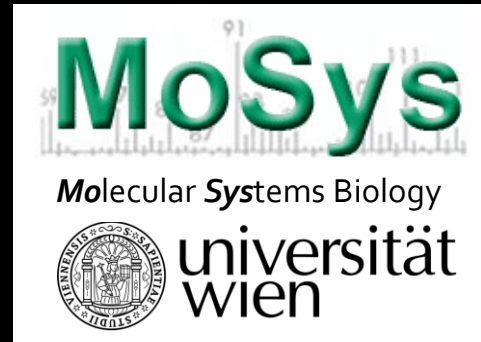


# Systemtheoretische Konzepte der genomweiten molekularen Analyse und Datenintegration in der Biologie



Wolfram Weckwerth  
Head of Department of Ecogenomics and Systems Biology  
MOSYS lab, University of Vienna



## Genotype -> Phenotype?

# Department of Ecogenomics and Systems Biology - PANOMICS platform



- From Model Organisms to Ecosystems
- Environmental samples: Genomics, Transcriptomics, Proteomics, Metabolomics



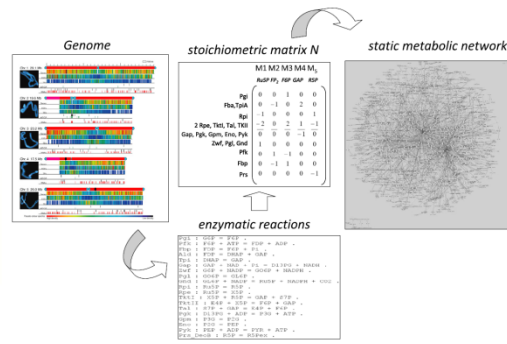
NGS



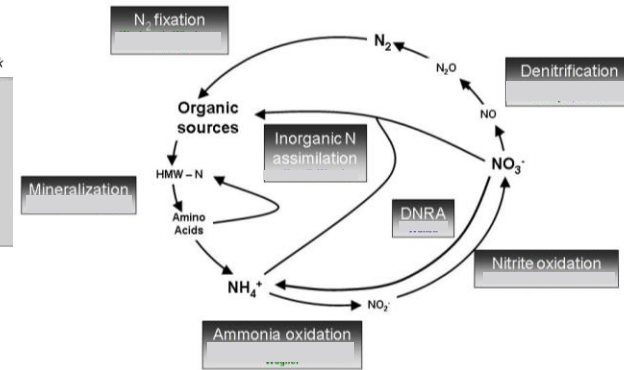
Proteomics



Metabolomics



Genome-scale modelling

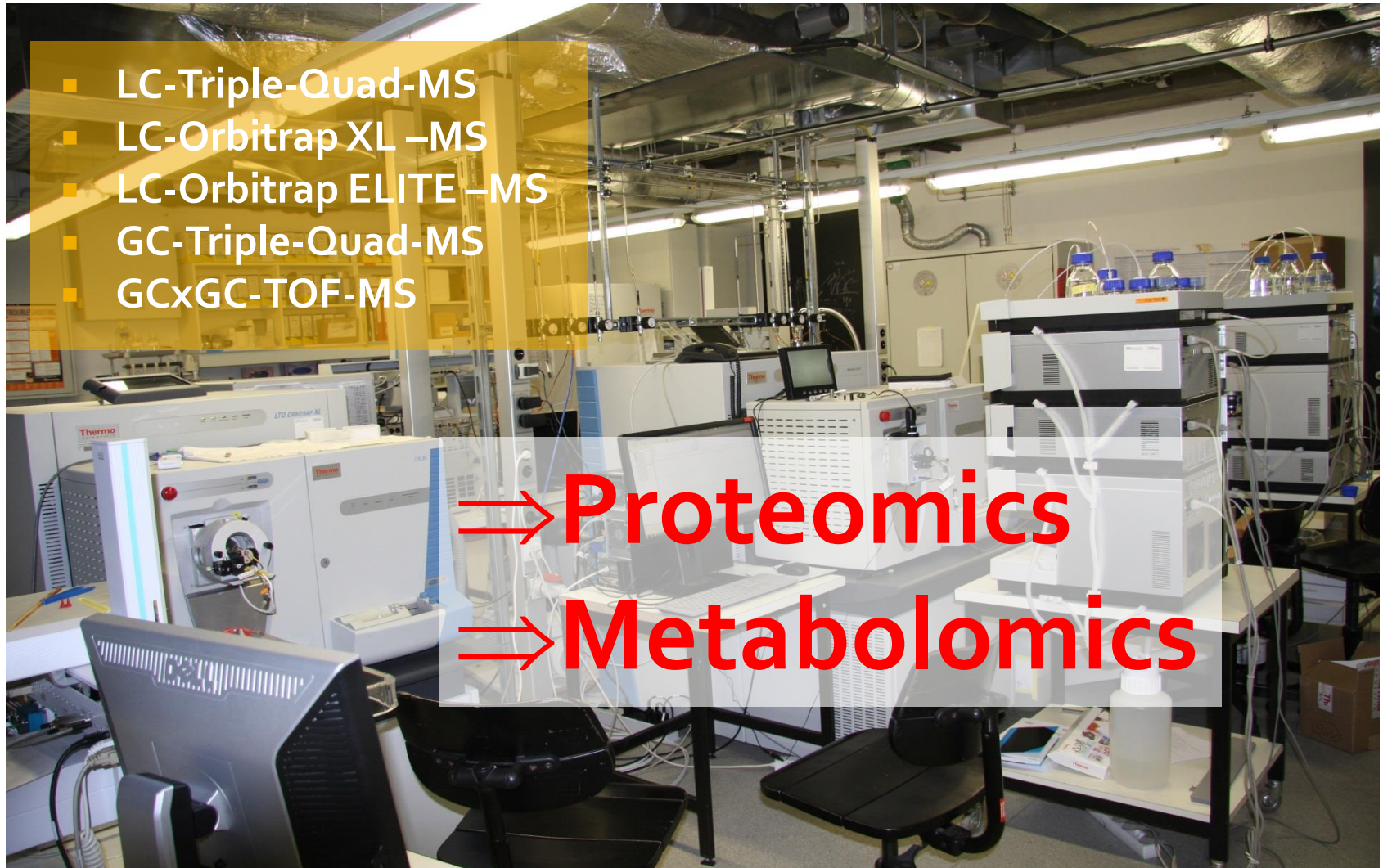


Ecosystems analysis

# Mass Spectrometry Facility

- LC-Triple-Quad-MS
- LC-Orbitrap XL –MS
- LC-Orbitrap ELITE –MS
- GC-Triple-Quad-MS
- GCxGC-TOF-MS

⇒ Proteomics  
⇒ Metabolomics

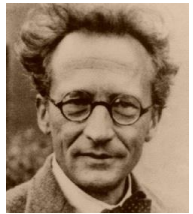
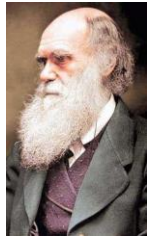


# Overview

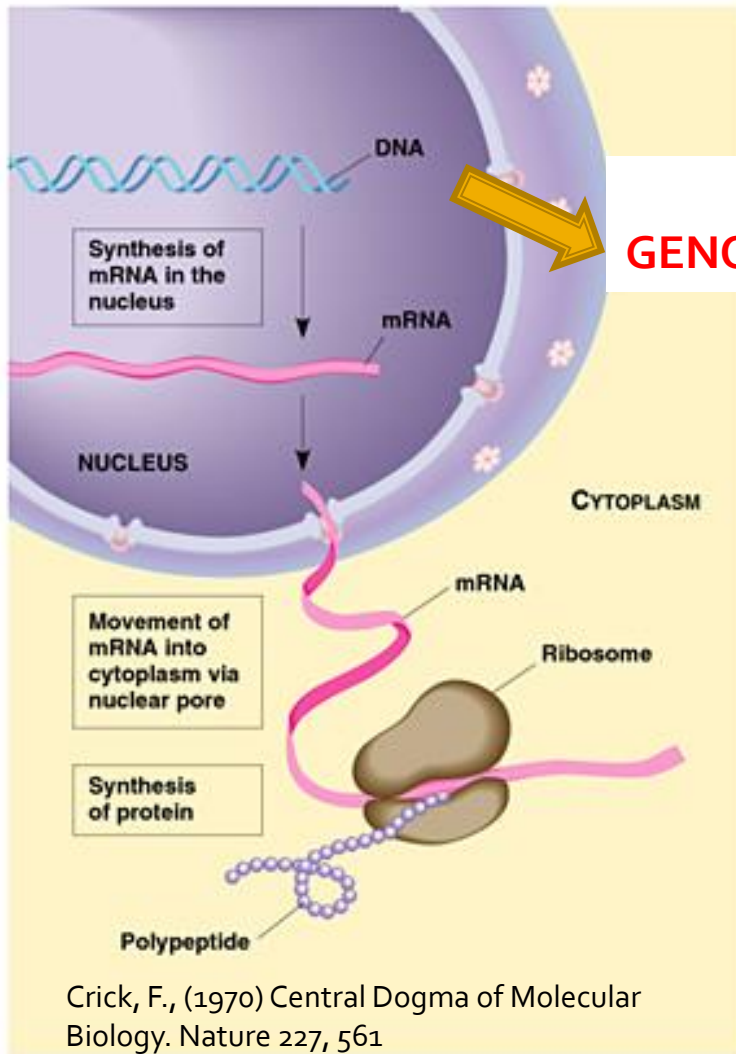
1. Motivation
2. PANOMICS platform
3. The genotype-phenotype equation

# What is life?

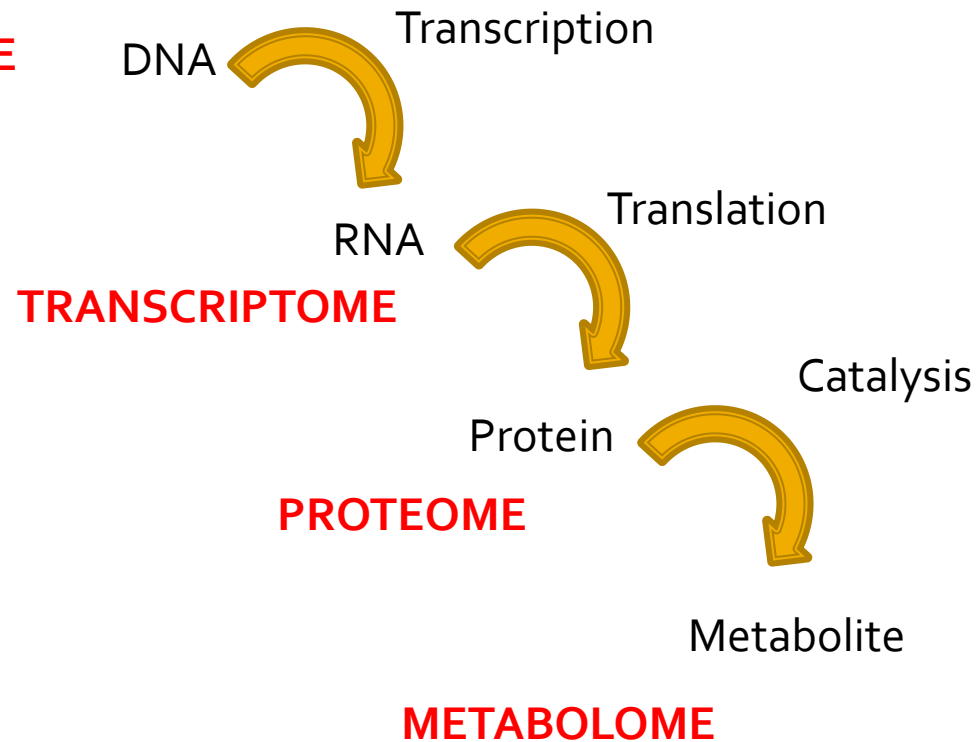
- Chemical-physical Organisation of Life:
  - Synthesis of Urea by Wöhler (1828)
  - Evolutionary Theory – Charles Darwin (1859)
  - Mendel's laws of inheritance - Gregor Mendel (1865) Versuche über Pflanzenhybride
  - Delbrück, Luria, Avery (1943/44); Erwin Schrödinger (1944) Was ist Leben? Postulate of a **molecular** genetic Code
  - Watson, Crick (1953) Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid  
Principles of information storage and transfer



# Central Dogma of Molecular Biology



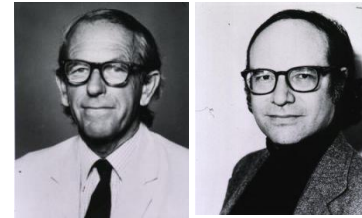
Principles of information transfer:



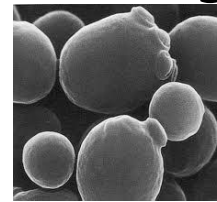
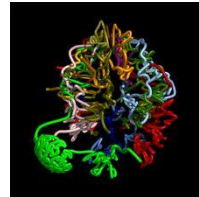
Crick, F., (1970) Central Dogma of Molecular Biology. Nature 227, 561

# Genome Sequencing

- 1976 DNA-Gilbert/Sanger-Sequencing

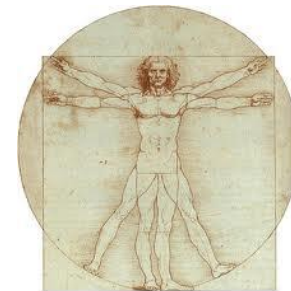


- 1996 Yeast genome



*Saccharomyces*

- 2000 Human- and higher plant genome

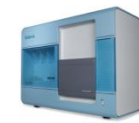


*Arabidopsis*

- 2010 "Next Generation Sequencing" (NGS)
- De novo/Resequencing in days



454  
Pyrosequencing



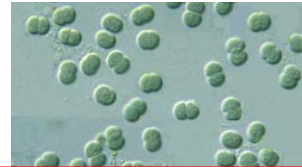
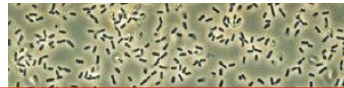
Genome Analyzer  
Reversible Terminator Chemistry



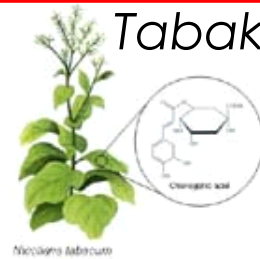
SOLID  
Ligation-based extension

# Next generation sequencing...

*Arabidopsis*



Seit 1995 ~80000 Genome projects... exponential...  
Metagenomes...



*Canis familiaris*



*Chlamydomonas*



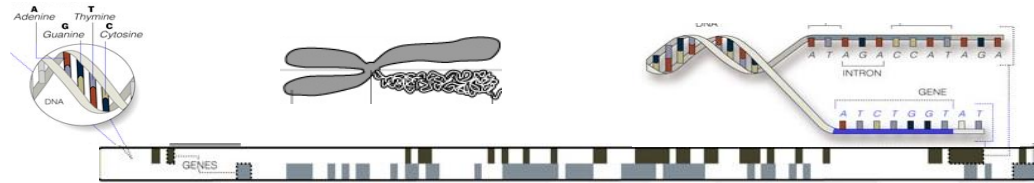
*Maus*

*Medicago*



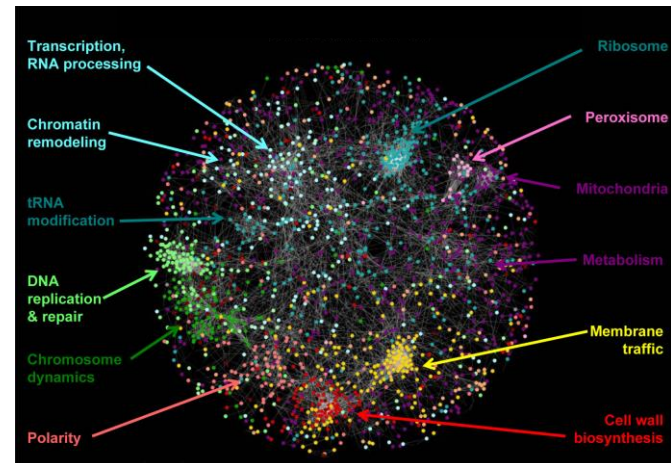
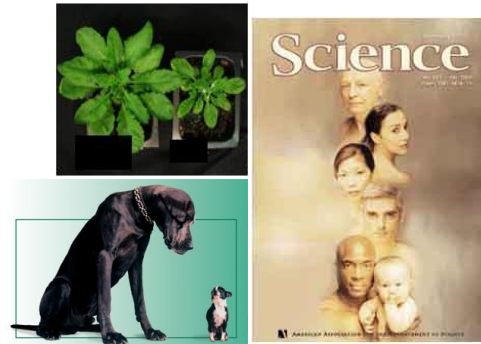
# Prediction of Phenotype?

GENOTYPE



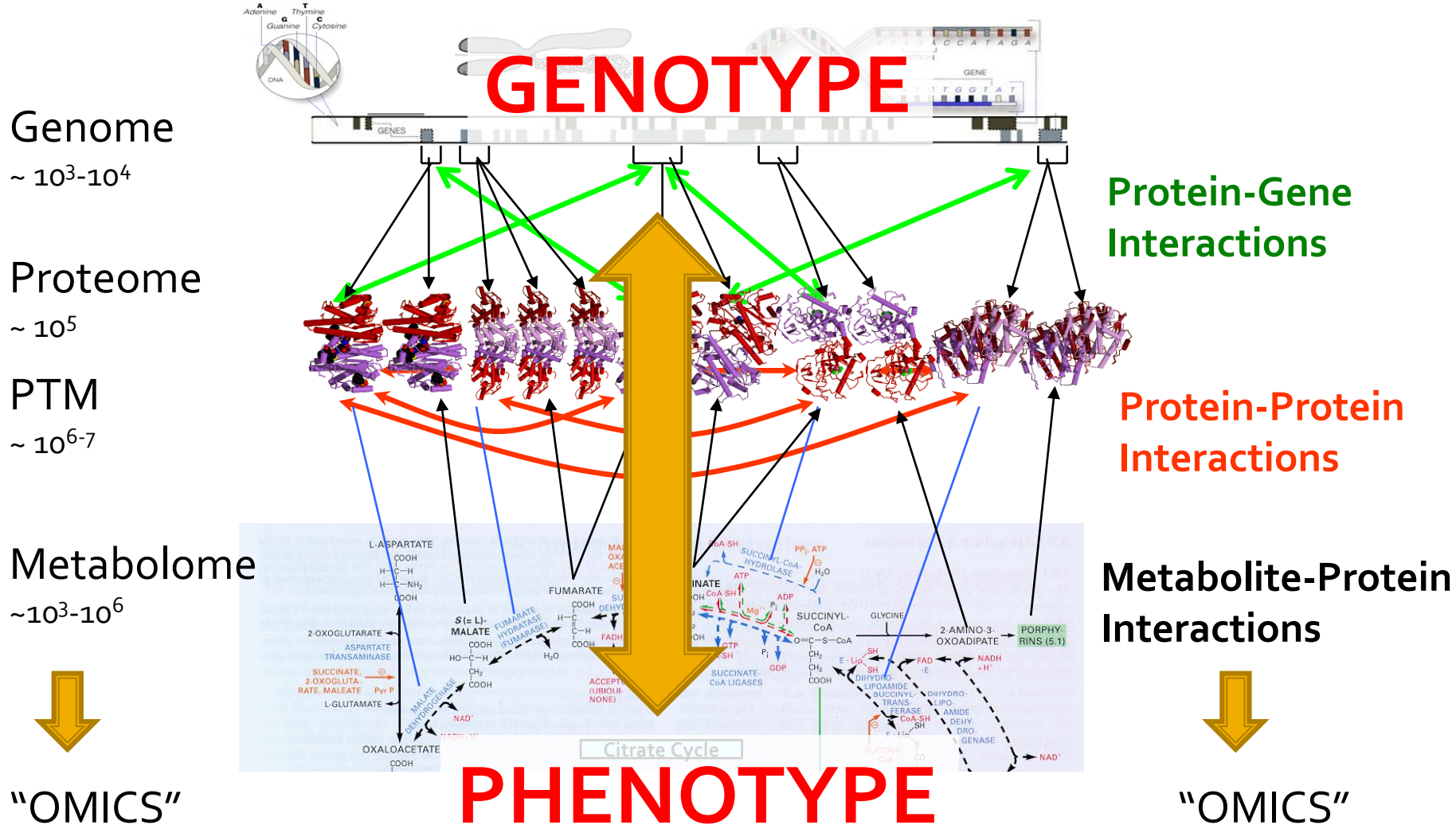
Prediction?

PHENOTYPE  
- dynamic



Weckwerth (2011) Unpredictability of metabolism.  
*Analytical and Bioanalytical Chemistry*

# How is complexity encoded?



# Genetischer Blueprint versus Cytoplasma-Gen-Interaktion

- Richard Dawkins Das egoistische Gen  
-> der genetische Blueprint des Lebens....???

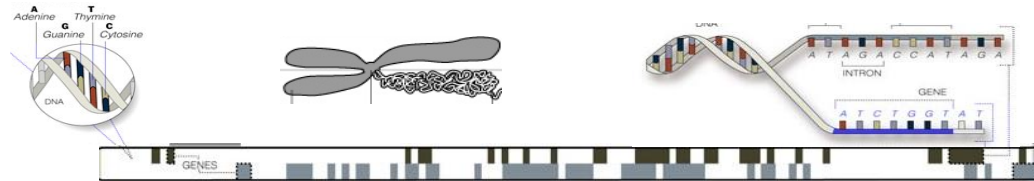
**Cytoplasmic Impact on Cross-Genus Cloned Fish Derived from Transgenic Common Carp (*Cyprinus carpio*) Nuclei and Goldfish (*Carassius auratus*) Enucleated Eggs**

Sun et al. *BIOLOGY OF REPRODUCTION* **72**, 510–515 (2005)



# Unpredictability of Metabolism

GENOTYPE

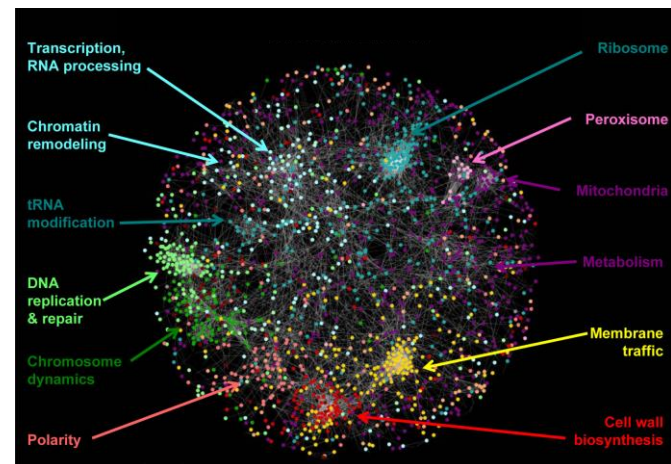


No! We have to *measure* metabolism to understand the genotype-phenotype equation



Prediction?

PHENOTYPE  
- dynamic

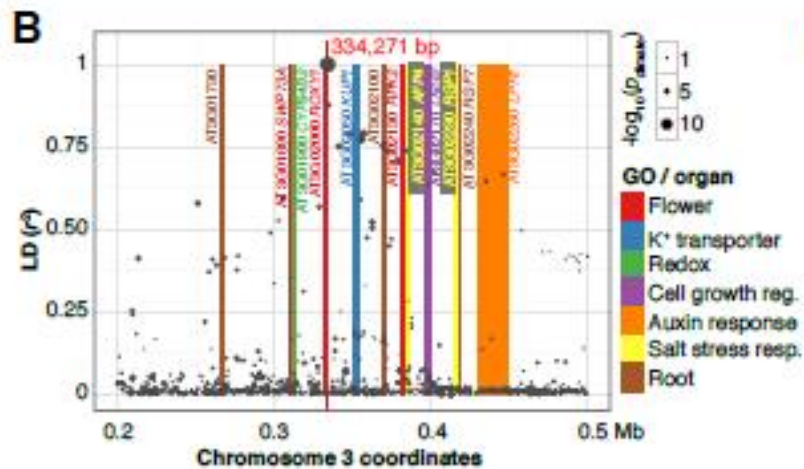


Weckwerth (2011) Unpredictability of metabolism.  
*Analytical and Bioanalytical Chemistry*

# 1001 Genomes Consortium <http://1001genomes.org/>



Polymorphism  
involved in  
climate  
adaptation



The 1001 Genomes Consortium (2016)  
**1135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana***  
*Cell* in print

# GWAS/Genotyping/NGS – GREAT, BUT....

- NO CAUSALITY
- NO PHYSIOLOGY

Genomic data will soon become a commodity; the next challenge — linking ... genetic variation with physiology ... — will be as great as the one genomicists faced a decade ago.

Craig Venter 2010

# Causal prediction of phenotype from Genotype?

- ⇒ Model design for prediction
- ⇒ Systems Biology

Weckwerth (2011) Unpredictability of metabolism.  
*Analytical and Bioanalytical Chemistry*

Weckwerth (2011) Green Systems Biology.  
*Journal of Proteomics*

Ludwig von Bertalanffy

# Vom Molekül zur Organismenwelt

Akademische Verlagsgesellschaft Athenaion  
Potsdam 1944



1901-1972

*geboren in Atzgersdorf bei Wien*

*Professor an der Universität*

*Wien 1934-48*



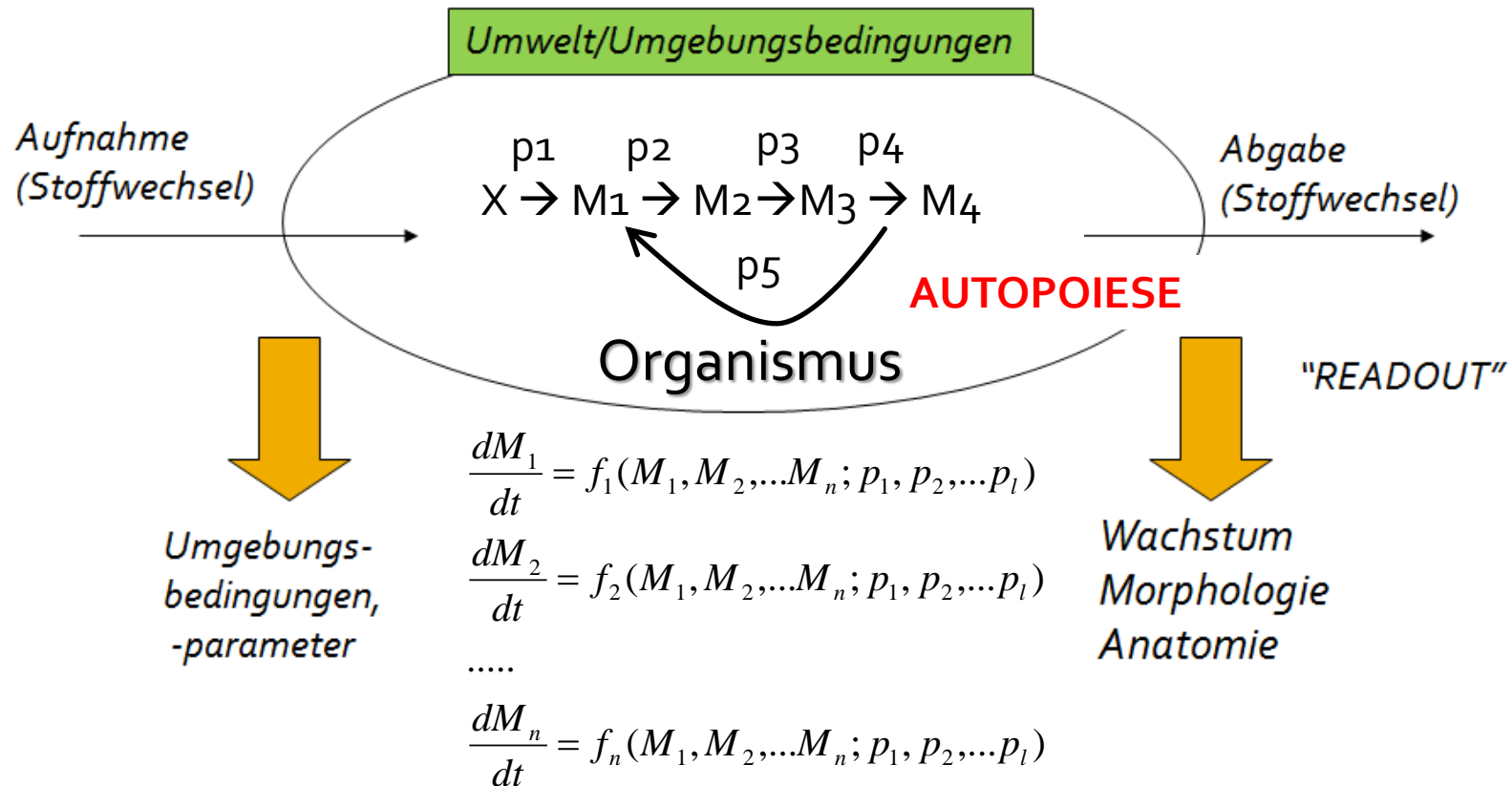
# Ludwig von Bertalanffy

PANTA RHEI! "Alles fließt!" HERAKLIT

1. Lebensformen sind **offene Systeme** und im Massenwirkungsgleichgewicht mit ihrer Umgebung (Fließgleichgewicht, stationärer Zustand (steady state), Stabilität)
2. "Allgemeine Systemtheorie"
  - kann auf alle komplexen – nichtlinearen Systeme angewendet werden (Biologie, Ökologie, Soziologie, Ökonomie, etc.)

# Der Organismus als physikalisches System betrachtet

Fließgleichgewicht = stationärer Zustand = Stabilität



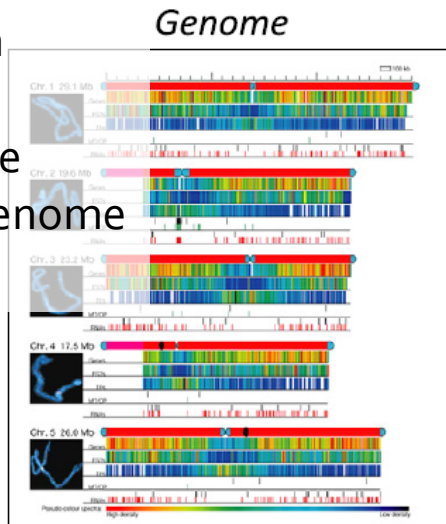
# Limitierungen Bertalanffys

- keine Genomsequenz – molekulare Aufklärung der Erbinformationskodierung erst 1953
- Organismen können nur durch numerische Simulation approximiert werden
- heutzutage möglich => Computerpower!!!
- das molekulare „Readout“ fehlt

# Genome-scale metabolic reconstruction

## Genome-scale metabolic reconstruction

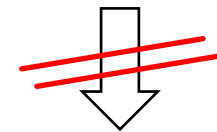
Human  
Plant  
Microbe  
Metagenome  
etc...



Weckwerth (2011)

**Unpredictability of  
metabolism.**

*Analytical and  
Bioanalytical Chemistry*



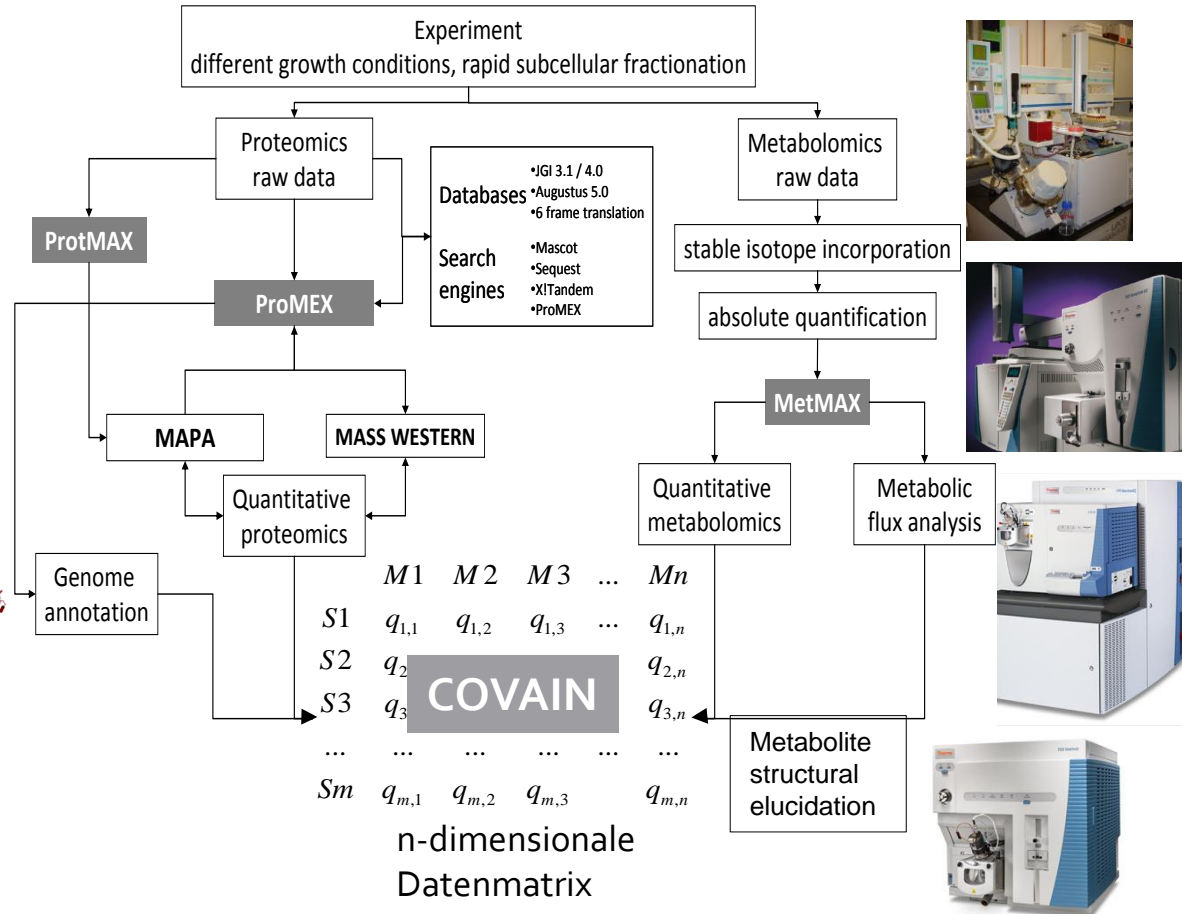
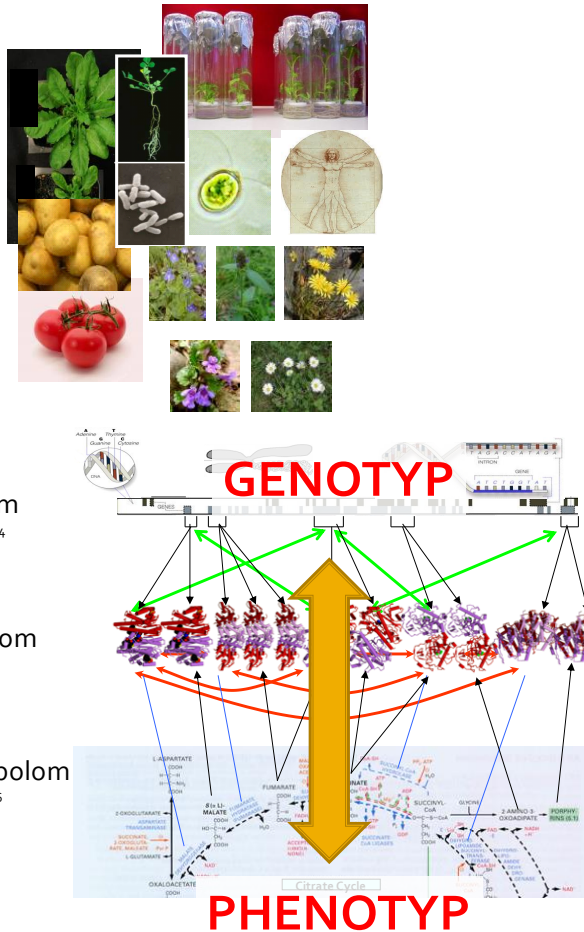
**Phenotypic plasticity**

# Systems Biology

- Genome sequencing
- Computer Science
  - => Statistics
  - => Modelling
  - => Internet
- OMICS
- Phenotype
  - => readout

Iteratively combining experiment (molecular data, phenotypical data) with modelling and prediction

# Integrative Metabolomics/Proteomics/ Phosphoproteomics platform links **static** Genotype and **dynamic** Phenotype



Weckwerth (2011) Green Systems Biology- from single genomes, proteomes and metabolomes to ecosystems research and biotechnology . J Proteomics

# PROTEOMICS

---



# FIELDOMICS

Hoehenwarter et al. 2008 Proteomics



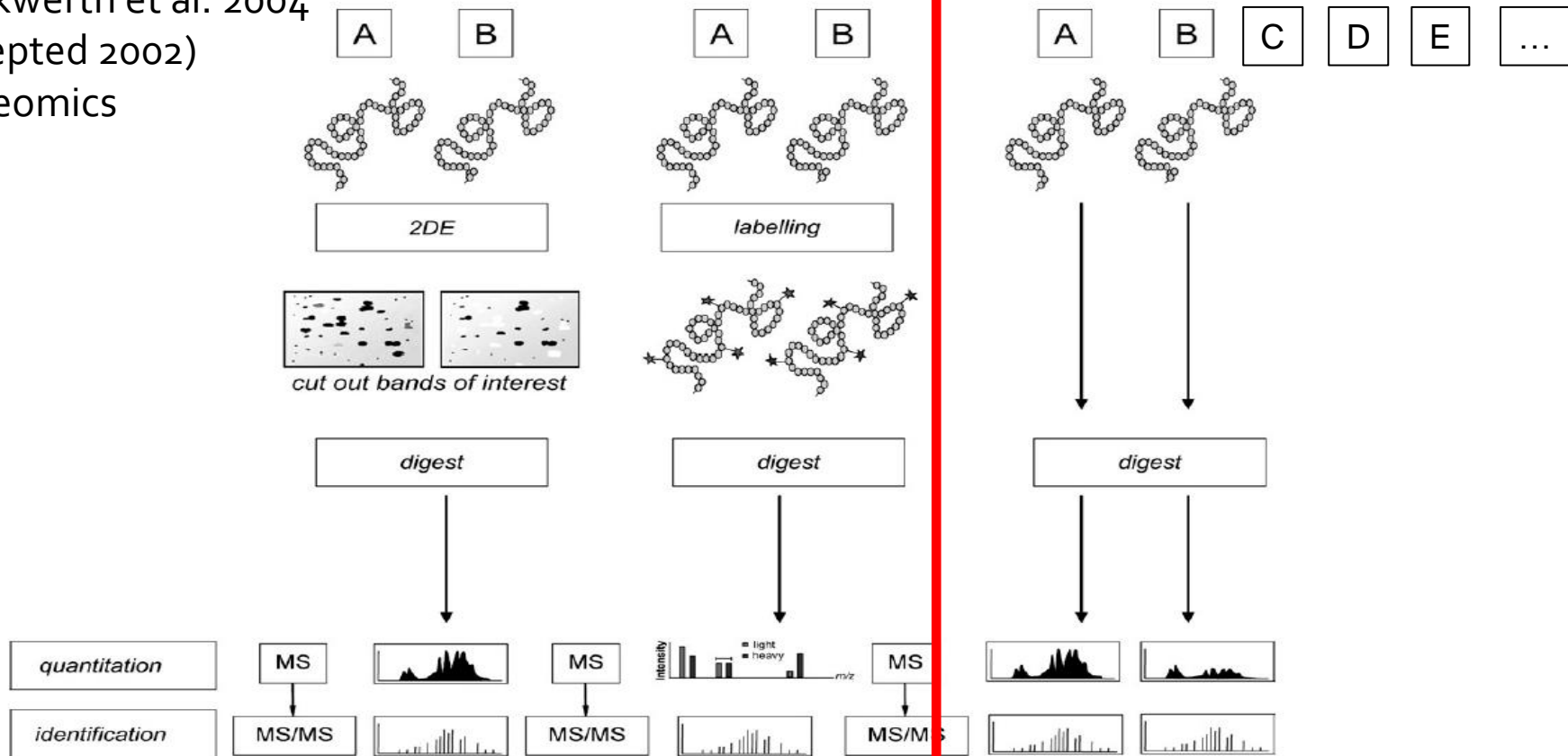
# 12 Potato Cultivars



- Mature tubers of tetraploid cultivars of *Solanum tuberosum*:
  - Ag, Agria, Al, Alliance; Ar, Arnika; F, Festien; G, Goldika; K, Kuras; LC, Lady Claire; Ma, Marabel; Mi, Milva; O, Omega; RF, Red Fantasy; T, Topas
- Grown on 3 plots each on 2 geographically separated fields in Germany
- 4 to 6 tubers of each cultivar were analyzed each three times = 216 shotgun proteomics analyses

# Non-gel label-free quantitative shotgun proteomics combined with multivariate statistics

Weckwerth et al. 2004  
(accepted 2002)  
Proteomics



- **Protein profiling with a high sample throughput**

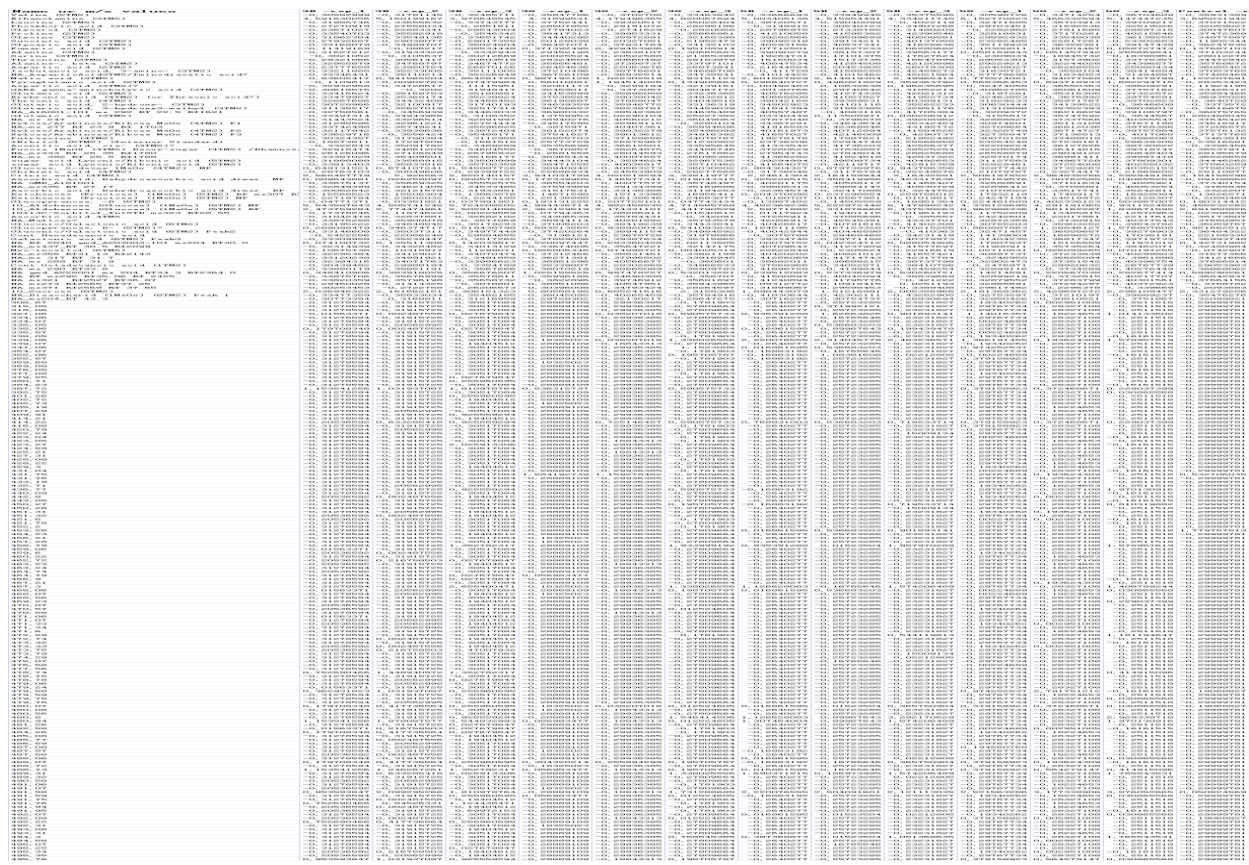
# Data Matrix

<-12 cultivars -> j = 216

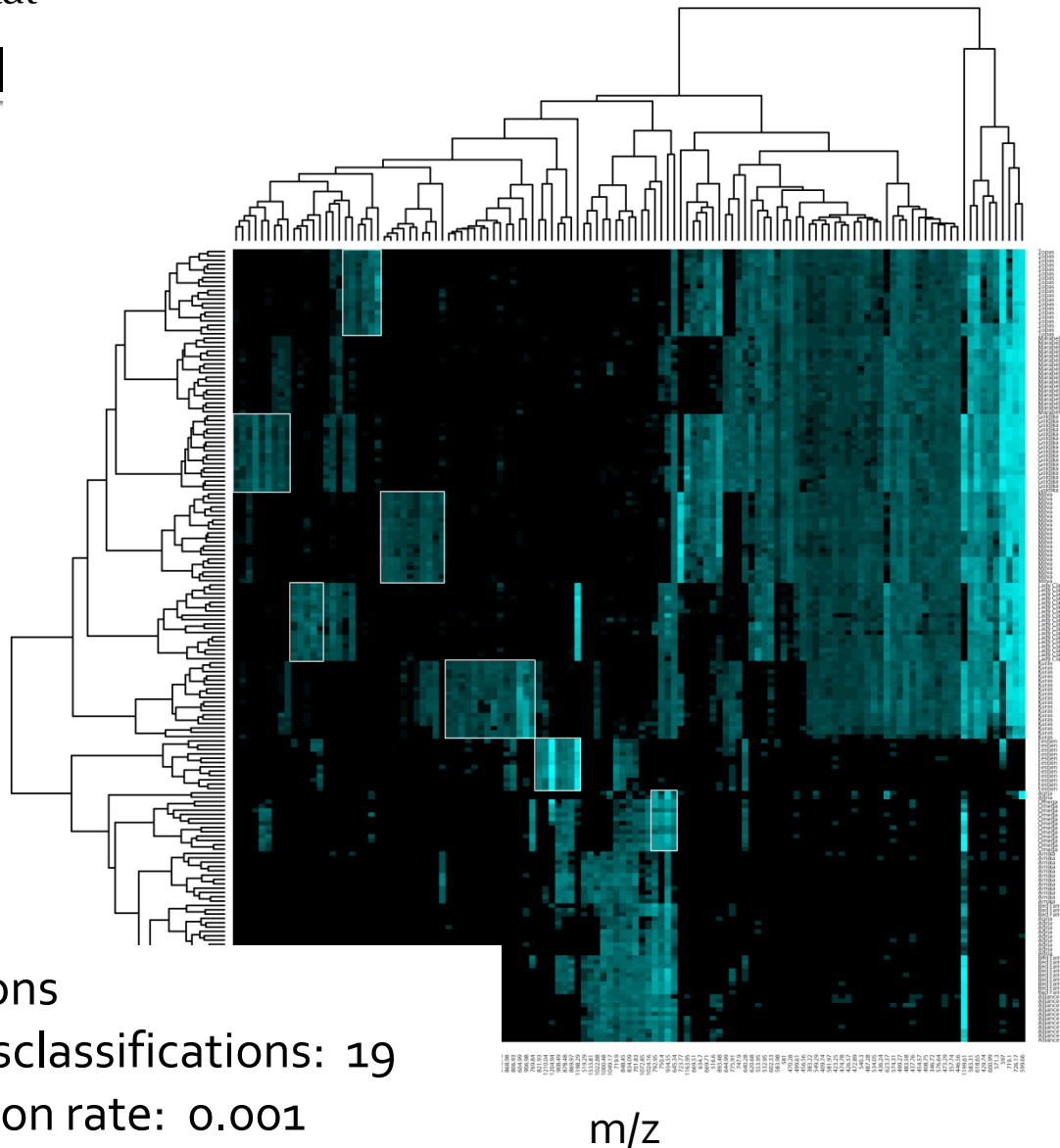
Analysis Binned m/z 0.01 Da		GOLDIKA ..... AGRIA			$\Sigma$
		Analysis 1	...	Analysis j	
Peptide 1	m/z 1	SC# 11	..	SC# 1j	SC# 1.
..	:	:	⊖	:	:
Peptide i	m/z i	SC# i1	..	SC# ij	SC# i.
	$\Sigma$	SC# .1	..	SC# .j	SC# ..

$i = 10.000$

Binned m/z 0.01 Da	Analysis 1	..	Analysis j	$\Sigma$
m/z 1	SC#11	..	SC#1j	SC#1.
:	:	⊛	:	:
m/z i	SC#i1	..	SC#ij	SC#i.
$\Sigma$	SC#.1	..	SC#.j	SC#..



# Cultivar Markers



- Samples
- Topas
  - Marabel
  - Goldika
  - Milva
  - Lady Claire
  - Kuras
  - Festien
  - Omega
  - Arnika
  - Agria
  - Red Fantasy
  - Alliance

18300 predictions  
 number of missclassifications: 19  
 missclassification rate: 0.001

# Protein Marker? VIP versus ICA

Independent component	Top 50 precursor ion <i>m/z</i>		Protein accession number	Protein name
	Identified	Not identified		
IC1	27	23	TA23344_4113	Patatin protein 07 ( <i>S. tuberosum</i> (potato))
			STRNA01	Patatin precursor ( <i>S. tuberosum</i> (potato))
			TA23357_4113	Patatin ( <i>S. tuberosum</i> (potato))
			TA23294_4113	Patatin ( <i>S. tuberosum</i> (potato))
			TA23358_4113	Patatin protein group A-3 ( <i>S. tuberosum</i> (potato))
			New patatin polymorphism <sup>a1</sup>	Previously undescribed Patatin isoform
			TA23798_4113	Kunitz-type protease inhibitor ( <i>S. tuberosum</i> (potato))
			TA23796_4113	Putative Kunitz-type proteinase inhibitor ( <i>S. tuberosum</i> (potato))
			CV286572	Serine protease inhibitor 5 precursor ( <i>S. tuberosum</i> (potato))
			TA24005_4113	Wound-inducible proteinase inhibitor I ( <i>S. tuberosum</i> (potato))
			TA24257_4113	Aspartic protease ( <i>Brassica napus</i> (Rape))
			TA23378_4113	Putative Kunitz-type tuber invertase inhibitor ( <i>S. tuberosum</i> (potato))
			TA24154_4113	Lipoxygenase ( <i>S. tuberosum</i> (potato))
			TA24161_4113	Lipoxygenase ( <i>S. tuberosum</i> (potato))
			STLOX1	Lipoxygenase 1 ( <i>S. tuberosum</i> (potato))
			TA24157_4113	5-Lipoxygenase ( <i>S. tuberosum</i> (potato))
			TA24257_4113	Enolase ( <i>Lycopersicon esculentum</i> (tomato))

## Rohbrei

- 1 Serine protease inhibitor 5 precursor [Solanum tuberosum (Potato)]
- 2 Lipoxygenase [Solanum tuberosum (Potato)]
- 3 Kunitz-type protease inhibitor [Solanum tuberosum (Potato)]
- 4 Patatin protein 07 [Solanum tuberosum (Potato)]
- 5 Patatin protein group A-3 [Solanum tuberosum (Potato)]
- 6 Patatin [Solanum tuberosum (Potato)]
- 7 Patatin precursor [Solanum tuberosum (Potato)]
- 8 Alpha-1,4 glucan phosphorylase, L-1 isozyme, chloroplast precursor [Solanum tuberosum (Potato)]
- 9 Putative Kunitz-type proteinase inhibitor [Solanum tuberosum (Potato)]

## Black spot

- 1 Patatin [Solanum tuberosum (Potato)]
- 2 Patatin protein group A-3 [Solanum tuberosum (Potato)]
- 3 Serine protease inhibitor 5 precursor [Solanum tuberosum (Potato)]
- 4 Patatin precursor [Solanum tuberosum (Potato)]
- 5 Putative Kunitz-type proteinase inhibitor [Solanum tuberosum (Potato)]
- 6 Patatin protein 07 [Solanum tuberosum (Potato)]
- 7 Kunitz-type protease inhibitor [Solanum tuberosum (Potato)]
- 8 Lipoxygenase [Solanum tuberosum (Potato)]
- 9 Aspartic protease inhibitor precursor

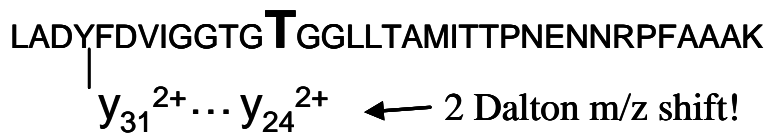
## Chips

- 1 Patatin [Solanum tuberosum (Potato)]
- 2 Patatin protein 07 [Solanum tuberosum (Potato)]
- 3 Patatin precursor [Solanum tuberosum (Potato)]
- 4 Putative Kunitz-type proteinase inhibitor [Solanum tuberosum (Potato)]
- 5 Lipoxygenase [Solanum tuberosum (Potato)]
- 6 Serine protease inhibitor 5 precursor [Solanum tuberosum (Potato)]
- 7 Patatin protein group A-3 [Solanum tuberosum (Potato)]
- 8 Enolase [*Lycopersicon esculentum* (Tomato)]
- 9 Kunitz-type protease inhibitor [Solanum tuberosum (Potato)]

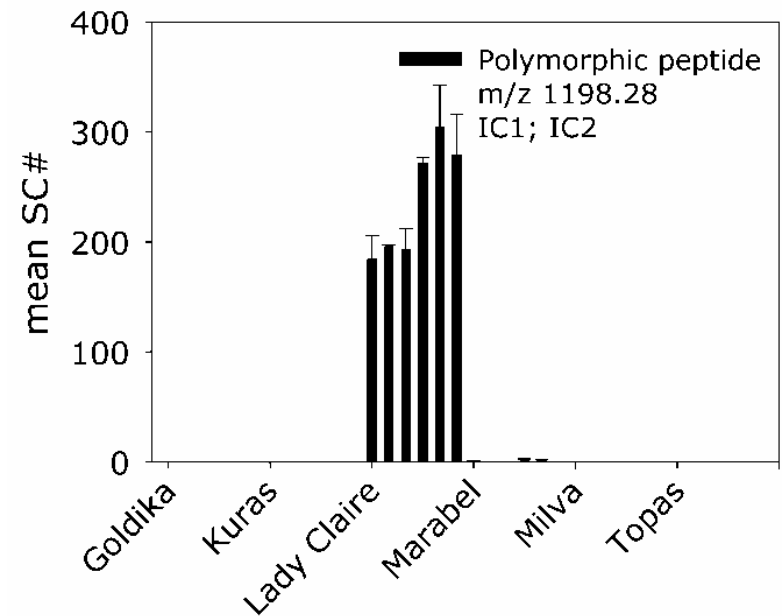
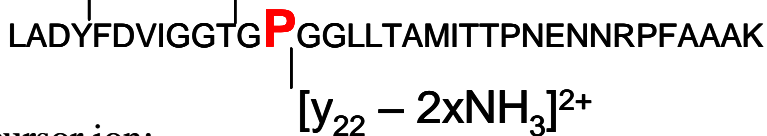
# Database-independent identification of cSNP/Protein polymorphism

m/z error < 2 ppm

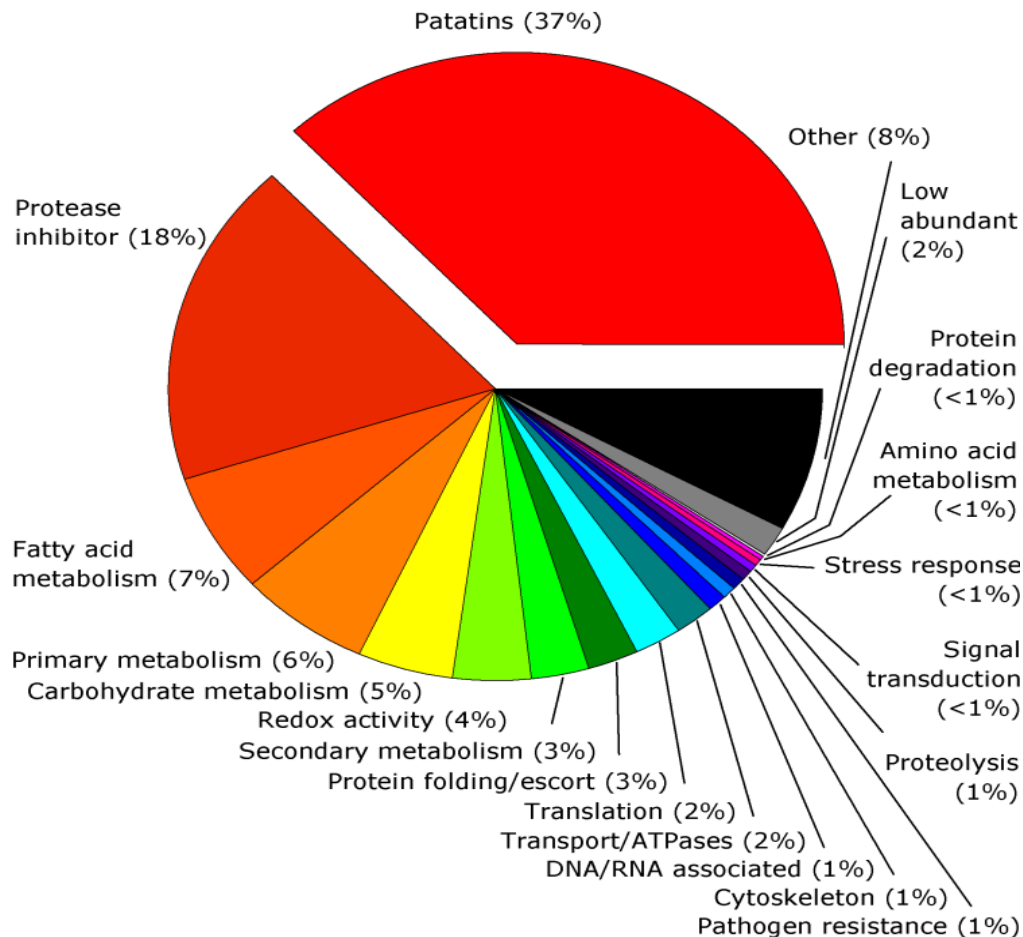
Precursor ion:  
m/z 1199.6112 (3+)



Precursor ion:  
m/z 1198.2795 (3+)



# Quantitative Functional Annotation of the Tuber Proteome



~ 4000 proteins  
=> Largest  
Potato  
Proteome to  
date

Hoehenwarter et al. 2011 Journal of  
Proteome Research



# PROMEX

## – Microbial, Plant and Animal Proteomics data base

<http://promexdb.org>

Hummel et al. 2007 BMC Bioinformatics

Egelhofer et al. 2012 Frontiers of Plant Proteomics

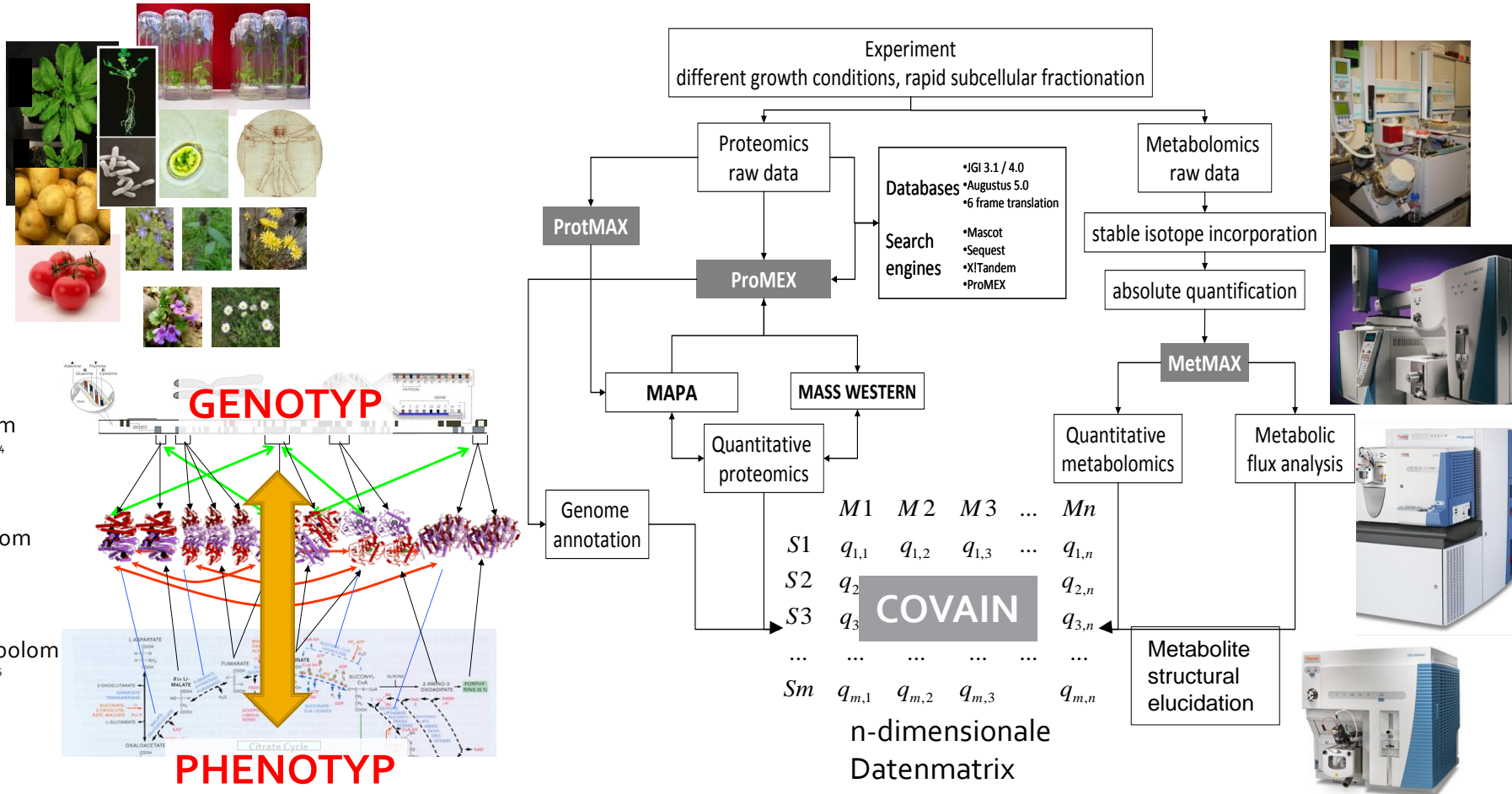
# Plant and microbial proteomics - ProMEX

(<http://promexdb.org>)



from 10-40% Proteome coverage

# Integrative Metabolomics/Proteomics/ Phosphoproteomics platform links **static** Genotype and **dynamic** Phenotype



Weckwerth (2011) Green Systems Biology- from single genomes, proteomes and metabolomes to ecosystems research and biotechnology . J Proteomics

# Metabolomics

---

# Erwin Schroedinger (1943)

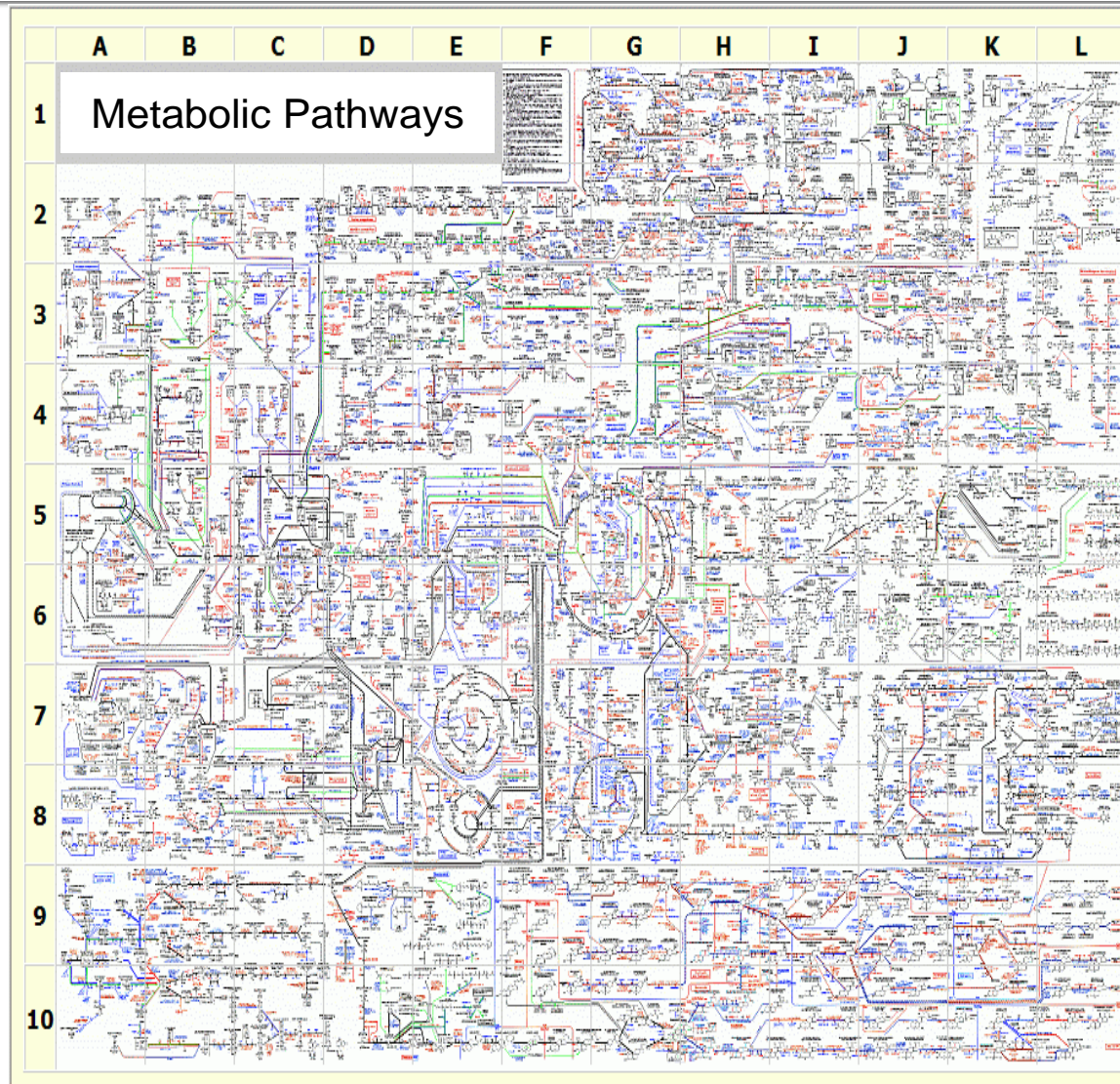
## What is life?

How can the organism escape death (second law of thermodynamics – maximal entropy)?

The answer is: through nutrition and respiration, in case of plants by assimilation. The terminus is

**Metabolismus.**

# Metabolomics

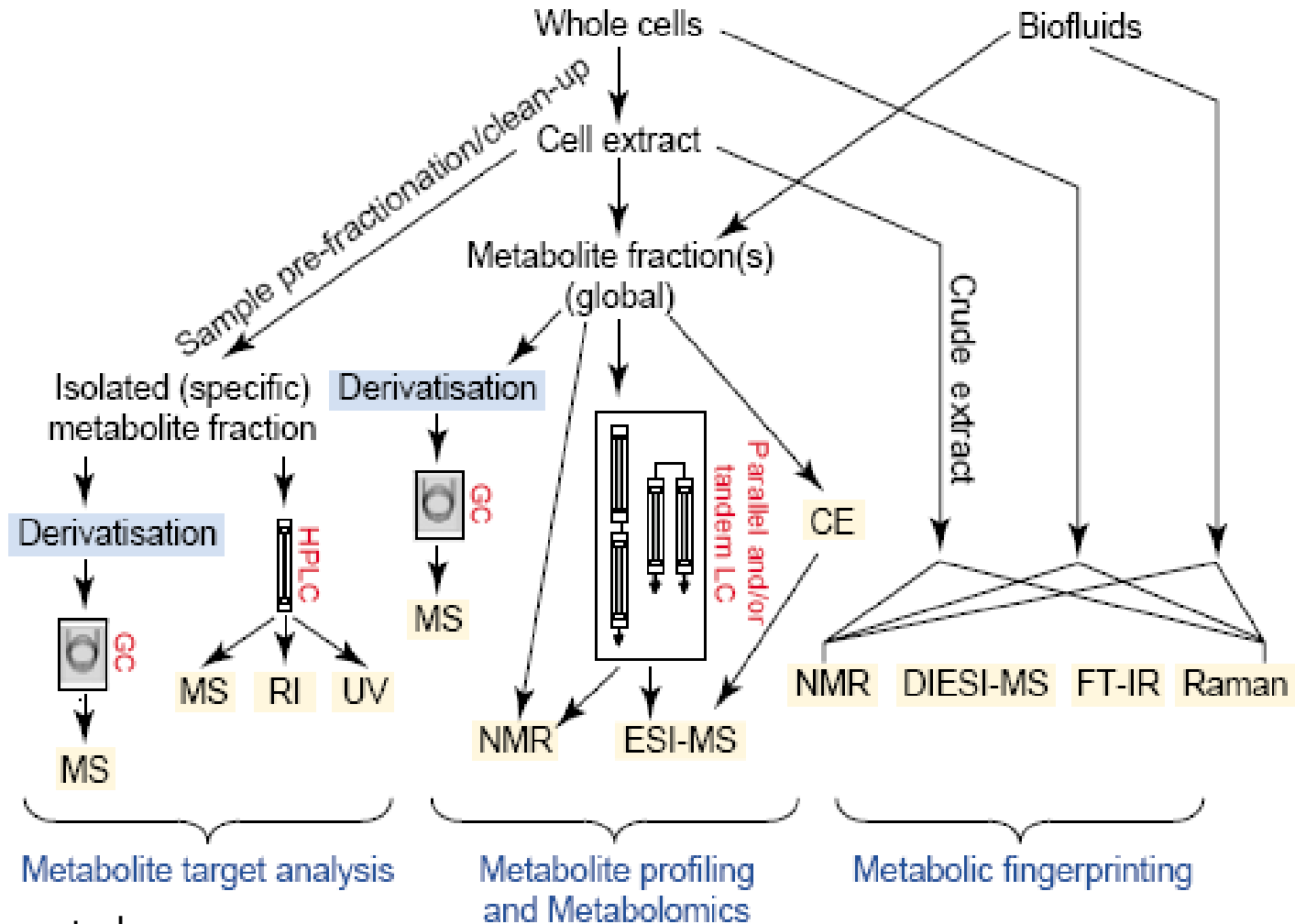


## Metabolomics

-non-targeted  
metabolite analysis  
- integrative part of  
systems biology

Weckwerth (2003)  
Metabolomics in  
Systems Biology.  
Annual Review of  
Plant Biology

# Metabolic fingerprinting - Metabolite profiling - Metabolomics



# APPLICATIONS OF Metabolomics

**HUMAN WELL-BEING**

HEALTH & NUTRITION

**CELL LINE ENGINEERING**

INDUSTRIAL BIOTECHNOLOGY

**PHARMACEUTICAL  
DEVELOPMENT**

HEALTH & NUTRITION

**BREEDING**

FOOD & AGRICULTURE

**COMPOUND IDENTIFICATION**

HEALTH & NUTRITION, INDUSTRIAL BIOTECHNOLOGY

**FLAVOR PROFILER™**

FOOD & AGRICULTURE

**DIAGNOSTICS**

HEALTH & NUTRITION

**CROP PROTECTION**

FOOD & AGRICULTURE

**PERSONALIZED MEDICINE**

HEALTH & NUTRITION

**TRANSLATIONAL RESEARCH**

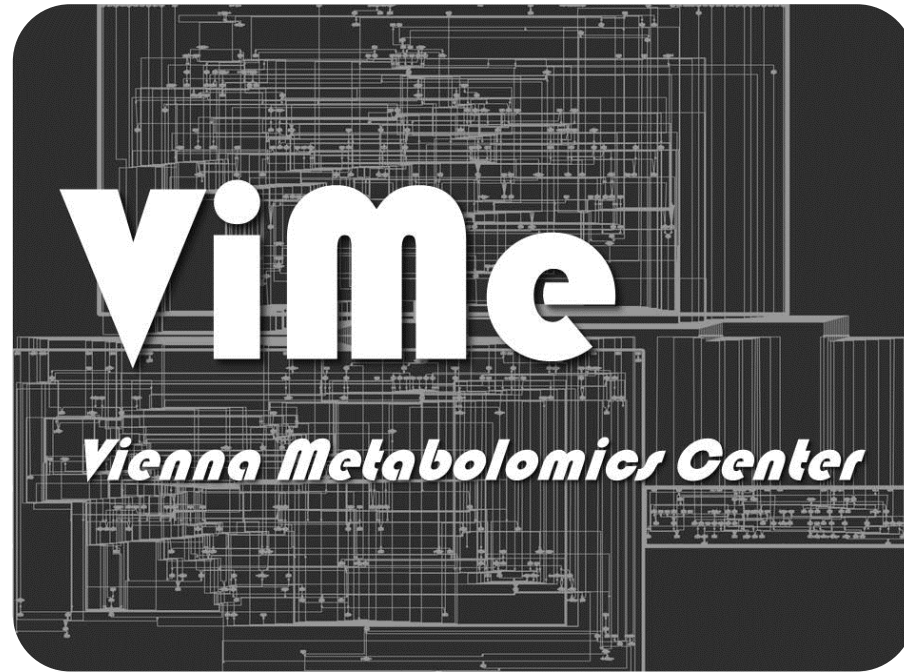
HEALTH & NUTRITION

BIO-BASED ECONOMY

PERSONALIZED MEDICINE



# Vienna Metabolomics Center



- Research Platform University of Vienna
- Coordinator: Univ.-Prof. Dr. Wolfram Weckwerth
- Faculty of Life Sciences



universität  
wien

# Vienna Metabolomics Consortium

- Faculty of Life Sciences
  - Verena Dirsch
  - Gerhard J. Herndl
  - Walter Jäger
  - Jürgen König
  - Thierry Langer
  - Christa Schleper
  - Karl-Heinz Wagner
- Faculty of Chemistry
  - Christopher Gerner
  - Gunda Koellensperger
  - Doris Marko
  - Veronika Somoza
- Faculty of Geosciences
  - Daniel Birgel
  - Stephan Krämer
  - Jörn Peckmann

# Impact of Genetic Polymorphisms on the metabolic response

Doris Marko

Department of Food Chemistry and Toxicology, Faculty of Chemistry

## SNP

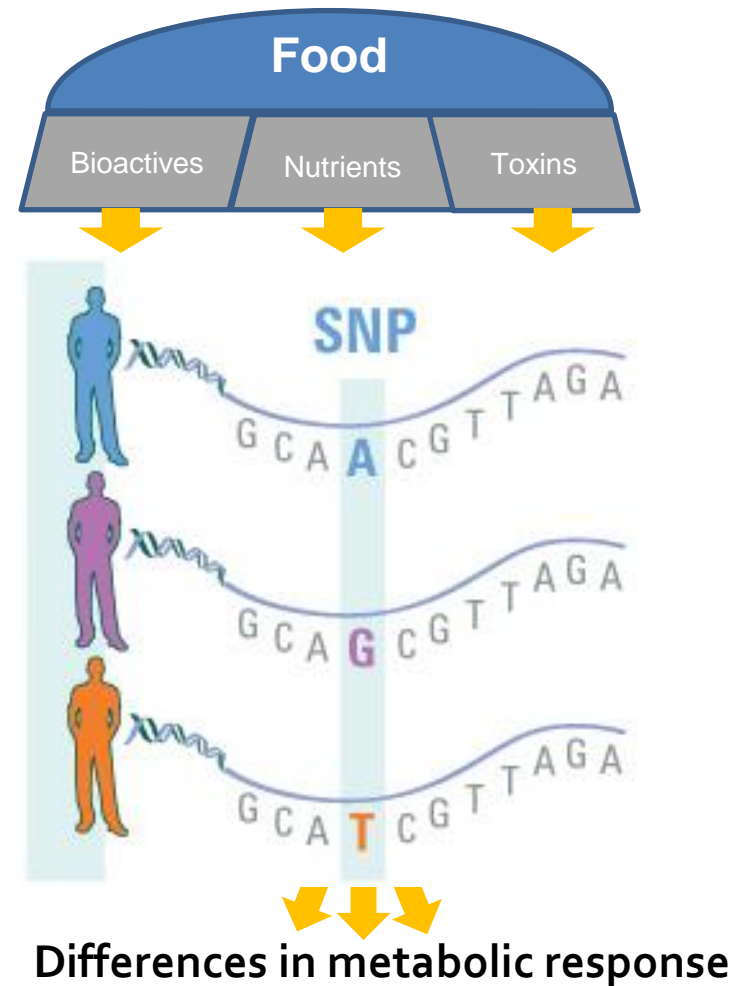
- Single nucleotide polymorphism
- Found in 90% of all human genes
- Potential impact in chemoprevention and disease/therapy

## Initial results:

Correlation between presence of SNPs in Nrf2/Nrf2-dependent genes and the transcriptional response to coffee consumption (Böttler et al., MetaGene, 2014)

## Aim:

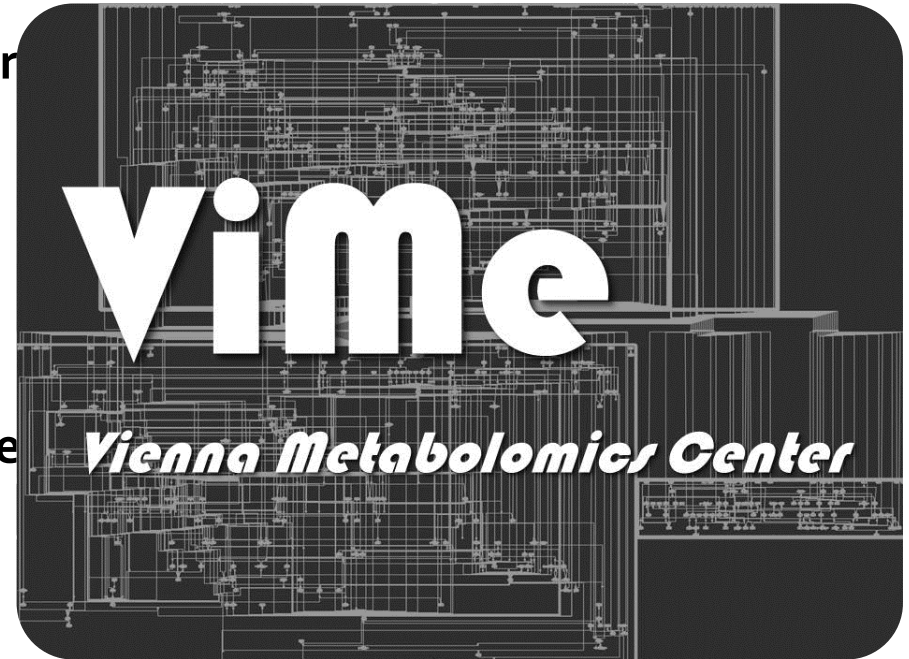
Nutrigenomic approach investigating the impact of genotype on the metabolome in response to a) functional food constituents (e.g. anthocyanins), b) contaminants (e.g. mycotoxins) and combinations thereof, focussing primarily on genes involved in cellular defense and antioxidative mechanisms



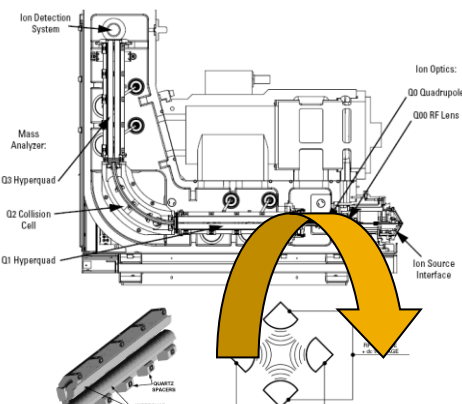
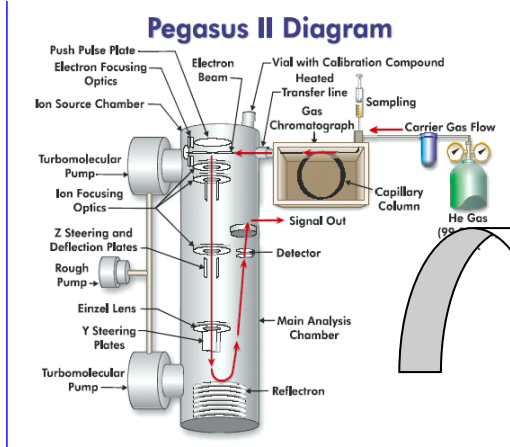
# Vienna Metabolomics Center

- ViMe – Vienna Metabolomics Center – is an open research platform for scientific collaborations
- Provides a full suite for metabolomic profiling combining GC-MS, LC-MS, structural elucidation, databases, multivariate statistics, data integration and metabolic modelling
- Provides a full suite for proteomic/phosphoproteomic profiling
- Contact:

[wolfram.weckwerth@univie.ac.at](mailto:wolfram.weckwerth@univie.ac.at)



# Metabolomics based on GC-MS



aliphatics  
alcohols  
acids

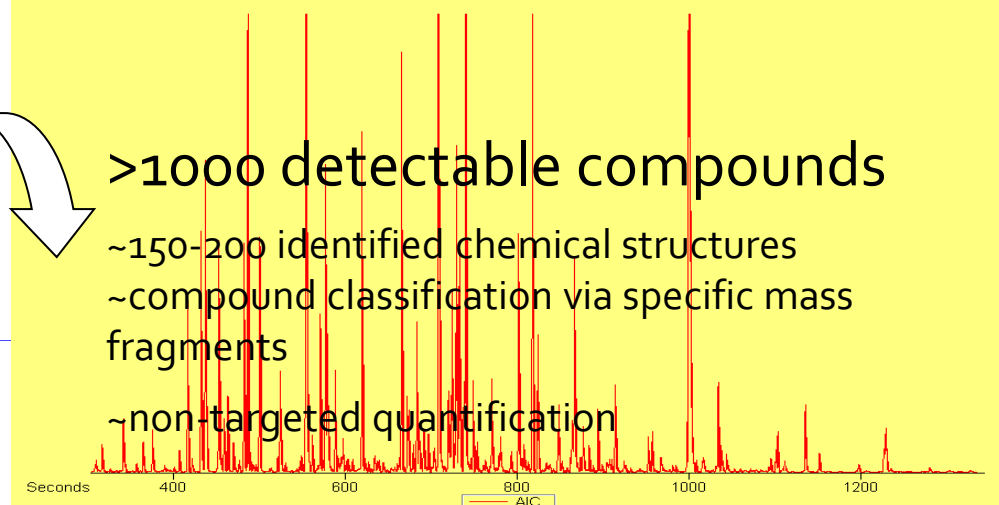
HO-acids  
NH<sub>2</sub>-acids

monosacch.  
fatty acids

sugar  
~phosph.  
~alcohols

disacch.  
HO-fatty acids

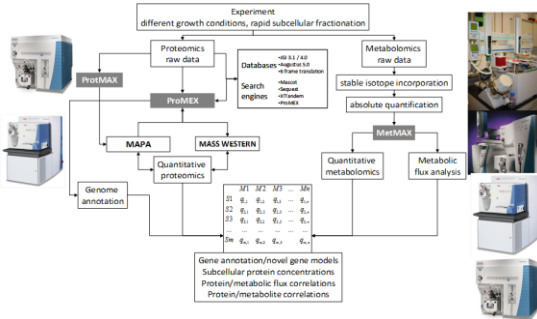
trisacch.  
sterols



Weckwerth et al. 2001 Weckwerth et al. 2004 Morgenthal et al. 2005, 2007  
 Boldt et al. 2005 Wienkoop et al 2008 Larrainzar et al 2009, Kempa et al  
 2009 Scherling et al 2010 Doerfler et al. 2013 Mari et al 2013, Naegele et al.  
 2014, Valledor et al. 2014, Albacete et al. 2015, Mair et al. 2015, Stare et al.  
 2015, Hasler-Sheetal et al. 2015, Nagler et al. 2015, Meijon et al. 2016,  
 Kogovsek et al. 2016, and many more ....

Comprehensive Identification and quantitation in short time  
 => „throughput“, many samples, time courses, statistics  
 Overall rel. standard deviation: 10-30 %

# Increasing metabolome coverage



Metabolomics -  
untargeted approach

GC/MS

LC/MS

GC-TOF-MS

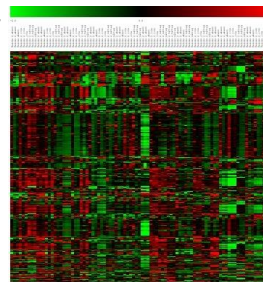
LC-FT-MS

~20000 MST

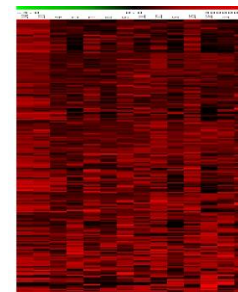
~20000  
features

smaller metabolites  
(sugars, amino  
acids, organic acids,  
others)  
lipids

larger  
metabolites  
(lipids,  
flavonoids,  
anthocyanines,  
others)



Spectral annotation



Sum formula and  
database search

Metabolite identification,  
sample pattern recognition and biological  
interpretation

# MAPA

<http://www.univie.ac.at/mosys/software.html>

The screenshot shows a web browser window with the URL `promex.pph.univie.ac.at/protmax/`. The page features the **ProtMAX** logo and a download link for ProtMAX2012 rev 2.12. Below the link, the system requirements are listed: a Pentium-compatible PC, Microsoft Windows (XP, Win7), 2 GB RAM minimum, and 4 GB RAM recommended. A 'Getting Started' section provides three steps: downloading the software, uncompressing it, and starting the installation. Contact information for Volker Egelhofer and Stefanie Wienkoop is provided at the bottom.

The ProtMAX 2012 preferences dialog box is open, showing the following settings:

- Main:** Method: Spectral Count; Quantification: Count; Decimals: 2; Cut: Cut
- Filter:** Charge states:  1+  2+  3+  4+  5+  6+; Environment:  +/- 1.0 RT [min]; Unite Neighbors:
- RT Cut off:** From: lowest; To: highest
- Output:** C:\Users\woweck\AppData\Local\Apps\2.0\Data\GAA41KAV;  Launch Excel

The status bar at the bottom of the dialog box displays: `Method Spectral Count Quantification Count Environment RT [min] +/- 1 Charge states any Decimals 2 RT full range`

# Extraction and Derivatisation

10-20mg homogenised FW

+ 1mL pre-chilled (-20°C) extraction mixture of  
MeOH : CHCl3 : H2O (2 : 1 : 1, v/v/v)  
+ 10µL Internal Standard (2mg/mL)

Shake for 10 min at 4°C

Centrifuge at 14,000 rpm for 5min



**Supernatant, contains soluble metabolites**

**Pellet, contains insoluble metabolites, cell walls, RNA, DNA and Proteins**

Human, Animal, Plant, Microbes, environmental samples

Methanol/H<sub>2</sub>O

Chloroform



Methoxyamine/MSTFA

TMSH



**GC/MS**

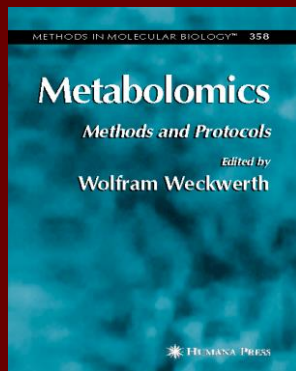
AA, sugars, organic acids, etc.

FAME

Shotgun proteomics  
Shotgun phosphoproteomics  
RNAseq  
DNA/Cell wall

Weckwerth et al. 2004 Proteomics





„Metabolomics: Methods and Protocols.“ Humana Press  
Wolfram Weckwerth (ed) October 2006

GC-MS

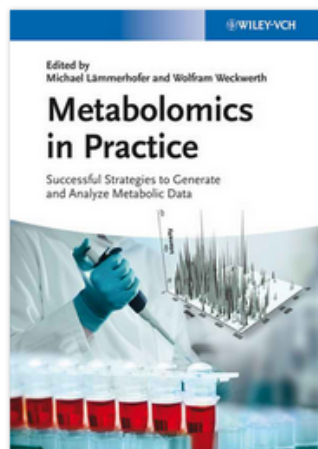
LC-MS

CE-MS

NMR

Data mining

Detailed protocols with tips & tricks

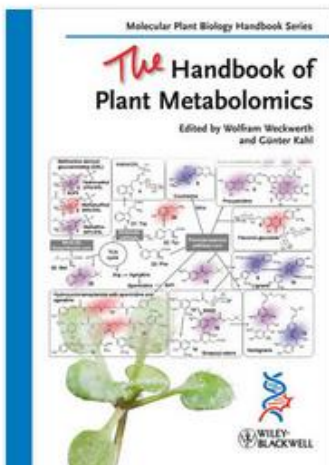


## Metabolomics in Practice: Successful Strategies to Generate and Analyze Metabolic Data

Michael Lämmerhofer (Editor), Wolfram Weckwerth (Editor)

ISBN: 978-3-527-33089-8

Hardcover  
442 pages  
January 2013



## The Handbook of Plant Metabolomics

Wolfram Weckwerth (Editor), Gunter Kahl (Editor)

ISBN: 978-3-527-32777-5

Hardcover  
448 pages  
April 2013, Wiley-Blackwell

# Workflow Metabolomics

Weckwerth 2011 Unpredictability of Metabolism.ABC

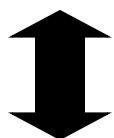


GC-MS  
 ▼  
 postprocessing  
 ▼  
 GC-PA

(extraction of RI, mass spectra, relative fragment intensities, sample versus variable alignment)

GMD

(metabolite database)



nanoLC-MS  
 ▼  
 mzXML  
 ▼  
 MAPA

(extraction of accurate precursor m/z, spectral count, sample versus variable alignment)



Data Matrix

	$M_1$	$M_2$	$M_3$	...	$M_n$
$S_1$	$q_{1,1}$	$q_{1,2}$	$q_{1,3}$	...	$q_{1,n}$
$S_2$	$q_{2,1}$	$q_{2,2}$	$q_{2,3}$	...	$q_{2,n}$
$S_3$	$q_{3,1}$	$q_{3,2}$	$q_{3,3}$	...	$q_{3,n}$
...	...	...	...	...	...
$S_m$	$q_{m,1}$	$q_{m,2}$	$q_{m,3}$	...	$q_{m,n}$

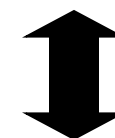
*n*-dimensional  
Data matrix

COVAIN

Sun & Weckwerth 2012 Metabolomics

MOSYS

(metabolite database)



sample pattern recognition/biological interpretation



# Data integration

---

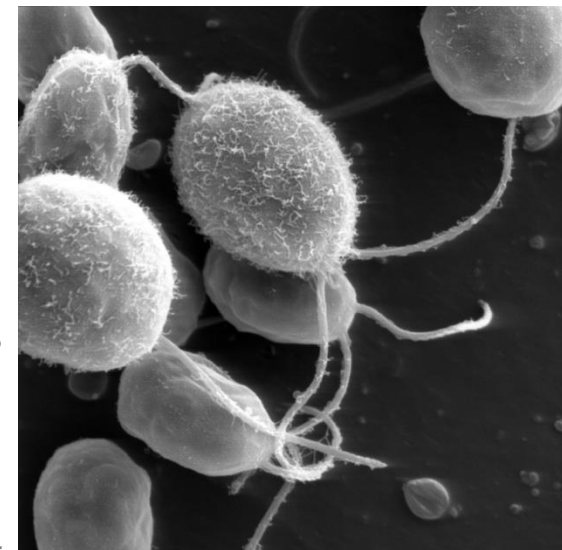
# *Chlamydomonas reinhardtii* – the “green yeast” – BIOFUEL THIRD GENERATION

RESEARCH ARTICLE

## The *Chlamydomonas* Genome Reveals the Evolution of Key Animal and Plant Functions

Sabeeha S. Merchant,<sup>1\*</sup> Simon E. Prochnik,<sup>2\*</sup> Olivier Vallon,<sup>3</sup> Elizabeth H. Harris,<sup>4</sup> Steven J. Karpowicz,<sup>1</sup> George B. Witman,<sup>5</sup> Astrid Terry,<sup>2</sup> Asaf Salamov,<sup>2</sup> Lillian K. Fritz-Laylin,<sup>6</sup> Laurence Maréchal-Drouard,<sup>7</sup> Wallace F. Marshall,<sup>8</sup> Liang-Hu Qu,<sup>9</sup> David R. Nelson,<sup>10</sup> Anton A. Sanderfoot,<sup>11</sup> Martin H. Spalding,<sup>12</sup> Vladimir V. Kapitonov,<sup>13</sup> Qinghu Ren,<sup>14</sup> Patrick Ferris,<sup>15</sup> Erika Lindquist,<sup>2</sup> Harris Shapiro,<sup>2</sup> Susan M. Lucas,<sup>2</sup> Jane Grimwood,<sup>16</sup> Jeremy Schmutz,<sup>16</sup> *Chlamydomonas* Annotation Team,† JGI Annotation Team,† Igor V. Grigoriev,<sup>2</sup> Daniel S. Rokhsar,<sup>2,6‡</sup> Arthur R. Grossman<sup>17‡</sup>

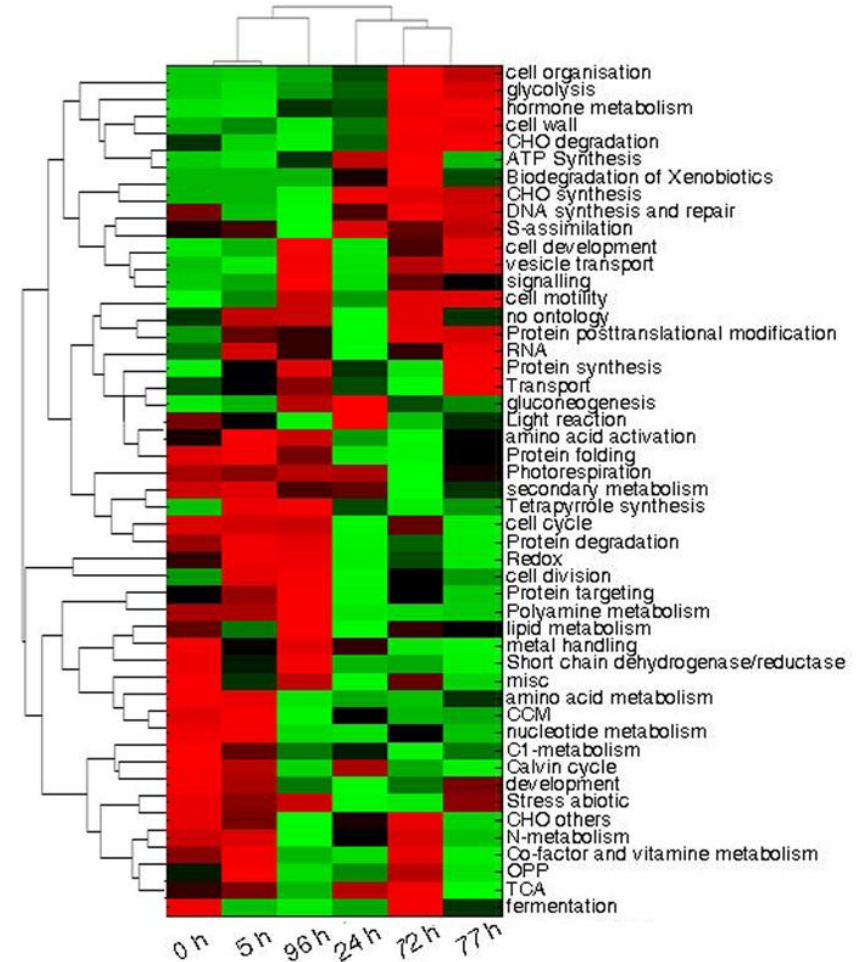
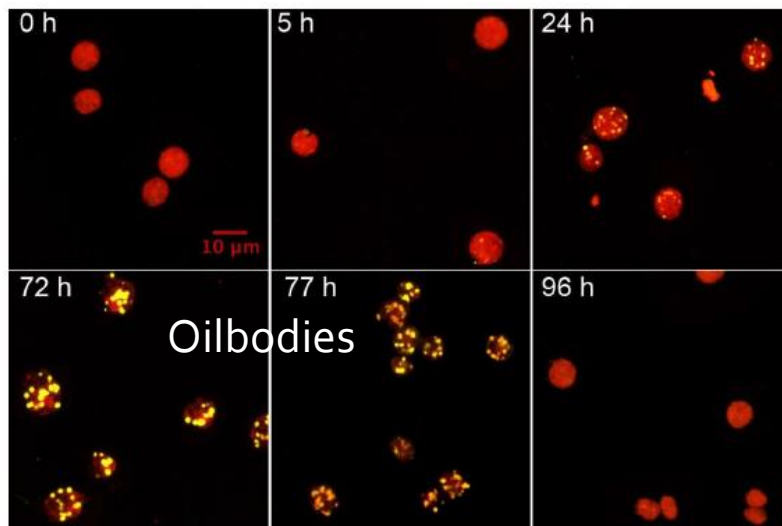
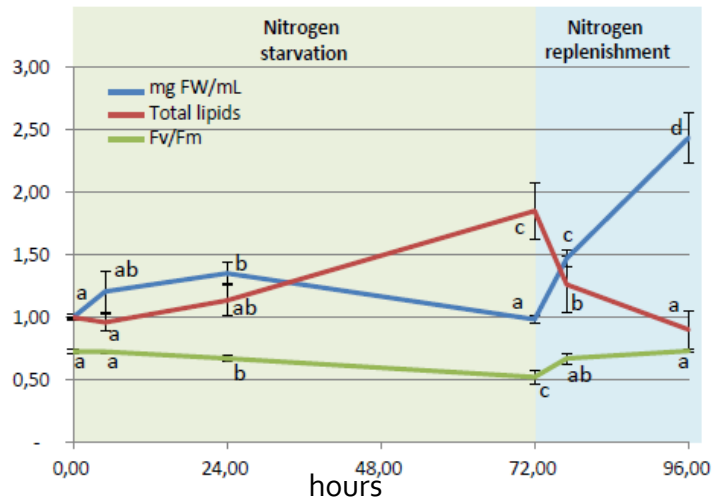
*Chlamydomonas reinhardtii* is a unicellular green alga whose lineage diverged from land plants over 1 billion years ago. It is a model system for studying chloroplast-based photosynthesis, as well as the structure, assembly, and function of eukaryotic flagella (cilia), which were inherited from the common ancestor of plants and animals, but lost in land plants. We sequenced the ~120-megabase nuclear genome of *Chlamydomonas* and performed comparative phylogenomic analyses, identifying genes encoding uncharacterized proteins that are likely associated with the function and biogenesis of chloroplasts or eukaryotic flagella. Analyses of the *Chlamydomonas* genome advance our understanding of the ancestral eukaryotic cell, reveal previously unknown genes associated with photosynthetic and flagellar functions, and establish links between ciliopathy and the composition and function of flagella.



**Genome sequencing and assembly.** The 121-megabase (Mb) draft sequence (10) of the *Chlamydomonas* nuclear genome was generated at 13× coverage by whole-genome, shotgun end-sequencing of plasmid and fosmid libraries, followed by assembly into ~1500 scaffolds (1). Half of the assembled genome is contained in 25 scaffolds, each longer than 1.63 Mb. The genome is unusually GC-rich (64%) (Table 1), which required modification of standard sequencing protocols. Alignment

org on December 14, 2007

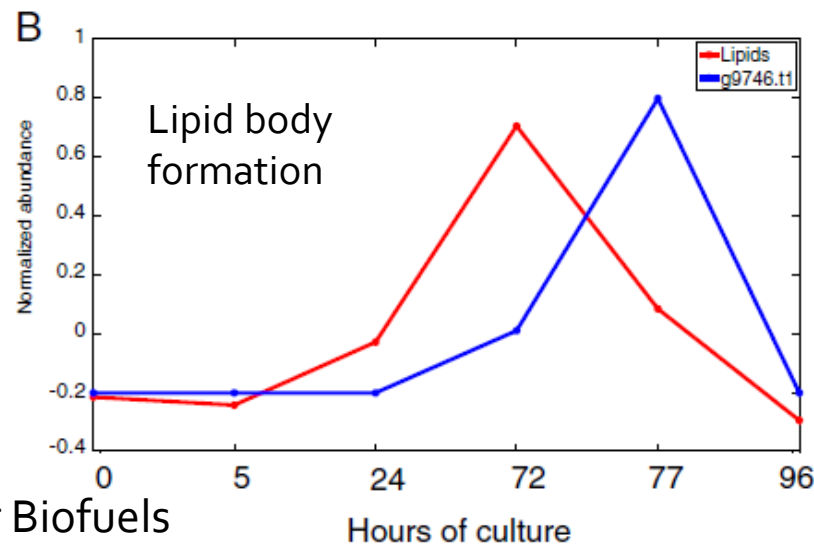
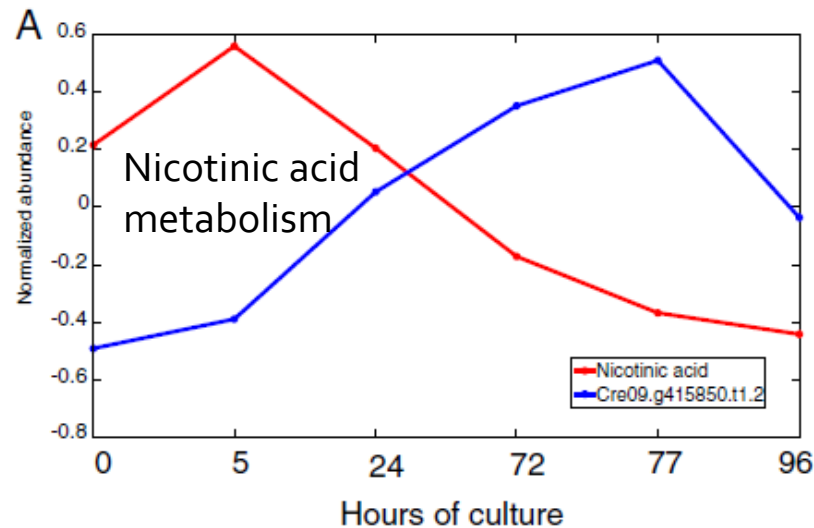
# Nitrogen starvation and recovery: formation of oil bodies and vegetative growth





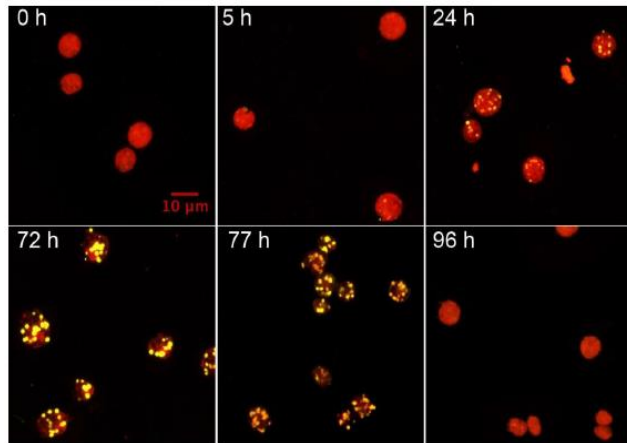
# Granger Causality during nitrogen starvation and recovery

- time-lagged correlation measures to identify potential cause–effect pairs
- Developed by economist and noble prize laureate Clive Granger
- Implemented in COVAIN

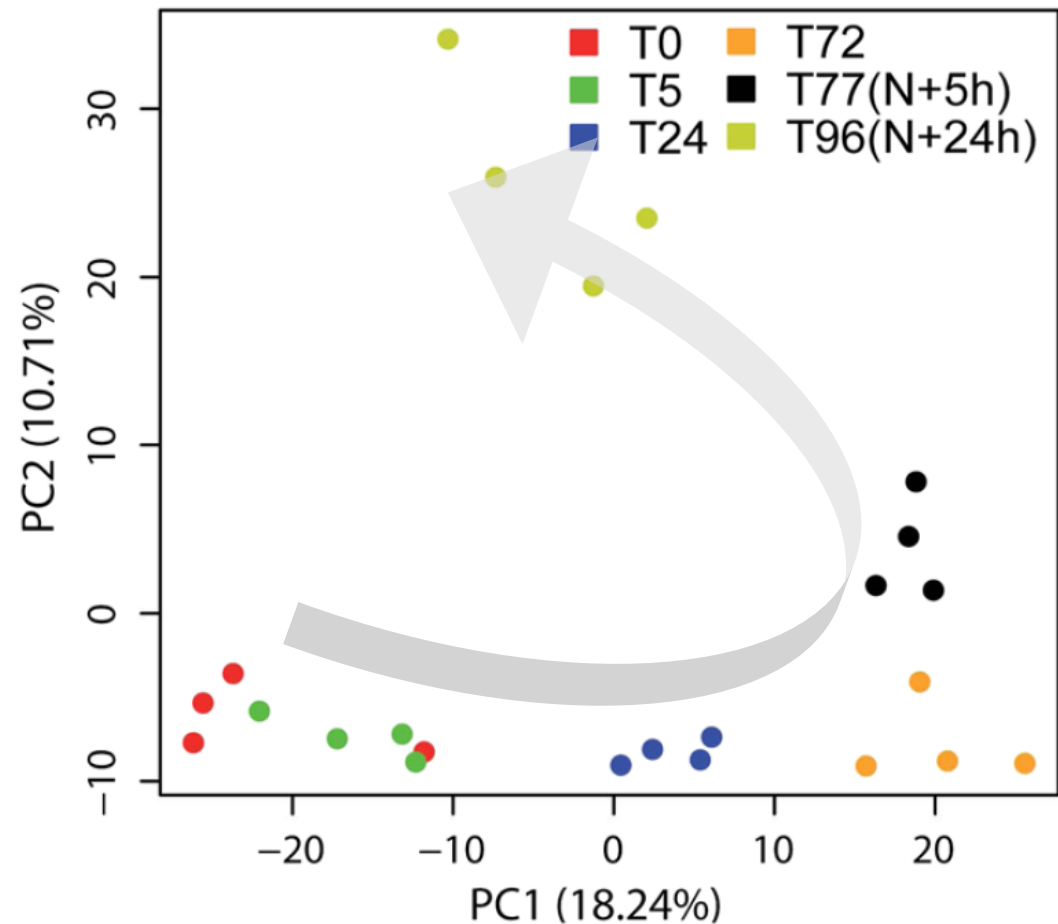




# Metabolic trajectory...

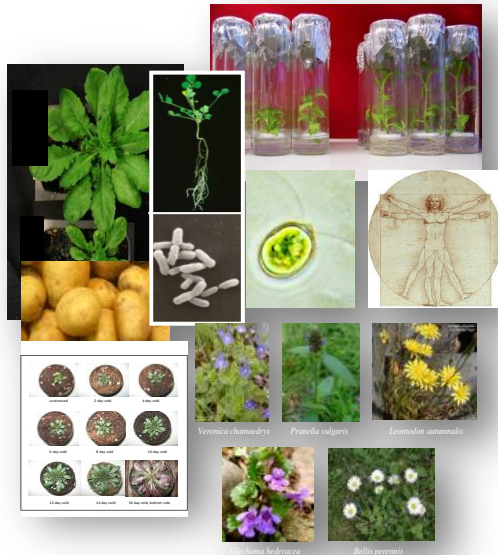


Valledor et al. 2014  
Biotechnology for  
Biofuels



# Is there a systematic relation?

Weckwerth (2011) *Green Systems Biology* Journal of Proteomics



	$M_1$	$M_2$	$M_3$	...	$M_n$
$S_1$	$q_{1,1}$	$q_{1,2}$	$q_{1,3}$	...	$q_{1,n}$
$S_2$	$q_{2,1}$	$q_{2,2}$	$q_{2,3}$	...	$q_{2,n}$
$S_3$	$q_{3,1}$	$q_{3,2}$	$q_{3,3}$	...	$q_{3,n}$
...	...	...	...	...	...
$S_m$	$q_{m,1}$	$q_{m,2}$	$q_{m,3}$	...	$q_{m,n}$



n-dimensional Data matrix

**PHENOTYPE**



Multivariate Statistics

**Trajectories**



$$\frac{dM_1}{dt} = f(M_1, M_2, \dots, M_n)$$

$$\frac{dM_2}{dt} = f(M_1, M_2, \dots, M_n)$$

.....

$$\frac{dM_n}{dt} = f(M_1, M_2, \dots, M_n)$$

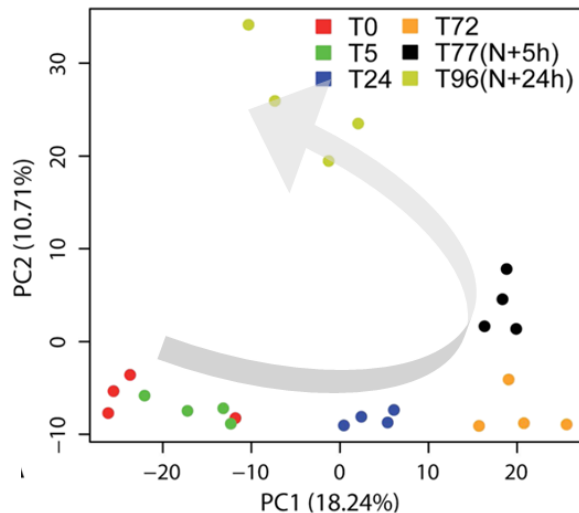
Systems equations

**Systems theory**



Genome sequence

**GENOTYPE**



Weckwerth 2011 Unpredictability of Metabolism. ABC  
Weckwerth 2011 Green Systems Biology

# **A system-theoretical derivation of a geno-phenotype-equation**

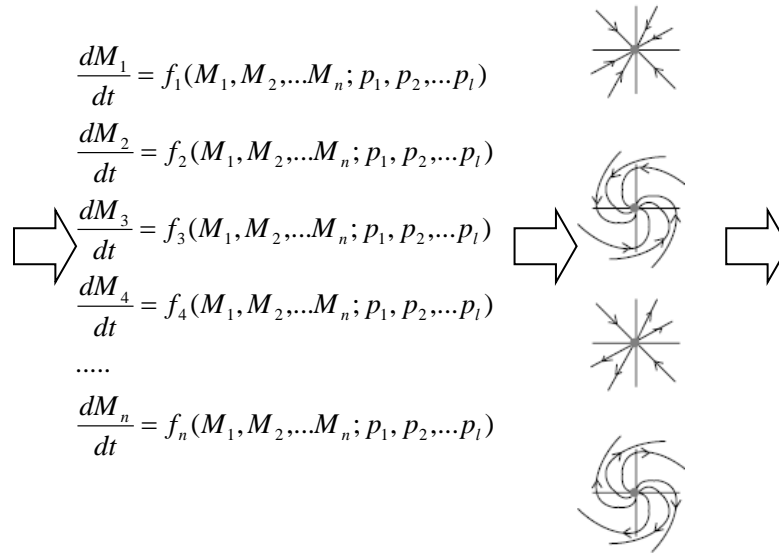
---

# Mathematical description of a complex nonlinear system and its trajectories...

complex system



stability analysis



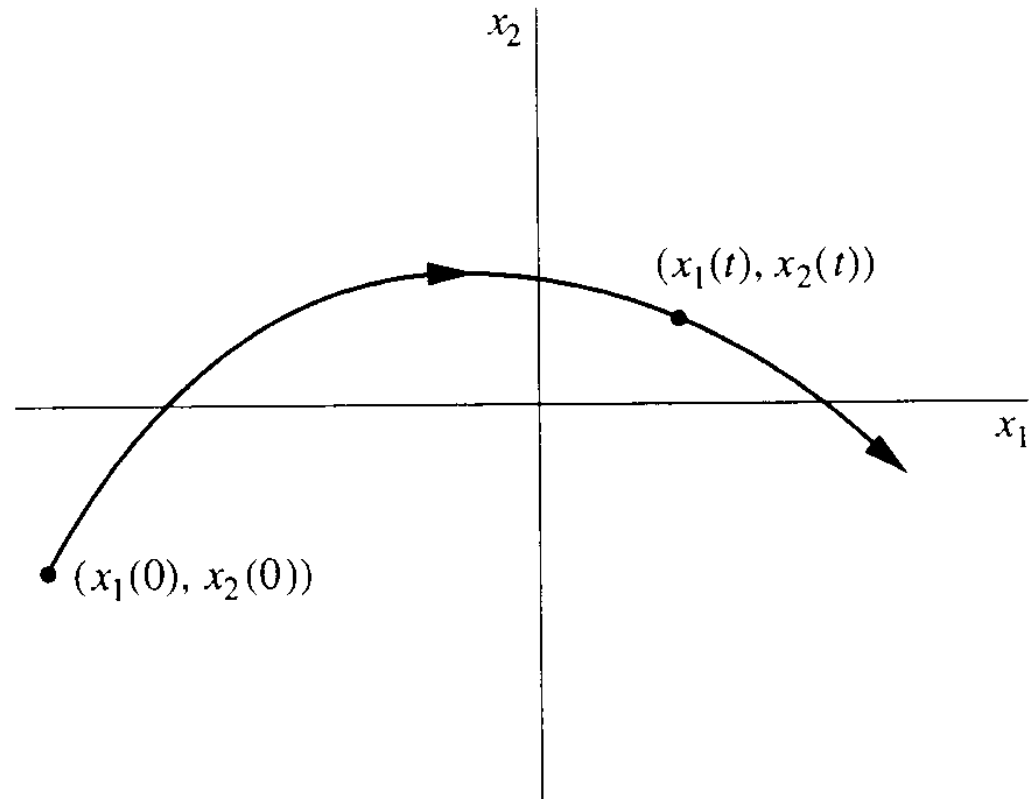
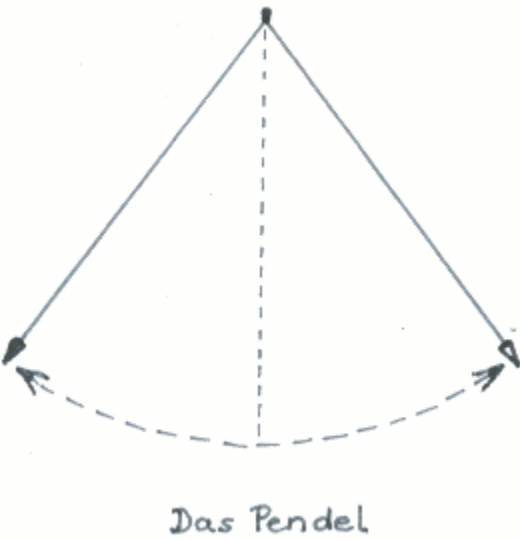
states



coupled  
differential  
equations

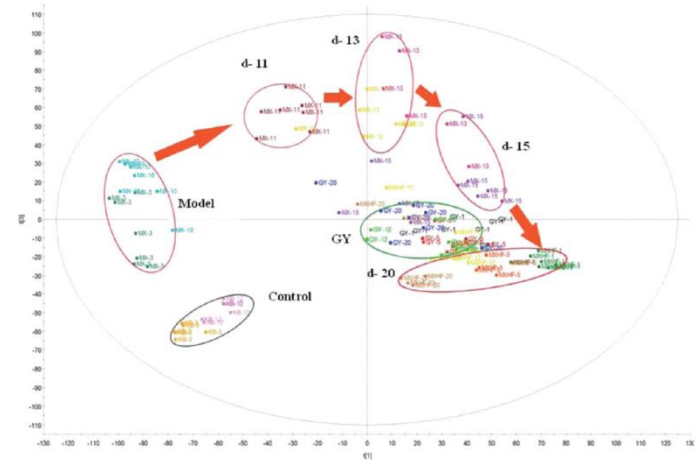
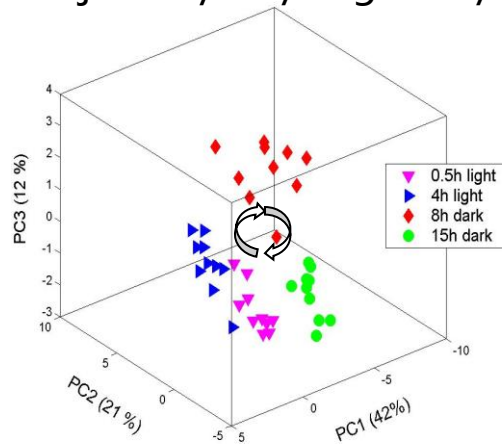
trajectories in  
multidimensional  
phase space

# Trajectory of a Pendular

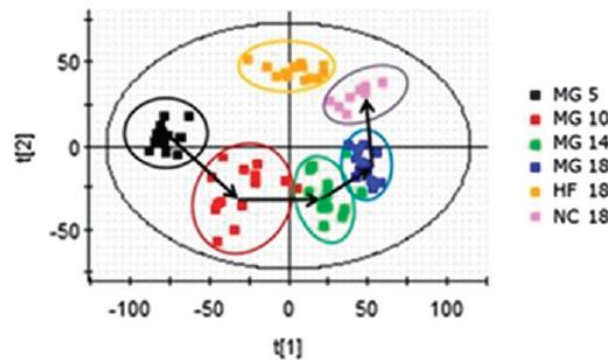


# Trajectories in living systems...

Trajectory day-night rhythm



Morgenthal et al. 2005



Wang et al. 2013

Zhou et al. 2015

# Correlation and Network Analysis

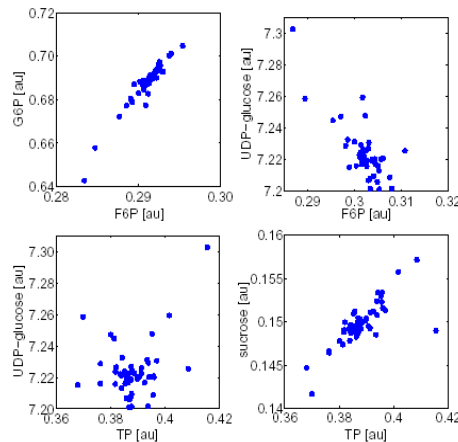
	$M_1$	$M_2$	$M_3$	...	$M_n$
$S_1$	$q_{1,1}$	$q_{1,2}$	$q_{1,3}$	...	$q_{1,n}$
$S_2$	$q_{2,1}$	$q_{2,2}$	$q_{2,3}$	...	$q_{2,n}$
$S_3$	$q_{3,1}$	$q_{3,2}$	$q_{3,3}$	...	$q_{3,n}$
...	...	...	...	...	...
$S_m$	$q_{m,1}$	$q_{m,2}$	$q_{m,3}$	...	$q_{m,n}$

n-dimensional  
Datenmatrix

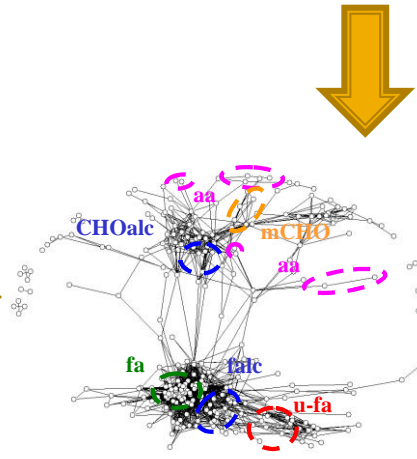


$$C = \begin{pmatrix} \text{cov}(M_1, M_1) & \text{cov}(M_1, M_2) & \cdots & \text{cov}(M_1, M_n) \\ \text{cov}(M_2, M_1) & \text{cov}(M_2, M_2) & \cdots & \text{cov}(M_2, M_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(M_n, M_1) & \text{cov}(M_n, M_2) & \cdots & \text{cov}(M_n, M_n) \end{pmatrix}_{n \times n}$$

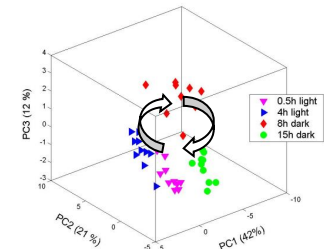
n-dimensional  
Covariance matrix



Metabolite correlations



Correlation network



Trajectories

# Correlation and Network Analysis

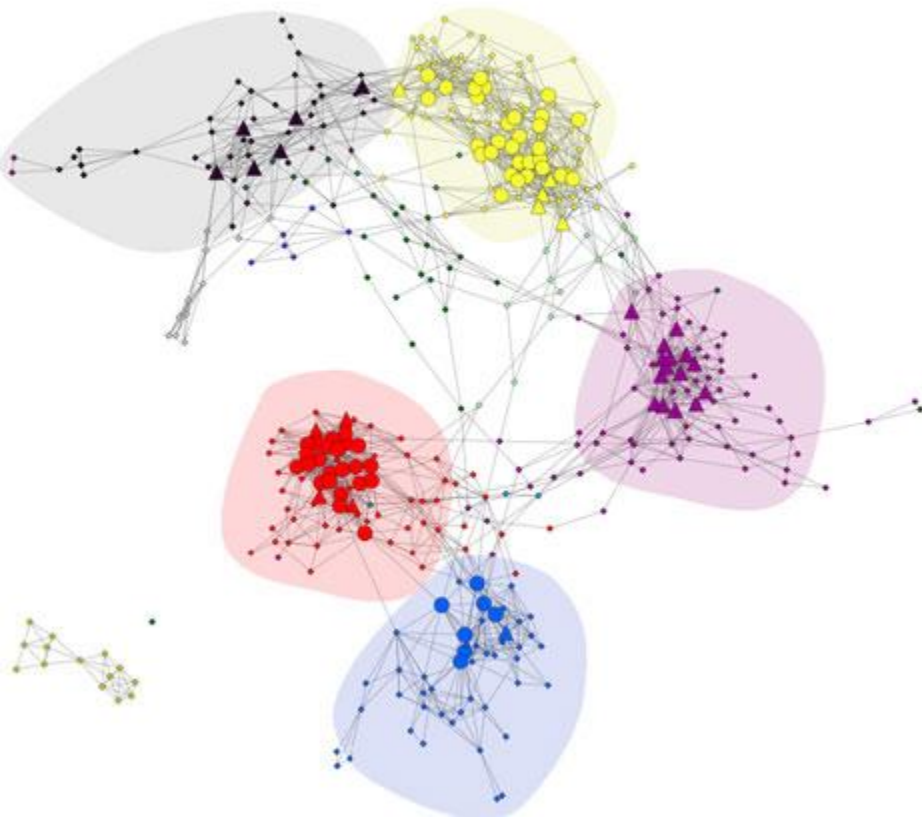
Data matrix  
Samples, Metabolites

WT 5h light	<sup>14</sup> C-Aniline	Alanine	Citric Acid	Aspartic	Asparagine	Aspartic	Biotin	Beta-Alanine	Citric Acid	Cinnoline	Ethanol	Fumaric
2.20219745	399.02873	305.02077	317.86837	310.86976	519.44528	4.7548126	742.02210	1.4410777	1.179			
1.22202401	303.04467	365.02829	38.02787	483.02413	1.5566466	5.8701883	11.89121	2.7102514	2.008			
5.0365428	303.90727	301.11720	31.7294283	241.20334	5.8281749	4.2331495	446.82229	1.6057388	1.0811328	11.0		
6.9820107	484.45254	263.48545	23.220199	357.18031	5.1316206	3.0471074	703.28249	1.9324001	7.9702371	1.988		
5.6126026	271.32545	287.70986	20.02727	159.48224	6.1314972	2.7776103	454.03232	5.9715642	5.9192544	1.465		
5.5763787	294.59605	270.39855	15.4495647	200.48134	7.1574621	2.4270964	446.40716	1.9388448	5.7576901	1.387		
7.1177070	131.41420	303.86265	24.7494489	133.89115	7.0444907	2.4976939	253.02202	5.9162955	5.9793449	0.613		
8.5920475	534.04875	263.22256	35.7714481	215.14839	10.314548	3.3823001	347.80792	2.1862742	4.8489520	1.373		
12.2626640	258.04262	267.82827	20.273992	227.29352	6.2251162	3.7842011	541.24116	1.8862540	1.9142628	1.768		
6.9284518	333.43346	306.48237	18.202272	184.75474	11.3214829	3.9471456	255.16916	1.4824153	5.8176933	2.118		
11.2526726	333.29216	165.8208	16.818229	289.72129	7.2039791	4.4264917	527.15844	1.13666	1.366			
6.9601771	139.01548	139.72465	24.687191	273.19161	6.1151978	4.4248161	54.02026	1.13666	1.366			
10.4419116	371.10682	163.82145	24.314971	202.78703	7.0774462	3.8022478	32.36767	1.13666	1.366			
6.8701481	272.23376	262.20126	13.824257	172.64621	6.9119789	6.0237019	258.91919	1.13666	1.366			
13.1247423	342.28677	149.34506	23.443047	201.18102	16.3822845	4.0898883	547.121	1.13666	1.366			
5.2232610	161.14814	107.20196	37.714007	124.38716	4.6298987	1.4882259	112.123	1.13666	1.366			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 4h light	8.3979206	402.16209	272.93497	21.620176	324.44814	10.034219	6.1614743	3.95	3.95			
WT 5h light	8.3979206	402.16209										

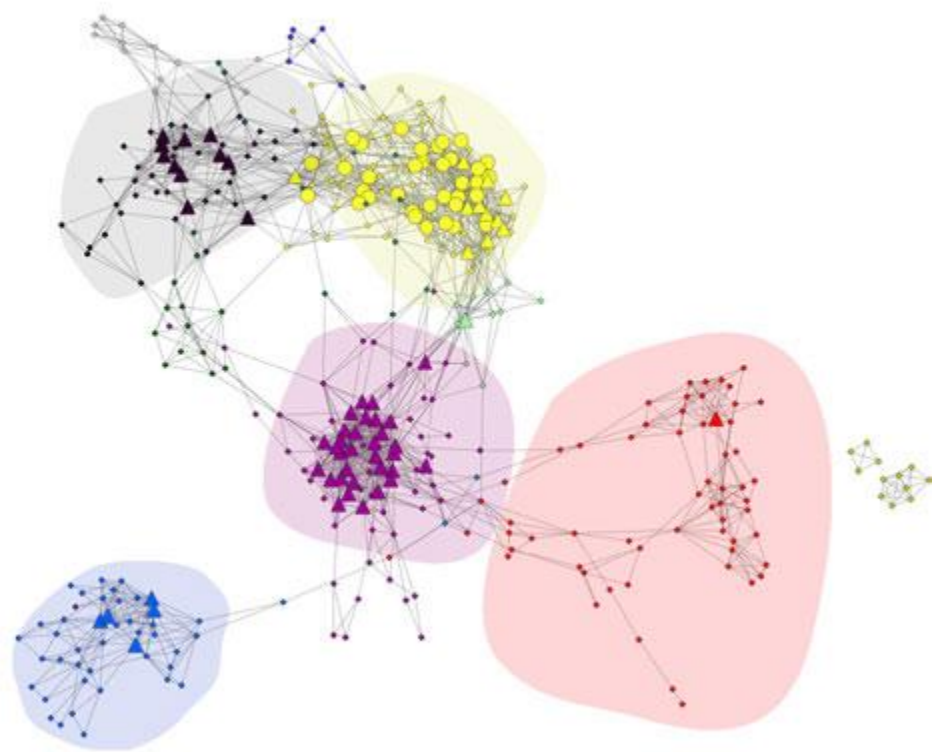


# Correlation and Network Analysis

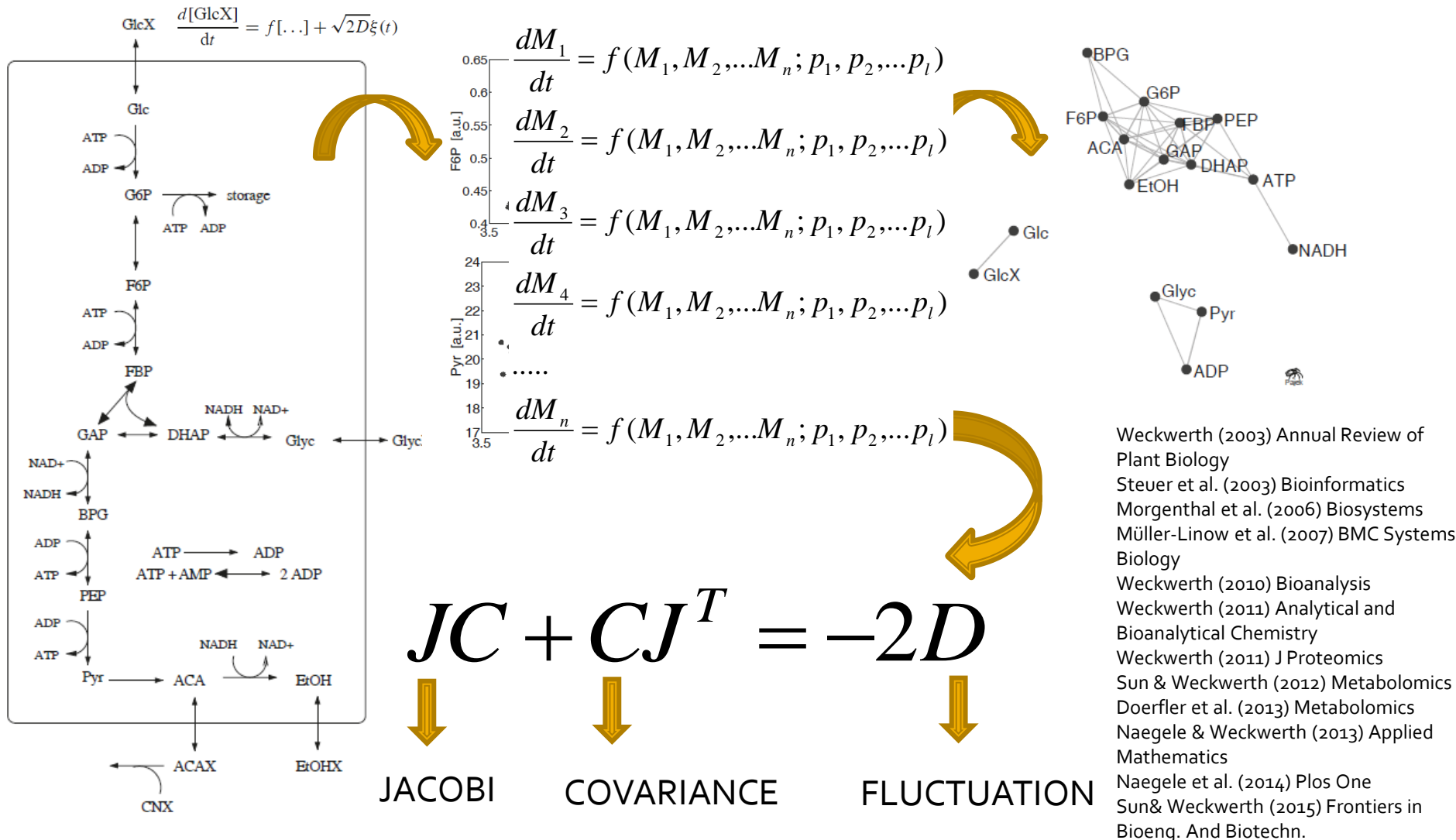
Tuesday (fasted/no caffeine)



Thursday (fed/caffeinated)



# Fundamental relation between network dynamics (statistics/correlation/covariance/network) and GENOME SEQUENCE



# Fundamental relation between network dynamics (statistics/correlation/covariance/network) and GENOME SEQUENCE

Jacobi entries are the Elasticities of the **enzymatic Reactions** dependent on metabolite concentration changes

$$\begin{aligned}
 \frac{dM_1}{dt} &= f_1(M_1, M_2, \dots, M_n) \\
 \frac{dM_2}{dt} &= f_2(M_1, M_2, \dots, M_n) \\
 \frac{dM_3}{dt} &= f_3(M_1, M_2, \dots, M_n) \\
 \frac{dM_4}{dt} &= f_4(M_1, M_2, \dots, M_n) \\
 &\dots \\
 \frac{dM_n}{dt} &= f_n(M_1, M_2, \dots, M_n)
 \end{aligned}
 \quad
 J = \begin{pmatrix}
 \frac{\partial f_1}{\partial M_1} & \frac{\partial f_1}{\partial M_2} & \dots & \frac{\partial f_1}{\partial M_n} \\
 \frac{\partial f_2}{\partial M_1} & \frac{\partial f_2}{\partial M_2} & \dots & \frac{\partial f_2}{\partial M_n} \\
 \vdots & \vdots & \ddots & \vdots \\
 \frac{\partial f_n}{\partial M_1} & \frac{\partial f_n}{\partial M_2} & \dots & \frac{\partial f_n}{\partial M_n}
 \end{pmatrix}_{n \times n}$$

$$\frac{\partial f}{\partial M} \begin{matrix} \nearrow \text{Change in reaction rate} \\ \searrow \text{Change in metabolite concentrations} \end{matrix}$$

$\partial f_i$   
 $\partial M_i$

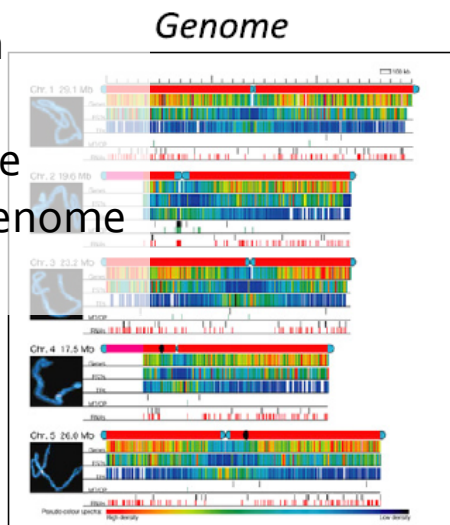
$$\underline{J = N * \partial f / \partial M}$$

- J: Jacobimatrix
- N: stoichiometric Matrix
- => GENOME SEQUENCE**
- v: Reaktion rate;
- M: metabolite

# Genome-scale metabolic reconstruction

## Genome-scale metabolic reconstruction

Human  
Plant  
Microbe  
Metagenome  
etc...



stoichiometric matrix  $N$

	M1	M2	M3	M4	M <sub>5</sub>
	Ru5P	FP <sub>2</sub>	F6P	GAP	R5P
Pgi	0	0	1	0	0
Fba,TpiA	0	-1	0	2	0
Rpi	-1	0	0	0	1
2 Rpe, TktI, Tal, TKII	-2	0	2	1	-1
Gap, Pfk, Gpm, Eno, Pyk	0	0	0	-1	0
Zwf, Pgl, Gnd	1	0	0	0	0
Pfk	0	1	-1	0	0
Fbp	0	-1	1	0	0
Prs	0	0	0	0	-1

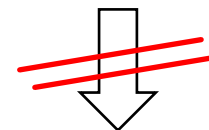
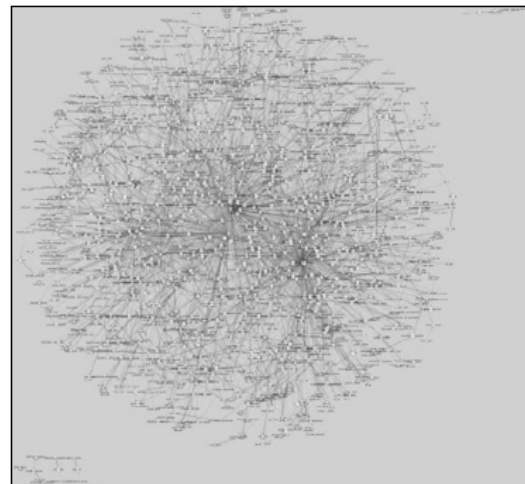


enzymatic reactions

```

Pgi : G6P = F6P .
Pfk : F6P + ATP = FDP + ADP .
Fbp : FDP = F5P + Pi .
Ald : FDP = DHAP + GAP .
Tpi : DHAP = GAP .
Gap : GAP + NAD + Pi = D13PG + NADH .
Zwf : G6P + NADP = GO6P + NADPH .
Pgl : GO6P = GL6P .
Gnd : GL6P + NADP = Ru5P + NADPH + CO2 .
Rpi : Ru5P = R5P .
Rpe : Ru5P = X5P .
TktI : X5P + R5P = GAP + S7P .
TktII : E4P + X5P = F6P + GAP .
Tal : S7P + GAP = E4P + F6P .
Pfk : D13PG + ADP = P3G + ATP .
Gpm : P3G = P2G .
Eno : P2G = PEP .
Pyk : PEP + ADP = PYR + ATP .
Prs_DeOB : R5P = R5Pex .
    
```

static metabolic network



Phenotypic plasticity

Weckwerth (2011)

**Unpredictability of metabolism.**

*Analytical and Bioanalytical Chemistry*

# A genotype-phenotype-equation

Genome sequence



Data derived dynamic model of the phenotype

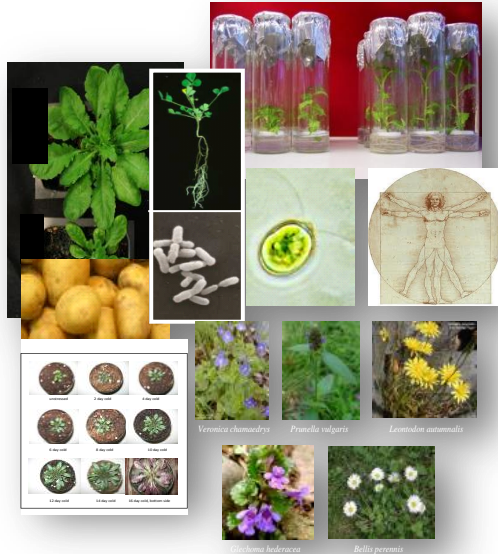
Structure of the Jacobi is defined by the stoichiometric Matrix derived from the genome

Environmentally driven stochastic Diffusionmatrix

$$JC + CJ^T = -2D$$

Covariance Matrix (dynamic network of the metabolic system)

# A system-theoretical derivation of a geno-phenotype-equation



	$M_1$	$M_2$	$M_3$	...	$M_n$
$S_1$	$q_{1,1}$	$q_{1,2}$	$q_{1,3}$	...	$q_{1,n}$
$S_2$	$q_{2,1}$	$q_{2,2}$	$q_{2,3}$	...	$q_{2,n}$
$S_3$	$q_{3,1}$	$q_{3,2}$	$q_{3,3}$	...	$q_{3,n}$
...	...	...	...	...	...
$S_m$	$q_{m,1}$	$q_{m,2}$	$q_{m,3}$	...	$q_{m,n}$

n-dimensional Data matrix

**PHENOTYPE C**



$$\frac{dM_1}{dt} = f(M_1, M_2, \dots, M_n)$$

$$\frac{dM_2}{dt} = f(M_1, M_2, \dots, M_n)$$

.....

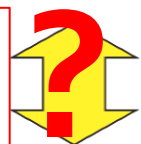
$$\frac{dM_n}{dt} = f(M_1, M_2, \dots, M_n)$$

Systems equations

**Systems theory J**



$$JC + CJ^T = -2D$$

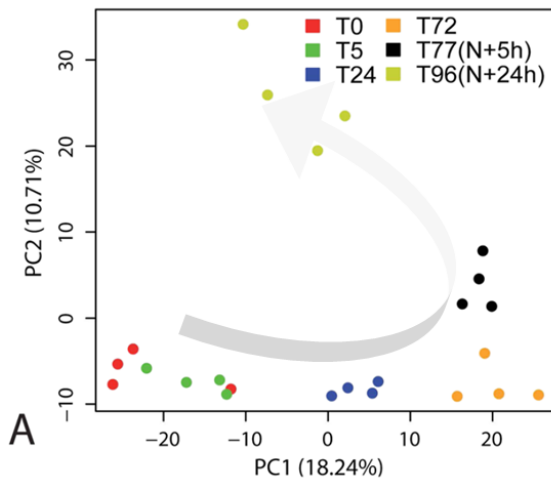


Multivariate Statistics

**Trajectories**



Genome sequence  
**GENOTYPE N**



Weckwerth 2011 Unpredictability of Metabolism. ABC  
Weckwerth 2011 Green Systems Biology

# COVAIN – COVariance *IN*verse: inverse calculation of the differential Jacobian

METABOLOMICS

DOI: 10.1007/s11306-012-0399-3 **Online First**

ORIGINAL ARTICLE

**COVAIN: a toolbox for uni- and multivariate statistics, time-series and correlation network analysis and inverse estimation of the differential Jacobian from metabolomics covariance data**

Xiaoliang Sun and Wolfram Weckwerth



Xiaoliang Sun

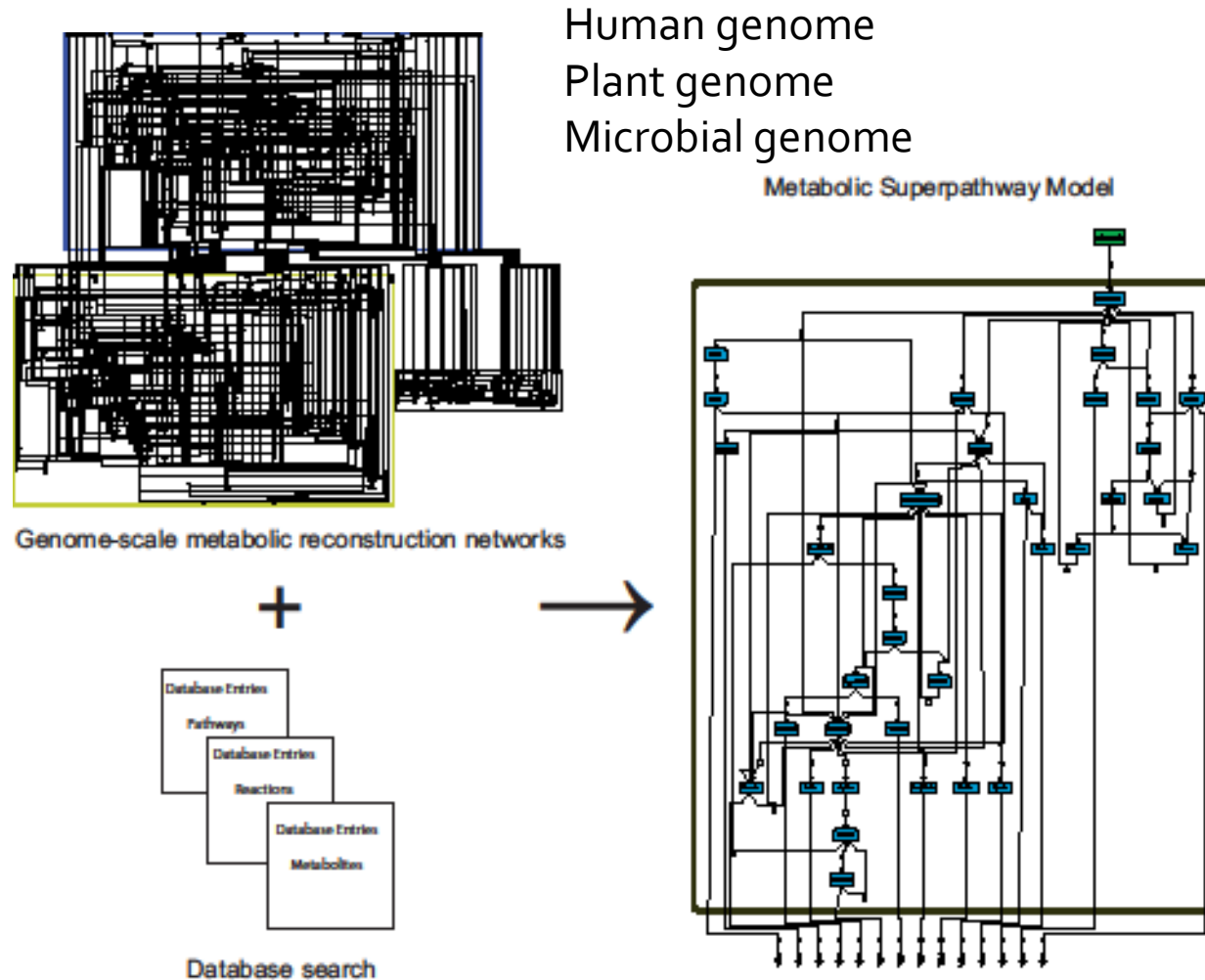
- PCA, ICA, HCA, etc.
- GRANGER CAUSALITY
- data integration
- data transformation
- data normalization
- metabolic modelling
- metabolomics
- proteomics
- transcriptomics
- other data

The screenshot displays the COVAIN Toolbox (Version 2012-03-22) interface. It is divided into several main sections:

- Data Information:** Includes a 'Choose data' dropdown menu, a 'Data preview' table, and buttons for 'Transpose', 'Normalise', 'Fill missing values', 'Adult outliers', 'ANOVA', 'Log transform', and 'Z transform'. The data preview table shows columns for 'Substrate', 'CHOL1', 'CHOL2', 'CHOL3', and 'CHOL4' with numerical values for various conditions like 'WT 5h light' and 'WT 4h light'.
- Data Analysis:** Contains sub-sections for 'Multivariate Statistics' (with options for PCA, ICA, Correlation, and Cluster) and 'Time Series' (with options for Correlation, Clustering, and Granger causation analysis).
- Network Analysis:** Features 'Network Inference' (with options for Correlation, Time Series, and Granger analysis), 'Network Property', 'Inverse Jacobian', and 'KEGG Pathway'.

On the right side of the interface, there are several control buttons: 'Load data...', 'Continue data...', 'My notes...', 'Name the results', 'Save', 'Options', and 'F10'.

# Metabolic reconstruction from a genome sequence

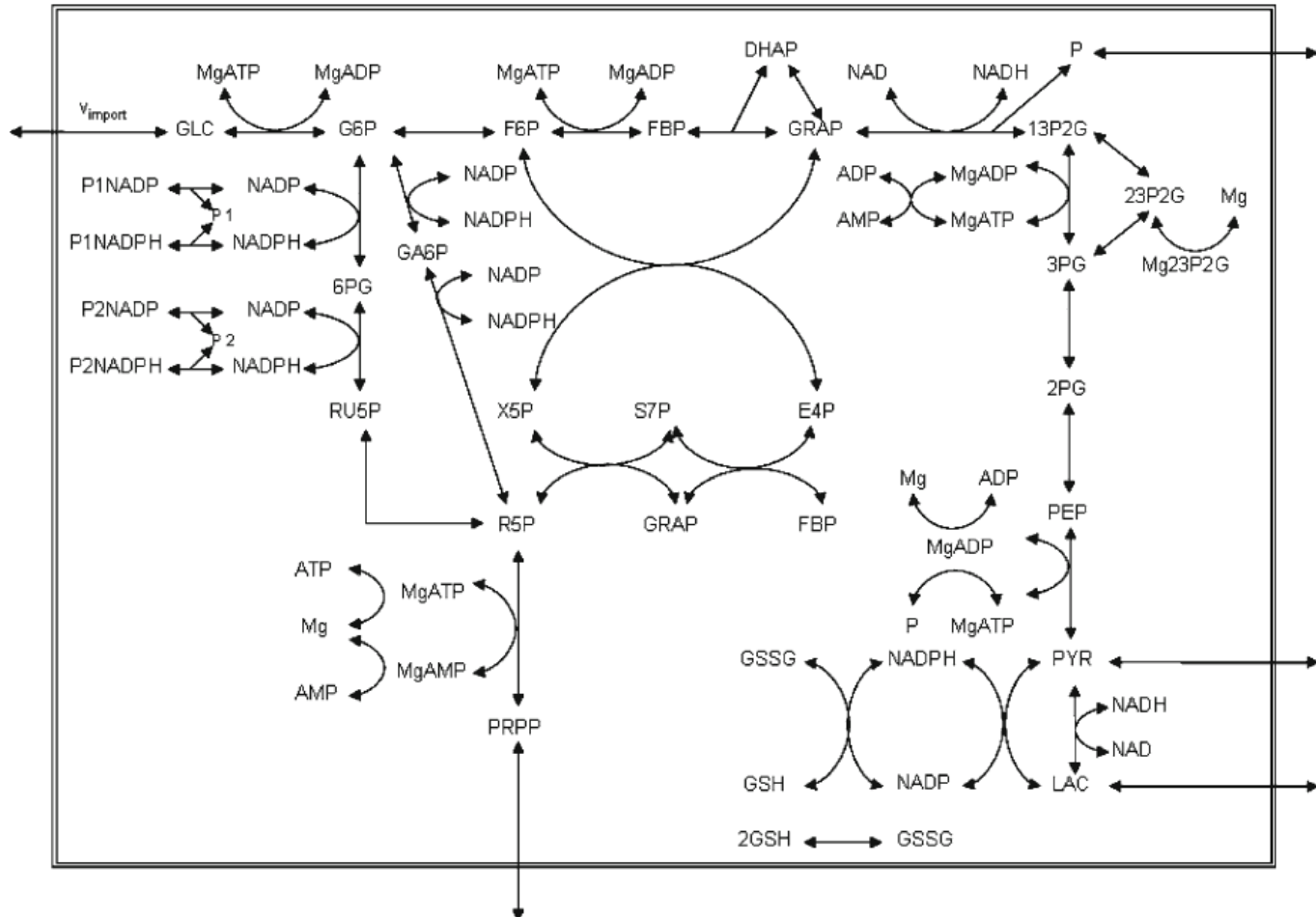


Weckwerth (2011) *Unpredictability of metabolism from genome sequences.*

Analytical and Bioanalytical Chemistry



# Calculation of the Differential Jacobian - Demonstration with an erythrocyte model



# Differential Jacobian - Demonstration with an erythrocyte model

- Erythrocyte Model
  - -high level glucose import
  - -low level glucose import
  - Using Covariance of metabolite data to calculate differential Jacobian

$$dJ_{ij} = \log_2 \left| \frac{J_{a,ij}}{J_{b,ij}} \right|$$

High level glucose import

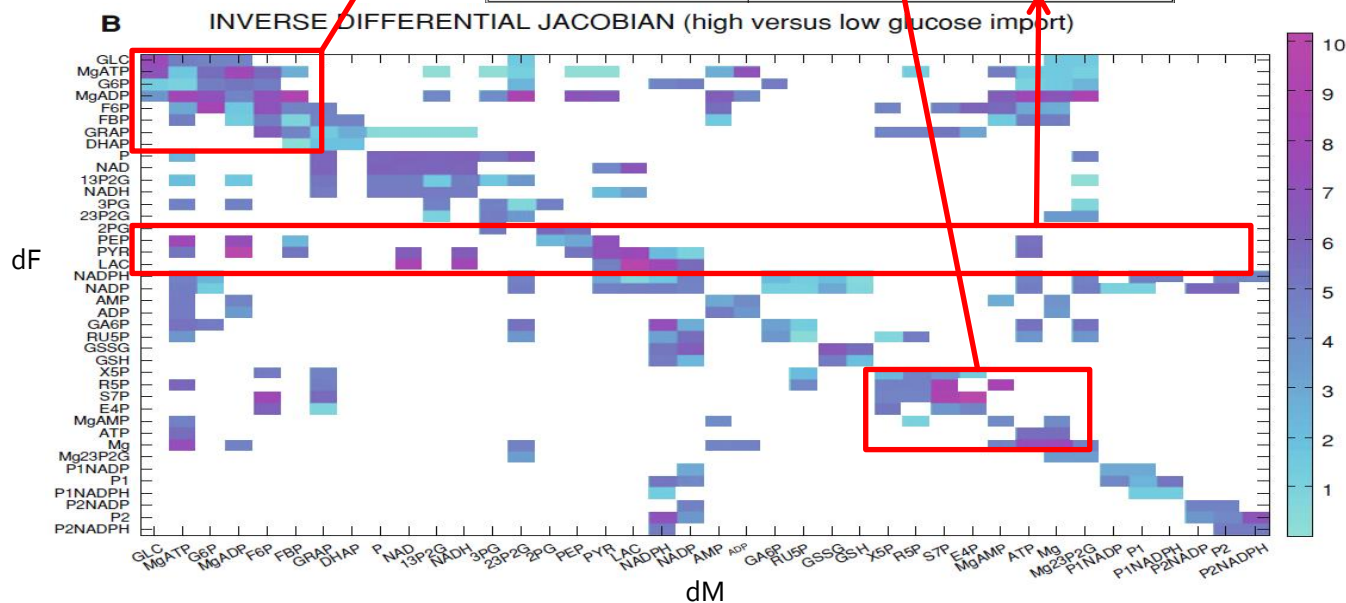
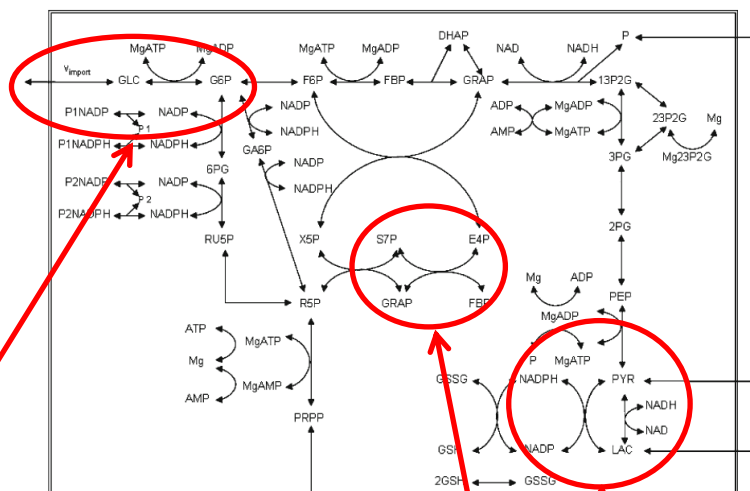
Low level glucose import

# Calculation of the differential jacobian - demonstration with an erythrocyte model

$$dJ_{ij} = \log_2 \left| \frac{J_{a,ij}}{J_{b,ij}} \right|$$

High level  
glucose  
import

Low level  
glucose  
import



# COVAIN – COVAriance INverse: toolbox for data mining

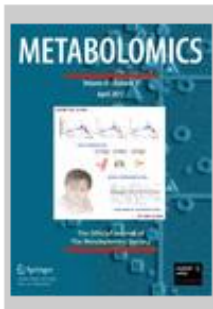
METABOLOMICS

DOI: 10.1007/s11306-012-0399-3 **Online First**

ORIGINAL ARTICLE

COVAIN: a toolbox for uni- and multivariate statistics, time-series and correlation network analysis and inverse estimation of the differential Jacobian from metabolomics covariance data

Xiaoliang Sun and Wolfram Weckwerth



Sun & Weckwerth  
2012 Metabolomics

$$JC + CJ^T = -2D$$

COVAIN GUI Matlab

- PCA, ICA, HCA, etc.
- GRANGER CAUSALITY
- data integration
- data transformation
- data normalization
- metabolic modelling
- metabolomics
- proteomics
- transcriptomics
- other data

	PGM 0.5h dark	PGM 0h dark	PGM 15.5h dark	PGM 0.5h light	PGM 0h light
*2,5-Diaminopentanoic acid (DTMS)	16.9086	17.1093	13.3326	20.1805	16.5804
*4-Aminobutyric acid (DTMS)	6.4533	6.9193	6.5746	6.4302	6.107
*Alanine (DTMS)	654.8559	303.4951	478.2449	757.2831	809.342
*Alanine (DTMS)	10.8419	9.9181	10.9686	16.6323	16.532
*Ascorbic acid minor 1273.6	21.0651	15.7917	12.6062	14.9128	15.436
*Ascorbic acid 1499.6	241.8182	241.8488	191.7987	286.9636	211.018
*Ascorbic acid minor	13.8556	15.1694	9.7207	12.9116	9.819
*Asparagine	60.7488	73.5348	122.3027	119.8638	86.486
*Asparagine (DTMS)	31.7554	46.1002	25.4189	59.5166	21.776
*Aspartic acid (DTMS)	194.8303	334.0679	236.7207	162.7911	164.876
*Benzoic acid (DTMS)	60.7488	73.5348	122.3027	119.8638	86.486
*CHO1	10.3086	10.5980	7.3127	10.1467	9.200
*CHO10	12.0973	12.6995	11.8871	14.7999	13.493
*CHO11	36.1957	36.5966	24.4786	27.2232	27.596
*CHO2	152.6357	147.0764	122.7431	163.0282	142.361
*CHO3	22.8198	17.1767	14.3129	18.8657	19.478
*CHO4	31.8055	31.2257	25.4971	29.3764	30.622
*CHO5	8.1542	4.3791	1.1063	2.6696	5.796
*CHO6	0.6553	1.6021	0.6533	1.2081	5.815
*CHO7	46.0822	81.2385	41.9604	43.4313	42.816
*CHO8	32.5477	56.9176	27.4907	31.0328	36.356
*CHO9	15.9181	15.7286	14.6513	19.1889	17.147
*Citric acid (DTMS)	243.6900	318.4196	207.2140	289.2163	286.947

# COVAIN – COVAriance *IN*verse: toolbox for data mining

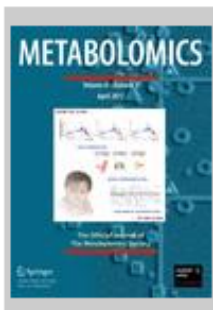
METABOLOMICS

DOI: 10.1007/s11306-012-0399-3 **Online First**

ORIGINAL ARTICLE

**COVAIN: a toolbox for uni- and multivariate statistics, time-series and correlation network analysis and inverse estimation of the differential Jacobian from metabolomics covariance data**

Xiaoliang Sun and Wolfram Weckwerth

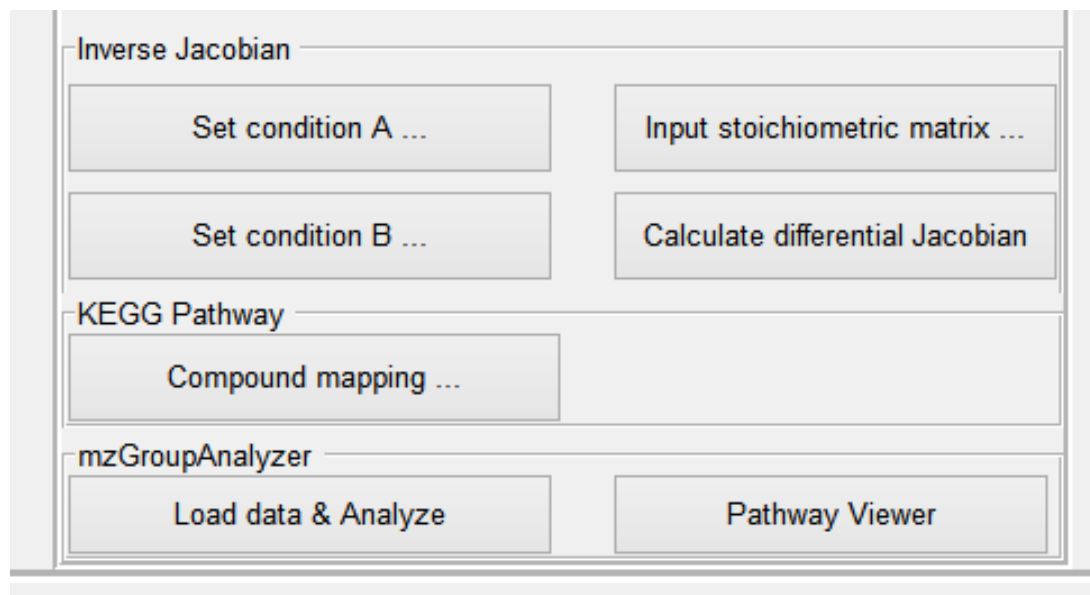


- PCA, ICA, HCA, etc.
- GRANGER CAUSALITY
- data integration
- data transformation
- data normalization
- metabolic modelling
- metabolomics
- proteomics
- transcriptomics
- other data

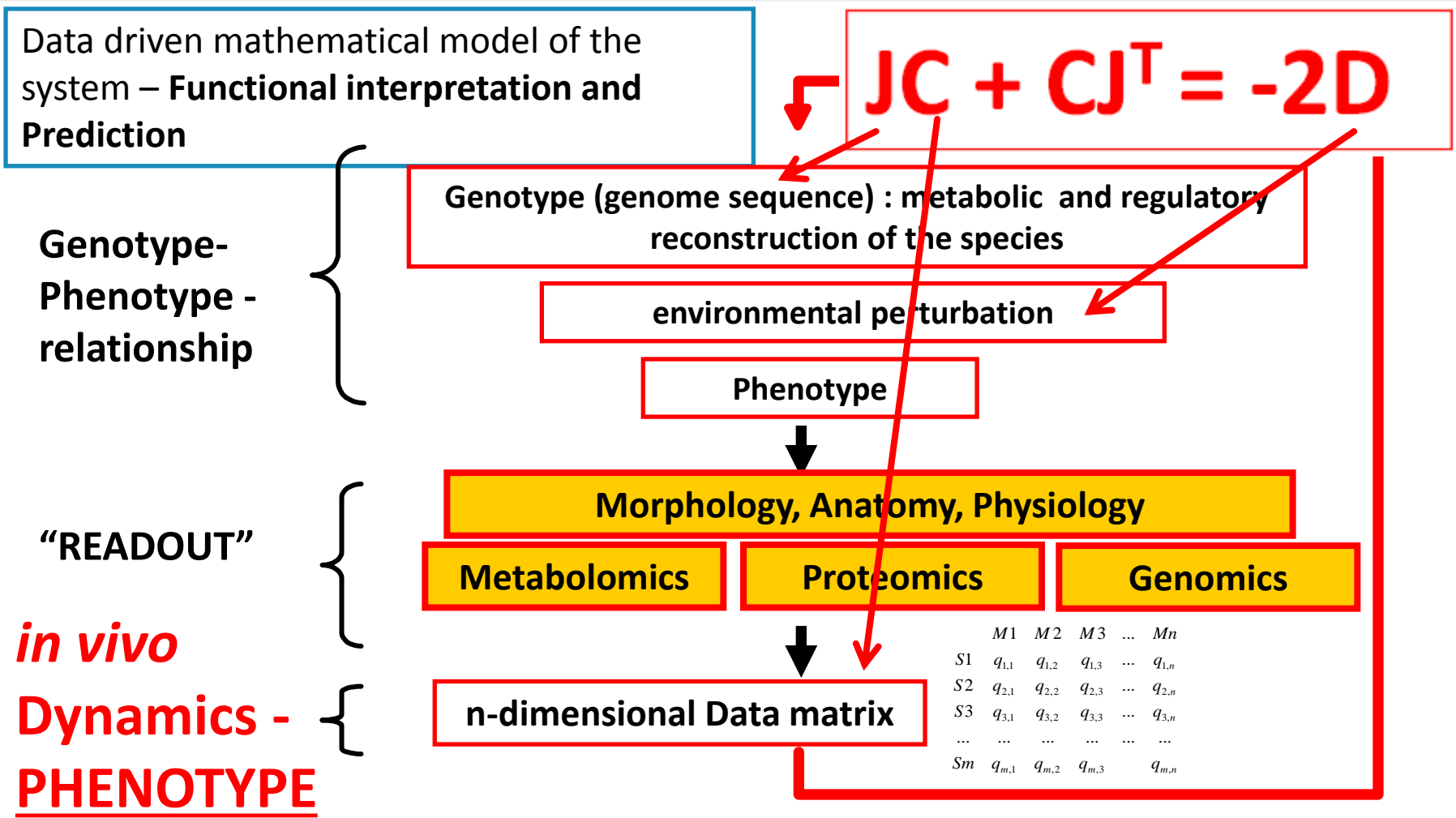
Sun & Weckwerth  
2012 Metabolomics

$$JC + CJ^T = -2D$$

COVAIN GUI Matlab



# Linking OMICS and Modelling Platform MoSys



# Calculation of biochemical regulation directly from metabolomics data

- Sun & Weckwerth (2012) Metabolomics
- Dörfler et al. (2013) Metabolomics
- Nägele et al. (2013) Applied Mathematics
- Nägele et al. (2014) Plos One
- Sun et al. (2015) Frontiers Bioeng. Biotechn.
- Nägele et al. (2016) Frontiers Metabolomics
- Nukarinen et al. submitted
- Nägele, Kirchmayer et al. In preparation

## Other groups

- Öksüz et al. (2013) Plos One
- Cakir et al. (2014) Frontiers in Bioeng. Biotechn.
- Kügler & Yang (2014) Journal of Mathematical Biology

# Summary and Outlook

- The genotype-phenotype equation links statistical features of metabolite dynamics with biochemical regulation and thus implements a novel level of biochemical causality in classical metabolomics analysis
- Novel pathway dynamics and fluctuations can be predicted.
- Integration of model-based Jacobian and other OMICS data opens up novel data integration strategies for PAN-OMICS approaches
- Applications envisaged from environmental sciences up to biomedical approaches including GWAS, Epigenetics, RNAseq, Proteomics and Phosphoproteomics for the molecular interpretation of the phenotype





Stefanie  
Wienkoop



Christiana  
Staudinger



Wolfgang  
Hoehenwarter



Till Ischebeck



Wiener Wissenschafts-, Forschungs- und Technologiefonds



Lena Fragner



Luiz Recuenco-  
Munoz



Vlora Mehmeti



Matthias Nagler



Thomas  
Naegele



Gert Bachmann



Anne-Mette  
Hanak



Wolfgang  
Postl



Alexander  
Seidel



QuiYang  
Fan



Yanmei Chen



Xiaoliang Sun



Volker  
Egelhofer



Peter Kropf



Doris Dellmour



Hannes Doerfler



Ella Nukarinen

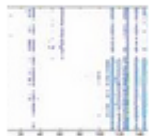
<http://www.univie.ac.at/mosys/team.html>

Submit your paper

Search in this section

## Articles

Latest Most viewed



### A data preprocessing strategy for metabolomics to reduce the mask effect in data analysis

Jun Yang, Xinjie Zhao, Xin Lu, Xiaohui Lin and Guowang Xu

**Original Research** Metabolomics is a booming research field. Its success highly relies on the discovery of different data sets (for example, pattern is that differences of the low abundance).

Published on 02 February 2015  
Front. Mol. Biosci. doi: 10.3389/fmolb

1,381 views 3 shares



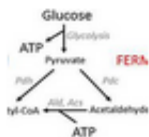
### Linking metabolomics data to underlying metabolic regulation

Thomas Nägele

**Review** The comprehensive experimental analysis of a metabolic constitution plays a central role in approaches of organismal systems biology. Quantifying the impact of a changing environment on the homeostasis of cellular metabolism has been the focus of ...

Published on 06 November 2014  
Front. Mol. Biosci. doi: 10.3389/fmolb.2014.00022

1,384 views 2 likes



### An evolutionary perspective on the Crabtree effect

Thomas Pfeiffer and Annabel Morley

**Perspective** The capability to ferment sugars into ethanol is a key metabolic

## Scope

Metabolomics publishes works that measure the physiological effects of the interactions between genes and the environment by profiling small molecules – including metabolic breakdown products – in massively parallel throughput, to infer gene-product function and regulation.

Read More >



## Editorial Board

ON BOARD EDITORS 63



**Wolfram Weckwerth**  
University of Vienna  
Vienna, Austria  
Specialty Chief Editor  
Metabolomics

Follow

12,965 views 122 publications 133 followers



**Eiichiro Fukusaki**

Follow



**David R Gang**  
Washington State University  
Pullman, USA  
Associate Editor  
Metabolomics

Follow

1,855 views 1 publications 15 followers



**Francois-Pierre Martin**  
Nestle Institute of Health Sciences  
Lausanne, Switzerland  
Associate Editor  
Metabolomics

Follow

2,984 views 47 publications 69 followers



**Thomas Nägele**  
University of Vienna  
Vienna, Austria  
Associate Editor  
Metabolomics

Follow

5,823 views 16 publications 26 followers



**Rima Obeid**

Follow

Submit your manuscripts...

## Topic Editors

Firas H Kobeissy, Hibah Omar Awwad and Saleh AlGhamdi

This Research Topic will consist of peer-reviewed original research papers, methods and review articles covering neuroscience topics, and the immunological and genetic mechanisms of pathological disorders, from a molecular biology perspective. The ...

Submission open

320 views

## Theoretical Concepts and Computational Methods in Metabolomics

### Topic Editors

Thomas Nägele, Thomas Pfeiffer and Wolfram Weckwerth

Metabolomics has become a central element of systems biology research. It ultimately aims at detecting, identifying, quantifying and interpreting the occurrence and the abundance of metabolites in context of organisms' biology. For this, various ...

Submission open

671 views 15 authors

View all >

Suggest a Topic

**END**

---