

Martina Emde-Rajaratnam
Dr. sc. hum.

Transcriptome Profiling Assessing Pathogenesis and Prognosis of Plasma Cell Dyscrasias –Bioinformatic Basis for Clinical Application

Fach/Einrichtung: Innere Medizin
Doktorvater: Priv.-Doz. Dr. med. Dr. biol. hom. Dirk Hose

Hintergrund. Das multiple Myelom ist eine hämatologische Erkrankung, die durch die Ansammlung maligner Plasmazellen im Knochenmark gekennzeichnet ist. Im Labor für Myelomforschung des Universitätsklinikums Heidelberg wurden Affymetrix U133 2.0-Microarray Genexpressionsdaten von 839 Patienten in erweiterter klinischer Routine untersucht. Die RNA-Sequenzierung wurde als neuartige Methode vorgestellt, die der Microarray Analyse überlegen sein soll, hinsichtlich Präzision, geringerer Menge an Input-RNA und der Möglichkeit Transkripte sowie Spleißvarianten zu analysieren. Daher bestand das Hauptziel dieser Arbeit darin, die bioinformatische Grundlage für die Implementierung der RNA-Sequenzierung in der angewandten translationalen Myelomforschung und erweiterten klinischen Routine zu schaffen. Das beinhaltet i) die Erstellung einer praktikablen Analysepipeline für RNA-Sequenzierungsdaten, ii) die Übertragung und die Verbindung von aktuellen Einteilungs- und Klassifizierungsmethoden basierend auf Microarrays zur zukünftigen RNA-Sequenzierungstechnologie, iii) die Entdeckung neuer prognostischer Gene und die Entwicklung einer auf RNA-Sequenzierung basierenden Risikoeinteilung, iv) die Analyse potentieller und insbesondere anwendbarer therapeutische Zielgene in Bezug auf Expression, alternatives Spleißen und Mutationen und v) die prospektive Testung dieser theoretischen Zielgen-Analysestrategien in einer großen Patientenkohorte in drei klinisch relevanten Beispielen: BCMA, bei der Entwicklung des gerade in klinischen Tests befindlichen, bispezifischen T-Zell-Antikörpers CC-93269, CD38, zur Analyse der klinischen Anwendbarkeit von Anti-CD38-Behandlung, und Krebs-Hoden-Antigenen als Impfziel unter Bewertung des möglichen Mechanismus der Vorabresistenz.

Methoden. Für die Risikostratifizierung wurden RNA-Sequenzierungs Proben von 535 Myelom Patienten in drei Gruppen eingeteilt, eine Trainings- (~40%), eine Validierungs- (~20%) und eine Testgruppe (~40%). Alle Proben wurden mit STAR aligniert und die resultierenden Read Anzahlen mit edgeR normalisiert. Jede Microarray-Klassifizierung wurde übertragen durch i) Translation der Microarray-Sondennamen in Ensembl-Gennamen, ii) Erzeugung einer RNA-Sequenzierungs-Klassifizierung, die so ähnlich wie möglich ist zur Microarray-Klassifizierung, iii) Berechnung neuer Grenzwerte und iv) die Einteilung der Patienten. Die Validierung der Klassifikationen erfolgte durch die Untersuchung der Proportionen der Gruppen und dem Vergleich der Überlebenszeitanalyse zu den Microarray-Klassifikationen von 534 Patienten mit multiplem Myelom. Zusätzlich wurden laufende Log-Rank-Algorithmen verwendet, um prognostische Gene zu bestimmen und eine neue auf RNA-Sequenzierung basierende Klassifizierung zu erstellen, den HDHRS. Für die Analyse potentieller Zielgene wurde die Expression pro Zielgen, z.B. *BCMA* oder *CD38*, bestimmt, eine Definition für "fehlende" und "vorhandene" Expression etabliert und alternative Spleißvarianten bewertet. Eine Pipeline wurde entwickelt und an 142 asymptomatischen, 69 rezidierten und 767 symptomatische Myelompatienten der CoMMpass-Kohorte getestet.

Ergebnisse. Die RNA-Sequenzierung kann bei 90% aller Patienten durchgeführt werden (im Gegensatz zu 80% bei Mikroarray) und die Qualität der RNA-Sequenzierungsdaten ist in 97% der Fälle ausreichend. Eine RNA-Sequenzierungspipeline wurde erfolgreich implementiert. Proliferation basierte Risikobewertung (GPI), Risikoklassifizierungen (RS, UAMS70, EMC92 und IFM15), sowie molekulare Klassifikationen (MC, TC) sind signifikant prognostisch und ebenso prädiktiv für die RNA-Sequenzierung wie auf Microarrays. Der neue HDHRS umfasst 53 prognostische Gene und unterscheidet drei Gruppen signifikant. Alle Klassifikationen wurden in der unabhängigen Testgruppe validiert. Die Klassifikationen sind auf der CoMMpass Kohorte signifikant, unter Berücksichtigung Labor- und bioinformatischen Variationen. Neunzehn potenzielle, fünf theoretische und das mutierte Zielgen *BRAF* wurden analysiert. Die angreifbaren Zielgene *BCMA*, *CD38* und *CD74* sind in allen Patienten exprimiert, die Krebs-Hoden-Antigene *MAGEA3* (33%), *RHAMM* (88%) und *NYESO1/2* (12%) in Subfraktionen von Myelompatienten und 2% zeigen die Expression des mutierten (V600E) *BRAF* Genes. Um einen möglichen Mechanismus der Vorabresistenz auszuschließen, wurden die Spleiß-Varianten von *BCMA* und *CD38* untersucht. Für beide Zielgene wird nur ein Haupttranskript exprimiert.

Diskussion. Die RNA-Sequenzierung wurde erfolgreich implementiert und in der Myelomforschung und der erweiterten klinischen Routine angewandt. Alle untersuchten Risiko-Einteilungen und Klassifikationen konnten erfolgreich übertragen werden und eine neue auf RNA-Sequenzierung basierende Klassifizierung (HDHRS) konnte entwickelt und validiert werden. Die Zielgen Analyse wurde am Beispiel von *BCMA*, *CD38* und Krebs-Hoden-Antigenen als Impfzielgene in Bezug auf Patientenkohorten beispielhaft dargestellt. Für *BCMA* wurde der Anteil der Patienten und Plasmazellvorläufer-Stadien bestimmt, die das Zielgen exprimieren, um mögliche klinische Nebenwirkungen des entwickelten Medikaments CC-63269 zu bestimmen. Dies veranschaulicht die Nützlichkeit von RNA-Sequenzierung bei der Bewertung potenzieller Zielgene für die Wirkstoffforschung. Ebenso ist RNA-Sequenzierung nützlich, um die Expression vor Beginn der Therapie zu bewerten, um Patienten von der Behandlung auszuschließen, denen die Expression des jeweiligen Zielgens fehlt. *BCMA* und *CD38* Spleißvariantenanalysen zeigten, dass alternatives Spleißen nicht die Ursache für Resistenzen gegen Anti-*BCMA*- und Anti-*CD38*-Behandlung sind. Beide Analysestrategien können und werden direkt auf zukünftige Wirkstoff Entwicklungen angewendet werden.

Somit ist diese Dissertation der Grundsatzbeweis, dass alle Schritte, die notwendig sind für eine zielgerichtete und risikoadaptierte Behandlungsstrategie beim Myelom, mit RNA-Sequenzierung durchführbar sind. Die Ergebnisse dieser Arbeit, einschließlich der entwickelten Pipeline, sind der grundlegende Schritt für die Implementierung von RNA-Sequenzierung für risikoadaptierte und zielgerichtete Behandlungsstrategien in der klinischen Routine.