Algorithmic Approaches to Study Mutational Processes in Cancer

Teresa M. Przytycka

National Center for Biotechnology Information, National Library of Medicine,
National Institutes of Health, Bethesda, MD, USA

Abstract

Mutations are the driving force of evolution, yet they underlie many diseases and, in particular, cancer. They are thought to arise from a combination of stochastic errors in DNA processing, naturally occurring DNA damage (e.g., the spontaneous deamination of methylated CpG sites), replication errors, carcinogenic exposures or cancer related aberrations of DNA maintenance machinery. These processes often lead to distinctive patterns of mutations, called "mutational signatures". Starting with the seminal work of Alexandrov at al. [1] several computational approaches have been developed to uncover such mutational signatures. However connecting mutational signatures to mutational processes is not always easy [3].

To gain insights into the relationships between mutational processes and computationally derived somatic mutation patterns (mutational signatures), we developed several complementary approaches that leverage different algorithmic techniques allowing us to link such patterns to their potential causes. For example, to investigate the genetic aberrations associated with mutational signatures, we took a network-based approach considering mutational signatures as phenotypes. Specifically, our analysis aims to answer the following two complementary questions: (i) what are functional pathways whose gene expression activities correlate with the strengths of mutational signatures, and (ii) are there pathways whose genetic alterations might have led to specific mutational signatures? To identify mutated pathways, we adopted an optimization method based on integer linear programming. Analyzing a breast cancer dataset, we identified pathways associated with mutational signatures on both expression and mutation levels. Our analysis captured important differences in the etiology of the APOBEC related signatures and the two clock-like signatures. In particular, it revealed that clustered and dispersed APOBEC mutations may be caused by different mutagenic processes. In addition, our analysis elucidated differences between two age related signatures – one of the signatures is correlated with the expression of cell cycle genes while the other has no such correlation but shows patterns consistent with the exposure to environmental/external processes [4].

Complementing this approach, we also developed a network-based method, named GENESIGNET that constructs an influence/information flow network connecting genes and mutational signatures [2]. The approach leverages sparse partial correlation among other statistical techniques to uncover dominant influence relations between the activities of network nodes. Applying GENESIGNET to cancer data sets, we uncovered important relations between mutational signatures and several cellular processes that can shed light on cancer-related processes. In particular, GENESIGNET exposed a link between the SBS8 signature of unknown etiology and the Nucleotide Excision Repair (NER) pathway.

Linking mutational signatures to molecular features can help understand the etiology and develop personalized cancer therapy. However, due to the complex and dynamic nature of tumor evolution, untangling the cause and effect relationship can be challenging and requires further integrated and comprehensive analyses.

2012 ACM Subject Classification Theory of computation → Design and analysis of algorithms; Applied computing → Life and medical sciences

Keywords and phrases Biological Networks, Cancer, Mutational Signatures, DNA Damage, DNA Repair

Digital Object Identifier 10.4230/LIPIcs.WABI.2023.1

Category Invited Talk

23rd International Workshop on Algorithms in Bioinformatics (WABI 2023).
Editors: Djamal Belazzougui and Aïda Ouangraoua; Article No. 1; pp. 1:1–1:2

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Funding *Teresa M. Przytycka*: Supported by the Intramural Research Program of the National Library of Medicine at the National Institutes of Health, USA.

References

1 Ludmil B Alexandrov, Serena Nik-Zainal, David C Wedge, Samuel Aparicio, Sam Behjati, et al. Signatures of mutational processes in human cancer. *Nature*, 500(7463):415–421, 2013. doi:10.1038/nature12477.

2 B. Amgalan, D. Wojtowicz, Y.-A. Kim, and T. M. Przytycka. Influence network model uncovers relations between biological processes and mutational signatures. *Genome Medicine*, 15, 2023.

3 Y.-A. Kim, M. D. M. Leiserson, P. Moorjani, R. Sharan, D. Wojtowicz, and T. M. Przytycka. Mutational Signatures: From Methods to Mechanisms. *Annu Rev Biomed Data Sci*, 4:189–206, July 2021.

4 Y.-A. Kim, D. Wojtowicz, R. Sarto Basso, I. Sason, W. Robinson, D. S. Hochbaum, M. D. M. Leiserson, R. Sharan, F. Vadin, and T. M. Przytycka. Network-based approaches elucidate differences within APOBEC and clock-like signatures in breast cancer. *Genome Med*, 12(1):52, May 2020.