

Elias Ulrich
Dr. med.

Analysis of non-coding regulatory variants in the genome of pediatric brain tumors based on whole-genome and RNA sequencing data

Fach/Einrichtung: Deutsches Krebsforschungszentrum (DKFZ)
Doktorvater: Prof. Dr. med. Stefan Pfister

Die Entstehung von Tumoren geht in vielen Fällen auf spezifische genetische Veränderungen zurück, welche die Entwicklung der Krebserkrankung antreiben (Treibermutationen). Die zuverlässige Erkennung und Analyse dieser Treibermutationen in Genomdaten verschiedener Tumorentitäten ist eine zentrale Herausforderung der Krebsgenomforschung und klinischen Bioinformatik. Für Treibermutationen in den rund 2% der menschlichen Erbinformation, welche für Proteine kodieren, ist die Erkennung relevanter genetischer Veränderung bereits etabliert. Da der anteilmäßig deutlich überwiegende, nicht proteinkodierende Bereich des Genoms jedoch zu 80% funktionell aktiv ist und für einen großen Anteil aller genetischer Erkrankungen verantwortlich ist, ist es von großer Wichtigkeit diese genomischen Anteile auch in die Suche krebstreibender genetischer Veränderungen mit einzubeziehen. Die Veränderung der genetischen Expression verschiedener oft räumlich-naher Gene ist dabei ein häufiger Mechanismus dieser nicht-kodierender Treibermutationen, welche dann auch als regulatorische Varianten bezeichnet werden und die Aminosäurestruktur der translatierten Gene nicht verändern.

Die Beschreibung der molekularen Landschaft und der typischerweise involvierten Treiber hat in der pädiatrischen Onkologie die Einführung einer zunehmend auf molekularen Eigenschaften basierten Klassifikation wichtiger pädiatrischer Tumorentitäten ermöglicht. Die Aufklärung der molekularen Treiber ist außerdem essenziell für die Entwicklung und Erprobung individueller Therapiekonzepten. Medulloblastome und pädiatrische hochgradige Gliome, sind ausgesprochen heterogene Tumorentitäten. Trotz moderner Therapieansätze ist die Prognose teils sehr schlecht. Für einen relevanten Teil dieser Tumore steht die Beschreibung der involvierten krebstreibenden molekularen Veränderungen noch aus. Die Analyse der nicht proteinkodierenden genetischen Bereiche erscheint dabei ein vielversprechender Ansatz diese Mechanismen für die verbleibenden Tumore aufzuklären.

Mit der Entdeckung des Enhancer Hijackings von GF11/GF11B und PRDM6 in Medulloblastomen konnte die pädiatrische Neuroonkologie in der Vergangenheit bereits einen wichtigen Beitrag in der Beschreibung möglicher Pathomechanismen regulatorischer Varianten leisten. Diese sind keineswegs auf diese Tumorentität beschränkt.

Ziel dieser Arbeit war es ein zuverlässiges Verfahren zur Analyse von Ganzgenom- und RNA-Sequenzierungsdaten auf regulatorische Varianten zu entwickeln und damit Medulloblastome sowie pädiatrische hochgradige Gliome zu untersuchen.

Verschiedene existierende bioinformatische Tools wie cis-X oder CESAM versprechen diese Aufgabe zu lösen, haben jedoch jeweils individuelle Stärken und Schwächen.

Die im Rahmen dieser Arbeit entwickelte Plattform Revana kombiniert mehrere Funktionen bestehender Methoden und ergänzt sie u.a. um die Berücksichtigung bekannter genspezifischer regulatorischer Regionen, Aggregation der analysierten Proben und umfassende grafische und interaktive Ergebnisbewertungsfunktionen. Das Tool wurde in dem Journal Bioinformatics veröffentlicht und steht als R-Package auf Github zum Download zur Verfügung.

Die Analyse von Sequenzierungsdaten von 170 Medulloblastomen und 172 pädiatrischen hochgradigen Gliomen mit Revana konnte vorbeschriebene regulatorische Varianten zuverlässig reproduzieren sowie weitere Kandidaten für nicht-kodierende regulatorische Treibermutationen offenlegen.

In der Kohorte der Medulloblastome wurden 10 von 11 der bekannten Enhancer Hijacking Events von PRDM6 und GFI1B wiedergefunden und diese bei weiteren 3 Proben beschrieben. Vorbeschriebene Genfusionen von PVT1 wurden außerdem als Kandidaten für regulatorische Treibermutationen identifiziert.

In der Kohorte der pädiatrischen hochgradigen Gliome konnten bekannte Genfusionen in MET und MN1 als regulatorische Varianten reproduziert werden. Außerdem wurden eine Reihe ähnlicher struktureller Varianten auf Chromosom 6 auffällig, welche systematisch zur Hochregulation des Gens NR2E1 führen. Eine wiederkehrend durch die genomischen Rearrangements in die räumliche Nähe von NR2E1 gebrachte Region auf Chromosom 6 wurde anschließend mithilfe von Chromatinimmunopräzipitationsdaten und Hi-C Daten auf regulatorisch aktive Enhancer-Bereiche untersucht. Diese könnten für die Veränderung der transkriptionellen Aktivität von NR2E1 verantwortlich sein und damit eine weitere Gruppe von Enhancer-Hijacking betroffener kindlicher Hirntumore begründen.

Revana verspricht mit seinen vielseitigen Analysen und Illustrationen, Einblicke in potenziell involvierte biologische Mechanismen der entdeckten Kandidaten für regulatorische Varianten. Als statistisches bioinformatisches Tool ist es nicht in der Lage, eindeutige Kausalitäten zu beweisen. Die abschließende Klärung der tumortreibenden Eigenschaften nicht-kodierender regulatorischer Varianten wie beispielsweise der gefundenen regulatorischen Varianten von NR2E1 unterliegt somit nach wie vor der Durchführung laborexperimenteller Studien. Dennoch kann es einen wichtigen Beitrag leisten, Treibermutationen in Tumoren zu identifizieren. Damit vermag Revana den Weg für die Untersuchung neuer pharmazeutischer Ziele und Medikamente zu ebnen, welche insbesondere in aktuell prognostisch ungünstigen Tumorentitäten wie pädiatrischen hochgradigen Gliomen dringend benötigt wird.