Faccioni *et al. BMC Methods* (2025) 2:13
https://doi.org/10.1186/s44330-025-00033-8

BMC Methods

METHODOLOGY Open Access

CloVarS: a simulation of single-cell clonal variability

Juliano L. Faccioni[1,2*], Frederico Kraemer-Mattos[1,2], Karine R. Begnini[1,2], Julieti H. Buss[1,2], Daphne Torgo[1,2], Camilla Brückmann de Mattos[1,2], Camila B. Cassel[1,2], Sophie Seidel[3], Leonardo G. Brunnet[4], Manuel M. Oliveira[5] and Guido Lenz[1,2*]

Abstract

Background High-throughput time-lapse microscopy has allowed researchers to monitor individual cells as they grow into colonies and react to treatments, but a deeper understanding of the data obtained after image analysis is still lacking. This is in part due to the biological and computational challenges related to long-running experiments and single-cell tracking.

Methods Clonal Variability Simulator (CloVarS) is a Python tool for generating synthetic data of single-cell lineage trees to model time-lapse microscopy experiments. After colony initialization, each individual cell is simulated for a given number of simulation frames. During simulation, cells can migrate, enter mitosis (divide), and enter apoptosis (die). These events are determined by distributions of cell division and death times, which can be inferred from and fit to experimental data. Colonies have an adjustable mother-daughter (MD) and sister-sister (SisSis) fitness memory (f_m), meaning that cell fitness can range from equal to the fitness of their parent or sibling (f_m =1), non-related (f_m =0) to anti-correlated (f_m =-1). Arbitrary treatments can be delivered to colonies at any time, modifying the division and death distributions.

Results We show examples of trees with different division and death curves and the resulting number of cells per tree. Values of f_m from -1 to 1 generated SisSis and MD Pearson correlations that were fit to correlations observed in different experimental data from normal and cancer cells.

Discussion CloVarS is an important asset for quickly exploring colony fitness dynamics, its heritability, testing biological hypotheses, benchmarking cell tracking algorithms, and ultimately improving our understanding of single-cell lineage data.

Keywords Single-cell tracking, Clonal variability, Cell fitness, Simulation, Clonal trees

*Correspondence:
Juliano L. Faccioni
julianofaccioni@gmail.com
Guido Lenz
lenz@ufrgs.br
[1]Departmento de Biofísica, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre 91501-970, Brazil
[2]Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre 91501-970, Brazil
[3]Department of Biosystems Science and Engineering, ETH Zurich, Klingelbergstrasse 48, Basel 4056, Switzerland
[4]Instituto de Física, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre 91501-970, Brazil
[5]Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre 91501-970, Brazil

Introduction

A human cell can be understood as a system of biochemical signals that collectively direct it towards a given phenotypic state of growth, migration, division, homeostasis or death, over time. Advances in time-lapse microscopy have allowed researchers to glimpse into these complex relationships through single-cell tracking [1–4], but this is not without its challenges. Maintaining cell viability during long-running experiments may be difficult or nearly impossible, depending on the experimental settings [5]. While much effort has been put into understanding single-cell tracking data in many different contexts, currently there is no methodological consensus on how to best analyze them [6]. Additionally, the variability generated in clonal cells can stem from different sources, and modeling how clones are generated is a key strategy for understanding the mechanisms underlying the onset of clonal cell heterogeneity [7–9].

Various simulation models attempt to encapsulate the intrinsic complexity of the variation in lineage trees with different modeling strategies. Most models focus on a specific questions, such as the overall variability of lineage maps [10] or the mother-daughter, sister-sister and cousin-cousin similarities in intermitotic time, apoptosis time and TP53 expression [11]. Others used lineage maps to model cellular heterogeneity and plasticity with a hidden Markov model to highlight the relevant states and its transitions of cancer cells treated with chemotherapeutic agents [12]. In spite of these efforts, a biologically-driven modeling of cell colonies to model the generation of heterogeneity, which is central for understanding key aspects of tumorigenesis and response to therapy [13], particularly one where users are able to tweak experimental parameters, is still lacking.

Here, we introduce the Clonal Variability Simulator (CloVarS), a simulator of the formation of cell colonies. The parameters governing cell division and death can be derived from experimental data of treated and untreated cells (see Supplementary Data for details), thus producing large amounts of biologically relevant data in seconds instead of weeks. CloVarS can be used to test hypotheses of cell growth dynamics and explore novel methods for analyzing single-cell tracking data without the need for a complex experimental setup. CloVarS is written in Python. Its source code, along with installation, execution instructions and a graphic user interface, can be found at https://github.com/jfaccioni/clovars.

Materials and methods

We briefly explain CloVarS, with further details provided in the Supplementary Data.

Simulation model

CloVarS simulates individual clonal cells over time. At each simulation time step (referred to as a *frame*), each cell is able to either: migrate to a new position; divide into two new cells; or die and vanish from the colony (Fig. 1A). At every frame, data from the current simulation state is written to the output files, resulting in a complete history of the lineage trees generated by each colony. The simulation ends when a stop condition is met (Supplementary Table S1). Output files can then be further processed or analyzed as desired. For convenience, CloVarS includes some basic visualization and analysis scripts of the output files.

Fitness thresholds and cell fate

Each cell has a fitness threshold for division (t_{div}) and death (t_{death}) (Fig. 1B). The fitness thresholds indicate how early a cell will trigger its division/death event; for example, if t_{div} = 0.1, then the cell divides at the age corresponding to the 0.1 percentile of the division curve (see the numerical examples in Supplementary Fig. S1). If the cell has not reached the age for neither of its division

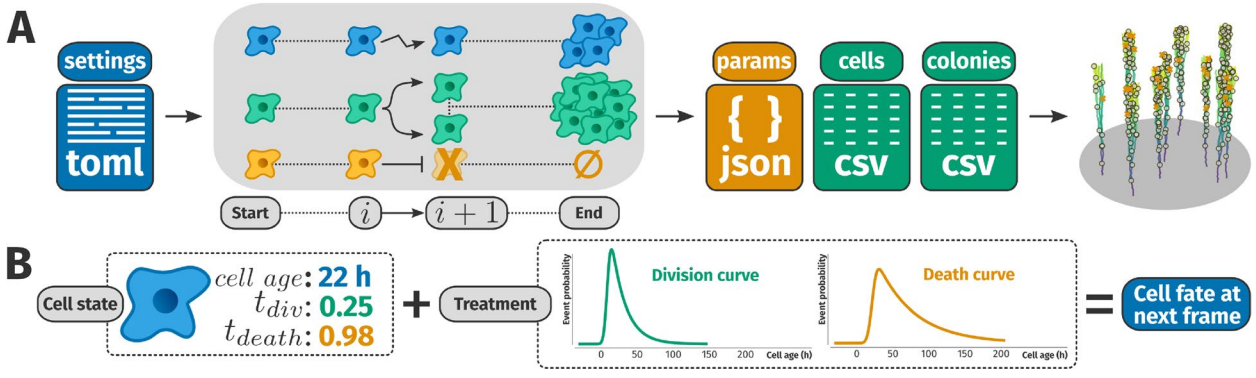

Fig. 1 CloVarS overview. **A** The simulation starts after reading the parameters from the settings file (see Supplementary Tables S1, S2, S3, S4 for parameter details). Between adjacent simulation frames i and $i+1$, cells are able to migrate, divide, or die. Once the simulation ends, the output files can be used for data visualization and analysis. **B** Cell fate at the next frame is determined by factoring in its age, division and death thresholds, and the treatment it is subject to (see also Supplementary Fig. S1)

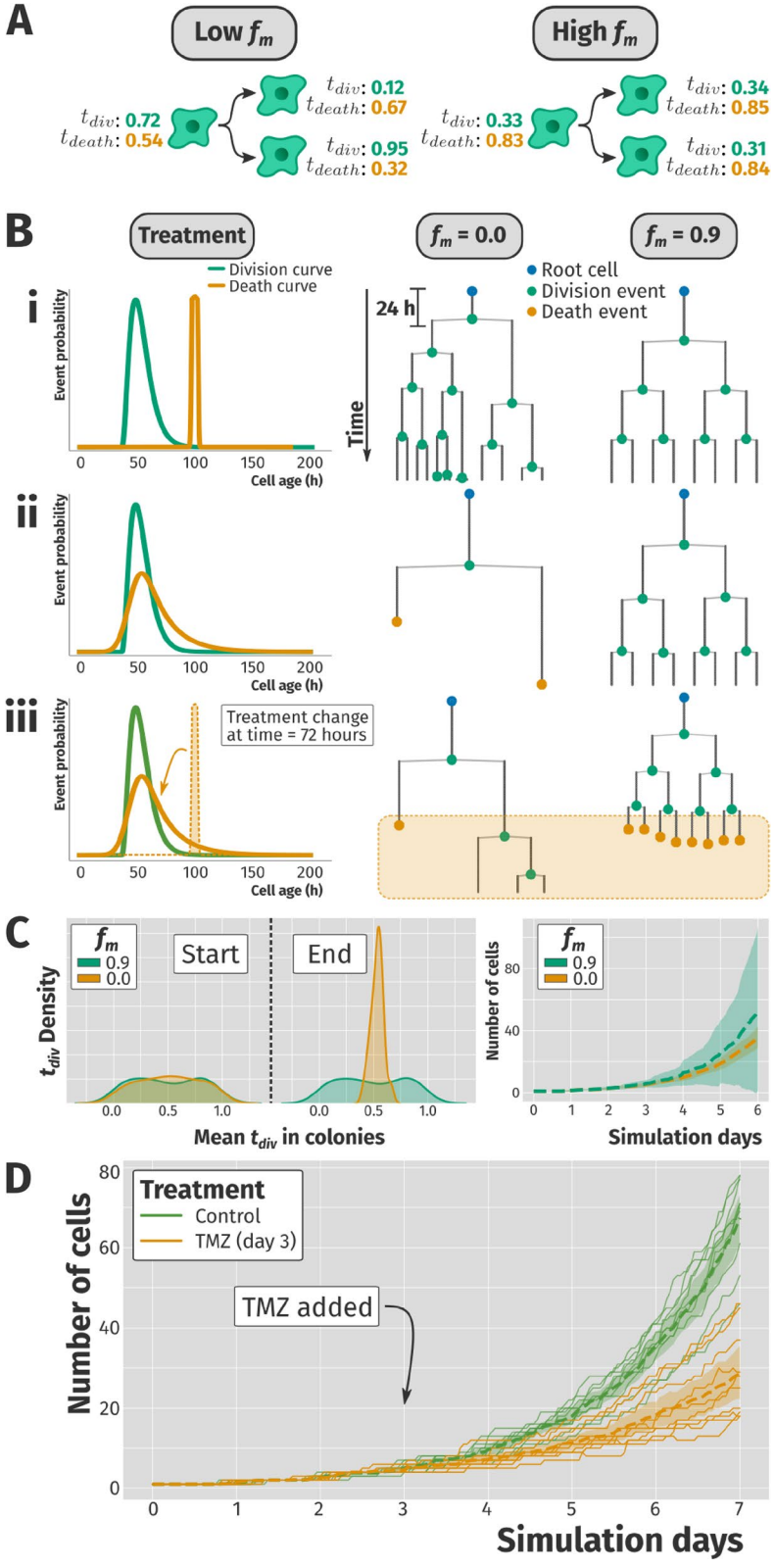

Fig. 2 (See legend on next page.)

(See figure on previous page.)

Fig. 2 CloVarS results. **A** Mother cells with high fitness memory (f_m) produce offspring with similar fitness thresholds, while daughter cells with low f_m have largely uncorrelated fitness thresholds among themselves and their mother. **B** Two colonies with a single initial cell each were simulated for 120 frames in the following scenarios: (*i*) no cell death; (*ii*) moderate cell death; (*iii*) moderate cell death after treatment change at 72 h. **C** Effect of f_m on colony size variability. For each f_m, 100 colonies were simulated for 144 frames. Left, distribution of mean t_{div} in colonies at the start and end of the simulation. Right, average number of cells + SD (shaded area) for each f_m over time. **D** Simulation results based on experimental data. For each treatment, 10 colonies with $f_m = 0$ were simulated for 168 frames. Dashed lines: treatment average. In Fig. 2B, C, and D, Δt between frames was 1 h. Code used to produce Fig. 2B, C, and D can be found in the CloVarS repository

or death events to trigger, its fate at the next frame is to migrate. Collectively, t_{div} and t_{death} represent the cell fitness, with high-fitness cells having low t_{div}(early occurrence of division events) and high t_{death}(resistance to death events).

Fitness memory and inheritance

At the simulation start, each cell draws its initial fitness thresholds from a uniform distribution. This is done to maximize the potential heterogeneity of the staring cells, since cells with the exact same fitness thresholds will have the same outcome regarding division and death. Upon cell division, these values may be partially or completely inherited by its daughter cells.

The inheritance of the fitness thresholds is determined by the fitness memory (f_m) of the colony: daughter cells from high f_m mothers tend to have almost identical t_{div} and t_{death}, both amongst themselves and their predecessor. On the other hand, f_m values close to 0.0 means that t_{div} and t_{death} values are largely uncorrelated between mother and daughter cells (Fig. 2A). Intermediate f_m values proportionally guide the colony towards fitness preservation ($f_m \approx 1.0$) or randomization ($f_m \approx 0.0$). f_m is constant and equal for all cells in a colony, although a treatment is able to modify it. The simulation also supports negative f_m values, which force daughter cells to have a fitness thresholds opposite to their mothers (e.g. a high fitness mother generating low fitness daughters).

Treatments

During the simulation, cells are under a given treatment (the term "treatment" is also used to refer to untreated/control scenarios). Treatments are used alongside an individual cell's t_{div} and t_{death} when defining its fate at the next frame (Fig. 1B). A treatment holds two probability density functions representing a division curve and a death curve, respectively. At population level, these curves describe the chance a cell has to divide or die as a function of its age. The division/death curves for a given treatment can be inferred from experimental data of cell age at division/death, respectively (see Supplementary Data for details). The current treatment can be modified during the course of the simulation.

Results

Simulation scenarios

The scenarios presented in Fig. 2B demonstrate the effect of treatments and f_m on colony growth patterns. In scenario *i*, no death events were triggered due to the death curve being dramatically shifted to the right. Cells from the low f_m colony divide at diverse ages, with no correlation between mother and daughter age at division. On the other hand, the offspring of the high f_m colony divides in regular intervals, mimicking the fitness of its initial cell (Fig. 2B, scenario *i*).

When cell death events are introduced, either from the simulation start (scenario *ii*) or after a treatment change (scenario *iii*), having low f_m increases the odds that at least some cells from the colony are able to survive the treatment. This is known as fractional killing in an in vitro experimental setting [14]. In contrast, a colony with high f_m faces an all-or-nothing scenario: as long as its initial cell is able to survive, its fitness thresholds are largely copied to its offspring, giving rise to a stably resistant colony (Fig. 2B, scenarios *ii* and *iii*).

Interestingly, even though low f_m leads to a higher heterogeneity in lineage tree structure, high f_m produces a higher variability in colony sizes (Fig. 2C). In fact, huge colonies are only achievable by high f_m colonies that preserve their fast division time and resistant phenotype throughout generations (Supplementary Video S1), as suggested by experimental data [9].

Experimental comparison

To evaluate how CloVarS compares to a traditional in vitro experiment, we fit the division and death curves from experimental data of manually tracked cells (see Supplementary Data for details). Two treatments were defined: *Control*, based on division times of A172 glioblastoma cells under unhindered growth, and *TMZ*, estimated from division and death times of A172 glioblastoma cells under concentrations of temozolomide that induce fractional killing. Similarly to in vitro experiments, cells under the *TMZ* condition were left to grow for 3 days prior to treatment addition. Switching to *TMZ* treatment reproduces the fractional killing effects observed in vitro (Fig. 2D, Supplementary Fig. S2).

For trees generated with Clovars, we calculated the Pearson correlation between the branch lengths of sister cells (SisSis) and of mother-daughter (MD) pairs generated with SisSis and MD f_m ranging from −1.0 to 1.0 and

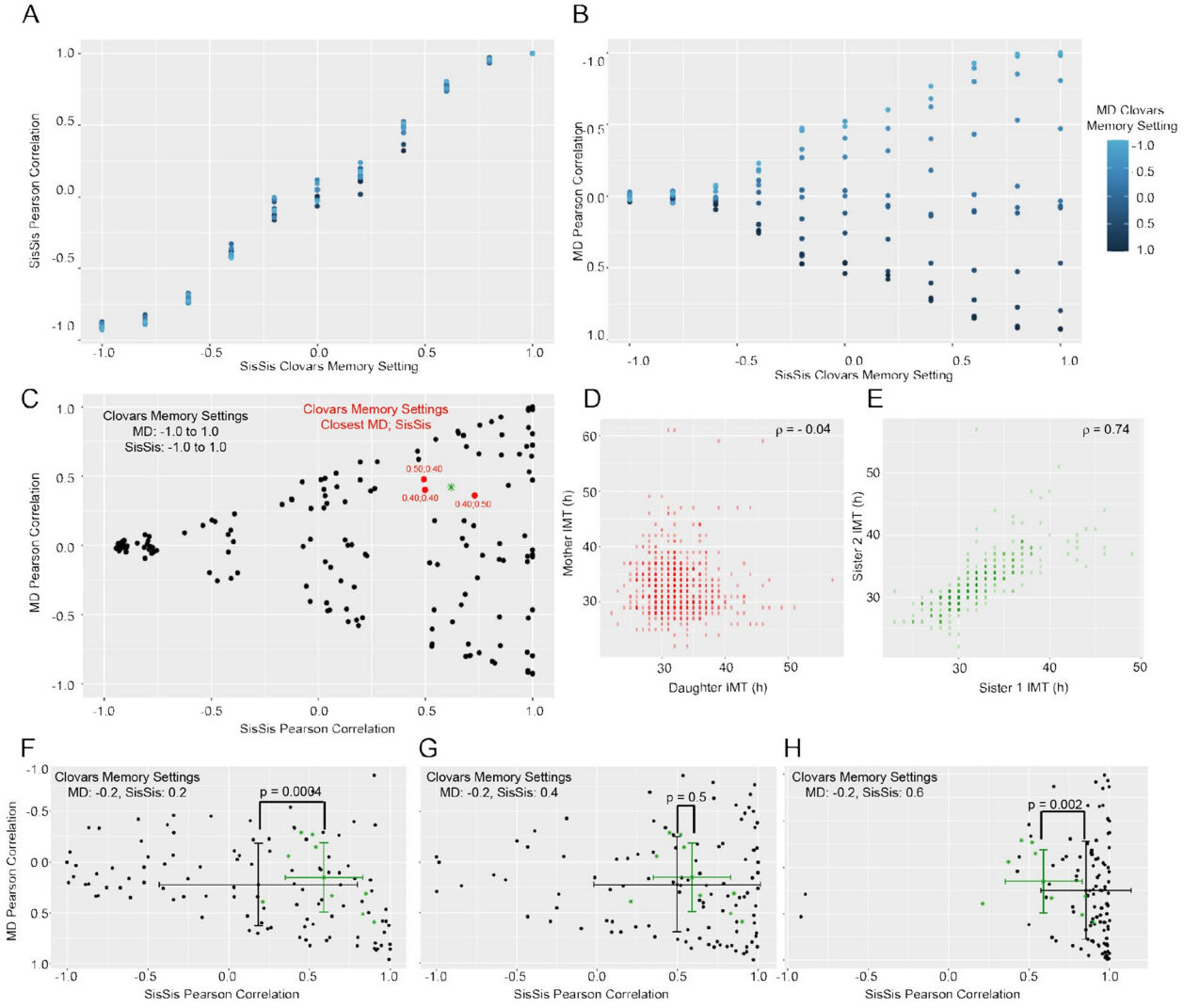

Fig. 3 Comparison to experimental results. Pearson correlation of **A** SisSis, **B** MD with different SisSis-f_m and MD-f_m in Clovars. **C** SisSis and MD Pearson correlations with f_m scan from −1.0 to 1.0 and experimental data from (Ulicna et al. 2021) highlighted in green and closest modeled tree in red, with the indicated MD-f_m and SisSis-f_m that generated these points. **D** MD and **E** SisSis intermitotic time (IMT) correlation of MD of −0.2 and SisSis of 0.6, *p*, Pearson correlation; **F-H** Clovars modeling of trees with the indicated f_m in comparison with experimental data from A172 glioma cells from our group. In black is the Average ± SD of the modeled trees and in green the data from the nine trees manually tracked of A178 glioma cells. p values are non-paired T Test

linked these to the correlations observed in Madin-Darby Canine Kidney (MDCK) epithelial cells [15]) (Fig. 3A-C). For HCT116 cells of human colon cancer [11]) (Fig. 3D and E) the calculated values generated with MD-f_m of –0.2 and SisSis-f_m of 0.6 were very close to the measured correlations of –0.03 for MD and 0.73 for SisSis correlations.

We also manually tracked cells to generate trees from A172 glioma cells and calculated the SisSis and MD correlations of each tree and then averaged these correlations. The MD-f_m of –0.2 and SisSis-f_m of 0.4 generated trees with correlations that were not statistically different from the measured trees, while changing the SisSis correlation 0.2 in both directions already generated correlations which were statistically different from the measured correlations (Fig. 3F-H). (For parameters, see Supplementary Table 5). These examples indicate that Clovars is a valuable tool to find memory conditions that create trees that describe the MD and SisSis correlations observed in real cell biology trees.

Discussion

Producing accurate and reliable single-cell microscopy data is experimentally and computationally challenging. In this sense, CloVarS can be a valuable tool for generating data for a virtually infinite number of colonies, and testing a variety of biological hypotheses. Cell signaling

fluctuations can also be explored and modeled using CloVarS (Fig. S3, Supplementary Video S2).

While other biological cell simulators exist in the literature [16–21], a major focus during CloVarS' development was to keep a biology-first mindset: the underlying rules of the simulation should not be overly complex to the point where the concepts cannot be understood by most cell biologists. The interdependence of cell fitness thresholds and treatment curves when choosing cell fate is a deliberate design choice for the simulation. It establishes a link between the individual cell fitness (t_{div} and t_{death}) and the population-derived division and death probability density functions. This is in accordance to experimental results that indicate that fitness is a dynamic phenotype that partially depends on the cell itself, but also on its surrounding environment [9, 22].

Conclusion

The interplay among individual cell fitness, colony fitness memory, and treatments are assembled into a program that is stochastic at its core, but reproduces the expected clonogenic behavior and variability as observed in vitro [23]. CloVarS allow researchers to explore biological hypotheses related to the formation of heterogeneous cell colonies and their fitness dynamics.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s44330-025-00033-8.

Supplementary Material 1.

Supplementary Material 2.

Supplementary Material 3.

Acknowledgements
This study was financed in part by CAPES - Finance Code 001; FAPERGS-Pronex 16/2551–0000473–0, FAPERGS-FAPESP, 2019/15477–3); CNPq (Universal 405033/2023-0). JLF, JHB, DT, CBM, MMO, GL are or were recipients of CNPq fellowships. KRB was a recipient of a CAPES/PNDP PD fellowship.

Authors' contributions
JLF: conceptualization, code writing, data generation, manuscript writing. GL: conceptualization, manuscript writing, project administration, JHB, KRB, DT, CBM, CBC, SS, FKM: experimental data generation for modeling, LGB, MMON, FKM: initial conceptualization and help with code writing.

Funding
This study was financed in part by CAPES - Finance Code 001; FAPERGS-Pronex 16/2551-0000473-0, FAPERGS-FAPESP, 2019/15477-3); CNPq (Universal 405033/2023-0). JLF, JHB, DT, CBM, MMO, GL are or were recipients of CNPq fellowships. KRB was a recipient of a CAPES/PNDP PD fellowship.

Data availability
General information are available at: www.ufrgs.br/labsinal/clovars. The source code for CloVarS is deposited at: https://github.com/jfaccioni/clovars.

Declarations

Ethical approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 2 March 2024 / Accepted: 29 May 2025
Published online: 16 June 2025

References

1. Spencer S, Cappell S, Tsai FC, Overton KW, Wang C, Meyer T. The proliferation-quiescence decision is controlled by a bifurcation in CDK2 activity at mitotic exit. Cell. 2013;155(2):369–83. https://doi.org/10.1016/j.cell.2013.08.062.
2. Aoki K, Kondo Y, Naoki H, Hiratsuka T, Itoh RE, Matsuda M. Propagating wave of ERK activation orients collective cell migration. Dev Cell. 2017;43(3):305-317.e5. https://doi.org/10.1016/j.devcel.2017.10.016.
3. Gagliardi PA, Dobrzyński M, Jacques MA, Dessauges C, Ender P, Blum Y, et al. Collective ERK/Akt activity waves orchestrate epithelial homeostasis by driving apoptosis-induced survival. Dev Cell. 2021;56(12):1712-1726.e6. https://doi.org/10.1016/j.devcel.2021.05.007.
4. Nathans JF, Cornwell JA, Afifi MM, Paul D, Cappell SD. Cell cycle inertia underlies a bifurcation in cell fates after DNA damage. Sci Adv. 2021;7(3):eabe3882. https://doi.org/10.1126/sciadv.abe3882.
5. Skylaki S, Hilsenbeck O, Schroeder T. Challenges in long-term imaging and quantification of single-cell dynamics. Nat Biotechnol. 2016;34(11):1137–44. https://doi.org/10.1038/nbt.3713.
6. Mattiazzi Usaj M, Styles EB, Verster AJ, Friesen H, Boone C, Andrews BJ. High-content screening for quantitative cell biology. Trends Cell Biol. 2016;26(8):598–611. https://doi.org/10.1016/j.tcb.2016.03.008.
7. Meir Z, Mukamel Z, Chomsky E, Lifshitz A, Tanay A. Single-cell analysis of clonal maintenance of transcriptional and epigenetic states in cancer cells. Nat Genet. 2020;52(77):709–18. https://doi.org/10.1038/s41588-020-0645-y.
8. Lenz G, Onzi GR, Lenz LS, Buss JH, Santos JAF, Begnini KR. The origins of phenotypic heterogeneity in cancer. Cancer Res. 2021. https://doi.org/10.1158/0008-5472.CAN-21-1940.
9. Lenz LS, Faccioni JL, Bracco PA, Santos JAF, Pereira LC, Buss JH, et al. Cancer Cell Fitness Is Dynamic. Cancer Res. 2021;81(4):1040–51. https://doi.org/10.1158/0008-5472.CAN-20-2488.
10. Hicks DG, Speed TP, Yassin M, Russell SM. Maps of variability in cell lineage trees. PLoS Comput Biol. 2019;15(2):e1006745. https://doi.org/10.1371/journal.pcbi.1006745.
11. Chakrabarti S, Paek AL, Reyes J, Lasick KA, Lahav G, Michor F. Hidden heterogeneity and circadian-controlled cell fate inferred from single cell lineages. Nat Commun. 2018;9(1):5372. https://doi.org/10.1038/s41467-018-07788-5.
12. Mohammadi F, Visagan S, Gross SM, Karginov L, Lagarde JC, Heiser LM, et al. A lineage tree-based hidden Markov model quantifies cellular heterogeneity and plasticity. Commun Biol. 2022;5(1):1258. https://doi.org/10.1038/s42003-022-04208-9.
13. Lenz G. Heterogeneity generating capacity in tumorigenesis and cancer therapeutics. Biochim Biophys Acta Mol Basis Dis. 2024;1870(5):167226. https://doi.org/10.1016/j.bbadis.2024.167226.
14. Paek AL, Liu JC, Loewer A, Forrester WC, Lahav G. Cell-to-cell variation in p53 dynamics leads to fractional killing. Cell. 2016;165(3):631–42.
15. Ulicna K, Vallardi G, Charras G, Lowe AR. Automated deep lineage tree analysis using a bayesian single cell tracking approach. Front Comput Sci. 2021;3. https://doi.org/10.3389/fcomp.2021.734559.
16. Anderson ARA, Weaver AM, Cummings PT, Quaranta V. Tumor Morphology and Phenotypic Evolution Driven by Selective Pressure from the Microenvironment. Cell. 2006;127(5):905–15. https://doi.org/10.1016/j.cell.2006.09.042.
17. Zhao B, Hemann MT, Lauffenburger DA. Modeling Tumor Clonal Evolution for Drug Combinations Design. Trends Cancer. 2016;2(3):144–58. https://doi.org/10.1016/j.trecan.2016.02.001.
18. Bandyopadhyay A, Wang H, Ray JCJ. Lineage space and the propensity of bacterial cells to undergo growth transitions. PLoS Comput Biol. 2018;14(8):e1006380. https://doi.org/10.1371/journal.pcbi.1006380.
19. Minussi DC, Henz B, Oliveira MdS, Filippi-Chiela EC, Oliveira MM, Lenz G. esiCancer: evolutionary in silico cancer simulator. Cancer Res. 2019;79(5):1010–3. https://doi.org/10.1158/0008-5472.CAN-17-3924.
20. Charlebois DA, Balázsi G. Modeling cell population dynamics. In Silico Biol. 2019;13(1–2):21–39. https://doi.org/10.3233/ISB-180470.

21. Opasic L, Scott JG, Traulsen A, Fortmann-Grote C. CancerSim: a cancer simulation package for Python 3. J Open Source Softw. 2020;5(53):2436. https://doi.org/10.21105/joss.02436.
22. Fennell KA, Vassiliadis D, Lam EYN, Martelotto LG, Balic JJ, Hollizeck S, et al. Non-genetic determinants of malignant clonal fitness at single-cell resolution. Nature. 2022;601(7891):125–31. https://doi.org/10.1038/s41586-021-04206-7.
23. Sandler O, Mizrahi SP, Weiss N, Agam O, Simon I, Balaban NQ. Lineage correlations of single cell division time as a probe of cell-cycle dynamics. Nature. 2015;519(7544):468–71.

Publisher's Note