

Molekularbiologische Datenbanken

Übungen
Aufgabe 2



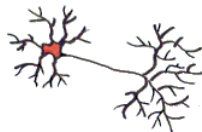
Silke Trißl
Ulf Leser



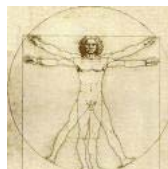
Wissensmanagement in der
Bioinformatik

Microarray oder Expressionsanalyse

- Gleiche Erbinformation in der Zelle (Genom), aber
 - viele unterschiedliche Zellen beim Menschen
 - Nervenzellen, Muskelzellen, Epithelzellen, ...

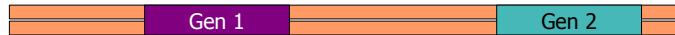


- viele Entwicklungsstadien
 - Embryo ... erwachsenen Menschen

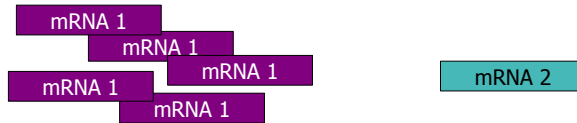


mRNA Gehalt

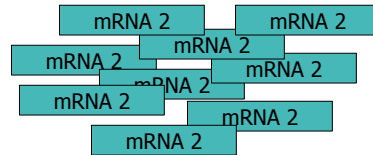
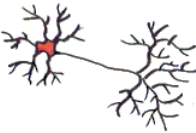
genomische
DNA



Muskelzellen



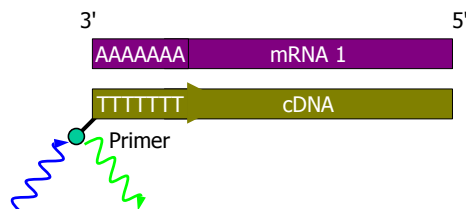
Nervenzellen



Silke Trißl, Ulf Leser: Molekularbiologische Datenbanken, SoSe 2004

Konzentration an mRNA

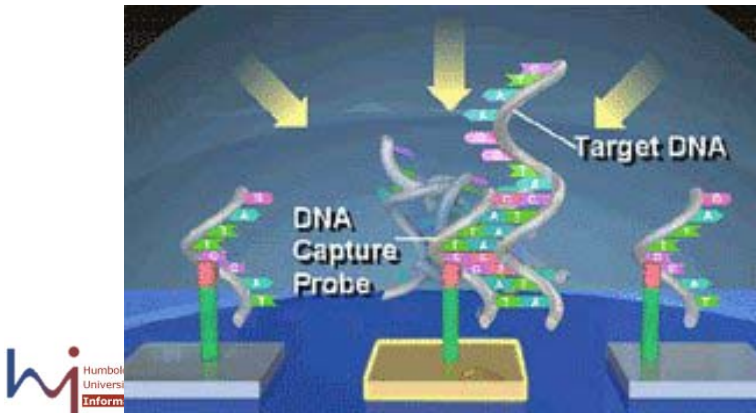
- cDNA – complementary DNA
 - chemisch stabiler als RNA
- hergestellt durch PCR (Polymerase chain reaction)
 - mRNA als Vorlage
 - poly-A als Startsequenz



Silke Trißl, Ulf Leser: Molekularbiologische Datenbanken, SoSe 2004

Microarray

- cDNA wird auf vorbereiteten Microarray aufgebracht
- 'erkennt' dort das passende Gegenstück
 - DNA Abschnitte hybridisieren



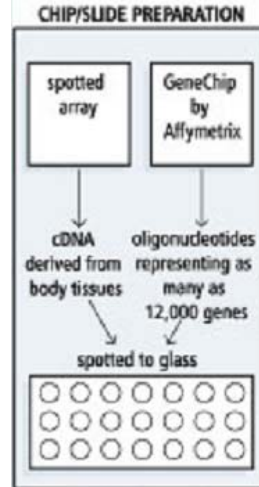
Was ist ein Microarray chip?

- Glasträger
 - mit Spots (Punkten) von DNA
 - bis zu 30 000 Spots auf einem chip in einer Matrix
- Spot (= probe)
 - besteht aus mehreren 'DNA capture probes' mit der gleichen Sequenz
 - 'erkennt' die cDNA eines Gens

⇒ Erfassung des mRNA Gehalts einer Zelle zu einer bestimmten Zeit

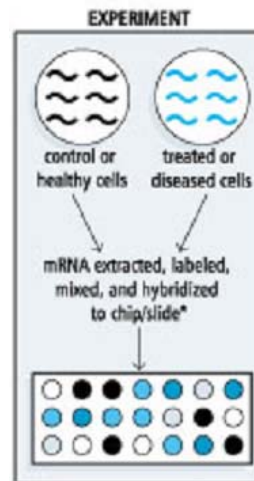
Chip Vorbereitung

- Chip Vorbereitung
 - DNA herstellen
 - DNA annotieren
 - GenBank
 - Gene Ontology (GO)
 - SwissProt
 - auf Glasträger aufbringen



Probenaufbereitung

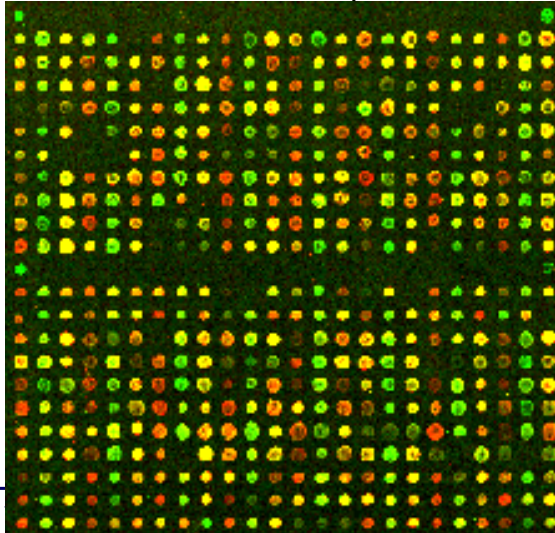
- cDNA Gewinnung
 - aus einer Kontrolle und den interessanten Zellen
 - Protokolle zur
 - Isolierung der mRNA
 - umschreiben in cDNA
 - markieren der cDNA
- Experiment
 - cDNA auf dem vorbereiteten Chip auftragen und hybridisieren



Microarray Analyse

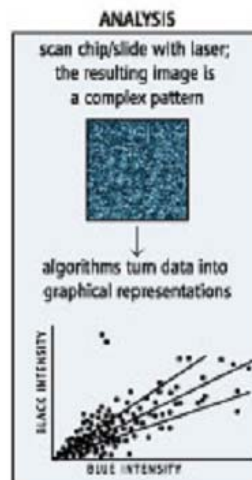
- Detektion durch Fluoreszenz-Analyse

rot: Sample
grün: Kontrolle
gelb: beide gleich



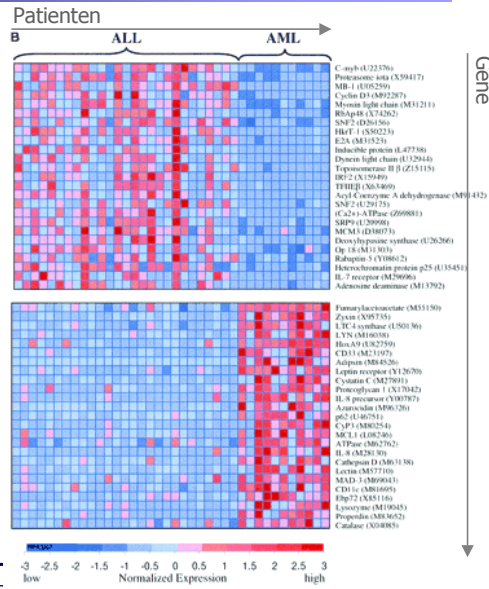
Microarray Analyse

- Analyse
 - Speichern von Werten für jeden einzelnen Punkt
- Viele verschiedene Samples
 - viele Patienten
 - Zeit-serien
- Auswertung
 - Clustern von Genen, die sich ähnlich verhalten
 - Clustern von Samples, die gleiches Verhalten zeigen



Microarray – Anwendung

- ALL
 - akute lymphatische **Leukämie**
- AML
 - akute myeloblastische **Leukämie**
- unterschiedliche
 - Erkrankungen
 - Therapien
- gleiche klinische Diagnose
 - schwer zu unterscheiden



Silke Trißl, Ulf Leser: Molekularbiologische Datenbanken, SoSe 2004

Anwendung

- Klinische Diagnose und Therapie (Erprobung)
 - Unterschiede zwischen Krebsarten
 - molekulare Ursache der Krankheit
 - mögliche Therapieansätze
- Systembiologie
 - Erstellen von Zeitverläufen
 - Reproduktionszyklus der Hefe
 - Reaktion auf Veränderungen der Umgebung
 - Ändern des Substrats (Nährstoffzufuhr)
 - Gärprozess der Hefe bei Sauerstoffmangel



Silke Trißl, Ulf Leser: Molekularbiologische Datenbanken, SoSe 2004

Aufgabe 2

- EU-Projekt, um Daten von Microarray-Experimenten zentral zu speichern
 - mehrere Labors nehmen teil
 - für jedes Labor gibt es eine oder mehrere Personen, die Daten übermitteln können
- Unsere Aufgabe (in Gruppen von 2 – 3 Studenten)
- Modellierung einer Datenbank, um
 - Herstellung bzw. den Einkauf des Chips zu dokumentieren
 - Verarbeitung der Zellen zu cDNA zu speichern
 - Ergebnisse zu archivieren

Daten

- Microarray Chip
 - Einkauf von verschiedenen Firmen
 - Art, Bezeichnung
 - welche Gene (probes) sind enthalten
- Samples
 - Art des Samples (Zelltyp, Organismus, ...)
 - Verschiedene Protokolle zur Verarbeitung von Zelle zur cDNA

Daten (cont.)

- Experiment
 - Aufbringen von einem behandelten Sample auf einen Array
 - dafür werden Protokolle benötigt
- Analyse
 - verschiedene Geräte und Protokolle zum Scannen und Bearbeiten der Daten
- Ergebnisse speichern
 - Werte, die sich für jeden Punkt ergeben

Literatur

- MIAME – Minimal information about microarray experiment
 - <http://www.mged.org/Workgroups/MIAME/miame.html>
- Array express
 - <http://www.ebi.ac.uk/arrayexpress>
- Gene Expression Omnibus – GEO
 - <http://www.ncbi.nlm.nih.gov/geo>

Anforderungen des Projekts

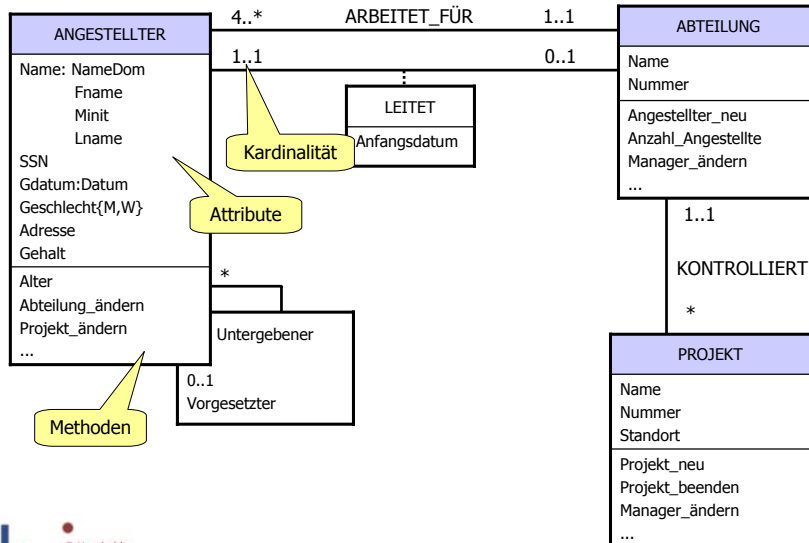
- Microarray – Datenbank
- Erste Planungsphase
 - Konzeptuelle Objektmodellierung in UML
 - Relationales Modell
 - Document Type Definition (DTD)

Kurze Wiederholung anhand eines Beispiels

- Aufgabenstellung
 - Eine Firma ist in **Abteilungen** organisiert. Jede Abteilung hat einen **Angestellten**, der die Abteilung **leitet**.
 - Die Abteilung kontrolliert eine Reihe von **Projekten**.
 - Ein **Angestellter** wird einer Abteilung zugewiesen und **kann einen Vorgesetzten haben**.

aus Grundlagen von Datenbanksystemen R Elmasri, S Navathe, Addison-Wesley

Konzeptuelles UML-Schema



Relationales Modell

- **Angestellter** (Vname, Initial, NName, SSN, Datum, Adresse, Geschlecht, Gehalt, orgSSN, Abt)
- **Abteilung** (AbtName, AbtNummer, ManagerSSN, ManagerAnfangsdatum)
- **Projekt** (PName, PNummer, PStandort, AbtNr)

Document Type Definition – 1

```
<!ELEMENT Abteilung (AbtName, AbtNummer, ManagerSSN,  
  ManagerAnfangsdatum, Angestellter+, Projekt+)>  
<!ELEMENT AbtName (#PCDATA)>  
<!ELEMENT AbtNummer (#PCDATA)>  
<!ELEMENT ManagerSSN (#PCDATA)>  
<!ELEMENT ManagerAnfangsdatum (#PCDATA)>  
  
<!ELEMENT Projekt (PName, PNummer, PStandort)>  
<!ELEMENT PName (#PCDATA)>  
<!ELEMENT PNummer (#PCDATA)>  
<!ELEMENT PStandort (#PCDATA)>
```

Document Type Definition – 2

```
<!ELEMENT Angestellter (Vname, Initial, NName, SSN, GDatum, Adresse,  
  Geschlecht, Gehalt, VorgSSN?)>  
<!ELEMENT VName (#PCDATA)>  
<!ELEMENT Initial (#PCDATA)>  
<!ELEMENT NName (#PCDATA)>  
<!ELEMENT SSN (#PCDATA)>  
<!ELEMENT GDatum (#PCDATA)>  
<!ELEMENT Adresse (#PCDATA)>  
<!ELEMENT Geschlecht (#PCDATA)>  
<!ELEMENT Gehalt (#PCDATA)>  
<!ELEMENT VorgSSN (#PCDATA)>
```

Fragen ?

- Aufgabe 2 über Goya oder auf Web-page
- eine Lösung je Gruppe
- Microarray – Folien auf Web-page
- Lösung bis 11.05., 17 Uhr per e-mail oder in RUD25, IV.104
- Präsentation der Ergebnisse am 12.05. jeder Gruppe

