

Ausprägungsfach Bioinformatik im Rahmen des Bachelor-Studiengangs Informatik

Heiko A. Schmidt, Arndt von Haeseler

Center for Integrative Bioinformatics Vienna (CIBIV)
Max F. Perutz Laboratories (MFPL)
Vienna, Austria
<http://www.cibiv.at>

Notes

Heiko A. Schmidt, Arndt von Haeseler Überblick Bioinformatik

CIBIV – Center for Integrative Bioinformatics Vienna

- Das CIBIV ist ein gemeinschaftliches Forschungsinstitut von
 - Universität Wien
 - Medizinischer Universität Wien
- Lehr-Veranstaltungen werden vor allem von folgenden Dozenten gehalten:
 - Prof. Dr. Arndt von Haeseler
 - Dr. Heiko Schmidt
 - aber auch von anderen Postdocs des CIBIV
- Am CIBIV arbeiten
 - 7 Postdocs
 - 6 Doktoranden
 - einige Studenten

Notes

Heiko A. Schmidt, Arndt von Haeseler Überblick Bioinformatik

Bioinformatik – eine Definition

Was ist Bioinformatik?

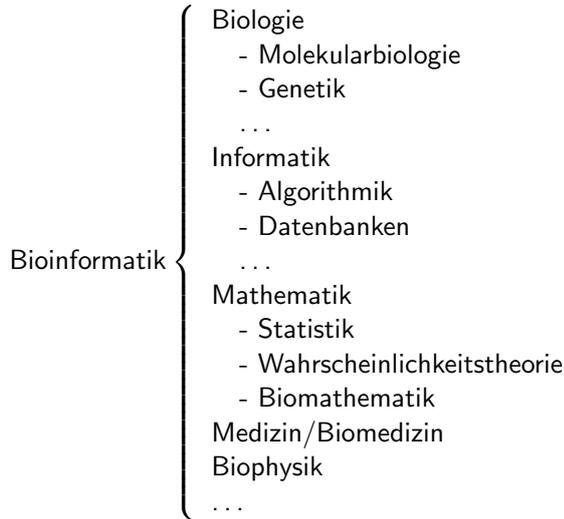
Die Bioinformatik im weitesten Sinne ist die Anwendung von **Informatik-Methoden** und Computern zur Beantwortung **biologischer Fragestellungen**.

(Es gibt viele weitere Definitionen, je nachdem, woher sie stammt.)

Notes

Heiko A. Schmidt, Arndt von Haeseler Überblick Bioinformatik

Bioinformatik ist ein interdisziplinäres Feld



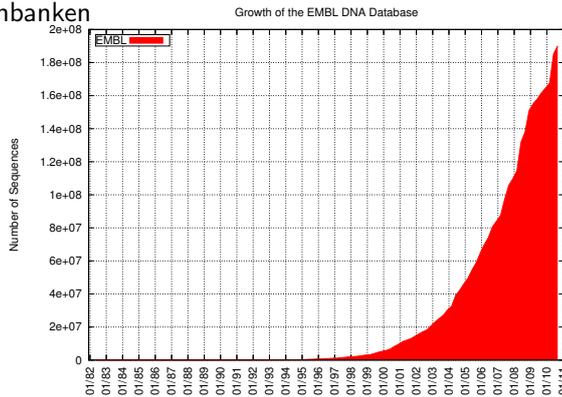
Notes

Heiko A. Schmidt, Arndt von Haeseler Überblick Bioinformatik

BI-Felder – Datenspeicherung/-aufbereitung/-integration

Biologische Datenbanken

- DNA-Datenbanken (EMBL, Genbank)
- Protein-Datenbanken (SwissProt, Uniprot)
- Struktur-Datenbanken (PDB)
- Literature-Datenbanken (PubMed)
- ... viele Spezialdatenbanken



Notes

Heiko A. Schmidt, Arndt von Haeseler Überblick Bioinformatik

BI-Felder – Sequenzvergleiche/Sequenzanalyse

- Muster-Suche (Wort gegen Sequenz)
- Datenbank-Suche (Sequenz gegen Datenbank)
- Sequenzvergleiche (Sequenz gegen Sequenz)
- Sequenzalignment (Ausrichten von Sequenzen gegeneinander, Grundlage für die meisten weiterführenden Analysen)

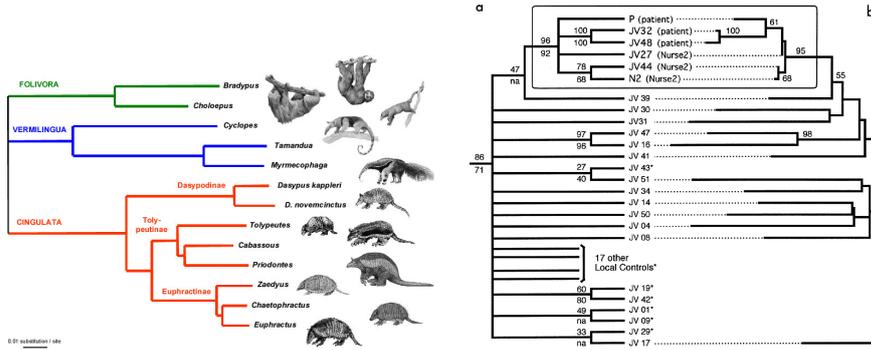
```
GAGGGCATGGCCGTCCAGCTCGTGTGTGAAATG
GAGGGCATGGCCGCCGTCCAGCTCGTACCGTGAGGCG
GAGGGCATGGCCGCCGTCCAGCTCGTACCGTGAGGCG
GTGGGCATGGCCGTCCAGCTCGTACCGTGAGGCG
GTGGGCATGGCCGTCCAGCTCGTACCGTGAGGCG
GAGGGCATGGCCGCCGTCCAGCTCGTACCGTGAGGCG
GAGGGCATGGCCGTCCAGCTCGTGTGTGAAATG
```

Notes

Heiko A. Schmidt, Arndt von Haeseler Überblick Bioinformatik

BI-Felder – Evolutionäre Bioinformatik/Phylogenetik

Rekonstruktion der Verwandtschafts- und Abstammungsverhältnisse



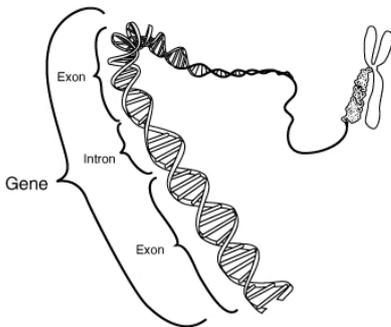
Grundlage für das Verständnis der meisten biologischen Vorgänge, da diese vor dem Hintergrund der Evolution entstanden sind.

Notes

BI-Felder – Genomik

Inzwischen können ganze Genome effizient sequenziert werden. Mit deren Sequenz allein wissen wir aber noch nichts konkretes. Was muss getan werden?

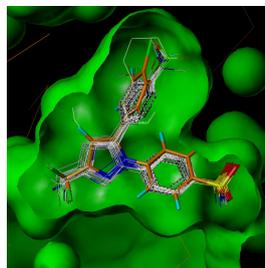
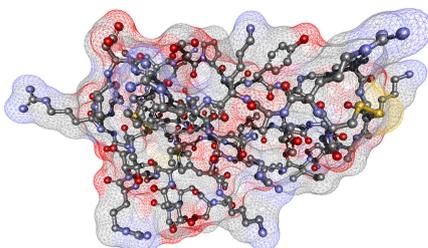
- **Genvorhersage** (Wo fängt ein Gen an/hört es auf?)
- Vorhersage **regulatorischer Elemente** (Die Steuerelemente der Genexpression)
- **Vergleich** mit andere Genomen (Wo sind Gemeinsamkeiten/Unterschiede)
- **Annotation** (Zuordnung von Funktion und Struktur zu den jeweiligen Bereichen des Genoms)
- . . .



Notes

BI-Felder – Strukturvorhersage

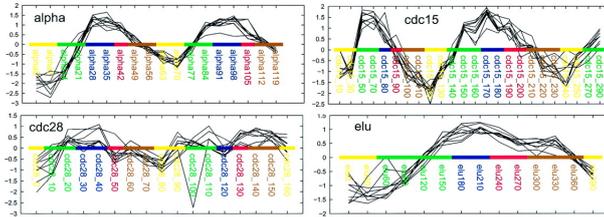
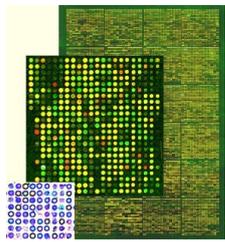
- RNA/DNA-Strukturen und Funktionseinschätzung
- Proteinstrukturen und Funktionseinschätzung
- Interaktion zwischen Molekülen
- Liganden-Docking (Wirkstoff-/Hemmstoff-Entwicklung in der Pharmaindustrie)



Notes

BI-Felder – Expressionsdatenanalyse

- Mit **Microarrays** läßt sich messen, wie häufig Gene *abgelesen* werden.
- Durch die Änderung dieser **Expressionsdaten** auf äußere Einflüsse, lassen sich Gene charakterisieren
- und evtl. Gruppen gemeinsam agierender Gene finden.

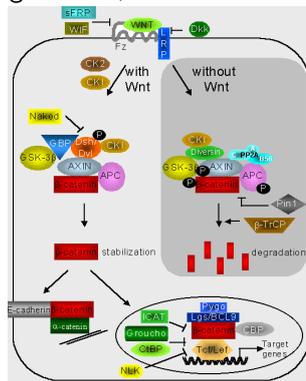
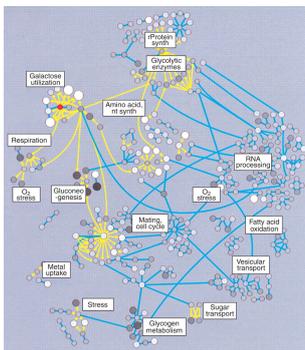


Notes

BI-Felder – Systems Biology

Systems Biology beschäftigt sich mit der Rekonstruktion und Modellierung komplexer Systems, wie z.B.

- Gen-regulatorischer Netzwerke
- Interaktions-Netzwerke
- Metabolischer Pfade
- oder gar ganzer Zellen, Geweben, Organismen, ...



Notes

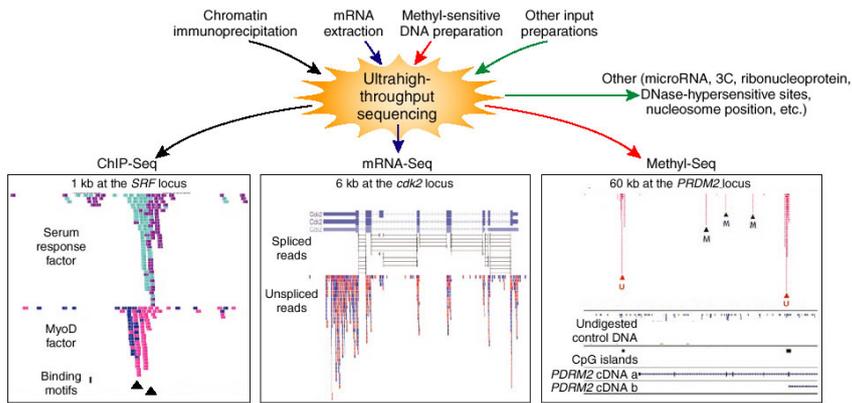
Classical vs. Next Generation Sequencing



- | | |
|---|--|
| <ul style="list-style-type: none"> • generiert 96 Reads • mit Längen bis zu 1000bp • gute Qualität | <ul style="list-style-type: none"> • generiert Millionen Reads • mit Längen bis zu 150bp • fehlerhaft |
|---|--|
- NGS kann billig Unmengen kurzer Reads generieren und erfordert damit neue Methodiken, die mit großen Mengen kurzer Sequenzen umgehen koennen.
- Hiermit können dann Fragen gestellt werden, die vorher nicht möglich waren.

Notes

Next Generation Sequencing



Quelle: Wold + Myers (2008) Nat. Meth.

- NGS-Verfahren werden auch immer häufiger in der klinischen Diagnostik verwendet. (Genetische Prädispositionen/Krankheiten, Krebsforschung etc.)
- Sie eröffnen Wege in die personalisierte Medizin.

Notes

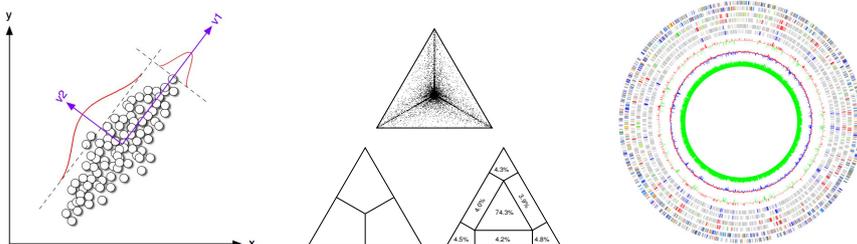
Weitere Felder

- Parallele Bioinformatik (Nutzung von Clouds, Grids, Clustern, Multi-CPU/Core, GPU-Plattformen)
- Populationsgenetik
- Biodiversität
- Biogeografie/Phylogeografie
- Simulationen in verschiedenen Bereichen
- Text-/Datamining
- ...

Notes

Eine wichtige Aufgabe ist immer

- Entwicklung effizienter Analyse Software
- Entwicklung von Analyse- und Visualisierungsmethoden



Notes
