

Nanoparticle risk assessment

practical solutions



systems toxicology

Mathematical Models for Analyzing Genomic Data Sets:
From Equations to Diagnosis ?

Knut Reinert

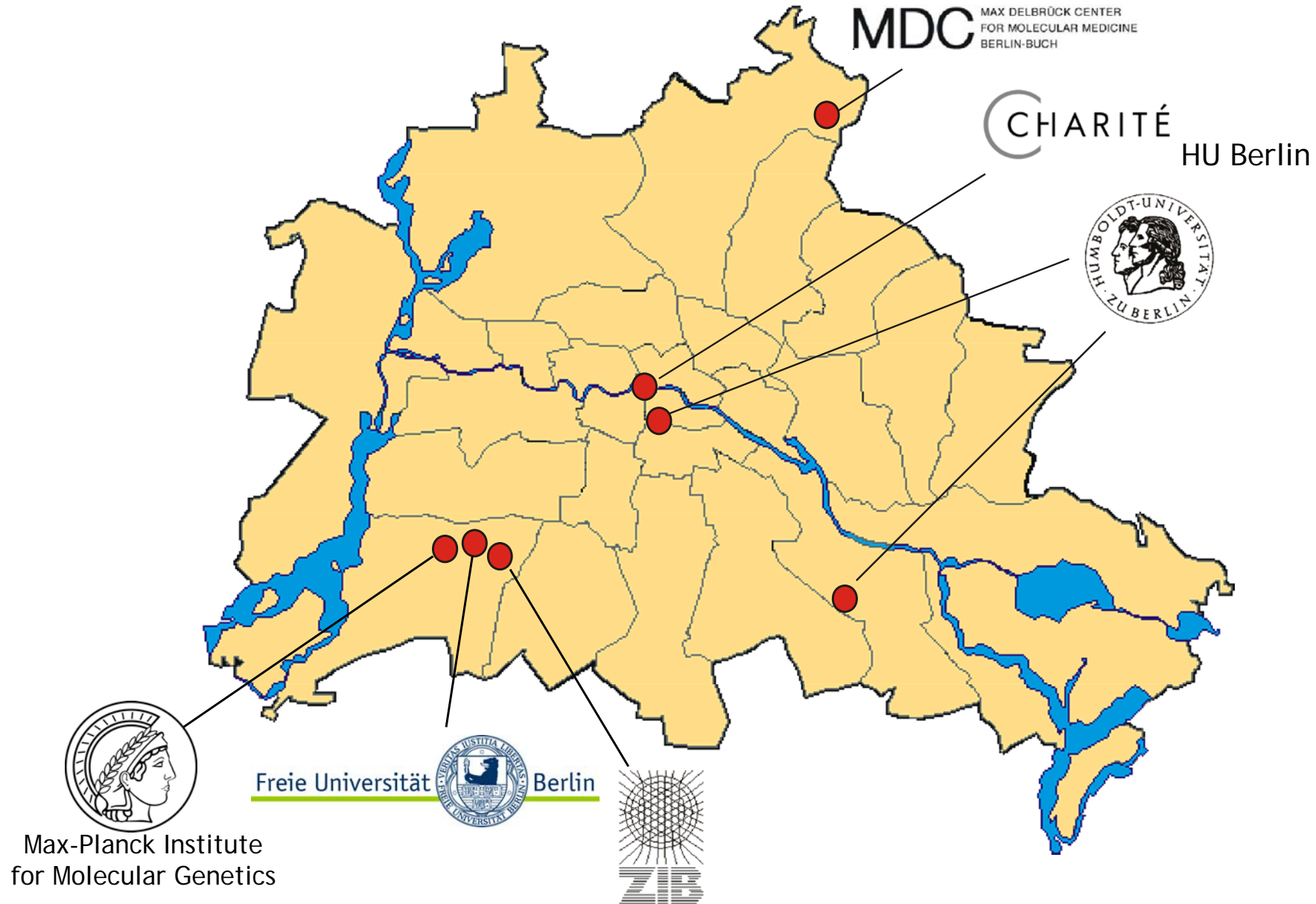
Algorithmische Bioinformatik

FU Berlin, Germany

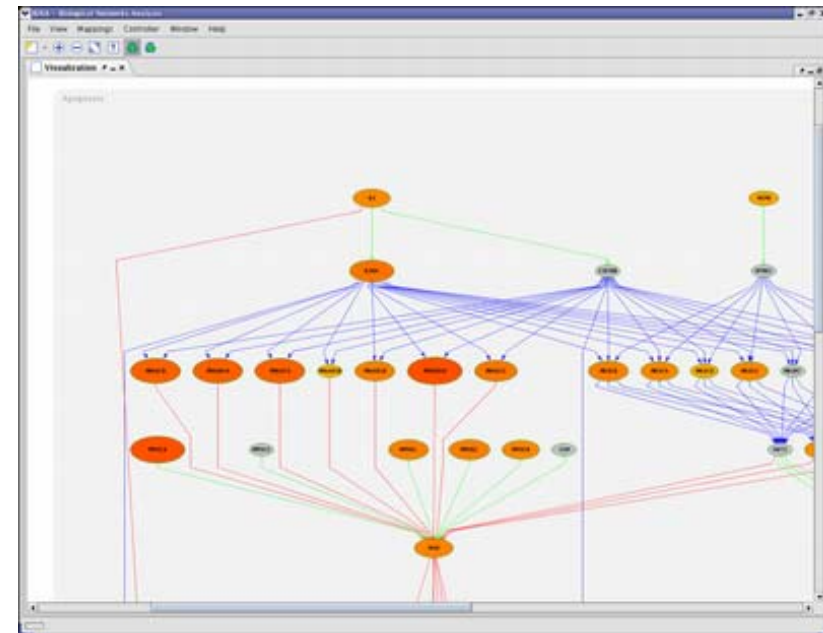
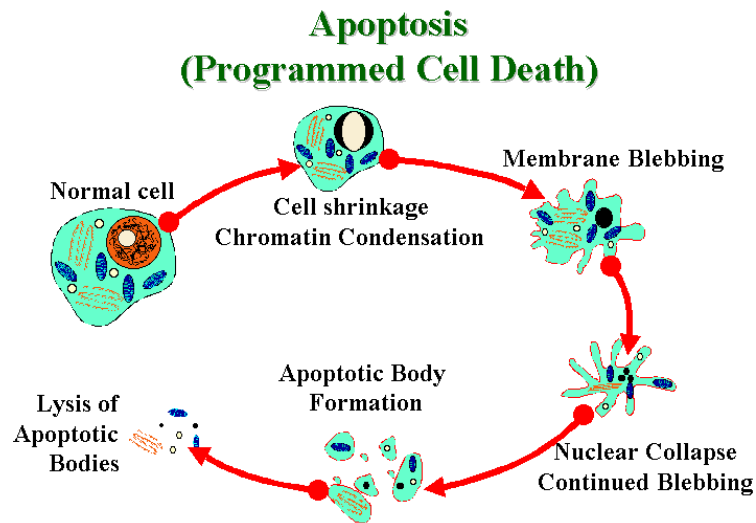
- My background
- Paradigm shift in toxicology testing
- HT -omics-technologies, computational analysis and implications
- Uncertainty and variability
- Conclusion

Bioinformatics in Berlin

Knut Reinert
AAAS, San Diego
February 2010



Model real life problems

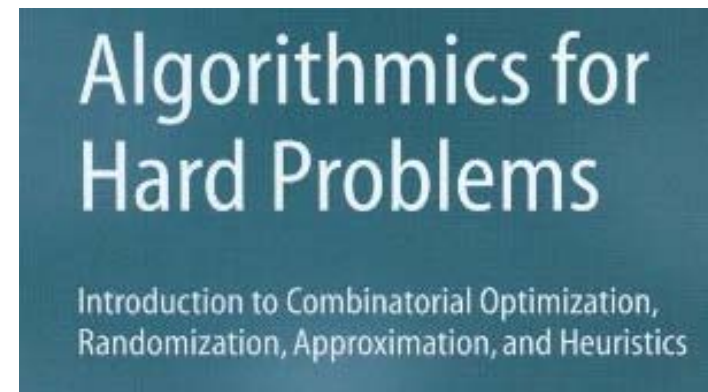
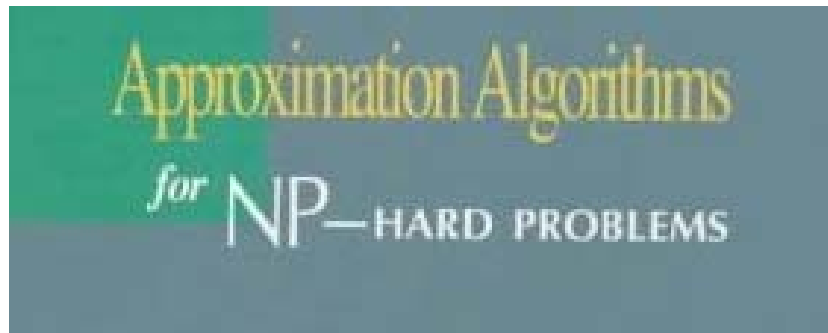


**formalization needed
for computational solution**

Assess hardness of formal problems

NP-hard?

=no efficient algorithm **POSSIBLE!**



**directs research efforts
in searching solutions**

Solve hard problems

Devise exact algorithms
necessarily exponential run time and/or space consumption

Devise approximation algorithms
yields bounds (e.g. guaranteed within 1.5x of optimum)

Devise heuristics
expert knowledge guides computation, **NO** guarantees

In computational biology many
hard problems have **heuristic** solution

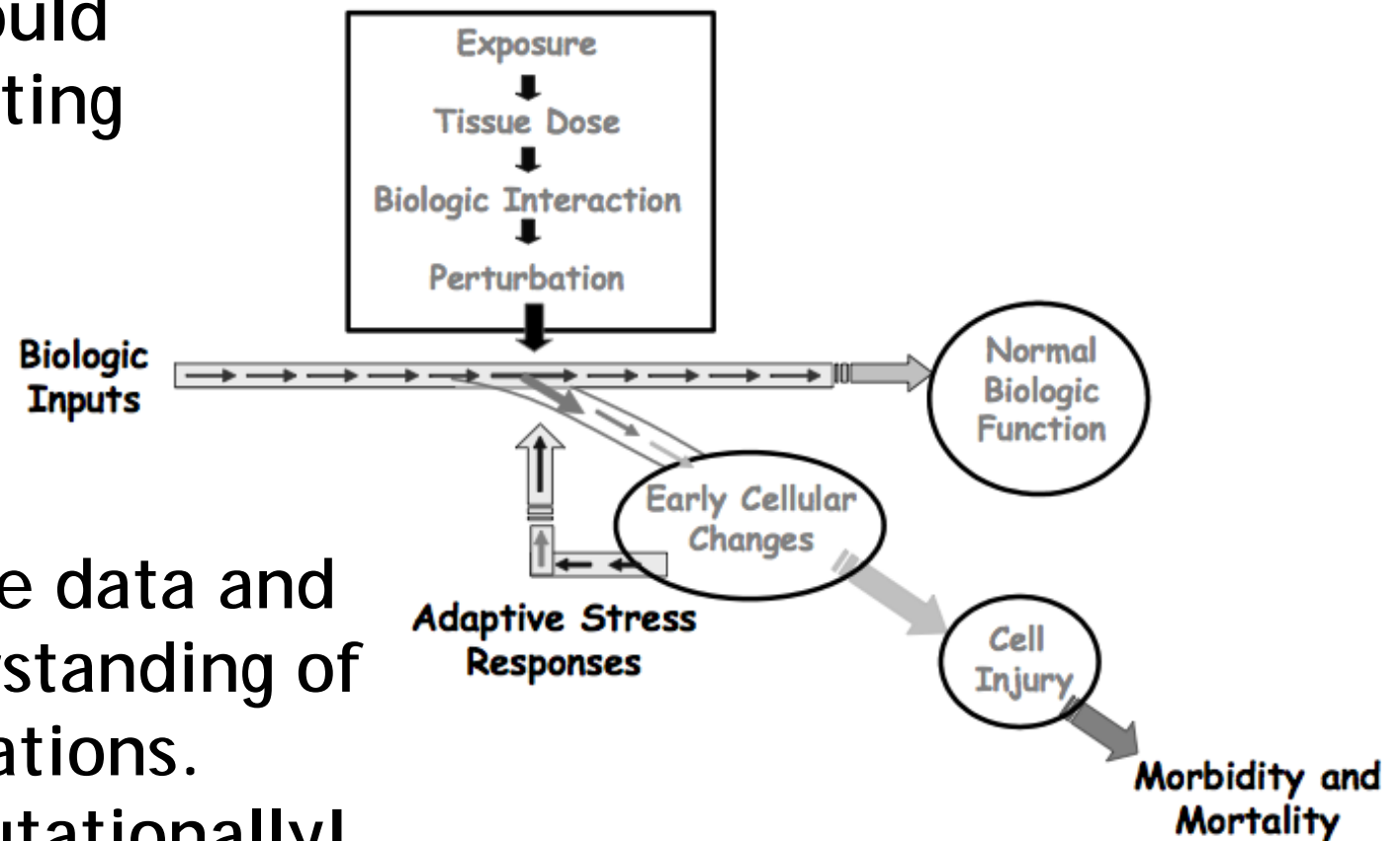
Demands on modern toxicity testing

- Test large number of **new** and existing chemicals, many of which lack basic toxicity data
- Evaluate potential adverse effects for all critical end points and life stages, assess uncertainty and variability
- minimize **animal use**, **time**, and **cost**
- Acquire **mechanistic understanding** and quantitative data to assist decision making

Paradigm change imminent

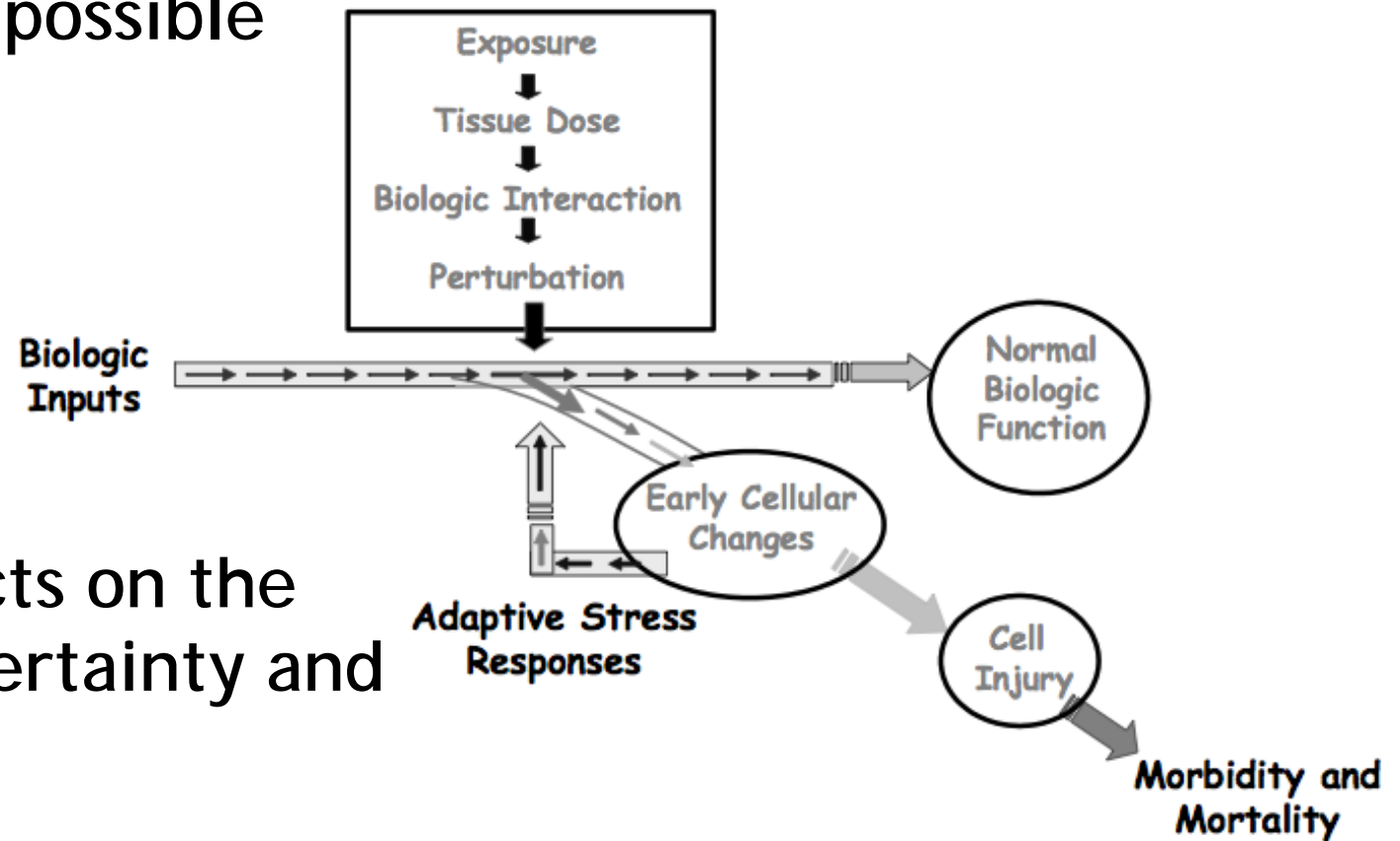
In vitro systems could replace in vivo testing

Goal:
Obtain quantitative data and mechanistic understanding of effects of perturbations.
Extrapolate computationally!



Paradigm change imminent

targeted research possible



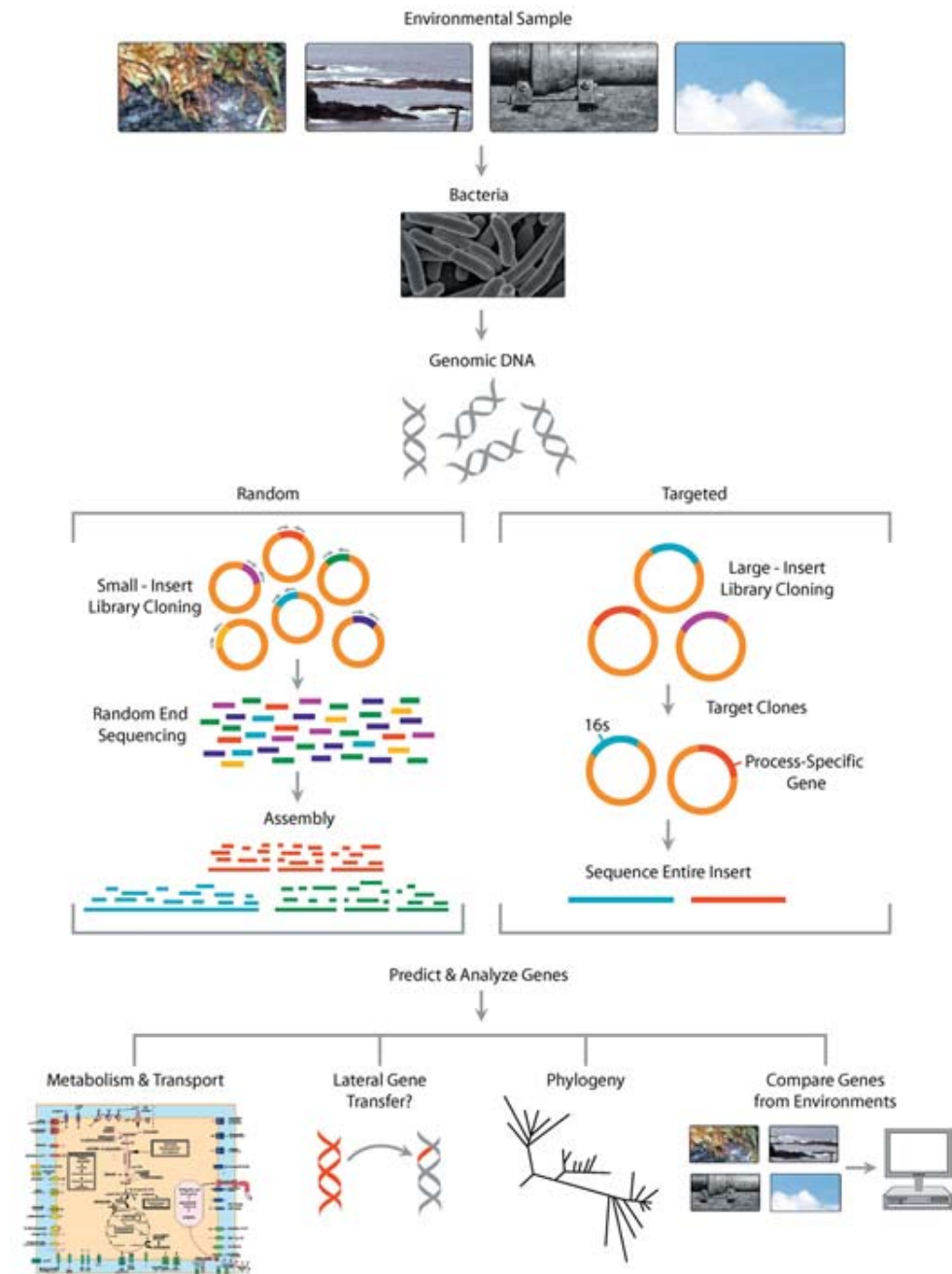
What are the effects on the assessment of uncertainty and variation?

Toxicity testing

Knut Reinert
AAAS, San Diego
February 2010

SYSTEM wide quantitative and qualitative analysis possible (Metagenomics)

What are the effects on the assessment of uncertainty and variation?



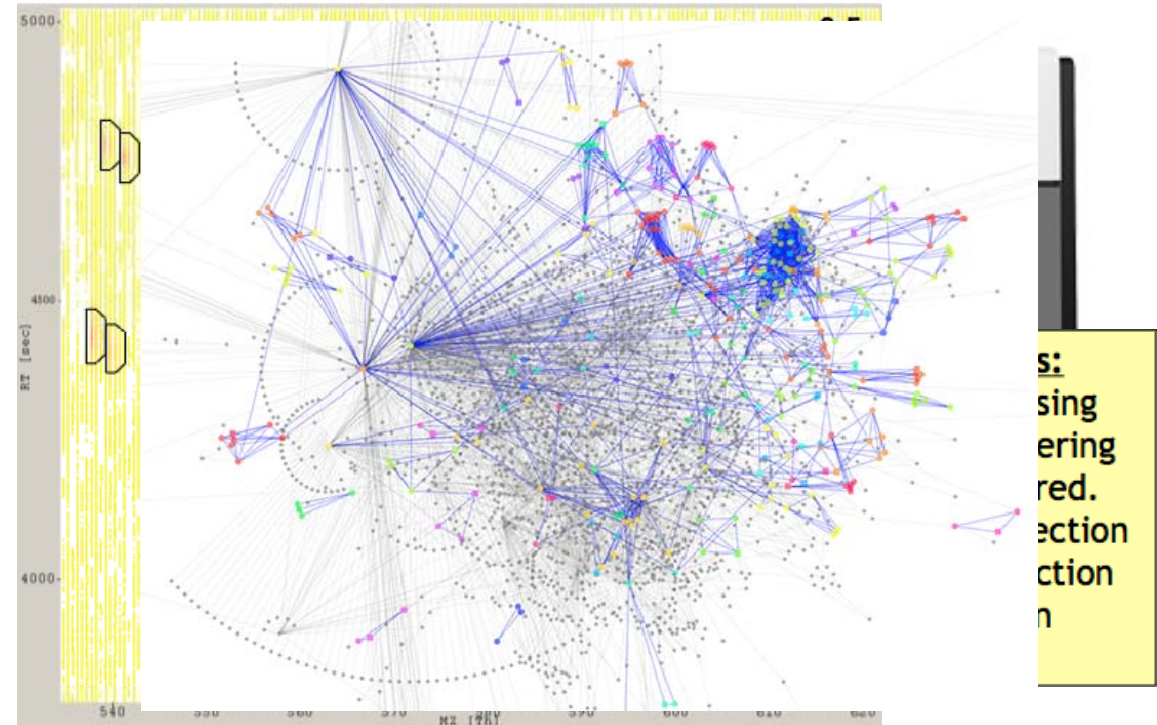
Computational treatment of high-throughput -omics techniques will be the work horse of Mode of action based risk assessment

Genomics

Proteomics

Metabolomics

Networks



More in Ivan Rusyn's talk

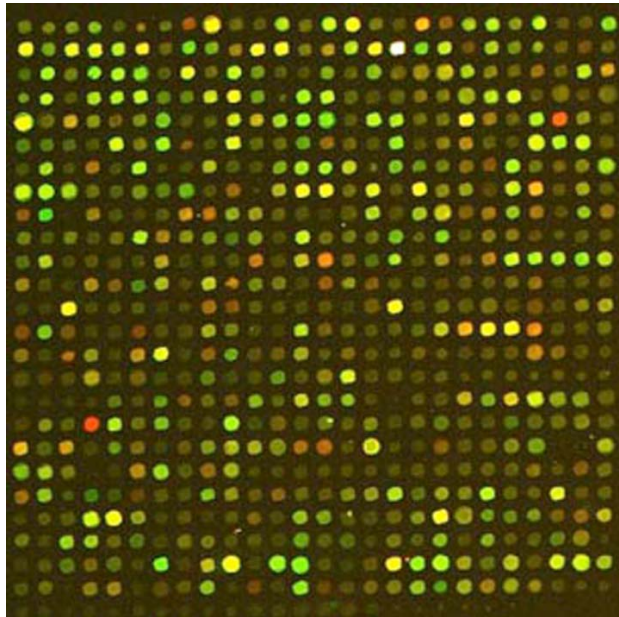
-omics techniques and “Computational uncertainty”

Hypothesis:
Computational treatment of High throughput -omics techniques can be divided in two groups

**Processing of data
yielding error models
(e.g. DNA sequencing reads)**

**Processing of data NOT
yielding error models
(examples to come)**

Microarrays

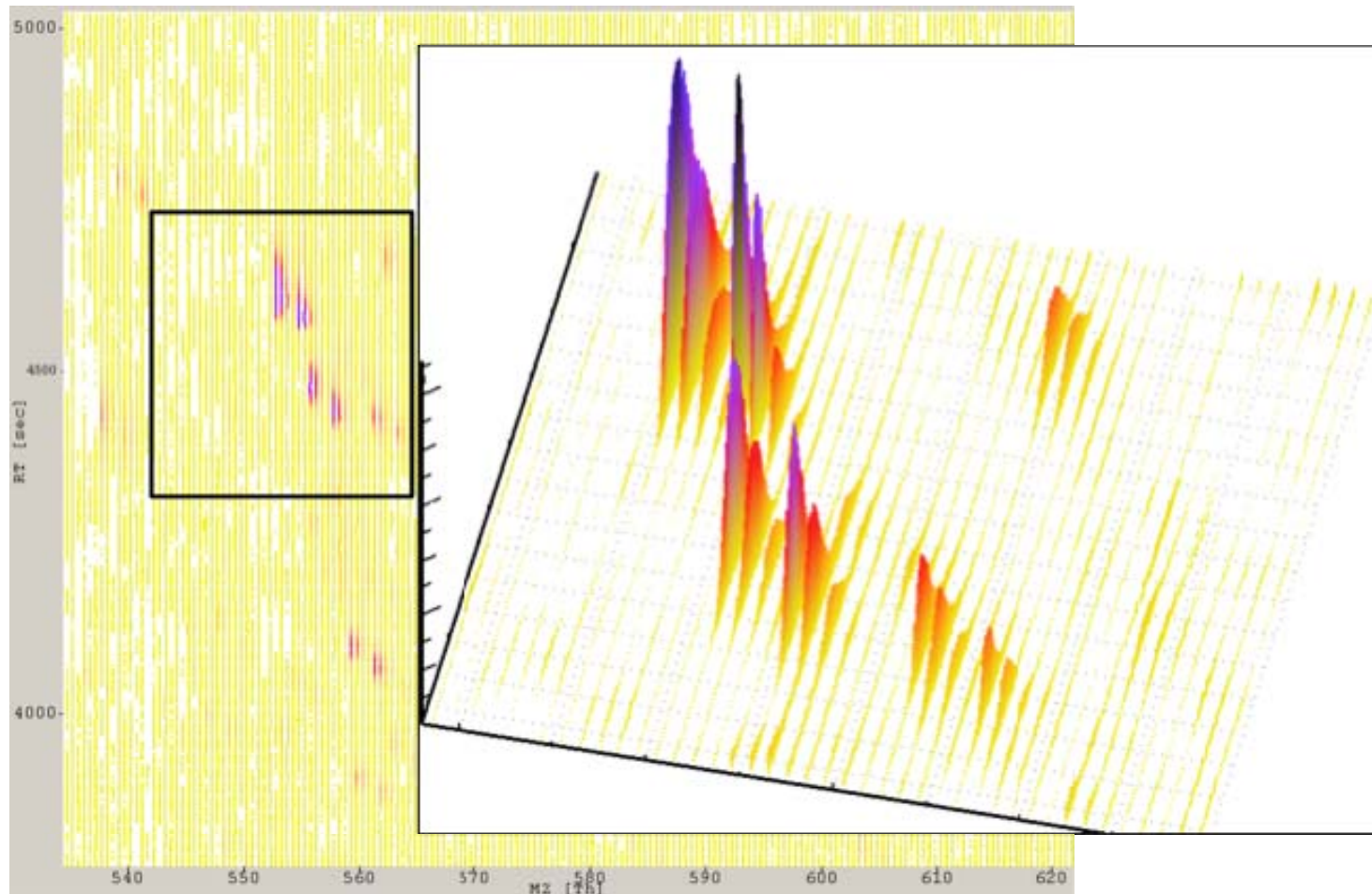


- + relatively stable technology
- + uncertainty and variation can be assessed
- + additional data processing
uncertainty negligible
- arrays can only test for what is spotted on them.

"Computational uncertainty"

Knut Reinert
AAAS, San Diego
February 2010

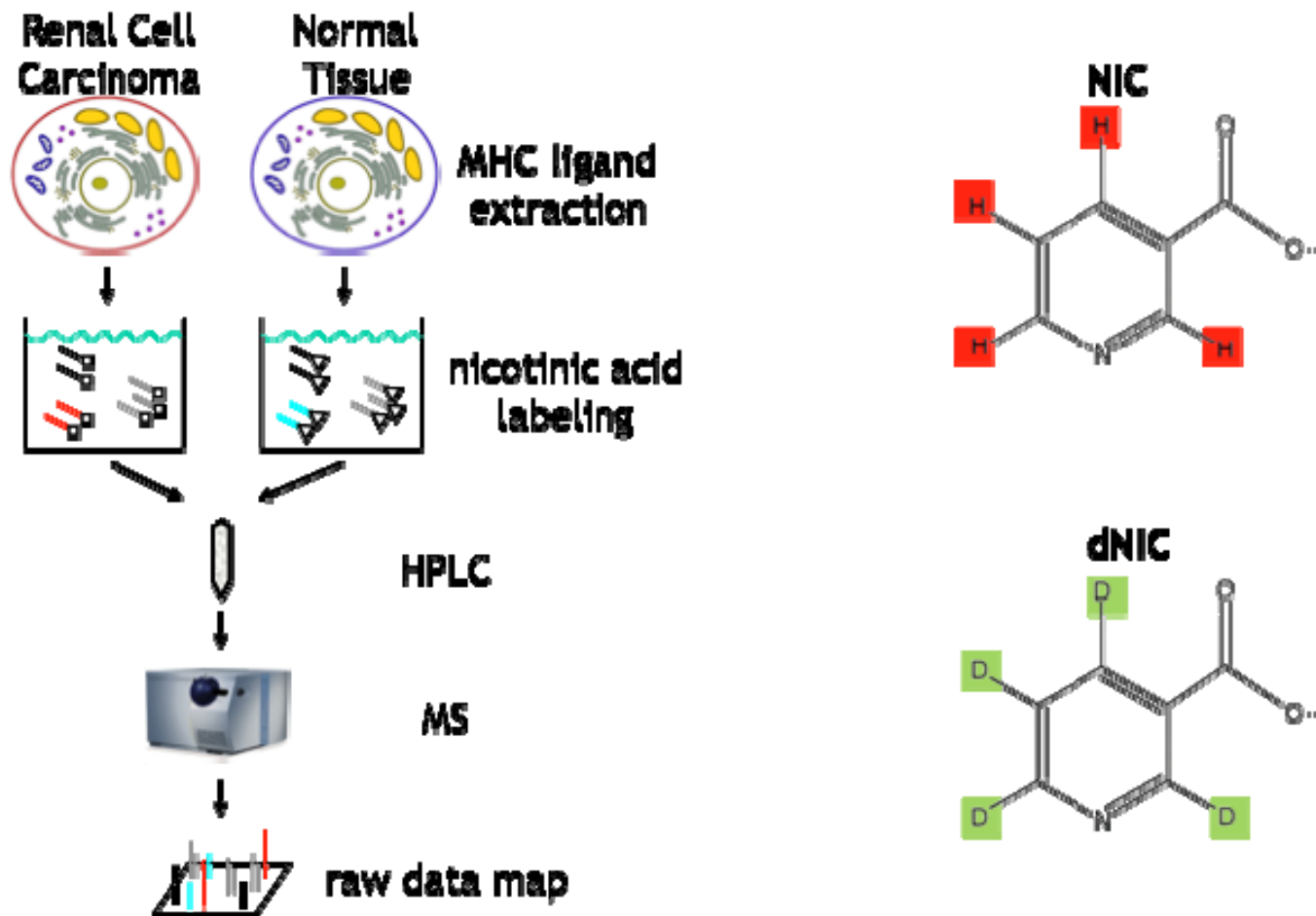
Mass spec



"Computational uncertainty"

Knut Reinert
AAAS, San Diego
February 2010

Mass spec



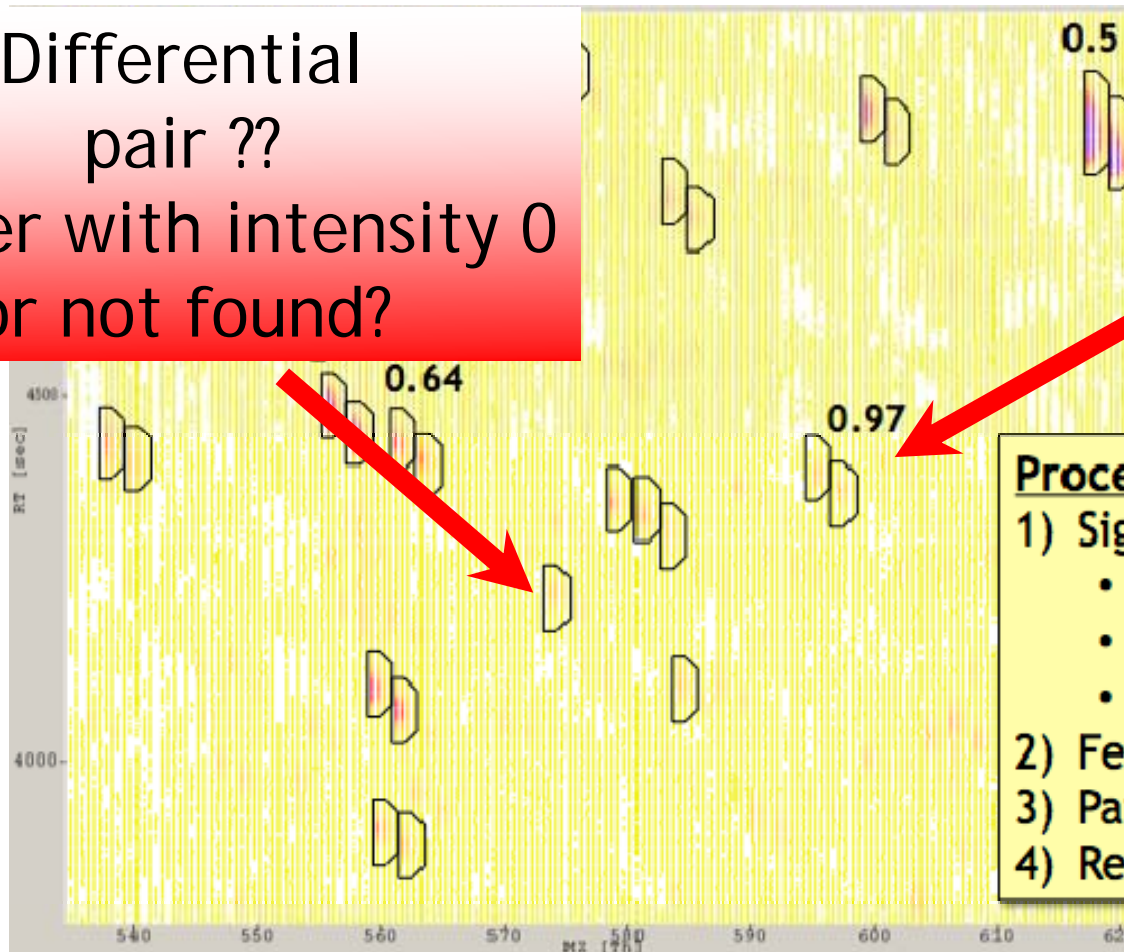
"Computational uncertainty"

Mass spec

Differential pair ??

Partner with intensity 0 or not found?

Differential pair

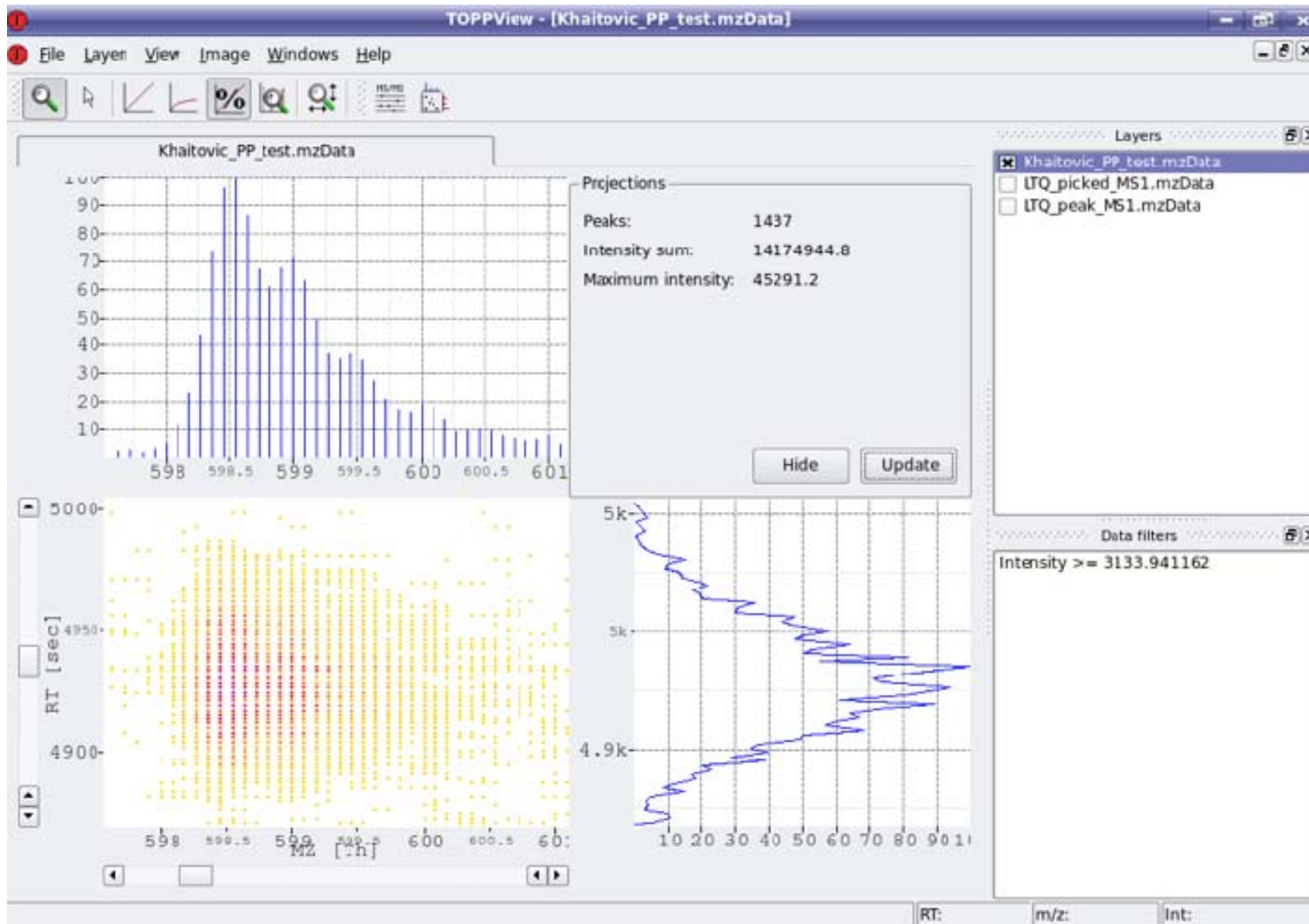


Processing steps:

- 1) Signal processing
 - Noise filtering
 - Baseline red.
 - Peak detection
- 2) Feature detection
- 3) Pair detection
- 4) Report ratios

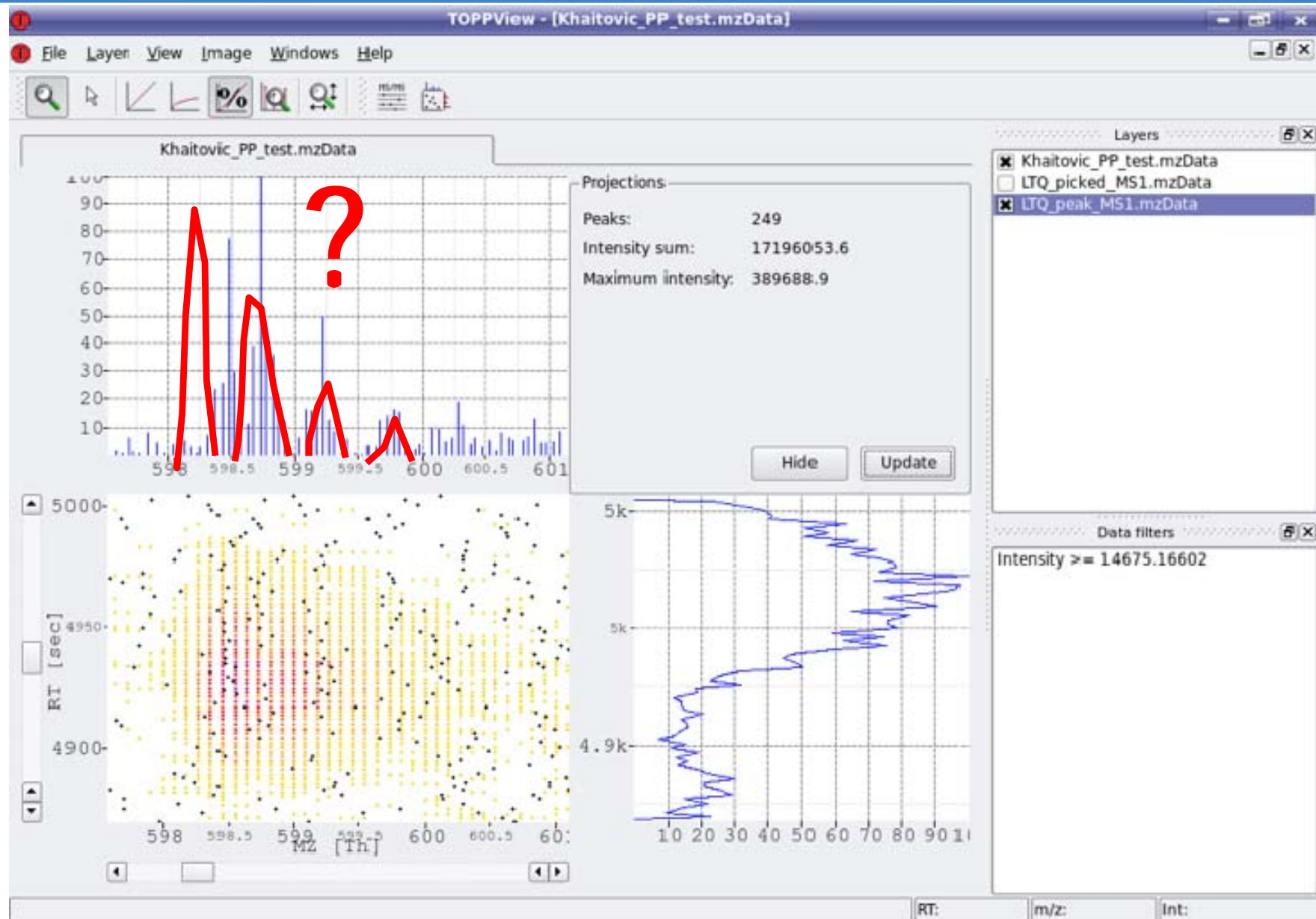
Signal Processing needed

Knut Reinert
AAAS, San Diego
February 2010



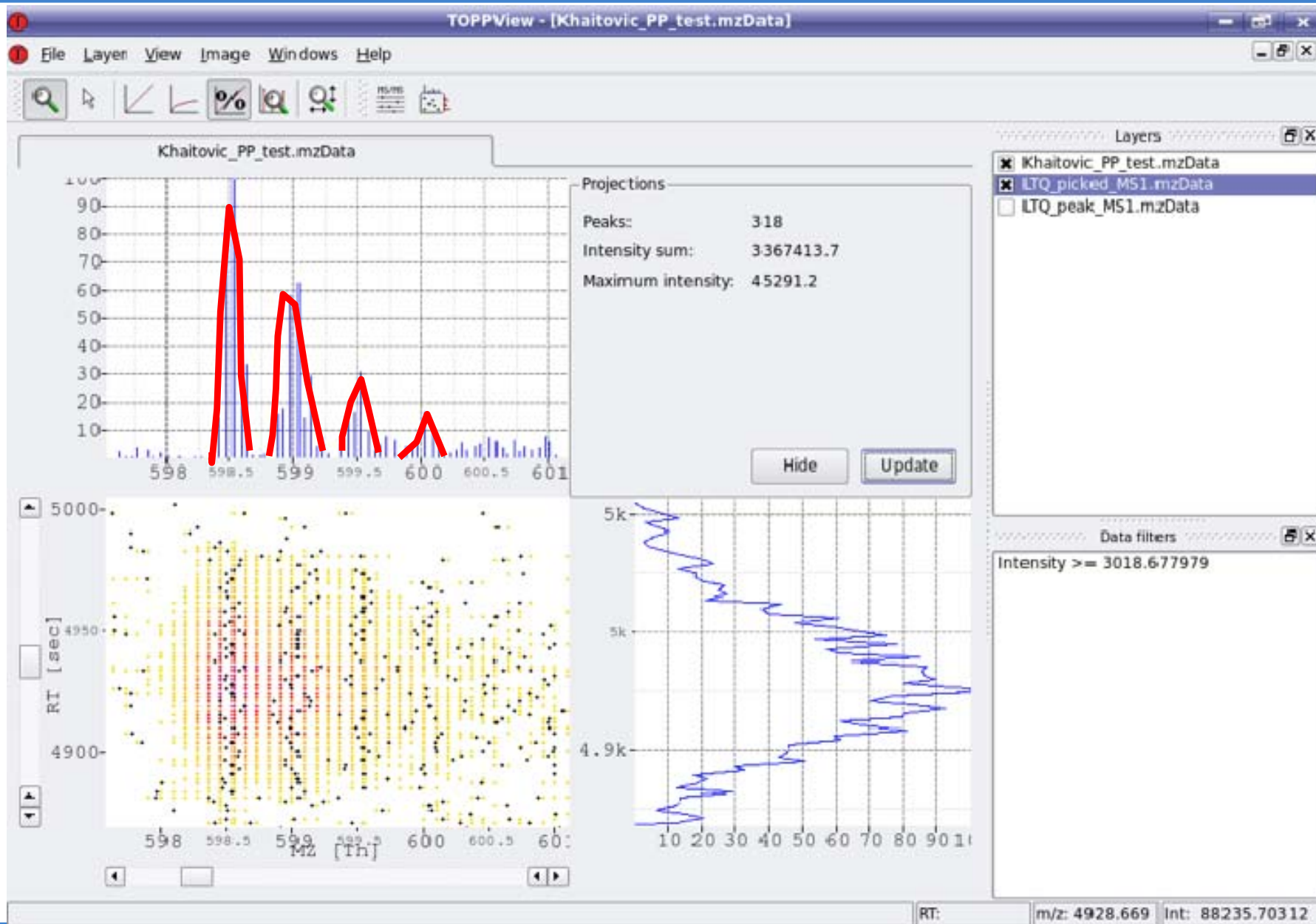
Algorithm A => quantitative signal bad

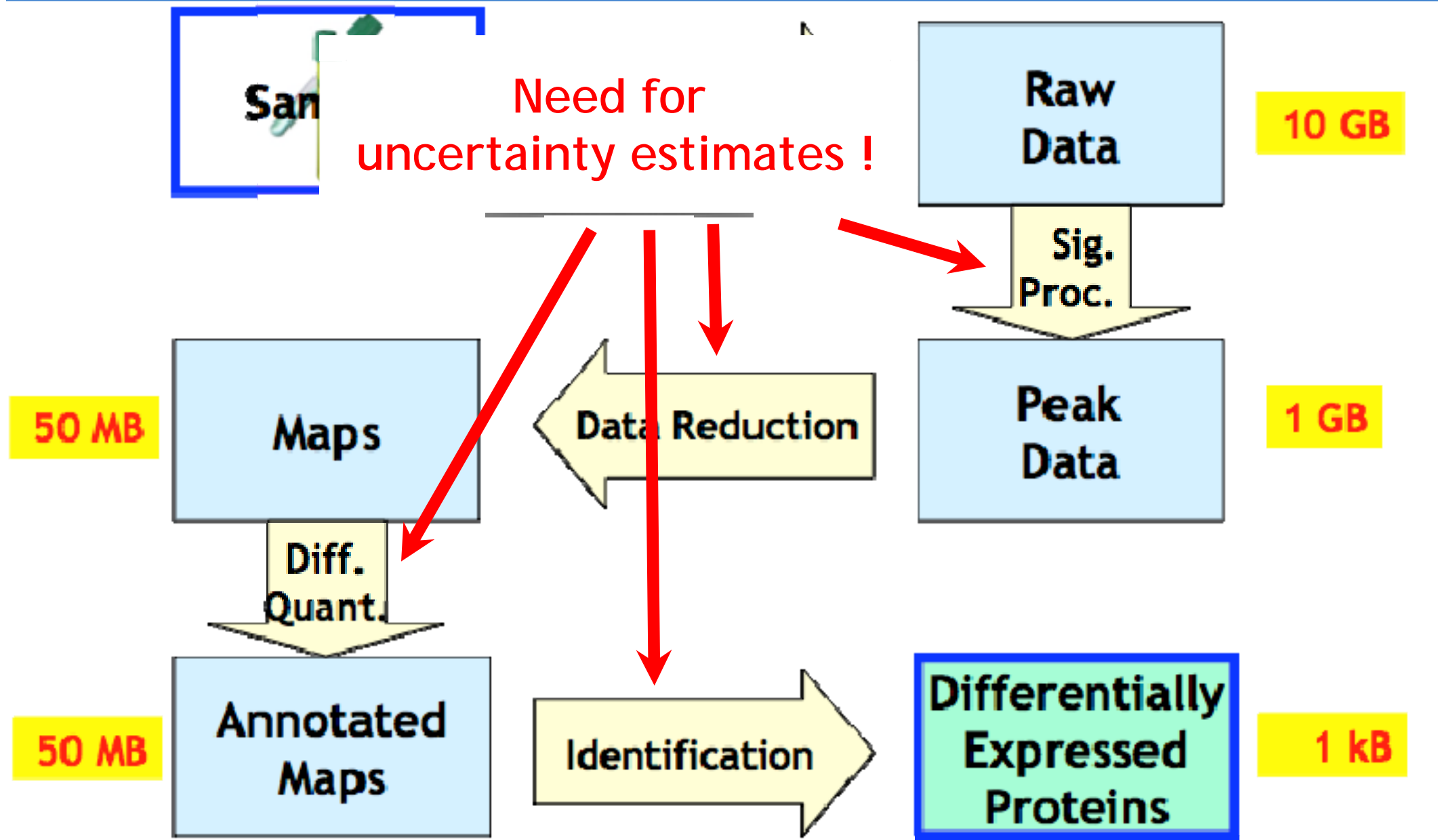
Knut Reinert
AAAS, San Diego
February 2010



Algorithm B => quantitative signal good

Knut Reinert
AAAS, San Diego
February 2010





Next generation sequencing



The Human genome project (started in 1990, finished in 2001) had costs of about **3 billion dollars** and took over 10 years to complete producing about **21 billion** base pairs

This little box can sequence in a month about **60 billion** base pairs costing a couple of thousand dollars (Illumina Ltd.) (Roche and ABI offer similar technologies)

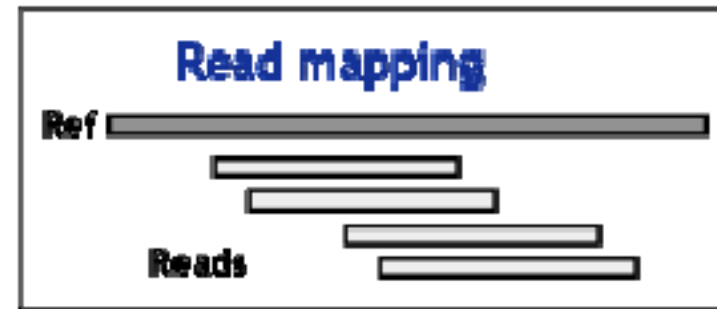
Next generation sequencing



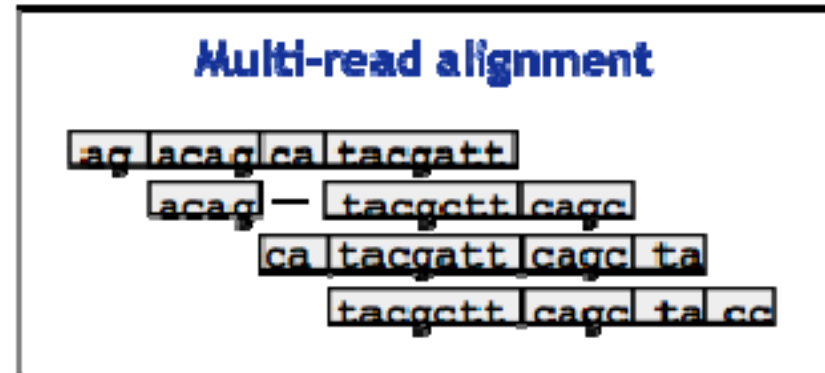
- + relatively stable technology
- + raw data quality can be assessed
- + opens a whole new world of possible tests (RNA-seq, ChipSeq, genome variability, Metagenomics etc.)
- + not limited in what can be found
- additional data processing uncertainty is **considerable** for many applications

Example: SNP calling

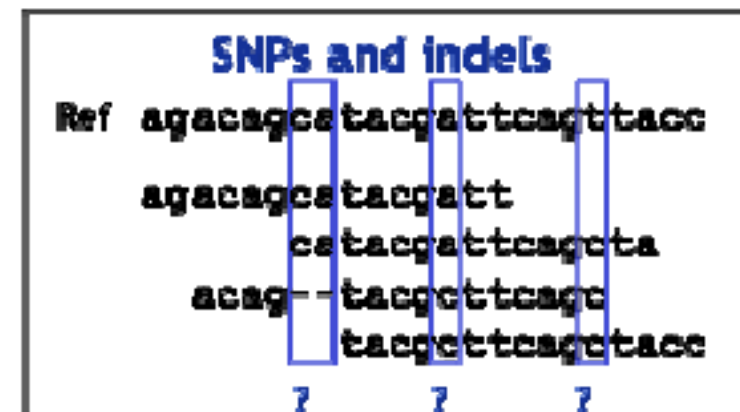
Mapping: relatively well understood



Alignment: NP-hard, can introduce **artificial** SNPs



Calling SNPs: relatively well understood



SNP or no SNP?



This is an alignment of the result of three different sequence assemblers on the SAME input. All marked differences would show as potential SNP.

Alignment program with sequences reversed

```
Seq1: TAAAGAAATACCTTTTACTTATATTTATCAAGATTCAACTTGTGTGATTTTACGTTGACCA  
Seq2: TAAAGAA-TACCTTTT-ACCTTATATTTATCAAGATTCAACTTGTGTGATTTT-ACGTTGACCA
```



```
Seq1: TAAAGAAATACCTTTTACTTATATTTATCAAGATTCAACTTGTGTGATTTTACGTTGACCA  
Seq2: TAAAG-AAATAC-CTTTACTTATATTTATCAAGATTCAACTTGTGTGA-TTTACGTTGACCA
```

```
Seq1: AATTGATTTTTTTTTTTTTTAAAGATTTATCAATTTGTTACGCAATAGACTTAAATATTAAGAT  
Seq2: AATTGATTTTTTTTTTTA--ATAATTTATCAATTTGTTACGCAATAGACTTAAATATTAAGAT
```



```
Seq1: AATTGATTTTTTTTTTTTTTAAAGATTTATCAATTTGTTACGCAATAGACTTAAATATTAAGAT  
Seq2: AATTGAA-TTTTTTTTTTAA-ATAATTTATCAATTTGTTACGCAATAGACTTAAATATTAAGAT
```

```
Seq1: AATTTAAAAAAATTTGAGGCACTCTCGTAA-CAAGATACTTTTACTTTA-TATTATCAAG  
Seq2: AATTTAAAAAAATTTGAGGCACTCTCGTAAAGATACTTTTACTTTAATAATTTATCAAG
```



```
Seq1: AATTT-AAAAAATTTGAGGCACTCTCGT-CAAGATACTTTTACTTT-ATAATTTATCAAG  
Seq2: AATTTAAAAAAATTTGAGGCACTCTCGTAAAGATACTTTTACTTTAATAATTTATCAAG
```

Toxicity testing has to be aware of potentially **MASSIV** additional uncertainty and variability introduced by computational processing

Computational biology has to address the problem of computing error estimates to incorporate them into the traditional treatment of uncertainty and variability

hard problems have heuristic solution

Thank you for your attention!

Sources used in presentation:

Transforming toxicology and implications for human health risk assessment

Peter Preuss, PPT Presentation Sept. 2009

Toxicity Testing in the Twenty-first Century: A Vision and a Strategy Committee on Toxicity and Assessment of Environmental

Science and Decisions: Advancing Risk Assessment Committee on Improving Risk Analysis Approaches Used by the U.S. EPA

Camera project

<http://camera.calit2.net/>