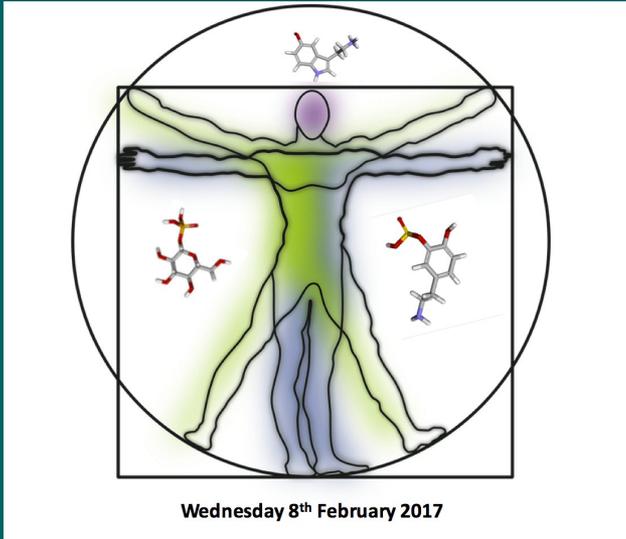


# Metabolomics: molecules of life

## An introduction



Reza Salek PhD

Metabolism and Molecular Informatics

The European Bioinformatics Institute (EMBL-EBI)

Email: [Reza.salek@ebi.ac.uk](mailto:Reza.salek@ebi.ac.uk)

# Some Definitions

## metabolome

/mɪˈtəbələʊm/

noun

BIOCHEMISTRY

noun: **metabolome**; plural noun: **metabolomes**

the total number of metabolites present within an organism, cell, or tissue.

ENGLISH

metabolism

ENGLISH

genome

metabolome

1990s

## metabolomics

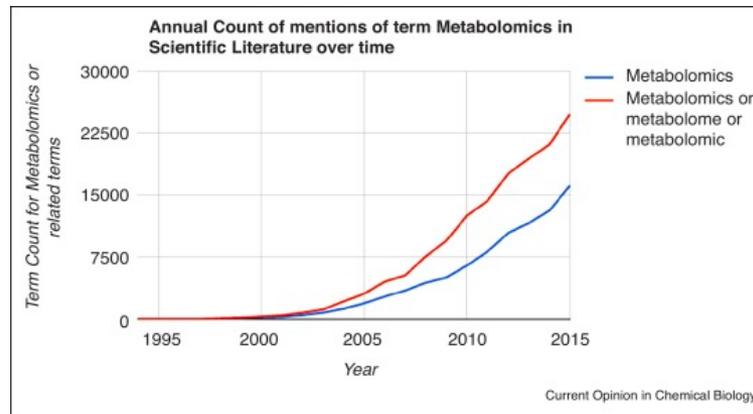
mɪˈtəbələmɪks/

Noun

BIOCHEMISTRY

noun: **metabolomics**

the scientific study of the set of metabolites present within an organism, cell, or tissue, often by measuring simultaneously (100s -10000s), many of which are not identified (features or analyts, )



ENGLISH

metabolome

ENGLISH

-ics

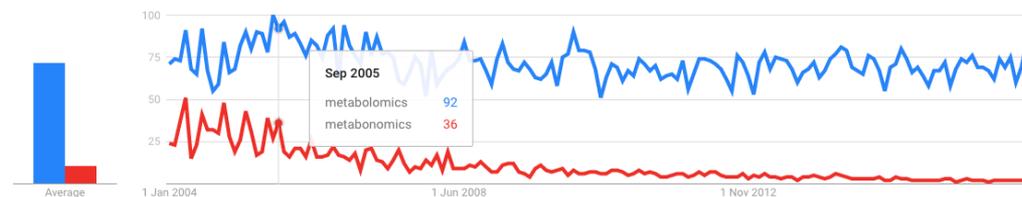
ENGLISH

genomics

metabolomics

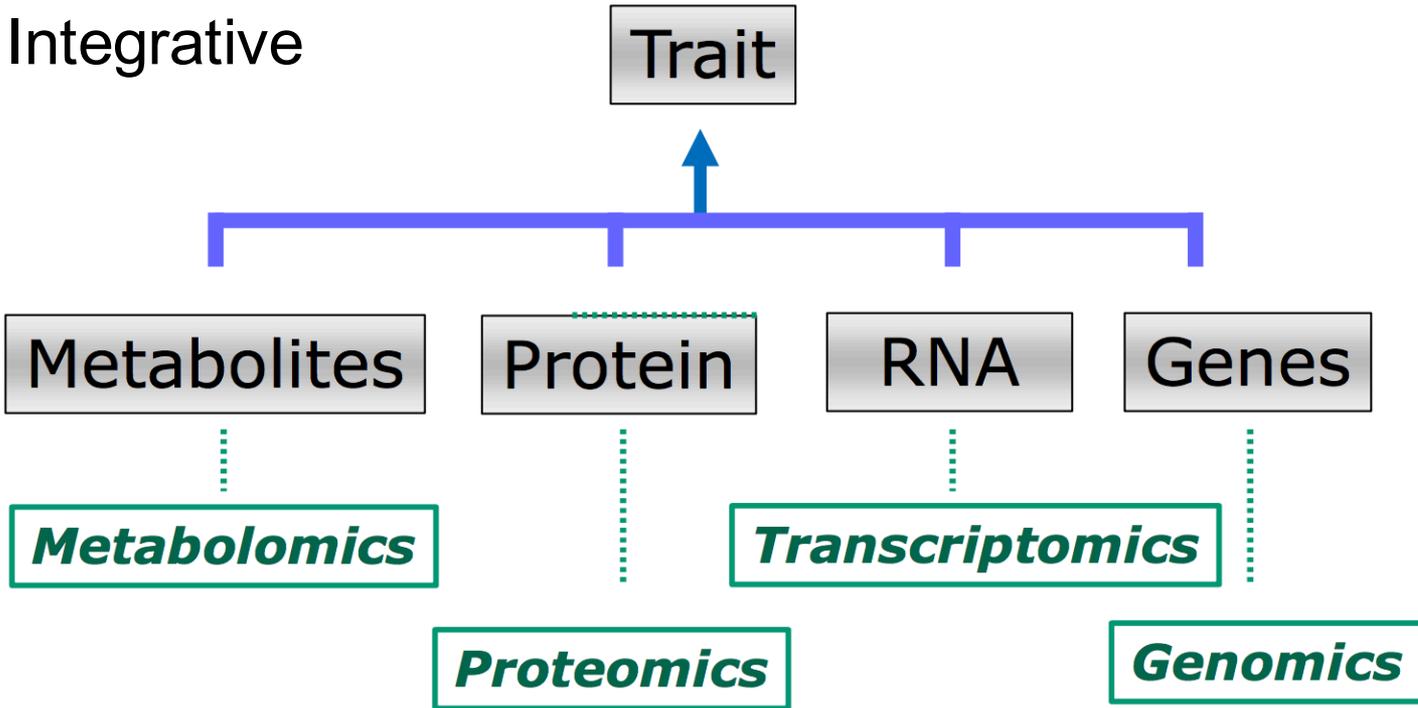
1990s

Interest over time

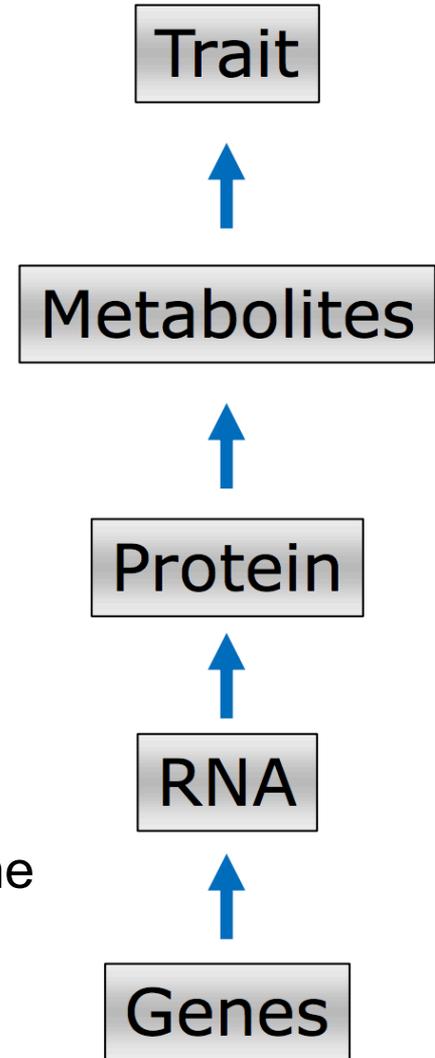


# Central dogma in Biology

Integrative



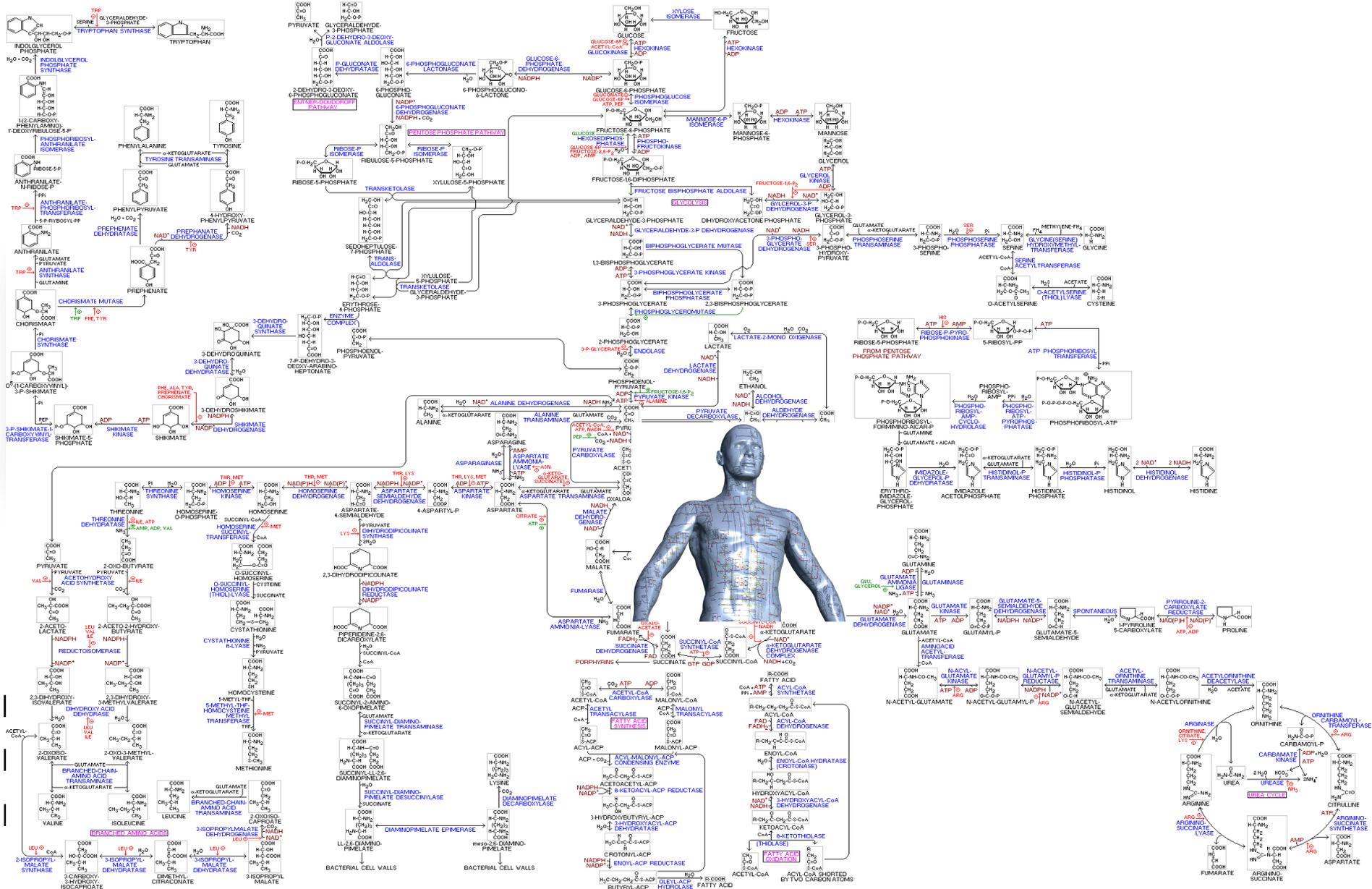
Reductionist



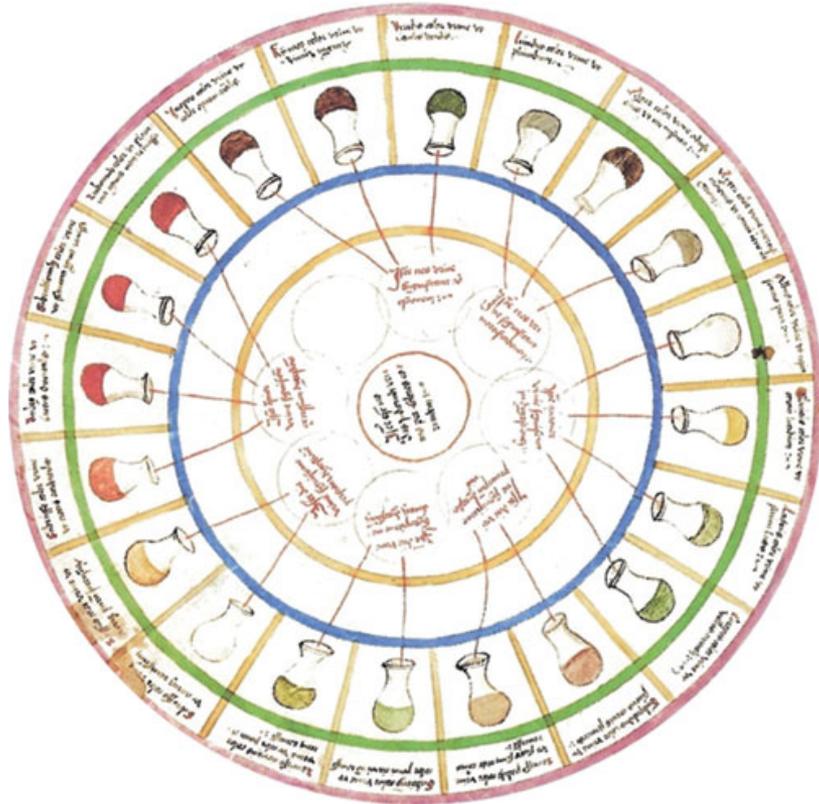
- The system (cell) is more than the sum of its parts
- To understand the system we must study the system not the parts
- e.g. Silkworm metamorphosis



# Complex Nature of biology



# Uroscopy – Early metabolomics

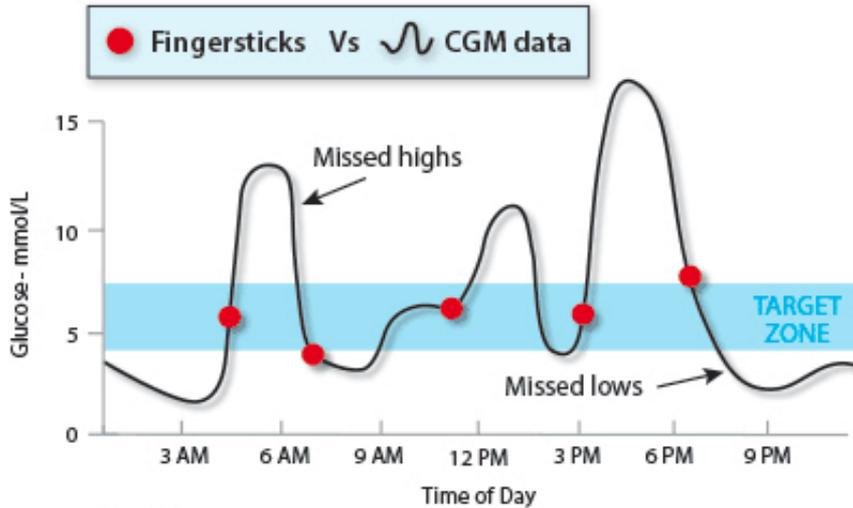


- It dates back to ancient [Egypt](#), [Babylon](#), and [India](#). It was particularly emphasized in [Byzantine medicine](#).
- The wheel describes the possible colors, smells and tastes of urine, and uses them to diagnose disease.

The Urine Wheel for diagnosing metabolic diseases, from *Epiphania Medicorum* by Ullrich Pinder in 1506



# Bio Marker Diagnosis and Monitoring

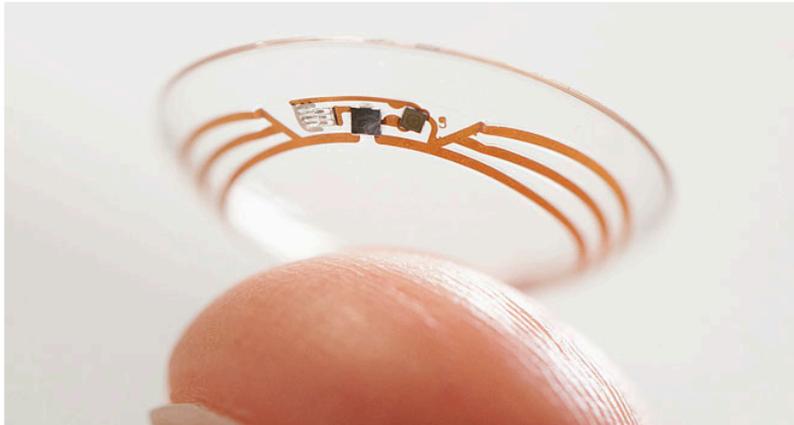


\*Illustration purposes only

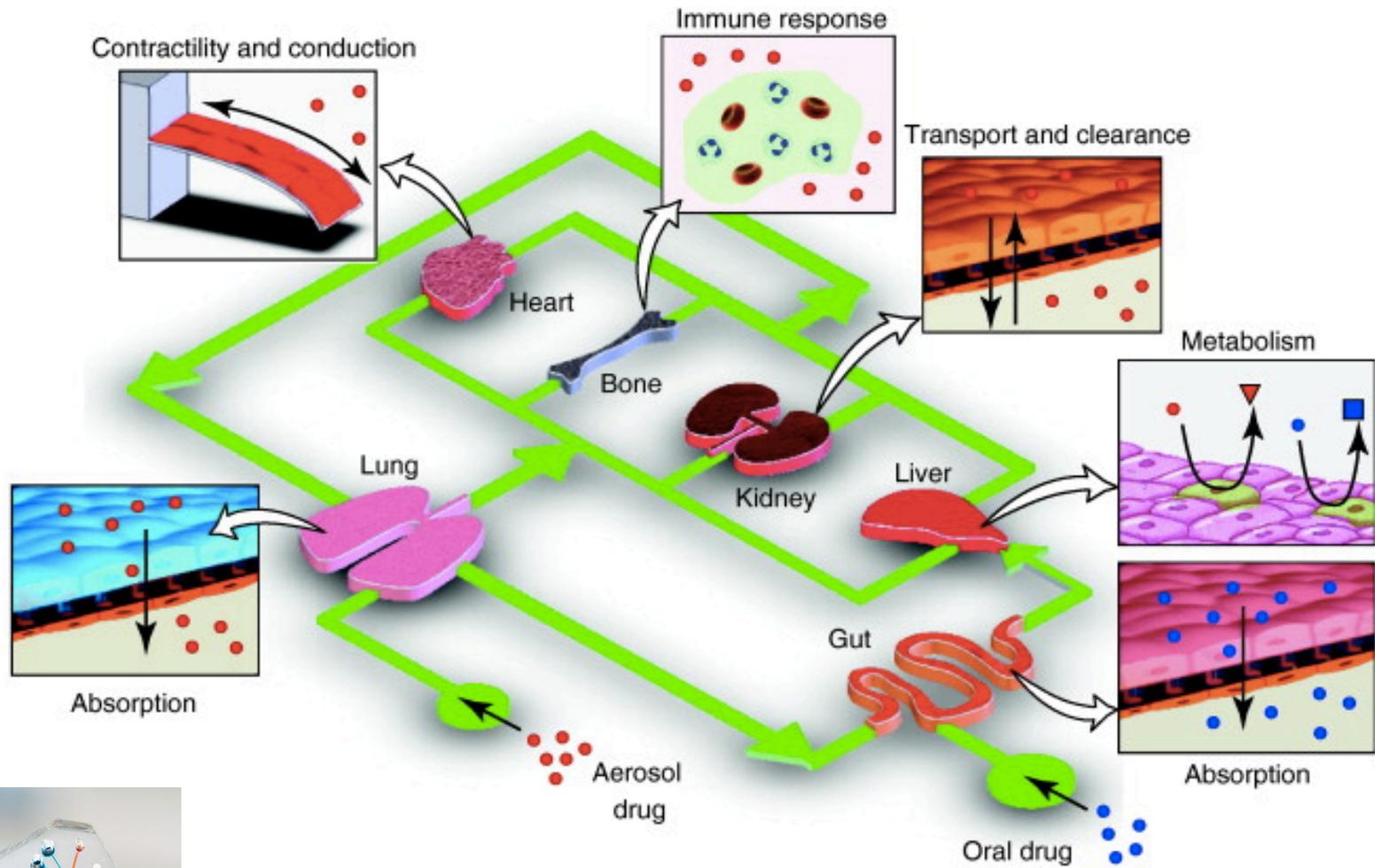


## Google announces 'smart' contact lenses that monitor glucose levels

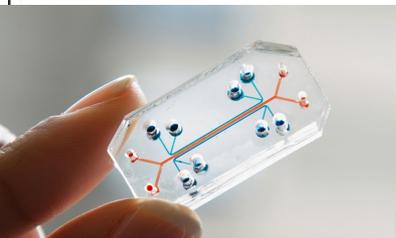
Published January 16, 2014 / FoxNews.com



# Metabolomics on a chip



TRENDS in Cell Biology



# Omics by numbers

'Omics'	results	Number
Genomics	DNA sequence	100,000-1,000,000
Transcriptomics	Gene expression	10,000-100,000
Proteomics	Protein expression	1,000-50,000
Metabolomics	Compounds	100-10,000*

\*It is estimated that all plant species contain 90,000 - 200,000 compounds. Each individual plant species contains about 5,000 – 30,000 compounds

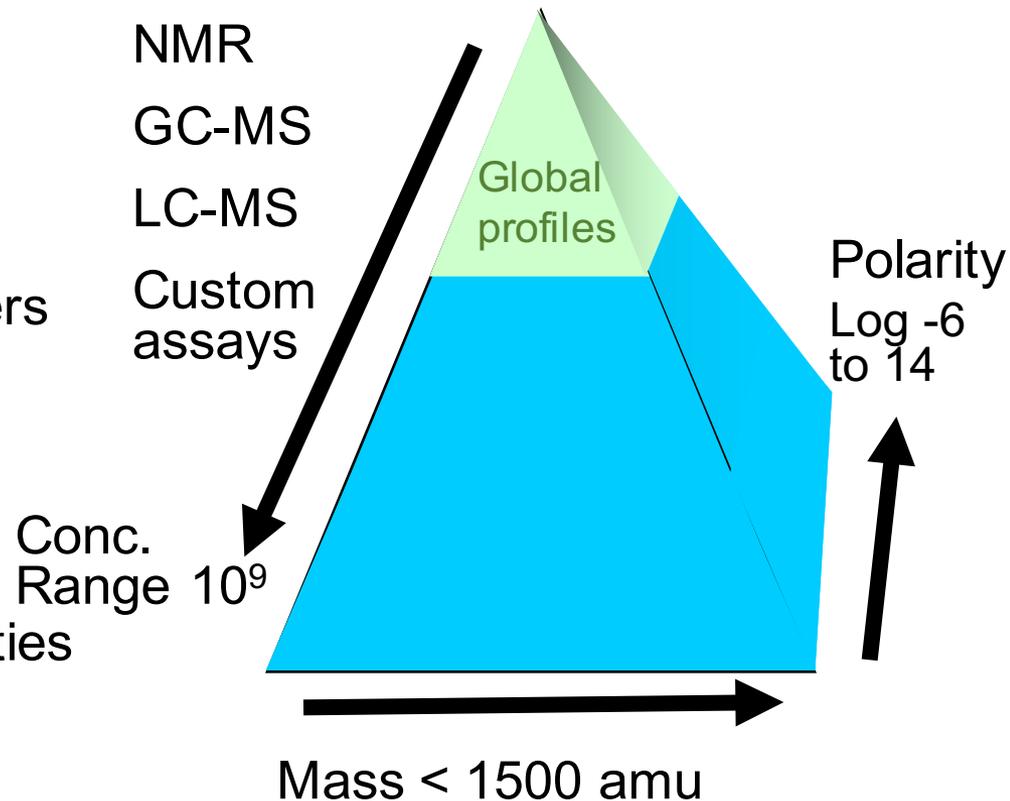
- Metabolites have a vast range of chemical structures and properties. Their molecular weights span two orders of magnitude
- No single extraction or analysis method works for all metabolites
- Many metabolite levels change with half times of minutes or seconds – far faster than nucleic acids or proteins

# Cell size vs Metabolite – by Lane et al,

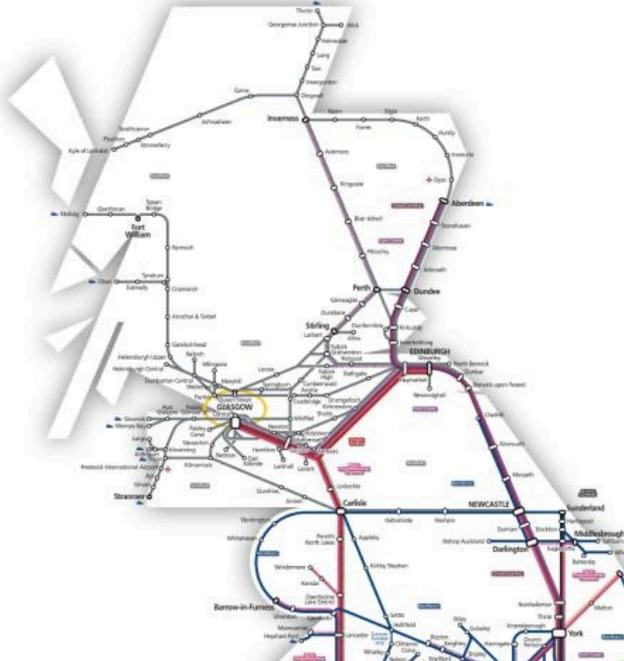
- Human cells range from 0.2-20 pL in volume
  - Biomass accounts for ca. 30 % of total mass (ash is very low)
  - Cells are ca. 20% w/v protein.
  - Human cells have 6 pg DNA
  - Variable RNA
- <10% “metabolites”
- 1 million cells of 1 pL volume make about 1 mm<sup>3</sup>
- Biomass of 1 million cells: 60-6000 pg biomass or 20-2000 pg metabolites
- **Effective metabolite concentrations:**
  - 1-10 mM (very abundant) = 1-10 fmol/ cell
  - 1-100 μM (typical) = 1-100 amol/cell
  - 1-100 nmol (rare) = 1-100 zmol/cell
- Concentrations depend on **pool sizes and compartmentation**: lymphocytes cell nucleus occupies >50% of internal volume, epithelial cells <10%.
- **Mitochondria numbers vary** from 1-1000 each ca. 1 fL, occupying a substantial fraction of the cell volume

# Dynamic range of metabolome

- ❑ Size varies “only” 2 orders of magnitude (~150 to 1500 Da)
- ❑ Abundance varies hugely (many orders of magnitude)
- ❑ Large number of structural isomers
  - ❑ i.e Lipids (TAG)
    - ❑ 40 common fatty acids
    - ❑ 3 fatty acids per TAG → 64000 species
- ❑ Hugely different chemical properties
  - Hydrophobicity
  - Polarity
  - Volatility
  - pKa

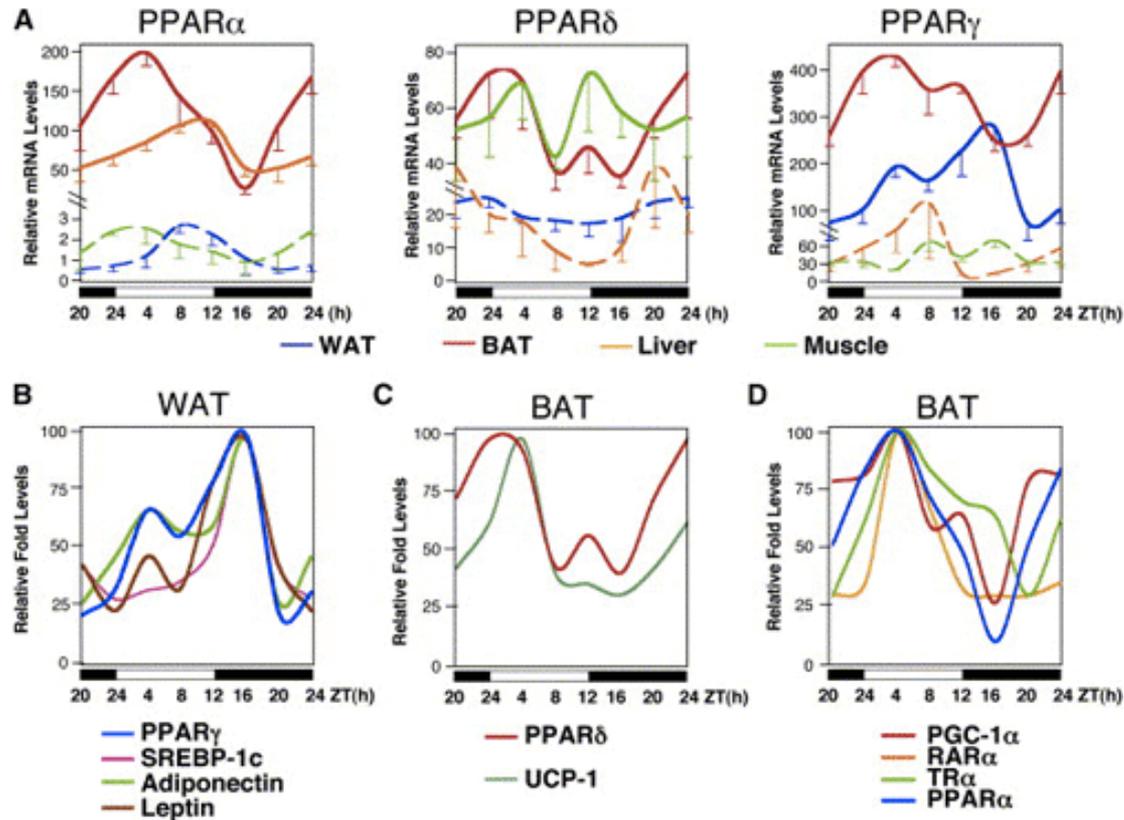






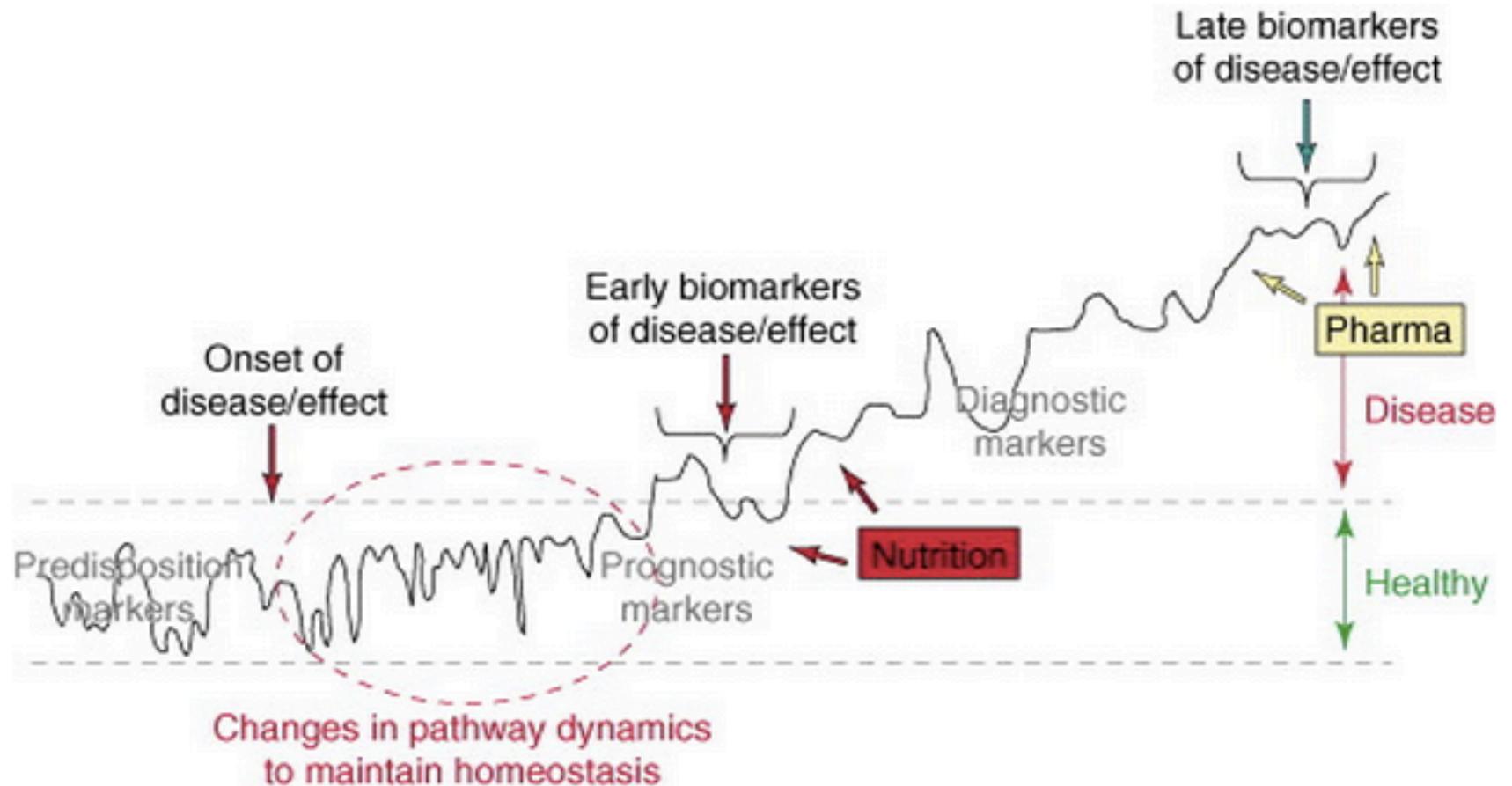
© 2012 Andrew Smithers  
1000 000 1 2011/12  
\*Locally set or fixed rates per day but there may be other costs

# Diurnal Rhythm



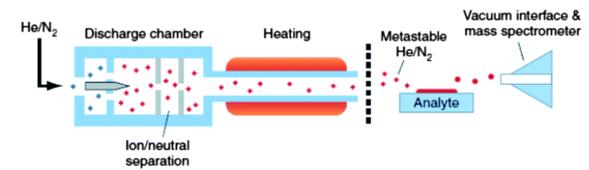
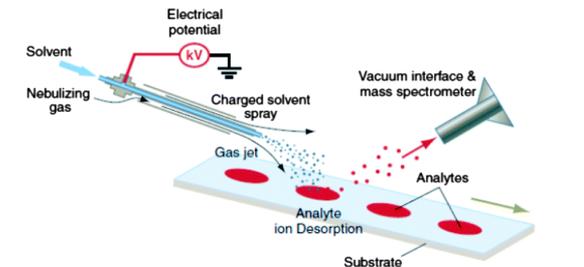
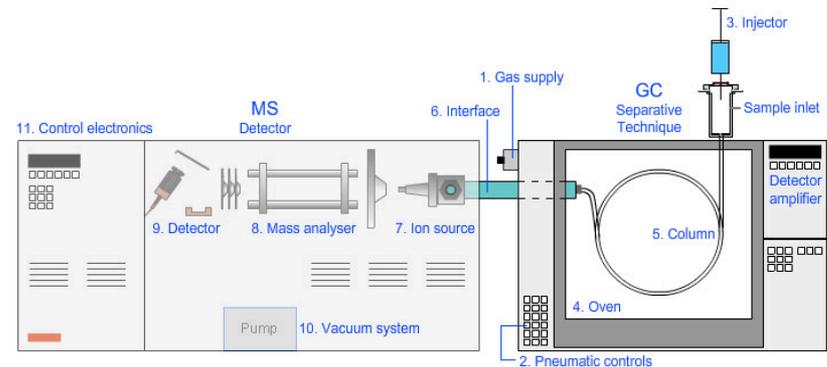
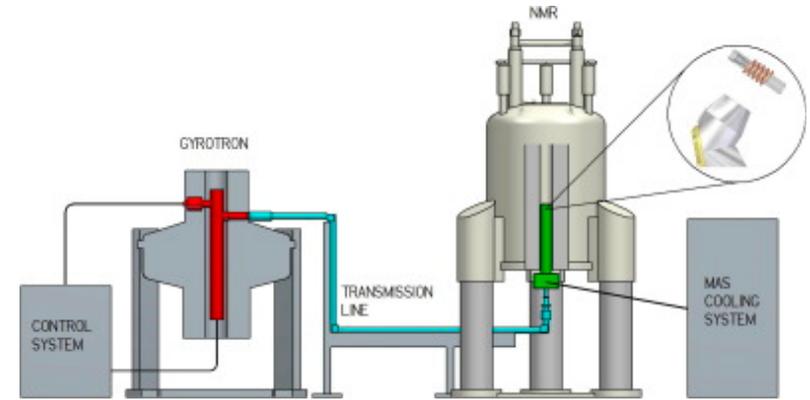
Tissue-Specific Diurnal Rhythm of PPARs and Target Gene Expression. From Yang et al. (2006) Cell 126:801-810.

# Why is it important



# Global profiling tool

- NMR spectroscopy
  - Solution state
  - Solid state (HR-MAS)
- GC- MS Spectrometry
- LC-MS spectrometry
  - TOF
  - QQQ
  - ESI
- Imaging-MS
  - MADLI -MS
  - DESI-MS



# Instrument progress

**FullHD**  
1920x1080

**4K**  
3840x2160

**8K UHD TV**  
7680x4320



# Exploiting High Mass Accuracy to ID Compounds

<u>Type</u>	<u>Mass Accuracy</u>
FT-ICR-MS	0.1 - 1 ppm
Orbitrap	0.5 - 1 ppm
Magnetic Sector	1 - 2 ppm
TOF-MS	3 - 5 ppm
Q-TOF	3 - 5 ppm
Triple Quad	3 - 5 ppm
Linear IonTrap	50-200 ppm (10 ppm in Ultra-Zoom)

$$\text{ppm} = \left( \frac{m_{\text{exp}} - m_{\text{calc}}}{m_{\text{exp}}} \right) 1 \text{E} + 6$$

# NMR pro and cons

## Advantages of NMR

- NMR is the only detection technique which does not rely **on separation** of the analytes. This means that the sample is **not destroyed**.
- All kinds of small molecule metabolite can be measured **simultaneously**.
- Sample analysis is **fast** (~7 minutes) and robust, enabling high throughput.

## Disadvantages of NMR

- It is **not as sensitive** as MS techniques.
- As a consequence high-resolution NMR of intact biofluids does not yet identify all the metabolites.
- Cannot analyse the **organic layer** from cell/tissue extracts very easily.

# Mass spectrometry pro and cons

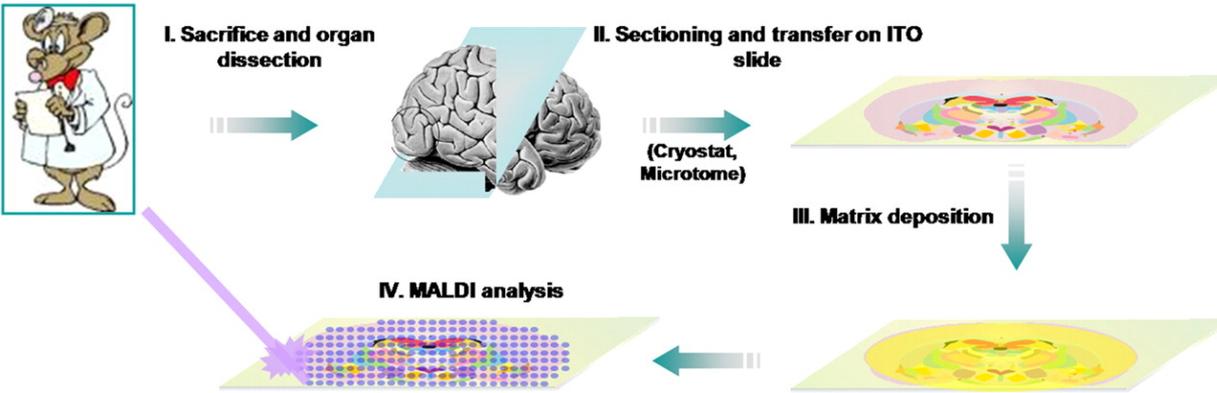
## Advantages of MS

- It is much more **sensitive** than NMR. This means we can study more metabolites.
- 30-100 metabolites can be identified from NMR data whilst GC-MS allowed the identification of >300 metabolites in the aqueous and > 1000 for LC-MS.
- Not limited to samples type/ potentially **more global**
- Takes up much **less space** than an NMR machine and **cheaper** to buy

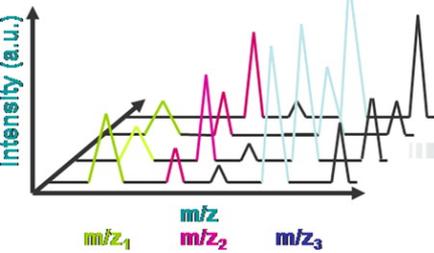
## Disadvantages of MS

- Similarity of **isomers** can make identification difficult. For LC-MS **libraries** are not complete
- Samples must be **derivatised** before analysis which increases the preparation time.
- Sample analysis time is **longer** than NMR and are less reproducible
- Some large molecules cannot be measured by this technique.

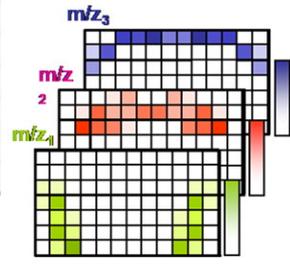
# Imaging Mass spectrometry



V. Data processing and image reconstruction



	x1	x2	x3	x4	...	...	xn
y1	$I_{1,1}$	$I_{1,2}$	$I_{1,3}$	$I_{1,4}$	...	...	$I_{1,n}$
y2	$I_{2,1}$	$I_{2,2}$	$I_{2,3}$	$I_{2,4}$	...	...	$I_{2,n}$
y3	$I_{3,1}$	$I_{3,2}$	$I_{3,3}$	$I_{3,4}$	...	...	$I_{3,n}$
y4	$I_{4,1}$	$I_{4,2}$	$I_{4,3}$	$I_{4,4}$	...	...	$I_{4,n}$
y5	$I_{5,1}$	$I_{5,2}$	$I_{5,3}$	$I_{5,4}$	...	...	$I_{5,n}$
y6	$I_{6,1}$	$I_{6,2}$	$I_{6,3}$	$I_{6,4}$	...	...	$I_{6,n}$
y7	$I_{7,1}$	$I_{7,2}$	$I_{7,3}$	$I_{7,4}$	...	...	$I_{7,n}$



METASPACE EMBL

# Applications of metabolomics is growing



# MetaboLights – metabolomics data sharing

MetaboLights > Search

## Search results

Filter your results

Type

- study  
 compound

Technology

- mass spectrometry  
 NMR spectroscopy

Organism

Organism Part

Validations Status

- green  
 red  
 amber

Validations Status Details

222 results, showing 1 to 10

Page 1 of 23

### Metabonomics of human fecal extracts characterize ulcerative colitis, Crohn's disease and healthy individuals.

Study Identifier	<a href="#">MTBLS237</a>
Study Size	45.66MB
Submitted by	<a href="#">Karen Atkins</a>

Organism	Homo sapiens
Study Factors	Diagnosis, Age, age at diagnosis, Years with disease , Family disposition, Smoker, Steroid response, Steroid dependency, HB-score, Mayo-score, Extraintestinal manifestations, Drug Treatment Asetylsalicylic acid, Drug Treatment Local Asetylsalicylic acid, Drug Treatment Azathioprine, Drug Treatment 6-mercaptopurine, Drug Treatment Salazopurine, Drug Treatment Systemic steroids, Drug Treatment Local steroids, Drug Treatment Infliximab , Disease Characterisation , Extension,

### N-glycosylation Profiling of Colorectal Cancer Cell Lines Reveals Association of Fucosylation with Differentiation and Caudal Type Homebox 1 (CDX1)/Villin mRNA Expression.

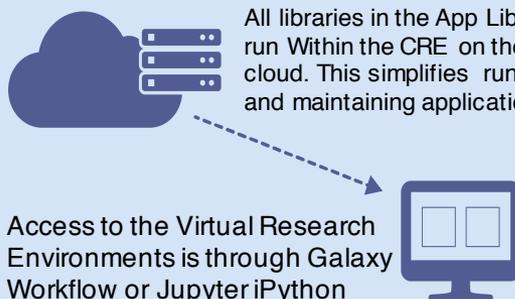
Study Identifier	<a href="#">MTBLS227</a>
Study Size	4.74GB
Submitted by	<a href="#">Stephanie Holst</a>

Organism	Homo sapiens
Study Factors	Biol. Replicate, Tech. replicate, Tumour site, Stage, CDX1/Villin, Sample type,

# Large scale computing for medical metabolomics



All libraries in the App Library run Within the CRE on the cloud. This simplifies running and maintaining applications



Access to the Virtual Research Environments is through Galaxy Workflow or Jupyter iPython coding web environment



App Library - Service Catalogue

App Library showcases our service catalogue listing applications that are available via Galaxy workflows and Jupyter libraries through the Cloud Research Environment.

Search for Apps

 <b>Iso2Flux</b> Open source software for steady state 13C flux analysis	 <b>MIDoor</b> MIDoor corrects raw isotopic isomers spectra for natural occurring isotopes and overlapping of peaks for several metabolites in m/z scale	 <b>W4M Multivariate</b> PCA, PLS(D), and OPLS(D) for multivariate analysis of omics data
 <b>rNMR</b> Open source software for identifying and quantifying metabolites in NMR spectra.	 <b>SOAP-NMR</b> An R package for 1H-NMR data pre-treatment.	 <b>W4M Univariate</b> Univariate statistics.

Functionality

- Preprocessing
- Annotation
- Post-processing
- Statistical Analysis
- Workflows
- Other Tools

Approaches

- Metabolomics
- Isotopic Labelling Analysis
- Lipidomics
- Glycomics

Instrument Data Types

- MS
- NMR
- IR
- Raman
- UVVIS
- DAD

The open-source App Library

An easy to use, cloud based scalable software infrastructure for metabolomics

PhenoMeNal Cloud Research Environment (CRE)  
All your favourite metabolomics applications under one hood.

# Some interesting references

- Brown, M. et al. A metabolome pipeline: From concept to data to knowledge. *Metabolomics*. 2005;1(1), 39-51
- Di Guida, R. et al. Non-targeted UHPLC-MS metabolomic data processing methods: a comparative investigation of normalisation, missing value imputation, transformation and scaling. *Metabolomics*. 2016;12(5), 1-14
- Cambiaghi, A. et al. Analysis of metabolomic data: tools, current strategies and future challenges for omics data integration. *Briefings in Bioinformatics*. 2016, 1-13
- Summer, L.W. et al. Metabolomics data analysis, visualization, and integration. *Methods in Molecular Biology*. 2007;406(4), 409-436
- Rocca-Serra, P. Data standards can boost metabolomics research, and if there is a will, there is a way. *Metabolomics*. 2016;12(1), 1-13
- Vinayavekhin, N. and Saghatelian, A. Untargeted Metabolomics. *Current Protocols in Molecular Biology*. 2010;30.1, 1-24
- Yi, L. et al. Chemometric methods in data processing of mass spectrometry-based metabolomics: A review. *Analytica Chimica Acta*. 2016;914, 17-34
- Lind, M. et al. The use of mass spectrometry for analysing metabolite biomarkers in epidemiology: methodological and statistical considerations for application to large numbers of biological samples. *European Journal of Epidemiology*. 2016;31(8), 717-733

## Possible other interesting resources

<http://metabolomicssociety.org/resources/tutorials>

<http://www.metabolomics-forum.com/>

[https://masspec.scripps.edu/mshistory/whatisms\\_toc.php](https://masspec.scripps.edu/mshistory/whatisms_toc.php)

<http://www.ebi.ac.uk/training/online/course/introduction-metabolomics/what-metabolomics>

<http://www.metabonews.ca/archive.html>

# Online course

## Train online

[Training](#)[Train online](#)[About Train online](#)[Glossary](#)[Support and feedback](#)[Login/register](#)

### Metabolomics: An introduction

#### What is metabolomics?

Small molecules

The metabolome and metabolic reactions

#### The importance of metabolomics

Some applications of metabolomics

#### Designing a metabolomics study

Key stages of a metabolomics study

Metabolomics resources at the EBI

Summary

Quiz: Metabolomics quiz

Your feedback

Learn more

References

Contributors

### Metabolomics: An introduction

This course provides a basic introduction into the rapidly emerging field of metabolomics and its importance and applications.

An undergraduate-level understanding of biology would be an advantage.

#### About this course

**Author(s):** [Reza Salek](#)

 [Systems](#)

 [Chemical biology](#)

 [Beginner](#)

 1 hour

[Start the course](#)

#### Learning objectives:

- Comprehend the purpose and importance of the field of metabolomics
- Describe some principles of metabolomic study design
- Evaluate advantages and limitations of some analytical techniques used in metabolomics studies
- Discuss some of the modern-day applications of metabolomics
- Access metabolomics resources at the EMBL-EBI

#### Rating:



# Upcoming webinars

See the full list of upcoming webinars at  
<http://www.ebi.ac.uk/training/webinars>

EMBL-EBI Services Research Training About us

## Training Webinars

[Overview](#) [Train at EMBL-EBI](#) [Train outside EMBL-EBI](#) [Train online](#) [Webinars](#) [About](#) [Contact](#)

[Training](#) > [Webinars](#)

### Filters

**Filter by: Topic**

- DNA & RNA
- Gene expression
- Proteins
- Structures
- Systems
- Chemical biology
- Ontologies
- Literature
- Cross domain

**Filter by: Year**

- 2017
- 2016
- 2015
- 2014

### Webinars

Our webinar series is a freely accessible collection of live seminars focused on EMBL-EBI resources. They provide both brief introductions to a range of databases and more in-depth coverage of new features and tools to assist you in your research. Each webinar is delivered by an EMBL-EBI expert and is followed by a question and answer session. The webinars are recorded and continue to be freely available through [Train online](#).

#### February 2017

**1** [EMBL-EBI resources: an introduction](#)

1st Feb

The EMBL-EBI is the home of the world's most comprehensive range of freely available molecular databases and resources. Our resources help researchers share and analyse data and perform complex queries in many different ways. Join Tom Hancocks in this webinar for a tour of the core data resources at EMBL-EBI.

First come, first served

[Bioinformatics](#) [Molecular biology](#)

**8** [Metabolomics: molecules of life, an introduction](#)

8th Feb

## Feedback

Tell us what you think