Proceedings of the International Conference on Industrial Engineering and Operations Management
Toronto, Canada, October 23-25, 2019

Computational Modeling Using Multi-omics to Extract Early Predictive Signatures of T-cells Quality

Valerie Y. Odeh-Couvertier and Wandaliz Torres-García
Department of Industrial Engineering
University of Puerto Rico
Mayaguyez, PR 00681
valerie.odeh@upr.edu, wandaliz.torres@upr.edu

Abstract

Chimeric Antigen Receptor (CAR) T-cell therapy involves the genetic modification of patient's autologous T-cells to find and attack cancer cells throughout the body. Establishing critical quality attributes (CQAs) and critical process parameters (CPPs) is a crucial task for ensuring the potency, safety, and consistency of this therapy. To better understand these therapies, a design of experiments evaluating the expansion of T cells was carried out. Through a supervised learning approach, multi-omics predictors (i.e. secretomes and NMR metabolomics features) can be used to understand T cell behavior based on growth and memory responses. The purpose of this work is to develop a computational pipeline that enables the characterization of multi-omics profiles that are predictive of quality responses at early stages of the manufacturing process. This includes the design of a computational tool that will measure the predictive power of omics variables and the sensitivity of these models to highly correlated predictors. The computational tool has been developed using mathematical modeling and machine learning techniques such as Random Forest (RF), Gradient Boosted Trees (GBT), Support Vector Machines (SVM), and Symbolic Regression (SR). A consensus measurement between the different models was used in order to identify potential CQAs and CPPs. The biological meaning of these features is then assessed through discussions with the domain experts. A degree of consensus was achieved throughout the models in identifying important variables when modeling the percentage of T cells that are CCR7+CD62L+CD4+ (CD4_memory_fraction). The models presented high prediction performance with R^2 ranging from 75% to 95% where SVM and SR provided the least and best prediction performance, respectively. The ranking of important features from these best-performing models can be misleading if these features are highly correlated. Hence, to mitigate the computational impact of highly correlated predictors, an approach involving the clustering of these variables is proposed before the model fitting process. Preliminary results showed that not only was this methodology able to rank groups of correlated predictors but was also able to improve performance for RF models. The findings of this work could enable the discovery of new knowledge necessary to achieve scalable biomanufacturing of CAR-T cell therapies in an automated manner.

Keywords

CAR T cells, chimeric antigen receptor, multi-omics, mathematical modeling, and machine learning.

Acknowledgements

We acknowledge that this work was possible by the collaboration of Nathan Dwarshuis, Danning Huang, Facundo Fernandez, and Krishnendu Roy from Georgia Institute of Technology, Maxwell Colonna and Arthur Edison from University of Georgia, and Theresa Kotanchek from Evolved Analytics.

This material is based upon work supported by the National Science Foundation under Grant No. EEC-1648035.

Biographies

Valerie Y. Odeh-Couvertier is an Industrial Engineering graduate student from the University of Puerto Rico-Mayaguez Campus, where she also earned a B.S. in Industrial Engineering. As an undergraduate student she had the opportunity of acquiring industry experience working as a manufacturing intern at HP Inc and as a warehouse intern at Lily del Caribe. Her research interest includes the application of Industrial Engineering methodologies on areas related to data mining, machine learning, and bioinformatics. She is currently working with the NSF Engineering Research Center for Cell Manufacturing Technologies to further efforts on transforming the manufacture of cell-based therapeutics into large scale, lower cost, and high-quality engineered process, for broad industry and clinical use.

Wandaliz Torres-García is an Associate Professor in the Department of Industrial Engineering at the University of Puerto Rico Mayagüez Campus. She has a Ph.D. in Industrial Engineering with concentrations in multivariate statistical learning and data mining from Arizona State University. She also has several years of research experience in the application and development of data mining methodologies to handle large amounts of multivariate data sets in the field of bioinformatics. Her two-year postdoctoral fellowship at The University of Texas MD Anderson deepened her computational proficiency by collaborating with a highly interdisciplinary research team in the area of cancer genomics and high-performance computing. Furthermore, her experiences in highly diverse research teams have made her an active advocate to participate in programs at various levels (K-12, undergraduate & graduate students and our local community) that promotes learning, training and mentoring within a collaborative and multidisciplinary environment needed to solve difficult problems.