



## XII Congresso Brasileiro de Fisiologia Vegetal

“Desafios para a Produção de Alimentos e Bioenergia”

7 a 12 de setembro de 2009

Fortaleza - CE

Logo da  
Instituição

# Metabolite profiling as a tool for functional analysis of plant genome

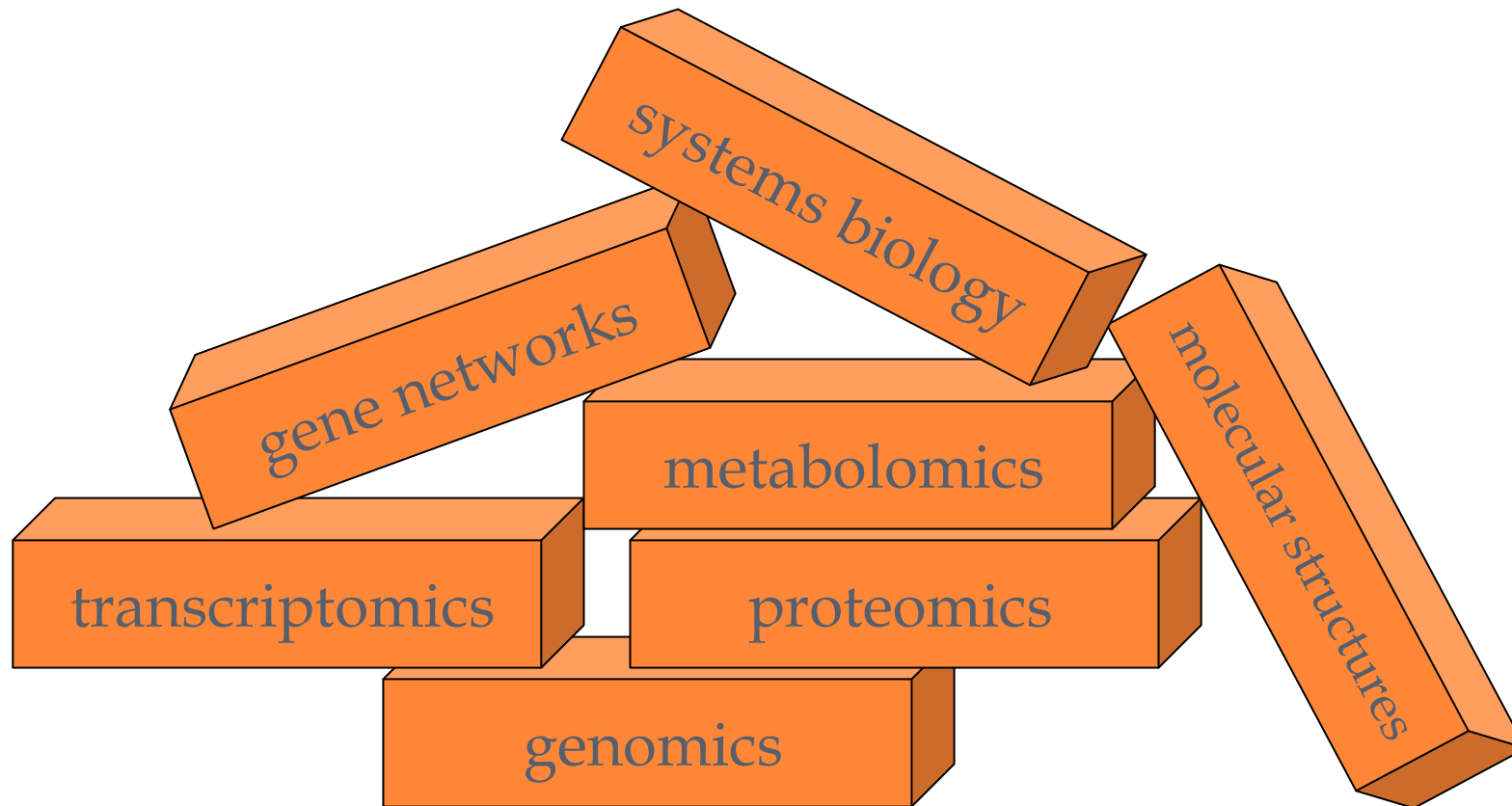
Prof. Marcelo Ehlers Loureiro  
Departamento de Biologia Vegetal  
Universidade Federal de Viçosa

[mehlers@ufv.br](mailto:mehlers@ufv.br)

031 38992576

# Genome

- *History of temporal succession of efforts*
- *Different tools, different use efficiencies...*



# Phenotype is controlled by *G x E x t x s*

Life Sciences



Genomics



Transcriptomics



Proteomics



Metabolomics

*Genotype x Environment*

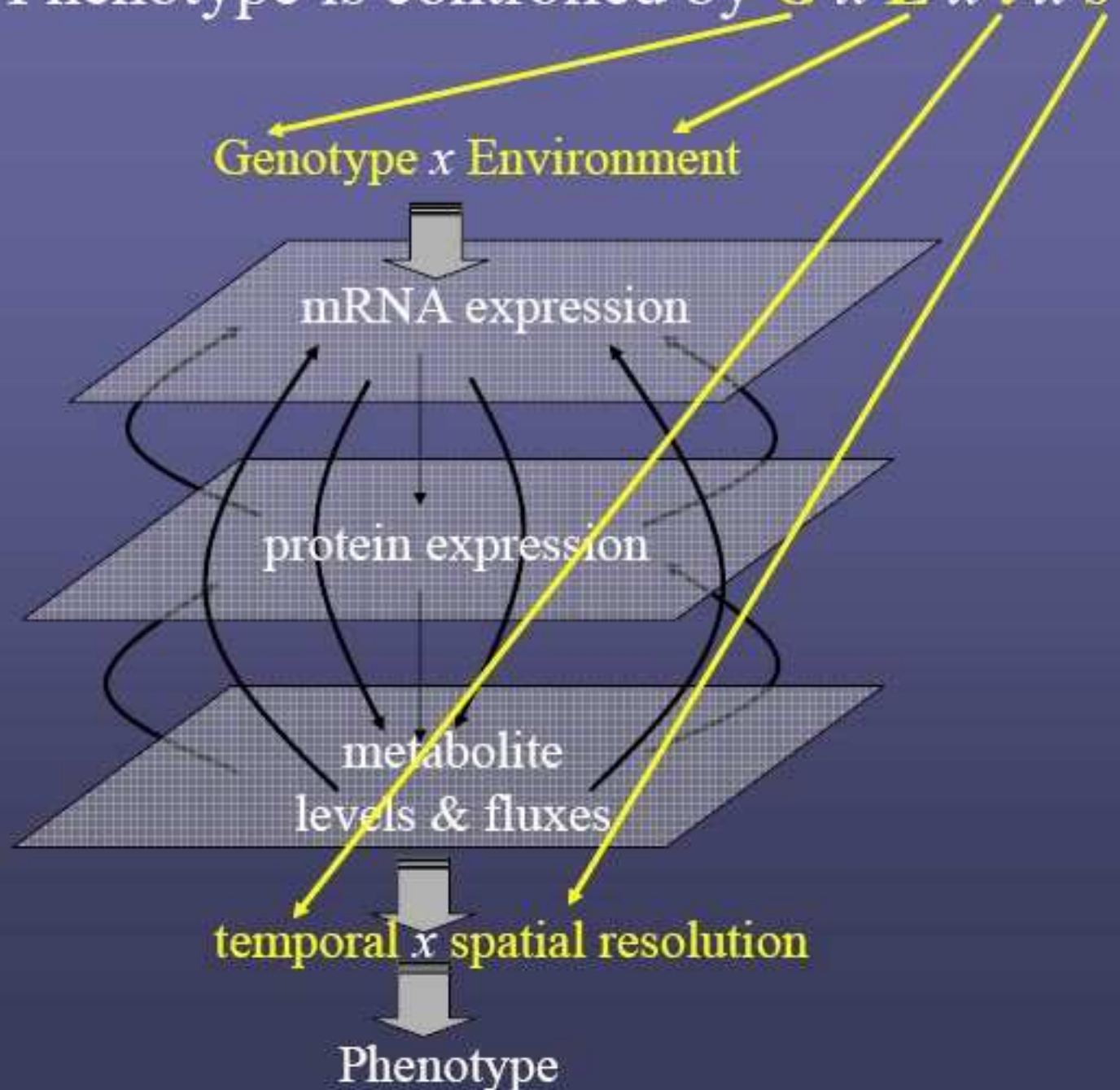
mRNA expression

protein expression

metabolite  
levels & fluxes

*temporal x spatial resolution*

Phenotype



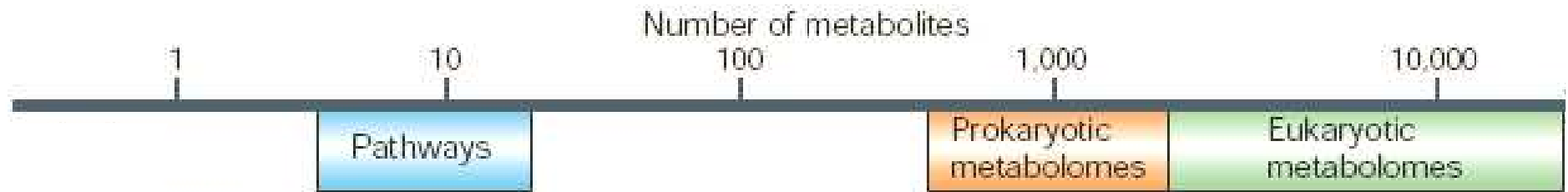
# Basic glossary about Metabolomics

**Metabolomics** “...the *complete set of metabolites/low-molecular-weight intermediates*, which are context dependent, varying according to the physiology, developmental or pathological state of the cell, tissue, organ or organism...” Oliver 2002

**Metabolic profiling** “...identification and quantification of a *selected number of pre-defined metabolites*, generally related to a *specific* metabolic pathways

**Metabolic fingerprinting** “high-throughput, *rapid, global* analysis of samples to provide sample *classification*. Quantification and identification are generally not employed.

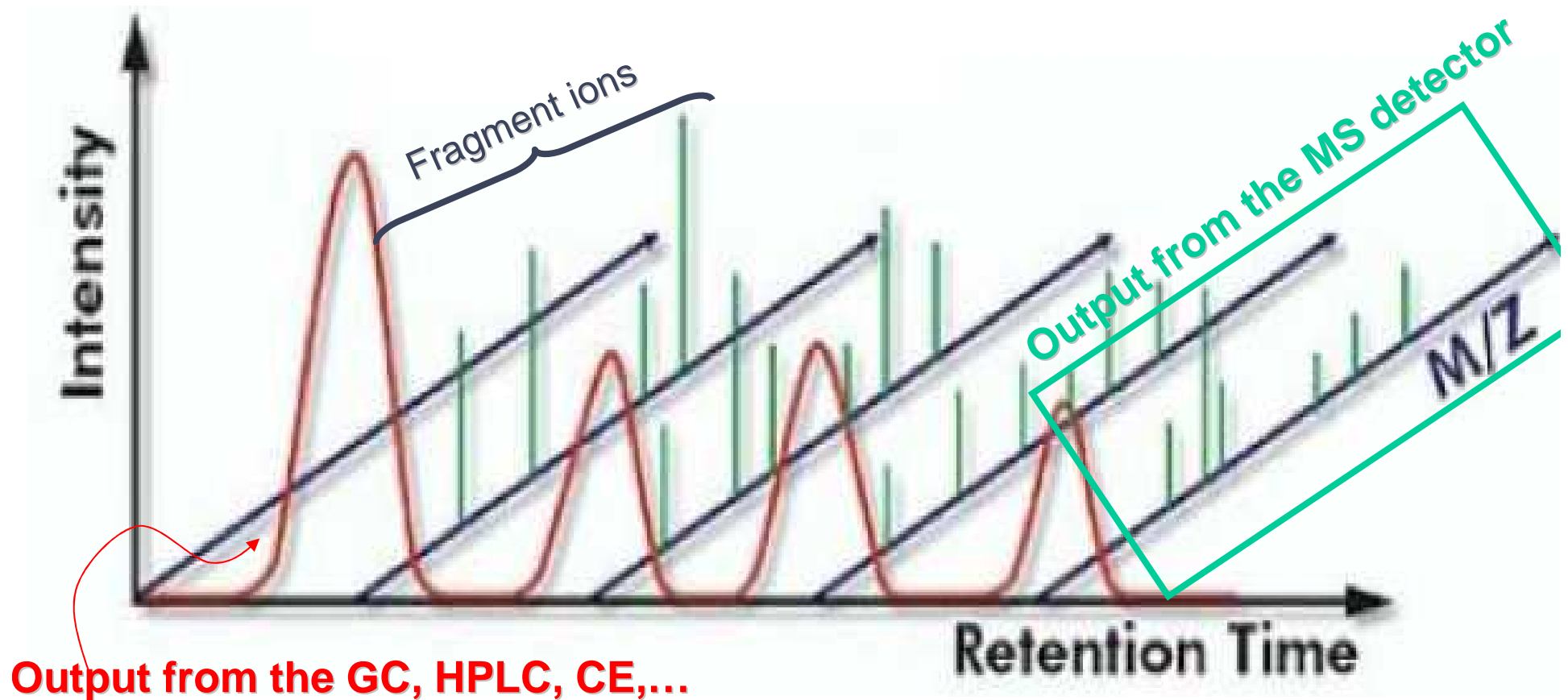
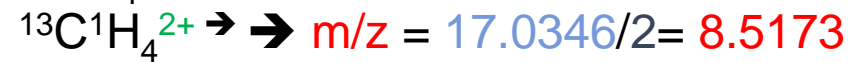
# Overview



200 000 metabolites are estimated to exist in the plant kingdom  
probably **~20.000 metabolites per plant**

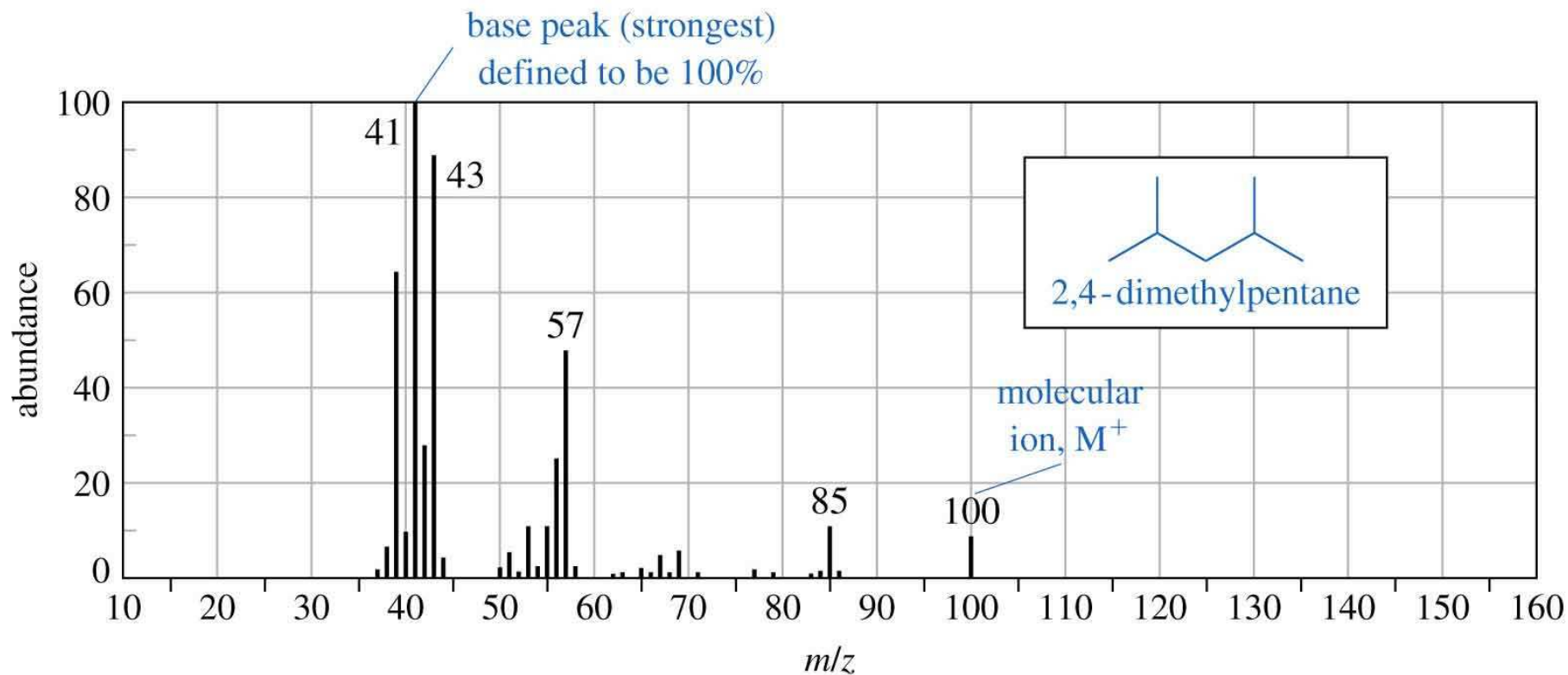
## $m/z$ = mass-to-charge-ratio

Unitless ratio of mass number to charge  $z$



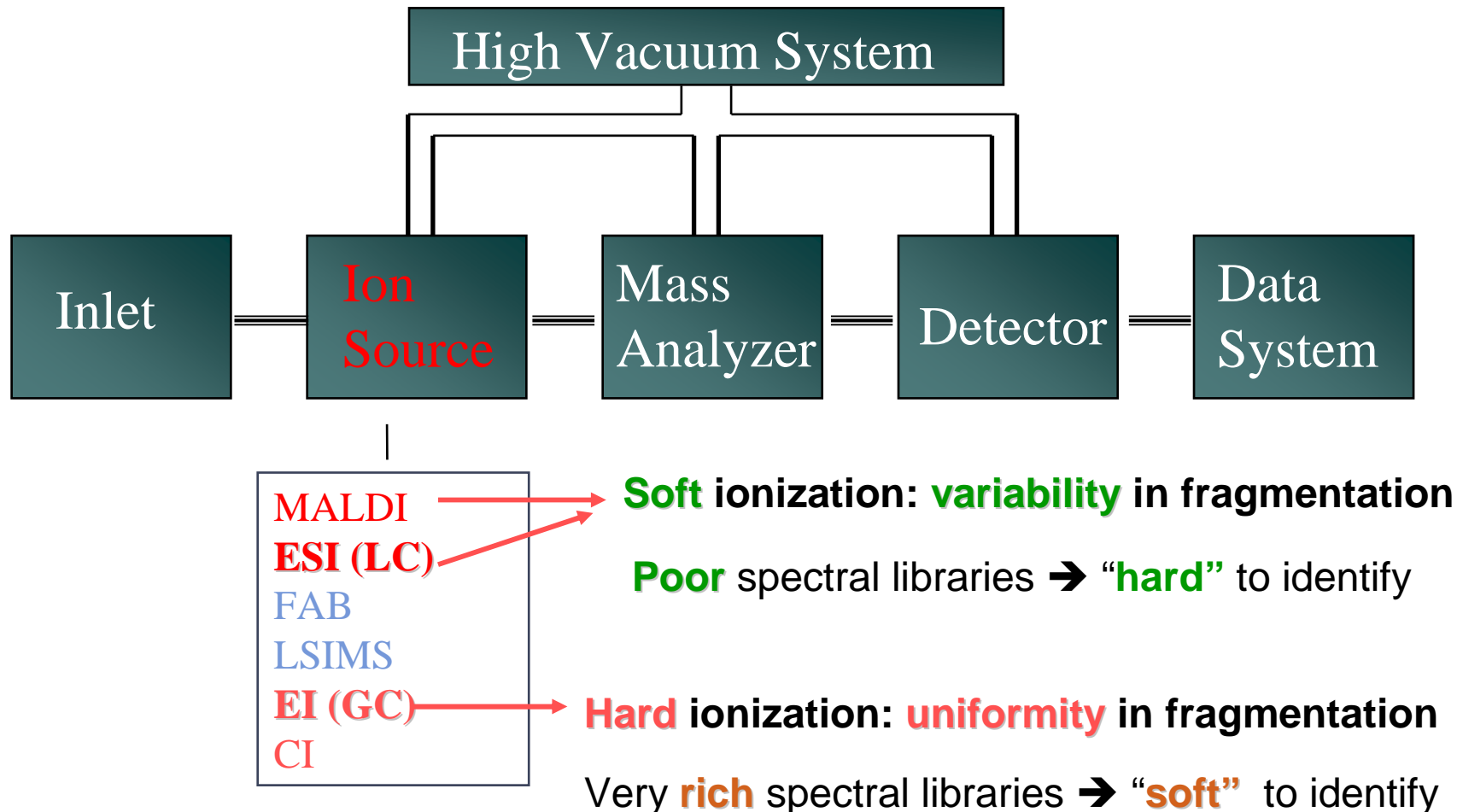
# The Mass Spectrum

Masses are graphed or tabulated according to their relative abundance.



=>

# Structure of a mass spectrometer



## Typical reactions in na **Electron-Impact** ionization (GC-MS; “hard)

Molecular ion formation

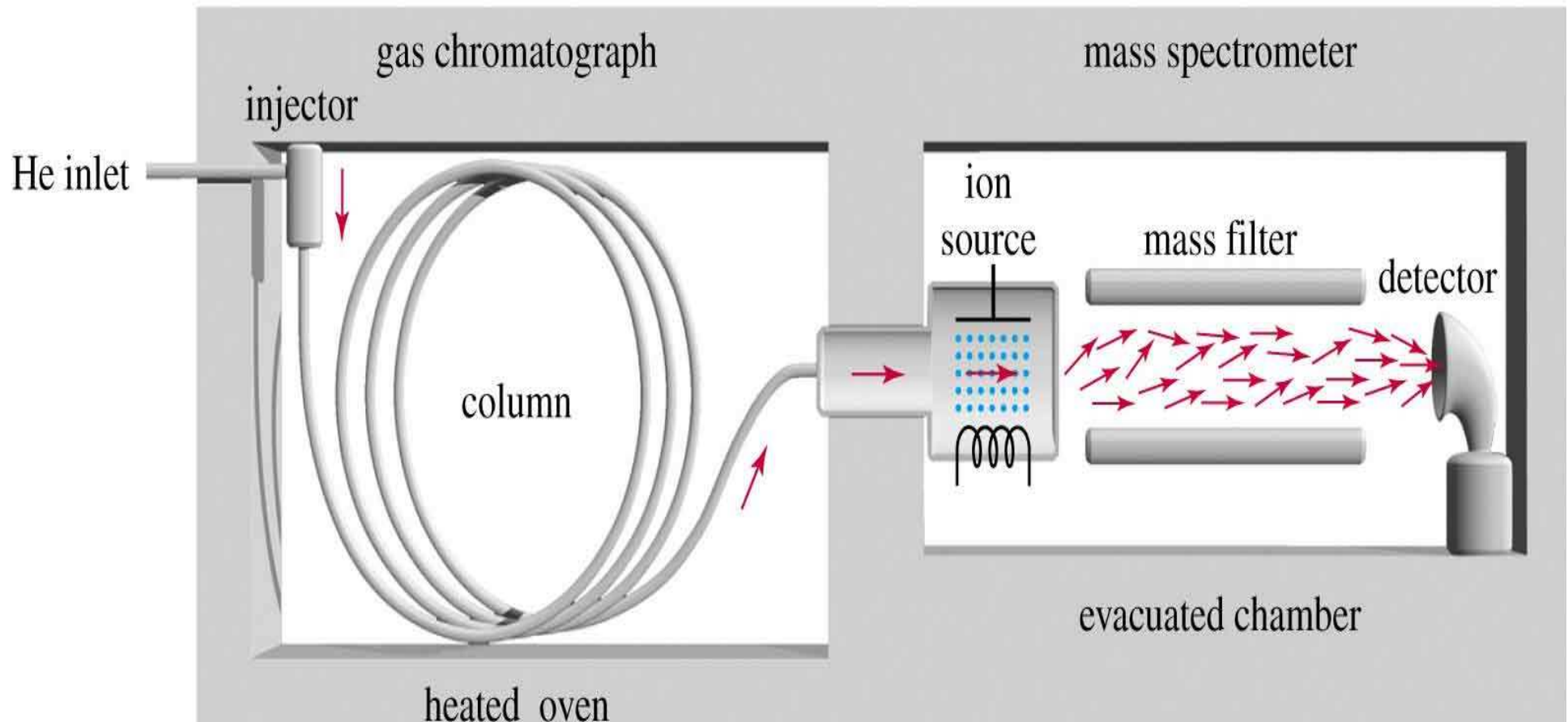
Fragmentation



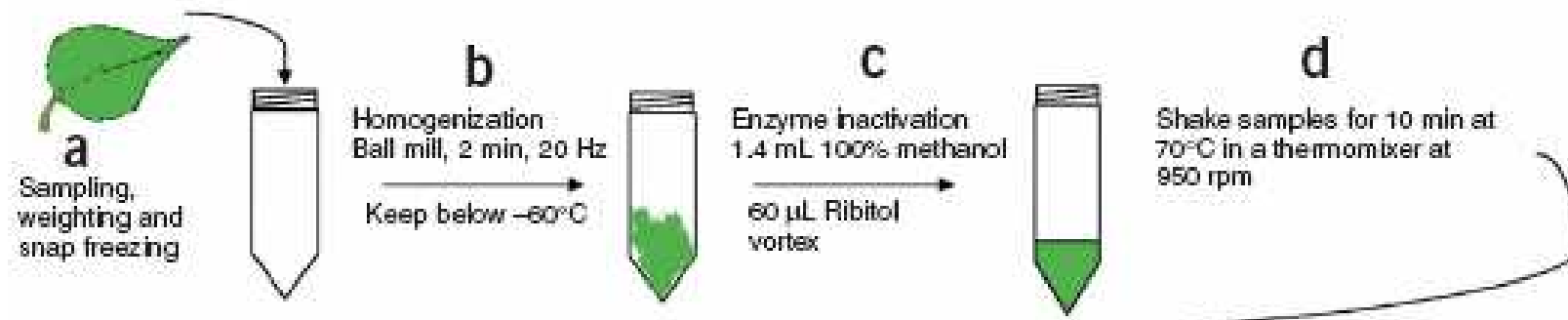
# GC-MS

*The more developed tool for metabolic profiling*

- Persuade the molecule to enter the vapor phase (derivatization)



## Extraction Protocol for Metabolic profiling of leaves



## Robotic derivatisation & full extract injection

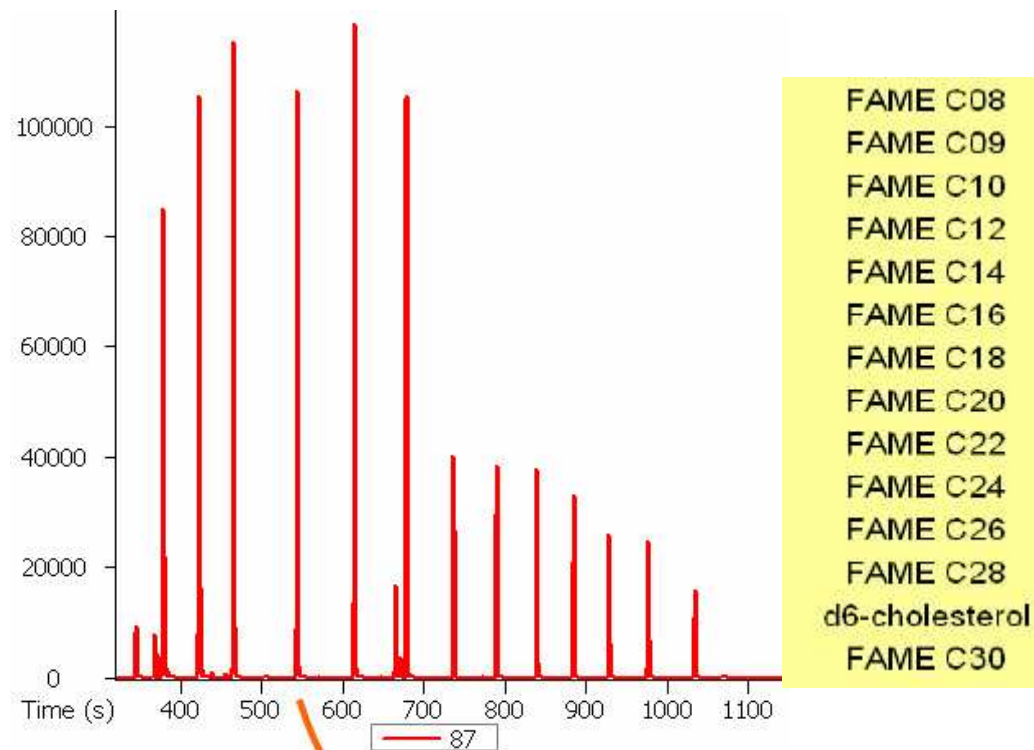
step reaction ovens: methoximation, silylation

2 robotic arms: 1  $\mu\text{L}$ , 100  $\mu\text{L}$

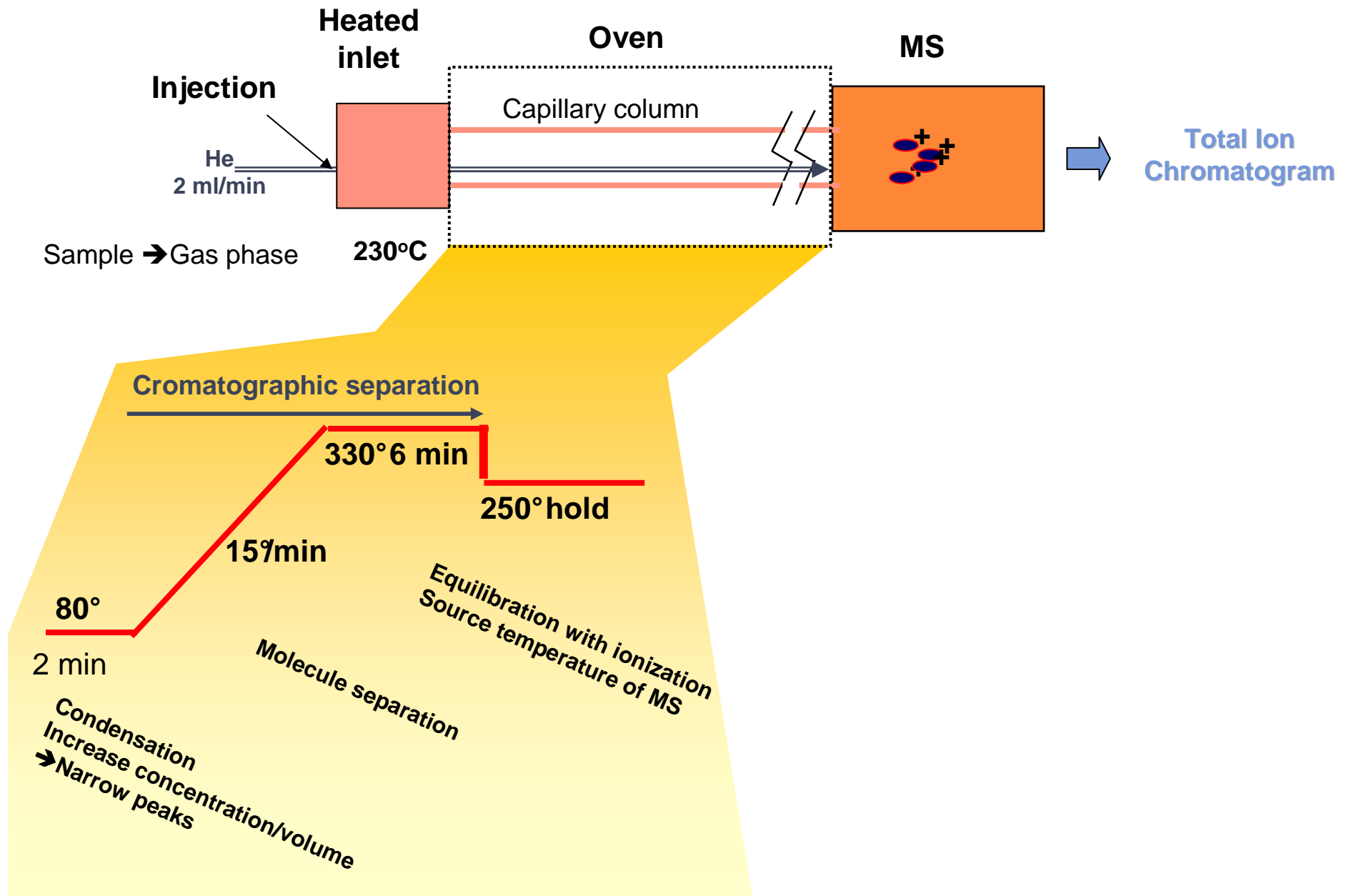
internal standards: RI, ribitol (or stable isotopes)

Retention index markers are mandatory.

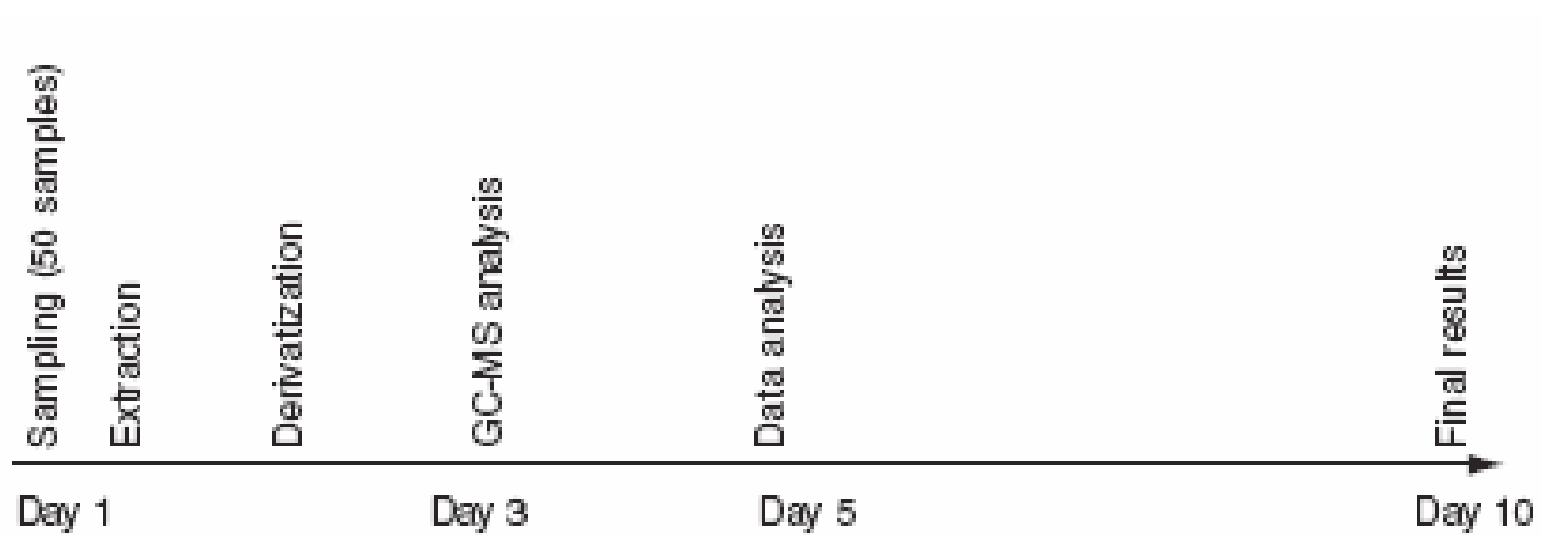
GCTOF: FAMES



# GC operation for our metabolic profiling protocol



## Timeline of standard operating procedure.



*Estimated amount of time from collecting 50 biological samples to final results*

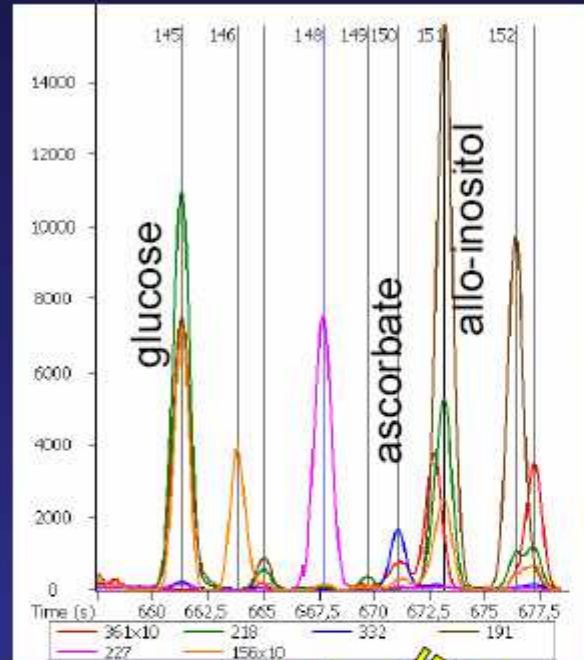
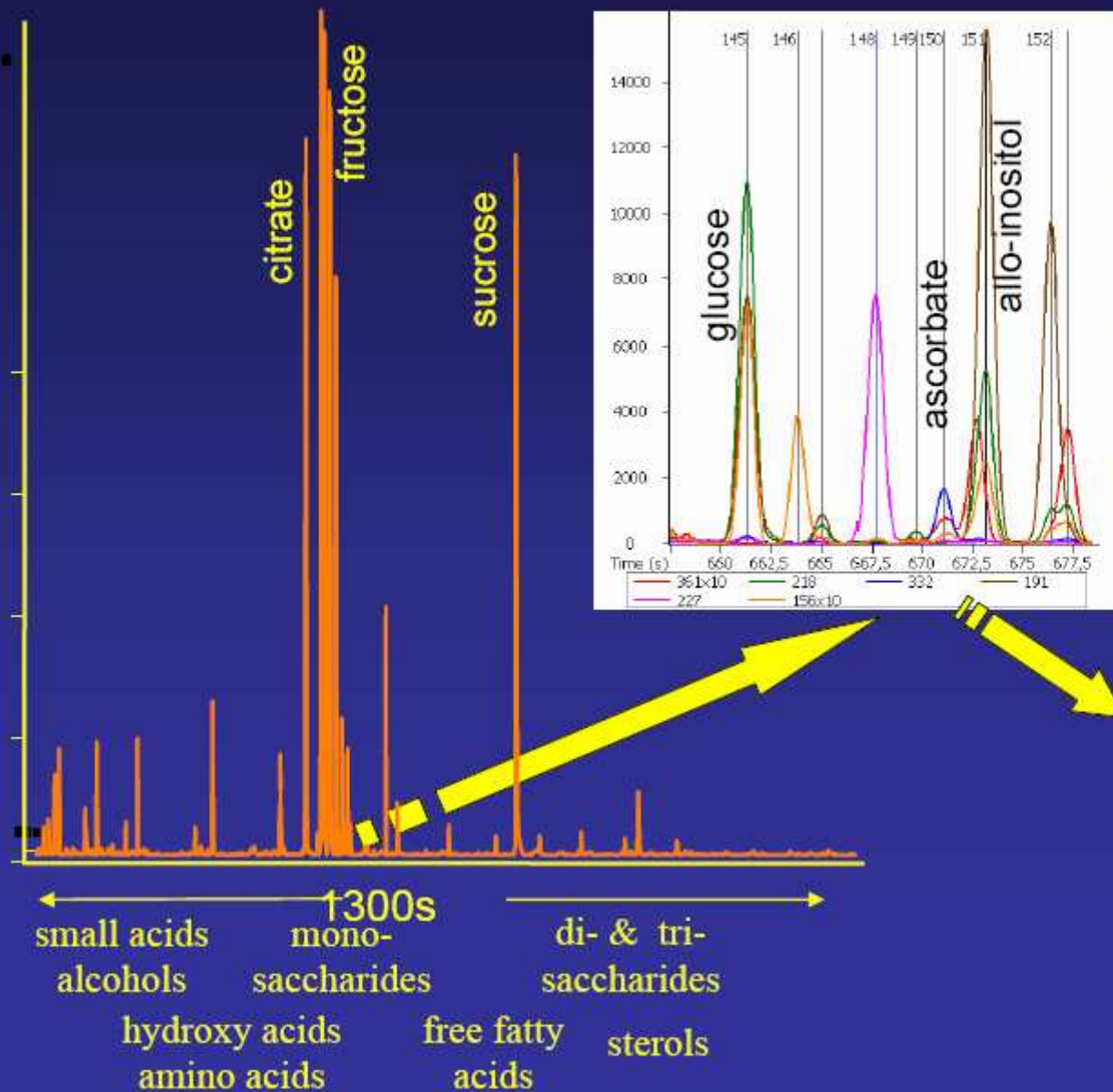
- ✓ Sampling + extraction + derivatization takes < 50% of the time,
- ✓ comprehensive data analysis: 2-5 days
- ✓ Manual inspection of chromatograms: 5 days: despite automatic algorithm, manual control is a highly advisable quality control

Metabolites identified from a methanol tissue extract of potato tubers

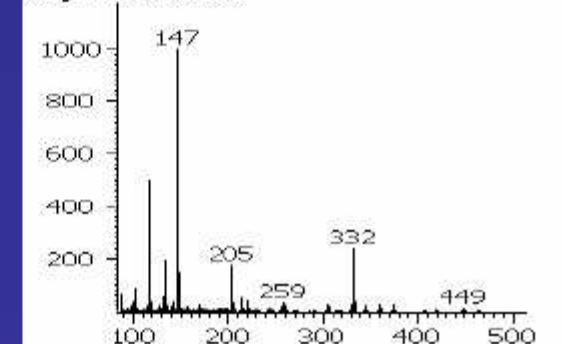
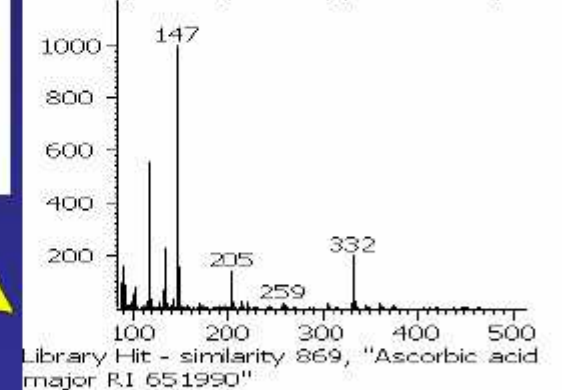
Amino acids	Organic acids	Sugars	Sugar alcohols	Aromatic amines
β-alanine (248)	2-aminoadipic acid (260)	erythrose (205)	erythritol (307)	dopamine (426)
alanine (116)	ascorbic acid (332)	fructose (307)	glycerol (205)	noradrenaline (355)
arginine (348)	benzoic acid (179)	fructose-6-phosphate (315)	inositol (305)	normetanephrine (297)
asparagine (188)	citric acid (465)	fructose (160)	maltitol (361)	octopamine (426)
aspartic acid (232)	fumaric acid (245)	galactose (160)	mannitol (319)	tyramine (338)
cystathionine (278)	glucoheptonic acid (435)	glucose (160)	xylytol/arabitol (319)	
cysteine (220)	gluconic acid (333)	glucose-6-phosphate (387)		
GABA (304)	glutaric acid (261)	isomaltose (480)		
glutamic acid (246)	glyceric acid (189)	mannose (160)		
glutamine (156)	hydroxybenzoic acid (267)	maltose (361)		
glycine (174)	isocitric acid (273)	raffinose (437)		
histidine (154)	α-ketoglutaric acid (198)	ribose (307)		
homocysteine (234)	lactic acid (117)	sucrose (451)		
homoglutamine (170)	malic acid (245)	trehalose (191)		
homoserine (218)	oxalic acid (147)	xylose/arabinose (307)		
isoleucine (158)	oxaloacetic acid (333)			
leucine (158)	6-phospho-gluconic acid (387)			
L-hydroxyproline (230)	3-phospho-glycerate (299)			
lysine (317)	phosphoric acid (299)			
methionine (176)	quinic acid (345)			
ornithine (142)	ribonic acid (307)			
phenylalanine (218)	shikimic acid (372)			
proline (142)	succinic acid (247)			
serine (116)				
threonine (291)				
tryptophane (202)				
tyrosine (218)				
valine (144)				

*Roesneer et al (2000)*

# MS deconvolution detects 500-1,000 peaks in leaves and 250-350 peaks in fruits. Screened against 420 (800) knowns.



Peak True - sample "060404abrsa58:2", peak 150, at 671,012 s (Spec # 6419)



GMD@CSB.DB: the Golm Metabolome Database -- Kopka et al. 21 (8): 1635 -- Bioinformatics - Microsoft Internet Explorer

Adresse <http://bioinformatics.oxfordjournals.org/cgi/reprint/21/8/1635>

## The Comprehensive Systems Biology Project (CSB)

# GMD@CSB.DB

## - The Golm Metabolome Database -

Hosted at Max Planck Institute of Molecular Plant Physiology

Databases: [Associated DB](#) | [Transcriptome DB](#) | [Metabolome DB](#) | [Co-Response DB](#)

[Home](#) | [What's new](#) | [Search](#) | [About CSB.DB](#) | [Site Map](#) | [Mail2Us](#) | [FAQ](#) | [Help](#)

**GMD@CSB.DB Navigation**

- GMD Page Tree
- GMD Analytics
- GMD MSRIs
- GMD Profiles
- GMD Tools
- GMD Service
- CSB.DB Team
- CSB.DB Service
- CSB.DB Contact
- CSB.DB Disclaimer

Last Modified: 2005-07-25  
© CSB.DB 2003-2005

**GMD@CSB.DB - MSRI: Mass Spectra Search**

**The Mass Spectra Search (sSPQ)**

[Info](#) | [Help](#)

Mass Spectra & Retention Time Index Libraries - Single Mass Spectrum Search Query (sSPQ): allows to search for similar mass spectra to get access to mass spectra libraries / information

**EASY**

1. **Mass Spectra**

**MS Type** NIST format: space-semicolon separated [e.g. 89\*40;\*101\*50;]

Default: NIST format (space-semicolon): space(=mass)-semicolon(=intensity) separated [e.g. 89\*40;\*101\*50;]

**ADVANCED**

2. **Platform Filter**

All technology platform: GC-TOF / GC-QUAD technology platform

Default: No restriction.

3. **RI Filter**

ignore retention time index (RI) based pre-search

**Contact**

For suggestions and questions feel free to contact the **CSB.DB curator**.

**Co-Response Module (affy2agi)**

**Update 2005-07-25**  
*A.thaliana*. Affymetrix to AGI Assignments updated. New co-response data matrices will be available soon... [query data](#) or [matrix info](#).

**Transcriptome Module (affy2agi)**

**Update 2005-07-25**  
*A.thaliana*. Affymetrix to AGI Assignments updated. AffymetrixID to AGI code converter will be available soon... [query data](#)

**Metabolome Module**

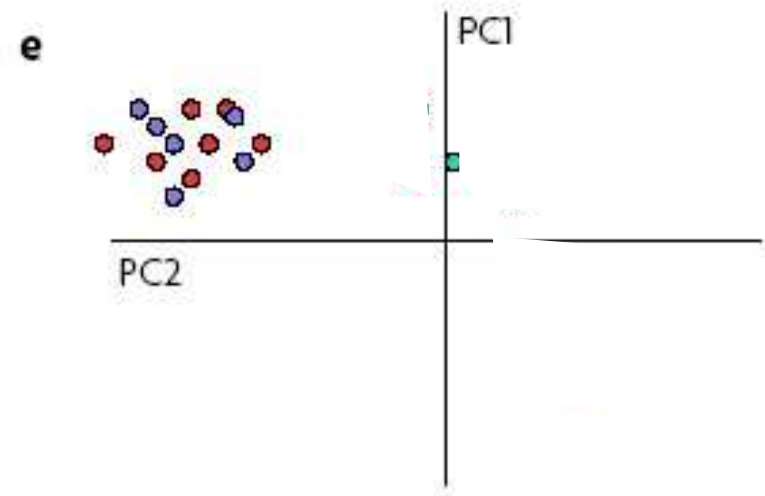
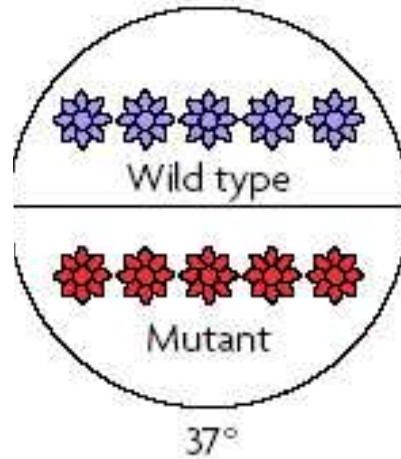
**Update 2005-07-25**  
GMD@CSB.DB - The Golm Metabolome Database: access to custom mass

Fertig Internet

Start Posteingang - ... CABBIO Brasil 2... class\_090807.ppt Class MdP.ppt GMD@CSB.DB: ... DE 03:35 a.m.

# Uses of metabolic profiling to study metabolism in plants

## Discovery of silent phenotypes

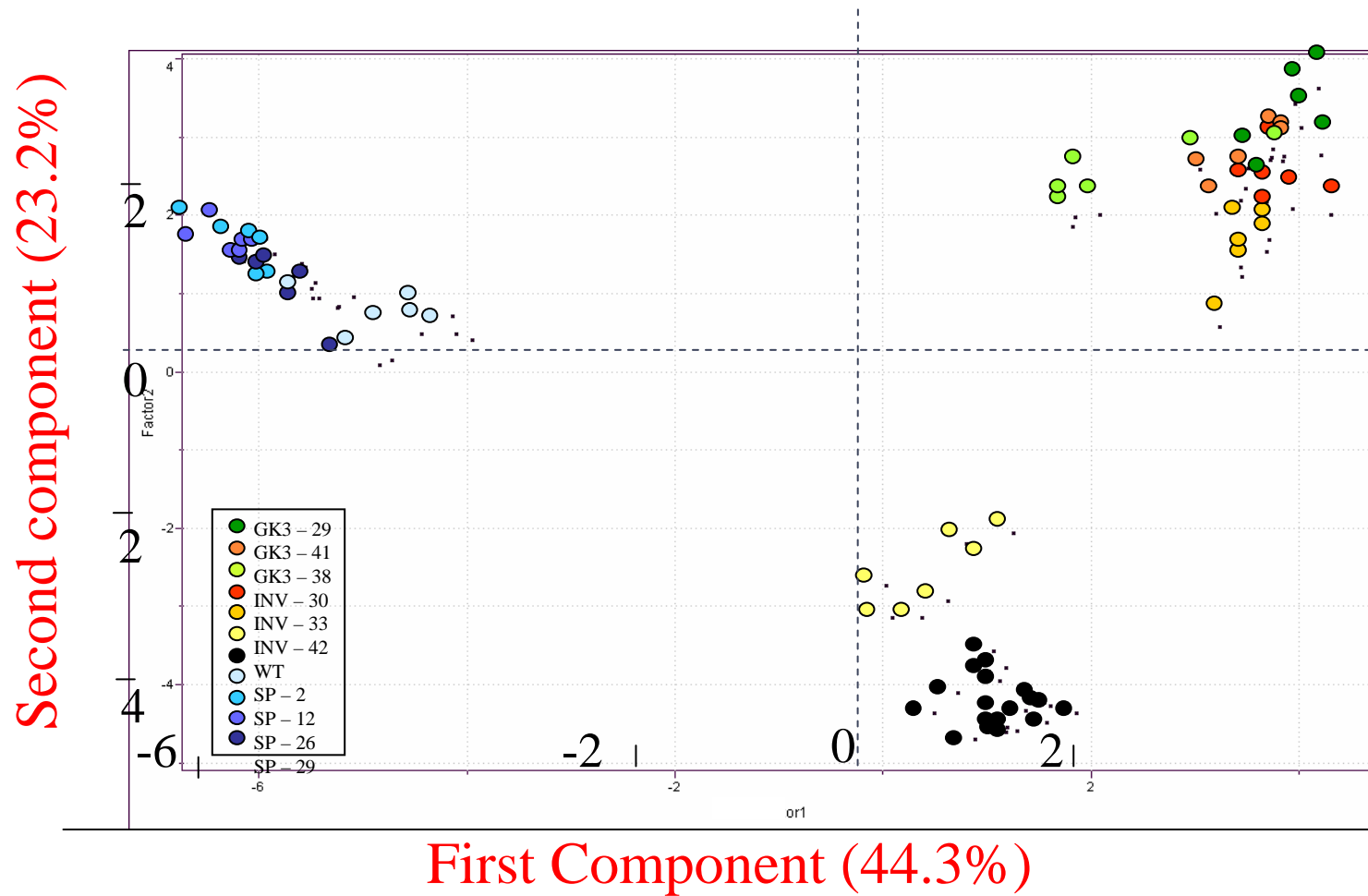


Large dataset produced by metabolic profiling analysis + multivariate analysis

→ **Discovery of a phenotype**

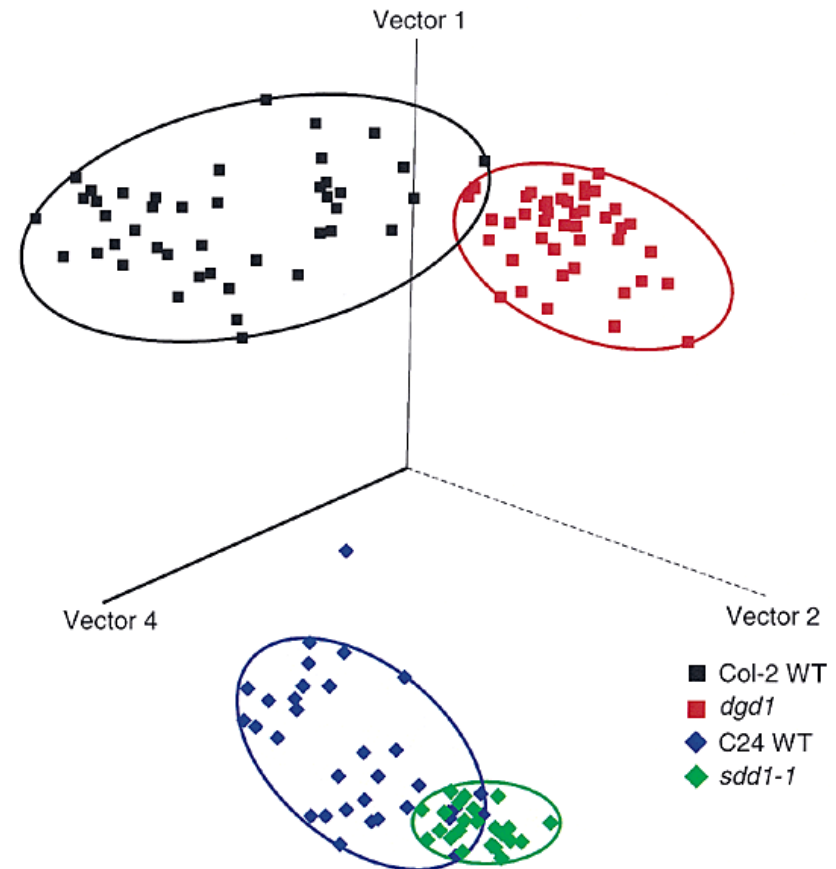
## Uses of metabolic profiling to study metabolism in plants

Several transgenic produced with increased levels of sucrose degradation:  
Enzymatic methods → similar phenotypes: low starch, high hexoses-phosphates...



## Uses of metabolic profiling to study metabolism in plants

- Fiehn et al. - GC-MS quantifies 326 distinct compounds in *Arabidopsis thaliana* leaf extracts



*Fiehn O et al. 2000 Nature Biotechnology, 18, 1157-1161.*

# LC-MS

## Main uses:

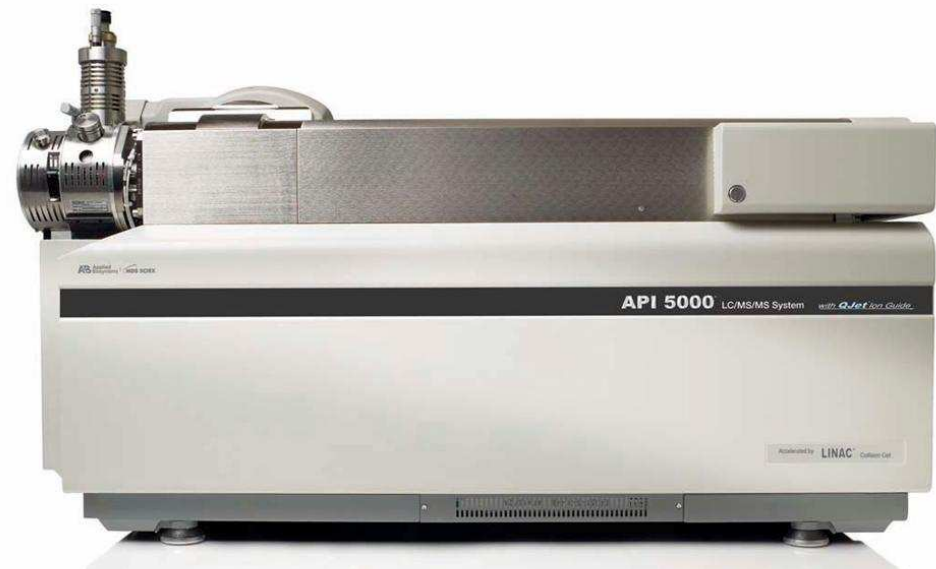
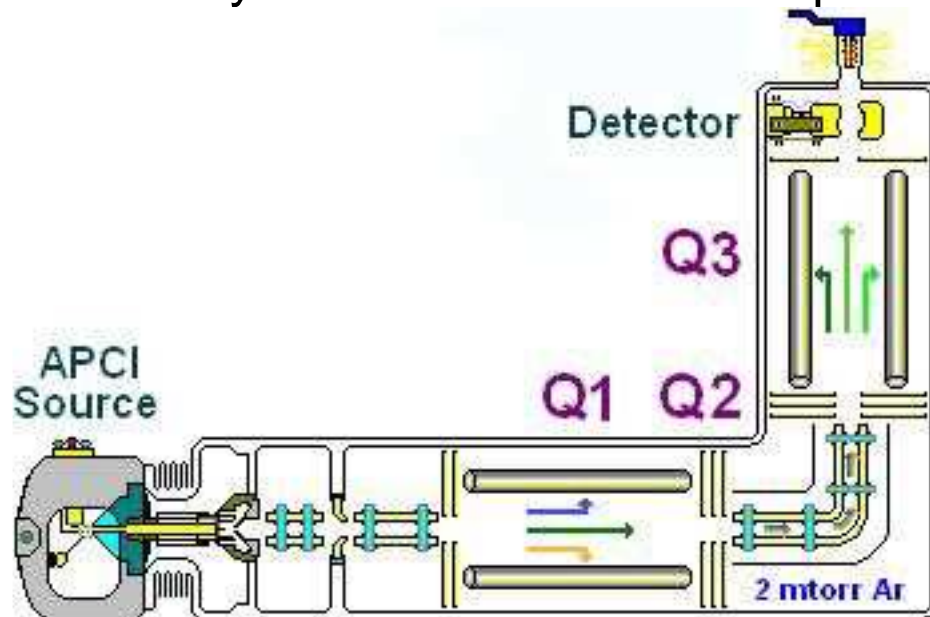
- to those analytes not amenable to GC-MS
  - Bigger molecules (>700 daltons) → non volatile
  - Unstable molecules: sugar phosphates, nucleosides, electron carries...

## Major analytical platforms:

LC- triple quadrupole, LC-Q-TOF, ...

## Actual significance:

analytical field with the more prominent development



### Amino acids

Alanine  
Arginine  
Asparagine  
Aspartate  
Glutamate  
Glutamine  
Glycine  
Histidine  
(Iso)leucine  
Lysine  
Methionine  
Phenylalanine  
Proline  
Serine  
Threonine  
Tryptophan  
Tyrosine  
Valine

### Amino acid derivatives/precursors

Citrulline  
Homoserine  
Ornithine  
*S-Adenosyl-L-methionine*

### Nucleoside bases

Adenine  
Guanine  
Hypoxanthine

### Nucleosides

Adenosine  
Cytidine  
Deoxyadenosine  
Guanosine  
Inosine  
Thymidine  
Uridine

### Nucleoside monophosphates

AMP  
CMP  
Cyclic-AMP  
dAMP  
dCMP  
dTMP  
GMP  
IMP  
UMP

### Xanthosine-5-P

### Nucleoside di/triphosphates

ADP  
ATP  
CDP  
CTP  
dTDP  
TTP  
UDP  
UTP

### Carbohydrate derivatives/precursors

3-Phosphoglycerate  
d-Hexose-P  
d-Rib(ul)ose-5-P  
Fructose-1,6-bis-P  
Glucosamine-6-P  
Glycerate  
Phosphoenolpyruvate  
Trehalose  
UDP-d-glucose  
UDP-N-acetyl-d-glucosamine

### Vitamins and derivatives

Biotin  
Carnitine  
*Folate*  
Nicotinamide  
Nicotinate  
Pantothenate  
Pyridoxine  
Riboflavin  
Thiamine  
5-Methyl-THF  
*7,8-Dihydrofolate*  
Thiamine-P

### Carboxylic acids

*Acetoacetate*  
Citrate  
Fumarate  
Malate  
*Oxaloacetate*  
Succinate  
Aconitate  
-Ketoglutarate

### Redox-electron-carriers and precursors

FAD  
FMN  
NAD<sup>+</sup>  
*NADH*  
NADP<sup>+</sup>  
*NADPH*  
Oxidized glutathione, Reduced glutathione  
Quinolate

## Metabolites quantified In *E. coli* using a LC-triple quadrupole

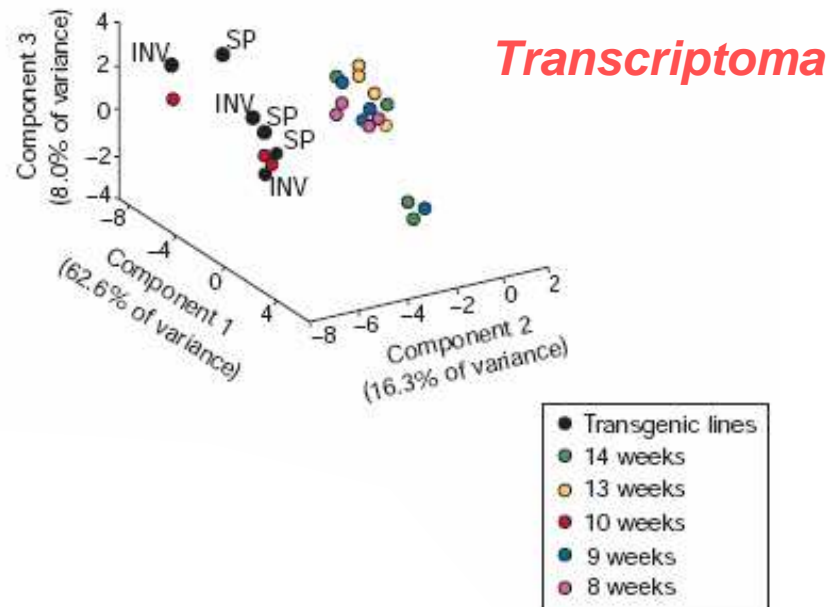
Bajad et al, 2006

# Parallel analysis of transcript and metabolic profiles: a new approach in systems biology

Ewa Urbanczyk-Wochniak, Alexander Luedemann, Joachim Kopka, Joachim Selbig, Ute Roessner-Tunali,  
Lothar Willmitzer & Alisdair R. Fernie<sup>+</sup>

Max-Planck-Institut für Molekulare Pflanzenphysiologie, Golm, Germany

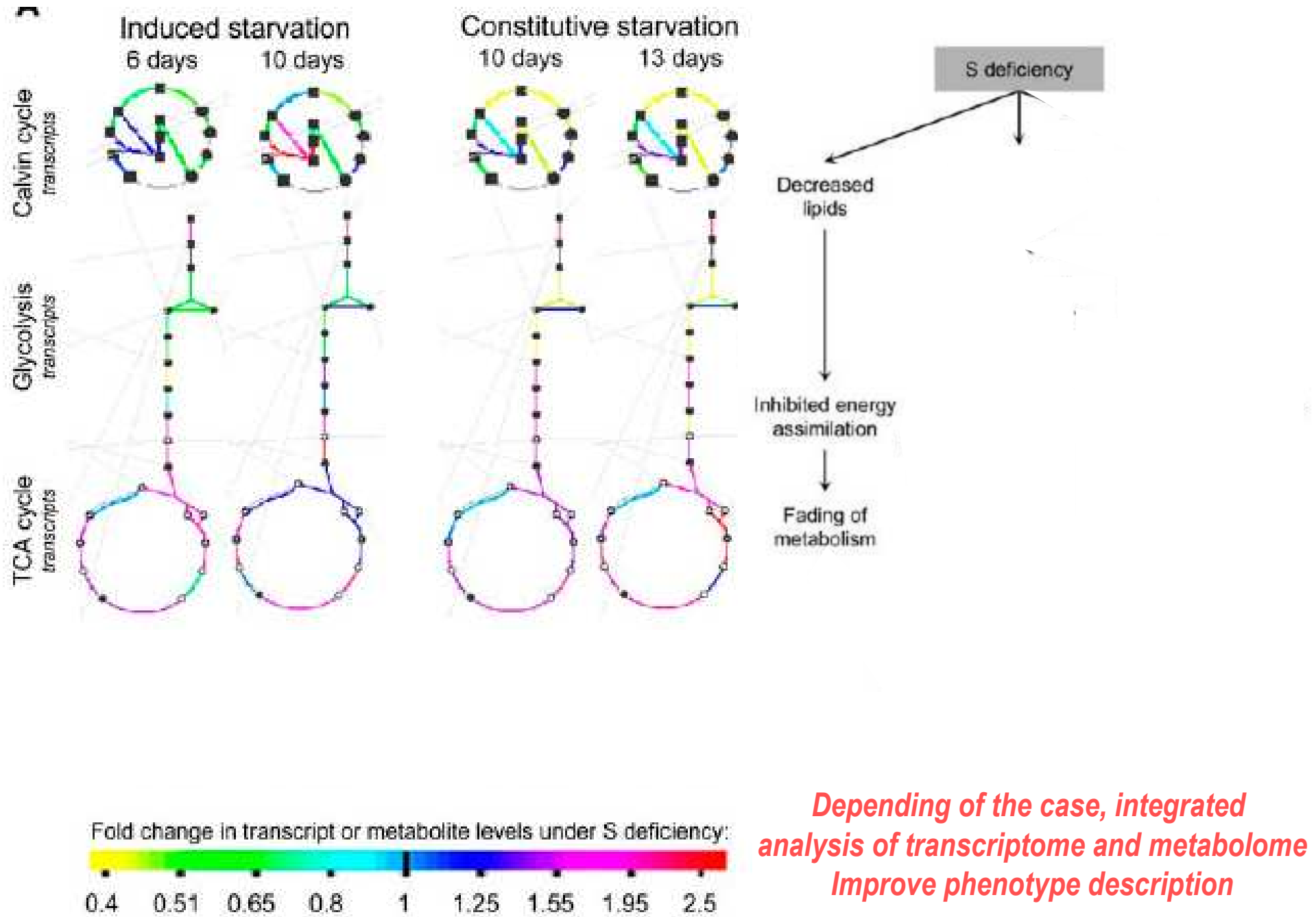
A



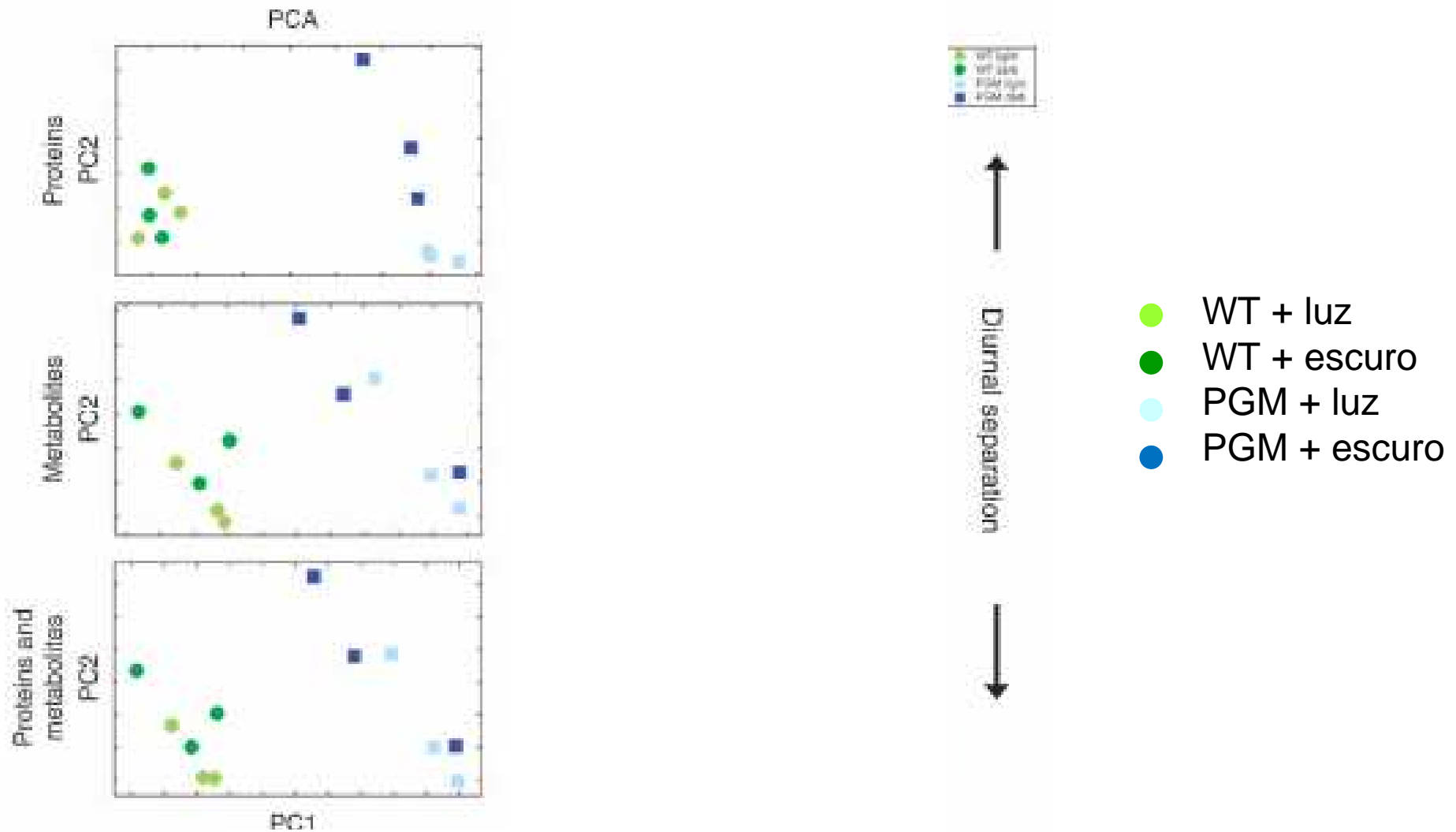
***Example of better performance of  
metabolomics over  
Transcriptomics por phenotype discovery.***

**General rule or particular case??**

# Sulphur deprivation: Integrated analysis of transcriptome and metabolome



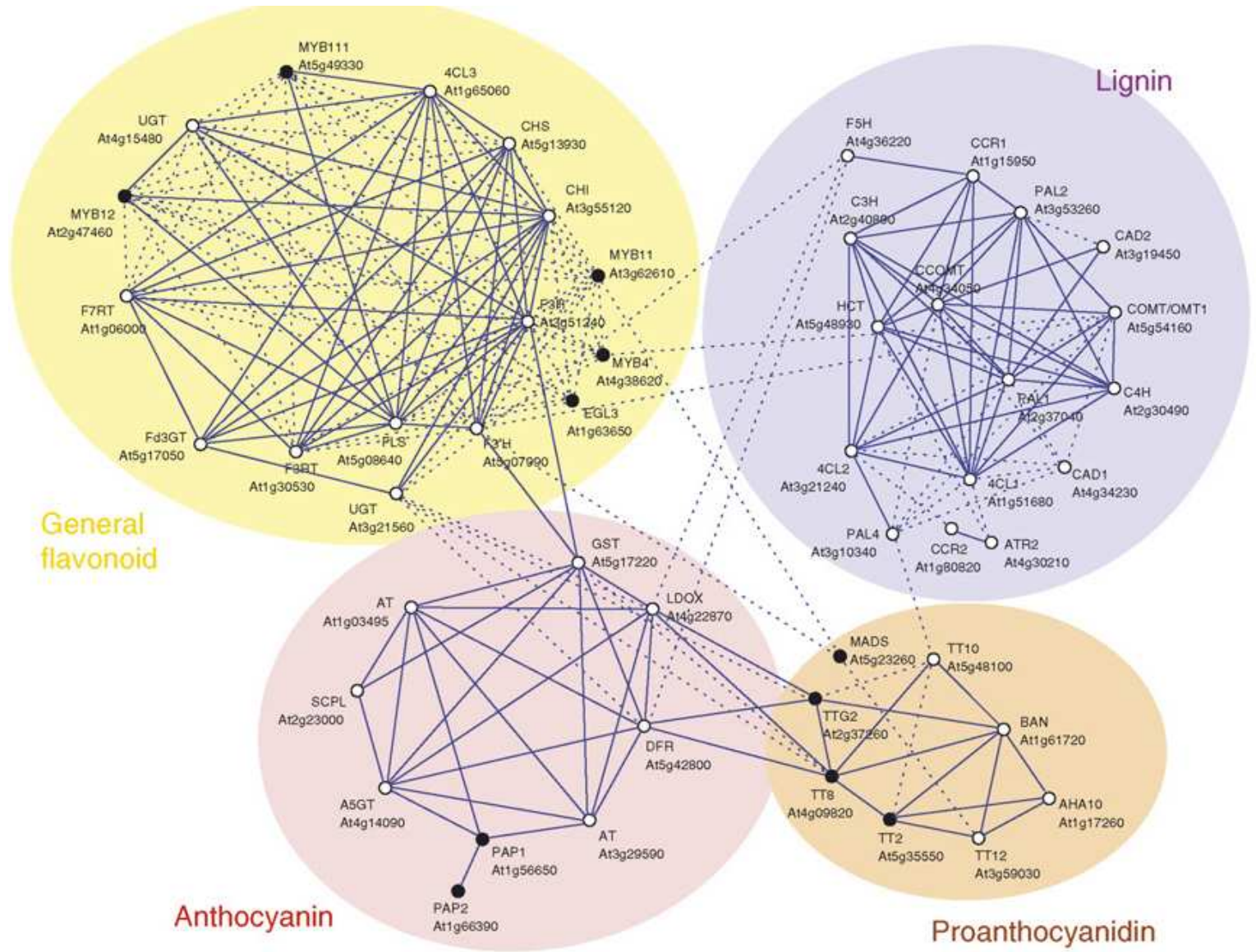
# Phenotype discrimination using metabolomic and proteomic data

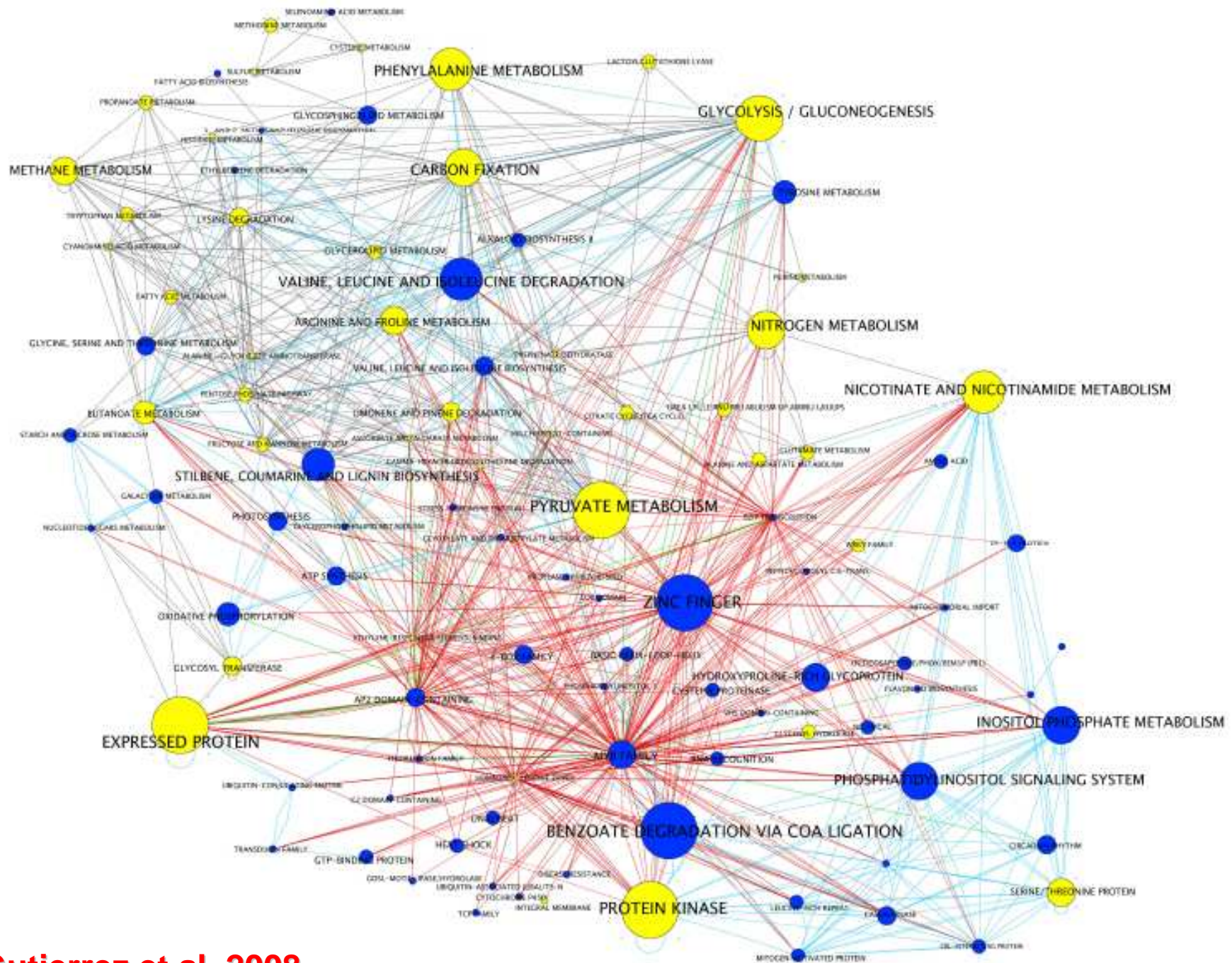


*Exemple of the need to develop more bioinformatic tools.*

# Coexpression analysis in plants

→ systemic gene expression network





Gutierrez et al, 2008

# Conclusão

## **Fatos atuais:**

✓ Tecnologias em rápido desenvolvimento: grandes avanços nos analisadores de massa, aumento das sequências em bibliotecas

## **Desafios atuais:**

✓ Apenas um pequeno número de metabólitos (1000/20000) pode ser detectado e separado com confiança, mas grande quantidade destes (60%) sem determinação de sua natureza química (fórmula química)

✓ Dificuldade de compreender grande quantidade de informação sobre alterações do metabolismo sem ter uma visão do controle e interação entre diferentes rotas

✓ Ausência de bibliotecas para LC-ESI-MS

## **Resultados importantes já demonstrados:**

- Capacidade de detectar fenótipos silenciosos

- Capacidade de distinguir fenótipos ou efeitos semelhantes antes considerados semelhantes

- Seu uso em conjunto com transcriptoma e proteoma permite uma análise sistêmica da biologia e geração de inúmeras hipóteses de trabalho

- Geração de um conjunto muito importante de informações para entender a regulação do metabolismo e o papel dos metabólitos como parte de uma rede de sinalização

## Our basic infraestructure:

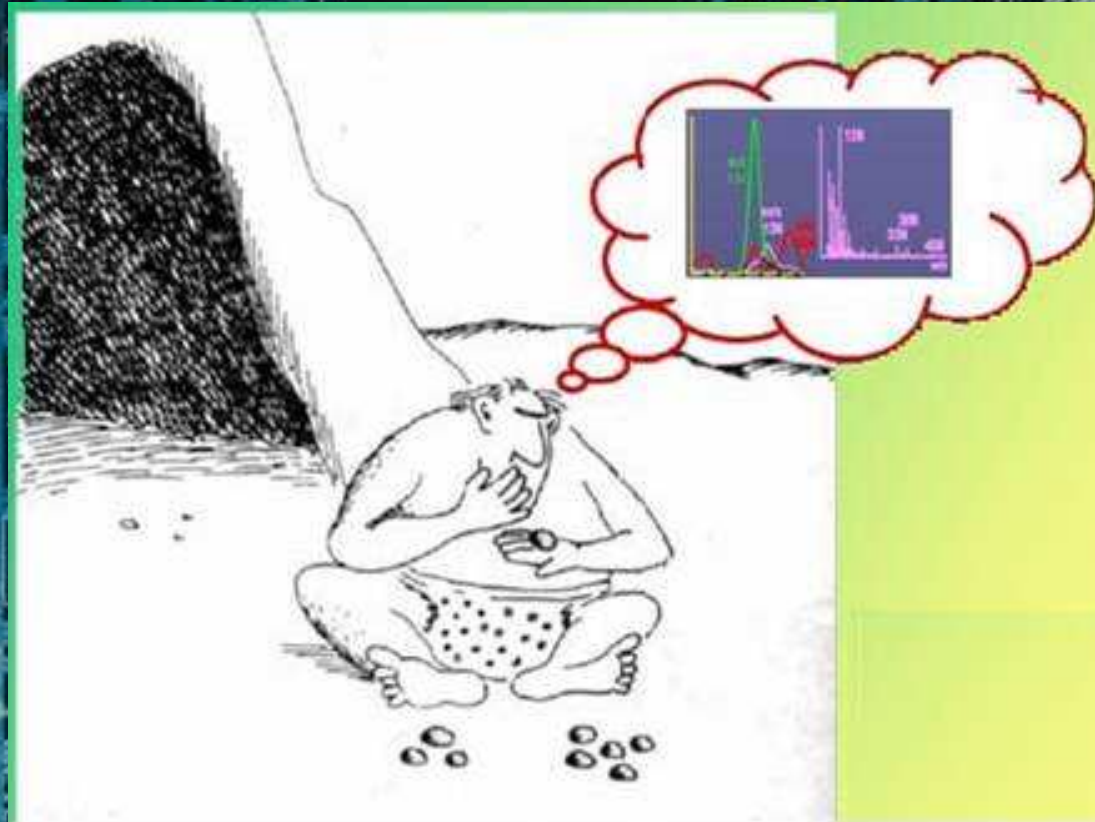


→ 5 greenhouses (1 for transgenics), and almost all equipment needed for basic Physiological characterizations: IRGAs (3), fluorometers (2), image fluorescence High resolution infrared camera, Scholander pumps, osmometer, etc



- Very good analytical equipment for metabolomics and proteomics
- Two tissue culture labs molecular physiology lab (real time PCR, nanoHPLC, elisa-readers...), ecophysiology lab, plant nutrition-metabolism lab, plant development, post-harvest labs.

## Ultimas palavras



→ Nova biologia ainda em sua infância.

→ Descobrir a função dos genes em larga escala, futurismo ainda hoje.

→ Há um rico acervo de informações nos bancos de mutantes, bancos ou publicações de transcriptoma, proteoma e metabolômicas ainda não explorados

Obrigado