

Visuelle Analyse klinischer Multiomikdaten in biologischen Netzwerken

Einführung

Das Universitätsklinikum Tübingen hat in laufenden Forschungsprojekten große Mengen unterschiedlicher Datensätze von Patienten gesammelt, die ein hohes Diabetes-Risiko aufweisen (Typ II) [1]. Diese Daten sind mit herkömmlichen Methoden nur schwer zu analysieren. Unter anderem ist die Fusion und Korrelation von Informationen aus verschiedenen Quellen zur Unterstützung einer umfassenden Analyse eine ungelöste Fragestellung. Insbesondere für große Datenmengen ist die visuelle Analytik hierfür ein leistungsstarkes Werkzeug. Gleichzeitig wurden durch systembiologische Forschung zahlreiche biologische Netzwerkkarten erstellt, die eine Vielzahl zellulärer Prozesse umfassend abbilden [2]. Diese statischen Karten sollen in diesem Projekt genutzt werden, um die Patientendaten in ihrem funktionalen Kontext darzustellen. Ein besonderer Schwerpunkt liegt auf Lipidomen und mitochondrialen Systemen, die in der aktuellen Literatur noch nicht ausreichend abgedeckt sind.

Zielsetzung

In einer früheren Arbeit an der Universität Tübingen wurde hierfür bereits ein auf Java basierendes Tool entwickelt, dessen Funktionsumfang zunächst analysiert werden soll und welches ggf. als Referenz genutzt werden kann¹.

Im Rahmen dieser Arbeit sollen anschließend die von den Projektpartnern zur Verfügung gestellten Multiomikdaten mittels gängiger Plots graphisch dargestellt werden, um einen ersten Eindruck von deren Beschaffenheit zu gewinnen [3]. Erst dann soll von gegebenen Netzwerkkarten relevante Teilkarten (sog. Pfade) identifiziert und nach Rang sortiert werden. Ausgehend von einer sehr umfassenden Karte könnten beispielsweise irrelevante Bereiche eliminiert werden, um die Daten im Kontext des verbleibenden Netzwerks übersichtlicher darzustellen und den Nutzern neue Erkenntnisse durch effiziente Exploration der Daten zu ermöglichen. Hierbei soll auf die zuvor erstellten Plots Bezug genommen werden.

Ziel der Arbeit ist die Implementierung eines neuen webbasierten Tools für die visuelle Analyse, das in Standard-Webbrowsern läuft und den Projektpartnern Möglichkeiten zur Analyse ihrer Daten in Verbindung mit interaktiv anpassbaren Netzwerken bietet.

¹ http://www.cogsys.cs.uni-tuebingen.de/forschung/terminiert/diabetes/welcome_e.html

Anforderungen

Interesse an biochemischen Netzwerken, Visualisierung und der Verarbeitung gängiger Dateiformate. Kenntnisse im Programmieren mit JavaScript und D3 sowie Grundkenntnisse über Multiomikdaten sind von Vorteil.

Betreuer

Jun.-Prof. Dr. Andreas Dräger draeger@informatik.uni-tuebingen.de

Jun.-Prof. Dr. Michael Krone michael.krone@uni-tuebingen.de

Literatur

- [1] X. Liu *et al.*, "Quality Control of Serum and Plasma by Quantification of (4E,14Z)-Sphingadienine-C18-1-Phosphate Uncovers Common Preanalytical Errors During Handling of Whole Blood," *Clinical Chemistry*, vol. 64, no. 5, pp. 810–819, May 2018.
- [2] H.-Y. Wu, M. Nöllenburg, and I. Viola, "A Visual Comparison of Hand-Drawn and Machine-Generated Human Metabolic Pathways," presented at the EG/VGTC Conference on Visualization, 2018.
- [3] T. Munzner, *Visualization Analysis and Design*. CRC Press, 2015.