Translational Bioinformatics in Drug Discovery

Michael Rebhan (Novartis Pharma AG, Basel, Switzerland)

ISMB July 2008, Toronto

BoF: Discussion on best practice for bioinformatics cores
What is best practice? How much does it vary?

- Non-profit institutes with service units
- Big pharma
- Medical research
- Small biotechs
- Agro
- Food

Bioinformatics research in academia
Bioinformatics in Biotech & Big Pharma

- Biomarkers
- Target discovery
- Understanding disease mechanisms
- Compound MoA
- IT infrastructure
- Bioinformatics
- Statistics

All 3 aspects well developed and working together closely!
Translational Bioinformatics

Transformation of genomic (and related biomedical) data into therapeutic (or diagnostic) concepts

- Staff with experience on both sides of the gap!
  - Relevant biology background plus several years of bioinformatics exposure (not afraid of scripting, careful usage of tools, curious!)

- Which datasets are most relevant? Which tools/methods?
  - Careful judgement of most relevant datasets, including public data for comparison

- Multiple lines of in silico evidence → justify costly experiments (e.g. using mouse models, patient samples)

- Less method development, more applied (using existing tools effectively)

- Expectation management within the organization
  - “Buy-in”: key partners, showcases, informal and formal communication, presence at biology discussions, …
Some Factors

- Ability to influence experimental design, early involvement

- Joint projects with wet lab units, finding key partners (benefits for those who are good partners)

  Service projects ↔ True collaborations
  (basic analysis) (extensive analysis)

- Critical mass for the core:
  - Expertise (monitoring of external developments, their relevance)
  - Links into other groups/departments, sufficient focus!
  - Internal tool landscape: critical tools? is the tool mature enough?

- Testable hypotheses
  (incl. some aspects of experimental design, e.g. which primers or probes to use)
Perspective

Don’t worry about:

1. *The amount of data produced by a single researcher will decrease*

2. *The data will become less heterogeneous*

3. *Communication will become easier*

![Diagram showing expectations and realistic zone with years 2000, 2003, 2008, and 2011? marked.]

*expectations*  
(non-bioinformaticians)
Approaching the “realistic zone”

Next gen sequencing
Epigenomics
Exon arrays
HT assays
etc. etc.

How many “in silico biologists” do we need per data-producing researcher in 2011?

Where will all those people come from??

expectations (non-bioinformaticians)

2000
2003
2008
2011?