BITS :: Call for Abstracts 2024 - Oral communication

Type	Oral communication
Session	Systems Biology
Title	Synthetic tumors: a data-driven approach to simulate the evolution of mutational profiles and clonality in cancer
All Authors	Daniela Volpatto [1,2], Roberta Sirovich [2], Simone Pernice [1], Sandro Gepiro Contaldo [1], Simone Ragaini [3], Elisa Genuardi [3], Simone Ferrero [3,4], Marco Ladetto [5], Riccardo Moia [6], Marco Beccuti [1], Francesca Cordero [1].

Affiliation

[1] Department of Computer Science, Università degli studi di Torino, Turin, Italy.
[2] Department of Mathematics, Università degli studi di Torino, Turin, Italy.
[3] Department of Molecular Biotechnologies and Health Sciences, Hematology Division, Università degli studi di Torino, Turin, Italy.
[4] Hematology Division, AOU Città della Salute e della Scienza di Torino, Turin, Italy.
[5] Hematology, Azienda Ospedaliera SS Antonio e Biagio e Cesare Arrigo, Alessandria, Italy.
[6] Division of Hematology, Department of Translational Medicine, Università del Piemonte Orientale, Novara, Italy

Motivation

Cancer is fundamentally a genetic disease driven by somatic evolution. Throughout somatic evolution, genetic and epigenetic changes propagate within a population of premalignant or cancerous cells. Unraveling the collective behaviors of cells within a tissue and their intricate interactions with the surrounding microenvironment is essential to grasp the intricacies of cancer development and progression fully [1]. Mathematical models have proven helpful in deriving a detailed understanding of mechanisms and processes in cancer and have been used to propose new experiments, suggest different treatment modalities and alter risk prognoses. In this work, we focus on understanding if data-driven approaches enable us better to understand the somatic evolutionary dynamics of human cancer. Specifically, we provide a mathematical description of the tumor's clonal evolution via stochastic modeling, considering passenger mutations and differentiating cells based on their genetic profile. Compared to pre-existing mathematical models, our approach considers both driver and passenger mutations, and it can distinguish individual driver mutations instead of just counting them.

Methods

We present an analytical workflow to elucidate cancer evolution through branching processes by integrating information obtained from variants calling for mutational profiling and longitudinal data analysis [2]. We employed a pipeline utilizing bcftools [3] for variant calling to identify and isolate variants from the collected samples. Variants with a base quality below 20, supported reads fewer than 20, and an allele frequency below 0.05 are excluded. Survival variants are further classified using CHASM [4] and CancerVar [5]. We exploited CONNECTOR [2] for longitudinal data analysis, an innovative automated computational framework designed to streamline the interpretation of longitudinal data. This framework facilitates the analysis of observations recorded at multiple time points, providing insights into system evolution. We formulated a multidimensional branching process to study cancer evolution to elucidate mutations that alter tumor fitness and influence further mutational events. This model relies on two key parameters: (i) the growth rate of cells linked to specific genetic profiles and (ii) the rate of acquiring new mutations. The selective advantage conferred by each driver mutation corresponds to an increase in the growth rate. When a driver mutation is expected to occur, it is randomly selected based on a probability distribution derived from gene dependency knowledge. Additionally, we incorporated considerations for tissue carrying capacity, which limits unchecked expansion and accounted for competitive interactions between clones sharing space and resources.

Results

The data from the FIL MCL0208 trial offer first-line high-dose chemotherapy and autologous transplantation to younger Mantle Cell Lymphoma patients. MCL0208 provided 66 patients for whom we collected data on the mutational profile based on the MCL-oriented NGS panel sequenced at diagnosis and ASO RQ-PCR minimal residual disease data generated from bone marrow (BM) and peripheral blood (PB) samples. The mutational profile data are analyzed using the variant calling pipeline described above. The patients are stratified into groups following the MRD kinetics analyzed by CONNECTOR. Specifically, CONNECTOR was run on MRD data characterized by at least three time points in bone marrow (median 6, min 3, max 9) and separately on MRD data of the patients with at least four time points in peripheral blood (median 7, min 4, max 9). CONNECTOR identified four clusters of patients with different MRD kinetics and highly significantly different disease prognosis.
The four clusters are viewed as four meta-patients, each characterized by a shared mutational profile and an average MRD curve within their respective groups.
Only the simulations obtained through the branching model matched in-silico tumors with the one created from the meta-patient of each cluster are considered. All the possible clonal evolutions and mutational events that might have resulted in the fixed profile emerge. The model analysis allows us to test theories regarding the interaction between driver mutations and predict the disease's evolution, looking for recurrent patterns within the different clusters. We implemented the model on R to simulate a tumor expansion and used mutational data from

lymphomas to calibrate the model and verify the results obtained. Our branching approach allows each subclone to (i) profile it from a genomic point of view and (ii) estimate the number of cells. To the best of our knowledge, this represents a step towards realizing a synthetic tumor to encompass the challenge of personalized medicine.

Info

[1] Sperling, Adam S., Christopher J. Gibson, and Benjamin L. Ebert. "The genetics of myelodysplastic syndrome: from clonal haematopoiesis to secondary leukaemia." Nature Reviews Cancer 17.1 (2017): 5-19.
[2] Pernice, Simone, et al. "CONNECTOR, fitting and clustering of longitudinal data to reveal a new risk stratification system." Bioinformatics 39.5 (2023): btad201.
[3] Danecek P, Bonfield JK, et al. Twelve years of SAMtools and BCFtools. Gigascience (2021) 10(2):giab008
[4] Wong WC, Kim D, Carter H, Diekhans M, Ryan MC, Karchin R. CHASM and SNVBox: toolkit for detecting biologically important single nucleotide mutations in cancer. Bioinformatics. 2011 Aug 1;27(15):2147-8. doi: 10.1093/bioinformatics/btr357. Epub 2011 Jun 17. PMID: 21685053; PMCID: PMC3137226.
[5] Li Q, Ren Z, Cao K, Li MM, Wang K, Zhou Y. CancerVar: An artificial intelligence-empowered platform for clinical interpretation of somatic mutations in cancer. Sci Adv. 2022 May 6;8(18):eabj1624. doi: 10.1126/sciadv.abj1624. Epub 2022 May 6. PMID: 35544644; PMCID: PMC9075800.

filename	-

Figure

-

Availability	-

Dissemination Material

Social

@QbioGroup

Summary

Decoding Cancer: Unveiling the mysteries of cancer progression through cutting edge methods that combine genomic profiling and cancer monitoring using complex mathematical branching processes.

Corresponding Author

Name, Surname	Simone, Pernice
Email	simone.pernice@unito.it

Submitted on	02.05.2024