

Felix Kokocinski

Verarbeitung und funktionelle Analyse von molekularen Profilen aus Microarray-Experimenten

Prof. Dr. P. Lichter (DKFZ, D) und PD Dr. Karsten Rippe (Univ. Heidelberg, D)

Der Einsatz von Hochdurchsatz-Methoden wie *matrixCGH* und *Expression Profiling* mit DNA-Microarrays in der molekular-medizinischen Forschung erlaubt die parallele Analyse von biologischen Messwerten in einer bisher unerreichten Größenordnung. Um die anfallende Menge an Daten optimal nutzen zu können bedarf es spezialisierter Software-Lösungen.

Für das Management einer Klonsammlung, welche den Ausgangspunkt der Experimente darstellt, wurde das *CloneBase*-System entwickelt. Es handelt sich um eine Datenbank mit Internetzugangsseiten und einer automatisierten Aktualisierung der Daten. Es führt damit die Archivierung, Annotation und Informationsbereitstellung für sämtliche Klone durch. Von den Nukleinsäurefragmenten selektierter Klone werden mit Hilfe von Robotern Microarrays produziert. Um den komplexen Herstellungsprozess steuern und überwachen zu können, wurde ein Labordatensystem (*Laboratory Information Management System, LIMS*) entwickelt. *QuickLIMS* arbeitet mit den per Barcode identifizierten Mikrotiterplatten, die die Klone enthalten, indem es sie schrittweise durch ein definiertes Protokoll führt. Der Prozessablauf ist flexibel gespeichert, die Benutzeroberfläche wird darauf basierend dynamisch generiert. Mit dem System kann eine größtmögliche Automatisierung und eine Rückverfolgung eventueller Fehler bei der Microarrayherstellung erreicht werden. Für die Interpretation der experimentellen Ergebnisse von dieser Art von Hochdurchsatz-Methoden stellt die Integration unterschiedlicher Informationsquellen eine Schlüsselrolle dar. Das Analysesystem *FACT (Flexible Annotation and Correlation Tool)* wurde für die explorative Analyse der Ergebnisse mit Hilfe von Annotationsdaten und die Zusammenführung von relevanten Informationen aus heterogenen Datenquellen entwickelt. Daten aus unterschiedlichen experimentellen Systemen, z.B. genomische Veränderungen aus *matrixCGH*-Versuchen und Expressionswerte von mit cDNA-Microarrays untersuchten Genen, können mit Hilfe von spezialisierten Einzelmodulen in Zusammenhang gebracht werden. Dies wird durch die flexible Speicherung der Datendefinitionen und die Transformation auf eine gemeinsame Datenbasis erreicht. In gleicher Art ist es möglich, unterschiedliche Annotationsquellen und neue Analysemethoden in das System einzubinden. Mit der Anwendung *AutoPrime* wird schließlich auch die Verifikation der erzielten Microarray-Ergebnisse mit der Methode der RQ-PCR zu einem weiteren Teil automatisiert. Es verbindet eine Sequenzdatenbank mit einem Primerdesignprogramm und erreicht dadurch eine schnellere und robustere Auswahl geeigneter Primersequenzen.

Mit den vorgestellten Anwendungen wurde ein System geschaffen, welches als Datenbank- und Analyse-Netzwerk den experimentellen Ablauf der biomedizinischen Hochdurchsatzforschung ermöglicht. Es erlaubte die Untersuchung der Pathomechanismen unterschiedlicher Hirntumore (Menigiome, Medulloblastome) und hämatopoetischer Fragestellungen (Differenzierung der Zelllinie HL60, Akute Myeloische Leukämie). Insbesondere konnte die Entstehung und Progression des *Non-Melanom* Hautkrebses mit Hilfe von Microarrays mit murinen Fragmenten intensiv analysiert werden. Die Identifikation von neuen Kandidatengenen für humanen Hautkrebs, unter anderem Gene des Epidermalen Differenzierungskomplexes (1q21.3), wurde dadurch ermöglicht.