Inherent evolutionary unpredictability in cancer model system

Subhayan Chattopadhyay[1*], Jenny Karlsson[1], Adriana Mañas[2], Ryu Kanzaki[2], Elina Fredlund[3,4,5], Andrew Jackson Murphy[6], Chris Morton[6], Natalie Andersson[1], Karin Hansson[2], Katarzyna Radke[2], Andrew M. Davidhoff[6], Sofie Mohlin[3,4,5], Kristian Pietras[2], Daniel Bexell[2], David Gisselsson[1,7]

[1]Division of Clinical Genetics, Department of Laboratory Medicine, Lund University, Lund, Sweden
[2]Division of Translational Cancer Research, Department of Laboratory Medicine, Lund University, Lund, Sweden
[3]Division of Pediatrics, Clinical Sciences, Translational Cancer Research, Lund University
[4]Lund Stem Cell Center, Lund University, Lund, Sweden.
[5]Lund University Cancer Center, Lund University, Lund, Sweden.
[6]Department of Surgery, St. Jude Children's Research Hospital, Memphis, USA
[7]Division of Oncology and Pathology, Department of Clinical Sciences, Lund University, Lund, Sweden

*Corresponding Author: Subhayan Chattopadhyay, Lund University, SE22184 Lund, Sweden. Phone: 467-369-78242; e-mail: subhayan.chattopadhyay@med.lu.se

MAIN

A quarter of a century ago, in 1993, Donald S. Coffey hypothesized that cancer is an 'abrupt' and 'emergent' phenomenon caused by the transformation of the cell proliferation machinery from an ordered to a disordered albeit self-organizing state [1]. In the following decades researchers have focused more and more on the characterization of the dysregulated (disordered) genomic and non-genomic elements of cancer. Yet the prevalent somatic mutation theory stating that near-random DNA lesions combined with selection cause cancer still indicate a failure to grasp Coffey's vision of using cancer's self-organizing features to identify unifying aspects across malignancies [2-4]. With the advent of precision cancer medicine, more efforts than ever are now being put on characterizing the broad spectrum of genetic variation among individual tumors instead of describing all cancers as one entity. This has resulted in a fine-tuned prognostication of many neoplasms and the identification of treatment targets based on each cancer's molecular signature. Yet, in many cases, the threat of a relapse and a consequent treatment resistant aggressive disease loom, reminding us that cancer shares features of resilience with many other self-organizing systems. Cancer relapse mechanism have also thoroughly been studied through the lens of genetic diversity, elucidating how tumors evolve along different evolutionary trajectories [5] and, how resistant clones often appear due to excessive branching early in the disease and remain dormant only to clonally fixate afterwards [6-7]. Despite the massive amount of data accumulated on the molecular routes to relapse, it still remains an essentially unpredictable phenomenon. However, as inferred by Coffey, cancer is not a purely stochastic phenomenon. Rather, it results from runaway dysregulations in a complex, dynamic and adaptive system. Can we then use generic knowledge from other systems in a state near chaos to better understand tumorigenesis?

In his seminal work on chaotic oscillations in dynamic systems, James A. Yorke candidly showed that a self-repeating process with the periodicity of three or higher experiences chaotic fluctuations. It was exemplified with the logistic function [8], a sigmoidal curve with a periodically oscillating slope (fig. 1A). Robert May further demonstrated how chaotic fluctuations could appear as a function of logistic growth in a bifurcation diagram [9] (fig. 1B) with a presentation of how a function can assume more than one value near its asymptotes (orbits). Today the generalized logistic model (with the Gompertz curve being a special case) is a commonly used construct in simulating spatially constricted (bonded) growth of species /

cell populations and it is popularly used to emulate solid tumor growth [10-12]. Several comparative studies analyzing *in-vivo* tumor growth have established the suitability of the logistic function in tumor growth estimation [12-13]. Additionally it has also been used to illustrate clonal selection and genetic drift in solid tumors in silico [14]. Simulations show that evolutionary trajectories of cancers are highly dependent on how cell populations grow and how they interact with the stromal boundary [15]. However, whether the logistic function as a mechanism for tumor growth can explain emergent clonal geographies in tumorigenesis remains to be explored. We here probe this question with a focus on the trends in evolution of the mutational landscapes.

First, we performed simulations based on the assumption that tumors prior to clonal expansions emerge from a uniform population of cells that are henceforth referred to as the ancestors. Ancestors went through clonal expansion adhering to certain parameters (i.e., a growth rate governed by birth and death rates, a rate of acquiring mutations at each cell division and, a probability of a mutation to be a driver mutation), which were set at initiation and remained unchanged, purely for the sake of simplicity. The selective advantages provided by a mutation was sourced from deleteriousness scores provided by the COSMIC database (detailed in Methods) [16] and the maximum number of cells at the end of the simulation was kept fixed. Instead of determining the ancestry of each cell, we focused on how cells underwent genetic diversification due to variations in simulation parameters making them genetically distant and heterogeneous progenies of the ancestors. All cells at the end of each simulation were clustered according to the number of acquired mutations where the least mutated cluster is considered to represent the most recent common ancestors. With this, we arrive at a simple measure of the percentage of ancestors remaining at the end of a simulation run.

With each simulation cycle the mutations were drawn at random from a set of known somatic mutations (see Methods). Using fixed initial conditions, we evaluated if the genetic diversification remained somewhat predictable given certain mutation parameters. Keeping the growth rate unchanged, we first varied the probability of acquiring a mutation at each cell division between 0.01 and 0.04 (fig. 1C) 14. Cellular growth rates were affected by mutation rate, changes in fitness etc. and we also observed a non-linear monotonically decreasing relationship between the percentage of ancestors and increasing mutation rate. The fraction of remaining ancestors at the end of the simulation varied on average between 92% (low

mutation rate) and 36% (high mutation rate, calculated over 100 simulation runs). We could not detect any canonical relationship between the median number of mutations and growth rates. Instead, mutation acquirement seems to undergo a dramatic step-like increment over an arbitrary span (fig. 1D). Clear jumps in median mutation aggregation can be seen near growth rates 3.0 and 3.4 indicating rapid changes in the distribution of mutation aggregation at certain values of growth rate. As the number of mutations increases, so should the percentage of cells that harbor a subset of these mutations. Indeed, the jumps in mutation aggregation corresponded to the classic bifurcation diagram of the logistic function (fig 1B 9) when plotting the percentage of the remaining ancestor against the growth rate (fig. 1E).

The shape of the bifurcation diagram predicts chaotic behavior as the growth constant increases over a certain threshold that also represents a one-to-many solution of the logistic map at the asymptotes and we see in our simulations, that the percentage of ancestors very closely resemble this orbit diagram. In our simulations faster growth rate resulted in a markedly heterogeneous mutational landscape in comparison to a slow growth rate (fig. 1F). This led us to conclude that chaotic growth (at least in our case) is a biologically emergent feature exclusively occurring in tumors following logistic growth and with a growth rate above 3.0.

Talkington and Durett evaluated *in vivo* growth characteristics of several cancer cell lines and found that numerous cell lines at least partly resemble logistic (specifically, Gompertzian) growth [13] and in addition, some of the earliest investigations on *in vitro* growth experiments have also pointed towards a similar pattern [17-18]. However, the extent to which fast logistic growth is present in tumors *in vivo* has remained unclear.

To evaluate whether such fast-growing tumors with potential non-linear clonal evolution exist, we assessed how often logistic growth is observed across *in vivo* model systems, in situations simulating relapse by implantation of a limited tumor population. Pediatric cancers are well known to have much higher growth rates compared to adult cancers. Neuroblastoma (NB) and Wilms tumor (WT) are two of the most common solid pediatric tumors, notorious for their fast growth. We estimated growth rates from previously published data from untreated patient-derived xenografts (PDX), derived from NB and WT [5,19-22]. As a comparison to adult tumors (i.e., slow growers) [23], we also evaluated uninhibited growth of several lung and breast cancer cell lines (unpublished data). Strikingly, 43% of the evaluated NB PDXs abided by logistic growth (73% of growth rates being more than 3.0 with a median

of 10.0), considerably above the chaotic bifurcation limit of 3.0 whereas 75% of the WT PDXs showed logistic growth (all growth rates over 3.0) with median growth rate of 31.0 (fig. 1G-H). PDXs from the H441 lung cancer and MCF7 breast cancer cell lines, 71% and 78%, respectively, experienced logistic growth but none over 3.0. The median growth rates were only 1.13 and 0.9 respectively, far below that for chaotic fluctuations. In addition to patient derived models, we also evaluated the growth of two lineages of the NB SK-N-BE(2)C cell line *in vivo*, which median growth rates of 5.0 and 4.0 respectively (Supplementary material). When combined, the growth rates from all NB replicates had a median of 6.0 and that for the WT was 24.0. All but one breast cancer replicates that adhered to a logistic growth had growth rates below 3.0 with a median of 0.99 and that for lung cancer replicates was 0.68 none reaching 3.0.

The implication of the present study is limited to tumors that demonstrated a logistic growth curve however not all tumors did so. This is possibly due to different inherent growth characteristics as the absolute volume of the tumors varied substantially between the pediatric and the adult tumors. By week four of observation the average volume for NB replicates exceeded 2000 mm^3. In the same time frame adult tumors grew only about 400 mm3 except for one replicate (m3) of the MDA-MB-231 breast cancer cell line (growth rate 3.6). The WT PDXs at implantation were all larger than 200 mm^3 with an average size of 308 mm^3 that grew up to 3800 mm^3 on average by week 3. Also notable is that NB PDXs underwent orthotopic implantation initially and then were grafted heterotopically[22] whereas the WTs were heterotopic. Often the reason that a curve failed to be reasonably fit to a logistic function was due to abrupt changes in growth pattern (measurement artifacts etc.) which is a regular phenomenon. Nevertheless, the presence of any systematic artifact being responsible for the low growth rates in the adult tumors was ruled out as the experiments yield growth rate within-variances of 0.023 and 0.012.

The simulations under chaotic growth resulted in a massively varied mutational landscapes implying that evolutionary trajectories, even in controlled model systems, are intrinsically unpredictable under certain conditions that invoke chaos. Our experimental data imply that this seems particularly applicable to fast-growing pediatric tumors. The mouse model data illustrate that some tumor types are associated with specific growth characteristics (non-logistic/logistic and slow/fast growing), which determines whether they evolve clonally in a chaotic or predictable fashion. This essentially puts an intrinsic limit to what can be predicted

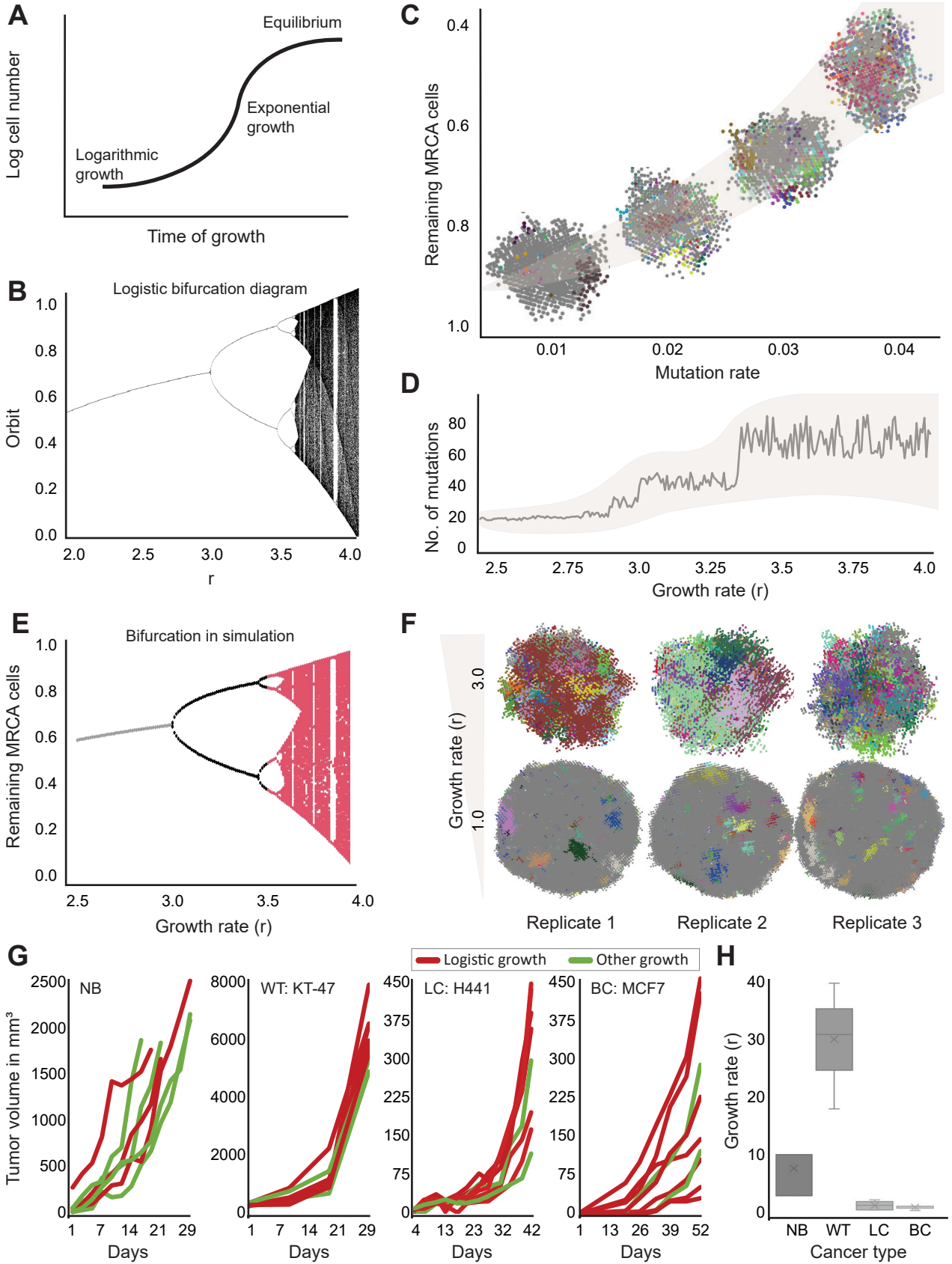
using today's popular approach of precision cancer medicine. However, observing the tumor growth rate can potentially become a noninvasive way to determine with what degree of predictability the tumor is expected to evolve, in turn providing insights to how often the clonal landscape of a tumor needs to be resampled to evaluate options for targeted therapy based on molecular profile.

REFERENCES

[1] DS Coffey, Nature Medicine **4** (8), 882 (1998).
[2] C Sonnenschein and AM Soto, PLOS Biology **18** (4), e3000670 (2020).
[3] H Ledford, Nature **528** (7582), 317 (2015).
[4] A Balmain, Nature Genetics **52** (11), 1139 (2020).
[5] J Karlsson et al., Nature Genetics **50** (7), 944 (2018).
[6] N Andersson et al., Cancer Research, canres.3468.2019 (2020).
[7] M Jamal-Hanjani et al., PLOS Biology **12** (7), e1001906 (2014).
[8] T-Y Li and JA Yorke, The American Mathematical Monthly **82** (10), 985 (1975).
[9] RM May, Nature **261** (5560), 459 (1976).
[10] L Heuser et al., **43** (5), 1888 (1979).
[11] EA Sarapata and LG de Pillis, Bulletin of Mathematical Biology **76** (8), 2010 (2014).
[12] P Gerlee, Cancer Research **73** (8), 2407 (2013).
[13] A Talkington and R Durrett, Bulletin of Mathematical Biology **77** (10), 1934 (2015).
[14] B Waclaw et al., Nature **525** (7568), 261 (2015).
[15] R Noble et al., Nature Ecology & Evolution **6** (2), 207 (2022).
[16] JG Tate et al., Nucleic Acids Research **47** (D1), D941 (2019).
[17] AK Laird, British Journal of Cancer **18** (3), 490 (1964).
[18] L Norton, Cancer Research **48** (24 Part 1), 7067 (1988).
[19] K Hansson et al., Science Translational Medicine **12** (562), eaba4434 (2020).
[20] K Radke et al., Translational Oncology **14** (8), 101149 (2021).
[21] AJ Murphy et al., Nature Communications **10** (1), 5806 (2019).
[22] A Mañas et al., bioRxiv, 2022.04.01.486670 (2022).
[23] F Hinman, Journal of Urology **145** (1), 126 (1991).

Figure legend

Figure 1. Simulation of tumor evolution. (**A**) Interpretation of the classical shape of a logistic growth curve. (**B**) Classical orbit diagram of logistic curve (logistic map) [9]. (**C**) Change in percentage of remaining MRCA (most recent common ancestors) cells was plotted against increasing mutation rate. Examples were overlayed at the estimated mean calculated with 100 simulation runs. Confidence interval with shaded overlay depicts 3σ limit. (**D**) Average number of total distinct mutations at the end of simulation is plotted against growth rates (r). All estimates are calculated over 100 separate runs. As logistic map experiences chaotic fluctuations starting at r = 3.0, the x-axis is terminated at 4.0 as it already provided adequate range (confidence interval depicts 3σ limit) (**E**) Cell growth rate was plotted against percentage of remaining ancestors at the end of simulations, grey section indicates $r < 3.0$, calculated by taking average; black section indicates $3.0 < r < 3.5$, calculated by cluster centroids; red section indicates $r > 3.5$ where all estimates are plotted. (**F**) Three replicates of simulation are shown for growth rates (r) 1.0 and 3.0 to draw attention to the fact that high r results in markedly different mutational pattern compared to that in low r. (**G**) Tumor growth data in mouse models are depicted; red: logistic growth and green: did not fit to logistic function. NB: neuroblastoma (Radke et. al.); WT: Wilms tumor (KT-47 sample, Murphy et. al.); LC: lung cancer (H441 cell line); BC: breast cancer (MCF7 cell line). (**H**) Box plots of growth rates of data shown in (1-G) plotted across cancer types.


A
Log cell number
Equilibrium
Exponential growth
Logarithmic growth
Time of growth
B
Logistic bifurcation diagram
Orbit
r
C
Remaining MRCA cells
Mutation rate
D
No. of mutations
Growth rate (r)
E
Bifurcation in simulation
Remaining MRCA cells
Growth rate (r)
F
Growth rate (r)
Replicate 1
Replicate 2
Replicate 3
G
Logistic growth
Other growth
NB
WT: KT-47
LC: H441
BC: MCF7
Tumor volume in mm³
Days
H
Growth rate (r)
NB
WT
LC
BC
Cancer type