

Serviceangebot der

Technologieplattform

für Mikrobielle Genomforschung (TPMG)

In den BMBF geförderten GenoMik- und GenoMik-Plus-Netzwerken hat sich in den vergangenen Jahren eine Forschungsinfrastruktur entwickelt, die auf drei Feldern zur Ausbildung von Zentren mit den Schwerpunkten Genomsequenzierung und -annotation, Bioinformatik und Proteomik geführt hat. Diese Zentren wurden durch die Gründung einer Technologieplattform für Mikrobielle Genomforschung weiterentwickelt, die aus den folgenden drei Bereichen besteht:

I. Bereich „Genomsequenzierung und –annotation“, Göttingen

II. Bereich „Bioinformatik“, Bielefeld

III. Bereich „Proteomik“, Greifswald

I. Bereich „Genomsequenzierung und –annotation“, Göttingen

Zusammenfassung

Im Rahmen der GenoMik und GenoMik-Plus Programme hat sich an der Georg-August-Universität Göttingen ein international renommiertes Zentrum für die Sequenzierung und die Annotation von mikrobiellen Genomen und Metagenomen herausgebildet. Dies schließt die Aufarbeitung der Sequenzdaten für die funktionelle Genomanalyse ein. Verfahren zur Größenbestimmung mikrobieller Genome, zur Herstellung von Plasmid-Banken mit definierter Insert-Größe sowie von Cosmid-, Fosmid- und cDNA-Banken sind etabliert sowie die mikrobielle Genomsequenzierung im Hochdurchsatz. Strategien und entsprechende bioinformatische Werkzeuge stehen zur Verfügung, um möglichst effektiv die Sequenzdaten zu assemblieren, zu editieren und Sequenzlücken zu schließen. Die Leistungsfähigkeit wird dokumentiert durch zahlreiche hochkarätige Publikationen.

Das Sequenzierzentrum verfügt über eine Datenbank, die alle sequenzierten mikrobiellen Genome enthält und die Annotation sequenzierter Genome und vor allem die Stoffwechselrekonstruktion und die vergleichende Genomanalyse ermöglicht. Darüber hinaus steht ein Transkriptom-Labor zur Verfügung, in dem genomweite DNA-Microarrays hergestellt und für physiologische Untersuchungen auch unter Einsatz von kontinuierlichen Kulturen eingesetzt werden können. Dies wird ergänzt durch Meta-Transkriptom-Analysen auf Basis der Sequenzierung von cDNA aus mikrobieller RNA. Die bisher erzielten Leistungen sind dargestellt unter <http://www.g2l.bio.uni-goettingen.de/>.

Leistungsbeschreibung

Der Bereich Göttingen der TPMG ist durch erfolgreich abgeschlossene nationale und internationale Kooperationen auf Arbeiten eingestellt, die eine moderne automatisierte Ausstattung für die Sequenzierung von DNA im Hochdurchsatz und die kompetente bioinformatische Auswertung beinhalten, wobei diese Arbeiten durch hochqualifiziertes wissenschaftliches und technisches Personal durchgeführt werden.

Darüber hinaus liegen spezielle Erfahrungen für die komparative Genomik und die sich anschließende Transkriptomik und Meta-Transkriptomik vor. Das Angebot umfasst folgende Leistungen:

- Herstellung von Genbanken aus genomischer und metagenomischer DNA bzw. RNA (cDNA)
- Durchführung der Sequenzierungen im Hochdurchsatz (Pyrosequenzierung: Roche FLX- und Titanium-Technologie, sowie Sanger-Sequenzierung z.B. ABI 3730)
- Bearbeitung der Rohdaten mit dem Ziel der Assemblierung, der Editierung der Datensätze, der einschlägigen Verfahren zum Lückenschluss (unter Verwendung von Phred, den Programmen des Stadenpackages, ERGO, Newbler)
- Automatische Annotation und manuelle Annotation unter Einsatz von YACOP, TIS, SIGI, COLOMBO, ERGO, BI-BLAST
- Stoffwechselrekonstruktion und komparative Genomik, z.B. zum Vergleich von Wildtypstämmen mit Hochleistungsstämmen oder von pathogenen und apathogenen, aber eng verwandten Bakterienarten, sowie von mikrobiellen Konsortien
- Herstellung von Oligonukleotiden für die Gewinnung von DNA-Microarrays
- Durchführung von Transkriptionsanalysen (z. B. mit Lucidea Array Spotter, Genepix 4000B Arrayscanner, ASP Hybridization System)
- Durchführung von Meta-Transkriptomanalysen
- Auswertung der DNA-Microarray-Experimente

PD Dr. Rolf Daniel
Prof. em. Dr. Gerhard Gottschalk

Laboratorium für Genomanalyse (G2L)
Institut für Mikrobiologie und Genetik
Georg-August-Universität Göttingen
Grisebachstraße 8, 37077 Göttingen

Tel: +49-551-39-3827
E-Mail: rdaniel@gwdg.de

II. Bereich „Bioinformatik“, Bielefeld

Zusammenfassung

Im Rahmen der Förderinitiative GenoMik-Plus wurde an der Universität Bielefeld die Technologieplattform-Bioinformatik eingerichtet. Diese bietet Forschungsk Kooperationen bei der bioinformatischen Auswertung von Daten aus Genom- und Postgenomanalysen an. Mittels eines Internet-basierten **GenoMik-Portals** erhalten Kooperationspartner Zugriff auf die in Bielefeld vorgehaltene Rechner- und Software-Infrastruktur (siehe www.cebitec.uni-bielefeld.de/brf/). Das Angebot auf dem Sektor **Genomik** umfasst die Assemblierung und Annotation bakterieller Genome. Die Funktionalität des Genomannotationsprogramms **GenDB** wird aktuell zur Annotation eukaryontischer Mikrobengenome weiterentwickelt. Daneben wird die Auswertung von Daten aus Metagenomanalysen mit etablierten Software-Anwendungen angeboten. Auf dem Gebiet **Transkriptomik** werden Programme wie **EMMA** zur Auswertung von Daten aus genomweiten Transkriptomanalysen zur Verfügung gestellt. Der Programm-übergreifende Zugriff auf gespeicherte Daten ist mittels eines spezifischen Daten-Integrationsprogramms möglich. Weitere Kooperationsangebote erstrecken sich auf die bioinformatische Auswertung experimenteller Daten aus **Proteom-** und **Metabolomanalysen** sowie auf experimentelle Arbeiten auf dem Sektor Transkriptomik.

Bioinformatik-Angebote mittels des GenoMik-Portals:

Die Bielefelder Bioinformatik-Softwareplattform wurde über ein Internet-basiertes GenoMik-Portal (www.cebitec.uni-bielefeld.de/groups/brf/software/portal-genomik) etabliert. Dieses ermöglicht einen geschützten und benutzerspezifischen, aber dennoch standardisierten Zugang zu den einzelnen mikrobiellen Genomprojekten bzw. zu den jeweiligen Software-Komponenten. Die Technologieplattform-Bioinformatik bietet die in Tabelle 1 aufgelisteten Leistungen für alle Projektpartner auf den Gebieten Genom- und Postgenomanalysen an. Das Angebot umfasst neben der Speicherung, Vorhaltung und Pflege der experimentellen Daten vor allem deren

bioinformatische Interpretation. So kann die automatische Annotation eines bakteriellen Genoms beispielsweise über Nacht durchgeführt werden.

Angebote für Genomik	Angebote für Transkriptomik
<ul style="list-style-type: none"> • Qualitätskontrolle von Sequenzdaten (Sanger, GS-FLX, Solexa, SOLiD) • Assemblierung von Genomsequenzen aus Sequenzierdaten inklusive Sequenzpolierung und Lückenschluss • Automatische Annotation von Genomsequenzen inklusive Visualisierung und Rekonstruktion von Stoffwechselwegen • Genomvergleiche ausgewählter Mikroorganismen • Speicherung aller erstellten Genomdaten • Komparative Genomanalysen wie z.B. die Bestimmung des Pangenoms • Speicherung und Hochdurchsatz-Analyse von Metagenomdaten inkl. taxonomischer Klassifikation 	<ul style="list-style-type: none"> • Berechnung spezifischer Oligonukleotide zur Erstellung von genomweiten Microarrays • Strukturierte Speicherung experimenteller Arbeitsabläufe bei Microarray-Experimenten • Erfassung, Auswertung und Visualisierung von Hybridisierungsdaten (Oligobasierend, Affymetrix, Agilent, Nimblegen) • Speicherung aller Daten aus Microarray-Experimenten • Erfassung, Speicherung und Visualisierung von genregulatorischen Netzwerken • Integration von Daten aus anderen „Omics“ Experimenten z.B. via Webservices • Auswertung von Sequenzbasierten Genexpressionsdaten (cDNAs)

Tabelle 1: Leistungen der Technologieplattform-Bioinformatik für Kooperationspartner auf den Gebieten Genomik und Transkriptomik

Die Technologieplattform-Bioinformatik bietet auch die bioinformatische Speicherung und Auswertung von quantitativen **Proteom-** und **Metabolomdaten** an. Die Programme **ProSE** bzw. **MeltDB** beinhalten jeweils eine LIMS Komponente zur strukturierten Speicherung von experimentellen Arbeitsabläufen und deren Ergebnissen. Weiterhin wird der Benutzer über eine Web-basierte Oberfläche bei der weitestgehend automatisierten Analyse der Daten unterstützt und es können unterschiedliche Visualisierungen zur Repräsentation der Resultate generiert werden. Neben etablierten Algorithmen die beispielsweise als R-Pakete zur Verfügung stehen, wurden auch eigene Methoden zur Datenanalyse in die beiden Programmpakete integriert, wie zum Beispiel verbesserte Verfahren zum multiplen Alignment von Chromatogrammen oder Algorithmen zur Proteinquantifizierung. Die Integration der verschiedenen Datensätze aus Genom-, Transkriptom-, Proteom- und

Metabolomanalysen mit den verschiedenen Software-Komponenten wird durch das **BRIDGE**-Programmpaket vermittelt, das die direkte Verknüpfungen von heterogenen Daten auf der Programm und Benutzerfrontend-Ebene ermöglicht. Für die Projekt- und Benutzerverwaltung steht das allgemeine Projekt-Management-System (GPMS) zur Verfügung. Darüber hinaus können in Kooperation mit der Technologieplattform-Bioinformatik auch neue bioinformatische Analyseverfahren, Algorithmen und Software-Applikationen zur Datenauswertung entwickelt werden.

Neben den beschriebenen Bioinformatik-Angeboten können Kooperationspartner auch Unterstützung bei experimentellen Arbeiten auf dem Gebiet Transkriptomik erhalten. Dies betrifft die Herstellung genomweiter Microarrays, die Organismus-spezifischen methodischen Anpassungen zur Herstellung und Markierung zellulärer RNA sowie deren Einsatz in Hybridisierungsexperimenten. Alternativ hierzu kann auf dem Sektor der experimentellen Transkriptomik auch die ultraschnelle Sequenzierung von cDNAs mittels der 454-Hochdurchsatztechnologie angeboten werden.

Prof. Dr. A. Pühler

Universität Bielefeld
CeBiTec
D-33594 Bielefeld
Tel.: 0521-106-8750
E-Mail: Puehler@Genetik.Uni-Bielefeld.DE

Dr. W. Selbitschka

Universität Bielefeld
CeBiTec
D-33594 Bielefeld
Tel.: 0521-106-8706
E-Mail: Werner.Selbitschka@
Genetik.Uni-Bielefeld.DE

III. Bereich „Proteomik“, Greifswald

Zusammenfassung

Mit der Vorlage der Genomsequenz ist lediglich der „Bauplan des Lebens“, nicht das Leben selbst beschrieben. Jetzt ist die funktionelle Genomforschung gefragt, wenn es gilt, das „virtuelle Leben der Gene in das reale Leben der Proteine“ umzuschreiben. Dafür wird insbesondere die Proteomanalyse benötigt, da sie sich wie keine andere Disziplin der funktionellen Genomforschung mit den eigentlichen „Spielern des Lebens“, den Proteinen beschäftigt. Die Methoden der Proteomanalyse entwickeln sich international mit einem rasanten Tempo: Verfahren der Hochdurchsatzproteomanalyse über Neuentwicklungen auf dem Gebiet der Gelevaluierung bis zur gelfreien Proteomics, verbunden mit Entwicklungen der Massenspektrometrie, sind hier zu nennen, die nicht nur höchst kostenaufwendig sind, sondern auch ein eingespieltes Team von Spezialisten erfordern. Heute ist es daher nicht mehr möglich, das breite Spektrum von Methoden der Proteomanalyse in einer einzelnen Arbeitsgruppe zu etablieren. Zentrenbildung heißt der internationale Trend, wodurch der Zugang der Arbeitsgruppen zu den Techniken der Proteomics auch in Zukunft kostengünstig garantiert wird. Im Rahmen der BMBF-Förderinitiative GenoMik-Plus ist an der Universität Greifswald eine Technologieplattform Proteomics eingerichtet worden. Diese bietet Kooperationsvorhaben auf den Gebieten der Proteomanalyse an.

Leistungsbeschreibung

In Greifswald sind die Methoden für Hochdurchsatz-Proteomics, der Massenspektrometrie sowie der gelfreien Proteomanalyse etabliert, zumal die AG auf eine mehr als 20-jährige Erfahrung in der mikrobiellen Proteomanalyse zurückblicken kann. Bei Vorlage der Genomsequenz sind Proteinidentifikationen nach 2D-gelelektrophoretischer Auftrennung der Proteine mit Hilfe der Hochdurchsatz-Proteomanalyse sehr effektiv möglich. Diese Studien schließen

vergleichende Proteomanalysen, Einsatz von Gelelektrophorese sowie die bioinformatische Auswertung von Proteomdaten ein. Die 2D-Gelelektrophorese kann in Greifswald durchgeführt werden, daneben können auch Proteine aus Gelen identifiziert werden, die im Partnerlabor hergestellt wurden.

In Kooperationen sind auch weitere, mehr in die Tiefe gehende Analysen zu vereinbaren, die aber bedeutend mehr Zeit in Anspruch nehmen:

- vergleichende Analysen der Proteinmenge bzw. Proteinsyntheserate
- vergleichende Proteom- und Transcriptomanalysen
- Analyse von Sub-Proteomfunktionen (z. B. Sekretome)
- gelfreie Proteom-Analysen (z. B. Membranproteine)
- posttranslationale Modifikationen oder Schädigungen von Proteinen (Phosphoproteom, Disulfidproteom u. dgl.)
- Analyse der Proteinstabilität (auf Ebene des Gesamtproteoms).

Darüber hinaus können auch quantitative Proteomdaten nach gelfreier Analyse erhoben werden.

Für die Kooperationen stehen Spezialisten sowie moderne Proteomanalysetechniken zur Verfügung, u. a. verschiedene Scanner, die Integrated Spot Handling Platform von Amersham Biosciences, die Proteom Work Station mit Scanner, Spot Cutter und Mass-Prep-Station von Biorad/Perkin Elmer, mehrere Massenspektrometer (Voyager DE-STR MALDI, Ettan-MALDI, QStar TM und MDLC, Applied Biosystems 4700 Proteomic Analyser, 4800 Proteomic Analyser), LTQ-FTICR, LTQ-Orbitrap XL) sowie verschiedene Systeme zur Peptidtrennung durch Nano-HPLC.

Prof. Dr. Michael Hecker
E.-M.-Arndt-Universität Greifswald
Institut für Mikrobiologie
F.-L.-Jahn-Straße 15
17489 Greifswald
Tel.: 03834-864200
E-Mail: hecker@uni-greifswald.de