METHODS IN MOLECULAR BIOLOGY

Series Editor
John M. Walker
School of Life and Medical Sciences
University of Hertfordshire
Hatfield, Hertfordshire, AL10 9AB, UK

For further volumes:
http://www.springer.com/series/7651

Computational Methods for Drug Repurposing

Edited by

Quentin Vanhaelen

Insilico Medicine, Inc., Rockville, MD, USA

Humana Press

Editor
Quentin Vanhaelen
Insilico Medicine, Inc.
Rockville, MD, USA

ISSN 1064-3745 ISSN 1940-6029 (electronic)
Methods in Molecular Biology
ISBN 978-1-4939-8954-6 ISBN 978-1-4939-8955-3 (eBook)
https://doi.org/10.1007/978-1-4939-8955-3

Library of Congress Control Number: 2018962410

This Humana Press imprint is published by the registered company Springer Science+Business Media, LLC, part of Springer Nature.
The registered company address is: 233 Spring Street, New York, NY 10013, U.S.A.

Preface

It is known that despite large R&D resources and expenses, the conventional drug discovery process has become increasingly time-consuming and elicits a relatively high attrition rate. With the current social and demographic trends, this results in an important disparity between the high R&D expenses, reduced number of new drugs, and unmet medical needs. To address these issues, the pharmaceutical sector continuously innovates to implement alternative approaches in order to optimize key steps of the drug discovery pipeline, essentially by focusing on a more accurate identification of promising candidates. Among them, drug repurposing is a well-known strategy to find alternative indications for drugs that have already undergone toxicology and pharma-kinetic studies but have failed in later stages during the development. Nevertheless, identifying new targets for previously approved drugs or repurposing candidates for a given indication or disease remains challenging. However, the availability of biological data of all kinds provides new opportunities to develop computational methods for accelerating the identification of potential target of interest for drug repurposing. These computational strategies attract much interest because they allow a fast identification of the most interesting candidate. The number of available methods for computational repurposing has quickly increased. All these methods have advantages and specific characteristics and requirement, for example, in terms of data required to perform the computational analysis. For scientists interested in using such techniques, identifying the appropriate type of algorithm and relevant technical information can be challenging.

The aim of this book is to provide an overview of the main techniques commonly used for performing computational drug repurposing. Each chapter has been designed by scientists whose research focuses on developing and using such techniques. In each chapter, the authors have used their experience in this field to describe in a comprehensive and accessible way the necessary steps required for the implementation and successful use of a specific repurposing method. In addition, several review chapters have been integrated in order to introduce from a larger perspective the main characteristics of the methods presented in this book.

The first chapter is a review about protein-protein interaction (PPI) interface targeting strategies. Chapter 2 presents a method combining structure-based virtual screening and molecular dynamics simulation. Chapter 3 covers a method based on the evolutionary relationships between targets of FDA-approved drugs and properties of proteins. Chapter 4 is a mining method using data from clinical trials. Chapter 5, a method based on connectivity mapping, illustrates how transcriptomic data from diseases can be reused to identify repositioning candidates.

Chapter 6 is an overview of the network-based methods, which rely on the assembly of networks to combine and exploit various kinds of information. The following chapters describe different network-based methods. The seventh chapter covers a method using bipartite graph to calculate drug pairwise similarity; whereas Chapter 8 presents a method combining disease-disease association and molecular simulation analysis. Chapter 9 describes a transcriptomics-based repurposing methodology. Chapter 10 is about the method CRAFTT, which combines transcription factor target gene sets with drug-induced expression profiling. Chapter 11 explains how to use a drug-drug interaction network to

infer pharmacological functions. The twelfth chapter presents the network propagation-based approach called DTINet for predicting clinical success of a drug target.

Chapter 13 is an introduction to the principles and types of ML algorithms. The following chapters describe methods based on ML techniques. Chapter 14 describes a machine learning method using ensemble learning for predicting drug-target interactions. The fifteenth chapter presents a regularization model for drug repurposing using electronic health records (EHRs). Chapter 16 explains a method based on support vector machine. Chapter 17 presents a machine learning algorithm, KronRLS-MKL, which integrates heterogeneous information sources into a single chemogenomic space represented by drug-target network to predict drug-target interactions. Chapter 18 describes the method Heter-LP, which can be used to predict drug-target, drug-disease, and disease-target interactions. Finally, the nineteenth chapter describes a method to compute similarities for drug-target prediction using a deep learning method.

All these methods presented here will certainly be of great interest for scientists looking into using computational drug repurposing.

To conclude, I would like to thank the Series Editor, Professor John Walker, for inviting me to edit this volume. This book would not exist without the work of all authors who collaborated for it. Their contributions are deeply acknowledged.

Rockville, MD, USA *Quentin Vanhaelen*

Contents

Contributors

Pankaj Agarwal • *Computational Biology, GSK R&D, Collegeville, PA, USA*

Salvatore Alaimo • *Department of Clinical and Experimental Medicine, University of Catania, Catania, Italy*

Hassan Aryapour • *Department of Biology, Faculty of Science, Golestan University, Gorgan, Iran*

Milad Bagheri • *Department of Biology, Faculty of Science, Golestan University, Gorgan, Iran*

Preethi Balasundaram • *Department of Biotechnology, School of Bio Sciences and Technology, Vellore Institute of Technology, Vellore, Tamil Nadu, India*

Yujia Bao • *The Massachusetts Institute of Technology, Cambridge, MA, USA*

Michael Caldwell • *The Marshfield Clinic, Marshfield, WI, USA*

Sohini Chakraborti • *Molecular Biophysics Unit, Indian Institute of Science, Bangalore, Karnataka, India*

Ivan G. Costa • *Institute for Computational Genomics, Centre of Medical Technology (MTZ), RWTH Aachen University Medical School, Aachen, Germany*

Olivier Elemento • *Department of Physiology and Biophysics, Institute for Computational Biomedicine, Weill Cornell Medicine, New York, NY, USA; Caryl and Israel Englander Institute for Precision Medicine, Weill Cornell Medicine, New York, NY, USA*

Ali Ezzat • *Biomedical Informatics Lab, School of Computer Science and Engineering, Nanyang Technological University, Singapore, Singapore*

Johannes M. Freudenberg • *Computational Biology, GSK R&D, Collegeville, PA, USA*

Yutaka Fukuoka • *Kogakuin University, Tokyo, Japan*

Kaitlyn Gayvert • *Department of Physiology and Biophysics, Institute for Computational Biomedicine, Weill Cornell Medicine, New York, NY, USA; Caryl and Israel Englander Institute for Precision Medicine, Weill Cornell Medicine, New York, NY, USA; Tri-Institutional Training Program in Computational Biology and Medicine, New York, NY, USA*

Nasser Ghadiri • *Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan, Iran*

James R. Green • *Department of Systems and Computer Engineering, Carleton University, Ottawa, ON, Canada*

Attila Gursoy • *Department of Computer Engineering, Koc University, Istanbul, Turkey*

Farideh Halakou • *Department of Computer Engineering, Koc University, Istanbul, Turkey*

Nivya James • *Department of Biotechnology, School of Bio Sciences and Technology, Vellore Institute of Technology, Vellore, Tamil Nadu, India*

Anil G. Jegga • *Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA; Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, USA; Department of Computer Science, University of Cincinnati College of Engineering, Cincinnati, OH, USA*

Xiao Ji • *Computational Biology, GSK R&D, Collegeville, PA, USA*

Rohini Kanagavelu • *Department of Biotechnology, School of Bio Sciences and Technology, Vellore Institute of Technology, Vellore, Tamil Nadu, India*

Evangelos Karatzas • *Department of Informatics and Telecommunications, University of Athens, Athens, Greece*
Ramanathan Karuppaswamy • *Department of Biotechnology, School of Bio Sciences and Technology, Vellore Institute of Technology, Vellore, Tamil Nadu, India*
Ozlem Keskin • *Department of Chemical and Biological Engineering, Koc University, Istanbul, Turkey*
George Kolios • *Laboratory of Pharmacology, Department of Medicine, Democritus University of Thrace, Alexandroupolis, Greece*
Zhaobin Kuang • *The University of Wisconsin, Madison, WI, USA*
Chee-Keong Kwoh • *Division of Software and Information Systems, School of Computer Science and Engineering, Nanyang Technological University, Singapore, Singapore*
Jiao Li • *Institute of Medical Information/Library, Chinese Academy of Medical Sciences/ Peking Union Medical College, Beijing, China*
Xiaoli Li • *Data Analytics Department, Institute for Infocomm Research, A-Star, Singapore, Singapore*
Hetong Ma • *Institute of Medical Information/Library, Chinese Academy of Medical Sciences/Peking Union Medical College, Beijing, China*
Sayoni Maiti • *Department of Biotechnology, School of Bio Sciences and Technology, Vellore Institute of Technology, Vellore, Tamil Nadu, India*
André C. A. Nascimento • *Department of Computing, UFRPE, Recife, Brazil*
Victoria Ngo • *Betty Irene Moore School of Nursing, University of California Davis, Sacramento, CA, USA*
Ruth Nussinov • *Cancer and Inflammation Program, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, National Cancer Institute at Frederick, Frederick, MD, USA; Department of Human Molecular Genetics and Biochemistry, Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel*
E. Sila Ozdemir • *Department of Chemical and Biological Engineering, Koc University, Istanbul, Turkey*
David Page • *The University of Wisconsin, Madison, WI, USA*
Peggy Peissig • *The Marshfield Clinic, Marshfield, WI, USA*
Ricardo B. C. Prudêncio • *Center of Informatics, UFPE, Recife, Brazil*
Alfredo Pulvirenti • *Department of Clinical and Experimental Medicine, University of Catania, Catania, Italy*
Gayatri Ramakrishnan • *Molecular Biophysics Unit, Indian Institute of Science, Bangalore, Karnataka, India; Indian Institute of Science Mathematics Initiative, Indian Institute of Science, Bangalore, India; Division of Genetics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA*
Maryam Lotfi Shahreza • *Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan, Iran*
Hon-Cheong So • *School of Biomedical Sciences, The Chinese University of Hong Kong, Shatin, Hong Kong; KIZ-CUHK Joint Laboratory of Bioresources and Molecular Research of Common Diseases, Kunming Zoology Institute of Zoology, Kunming, China*
Farzin Sohraby • *Department of Biology, Faculty of Science, Golestan University, Gorgan, Iran*
George M. Spyrou • *Bioinformatics ERA Chair, The Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus*
Narayanaswamy Srinivasan • *Molecular Biophysics Unit, Indian Institute of Science, Bangalore, Karnataka, India*

Ron Stewart • *The Morgridge Institute for Research, Madison, WI, USA*
Eric Wen Su • *Advanced Analytics Hub, Eli Lilly and Company, Indianapolis, IN, USA*
James Thomson • *The Morgridge Institute for Research, Madison, WI, USA*
Lucreţia Udrescu • *Faculty of Pharmacy, "Victor Babeş" University of Medicine and Pharmacy Timişoara, Timişoara, Romania*
Mihai Udrescu • *Department of Computer and Information Technology, Politehnica University of Timişoara, Timişoara, Romania; Timişoara Institute of Complex Systems, Timişoara, Romania*
Shanthi Veerappapillai • *Department of Biotechnology, School of Bio Sciences and Technology, Vellore Institute of Technology, Vellore, Tamil Nadu, India*
Jiayang Wang • *Institute of Medical Information/Library, Chinese Academy of Medical Sciences/Peking Union Medical College, Beijing, China*
Yunguan Wang • *Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA*
Rebecca Willett • *The University of Wisconsin, Madison, WI, USA*
Rachael Sze Nga Wong • *Department of Biomedical Informatics, School of Medicine, University of California San Diego, San Diego, CA, USA*
Min Wu • *Data Analytics Department, Institute for Infocomm Research, A-Star, Singapore, Singapore*
Jaswanth Yella • *Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA; Department of Computer Science, University of Cincinnati College of Engineering, Cincinnati, OH, USA*
Kai Zhao • *School of Biomedical Sciences, The Chinese University of Hong Kong, Shatin, Hong Kong*
Si Zheng • *Institute of Medical Information/Library, Chinese Academy of Medical Sciences/ Peking Union Medical College, Beijing, China*
Nansu Zong • *Department of Biomedical Informatics, School of Medicine, University of California San Diego, San Diego, CA, USA*