

AF. Bioinformatik im Bachelor-Studiengang Informatik

Heiko A. Schmidt, Ingo Ebersberger, Arndt von Haeseler

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

April 11, 2011

Notes

Heiko A. Schmidt, Ingo Ebersberger, Arndt von Haeseler Überblick Bioinformatik

CIBIV – Center for Integrative Bioinformatics Vienna

- Das CIBIV ist ein gemeinschaftliches Forschungsinstitut von
 - Universität Wien
 - Medizinischer Universität Wien
 - Veterinärmedizinische Universität Wien
- Die Veranstaltungen werden vor allem von folgenden Dozenten gehalten:
 - Prof. Dr. Arndt von Haeseler
 - Dr. Ingo Ebersberger
 - Dr. Heiko Schmidt
- Weitere Informationen:
 - <http://www.cibiv.at>

Notes

Heiko A. Schmidt, Ingo Ebersberger, Arndt von Haeseler Überblick Bioinformatik

Bioinformatik – eine Definition

Was ist Bioinformatik?

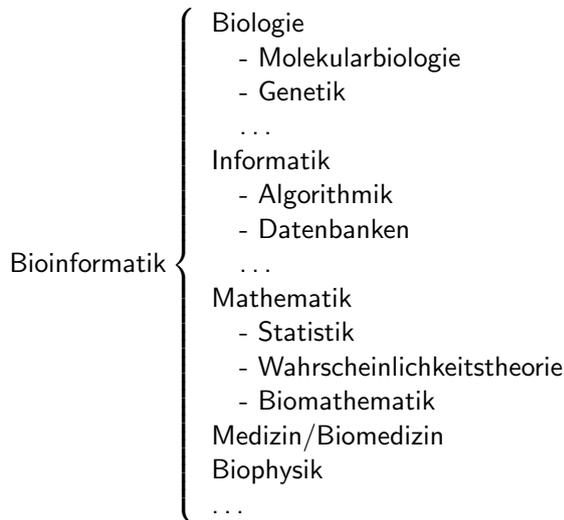
Die Bioinformatik im weitesten Sinne ist die Anwendung von **Informatik-Methoden** und Computern zur Beantwortung **biologischer Fragestellungen**.

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

Notes

Heiko A. Schmidt, Ingo Ebersberger, Arndt von Haeseler Überblick Bioinformatik

Bioinformatik ist ein interdisziplinäres Feld

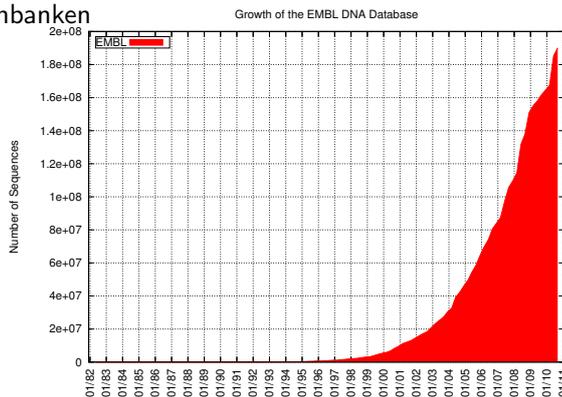


Notes

BI-Felder – Datenspeicherung/-aufbereitung/-integration

Biologische Datenbanken

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



Notes

BI-Felder – Sequenzvergleiche/Sequenzanalyse

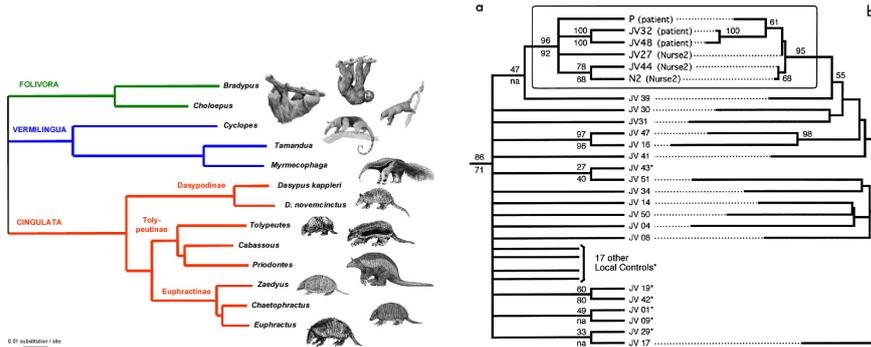
- Muster-Suche
(Wort gegen Sequenz)
- Datenbank-Suche
(Sequenz gegen Datenbank)
- Sequenzvergleiche
(Sequenz gegen Sequenz)
- Sequenzalignment
(Ausrichten von Sequenzen gegeneinander, Grundlage für die meisten weiterführenden Analysen)

```
GAGGGCATGGCCGTCCGAGCTCGTGTGTGAAATG
GAGGGCATGGCCGCGTCAGCTCGTACCGTGAGGCG
GAGGGCATGGCCGCGTCAGCTCGTACCGTGAGGCG
GTGGGCATGGCCGTCGTCAGCTCGTACCGTGAGGCG
GTGGGCATGGCCGTCGTCAGTTCTACTGTGAAGCA
GAGGGCATGGCCGCGTCAGCTCGTACCGTGAGGCG
GAGGGCATGGCCGTC---AGCTCGTGTGTGAAATG
GAGGGCATGGCCGTC---GGCTCGTCCGTGAGGTG
GAGGGCATGGCCGTC---AGCTCGTCCGTGAGGTG
GAGGGCATGGCCGTCGCCAGCTCGTGTGAAATG
GAGGGCATGGCCGTC---AGCTCGTACCGTGAGGCG
GAGGGCATGGCCGCGTCAGCTCGTACCGTGAGGCG
GTGGGCATGGCCGTCGTCAGCTCGTACCGTGAGGCG
GAGGGCATGGCCGTCGTCAGTTCTACTGTGAAGCA
```

Notes

BI-Felder – Evolutionäre Bioinformatik/Phylogenetik

Rekonstruktion der Verwandtschafts- und Abstammungsverhältnisse



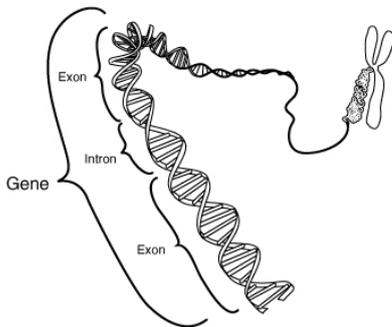
Grundlage für das Verständnis der meisten biologischen Vorgänge, da diese vor dem Hintergrund der Evolution entstanden sind.

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?

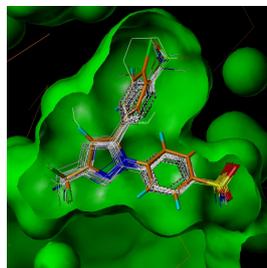
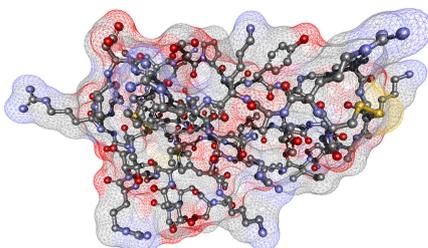
- Genvorhersage (Wo fängt ein Gen an/hört es auf?)
- Vorhersage regulatorischer Elemente (Die Steuerelemente der Genexpression)
- Vergleich mit andere Genomen (Wo sind Gemeinsamkeiten/Unterschiede)
- Annotation (Zuordnung von Funktion und Struktur zu den jeweiligen Bereichen des Genoms)
- . . .



Notes

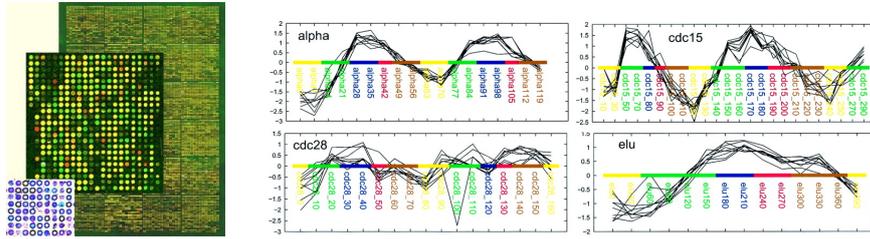
BI-Felder – Strukturvorhersage

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



Notes

- 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.

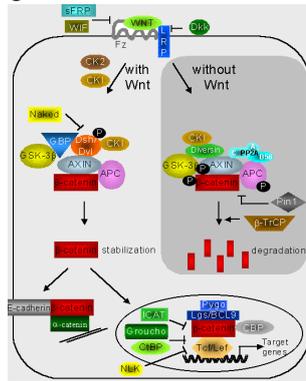
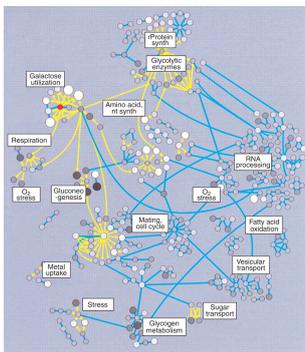


Notes

BI-Felder – Systems Biology

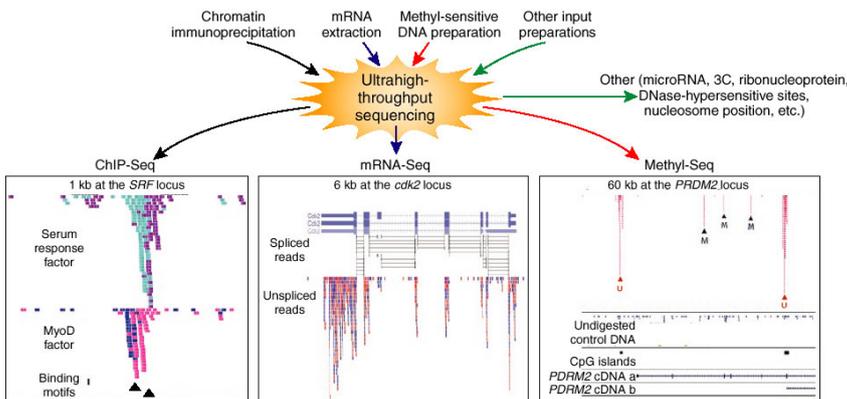
Systems Biology beschäftigt sich mit der Rekonstruktion und Modellierung komplexer Systems, wie z.B.

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



Notes

Next Generation Sequencing



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

- 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.

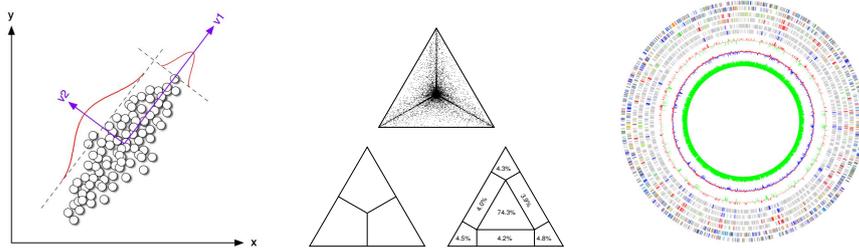
Notes

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

Notes

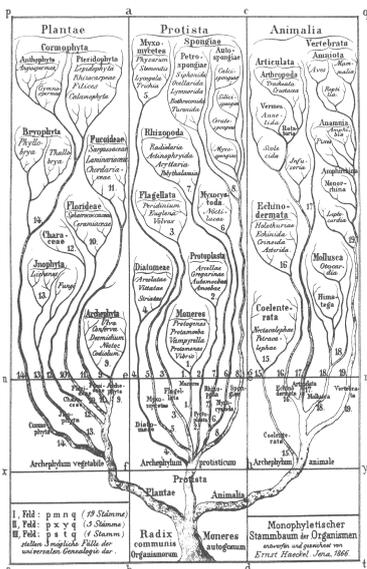
Eine wichtige Aufgabe ist immer

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



Notes

Rekonstruktion der Evolution der Lebewesen



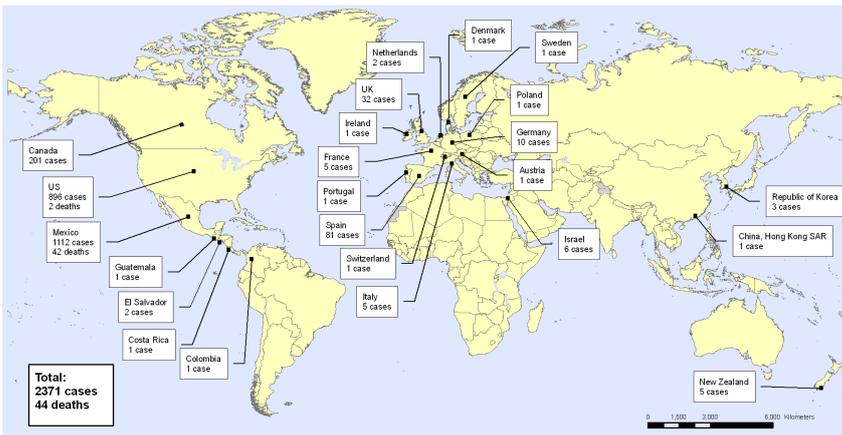
- Die Herkunft und Entstehung der Vielfalt des Lebens war lange schon von hohem Interesse.
- Grundlagen für unser heutiges Verständnis legte **Charles Darwin**, 1859.
- Frühe Entwürfe wie der *Stammbaum der Organismen* von Ernst Haeckel (1866) sind heute mittels molekularer Methoden weitgehend wiederlegt.
- Dennoch ist heute allgemein belegt und akzeptiert, dass die Entstehung dieser Vielfalt auf Vererbung beruht.

Notes

Die H1N1 Schweinegrippe-Pandemie 2009

New Influenza A (H1N1),
Number of laboratory confirmed cases and deaths as reported to WHO

Status as of 7 May 2009
18:00 GMT



The boundaries and names shown and the designations used on this map do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted lines on maps represent approximate border lines for which there may not yet be full agreement.

Data Source: World Health Organization
Map Production: Public Health Information and Geographic Information Systems (GIS)
World Health Organization

World Health Organization
© WHO 2009. All rights reserved

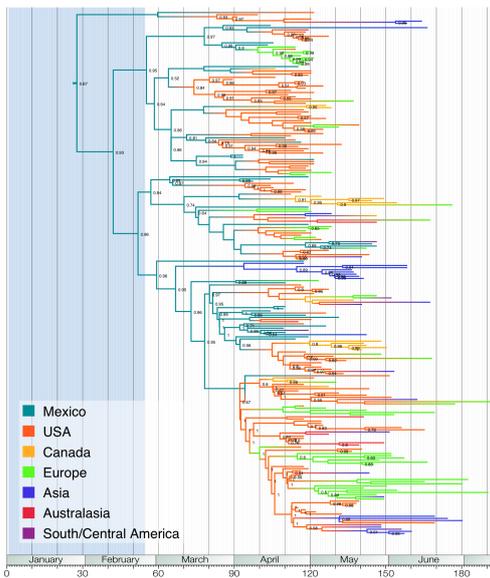
Map produced: 7 May 2009 18:05 GMT

Notes

Heiko A. Schmidt, Ingo Ebersberger, Arndt von Haeseler

Überblick Bioinformatik

H1N1 Stammbaum zu Beginn der Pandemie



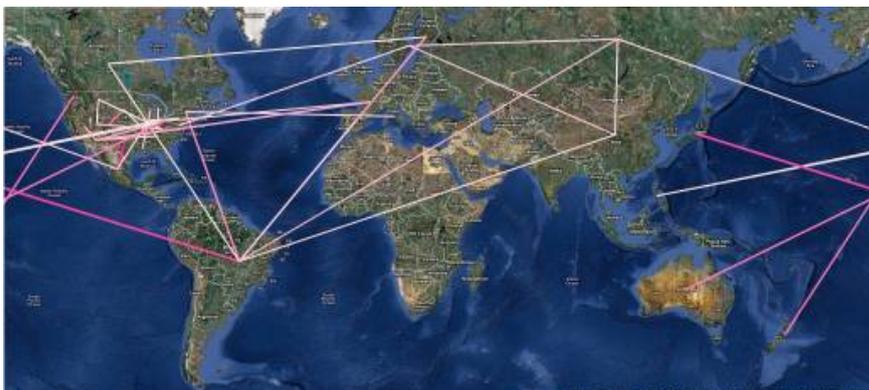
- Aus den H1N1 Genomen sowie Sammelort und -datum lassen sich Stammbäume berechnen, die die Historie der Pandemie widerspiegeln.
- Mit diesem Wissen lassen sich die genauen Verbreitungswege schätzen.
- Quelle: Lemey et al. (2009) PLoS Curr.

Notes

Heiko A. Schmidt, Ingo Ebersberger, Arndt von Haeseler

Überblick Bioinformatik

Ergebnis aus Epidemiologie und Phylogeographie



Notes

Für einen zeitlichen Ablauf:

http://www.youtube.com/watch_popup?v=3_Abd9I402Y&vq=medium#t=27

Heiko A. Schmidt, Ingo Ebersberger, Arndt von Haeseler

Überblick Bioinformatik

Studienplan mit Fernfachkombination Bioinformatik

- Ringvorlesung *VU Einführung in die Bioinformatik* (2 SWS, 3 ECTS, WiSe)
- *VU Grundlagen Bioinformatik* (4 SWS, 6 ECTS, WiSe)
- *VU Angewandte Bioinformatik* (4 SWS, 6 ECTS, WiSe)
- *PS Erweiterte Grundlagen der Bioinformatik* (2 SWS, 3 ECTS, WiSe)
- *VU Statistische Methoden in der Bioinformatik* (2 SWS, 3 ECTS, WiSe)
- *PR Praktikum aus Bioinformatik* (4 SWS, 6 ECTS, SoSe)
- *PR Praktikum Bioinformatik mit Bachelorarbeit* (4 SWS, 18 ECTS, SoSe)

Notes

Notes

Notes
