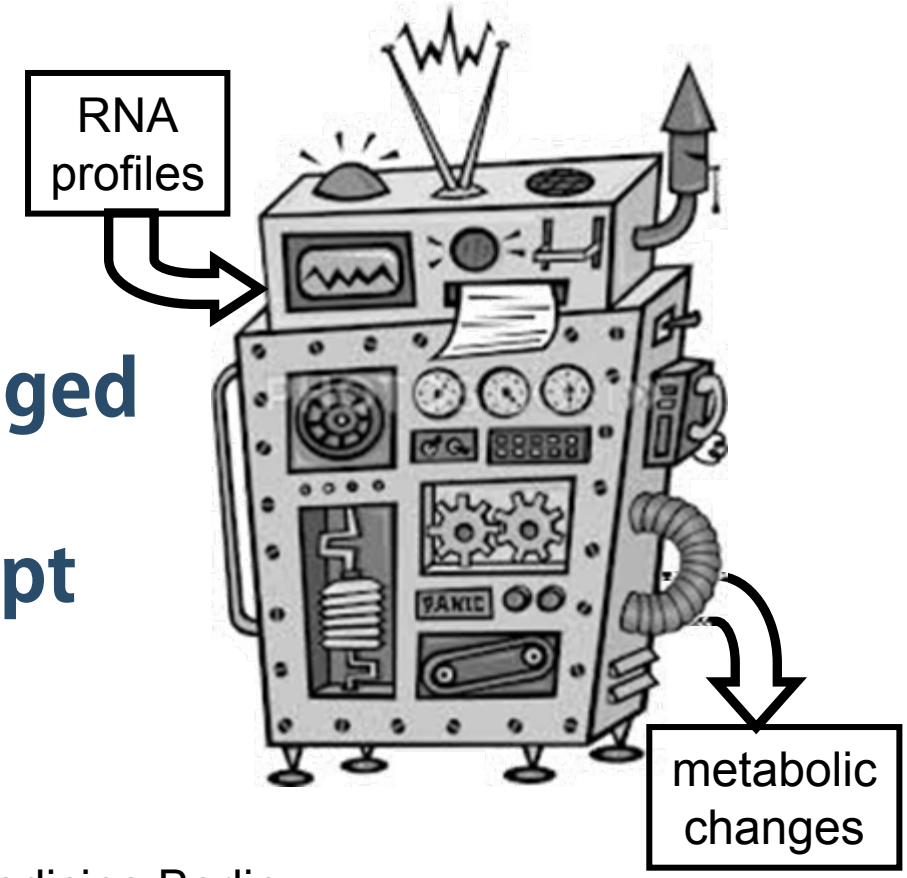
virtual liver network

ModeScore

A method to infer changed activity of metabolic functions from transcript profiles



Andreas Hoppe, Charité University Medicine Berlin
Computational Systems Biochemistry Group

COMPUTATIONAL SYSTEMS BIOCHEMISTRY

CHARITÉ
UNIVERSITÄTSMEDIZIN BERLIN

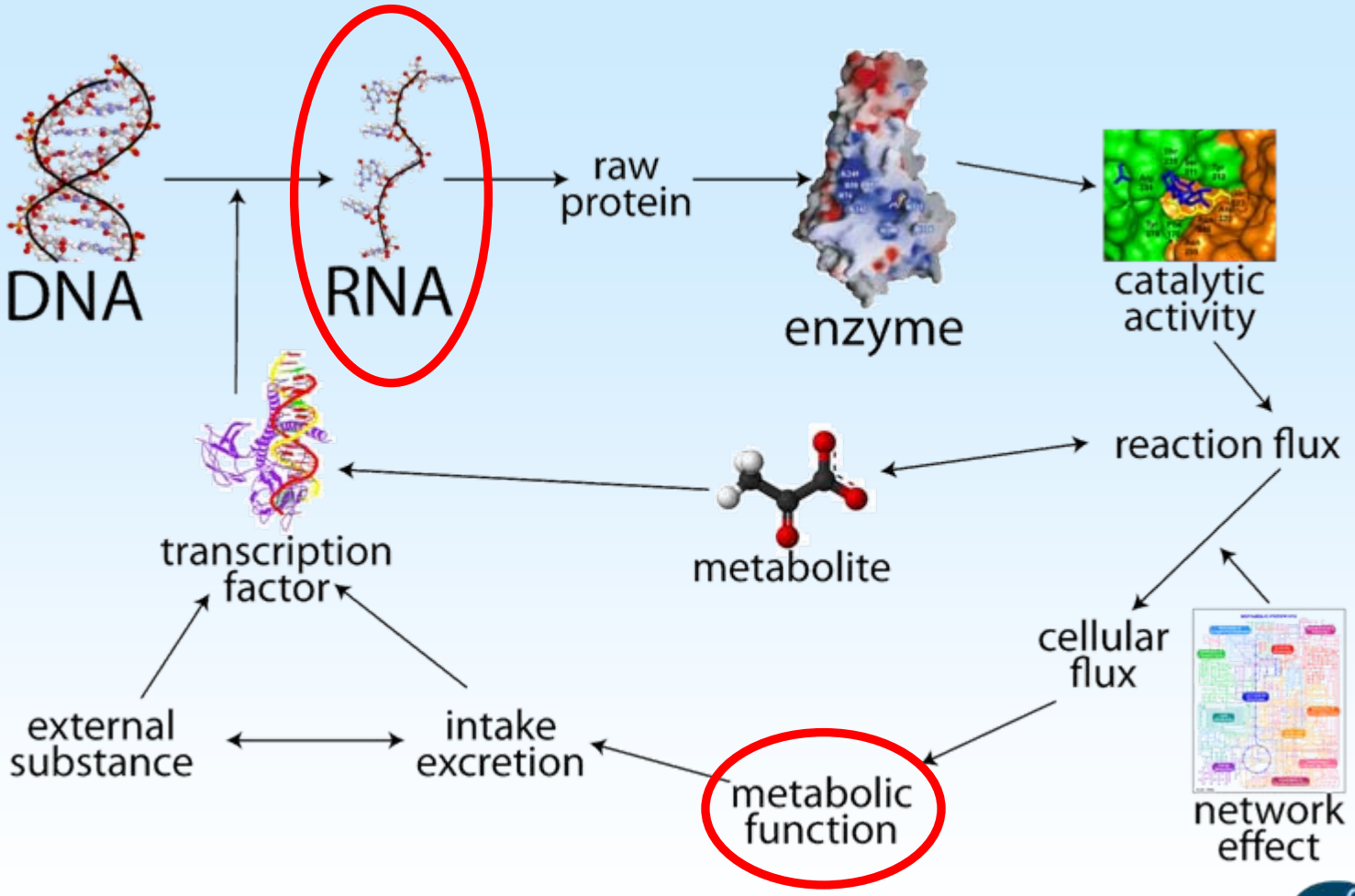
Outline

- Introduction
- ModeScore method
- Application example
- Implementation/Summary

Outline

- Introduction
- ModeScore method
- Application example
- Implementation/Summary

Functional layers of cells



virtual liver network

Functional layers of cells

Many intermediate levels

- many modifying factors
- quantitative predictivity low
- knowledge must be integrated

Gygi et al., 1999, Mol. Cell Biol.

Blazier & Papin. 2012, Front Physiol.
Hoppe, 2012, Metabolites 2.

Why transcripts then?

- large information gain per material & money
- easy measurement (compared with metabolites, fluxes, proteins)
- multitude of available datasets

virtual liver network

Objective of method

Given: measured transcript abundances

- Select metabolic function with the most remarkable pattern
- Select the genes that
  - are related to this metabolic function
  - significantly change
  - have sufficiently high expression
  - show remarkable pattern of change

virtual liver network

Outline

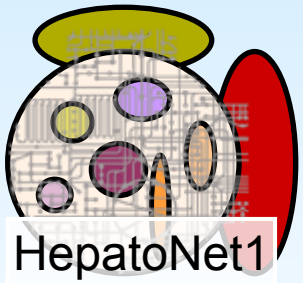
- Introduction
- ModeScore method
- Application example
- Implementation/Summary

Prediction idea

We know

- the way enzymes cooperatively work
- the cell's metabolic functions

→ reference flux distributions



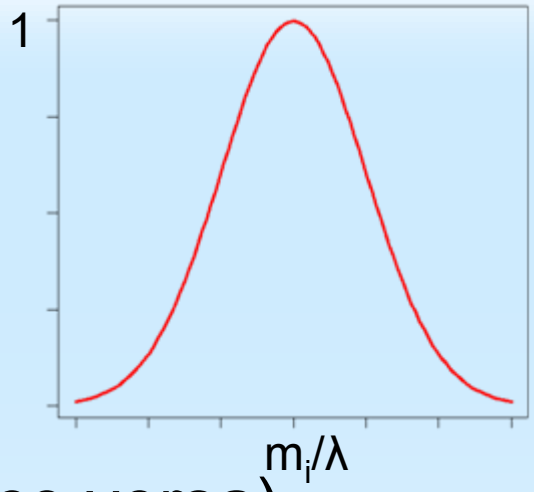
Metabolic function definition

Reference flux mode

HepatoNet1 … , Gille et al., 2010, Mol Syst Biol
FASIMU … Hoppe et al., 2011, BMC Bioinf

virtual liver network

Prediction idea



Assumptions:

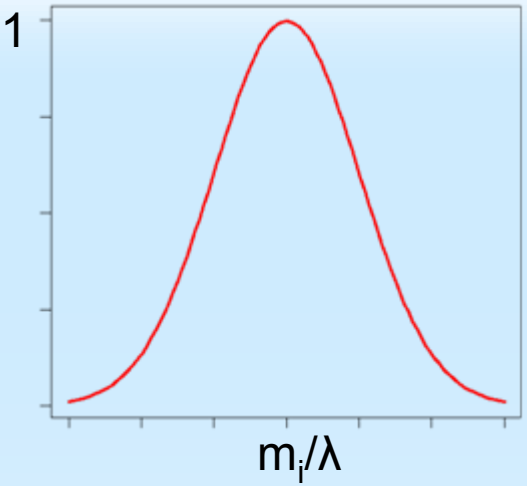
- Gene up → flux value up (& vice versa)
- Normal distribution
- Probability maximum: flux/scaling factor λ

Pattern match

Abundance change — Flux mode

virtual liver network

Mode set scoring



$$\text{Score}(M_k, V) = \frac{\sum_{i \in I_k} w_i \text{score}_i(m_i, v_i)}{\sum_{i \in I_k} w_i}$$ total score

$$I_k = \{i | \, m_i \neq 0\}$$ indices of nonzero values

$$w_i = \sqrt{|m_i| \omega_i}$$ weights

ω_i weight adjustment

$$\text{score}_i(m_i, v_i) = e^{-\frac{1}{2}\left(2 \frac{\lambda v_i - m_i}{|m_i|}\right)^2}$$ score component

$M_k = (m_i)_i$ k-th reference mode

$V = (v_i)_i$ relative expression profile

λ chosen such that $\text{Score}(M_k, V)$ is maximal

virtual liver network

ModeScore amplitude $(1/\lambda)$

- Measures strength of regulation for the function
- Compatible to $\log_2$ fold change
- Cluster point (not average) of gene changes

Contribution scores ($score_i$)

- Measures how good a gene change represents the function's amplitude

ModeScore analysis

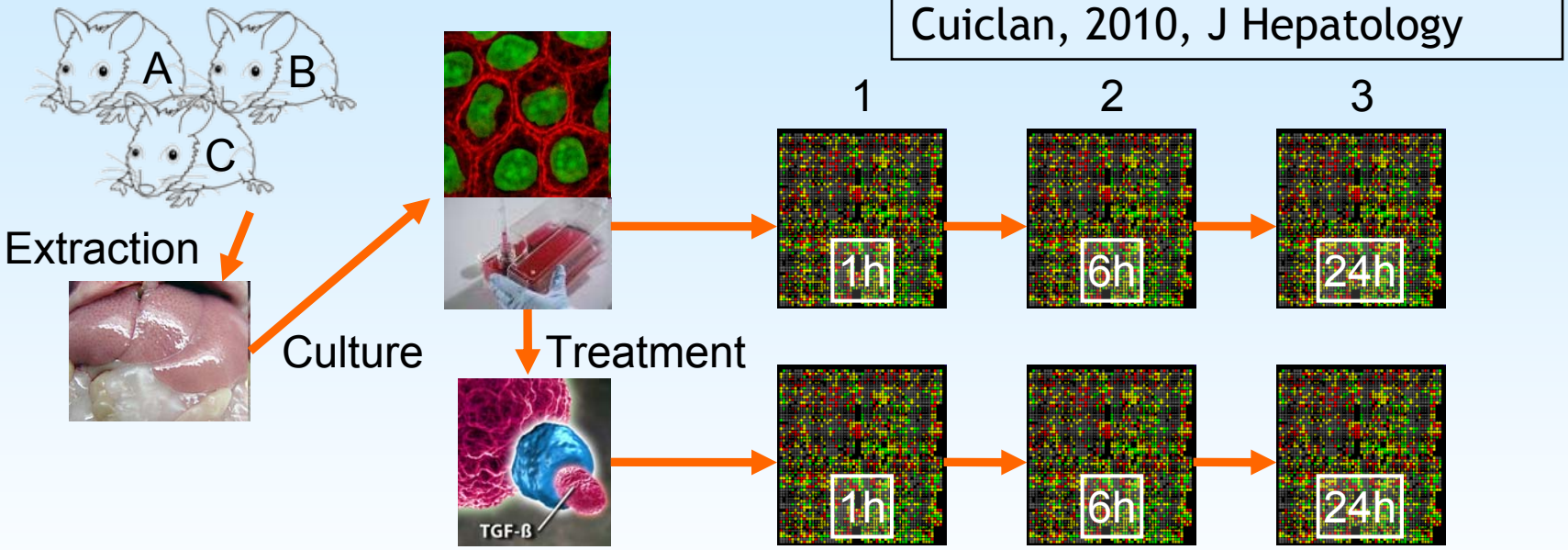
1. Ranking of functions by amplitude for each relative profile
2. Collect similar functions
3. Select remarkable functions
4. For each function, rank the genes by their contribution
5. Select set of genes representing the remarkable pattern

Outline

- Introduction
- ModeScore method
- Application example
- Implementation/Summary

Hepatocyte culture/TGFβ treatment

Dooley, 2008, Gastroenterology
Godoy, 2009, Hepatology
Cuiclan, 2010, J Hepatology



Ilkavets, Dooley

virtual liver network

Ranking of functions, example

Bottom up, C24h higher Simulation	ampl	score
448 Ethanol degr	-3.56	0.54
489 Tyrosine	-2.86	0.65
54 Glucose-6P	-2.74	0.52
925 Saccharopine	-1.92	0.63
658 4-Hydroxyphenylpyruvate	-1.87	0.73
667 5-Formiminotetrahydrofolate	-1.82	0.46
821 Homogentisate	-1.76	0.65
855 Mercaptopyruvate	-1.75	0.37
89 Alanine from Phenylalanine	-1.74	0.47
896 Pantetheine	-1.68	0.52
907 Phosphopantetheine	-1.67	0.45
272 beta-Alanine from Alanine	-1.67	0.29
78 Alanine from Aspartate	-1.65	0.43
343 Collagen CD36(c) synthesis	-1.64	1
660 4-Maleylacetoacetate	-1.62	0.7
118 Asparagine from Alanine	-1.6	0.39
407 Alanine degr	-1.54	0.43

TGFβ treated at 24h vs. control 24h

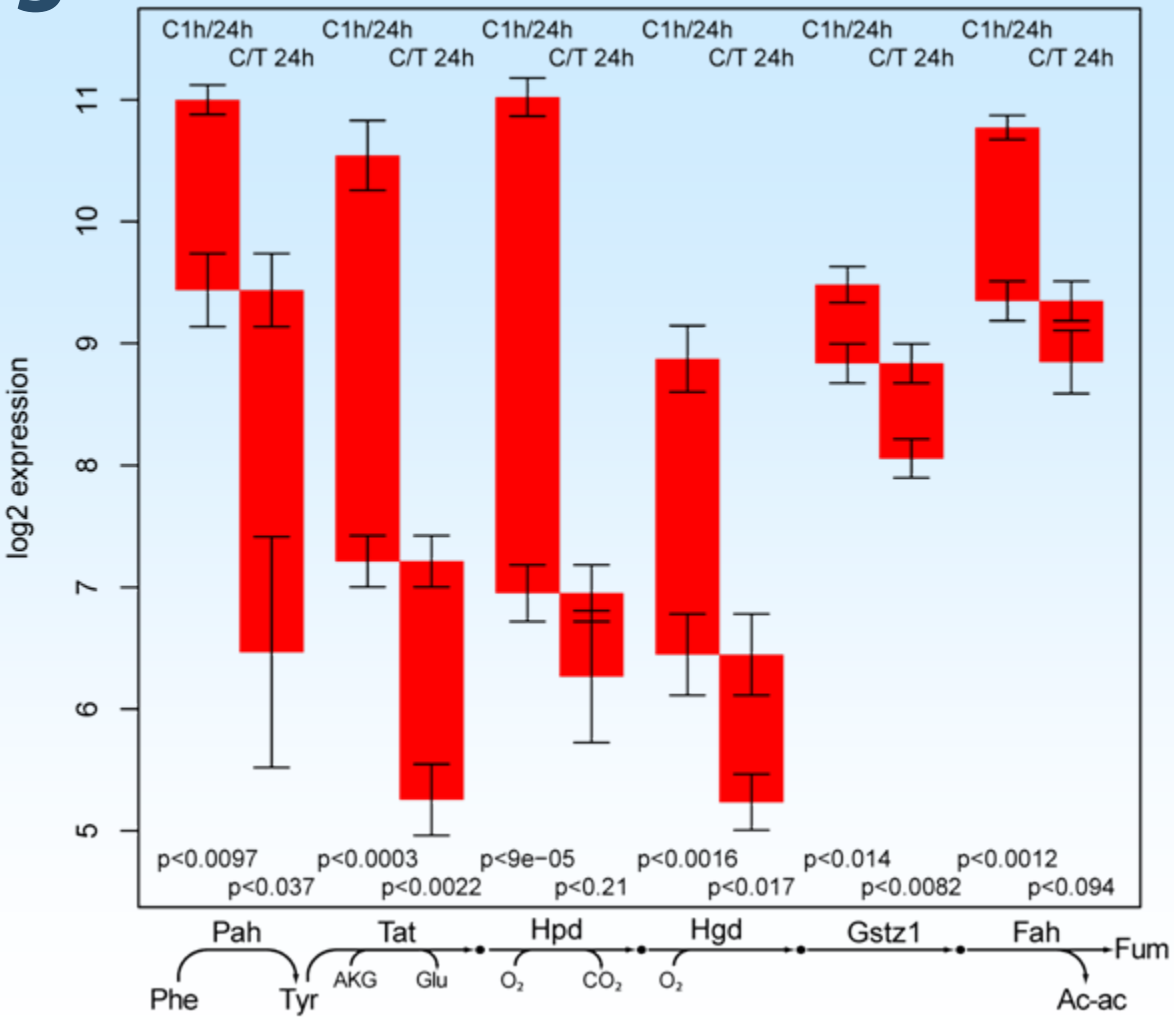
virtual liver network

Selection of genes, example

Rea ID	control 1/24h Score *0.41*	Expr Δ	Expr C1h	Expr C24h	TGFβ/C 24h Score *0.43*	Expr Δ	Expr C24h	Expr T24h	Reaction
r0544	0.87	-4.07	11	6.95	0.98	-0.69	6.95	6.27	4-Hydroxyphenylpyruvate(c) + O2(c) ⇌ Homogentisate(c) + CO2(c)
Hpd									ENSMUSG00000029445
r0183	1	-3.33	10.5	7.21	0.01	-1.96	7.21	5.26	Tyrosine(c) + AKG(c) ⇌ 4-Hydroxyphenylpyruvate(c) + Glutamate(c)
Tat									ENSMUSG00000001670
r0605	0.28	-0.65	9.48	8.84	1	-0.78	8.84	8.06	4-Maleylacetoacetate(c) ⇌ Fumarylacetoacetate(c)
Gstz1									ENSMUSG00000021033
r0398	0.19	-0.29	10.2	9.92	0.89	-0.58	9.92	9.34	Dihydrobiopterin(c) + NADPH(c) ⇌ Tetrahydrobiopterin(c) + NADP+(c)
Qdpr (2 sp.)									ENSMUSG00000015806
r0399	0.59	-1.56	11	9.44	0	-2.97	9.44	6.47	Tetrahydrobiopterin(c) + Phenylalanine(c) + O2(c) ⇌ Dihydro-biopterin(c) + Tyrosine(c) + H2O(c)
Pah									ENSMUSG00000020051
r0320	0.11	0.16	6.45	6.61	0.2	-0.08	6.61	6.53	ATP(c) + CoA(c) + Acetoacetate(c) → AMP(c) + PPi(c) + Acetoacetyl-CoA(c)
Aacs (3 sp.)									ENSMUSG00000029482
r0543	0.88	-2.43	8.87	6.45	0.49	-1.21	6.45	5.24	Homogentisate(c) + O2(c) ⇌ 4-Maleylacetoacetate(c)
Hgd									ENSMUSG00000022821
r0324	0.53	-1.42	10.8	9.35	0.79	-0.5	9.35	8.85	Fumarylacetoacetate(c) + H2O(c) ⇌ Fumarate(c) + Acetoacetate(c)
Fah									ENSMUSG00000030630

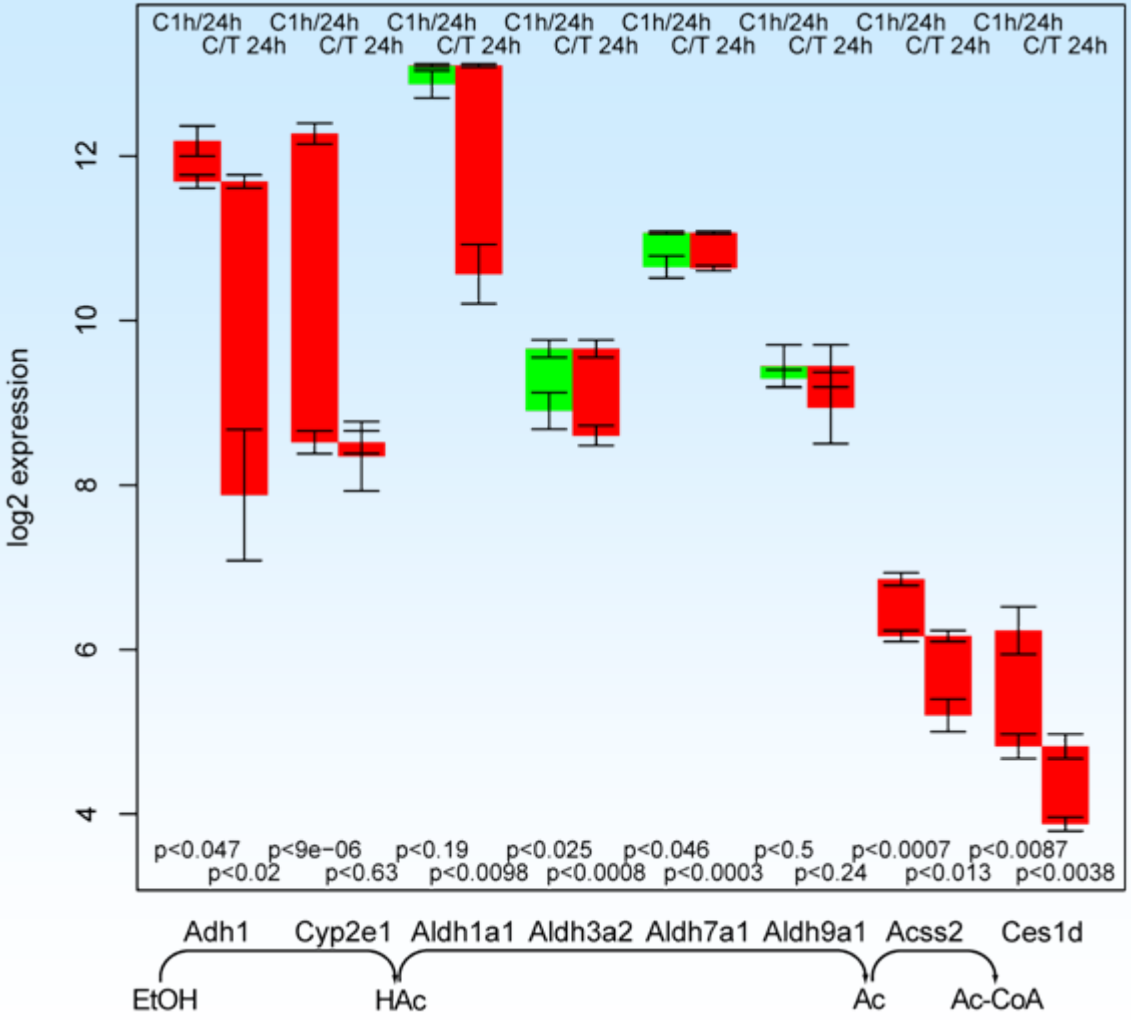
virtual liver network

Phenylalanine/Tyrosine degradation



virtual liver network

Ethanol degradation



virtual liver network

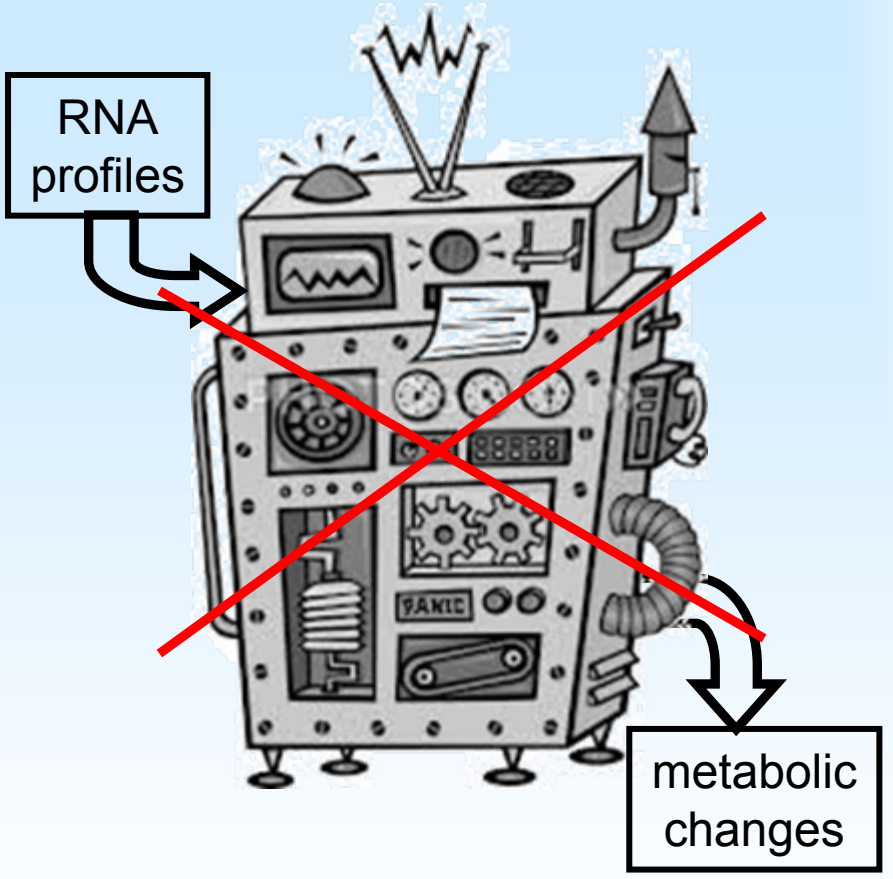
Outline

- Introduction
- ModeScore method
- Application example
- Implementation/Summary

Implementation

- Reference flux mode computation
  - www.bioinformatics.org/fasimu
- ModeScore computation
  - data handling: bash/gawk
  - scaling factor optimization: octave
  - table generation: LaTeX
  - bargraphs, t-test: R

Summary



virtual liver network

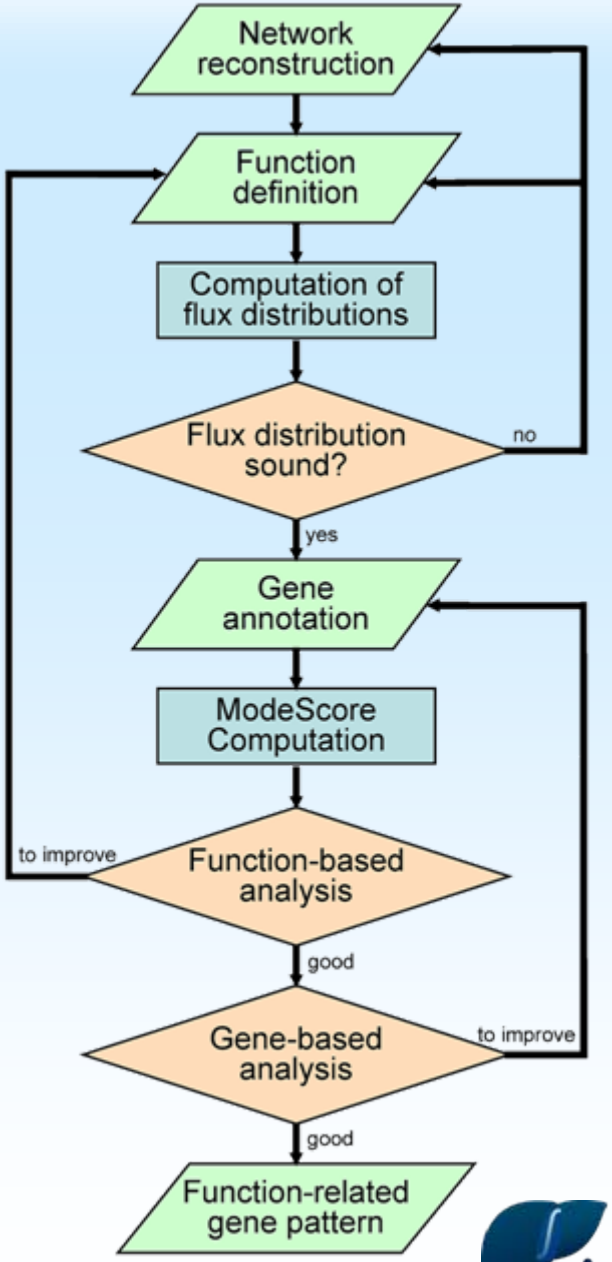
Summary

Semi-automatic process

- refinement of network, functions, annotations
- scoring/ranking
- manual selection

Selection of changed genes

Testable hypothesis



virtual liver network

Acknowledgements

Hermann-Georg Holzhütter, Berlin

Sebastian Vlaic, Jena

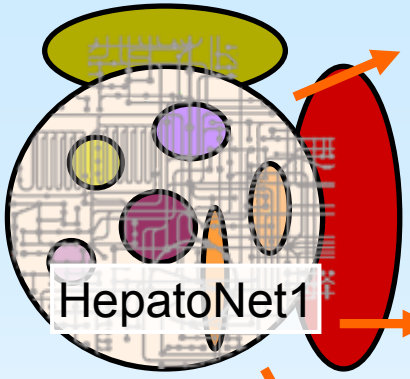
Iryna Ilkavets Mannheim

Matthias König, Berlin

Patricio Godoy, Dortmund

virtual liver network

HepatoNet gene annotation



815 (RN)

782 (KO)

738 (HSA)

915 (HSA)

914 (ENSG)

242 (EC)

177 (HSA)

2539

7 (ENSxxxG)

389 (Uniprot)

837 (ENSG)

2074 (ENSMUSG)

62 (ENSMUSG)

1225 (Uniprot)

289 (ENSRNOG)

802 (TCDB)

800 (Uniprot)

2 (ENSBTAG)

1 (ENSOCUG)

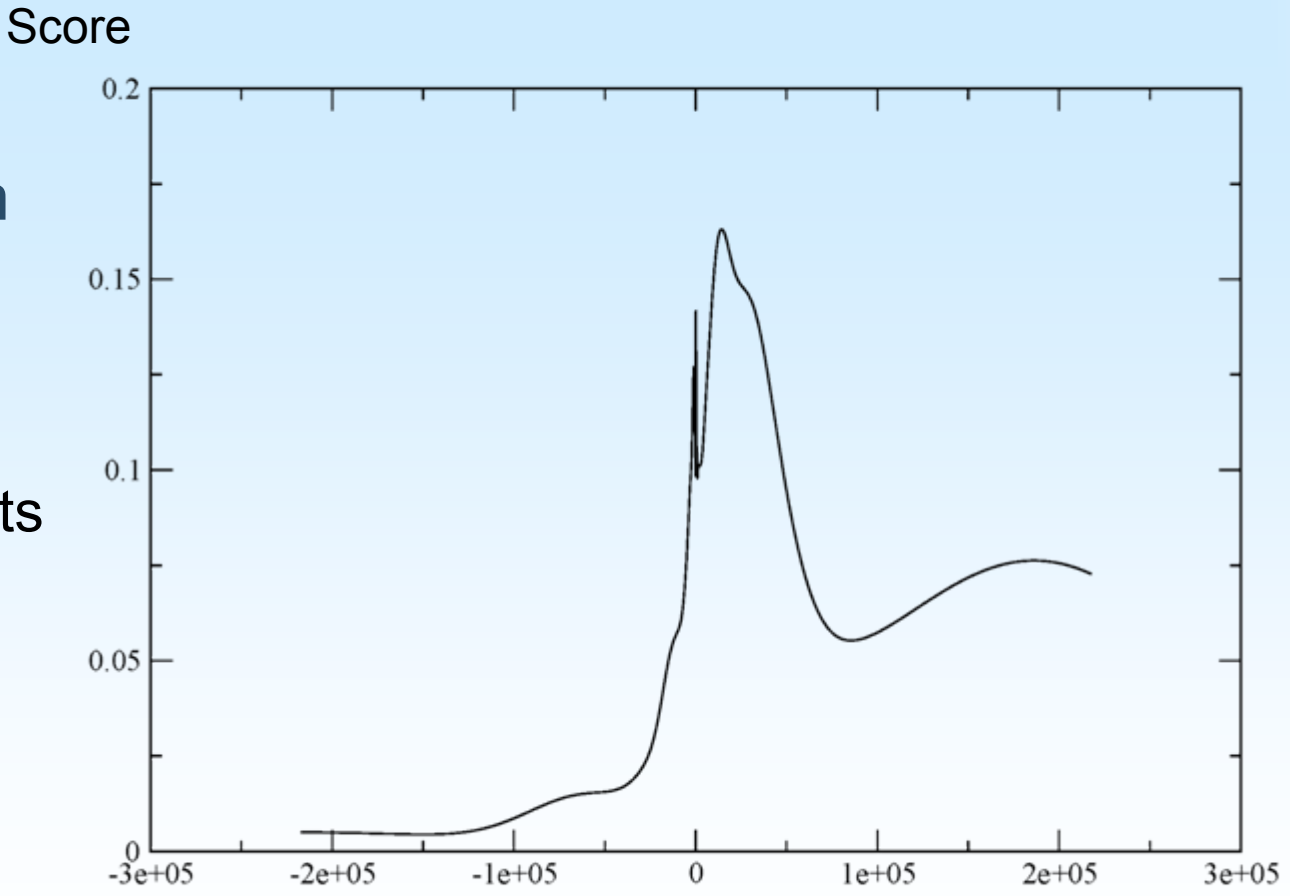
Transporter Classification Database

UniProt

virtual liver network

Optimization of scaling factor

- Optimization
  - Explicit derivation
  - Maxima of components
- Local maxima: alternative matches



virtual liver network