

## **5. NGFN Qualitätsmanagement-Workshop**

13. Juni 2006, 11:00 bis 17:30 Uhr

Projektträger DLR, Heinrich-Konen-Str. 1, 53227 Bonn, Raum E.059

**Moderation:** Herr Dr. Timm Jessen (Mitglied des NGFN-Lenkungsgremiums)

**Teilnehmer:** Frauke Behrens (RZPD Berlin), Amalia Diaz Lacava (Universität Bonn), Huberta von Eberstein (PopGen Kiel), Rolf Fimmers (Universität Bonn), Ronald Frank (GBF Braunschweig), Helmut Fuchs (GSF München), Thomas Häupl (Charité Berlin), Jörg Hoheisel, (DKFZ Heidelberg), Claus Hultschig (MPI MG Berlin), Ruprecht Kuner (DKFZ Heidelberg), Olaf Krüger (Projektmanagement NGFN, Bonn), Chris Lawerenz (DKFZ Heidelberg), Simon Little (Universitätsklinikum Gießen), Michael Majores (Uni Bonn), Michael Putz (GSF München), Christina Schröder (RZPD Berlin), Birte Sönnichsen (Cenix Bioscience GmbH), Uta Strasser (Projektmanagement NGFN, Bonn), Holger Sülthmann (DKFZ Heidelberg), Patrick Umbach (MDC Berlin-Buch), Dieter Weichenhan (Universitätsklinik Heidelberg), Tanja Weis (Universitätsklinik Heidelberg), Ruth Wellenreuther (DKFZ Heidelberg), Stefan Wiemann (DKFZ Heidelberg)

### **Zusammenfassung der Vorträge und Berichte**

Vortrag **“Datenstrukturen und -integration“**

Referent: Christian Lawerenz  
Deutsches Krebsforschungszentrum  
Theoretische Bioinformatik  
Im Neuenheimer Feld 580, 6912 Heidelberg  
[c.lawerenz@dkfz.de](mailto:c.lawerenz@dkfz.de)

Das Management der enorm großen und weiter wachsenden Datenmengen, die durch verschiedene Hochdurchsatz-Technologien erzeugt werden, stellt einen besonderen Schwerpunkt der QM-Aktivitäten im NGFN dar. Um eine effiziente und projektübergreifende Nutzung dieser Daten zu gewährleisten, gibt es im NGFN mehrere Datenbank-Initiativen. Die Bandbreite der verwalteten Datenbestände umfasst Daten aus Microarray-, RNAi-, Massenspektrometrie-, Hefe-Zwei-Hybrid- und zellbasierten Experimenten sowie Daten zu Open Reading Frame (ORF) und Gewebe-Ressourcen. Bereits etablierte und anerkannte internationale Standards wie z.B. mzDATA, MIAME bzw. MAGE\_ML wurden in diese Datenbanken integriert. Einige wichtige Elemente der Datenbank-Infrastruktur im NGFN sind nachfolgend aufgeführt:

- Die Datenbank „Central Research Infrastructure for molecular Pathology“ („CRIP“, <https://crip.rzpd.de>) am RZPD vernetzt Gewebekbanken und macht die dort für die Forschung zur Verfügung stehenden anonymisierten Daten zu humanen Gewebe-Proben einer Internet-basierten Suche zugänglich. Die anonymisierten Daten werden von den

vertraglich angebotenen Partner-Instituten in die RZPD-Datenbank eingespeist. Die registrierten Nutzer können über eine interaktive Suche auf die Daten zugreifen und abfragen, wie viele Fälle vorhanden sind, die einer bestimmten wissenschaftlichen Fragestellung entsprechen.

- LIFEdb am Deutschen Krebsforschungszentrum (DKFZ) ist eine Datenbank, die von der SMP Cell etabliert wurde und in der Daten aus der molekularen und zellulären Genfunktionsanalyse publiziert werden (<http://www.lifedb.de>). Über ein Web-Interface kann die Datenbank nach Daten zu Genen und Proteinen der SMP Cell abgefragt werden (u.a. zur Expression von Genen sowie zur Proteinlokalisierung, -funktion und –interaktion). Zusätzlich werden Links zu externen Datenquellen bereitgestellt, die weitere Informationen liefern.
- Die Datenbankplattform iCHIP am Deutschen Krebsforschungszentrum (DKFZ) stellt die klinische und experimentelle Annotation mit den prozessierten Werten und Rohdaten strukturiert und entsprechend internationaler Standards zur Verfügung ([www.ichip.de](http://www.ichip.de)). Die kooperierenden krankheitsorientierten deutschen und europäischen Konsortien nutzen iCHIP als die zentrale Datenbank für ihre aktuellen Technologien (z.B.: Microarray, aCGH, Massenspektrometrie, Gelelektrophorese, RNAi). Die Integration der verschiedenen Verfahren aus unterschiedlichen Studien innerhalb von iCHIP ermöglicht studienübergreifende Analysen und komplexe Aussagen über die krankheitsrelevanten Funktionen von Gen- und Proteingruppen.
- Die im NGFN generierten Ergebnisse ermöglichen die internationale Sichtbarkeit des NGFN durch das webbasierte NGFN-Portal am GSF-Forschungszentrum. Die veröffentlichten, im NGFN erzeugten Forschungsergebnisse und die angebotenen Services werden den internationalen Wissenschaftlern übersichtlich präsentiert. Dieses Portal bietet den im NGFN eingebundenen Gruppen die Möglichkeit, in einem einfachen Verfahren ihre Daten semantisch mit den jeweiligen Ergebnissen anderer Gruppen zu verknüpfen.
- In der Datenbank der krankheitsorientierten Netze und der Genetisch-Epidemiologischen Methodenzentren (KN-GEM-DB) werden die im NGFN verfügbaren Patientenkollektive, Proben und klinische Daten über einen datenschutzrechtlich unbedenklichen Variablensatz transparent gemacht und in aggregierter Form dargestellt. Die Datenbank ist somit ein Übersichtswerkzeug für die Forscher im NGFN, um zu recherchieren, ob bestimmte klinische Daten erhoben wurden und an welchen Standorten und wie häufig diese Proben/Daten im NGFN vorhanden sind.

Die verschiedenen Datenmanagement-Ansätze im NGFN sind bereits gut vernetzt. In Zusammenarbeit mit den Klinikern der Krankheitsnetze wurden umfangreiche Spezifikationen für die klinische Beschreibung entwickelt. Diese wurden in die jeweiligen Strukturen der wesentlichen Datenbanklösungen integriert und gewährleisten die Präsentation und Weiterverwendung der Studiendaten sowie den harmonisierten und validierbaren XML-basierten Austausch innerhalb der klinischen Netze und NGFN-weit.

## **Vorschläge zur weiteren Optimierung der Datenmanagement und Datenbank-Infrastruktur**

Seit 2005 gibt es einen jährlich stattfindenden Datenbank-Workshop, um das Datenmanagement im NGFN zu koordinieren und um redundante Entwicklungen zu vermeiden. Im Rahmen des letztjährigen Datenbankworkshops, der am 31. Mai 2006 stattfand, wurden Vorschläge zur weiteren Optimierung der Datenmanagement- und Datenbank-Infrastruktur erarbeitet:

- NGFN-weit sollten einheitliche web-basierte Datenintegrations-Technologien eingeführt werden, um eine allgemeine Verfügbarkeit der Daten und einen einfachen Datenaustausch zu ermöglichen. Die aktuellen Webstandards (z.B. XML, Kernvokabularien, Thesauri, Taxonomien, Ontologien, etc.), Service-orientierte Technologien und eine mehrschichtige Systemarchitektur sollten verwendet werden.
- Datenstrukturen und Werkzeuge zur Beschreibung von Dynamiken und Modellen (e.g. MIACA, SBML, CellML) sollten integriert werden. Diese Datenstrukturen ermöglichen eine direkte Zusammenarbeit der Entwickler von theoretischen Modellen und der praktisch experimentierenden NGFN-Wissenschaftler. Elemente der Systembiologie sollten etabliert werden, um diese zur Bestätigung von Modellen, Komponenten und Reaktionen innerhalb von biologischen Systemen sowie zur Validierung von mathematischen Modellen zu nutzen.
- Integrations-Plattformen für die translationale Forschung sollten etabliert werden. Klinische und molekularbiologische Experimente könnten besser verknüpft werden, um eine effektivere und produktivere Forschung an den klinischen Einrichtungen zu ermöglichen. Die wesentlichen Parameter für die Bestimmung von Krankheitsmerkmalen sind zu bestimmen und zu verwenden. Klinische Komplexität ist nur insofern abzubilden, falls sie statistisch auswertbar und medizinisch relevant ist. Der Harmonisierungsprozess würde auch die Verwendung von klinischen Daten-Pools und Hochdurchsatz-Experimenten mit Proben von Biomaterialbanken außerhalb des NGFN ermöglichen.
- Die im NGFN verfügbaren Standardprotokolle sollten mit Metainformationen (z.B. Kurzbezeichnung und Autor(en)) verknüpft und in Datenbankmanagementsysteme integriert werden. Um die Standardprotokolle dauerhaft nutzen zu können und um auf diese als Referenz verweisen zu können, sollte eine persistente URL-Adresse eingerichtet werden, unter der die Protokolle abgelegt werden.
- In das Management von personenbezogenen Daten ist die Miteinbeziehung eines Datenschutzkonzeptes unerlässlich. Die Institutionen und Arbeitsgruppen, die über Datenbanken verfügen, sollten in enger Verbindung mit den Datenschutzorganen (Ethikkommission, Landesdatenschutzbeauftragte, AGs „Datenschutz“ und „Biomaterialdatenbank“ der TMF) stehen. Um die Eindeutigkeit bei der Vergabe von Patientenschlüsseln in klinischen Netzen zu ermöglichen und die geforderte und notwendige Pseudonymisierung von Daten zu gewährleisten, sollte der PersonenIdentifikator (PID)-Generator der TMF eingesetzt werden.

Vortrag

## **„Übersichtsdatenbanken und Webanwendungen der Krankheitsnetze im NGFN“**

Referent: Dr. Michael Putz

GSF – Forschungszentrum für Umwelt und Gesundheit

Institut für Epidemiologie

Ingolstaedter Landstr. 1, 85764 Neuherberg

michael.putz@gsf.de

In der Krankheitsnetz-GEM-Datenbank werden die im NGFN verfügbaren Patientenkollektive, Proben und klinische Daten über einen datenschutzrechtlich unbedenklichen Variablensatz transparent gemacht und in aggregierter Form dargestellt. Die Datenbank ist somit ein Übersichtswerkzeug für die Forscher im Nationalen Genomforschungsnetz, um zu recherchieren, ob bestimmte klinische Daten erhoben wurden und an welchen Standorten und wie häufig diese Proben/Daten im NGFN vorhanden sind.

Außerdem ermöglicht die Datenbank eine verbesserte Berichterstattung innerhalb des NGFN.

Die Datenbank verwendet das Betriebssystem Linux, als Datenbankverwaltungssystem wird MySQL zusammen mit dem Webserver Apache eingesetzt und die Benutzeroberfläche wird mit der Programmiersprache PHP realisiert.

Die Daten werden von den Projektverantwortlichen zum leichteren Import in die Datenbank in einem Excel-Formular gesammelt. Mit dem Mauscursor können Beschreibungen zu den entsprechenden Feldern abgerufen werden.

Die Datenauswertung ist nach verschiedenen Kriterien möglich. Es können Überblickstabellen und Statistiken über Krankheiten, Zentren, Populationen und Bioproben erzeugt werden. Über ein Reporting-Tool können Berichte erstellt werden. Weitere Funktionalitäten sind die dynamische Gestaltung und Erstellung von Menüs, Diagrammen und Statistiken. Außerdem ist ein halbautomatisiertes Qualitätsmanagementsystem (im Excel-Importfile) integriert, mit dem z.B. Daten-Ausreißer erkannt werden können.

Mittlerweile ist die Implementierung der Datenbank und des web-basierten Informationssystems abgeschlossen. Zusammen mit den Koordinatoren der Krankheitsnetze wird derzeit der Transfer der Datenbestände an die GEMs und der Import der Daten in die Datenbank vorbereitet.

Jedes Krankheitsnetz wird durch ein GEM (KN-GEM) betreut. Die Dateneingabe kann über das zuständige GEM erfolgen, so dass die Arbeitsbelastung in den Netzen gering gehalten wird. Das KN-GEM nimmt die Daten-Meldungen der Projektleiter entgegen und kontrolliert auf Plausibilität und Vollständigkeit, sowie auf Mehrfachmeldungen. Gemeldete Daten werden aggregiert und in Form eines datenschutzrechtlich unbedenklichen, anonymisierten Variablensatzes in die Datenbank eingegeben. In diesen Variablensatz wurde der Minimaldatensatz für klinische Daten im NGFN (entwickelt durch die QM-AG „Klinik, Datenmanagement und Datenanalyse“) integriert. Einige Beispiele für Variablen: NGFN Projekt-ID, Kontakt, Fall-Kontroll-Status, Geschlecht, Geburtsjahr, Ethnizität, Raucherinformationen (ja/nein), Familieninformationen (ja/nein), Diagnose, Art der Bioprobe, vorhandene Genotypinformation.

Die Webanwendung ist über das SSL-Verschlüsselungsprotokoll für Datenübertragungen abgesichert, so dass höchste Datentransfersicherheit gewährleistet wird.

## Vortrag „Datenmanagement in der Deutschen Mauslinik“

Referent: Dr. Helmut Fuchs

GSF-Forschungszentrum für Umwelt und Gesundheit, München

Institut für Experimentelle Genetik

Ingolstädter Landstr. 1, 85764 Neuherberg

[hfuchs@gsf.de](mailto:hfuchs@gsf.de)

Die Deutsche Mauslinik ist eine Plattform für die systematische, standardisierte und umfassende Phänotypisierung von mutierten Mäusen. In einem Primärscreen werden mehr als 240 Parameter analysiert. Sekundär- und Tertiär-Untersuchungen ermöglichen eine tiefergehende Analyse von weiteren Parametern. Bis zu 1500 Mäusen werden pro Jahr im Primärscreen untersucht, zusätzliche 1500 Mäuse pro Jahr im Sekundär- und Tertiär-Screen. Ein Projekt von solcher Dimension ist nur mithilfe eines zentralen IT Systems durchführbar.

Beim Projektstart (2001) wurde zunächst ein kommerzielles klinisches Informationssystem (Galilei Clinical Research System) verwendet, das an die besonderen Erfordernisse der Mauslinik angepasst wurde. Um das Datenmanagement der Mauslinik weiter zu optimieren, wurde in-house ein maßgeschneidertes Informationssystem entwickelt, das weitere Funktionalitäten besitzt, die in der kommerziellen Software fehlen oder nur zum Teil enthalten sind. Dazu gehört z. B. die Projektsteuerung (Koordinationssoftware, Zuchtbuch und Versuchstiermeldungen), erweiterte Dokumentation der Phänotypisierungen (Verknüpfung von Messwerten und SOPs) und eine verbesserte Integration von Genotypen und Mauslinien mit Links zur öffentlich verfügbaren Literatur und zu anderen Datenbanken. Das neue Mausverwaltungssystem MausDB wurde ausschließlich auf der Grundlage von open source Produkten entwickelt (Linux als Betriebssystem, Perl als Programmiersprache, Apache Webserver, MySQL Datenbank). Diese Kombination ist weltweit vielfach und erfolgreich im Einsatz.

Die Mausdatenbank besitzt folgende Eigenschaften und Funktionen:

- eine grafische Oberfläche mit Formularen für die assistierte Dateneingabe (Verpaarungen, Würfe und Aufzuchten, Genealogie) sowie für die Präsentation von Ergebnissen der Phänotypanalyse für eine bestimmte Maus
- Zentrale Datenhaltung mit eindeutigen Identifikationsnummern für Mäuse, Regale, Käfige, Linien etc. Farbkodierte Käfigkarten werden verwendet, um die Mäuse im jeweiligen Käfig zu beschreiben.
- Einlesen der Phänotypisierungsdaten aus Exceldateien
- Erstellen und Ausdrucken von Arbeitslisten für die offenen Tätigkeiten (Absetzen von Würfen, Listen für die Phänotypisierung von Mäusen.
- Historie aller Aufenthaltsorte einer Maus unter Angabe der Käfignachbarn
- Ausgabe der aktuellen Tierbestandszahlen als Exceltabelle für den Datenaustausch
- Einheitliches, standardisiertes Vokabular für die Stammdaten
- Verbindung zur Datenbank des Kollaborationspartners über die ForeignID einer Maus
- standardisierte Excel-Schnittstelle für den Import von Mausdaten (z.B. Geburtsdatum, Geschlecht, genetischer Hintergrund, Ohrmarke, Fellfarbe usw.) aus Datenbanken anderer Einrichtungen in die Datenbank der Mauslinik.
- Link auf eingescanntes Gesundheitszeugnis für importierte Mäuse,

- nutzerfreundliche Suche nach bekannten Aspekten von Genen und Phänotypen in der Literatur.

Rohdaten, die von den Analysegeräten in der Mauslinik erzeugt werden, durchlaufen zuerst eine Analyse und werden durch verschiedene Scripts automatisch auf Eingabefehler, Plausibilität und referentielle Integrität überprüft, bevor sie in die Datenbank übertragen werden. Die Endvalidierung der Daten wird von den Wissenschaftlern auf der Ebene von Excel-Dateien durchgeführt. Derzeit sind etwa 590 000 Ergebnis-Datensätze von 5432 Mäusen in der Datenbank abgelegt. Das sind im Durchschnitt 108 Parameter pro Maus. Die Ergebnisse der Primärscreen-Untersuchungen sind für alle Wissenschaftler der Untersuchungsplattformen der Mauslinik einsehbar. Deshalb können die Ergebnisse von unterschiedlichen wissenschaftlichen Perspektiven beurteilt werden.

Vielfältige Suchfunktionen erlauben die Suche nach Mäusen, Verpaarungen, Würfen, Importen, Phänotypisierungsaufträgen, Käfigen etc. Bei der Suche nach Mäusen können einzelne Suchkriterien wie das Geburts- oder Todesdatum, die Linie, das Geschlecht, das Kommentarfeld, die ID der pathologischen Untersuchung, das Alter und der Genotyp einzeln oder in Kombination miteinander über die Oberfläche abgefragt werden. Das Ergebnis der Abfrage, das Resultset, kann dann in einem Cart gespeichert und weiterverwendet oder nach Excel exportiert werden. In Übersichten werden alle Importe und alle Verpaarungen mit den relevanten Begleitdaten chronologisch dargestellt.

Um das Koordinationsteam der Mauslinik bei seinen vielfältigen Aufgaben zu unterstützen und um alle Experimente einige Wochen im voraus planen zu können werden, wurde eine web-basierte Management-Software etabliert. Die potentiellen Kooperationspartner füllen ein detailliertes Formular auf der GMC-Homepage ([www. mouseclinic.de](http://www.mouseclinic.de)) aus. Dort werden die Informationen abgefragt, die benötigt werden, um über die Anfrage zu entscheiden. Die Daten werden in einer relationalen Sybase ASE 12.5 Datenbank gespeichert und können durch die GMC Manager über passwort-geschützte web-basierte Formulare abgerufen und editiert werden. Nachdem eine Anfrage akzeptiert wurde, wird automatisch vom System bestimmten Mitgliedern des GMC Managements standardisiert eine Reihe von Aufgaben mit einer Deadline zugeteilt. Ein paar Tage bevor die Deadline einer Aufgabe erreicht ist, sendet ein automatisches Erinnerungssystem eine E-Mail an den verantwortlichen Manager. Wird bis zum Stichtag die Aufgabe in der Datenbank nicht als abgearbeitet eingetragen, werden alle Mitglieder des Managementteams per e-Mail informiert. So wird verhindert, daß wichtige Aufgaben verspätet bearbeitet werden oder in Vergessenheit geraten.

Die Koordinationssoftware ermöglicht einen schnellen Überblick über alle laufenden bzw. abgeschlossenen Projekte und deren Status sowie über den Status aller einzelnen Aufgaben, deren Deadlines und Zuständigkeiten. Auch die Phänotypisierungs-Starttermine werden zentral über das System verwaltet. Die Vorteile des Koordinationssystems liegen in der zentralen und ausgeklügelten Verwaltung der standardisierten Arbeitsabläufe und dem Remindersystem, das die Mitglieder bei den Routinearbeiten entlastet.

## Vortrag „Weiterentwicklung des Sanitäts-Monitoring Systems der Deutschen Mauslinik

Referent: Dr. Helmut Fuchs

GSF-Forschungszentrum für Umwelt und Gesundheit, München

Umfangreiche Hygiene-Maßnahmen, Barriere- und Importregeln sowie ein effizientes Sanitäts-Monitoring-System gewährleisten in der Deutschen Mauslinik einen möglichst optimalen Hygiene-Status und verhindern das Einschleppen von Krankheitserregern.

Die Mäuse werden in einzelbelüfteten Käfigsystemen (sogenannte IVC, VentiRack Bioscreen TM, Biozone, Margate, UK) gehalten, deren Entlüftungssystem mit dem Ventilationssystem (HVAC) verbunden ist. Durch strikte Einhaltung von Barrierebedingungen (u.a. Kleidungswechsel bei Eintritt in die Mauslinik; Tragen von Handschuhen, Kopfbedeckung, Masken; Autoklavieren bzw. Desinfizieren von zu importierendem Material) und durch Verwendung eines elektronischen Zugang-Kontrollsystems (Cerpas TM, Siemens) werden die Tiere vor Keimen geschützt.

Die Käfige sind mit dem neu entwickelten Bioscreen TM System ausgestattet, das in der Lage ist über sogenannte Wächter-Mäuse den Status der Mäuse eines Mausregales zu überwachen. Dazu wird die Abluft aus den Käfigen des zu überwachenden Regales durch den Käfig der Wächtermäuse geleitet. Zusätzlich bekommen die Wächtermäuse kleine Mengen des benutzten Streus der anderen Käfige.

Wenn neue Pathogene entdeckt werden, können mithilfe der in der GMC-Datenbank gespeicherten Informationen alle Mäuse identifiziert werden, die von den Krankheitserregern ebenfalls befallen sein könnten, denn in der GMC Datenbank ist jeder Käfig-Transfer der Mäuse eingegeben. So kann der Ursprung einer Infektion schnell ausfindig gemacht werden und eine weitere Ausbreitung der Krankheitserreger auf andere Mäuse wird verhindert.

Ein Sentinel-Monitoring-System wurde implementiert, um den aktuellen Hygiene-Status der Mäuse zu überwachen. Das System verwendet Indikator-Tiere (Sentinels), um den Verbrauch wertvoller Versuchsmäuse für routinemäßige Hygieneuntersuchungen zu vermeiden: Je zwei Sentinel-Mäuse werden in den Bioscreen-Käfig gesetzt. Der Gesundheitsstatus dieser Sentinel-Mäuse wird in regelmäßigen Abständen von drei Monaten überprüft.

Zum Schutz des Versuchstierbestands der GSF vor potenziellen Infektionserregern der importierten Tiere wird für jeden dieser Importe ein aktuelles Gesundheitszeugnis für die entsprechenden Tiere angefordert, das durch die Veterinärabteilung der GSF geprüft wird. Nach Ende der Untersuchungen der Mäuse im Primärscreen der German Mouse Clinic wird eine Serumprobe der Mäuse eingefroren, die für spätere Nachuntersuchungen der Mäuse auf Viren verwendet werden kann.

Dieses Konzept könnte im Hinblick auf folgende Aspekte weiter verbessert werden:

- Die Bewertung des Gesundheitsstatus der zu importierenden Mäuse hängt im vollen Umfang vom Gesundheitszeugnis ab.
- Infektionen kurz vor oder während des Transportes werden nicht detektiert.

- Eine Messung der archivierten Proben erfolgt nur bei Verdacht.
- Es sind keine Daten für die individuelle Maus verfügbar.

Um das Hygiene-Monitoring im Hinblick auf die genannten Aspekte zu optimieren, könnte folgendes Verfahren etabliert werden:

- Nach Ankunft werden die importierten Tiere für 4 Wochen in einem „Hygiene-Modul“ gehalten
- Danach werden bei allen Tieren folgende Tests durchgeführt:
  - Blutprobe
  - Abklatsch und Kot
  - Ektoparasiten Adspektion
  - Abstrich von Rachen und Genitalen
- Nach Durchführung des Primärscreens und Durchlaufen der Untersuchungsstationen wird noch mal ein Endtest durchgeführt:
  - Mikrobiologische Sektion
  - Bluttests
  - Bakteriologische Untersuchung
  - Untersuchung auf Ekto- und Endoparasiten

Derzeit wird überprüft, inwiefern dieses Verfahren ohne gravierende Einschnitte in den bestehenden Arbeitsfluss durchgeführt werden kann, und ob die Kosten in einem akzeptablen Rahmen gehalten werden können. Die Möglichkeit einer Zusammenarbeit mit anderen Institutionen und der Kosteneinsparung bei Durchführung von Tests außer Haus wird ebenfalls derzeit geklärt.

#### Vortrag „Qualitätsmanagement der SMP RNAi“

Referentin: Dr. Birte Sönnichsen  
 Cenix BioScience GmbH  
 Tatzberg 47-51, 01307 Dresden  
 soennichsen@cenix-bioscience.com

Qualitätsmanagement ist ein wichtiger Bestandteil der SMP RNAi und gewährleistet die Optimierung von Arbeitsabläufen und die Weiterentwicklung der RNAi-Technologie zu einem Routineverfahren, das die Hochdurchsatz-Untersuchung der Gen-Funktionen bei menschlichen Zellen *in vitro* und bei Mäusen *in vivo* ermöglicht.

Um die Verwendung der RNAi-Techniken auf ein einheitlich hohes Qualitätsniveau zu bringen und Transparenz über die experimentelle Vorgehensweise herzustellen, wurden Standardprotokolle zu *in vitro* und *in vivo* RNAi-Techniken erarbeitet, die auf den Webseiten der SMP RNAi (<http://www.rnai.ngfn.de>) abrufbar sind.



#### a) QM im Teilbereich **Screening Reagent esiRNA**

Zwei Teilprojekte der SMP RNAi etablieren und optimieren die Herstellung von esi-RNA-Bibliotheken. EsiRNAs werden durch einen Verdau von 300-500bp langen dsRNAs mit RNase III hergestellt und sind eine effiziente und kostengünstige Alternative zur chemisch synthetisierten siRNA. Ein wichtiger Unterschied zwischen siRNA und esiRNA ist der, dass bei der siRNA alle dsRNA-Moleküle die gleiche Sequenz haben, während die esiRNA eine Mischung von verschiedenen dsRNA-Molekülen darstellt.

Bei der Herstellung von esi-Bibliotheken in der SMP RNAi werden alle wichtigen Schritte qualitätskontrolliert. Die Qualitätskontrollen umfassen Gel-Checks nach der Gen-spezifischen PCR-Amplifikation, nach dem RNaseIII-Verdau und nach der Konzentrationseinstellung der Bibliothek. Um Sequenzen mit guter Silencing-Fähigkeit zu finden, wurde das Programm DEQOR entwickelt. Dieses Programm errechnet, mit welcher Gen-Region ein optimales Silencing erreicht werden kann und es erkennt und minimiert Cross-Silencing Effekte. Um die experimentellen Schritte zu dokumentieren und um den Herstellungsprozess jeder RNAi-Ressource nachverfolgen zu können, wurde ein Labor Informationsmanagement System (LIMS) entwickelt. Außerdem wurden zahlreiche Schritte zur Herstellung einer esiRNA Bibliothek rationalisiert und automatisiert, so dass jetzt 2400 esiRNAs pro Tag hergestellt werden können. Die esiRNA-Bibliothek deckt ca. 17 000 humane Gene ab.

Das esi-RNA Reagenz wird derzeit in mehreren Laboren mit verschiedenen Zelllinien getestet. Über eine quantitative PCR konnte gezeigt werden, dass chemisch synthetisierte siRNAs und esiRNAs eine vergleichbare Knockdown-Effizienz in unterschiedlichen Zellen aufweisen. In einem weiteren QM-Schritt wird untersucht, inwieweit unspezifische Effekte - so genannte „Off-Target Effects - auftreten. Erste Ergebnisse deuten darauf hin, dass die Zahl der Off-Target-Effects von esiRNAs im Vergleich zu siRNAs deutlich kleiner sind: In Expressionsarrays behandelter Zellen zeigen siRNAs eine bis zu 20fach höhere Anzahl differenziell exprimierter Gene im Vergleich zu esiRNAs. Das könnte daran liegen, dass esiRNAs aus einem Pool von vielen individuellen siRNA-Molekülen bestehen. Jede dieser individuellen siRNA-Molekülsorten hat einen unterschiedlichen off-target-Effekt, der allerdings relativ klein ist, da jede siRNA-Molekülsorte in relativ geringer Konzentration vorhanden ist.

In weiteren Experimenten soll überprüft werden, in welchen Zelllinien und in welchem Ausmaß die Transfektion mit esiRNAs toxisch ist, ob die Zytotoxizität Chargen-spezifisch ist, und ob die Zytotoxizität durch eine Interferon-Reaktion oder durch einen anderen Mechanismus ausgelöst wird.

#### b) QM im Teilbereich **Cell based RNAi screening**

Die Zellbasierte RNAi, sei es in Mikrotiterplatten oder mit Hilfe der Zellarray-Technologie auf Glaträgern wird kontinuierlich optimiert und qualitätsgeprüft.

Um einen RNAi Screening Assay zu optimieren, müssen die Bedingungen für das Gen-Silencing durch RNAi mit den Assay Readout Bedingungen kombiniert werden. Folgende Schritte der Qualitätskontrolle werden durchgeführt:

- Verschiedene Transfektionsreagenzien, Ausplattungsdichten für Zellen sowie Wachstumsbedingungen werden ausgetestet.
- Positive Kontrollgene für die spezifischen Assay Readouts werden ausgewählt und zur Überprüfung der Datenqualität sowie zur Optimierung der Silencing-Effekte und des Signal-zu-Hintergrund-Verhältnisses eingesetzt.
- Verschiedene siRNAs, die als Negativkontrolle dienen sollen, werden getestet. Negativkontrollen sind wichtig für die Bewertung von unspezifischen Auswirkungen, die durch die Transfektion mit dsRNA entstehen, und sie ermöglichen die Normalisierung der Daten-Untergruppen der Experimente.
- Mit Loss-of-Function Kinetiken, bei denen die Knockdown-Effizienz mit Hilfe von quantitativer RT-PCR ermittelt wird, werden geeignete Zeitfenster gewählt.
- die Faktoren, die zur Variabilität des Experimentes führen könnten, werden charakterisiert (z.B. die Reagenzien-Chargen, die Zell-Chargen usw.).
- die Variabilität zwischen und innerhalb der 96-well-Platten bzw. der Glaträger, sowie die Reproduzierbarkeit der Experimente wird untersucht.

Aus den Ergebnissen und Beobachtungen dieser Kontrollen werden Qualitätskriterien definiert (z.B. Schwellenwerte für die Werte der Positiv- und Negativkontrollen)

Jedes Durchmusterungs(Screening)-experiment umfasst drei Runden:

- (I) In der ersten Runde wird das zuvor optimierte Verfahren mit Fokus auf maximales Silencing durchgeführt. Hierfür werden die Gen spezifischen Silencing Moleküle in hohen Konzentration eingesetzt. Im Falle von siRNA Molekülen werden 3 verschiedene siRNAs pro Gen getestet. Bei diesem Schritt sind noch viele Falsch-Positive zu erwarten.
- (II) In der zweiten Runde werden alle Kandidatengene, die bei der ersten Durchmusterungsrunde mit mehr als einem siRNA Molekül gefunden wurden, bestätigt. Dadurch werden die Falsch-Positiven eliminiert, da die Wahrscheinlichkeit sehr gering ist, dass mehrere siRNAs mit unterschiedlicher Sequenz die gleichen Off-target-Effects haben.
- (III) In der dritten Runde werden die Ergebnisse der zweiten Durchmusterungsrunde nochmals bestätigt und der Silencing-Effekt wird quantitativ gemessen.

In der zweiten und dritten Runde kommen häufig auch Sekundärassays zum Einsatz, die helfen, die gefundenen zellulären Phänotypen besser zu definieren.

### c) QM im Teilbereich **In vivo RNAi in Mice**

Der SMP-RNAi Teilbereich „*in vivo* RNAi in Mäusen“ entwickelt die RNAi-Technologie als ein standardisiertes und zeitsparendes Verfahren für das Gen Silencing in Mäusen. Dieser Ansatz ermöglicht die Herstellung von Knock-Down Mutanten in größerem Maßstab und mit weniger Zeitaufwand als die herkömmliche Gen-Knock-Out-Methode.

Zu den QM-Maßnahmen in diesem Teilbereich gehören u.a. das Testen der Konstrukte auf Silencing Aktivität in ES-Zellen und die Durchführung von quantitativen RT-PCRs und Northern Blots aus Gewebeproben der Mäuse, um die Effizienz des shRNA-vermittelten Silencing in Mäusen zu validieren.

#### d) **Computation, Database, Bioinformatics**

Die Daten, die mithilfe der RNAi-Experimente gewonnen werden, sollen in eine Datenbank eingehen, die auf iCHIP basiert. Eine Suchfunktion wird zu jeder eingetragenen siRNAi-ID die Abfrage von numerischen Auswertungen, Bildern, Gen-Annotationen, SOPs usw. ermöglichen. Es ist äußerst wichtig, gemeinsame Standards für den Austausch von RNAi-Daten zu definieren, um zu gewährleisten, dass die Daten verschiedener Arbeitsgruppen später miteinander vergleichbar sind. Deshalb wurde in Anlehnung an MIAME innerhalb des SMP-RNAi in 2005 die Initiative **MARIE** (=Minimal Annotation for RNAi Experiments) ins Leben gerufen. Diese Initiative arbeitet an der Auswahl von Kriterien für den Datenaustausch von RNAi Experimenten. Sie arbeitet eng mit der Initiative MIACA der SMP cell zusammen, um sicherzustellen, dass derartige Standards nicht allein stehen, sondern weitgehend integrierbar sind. Mittlerweile gibt es auf Anregung der Firma Dharmacon auch eine weltweite Initiative (MIARE) zur Erarbeitung von Standards für RNAi Experimente. Anstrengungen zur Harmonisierung aller drei Konzepte sind laufen derzeit.

#### Vortrag „**Qualitätsmanagement der SMP Epigenetik**“

Referent: Dr. Jörg Hoheisel

Deutsches Krebsforschungszentrum

Functionelle Genomanalyse

Im Neuenheimer Feld 580, 69120 Heidelberg

[j.hoheisel@dkfz.de](mailto:j.hoheisel@dkfz.de)

Die SMP Epigenetik besteht aus 8 Partnern (DKFZ Heidelberg; Internationale Universität Bremen; Universität des Saarlandes; Universität Bonn, Max-Planck-Institut für molekulare Genetik, Berlin; febit, Heidelberg; Leibniz Institut für Altersforschung; Epigenomics, Berlin). Außerdem ist die Firma Siemens assoziiert. Alle Partner tauschen regelmäßig Proben und Ergebnisse aus, um einen hohen technischen Standard aller Experimentierschritte und eine Vergleichbarkeit der Ergebnisse zu gewährleisten. Extensive Kooperationen bestehen auch mit anderen NGFN Bereichen, mit Tumorbanks sowie mit internationalen Projekten und wissenschaftlichen Gruppen (davon allein 11 Partner auf europäischer Ebene).

Methylierungsänderungen werden meist mithilfe von Techniken untersucht, die auf der Natriumdisulfit Reaktion beruhen. Hierbei wird unmethyliertes Cytosin in Uracil umgeformt, während methyliertes Cytosin nicht verändert wird. Diese chemische Reaktion produziert somit einen methylierungsabhängigen Einzelbasen-Polymorphismus, der über verschiedenen Methoden nachgewiesen werden kann.

Die Einführung der Microarrays-Technik hat die Untersuchung von Methylierungsmustern rationalisiert und automatisiert, da mit dieser Technologie der Methylierungsstatus von tausenden individueller Methylierungsstellen (CpGs) gleichzeitig untersucht werden kann. Bei der hochauflösenden Analyse werden die Microarrays mit der *Geniom One* Technologie der febit biotech synthetisiert. Genomische DNA wird mit Natriumdisulfit behandelt und dann über eine PCR vervielfältigt. Die PCR-Produkte werden fluoreszent markiert. Die Produkte binden auf den Microarrays entweder an die Oligonukleotide für methylierte DNA oder an Oligonukleotide, die die nicht-methylierte DNA repräsentieren, woraus man den Methylierungsstatus der ursprünglichen genomischen DNA ableiten kann.

Jeder Microarray enthält Positiv- und Negativ-Kontrollen. Mithilfe der Positivkontrolle wird die korrekte Synthese und Hybridisierung überprüft. Mithilfe von Negativkontrollen, die eine vollständige unabhängige Sequenz repräsentieren, kann kontrolliert werden, inwieweit es ein Hintergrundsignal durch unspezifische Bindung gibt. Jedes Experiment wird gleichzeitig auch mit *in vitro* teilweise oder vollständig methylierter sowie mit unmythilyerter genomischer DNA durchgeführt. Auf diese Weise können die Daten der eigentlichen Probe immer auch mit den Daten verglichen werden, die mit nicht, partiell oder voll methylierter genomischer DNA gewonnen wurden. Die Qualität der Ergebnisse wird mit anderen etablierten Techniken der epigenetischen Analyse (z.B. Disulfit-Restriktionsanalysen, Disulfit-Sequenzierung) validiert.

Alternativ zur Hybridisierung mit methylierungsspezifischen Oligonukleotiden kann auch eine Primer-Verlängerungsreaktion (Primer Extension Reaction) durchgeführt werden, um den Methylierungsstatus von CpG zu analysieren. Bei dieser Methode hybridisieren die Primer direkt neben der zu untersuchenden Base. Nach Inkubation mit markierten Dideoxynukleotiden wird das passende Nukleotid durch die Polymerase eingesetzt und kann über die spezifische Markierung detektiert werden. Dieses Verfahren ist sehr robust und die Zahl der Analysen auf einem Array kann deutlich vergrößert werden. Außerdem kann die Hybridisierung der Primer bei einer höheren Temperatur stattfinden, so dass intramolekulare Faltungen und unspezifische Bindungen vermieden werden. Die Primer-Verlängerungsreaktion wird auch beim universell einsetzbaren Microarrays verwendet, einem Verfahren, bei dem verschiedene molekularbiologische Analysen (Epigenetische Studien, SNP Genotypisierung, RNA-Expressionsprofile, Splice-Analysen, Protein-DNA Wechselwirkungen, usw.) in frei wählbarer Kombination auf einem Microarray durchgeführt werden können.

Bei allen Experimenten werden die Leitlinien der Guten Laborpraxis (Good Laboratory Practice) und der Good Clinical Practice (bei klinischen Projekten zur Untersuchung des Methylierungsmusters von Tumormaterialien) angewendet. Qualitätssicherungsverfahren werden zwischen den Gruppen der SMP Epigenetik und mit internationalen Partnern verglichen und ausgetauscht. Ein Fokus liegt dabei auf der Qualitätssicherung der verwendeten Probenmaterialien, die sehr gut charakterisiert sein müssen.

Ziel des Gesamtprojektes der SMP Epigenetik ist die Etablierung eines Basisdatensatzes, der für diagnostische Zwecke, für die Prognose der Krankheitsentwicklung und als Hilfsmittel für die Therapieplanung eingesetzt werden kann.

## Vortrag „NGFN-Ringversuch Hochdurchsatz-Genotypisierung“

Referent: Dr. Rolf Fimmers

Universität Bonn

Institut für Medizinische Biometrie, Informatik und Epidemiologie

Sigmund-Freud-Straße 25, 53105 Bonn

rolf.fimmers@ukb.uni-bonn.de

Um die Qualität der Genotypisierungen an den Genotypisierungszentren im NGFN zu evaluieren, werden Ringversuche durchgeführt. Ein erster Durchlauf testete die Qualität der Genotypisierungen von Gen-Chips, die einen festgelegten Satz von genomweit verteilten Markern tragen. An dem Ringversuchdurchlauf beteiligten sich fünf Genotypisierungszentren. Vier Zentren (CCG Köln, GSF München, MDC Berlin, RZPD Berlin) genotypisieren mit einer Affymetrix-Plattform und verwendeten während des ersten Ringversuch-Durchlaufes den Affymetrix GeneChip 100k. Das Genotypisierungszentrum GSF/MPI München besitzt eine Illumina-Plattform und verwendete während des ersten Ringversuch-Durchlaufes den Linkage IV Panel Genchip.

Wiederholungen des Ringversuchs sollen halbjährlich laufen. Da die Zentren derzeit auf größere Genchips umsteigen (Affymetrix 500K Set, Illumina HumanHap 300), wird im zweiten Ringversuch-Durchlauf die Genotypisierungsqualität dieser größeren Genchips getestet. In dieser zweiten Runde ist auch ein Vergleich zwischen verschiedenen GenChip-Chargen vorgesehen, um zu überprüfen, ob es Chargen-abhängige Qualitätsschwankungen gibt. Bei den nächsten Ringversuch-Durchläufen wird noch ein weiteres Illumina-Genotypisierungszentrum (L&B Bonn) hinzukommen. Die Markersätze aller vier Chip-Typen überlappen sich teilweise, so dass auf diesen überlappenden Bereichen auch ein Quervergleich Affymetrix vs. Illumina, GeneChip 100k vs. 400K Set oder Linkage IV Panel vs. HumanHap 300 möglich ist. Es kann also z.B. geprüft werden, ob ein bestimmter Marker mit der Affymetrix- bzw. mit der Illumina-Technik mit ähnlicher Genauigkeit genotypisiert wird, oder ob die großen Gen-Chips eine ähnliche Qualität besitzen wie die in der ersten Runde getesteten kleineren Gen-Chips

Alle Zentren haben vom Genetisch-Epidemiologischen Methodenzentrum (GEM) Bonn die gleiche, ihnen fremde DNA zur Genotypisierung erhalten. Im Nachhinein stellte sich heraus, dass zwei der 6 Blutproben genetisch identisch waren, weil sie von eineiigen Zwillingen abstammten.

Die aus Blut gewonnenen DNA-Proben wurden vom Institut für Experimentelle Hämatologie und Transfusionsmedizin (IEHT) Bonn zur Verfügung gestellt. Zuvor wurden die Proben anonymisiert. Alle Schritte des Versuches wurden mit der Ethikkommission in Bonn abgestimmt.

Die Genotypisierungs-Ergebnisse wurden ans GEM Bonn geschickt und hinsichtlich Reproduzierbarkeit und Typisierungsrate ausgewertet. Die Auswertung des ersten Ringversuch-Durchlaufes hat ergeben, dass insgesamt mit sehr hoher Genauigkeit genotypisiert wurde. Die Genotypisierungsrate und Übereinstimmung der Ergebnisse war generell sehr hoch. Die Typisierungsraten (= Anzahl der bestimmten Genotypen/Anzahl der durchzuführenden Genotypisierungen) liegen zwischen 97,63% und 99,61%. Die Fehlerrate

wurde aus den 4875 Markern geschätzt, die bei den Illumina- und Affymetrix-Chiptypen überlappen. Dabei wurde das Maximum-Likelihood-Modell verwendet. Die so geschätzte Fehlerrate lag bei allen Genotypisierungszentren und Proben unter 0,3% (zwischen 0,0096% und 0,211%). Eines der Genotypisierungszentren hatte im Vergleich zu den anderen Zentren eine leicht erhöhte Fehlerrate. Das könnte an einem erhöhten Durchsatz liegen. Aber auch dieses Zentrum hatte im Vergleich zu dem sonst üblicherweise angelegten Maßstab ein sehr hohes Qualitätsniveau.

Zusammenfassend zeigte der Ringversuch im ersten Durchlauf die hohe Qualität der Genotypisierungen im NGFN. Die Auswertung ergab auch, dass mit der Affymetrix- und der Illumina-Plattform vergleichbar gute Genotypisierungs-Ergebnisse erzielt werden.

Die Kodierung der Polymorphismen wird von den Zentren unterschiedlich gehandhabt. Daher sind die Genotypisierungs-Ergebnisse in drei unterschiedlichen Formaten an das GEM Bonn geschickt worden: (a) Format der NCBI Datenbank dbSNP, (II) Ergebnisse als AA, BB homozygot, AB heterozygot, (III) ACGT-Kodierung (proprietär). Eine Transformation der Kodierung ist zwar möglich, allerdings mit einem Problem behaftet: Affymetrix und Illumina geben für die verschiedenen Marker die Allele nicht immer im Bezug auf den gleichen der beiden komplementären DNA-Stränge an. Um das Problem zukünftig von vorneherein zu umgehen, wird eine von allen Zentren einheitlich vorgenommene, eindeutige, standardisierte Kodierung angestrebt.

In einer weiteren Genotypisierungs-Runde ist der Vergleich von variablen und frei definierbaren Markersätzen vorgesehen. Diese Runde ist zeitlich abhängig von den an den Zentren laufenden Genotypisierungsprojekten. In jedem Experiment mit variablen SNP-Markersätzen sollen zur Kontrolle bis zu 10 ausgewählte SNPs („Kontroll-SNPs“) mit bis zu 10 Proben zusätzlich typisiert werden. Die DNA-Proben werden vom GEM Bonn zur Verfügung gestellt.

In einem dritten Schritt soll die Qualität der Genotypisierung von konkreten, einzelnen SNPs kontrolliert werden.

Das Ergebnis dieser Auswertungen wird den Zentren in einem internen Bericht mitgeteilt. Anschließend werden Verfahren und Maßnahmen zur Optimierung der Genotypisierungsqualität diskutiert. Eine Zusammenfassung der Auswertung wird auch dem Projektkomitee vorgelegt. Die Ergebnisse der Ringversuche könnten, wenn die Zentren einverstanden sind, später publiziert werden, um die Qualität der Genotypisierung zu dokumentieren.

## Vortrag „Die NGFN Sample ID: Schnittstelle für den Material- und Datenaustausch in genetisch-epidemiologischen Projekten“

Referentin: Amalia Diaz Lacava

Universität Bonn

Institut für Medizinische Biometrie, Informatik und Epidemiologie

Sigmund-Freud-Straße 25, 53105 Bonn

diaz-lacava@imbie.meb.uni-bonn.de

Die Genetische Epidemiologie befasst sich mit der Verbreitung von Krankheiten und deren wichtigsten genetischen Risikofaktoren in der Bevölkerung. Bei diesen groß angelegten Projekten, bei denen Experten verschiedener Disziplinen (Kliniker, Humangenetiker, Statistiker u.a.) an unterschiedlichen Standorten zusammenarbeiten, ist der Austausch von Material und Daten erforderlich. Dieser Austausch erfolgt in entweder anonymisierter oder pseudonymisierter Form mit Unterstützung von Identifikationsnummern (IDs). Aus Datenschutzgründen sollten zwei separate Identifikationsnummern für Personen (Probanden, Patienten, Familienangehörige, Kontrollpersonen) und Material (Blut, Gewebe, DNA) vergeben werden. Bei pseudonymisierten Daten wird mithilfe eines Datenerfassungssoftwaresystems eine logische Verbindung von Person- und Sample-ID hergestellt (Pseudonym-Server).

Um die Dokumentation und die Nachverfolgung von Proben während des Genotypisierungsprozesses zu erleichtern, sollte die Proben-Identifikationsnummer (Sample-ID) auf internationalen Standards (nach ISO/IEC) basieren und global eindeutig sein ("unique identification mark", UIM). Jedes Probengefäß sollte mit einem Etikett versehen sein, das die Proben-Identifikationsnummer sowohl in maschinenlesbarer Form als auch als im Klartext enthält. Der maschinenlesbare Teil kann – je nach Größe der Probengefäße – entweder als linearer Barcode oder als 2D Datamatrix dargestellt werden. Ein 2D Datamatrix Code benötigt nur einen Platz von  $3,2 \times 3,2 \text{ mm}^2$  und kann deshalb sogar auf den kleinsten Probengefäßen angebracht werden. So etikettierte Probengefäße können durch maschinelle Lesegeräte schnell und einfach identifiziert und in eine automatische Prozesskette integriert werden. Ein 2D Datamatrix Code enthält redundante Daten und hat deshalb den Vorteil, dass die Identifikationsnummer selbst bei teilweiser Beschädigung des Etiketts über einen Fehlerkorrektur-Algorithmus (EC-200) noch lesbar ist.

Die Identifikationsnummern sollten Prüfziffern enthalten, um Fehler bei der manuellen Eingabe zu erkennen. Die Prüfziffer wird über einen Algorithmus generiert und die Sample-ID besteht dann aus Form DEnn.cc.xx...xx, wobei cc die zweistellige Prüfziffer ist. Die richtige Eingabe der Identifikationsnummer kann über einen Kontroll-Algorithmus validiert werden. Wird die Identifikationsnummer falsch eingegeben, errechnet der Algorithmus eine ungültige Zahl und es erfolgt eine Meldung, dass die Eingabe nicht korrekt war. Das bedeutet, daß ungültige Eingaben sofort zurückgewiesen werden.

Im IMBIE wurde ein Datenerfassungs-System (DES) entwickelt, die es ermöglicht, die Daten effizient und in hohe Qualität zu erfassen, wobei eine logische Verbindung von Person- und Sample-ID hergestellt wird. Plausibilitäts-Checks ermöglichen es dem rekrutierenden Arzt, die Eingabe auf Richtigkeit zu überprüfen. So werden Fehler beim Ausfüllen von Formularen

auf ein Minimum reduziert. Weitere Plausibilitäts-Checks sind auf verschiedenen anderen Ebenen möglich, z.B. bei der Datenabfrage, bei der Datenzusammenfassung und durch die graphische Darstellung des Familien-Stammbaumes.

Um die Qualität der Genotypisierungen zu evaluieren, werden die Genotypisierungsergebnisse einem Set von Qualitätsindikatoren unterworfen.

- (I) Die Genotypisierungsrate (= Anzahl der korrekten Genotypisierungen/Anzahl der zur Genotypisierung bereitgestellten Proben) ermöglicht eine Abschätzung des Erfolges und der Qualität des Genotypisierungsprozesses.  
Die Genotypisierungsrate pro Person ist hilfreich, um ungeeignetes oder unzureichendes Material zu identifizieren.  
Die Genotypisierungsrate pro Marker ermöglicht die Identifikation von unzuverlässigen Markern, bzw. Assays.
- (II) „Allel-sharing“-Methoden testen die angegebenen Verwandtschaftsbeziehungen zwischen den Probanden
- (III) Das „De Finetti Diagramm“ ermöglicht eine anschauliche Validierung der einzelnen Marker und testet das „Hardy-Weinberg Gleichgewicht“ (HWE) der Genotypverteilung in einer Stichprobe. Abweichungen vom HWE sind oft als starke Hinweise auf Genotypisierungsfehler zu werten, denen dann gezielt nachgegangen werden kann.

#### Vortrag „**Qualitätsmanagement der AG Microarrays**“

Referent: Dr. Ruprecht Kuner  
Deutsches Krebsforschungszentrum  
Molekulare Genomanalyse  
Im Neuenheimer Feld 580, 69120 Heidelberg  
r.kuner@dkfz-heidelberg.de

Die Microarray-Technologie ist eine leistungsfähige und effiziente Methode zur Analyse von komplexen Expressionsmustern. Allerdings können hier bereits kleine Unterschiede im experimentellen Design, der Probenaufbereitung, der Plattform-Technologie und der Datenanalyse dazu führen, dass Daten aus verschiedenen Microarray-Experimenten nicht miteinander vergleichbar sind. Vorrangiges Ziel des Qualitätsmanagements zur Micro-Array-Technologie im NGFN ist deshalb die Definition von minimalen Standards unter Einhaltung der internationalen Vereinbarungen (MIAME, MAGE-ML, usw.), um die Vergleichbarkeit und Reproduzierbarkeit von Microarray-Experimenten zu gewährleisten.

Die Arbeitsgruppe "Microarrays", die alle Microarray-Plattformen im NGFN repräsentiert, definierte Minimalstandards für etwa 200 relevante Parameter, die in Form von 26 Protokollen zur Verfügung stehen. Die Formulierung der Protokolle hat zur Qualitätssicherung einen Konsensprozess durchlaufen, ihre Nutzung beruht auf freiwilliger Selbstverpflichtung. Die Protokolle sind auf der NGFN-Webseite ([http://www.science.ngfn.de/index\\_455.htm](http://www.science.ngfn.de/index_455.htm)) und auf der Webseite der SMP-RNA (<http://www.dkfz.de/smp-rna/rna.org/groups.asp?siteID=23>) abrufbar. Protokolländerungen



werden kontinuierlich nach dem Stand der Technologie vorgenommen und durch entsprechende Eintragungen in letzten Abschnitt des Protokolls („Track of Changes“) sowie durch die Vergabe von Versionsnummern dokumentiert.

Die Protokolle sind die Grundlage für die Vergleichbarkeit von Expressions-Daten durch eine standardisierte Ausführung der Microarray-Experimente und durch eine einheitliche Datenanalyse. Sie spezifizieren außerdem Maßnahmen zur Qualitätssicherung auf verschiedenen Ebenen des Microarray-Experimentes. Dies umfasst

- (I) die Qualitätskontrolle der Ausgangsmaterialien (RNA/DNA) durch UV-Spektroskopie, Gelelektrophorese und quantitative RT-PCR.
- (II) Strichkode-Markierung der Mikrotiter-Platten und Nachverfolgung aller zugehörigen Verarbeitungsschritte durch ein Laborinformations-Managementsystem (LIMS).
- (III) Sorgfältige Überprüfung, Dokumentation und Analyse der Microarray-Rohdaten durch die Software, arrayMagic. Die automatisierte Analyse-Pipeline von arrayMagic umfasst Datenimport, Quantifizierung und Visualisierung der Daten, Qualitätskontrolle, Normalisierung, statistische Auswertung und Datenexport. Die Software ermöglicht so die schnell und effiziente Bewertung der Hybridisierungsqualität und gewährleistet qualitätskontrollierte Microarray-Daten.

Es ist geplant, die Standardprotokolle als Teil einer NGFN-weiten Initiative, die die Dokumentation aller im NGFN eingesetzten Hochdurchsatztechnologien umfasst, in einem wissenschaftlichen Journal veröffentlicht werden. Die Publikation einer derartigen technikübergreifenden SOP-Sammlung aus einem Forschungsnetz wäre bisher einmalig.

Als zukünftige Aktivitäten der Arbeitsgruppe "Microarrays-QM" sind (I) der Datentransfer von arrayMagic in die iChip-Datenbank, (II) die automatisierte Nutzung von öffentlichen Datensätzen, (III) die Anbindung an weitere QM-Arbeitsgruppen, (IV) Microarray-Ausbildungskurse und (V) die Steigerung der Transparenz dieser Aktivitäten geplant.

## Vortrag „**Qualitätsmanagement bei der Open Reading Frame (ORF)-Klonierung**“

Dr. Ruth Wellenreuther  
Deutsches Krebsforschungszentrum  
Molekulare Genomanalyse  
Im Neuenheimer Feld 580, 69120 Heidelberg  
[r.wellenreuther@dkfz-heidelberg.de](mailto:r.wellenreuther@dkfz-heidelberg.de)

Die SMP Cell verfolgt die Strategie der Open Reading Frame (ORF)-Klonierung, um neue Gene zu identifizieren und zu klonieren, und um Spleißvarianten zu finden, die bislang noch nicht erkannt wurden.

Die ORF-Amplifikation erfolgt entweder mittels PCR ausgehend von cDNA- oder EST-Klonen, die den gewünschten ORF enthalten, oder über RT-PCR unter Verwendung von RNA aus ausgewählten Zell-Linien oder Geweben. Die PCR wird mit zwei geringfügig

unterschiedlichen 3'-Primern durchgeführt, so dass zwei Varianten von PCR-Produkten entstehen, und zwar ein PCR-Produkt mit Stopcodon und eines ohne Stopcodon. Die vervielfältigten ORFs werden dann mit dem Gateway® cloning System (Invitrogen) in einen Entry Vektor rekombiniert. In einem weiteren Schritt werden die ORFs in GFP-Expression-Vektoren subkloniert. Die PCR-Produkte werden sowohl in N-terminal als auch in C-terminal markierte GFP-Expressions-Vektoren (N- and C-terminally tagged GFP expression vectors) kloniert, wobei die PCR-Produkte mit Stopcodon in den N-terminal-markierten GFP-Expressionvektor eingefügt werden, so dass die exprimierten Fusionsproteine an dem natürlichen C-Terminus des zu untersuchenden Proteins enden.

Standardprotokolle, Nomenklaturstandards und Informationen zu den Expressionsvektoren gibt es auf den NGFN-Webseiten ([http://www.science.ngfn.de/index\\_457.htm](http://www.science.ngfn.de/index_457.htm)), auf den Webseiten der SMP-Cell (<http://www.smp-cell.org/>) und im NGFN-Intranet. Die Protokolle beschreiben den Hintergrund eines Experimentes und sie listen detailliert die Arbeitsschritte sowie Versions-Änderungen auf. Letztere sind erforderlich, um schnell auf veränderte Anforderungen innerhalb der kooperierenden Projekte reagieren zu können, und um verbesserte Techniken, Anwendungen und Methoden zu berücksichtigen.

Kürzlich wurden ein Labor Informationsmanagement System (LIMS) und ein Datenbanksystem fertiggestellt. Mit diesem LIMS ist es möglich, lückenlos den Arbeitsfluss und den Status der bearbeiteten ORFs zu verwalten und zu verfolgen (Fig. 1). Das LIMS vergibt automatisch einen standardisierten Namen an jeden ORF. Diese einheitliche Nomenklatur ist essentiell, um Verwechslungen bei späteren Arbeitsschritten auszuschließen. Über ein Fenster des Anwendungsprogramms SCISSORS kann man sich die erzeugten Entry-Klone und die Ergebnisse der Qualitätskontrolle anzeigen lassen.

Alle Entry-Klone werden daraufhin überprüft, ob sie ein Insert enthalten. Wenn ja, dann werden zunächst 4 dieser Klone sequenziert. Wenn die Klone fehlerhaft ist, werden weitere Klone sequenziert. Diese Schritte werden im LIMS genau dokumentiert und verwaltet und die daraus resultierenden Folgeschritte für den jeweiligen ORF angezeigt. Wenn die Entry-Klone die Qualitätskontrolle nicht bestehen (z.B. weil sie eine Mutation enthalten, die das Leseraster verschiebt), dann fallen sie aus dem weiteren Prozess heraus und die erforderlichen Arbeitsschritte werden von dem LIMS automatisch angezeigt.

Insgesamt sind die Arbeitsschritte bei der Herstellung von ORFs weitgehend automatisiert, so dass im Idealfall nur ein minimaler manueller Aufwand erforderlich ist. Aufwändig ist das Reorganisieren von Arbeitsschritten und damit auch von Proben im Labor, wenn es Ausfälle an den verschiedenen Arbeitsschritten des Prozesses gibt und diese wiederholt werden müssen. Ausfälle gibt es vor allem beim RT-PCR-Schritt auf RNA, oft wenn es sich um ein sehr gering exprimiertes Gen handelt, und bei der Rekombination in den Entry-Vektor, besonders wenn es sich um einen sehr großen ORF handelt.

## Vortrag „Der MIACA Standard für Hochdurchsatz-Zell-Assays“

Referent: Dr. Stefan Wiemann  
Deutsches Krebsforschungszentrum  
Molekulare Genomanalyse  
Im Neuenheimer Feld 580, 69120 Heidelberg  
s.wiemann@dkfz-heidelberg.de

Eine der Herausforderungen der systematischen Genomforschung besteht darin, die Funktionen der Gene und Proteine sowie deren Rolle in krankheitsrelevanten Prozessen aufzuklären. Solche Funktionsanalysen können inzwischen in großem Maßstab und effizient mit Hochdurchsatz-Zell-Assays durchgeführt werden. Hier werden biologische Vorgänge in der Zelle durch die Unter- oder Überexpression von Proteinen systematisch gestört (Perturbation), und die dadurch induzierten Effekte im Anschluss gemessen. So können neue Targets identifiziert, und diese zellulären Prozessen zugeordnet werden.

Um die Daten von verschiedenen Hochdurchsatz-Zell-Assays auszuwerten, zu vergleichen und in einen Gesamtkontext zu integrieren, ist es notwendig, die Beschreibung von Zell-basierten Assays international zu standardisieren und die minimal notwendigen Experimentannotationen zu spezifizieren. Nur so ist es möglich die Qualität und Relevanz von Datensätzen verschiedener Labore zu evaluieren, und letztlich über deren Verwendbarkeit entscheiden zu können. Zu diesem Zweck haben NGFN-Wissenschaftler der SMP Cell in Anlehnung an die Entwicklungen von MIAME und MAGE-ML eine internationale Initiative gestartet - **MIACA** (**M**inimal **I**nformation **A**bout a **C**ellular **A**ssay). Im Rahmen dieser Initiative werden Minimal-Informationsrichtlinien (minimum information guidelines), ein standardisiertes Vokabular sowie ein einheitliches Datenformat erarbeitet. Diese Arbeiten werden zu einer Vereinheitlichung von Daten-Beschreibungen führen und dienen als Grundlage für den Aufbau von öffentlichen Datenbanken. Strukturen und Ontologien von bereits bestehenden Standards (FuGE, OBI, GO, MGED, HUPO-PSI) werden dabei weitgehend genutzt und integriert, um MIACA kompatibel mit etablierten Standards zu machen.

MIACA ist modular konzipiert, um die Vielfalt von zellulären Assays sinnvoll und umfassend abbilden zu können. Folgende Module sind derzeit vorgesehen:

- (I) Studien Design (administrative Angaben, Projektbeschreibung, Zelllinie, Materialien & Geräte)
- (II) Behandlung der lebenden Zellen (Zellkulturbedingungen vor der Perturbation, Zellpassage, Anzahl der Zellen, usw.)
- (III) Bedingungen und Art der Perturbation
- (IV) Nachbehandlung der Zellen (Zellkulturbedingungen nach dem Experiment, Art der Fixierung, Antikörper-Inkubationen, Waschschrte. usw.)
- (V) Datenerfassung (Art der Datenerfassung, Geräteeinstellungen usw.)
- (VI) Datenanalyse (z.B. Verfahren für das Filtern und Normalisieren der Daten)

Die Experimentellen Module II-V können nahezu beliebig miteinander kombiniert werden, einzelne Module können auch mehrfach implementiert werden. Dies erlaubt auch komplexe

und mehrstufige Assays vollständig zu erfassen. Die Angaben werden schließlich über ein XML-Austauschformat in eine Datenbank transferiert.

Es ist geplant, den Fortschritt der MIACA-Initiative in einem hochrangigen Journal zu veröffentlichen, um die Sichtbarkeit und Akzeptanz von MIACA als weltweiten Standard für die funktionelle Genomik und Systembiologie weiter zu erhöhen. Informationen zu den MIACA-Standards sind unter <http://miaca.sf.net> und <http://sf.net/projects/miaca> abrufbar.

### Vortrag „**TMF Projekt Biomaterialbanken - Zusammenfassung der auf dem Symposium in Berlin am 27.04.2006 vorgestellten Ergebnisse**“

Referent: Dr. Olaf Krüger (in Vertretung von Prof. Dr. Michael Krawczak, Kiel)  
NGFN Projektmanagement, Projektträger DLR  
Heinrich-Konen-Str. 1, 53227 Bonn  
Email:olaf.krueger.1@dlr.de

Am 27. April 2006 stellte die TMF-Arbeitsgruppe „Biomaterialbanken“ im Rahmen eines Symposiums in Berlin die Ergebnisse ihres umfangreichen Projektes vor.

Ziel dieses TMF-Projektes war die Erstellung eines umfassenden Konzeptes für den Betrieb von Biomaterialbanken mit Klärung der rechtlichen Grundlagen, insbesondere der datenschutzrechtlichen Rahmenbedingungen. Im Rahmen des Projektes sind Rechtsgutachten, Datenschutzkonzepte und Musterverträge erstellt worden. Um die konkrete Umsetzung der Ergebnisse in die Praxis zu erleichtern, wurden darüber hinaus Checklisten und Textvorlagen zu Patienteneinwilligungen sowie Checklisten zur Qualitätssicherung beim Probenmanagement in Biobanken erarbeitet.

Nachfolgend sind die wichtigsten Aspekte zu den Themen Persönlichkeitsrechte und Patienteneinwilligung zusammengefasst, die in den Vorträgen von Prof. Dr. Jürgen W. Goebel (RAe Goebel & Scheller, Bad Homburg), von Peter Ihle (Kompetenznetz Maligne Lymphome, Universitätsklinikum Köln) und Prof. Dr. Michael Krawczak (Christian-Albrechts-Universität Kiel) dargestellt wurden.

#### • **3-Schichten-Modell der Rechte an Bioproben**

In Bezug auf die Rechte des Probanden an seiner Probe sollten drei Ebenen berücksichtigt werden:

(I) Auf der ersten Ebene werden die Biomaterial-Proben als Sachen im Sinne des §90 BGB angesehen und befinden sich zunächst gemäß §953 BGB im Eigentum des Probanden.

Die Biomaterialbank kann ein Bioprobe in ihrer Sacheigenschaft als Eigentum erwerben (durch ein Rechtsgeschäft). Dies ist bei der anonymisierten Verarbeitung einer Probe erforderlich, denn die Probe ist nach der Anonymisierung dem Zugriff des Patienten dauerhaft entzogen.

Bei einer pseudonymisierten Verarbeitung der Probe ist die Vereinbarung einer Nutzungsübertragung sinnvoll, denn der Patient kann jederzeit seine Probe ohne Angabe von Gründen zurückfordern. Die Nutzungsübertragung erfolgt ebenfalls durch einen Vertrag.

- (II) Die zweite Ebene betrifft das Grundrecht auf „informationelle Selbstbestimmung“, aufgrund dessen der Spender das Recht hat, grundsätzlich selbst darüber zu entscheiden, ob und wem er seine personenbezogenen Daten zu welchem Zweck überlässt.
- (III) Die dritte Ebene umfasst das Allgemeine Persönlichkeitsrecht. Die Probe wird hier nicht nur als Sache, sondern auch als Teil einer Person wahrgenommen. Die Probe als Eigentum kann zwar abgetreten werden, Persönlichkeitsrechte aber nicht. Deshalb darf eine Bioprobe auch nach Eigentumsübertragung nicht beliebig verwendet werden. In keinem Fall dürfen bei der Verwendung der Probe die allgemeinen Persönlichkeitsrechte des Probanden nicht verletzt werden.

Aus diesen Rechtsgrundlagen und weiteren Gesetzen (z..B. Bundesdatenschutzgesetz und Länderdatenschutzgesetze) folgt, dass für die Nutzung einer Probe zum Zweck der medizinischen Forschung eine schriftliche Einwilligung des Patienten/Probanden vorliegen muss. Außerdem muss der Proband über die Verwendung seiner Probe schriftlich und ergänzend mündlich umfassend informiert werden, u.a. über

- das Ziel der Forschung,
- die Dauer der Aufbewahrung und Nutzung der Proben und Daten
- die Form der Aufbewahrung (anonymisiert oder pseudonymisiert)
- mögliche Anlässe und Berechtigte zu einer Re-Identifizierung des Probanden (bei pseudonymisierter Aufbewahrung)
- den Hinweis auf die Freiwilligkeit der Einwilligung und darauf, dass eine Ablehnung keinerlei Einfluss auf die medizinische Behandlung hat,
- den Hinweis auf das Recht des Spenders, die Einwilligung für die Zukunft zu widerrufen und in diesen Fällen eine Herausgabe oder Vernichtung der Probe zu verlangen (bei pseudonymisierter Aufbewahrung).

- **Reichweite der Einwilligung**

Die Probe unterliegt in der Regel der Zweckbindung und darf nicht ohne Zustimmung des Probanden für weitere Forschungsfragen genutzt oder an Dritte weitergegeben werden. Oft ist es aber sinnvoll, dass die Proben für ein großes Spektrum von Forschungsfragen zur Verfügung stehen, auch für solche, die zum Zeitpunkt der Einwilligung noch nicht bekannt sind (z.B. im Falle von Biomaterialbanken). Hier ist nach Auffassung des Nationalen Ethikrates eine Ausnahme vertretbar und die Einwilligungserklärung kann im Hinblick auf den Forschungszweck, die Nutzungsdauer und den Nutzerkreis offen formuliert werden. Allerdings müssen die mit dieser Erleichterung zugunsten der Forschung einhergehenden Risiken durch festgelegte Rahmenbedingungen und zusätzliche Sicherheitsmaßnahmen kompensiert werden, um so zu verhindern, dass Daten an unberechtigte Dritte gelangen.

Der Patient hat auf jeden Fall ein Recht auf Widerruf seiner Einwilligung und behält auf diese Weise die Kontrolle über seine Daten.

Alternativ kann der Patient mit einer abgestuften Einwilligungserklärung die Reichweite seiner Einwilligung individuell festlegen:

- **Pseudonymisierung/Anonymisierung von Proben**

Proben und Daten sind nach der Einwilligung und vor der Aufnahme in die Bio- / Datenbank entweder vollständig zu anonymisieren oder - wenn aus Forschungsgesichtspunkten eine Zuordnung zum Spender / Betroffenen möglich bleiben soll – sicher zu pseudonymisieren,

d.h. der Name und andere Identifikationsmerkmale sind durch einen Code zu ersetzen, den ausschließlich berechnete Personen in vorher festgelegten Bedarfsfällen dem Probenspender / der betroffenen Person wieder zuordnen können. Je nach Umfang und Sensitivität der Daten kann eine externe Pseudonymverwaltung durch einen Datentreuhänder geboten sein.

Eine absolute Anonymisierung von humanem Biomaterial ist in der Regel nicht möglich, denn wenn eine Referenzprobe vorliegt, ist eine Rückidentifizierung auf den Spender grundsätzlich möglich. Das Bundesdatenschutzgesetz fordert allerdings lediglich die „faktische“ Anonymisierung, die durch die Ausgestaltung der Rahmenbedingungen zu gewährleisten ist.

Bei einer geplanten Anonymisierung der Daten/Materialien ist der Patient darauf hinzuweisen, dass er bestimmte Rechte dadurch nicht mehr wahrnehmen kann, beispielsweise das Auskunftsrecht, die Mitteilung von individuellen Ergebnissen, das Recht auf Berichtigung/Löschung/Vernichtung und der Widerruf der Verarbeitung.

Eine anonymisierte Probe kann ohne Einschränkung an Dritte übereignet werden. Sie unterliegt jedoch in Analogie zu §40 BDSG weiterhin der Zweckbindung.

Für eine langfristige Nutzung von Biomaterialien ist eine Pseudonymisierung zu bevorzugen. So können neu erhobene mit bereits erfassten Daten/Proben zusammengeführt werden. Außerdem kann der Patient später noch einmal kontaktiert werden, z. B. zum Zwecke der Zusatzerhebung weiterer Daten. Bei einer Pseudonymisierung ist aber Datensicherheit zu gewährleisten und das Rückidentifizierungsrisiko muss auf ein Minimum reduziert werden. Dies kann durch technische Maßnahmen (Verschlüsselung, Firewalls, Zugangsberechtigungen, Monitoring) und organisatorische Maßnahmen (Informationstrennung, d.h. Speichern der Personendaten, medizinischen Daten und Biomaterialdaten in physikalisch getrennten Datenbanken, getrennte Lagerung der Proben, vertragliche Regelungen, SOPs usw.) erreicht werden.

Der Proband muss seine Einwilligung dafür erteilen, dass seine Materialien und Daten in pseudonymisierter Form gelagert bzw. gespeichert werden. In diesem Zusammenhang muss er auch Auskunft erhalten, wer ihn wie und wann zu welchem Zweck kontaktieren darf.

- ***Recht des Patienten/Probanden auf Auskunft***

Der Proband hat das Recht, durch aktive Nachfrage jederzeit zu erfahren, wer welche Daten über ihn speichert. Die Daten speichernden Stellen bzw. die Proben lagernden Stellen müssen, solange die Daten personenbezogen (pseudonymisiert) vorliegen, Auskunft erteilen. In Bezug auf die Herausgabe persönlicher Untersuchungsergebnisse müssen neben dem Recht auf Auskunft aber auch die bestehenden klinischen Standards beachtet werden. So sollte die Auskunft über persönliche genetische Untersuchungsergebnisse immer auch mit einer ärztlichen Beratung bzw. Betreuung verknüpft sein. Wenn ein Proband ausdrücklich Auskunft über persönliche Untersuchungsergebnisse verlangt, wird empfohlen, diese dem Probanden über den behandelnden Arzt zugänglich zu machen.

- ***Recht des Patienten/Probanden auf Widerruf einer Einwilligung***

Im Rahmen einer Verarbeitung auf der Grundlage einer Einwilligung hat der Patient jederzeit das Recht auf Widerruf seiner Einwilligung und er sollte in jedem Fall auf dieses Recht hingewiesen werden. Bei Widerruf müssen die erhobenen Daten entweder gelöscht

oder anonymisiert werden. Verlangt der Proband anstelle der Anonymisierung ausdrücklich die Löschung seiner Daten und die Vernichtung seiner Proben, so muss dem Folge geleistet werden.

Das Recht auf Widerruf kann durch Spezialgesetze eingeschränkt werden (z. B. AMG §40): Die Einwilligung ist dann unwiderruflich, wenn die Daten notwendig sind,

- a) die Wirkungen des zu prüfenden Arzneimittels festzustellen,
- b) sicherzustellen, dass schutzwürdige Interessen der betroffenen Person nicht beeinträchtigt werden,
- c) der Pflicht zur Vorlage vollständiger Zulassungsunterlagen zu genügen

- **Zentrale Kontaktstelle**

Im Rahmen einer Biomaterialbank kann die Zahl der Daten speichernden und Proben lagernden Stellen relativ groß werden. Die Einrichtung einer zentralen Kontaktstelle als primäre Ansprechstelle für den Probanden ist daher empfehlenswert.

- **Das Problem der »Altproben«**

Es gibt zahlreiche Proben (z. B. im Labor einer Klinik), die im Rahmen einer Behandlung entnommen wurden und für die keine Einwilligung des Betroffenen für eine Verwendung zu Forschungszwecken vorliegt. Der Umgang mit solchen „Altproben“ ist rechtlich nicht eindeutig geregelt. Durch (einige) Landeskrankenhausgesetze wird die eigene Forschung der Klinik an diesen Altproben gedeckt. Die Weitergabe in anonymisierter Form wird von einigen Landesbeauftragten für Datenschutz vorläufig geduldet, von anderen nicht.

Aufgrund der sehr unsicheren Rechtslage bei Verwendung von Altproben zu Forschungszwecken sollte, sofern der Proband bekannt oder rückidentifizierbar ist, eine nachträgliche Einwilligungserklärung eingeholt werden.

- **Publikation der Ergebnisse in der TMF-Schriftenreihe**

Die mit den Datenschutzbeauftragten des Bundes und der Länder abgestimmten Ergebnisse der Arbeitsgruppe „Biomaterialbanken“ wurden in der TMF-Schriftenreihe publiziert:

Simon / Paslack / Robiński / Goebel / Krawczak

**Biomaterialbanken - Rechtliche Rahmenbedingungen**

*Schriftenreihe der Telematikplattform für Medizinische Forschungsnetze*

U. Harnischmacher / P. Ihle / B. Berger / J. Goebel / J. Scheller

**Checkliste und Leitfaden zur Patienteneinwilligung**

Grundlagen und Anleitung für die klinische Forschung