

Peter Bengert  
Dr. med.

## **Systematische computergestützte Identifizierung biomedizinisch wichtiger regulatorischer Ribonukleinsäure-Elemente**

Geboren am 10.02.1978 in Speyer  
Staatsexamen am 7.6.2005 an der Universität Heidelberg

Promotionsfach: Biochemie  
Doktorvater: Prof. Dr. med. Thomas Dandekar

Die Regulation auf der Ebene der RNA ist ein Schlüssel zum Verständnis zahlreicher medizinischer Prozesse, wie beispielsweise dem Eisenstoffwechsel (Iron-responsive elements in mRNAs), dem Stoffwechsel pathogener Bakterien (Riboswitch RNAs) und parasitärer Erreger (trans-spleißende RNAs) oder der Differenzierung und Regulation von Zellen (microRNAs und Zielsequenzen). Zusätzlich brachten die umfangreichen Genomsequenzierungsprojekte der letzten Jahre eine große Menge an Sequenzdaten hervor.

Das Ziel dieser Doktorarbeit bestand darin, neue bioinformatische Methoden zur Analyse von RNA-Sekundärstrukturelementen zu erstellen, um damit diese Sequenzdaten nach solchen Elementen abzusuchen.

Es wurde eine Suchstrategie entworfen, die eine Nukleotidsequenz mit Hilfe von Methoden der Mustererkennung auf ihre Primärsequenz hin untersucht. In einem zweiten Schritt werden anhand von bekannten Methoden der Energieminimierung die Sekundärstruktur eines bestimmten Teils der Primärsequenz ermittelt und verschiedene Parameter errechnet. Ein Abgleich dieser Parameter mit einem Katalog an Filterregeln führt dann im Erfolgsfalle zur Ausgabe der Sequenz mit der dazugehörigen Sekundärstruktur als Treffer.

Dann wurde diese Suchstrategie auf verschiedene RNA-Sekundärstrukturelemente wie Riboswitches vom Guanintyp, Iron-responsive elements, trans-spleißende Strukturen und microRNAs angewendet und eine Reihe von Vorhersagen getroffen.

So konnten neue Riboswitches in 19 bakteriellen Genen von insgesamt neun verschiedenen Bakterien vorhergesagt werden. Um diese Suchroutine Interessierten zur Verfügung zu stellen, wurde ein Webserver „Riboswitch Finder“ programmiert, der über das Internet erreichbar ist. Dieser untersucht Sequenzen des Nutzers auf Riboswitches vom Guanintyp und gibt im Erfolgsfall eine graphische Übersicht über den Treffer und die Sekundärstruktur aus.

Bei der Suche nach Iron-responsive elements konnten acht sequenz- und strukturkonservierte potentielle neue Iron-responsive elements mit Konservierung in den Genomdaten des Menschen und der Maus vorhergesagt und für eine experimentelle Testung zur Verfügung gestellt werden. Die Testung am Europäischen Laboratorium für Molekularbiologie in Heidelberg konnte hierbei ein neues Iron-responsive element experimentell verifizieren.

Die Suche nach trans-spleißenden Strukturen vom „spliced leader“ – Typ zeigt, dass eine große Anzahl solcher Strukturen im menschlichen Genom vorhanden sind und möglicherweise eine regulatorische Funktion haben.

Mit dem RNA Analyzer wurde ein Webserver erstellt, der als universelles Softwaretool zur Analyse beliebiger RNA- oder DNA-Sequenzen dient und ca. 15000 Anfragen pro Jahr verarbeitet. Die eingegebene Sequenz wird auf das Vorliegen von Exons, einer kodierenden Sequenz, Promotoren, AU-reiche Regionen, Polyadenylierungssignale und snRNP-Motive hin untersucht. Weiterhin wird überprüft, ob es sich um eine Transfer-RNA handelt und ob potentielle Iron-responsive elements oder trans-spleißende Strukturen vorhanden sind. Kodierte die Sequenz für ein Protein, so wird dieses vorhergesagt und per Link mit einem Server zur Analyse von Proteindomänen verlinkt. Werden keine Exons oder keine kodierende Sequenz gefunden, so wird die Sequenz auf Bereiche hin untersucht, in denen sich Auffälligkeiten im Bereich der Sekundärstruktur zeigen und der Benutzer erhält weitergehende Hinweise.

Zur Suche nach microRNAs wurde eine neue Suchstrategie erstellt, um auch Sekundärstrukturelemente ohne konservierte Primärstruktur zu finden. Eine Suche in der Sequenz des menschlichen Chromosoms 13 zeigt, dass hiermit bekannte microRNA-Cluster gefunden werden können. Weiterhin konnten mehrere potentielle microRNA-Cluster vorhersagt werden. Damit konnten wir zu jedem der eingangs genannten Themen neue medizinisch und biologisch relevante RNAs und neue Möglichkeiten für ihre Identifikation aufzeigen.