the effect of the cell populations on themselves a_T , the interaction strength σ , and the relevance of the system C_1 ; please see Figure 6 and Figure SM9 given in the supplemental material. In addition, Figures 2(c) and (d) reveal that, systems modeling malignant tumors approach instability with smaller S , which implies that malignant tumors have a lower threshold of stability on heterogeneity compared with normal tissues.

3. Discussion. Our theoretical analysis and biological experiments showed that the system modeling normal tissues and benign tumors may approach instability as the “species number” S increases while keeping other test conditions constant, which indicates that the malignancy may increase as the heterogeneity or complexity

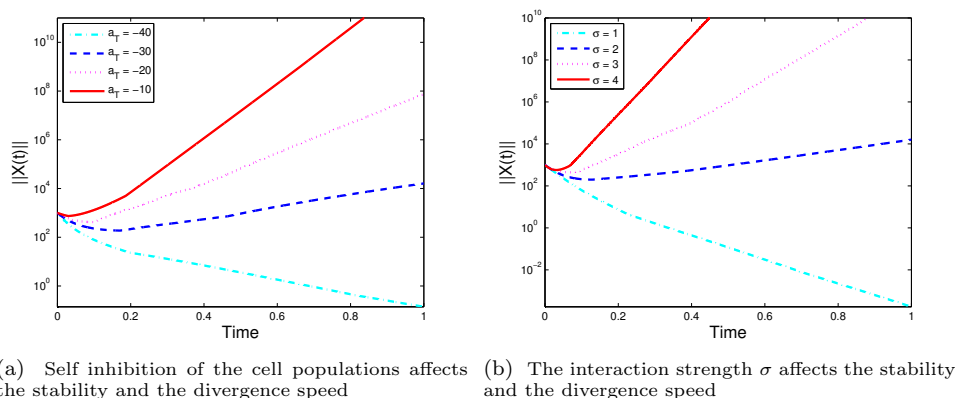


FIG 6. The effect of the cell populations on themselves a_T and the interaction strength σ affect the stability and the divergence speed of the system (RODEs (2.18)) describing malignant tumors. (a): Dynamics of RODEs (2.18) with different a_T are given. Here, $S = 200, \sigma = 2, C_1 = 0.6$. (b): Dynamics of RODEs (2.18) with different σ are given. Here, $S = 200, C_1 = 0.6, a_T = -30$. $\|X(t)\|$ is the infinity norm of the solution.

increases in tissues. May's theory revealed that it is more impossible for complex ecosystems to recover from small perturbations than from simpler ones [28]. In natural tissues, pairs of cell populations have well-defined interactions, which are very different when compared with ecosystems. Here we extend May's results to natural tissues and give an analytic stability criterion for the new case by matrix perturbation theory and the eigenvalue distribution of large scale random matrices. As far as we know, this is the first time that May's theory was extended and applied to tissue modeling. Some work, for example, Jacquet et al. [18], stated that complexity is not related to stability in empirical ecosystems. They found that the intrinsic energy organization within the food web is highly stable, allowing complex ecosystems to recover from perturbations. Authors concluded that May's stability criterion does not apply to small empirical ecosystems because of their structure, which has several stabilizing nonrandom properties. These nonrandom properties contributed to stability of complex systems. This study yields new insight on the complexity-stability debate compared with May's random matrix theory. Random matrix theory is carried out in the limit of infinitely large matrices, and all distributions are expected to converge in systems of several hundreds of species. But the community matrices that Jacquet et al. investigated had only between 6 and 54 species. Random matrix theory provides a sufficient approximation for large systems, but might need to be refined for smaller food webs. Since there are many functionally different subpopulations in one cell population, S describing species number in our model may be very large. Thus random matrix theory can be employed in the analysis.

A study of PCCs and sPGLs indicated that malignant tumors indeed have higher heterogeneity than benign tumors under certain circumstances [19]. The increase of heterogeneity is a signal in cancer genesis and progression. In addition, with the increase of age, the accumulation of random mutations eventually leads to the increase of the number of cell populations S as mentioned in the neutral theory. According to (2.13) and (2.14), there is a tendency to cause systems to become unstable with the increase of S . Therefore, older people are more likely to get sick and get cancer. Our theoretical findings and experimental results are in agreement with some existing viewpoints of cancer pathogenesis to a certain extent, for example, ideas from

references [6] and [40] as previously mentioned. Furthermore, our theory may provide a solution to Peto's paradox which claims that the cancer risk of different species does not increase with the number of somatic cells and the life span of the organism [36]. For instance, a man's cells are 1000 times the size of a mouse, and we usually live at least 30 times longer. However, in the wild, there seems to be little difference in the probability of cancer induction between mice and humans. The interactions between different animal cell populations are different, so their community matrices are different. Therefore, thresholds causing phase transitions of the system from a stable state to an unstable one will be different. The threshold is determined by the self-inhibition of the cell populations, the long-range interaction strength, and the heterogeneity. Consequently, different animals may have different tolerances to heterogeneity.

Furthermore, we obtain the stability criterion for RODEs modeling malignant tumors and compare the divergence speed of RODEs with different S . The divergence speed of ODE systems and the eigenvalue distribution of large scale random matrices are utilized in the analysis. Our theory show that the the RODEs system diverges faster with a larger species number S , which implies that malignant tumors with higher heterogeneity may grow faster. This theoretical result can explain our study result of tumor resistance in mice from a mathematical viewpoint. Furthermore, clinical data and mutation information from TCGA revealed that patients with more SNVs or gene clones have poorer survival for many cancer cases. More SNVs or gene clones imply higher heterogeneity in the tissue, and they are of higher malignancy and have shorter survival time. We consider other factors affecting growth speed of tumors, i.e., the effect of the cell populations on themselves a_T , the relevance of the system C_1 , and the interaction strength σ , are the same for all samples in the analysis. Our results of TCGA database suggests that intratumor heterogeneity is widespread in different tumor types, and more than 75% of cancer patients have at least two subclones. Overall survival of patients with higher heterogeneity was shown to be significantly worse than those patients with lower heterogeneity. Moreover, the heterogeneity of each tumor subtype is strongly correlated with the characteristics of each subtype, higher malignant subtypes tend to exhibit a higher heterogeneity. Together with previous reports, heterogeneity allows cancer cells to cope with microenvironmental assaults, immune surveillance, and therapy, eventually resulting in tumor invasion and metastasis [35]. Given the importance of heterogeneity in cancer research and treatment, we built a model in order to prove why increase of heterogeneity signals predeterioration during tumor progression. Compared with traditional heterogeneity research methods, including PyClone and EXPANDS, our model will be a potential alternative heterogeneity predictive model especially because of the rapid development of single-cell sequencing technology.

Malignant tumors develop and grow in complex microenvironments, which can be considered as ecosystems. Tumor cells interact frequently with their surrounding microenvironments, including noncellular and cellular components. These interactions can result in loss of tissue integrity, carcinogenesis, and further progress. They effectively control most stages of tumorigenesis, including migration, invasion, metastasis, and chemotherapeutic drug resistance [54]. Understanding these interactions would be helpful for developing better therapy strategy for cancer treatment. We suppose that changing some important intergroup or intragroup interactions between cell populations in the tumor is a promising strategy for future cancer treatment, such as reducing tumor growth by reprogramming TAMs to an M1 profile like paclitaxel [56] and enabling T cells to recognize and kill tumor cells by restarting and

maintaining the tumor-immune cycle and restoring the body's normal antitumor immune response in cancer immunotherapy [10]. Studying the mechanism behind the contact inhibition and restoring the contact inhibition in malignant tumors would be an interesting cancer treatment [44, 20]. $D_{Tii} = a_T$ is considered to be reduced in the process of contact inhibition being restored, thus the tumor growth speed may become slower as mentioned previously. Community matrices are changed while interactions change. Therefore, the dynamic behavior of the tumor will change. However, further research is still needed for full understanding of tumor cell interactions. Furthermore, we should pay attention to the importance of gap junctions related connexins in the prevention of tumor occurrence and early disease progression, because prevention may be more important than treatment. In addition, targeting the heterogeneity may be a potential strategy for cancer treatment. We expect that antiviral treatment can block the heterogeneity in the organ. Therefore, accepting antiviral treatment would probably work well on preventing the occurrence and recurrence of virus related cancers. Long-term studies on chronic hepatitis B patients at various stages showed that effective antiviral treatment consistently reduced the incidence of hepatocellular carcinoma (HCC) [66, 21]. The randomized controlled trials on postoperative antiviral treatment showed that antiviral treatment significantly decreased HCC recurrence and reduced HCC related death. Antiviral treatment also significantly decreased early recurrence and improved liver function after surgery [64].

It should be noted that, in this work we try to give a model for tissues (normal and tumor) in which interactions between cell populations are very complex. It is quite difficult to obtain the true values of parameters in the model. Therefore, the size S of cell populations cannot be roughly assumed for normal, benign, and malignant cells according to the stability conditions. Quantitatively, it depends on the tissue and disease. But qualitatively, we point out that the increase of heterogeneity signals predeterioration or tumor initiation during tumor progression. Our model for tissues (normal and tumor) has its specific characteristics. The application scenario is different from that of the existing articles using eigenvalues for stability analysis, such as [2, 13, 18]. We extend May's results to natural tissues in which pairs of cell populations have well-defined interactions and give an analytic stability criterion for the new case. In addition, we obtain the stability criterion for RODEs modeling malignant tumors and compare the divergence speed of RODEs with different S in order to study the speed of disease development. Hence, this work is not trivial from the perspective of applied mathematics. Moreover, we point out that increase of heterogeneity signals predeterioration during tumor progression, which can be used for early diagnosis and prediction of diseases. Therefore, the work is meaningful from the perspective of biomedicine. Besides, although we cannot experimentally prove that the increase of heterogeneity drives the cancer progression, we mathematically prove that heterogeneity is one of the possible causes of cancer. It is considered as an important signal for predeterioration or malignancy. Other biological sources may also destroy the stability of the system, for example, self-inhibition of the cell populations and the interaction strength also affect the stability of the system as mentioned in section 2. Tumor formation results from systematic changes in organisms, and its causality is complex. Heterogeneity may be one of the key factors under some conditions. Thus, the work is meaningful from the perspective of modeling. We used simple mathematical tools for stability analysis through eigenvalues in our paper, which, however, provides useful results that could be understood by more readers. The conclusions drawn in this paper are dependent on the model. However, our model is established under reasonable assumptions where the references on

biomedicine are listed to support these assumptions. Furthermore, the model is biologically plausible under many circumstances. The biological experiments are consistent with our theoretical findings. Although our model cannot cover all cases, it reflects important aspects which are observed in various real datasets. What is more, in the early stage, the disease is in a relatively stable and controllable state compared with later stages of tumors. At this time, the heterogeneity is still relatively low. Then, with the increase of heterogeneity, the system faces a tipping point [11, 63, 45, 65] when $Re(\lambda_{max}^{M_T(\omega)}) \approx a_T + \sigma\sqrt{SC_1}(1 + \frac{2}{\pi}) \approx 0$. If the heterogeneity continues to increase, the system becomes unstable and the disease enters the deteriorating state. The initial stage of cancer is such a process, and there is a similar process when cancer comes back. Some clones may be dominant due to factors such as drug resistance and some other clones may be replaced [27], thus reducing the heterogeneity of the tumor. The system will face a tipping point again with the increase of heterogeneity. It becomes unstable and the disease enters a deteriorating state again with the increase of heterogeneity. Therefore, a tumor could also be regarded as such a multistage process, and the predeterioration or malignancy of each stage can be signaled by the increase of heterogeneity. Actually, such a concept or explanation on “heterogeneity” (for cell populations or species of a tissue) is also consistent with (or similar to) dynamic network biomarker (DNB) theory on “fluctuation” (for cellular or molecular states of a tissue) [11, 63, 45, 65], which shows that increase of collective fluctuation in terms of cellular or molecular variations implies the imminent critical transition [25] from one state to another during a disease progression or during a biological process.

Acknowledgment. The authors are grateful to the anonymous referees whose insightful comments and valuable suggestions were crucial to the improvements of this article.

REFERENCES

- [1] T. AASEN, M. MESNIL, C. C. NAUS, P. D. LAMPE, AND D. W. LAIRD, *Gap junctions and cancer: Communicating for 50 years*, Nature Rev. Cancer, 16 (2016), pp. 775–788.
- [2] S. ALLESINA AND S. TANG, *Stability criteria for complex ecosystems*, Nature, 483 (2012), pp. 205–208.
- [3] N. ANDOR, J. V. HARNESS, S. MUELLER, H. W. MEWES, AND C. PETRITSCH, *Expands: Expanding ploidy and allele frequency on nested subpopulations*, Bioinformatics, 30 (2013), pp. 50–60.
- [4] R. BAGHBAN, L. ROSHANGAR, R. JAHANBAN-ESFAHLAN, K. SEIDI, A. EBRAHIMI-KALAN, M. JAYMAND, S. KOLAHIAN, T. JAVAHERI, AND P. ZARE, *Tumor microenvironment complexity and therapeutic implications at a glance*, Cell Commun. Signal., 18 (2020), pp. 1–19.
- [5] S. BENZEKRY, C. LAMONT, A. BEHESHTI, A. TRACZ, J. M. L. EBOS, L. HLATKY, AND P. HAHNFELDT, *Classical mathematical models for description and prediction of experimental tumor growth*, PLOS Comput. Biol., 10 (2014), e1003800.
- [6] F. BLOKZIL, ET AL., *Tissue-specific mutation accumulation in human adult stem cells during life*, Nature, 538 (2016), pp. 260–264.
- [7] S. BUNIMOVICH-MENDRAZITSKY, S. HALACHMI, AND N. KRONIK, *Improving bacillus calmette-guérin (BCG) immunotherapy for bladder cancer by adding interleukin 2 (IL-2): A mathematical model*, Math. Med. Biol., 33 (2016), pp. 159–188.
- [8] S. CAO, X. ZHU, C. ZHANG, H. QIAN, H.-B. SCHUTTLER, J. GONG, AND Y. XU, *Competition between DNA methylation, nucleotide synthesis, and antioxidation in cancer versus normal tissues*, Cancer Res., 77 (2017), pp. 4185–4195.
- [9] S. L. CARTER, ET AL., *Absolute quantification of somatic DNA alterations in human cancer*, Nature Biotechnol., 30 (2012), pp. 413–421.
- [10] D. S. CHEN AND I. MELLMAN, *Oncology meets immunology: The cancer-immunity cycle*, Immunity, 39 (2013), pp. 1–10.

- [11] L. CHEN, R. LIU, Z.-P. LIU, M. LI, AND K. AIHARA, *Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers*, Sci. Rep., 2 (2012), 342.
- [12] M. GERLINGER, ET AL., *Intratumor heterogeneity and branched evolution revealed by multiregion sequencing*, N. Engl. J. Med., 366 (2012), pp. 883–892.
- [13] D. GRAVEL, F. MASSOL, AND M. A. LEIBOLD, *Stability and complexity in model meta-ecosystems*, Nature Commun., 7 (2016), 12457.
- [14] R. M. GRAY, *Toeplitz and circulant matrices: A review*, Found. Trends Commun. Inform. Theory, 2 (2005), pp. 155–239.
- [15] D. HANAHAN AND R. A. WEINBERG, *Hallmarks of cancer: The next generation*, Cell, 144 (2011), pp. 646–674.
- [16] D. HART, E. SHOCHAT, AND Z. AGUR, *The growth law of primary breast cancer as inferred from mammography screening trials data*, British J. Cancer, 78 (1998), pp. 382–387.
- [17] K. A. HOADLEY, ET AL., *Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer*, Cell, 173 (2018), pp. 291–304.
- [18] C. JACQUET, C. MORITZ, L. MORISSETTE, P. LEGAGNEUX, F. MASSOL, P. ARCHAMBAULT, AND D. GRAVEL, *No complexity stability relationship in empirical ecosystems*, Nature Commun., 7 (2016), 12573.
- [19] E. KORPERSHOEK, C. K. STOBBE, F. H. VAN NEDERVEEN, R. R. DE KRIJGER, AND W. N. DINJENS, *Intra-tumoral molecular heterogeneity in benign and malignant pheochromocytomas and extra-adrenal sympathetic paragangliomas*, Endocr. Relat. Cancer, 17 (2010), pp. 653–662.
- [20] B. KRUST, D. E. KHOURY, C. SOUNDARAMOURTY, I. NONDIER, AND A. G. HOVANESSIAN, *Suppression of tumorigenicity of rhabdoid tumor derived G401 cells by the multivalent HB-19 pseudopeptide that targets surface nucleolin*, Biochimie, 93 (2011), pp. 426–433.
- [21] C.-L. LAI AND M.-F. YUEN, *Prevention of hepatitis B virus-related hepatocellular carcinoma with antiviral therapy*, Hepatology, 57 (2013), pp. 399–408.
- [22] A. K. LAIRD, *Dynamics of tumour growth: Comparison of growth rates and extrapolation of growth curve to one cell*, British J. Cancer, 19 (1965), pp. 278–291.
- [23] Y. LAVIN, ET AL., *Innate immune landscape in early lung adenocarcinoma by paired single-cell analyses*, Cell, 169 (2017), pp. 750–765.
- [24] J. LEI, S. A. LEVIN, AND Q. NIE, *Mathematical model of adult stem cell regeneration with cross-talk between genetic and epigenetic regulation*, Proc. Natl. Acad. Sci. USA, 111 (2014), pp. E880–E887.
- [25] R. LIU, K. AIHARA, AND L. CHEN, *Collective fluctuation implies imminent state transition: Comment on “Dynamic and thermodynamic models of adaptation” by A. N. Gorban et al.*, Phys. Life Rev., 37 (2021), pp. 103–107.
- [26] W. R. LOEWENSTEIN, *Junctional intercellular communication and the control of growth*, Biochim. Biophys. Acta, 560 (1979), pp. 1–65.
- [27] Z. LU, B. NIE, W. ZHAI, AND Z. HU, *Delineating the longitudinal tumor evolution using organoid models*, J. Genet. Genomics, 48 (2021), pp. 560–570.
- [28] R. M. MAY, *Will a large complex system be stable?*, Nature, 238 (1972), pp. 413–414.
- [29] L. MELLAL, D. FOLIO, K. BELHARET, AND A. FERREIRA, *Modeling of optimal targeted therapies using drug-loaded magnetic nanoparticles for liver cancer*, IEEE Trans. Nanobiosci., 15 (2016), pp. 265–274.
- [30] L. M. MERLO, J. W. PEPPER, B. J. REID, AND C. C. MALEY, *Cancer as an evolutionary and ecological process*, Nature Rev. Cancer, 6 (2006), pp. 924–935.
- [31] M. MEYERSON, N. L’OPEZ-BIGAS, N. NAVIN, A. FEINBERG, K. POLYAK, R. A. WEINBERG, A. LEVCHENKO, E. MARDIS, AND V. QUARANTA, *What fundamental questions in cancer would most benefit from math, engineering, and AI?*, Cell Systems, 6 (2018), pp. 263–265.
- [32] L. OESPER, G. SATAS, AND B. J. RAPHAEL, *Quantifying tumor heterogeneity in whole-genome and whole-exome sequencing data*, Bioinformatics, 30 (2014), pp. 3532–3540.
- [33] S. PALMER, L. ALBERGANTE, C. C. BLACKBURN, AND T. J. NEWMAN, *Thymic involution and rising disease incidence with age*, Proc. Natl. Acad. Sci. USA, 115 (2018), pp. 1883–1888.
- [34] M. PAVEL, ET AL., *Contact inhibition controls cell survival and proliferation via YAP/TAZ-autophagy axis*, Nature Commun., 9 (2018), pp. 1–18.
- [35] E. PERSI, Y. I. WOLF, H. DAVID, E. RUPPIN, D. FRANCESCA, R. A. GATENBY, R. J. GILLIES, AND E. V. KOONIN, *Mutation-selection balance and compensatory mechanisms in tumour evolution*, Nature Rev. Genet., 22 (2021), pp. 251–262.
- [36] R. PETO, *Epidemiology, multistage models, and short-term mutagenicity tests*, Int. J. Epidemiol., 45 (2016), pp. 1403–1428.
- [37] X. QIU, *Pathology*, Shanghai Science and Technology Press, Shanghai, 2016 (in Chinese).

- [38] X. REN, B. KANG, AND Z. ZHANG, *Understanding tumor ecosystems by single-cell sequencing: Promises and limitations*, Genome Biol., 19 (2018), 211.
- [39] A. ROTH, J. KHATTRA, D. YAP, A. WAN, E. LAKS, J. BIELE, G. HA, S. APARICIO, A. BOUCHARD-CÔTÉ, AND S. P. SHAH, *PyClone: Statistical inference of clonal population structure in cancer*, Nature Methods, 11 (2014), pp. 396–398.
- [40] A. I. ROZHOK AND J. DEGREGORI, *Toward an evolutionary model of cancer: Considering the mechanisms that govern the fate of somatic mutations*, Proc. Natl. Acad. Sci. USA, 112 (2015), pp. 8914–8921.
- [41] S. SAKAGUCHI, T. YAMAGUCHI, T. NOMURA, AND M. ONO, *Regulatory T cells and immune tolerance*, Cell, 133 (2008), pp. 775–787.
- [42] B. SALOMON, D. J. LENSCHOW, L. RHEE, N. ASHOURIAN, B. SINGH, A. SHARPE, AND J. A. BLUESTONE, *B7/CD28 costimulation is essential for the homeostasis of the CD4⁺ CD25⁺ immunoregulatory T cells that control autoimmune diabetes*, Immunity, 12 (2000), pp. 431–440.
- [43] R. SCHWARZ, ET AL., *Spatial and temporal heterogeneity in high-grade serous ovarian cancer: A phylogenetic analysis*, PLOS Med., 12 (2015), e1001789.
- [44] A. SELUANOV, C. HINE, J. AZPURUA, M. FEIGENSON, M. BOZZELLA, Z. MAO, K. C. CATANIA, AND V. GORBUNOVA, *Hypersensitivity to contact inhibition provides a clue to cancer resistance of naked mole-rat*, Proc. Natl. Acad. Sci. USA, 106 (2009), pp. 19352–19357.
- [45] J. SHI, K. AIHARA, AND L. CHEN, *Dynamics-based data science in biology*, Natl. Sci. Rev., 8 (2021), nwab029.
- [46] M. SIMEONI, P. MAGNI, C. CAMMIA, N. G. DE, V. CROCI, E. PESENTI, M. GERMANI, I. POGGESI, AND M. ROCCHETTI, *Predictive pharmacokinetic-pharmacodynamic modeling of tumor growth kinetics in xenograft models after administration of anticancer agents*, Cancer Res., 64 (2004), pp. 1094–1101.
- [47] J. A. SPRATT, D. V. FOURNIER, J. S. SPRATT, AND E. E. WEBER, *Decelerating growth and human breast cancer*, Cancer, 71 (1993), pp. 2013–2019.
- [48] G. W. STEWART AND J. G. SUN, *Matrix Perturbation Theory*, Academic Press, Boston, MA, 1990.
- [49] H. SUNG, J. FERLAY, R. L. SIEGEL, M. LAVERSANNE, I. SOERJOMATARAM, A. JEMAL, AND F. BRAY, *Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries*, CA Cancer J. Clin., 71 (2021), pp. 209–249.
- [50] K. SWANSON, R. ROSTOMILY, AND E. ALVORD, JR., *A mathematical modelling tool for predicting survival of individual patients following resection of glioblastoma: A proof of principle*, British J. Cancer, 98 (2008), pp. 113–119.
- [51] T. TAO, V. VU, AND M. KRISHNAPUR, *Random matrices: Universality of ESDs and the circular law*, Ann. Probab., 38 (2010), pp. 2023–2065.
- [52] C. TOMASETTI AND B. VOGELSTEIN, *Variation in cancer risk among tissues can be explained by the number of stem cell divisions*, Science, 347 (2015), pp. 78–81.
- [53] K. M. TURNER, V. DESHPANDE, D. BEYTER, T. KOGA, J. RUSERT, C. LEE, B. LI, K. ARDEN, B. REN, AND D. A. NATHANSON, *Extrachromosomal oncogene amplification drives tumor evolution and genetic heterogeneity*, Nature, 543 (2017), pp. 122–125.
- [54] H. UNGEFROREN, S. SEBENS, D. SEIDL, H. LEHNERT, AND R. HASS, *Interaction of tumor cells with the microenvironment*, Cell Commun. Signal., 9 (2011), pp. 1–8.
- [55] I. VITALE, E. SHEMA, S. LOI, AND L. GALLUZZI, *Intratumoral heterogeneity in cancer progression and response to immunotherapy*, Nature Med., 27 (2021), pp. 212–224.
- [56] C. W. WANDERLEY, ET AL., *Paclitaxel reduces tumor growth by reprogramming tumor-associated macrophages to an M1 profile in a TLR4-dependent manner*, Cancer Res., 78 (2018), pp. 5891–5900.
- [57] D. C. WANG AND X. WANG, *Systems heterogeneity: An integrative way to understand cancer heterogeneity*, Semin. Cell Dev. Biol., 64 (2017), pp. 1–4.
- [58] J. A. WEIS, M. I. MIGA, L. R. ARLINGHAUS, X. LI, V. ABRAMSON, A. B. CHAKRAVARTHY, P. PENDYALA, AND T. E. YANKEELOV, *Predicting the response of breast cancer to neoadjuvant therapy using a mechanically coupled reaction-diffusion model*, Cancer Res., 75 (2015), pp. 4697–4707.
- [59] J. WEST AND P. K. NEWTON, *Cellular interactions constrain tumor growth*, Proc. Natl. Acad. Sci. USA, 116 (2019), pp. 1918–1923.
- [60] K. WING, Y. ONISHI, P. PRIETO-MARTIN, T. YAMAGUCHI, M. MIYARA, Z. FEHERVARI, T. NOMURA, AND S. SAKAGUCHI, *CTLA-4 control over Foxp3⁺ regulatory T cell function*, Science, 322 (2008), pp. 271–275.
- [61] I. P. WITZ AND O. LEVY-NISSENBAUM, *The tumor microenvironment in the post-PAGET era*, Cancer Lett., 242 (2006), pp. 1–10.

- [62] T. WU AND Y. DAI, *Tumor microenvironment and therapeutic response*, Cancer Lett., 387 (2017), pp. 61–68.
- [63] B. YANG, M. LI, W. TANG, W. LIU, S. ZHANG, L. CHEN, AND J. XIA, *Dynamic network biomarker indicates pulmonary metastasis at the tipping point of hepatocellular carcinoma*, Nature Commun., 9 (2018), 678.
- [64] J. YIN, N. LI, Y. HAN, J. XUE, Y. DENG, J. SHI, W. GUO, H. ZHANG, H. WANG, S. CHENG, AND G. CAO, *Effect of antiviral treatment with nucleotide/nucleoside analogs on postoperative prognosis of hepatitis B virus-related hepatocellular carcinoma: A two-stage longitudinal clinical study*, J. Clin. Oncol., 31 (2013), pp. 3647–3655.
- [65] C. ZHANG, H. ZHANG, J. GE, T. MI, X. CUI, F. TU, X. GU, T. ZENG, AND L. CHEN, *Landscape dynamic network biomarker analysis reveals the tipping point of transcriptome reprogramming to prevent skin photodamage*, J. Mol. Cell Biol., 13 (2021), pp. 822–833.
- [66] Y.-Q. ZHANG AND J.-S. GUO, *Antiviral therapies for hepatitis B virus-related hepatocellular carcinoma*, World J. Gastroenterol., 21 (2015), pp. 3860–3866.
- [67] J. X. ZHOU, A. O. PISCO, H. QIAN, AND S. HUANG, *Nonequilibrium population dynamics of phenotype conversion of cancer cells*, PLOS One, 9 (2014), e110714.