

D Summary translation



please translate this into german:

This PhD thesis integrates biomedical research and data science, focusing on an *in vitro* model for studying myeloma cell dissemination and a Python-based tool, *plotastic*, for semi-automated analysis of multidimensional datasets. Two major challenges are addressed: (1) understanding the early steps of myeloma dissemination and (2) improving data analysis efficiency to address the complexity- and reproducibility bottlenecks currently present in biomedical research.

In the experimental component, primary human mesenchymal stromal cells (hMSCs) are co-cultured with INA-6 myeloma cells to study cell proliferation, attachment, and detachment via time-lapse microscopy. Key findings reveal that detachment often follows cell division, predominantly driven by daughter cells. Novel separation techniques were developed to isolate myeloma subpopulations for further characterization by RNAseq, cell viability, and apoptosis assays. Differential expression of adhesion and retention factors upregulated by INA-6 cells correlates with patient survival. Overall, this work provides insights into myeloma dissemination mechanisms and identifies genes that potentially counteract dissemination through adhesion, which could be relevant for the design of new therapeutics.

To manage complex data, a Python-based software named *plotastic* was developed that streamlines analysis and visualization of multidimensional datasets. *plotastic* is built on the idea that statistical analyses are performed based on how the data is visualized. This approach not only simplifies data analysis, but semi-automates analysis in a standardized statistical protocol. The thesis becomes a case study as it reflects on the application of *plotastic* to the *in vitro* model, demonstrating how the software facilitates rapid adjustments and refinements in data analysis and presentation. Such efficiency is crucial for handling semi-big data transparently, which *despite being manageable* is complex enough to complicate analysis and reproducibility.

Together, this thesis illustrates the synergy between experimental methodologies and advanced data analysis tools. The *in vitro* model provides a robust platform for studying myeloma dissemination, while *plotastic* addresses the need for efficient data analysis. Combined, they offer a comprehensive approach to handling complex experiments,

advancing both cancer biology and research practices, in support of exploratory and transparent analysis of challenging phenomena.



Diese Doktorarbeit integriert biomedizinische Forschung und Datenwissenschaft und konzentriert sich auf ein *in vitro*-Modell zur Untersuchung der Ausbreitung von Myelomzellen sowie ein auf Python basierendes Werkzeug namens *plotastic* zur halbautomatisierten Analyse multidimensionaler Datensätze. Zwei Hauptprobleme werden angegangen: (1) das Verständnis der frühen Schritte der Myelomausbreitung und (2) die Verbesserung der Effizienz der Datenanalyse, um die derzeitigen Engpässe bei Komplexität und Reproduzierbarkeit in der biomedizinischen Forschung zu bewältigen.

Im experimentellen Teil werden primäre humane mesenchymale Stromazellen (hMSCs) gemeinsam mit INA-6-Myelomzellen kultiviert, um Zellproliferation, Anhaftung und Ablösung mittels Zeitraffer-Mikroskopie zu untersuchen. Die wichtigsten Ergebnisse zeigen, dass die Ablösung oft auf die Zellteilung folgt, die hauptsächlich durch Tochterzellen vorangetrieben wird. Neue Trenntechniken wurden entwickelt, um Myelom-Subpopulationen für weitere Charakterisierung durch RNAseq, Zellviabilitäts- und Apoptosetests zu isolieren. Die differentielle Expression von Adhäsions- und Retentionsfaktoren, die von INA-6-Zellen hochreguliert werden, korreliert mit dem Überleben der Patienten. Insgesamt liefert diese Arbeit Einblicke in die Mechanismen der Myelomausbreitung und identifiziert Gene, die potenziell der Ausbreitung durch Adhäsion entgegenwirken, was für die Entwicklung neuer Therapeutika relevant sein könnte.

Zur Verwaltung komplexer Daten wurde eine auf Python basierende Software namens *plotastic* entwickelt, die die Analyse und Visualisierung multidimensionaler Datensätze vereinfacht. *Plotastic* basiert auf der Idee, dass statistische Analysen basierend darauf durchgeführt werden, wie die Daten visualisiert werden. Dieser Ansatz vereinfacht nicht nur die Datenanalyse, sondern automatisiert die Analyse auch teilweise nach einem standardisierten statistischen Protokoll. Die Arbeit wird zu einer Fallstudie, da sie sich auf die Anwendung von *plotastic* auf das *in vitro*-Modell bezieht und zeigt, wie die Software schnelle Anpassungen und Verfeinerungen in der Datenanalyse und -präsentation erleichtert. Eine solche Effizienz ist entscheidend für die transparente Handhabung von halb-großen Daten, die trotz ihrer Handhabbarkeit komplex genug sind, um Analyse und Reproduzierbarkeit zu erschweren. Zusammen zeigt diese Dissertation die Synergie zwischen experimentellen Methoden und fortgeschrittenen Datenanalysewerkzeugen auf. Das *in vitro*-Modell bietet eine robuste Plattform zur Untersuchung der Myelomausbreitung, während *plotastic* den Bedarf an effizienter Datenanalyse deckt. In Kombination bieten sie einen umfassenden Ansatz zur Bewältigung komplexer Experimente, der sowohl die Krebsbiologie als auch die Forschungspraktiken vorantreibt und die explorative und transparente Analyse herausfordernder Phänomene unterstützt.

ChatGPT kann Fehler machen. Überprüfe wichtige Informationen.