

## Terminology

- **Metabolite**: substance produced or used during metabolism such as lipids, sugars and amino acids
- **Metabolome**: the quantitative complement of all the low molecular weight molecules present in cells or biofluids in a particular physiological or developmental state
- **Metabolomics**: a comprehensive analysis of the whole metabolome under a given set of conditions

## Metabo\*omics

- **Metabonomics**:

The quantitative measurement of the dynamic multiparametric response of living systems to pathophysiological stimuli or genetic modification  
(*Nicholson et al., Xenobiotica 1999*)

holistic analysis of biofluids and tissues in order to determine metabolic composition

deals with integrated, multicellular, biological systems including communicating extracellular environments in animal and human biochemistry

- **Metabolomics**:

Measurement of metabolite concentrations and fluxes in isolated cell systems (*Nicholson et al., Xenobiotica 1999*)

deals with simple cell systems and mainly intracellular metabolite concentrations in microbial and plant biochemistry

## Metabolomics vs genomics and proteomics

Genomics and proteomics tell you what **might** happen, but metabolomics tells you what actually **did** happen  
(*Bill Lasley, UC Davis*)

Although changes in the quantities of individual enzymes might be expected to have little effect on metabolic fluxes, they can and do have significant effects on the concentrations of numerous individual metabolites.

The metabolome is further down the line from gene to function and so reflects more closely the activities of the cell at a functional level. Thus, as the 'downstream' result of gene expression, changes in the metabolome are expected to be amplified relative to changes in the transcriptome and the proteome.

Metabolic fluxes (at least as exemplified by glycolysis in trypanosomes) are not regulated by gene expression alone.

## General applications

- Assessing gene function and relationships to phenotypes
- **Understanding metabolism and predicting novel pathways**
- To increase metabolite fluxes into valuable biochemical pathways using metabolic engineering
- **To compare genetically modified organisms**
- To assess the effect of environmental/stress/temperature changes that lead to changes in gene expression, flux pathways, and extent of carbon and electron flow through them

## Who believes in metabolomics?

### NIH Roadmap

<http://nihroadmap.nih.gov/initiatives.asp>

### Building Blocks, Pathways, and Networks Implementation Group

**Metabolomics Technology Development.** This initiative will promote development of novel technologies to study cellular metabolites, such as lipids, carbohydrates, and amino acids. Knowledge gained from these studies will be used to understand more precisely the role of metabolites in the context of cellular pathways and networks.

### RFA for "Metabolomics Technology Development"

<http://grants.nih.gov/grants/guide/rfa-files/RFA-RM-04-002.html>

## Characteristics of the metabolomes

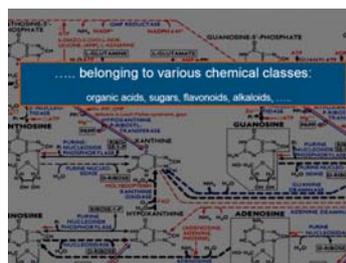
### Metabolome size:

- *S. cerevisiae*: about 600 metabolites
- Plants: estimated 200,000 primary and secondary metabolites
- Mammals: ?

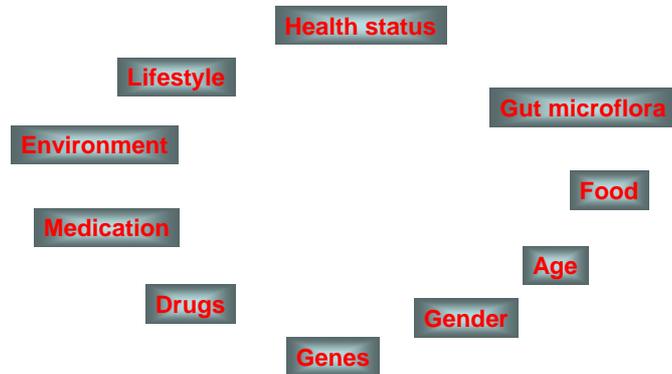


### Metabolite chemical diversity:

- the metabolome extends over an estimated 7–9 magnitudes of concentration (pmol–mmol)
- wide variations in chemical (molecular weight, polarity, solubility) and physical (volatility) properties



## Factors affecting the human metabolome



## Dietary contributions to the human metabolome

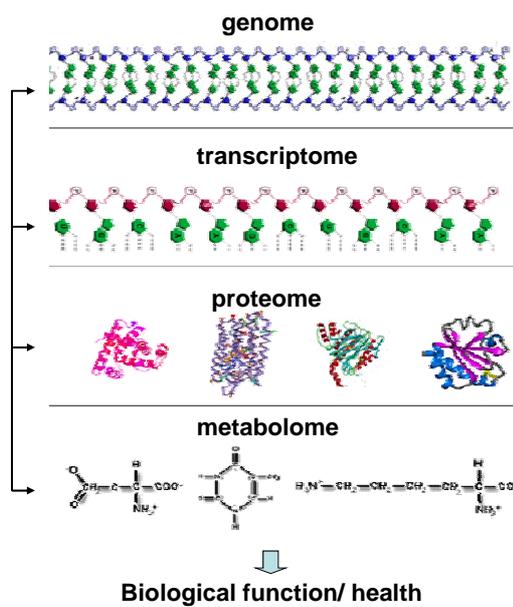
Macronutrient energy Sources

Essential micronutrients

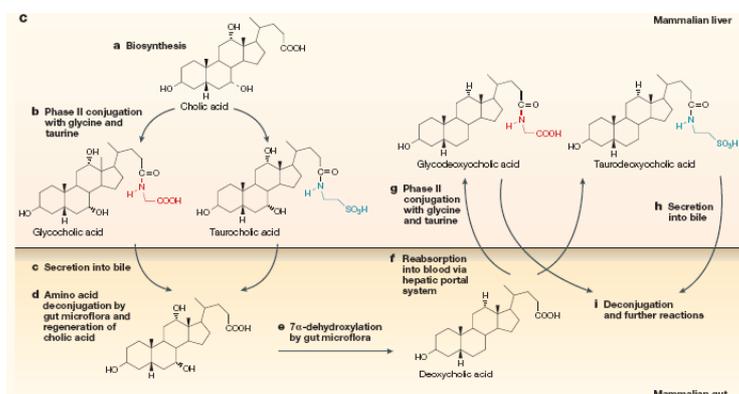
Non-essential, beneficial dietary components

Metabolically neutral dietary components

Dietary toxins and toxicants



## interacting metabolomes



Example of sym-xenobiotic metabolism occurring in mammals

Cholic acid and other bile acids biosynthesized in the liver undergo a series of conversions in both the host liver and inside the gut microflora; these compartments are connected by the entero-hepatic circulation

## Classification of metabolomics approaches

**Metabolomics is the study of metabolic changes. It encompasses metabolomics, metabolite target analysis, metabolite profiling, metabolic fingerprinting, metabolic profiling, and metabonomics – the *Metabolomics Society***

**Metabolite target analysis:** analysis restricted to metabolites of, for example, a particular enzyme that would be directly affected by abiotic or biotic perturbation

**Metabolite profiling:** analysis focused on a group of metabolites, for example, a class of compounds such as carbohydrates, amino acids or those associated with a specific pathway

**Metabolomics:** comprehensive analysis of the whole metabolome under a given set of conditions

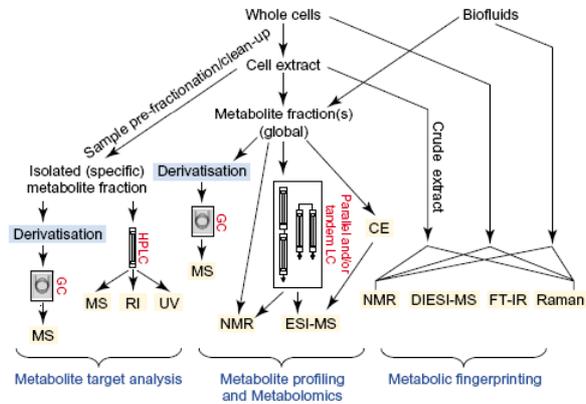
**Metabolic fingerprinting:** classification of samples on the basis of provenance of either their biological relevance or origin

**Metabolic profiling:** often used interchangeably with 'metabolite profiling'; m.p. is commonly used in clinical and pharmaceutical analysis to trace the fate of a drug or metabolite

**Metabonomics:** measure the fingerprint of biochemical perturbations caused by disease, drugs and toxins



## Technologies for metabolome analysis



General strategies for metabolome analysis.

CE, capillary electrophoresis; DIESI, direct-infusion ESI, which can be linked to Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS); NMR, nuclear magnetic resonance; RI, refractive index detection; UV, ultraviolet detection

## Differenti tecniche separative

Basate su...

Solubilità



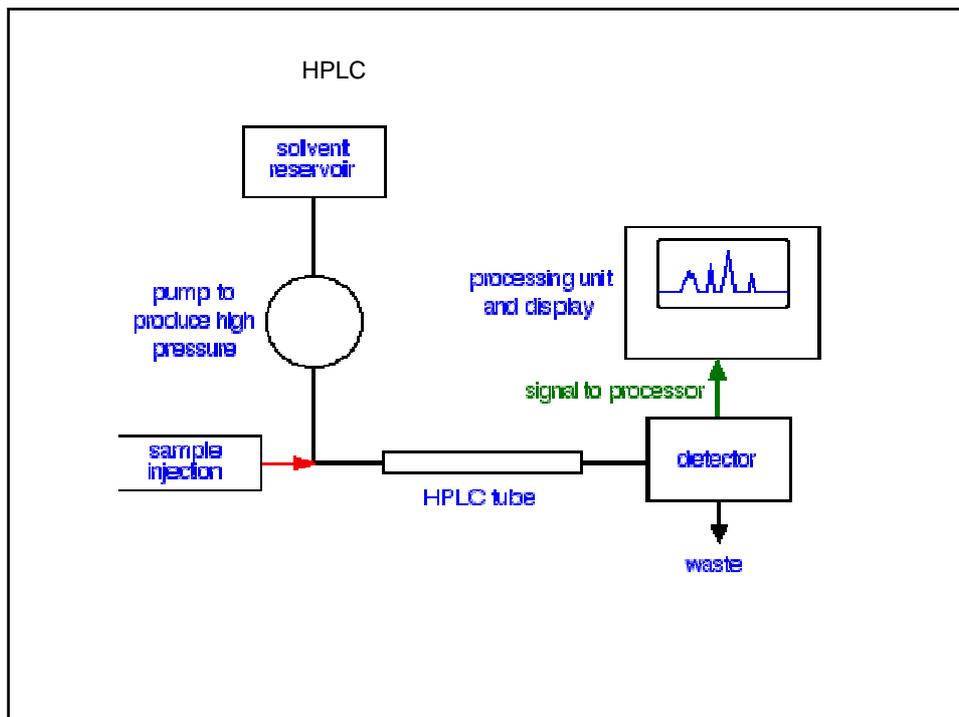
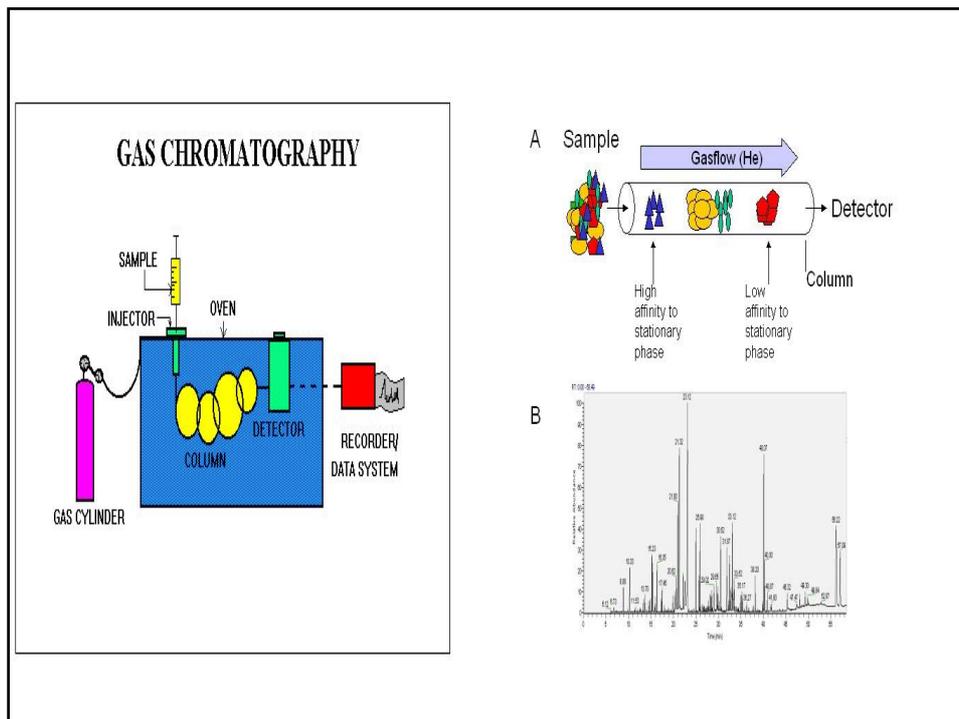
Liquid chromatography (HPLC)

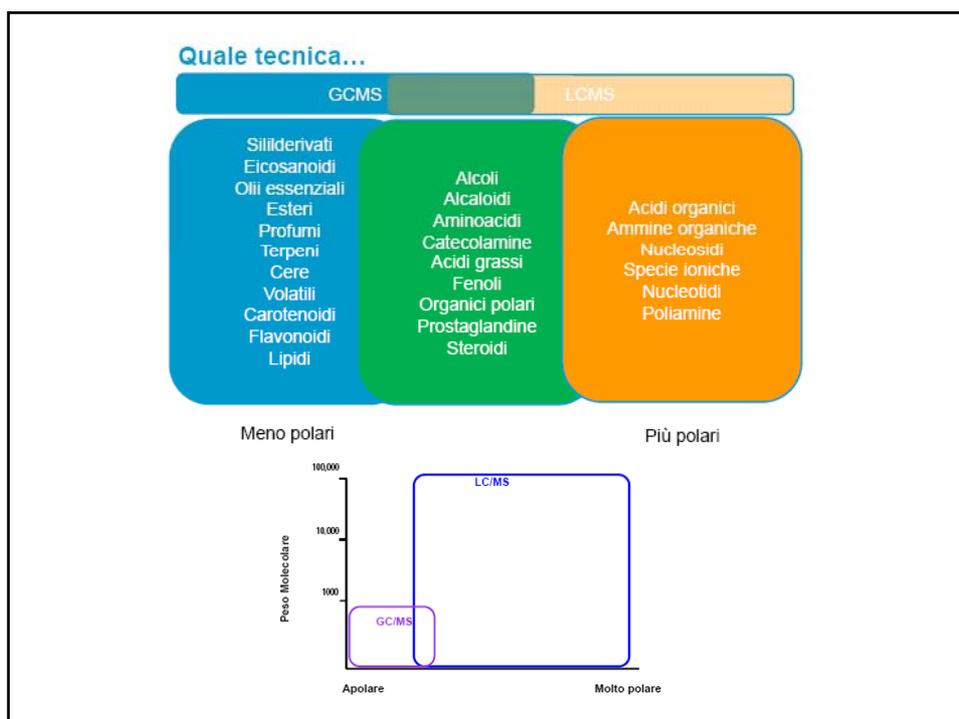
Polarità

Capillary electrophoresis (CE)

Volatilità

Gas chromatography (GC)





## Analisi dei campioni – GC/MS o LC/MS

### GC/MS

#### Vantaggi

- Separazione ad alta efficienza
- Alta sensibilità
- Librerie EI
- No soppressione ionica
- Economico

#### Svantaggi

- Volatilità analiti
  - Derivatizzazione
- No ione molecolare



## Analisi dei campioni – GC/MS o LC/MS

### LC/MS

#### Vantaggi

- Più ampio spettro di analiti
  - Non richiede derivatizzazione
- Alta sensibilità
- Ione pseudomolecolare

#### Svantaggi

- Risoluzione media
- Soppressione ionica
  - ESI o APCI
- Non ci sono librerie di spettri
- Investimento maggiore



## Gas chromatography-mass spectrometry (GCMS)

Ionization of the molecules in GCMS can be done in different ways:

- Electron ionization (positive and negative)
- Chemical ionization (positive and negative)

The fragment ions are detected by time-of-flight (TOF) or by quadrupole mass spectrometry (MS). Often derivatisation methods are used to make metabolites more volatile.

GCMS mainly separates metabolites that are smaller than 500 Dalton. Separation is based on boiling point and binding to the column. Many metabolites can be identified in the GCMS, such as sugars, fatty acids, organic acids and amino acids. GC-MS is poor for the analysis of substances, which are non-volatile due to their high molecular weight and/or polarity. GCMS is suitable as a broad metabolic profiling technique.

## Liquid chromatography-mass spectrometry (LCMS)

Different types of LCMS approaches:

- lipid LCMS
- ion pair LCMS
- polar (derivatised) LCMS

LCMS is the better choice for (semi) polar and non-volatile compounds. It can also be applied to profiling of polar compounds, but special (ion pair) agents need to be used or derivatisation in order to retain polar compounds on the column. In LCMS, more combinations of LC (Normal Phase, Reversed Phase, Ion Pair, HILIC,..) and MS (TOF, Ion trap, Quadrupole, FTMS instruments...) parts are available for different applications. However, the identification of metabolites is more difficult than with GCMS. Often derivatisation methods are used to make metabolites better solvable. LCMS polar is a suitable technique when you would like to apply a fingerprinting procedure specifically on polar compounds. Lipid LCMS techniques are suitable as metabolic profiling techniques when you are specifically interested in lipid metabolism.

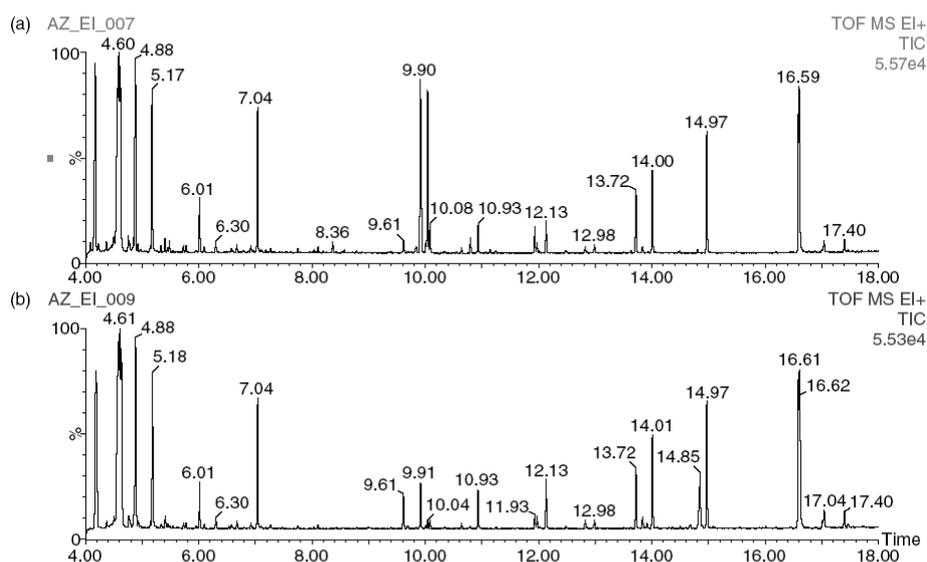


Figure 5.1. Typical TIC traces obtained from GC-EI-MS analysis of plasma obtained from (a) Wistar-derived and (b) Zucker (fa/fa) obese rats.

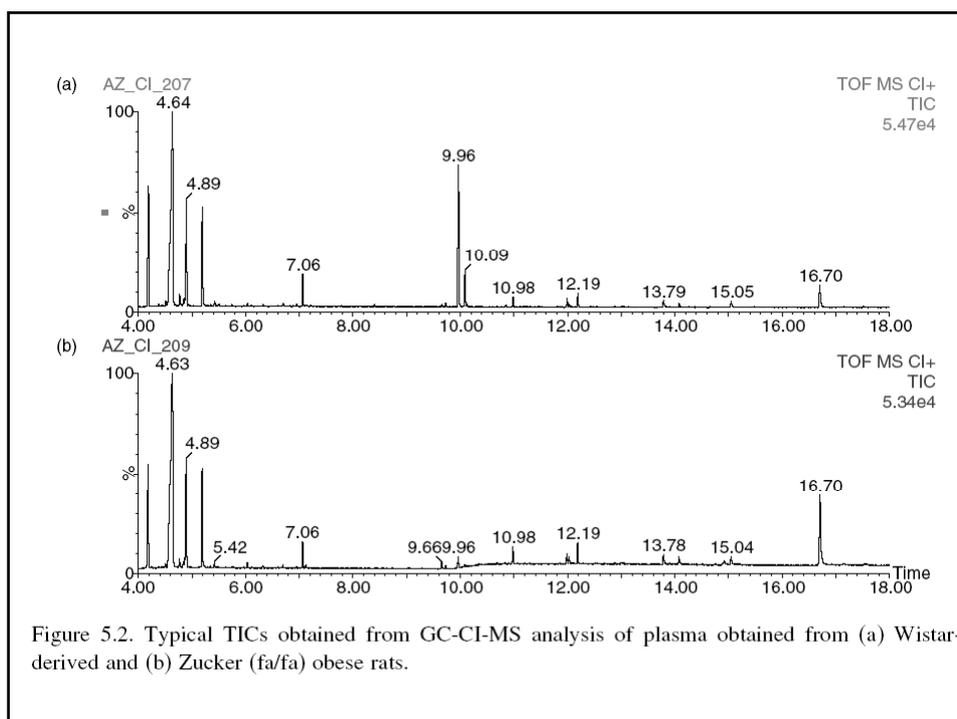


Figure 5.2. Typical TICs obtained from GC-MS analysis of plasma obtained from (a) Wistar-derived and (b) Zucker (fa/fa) obese rats.

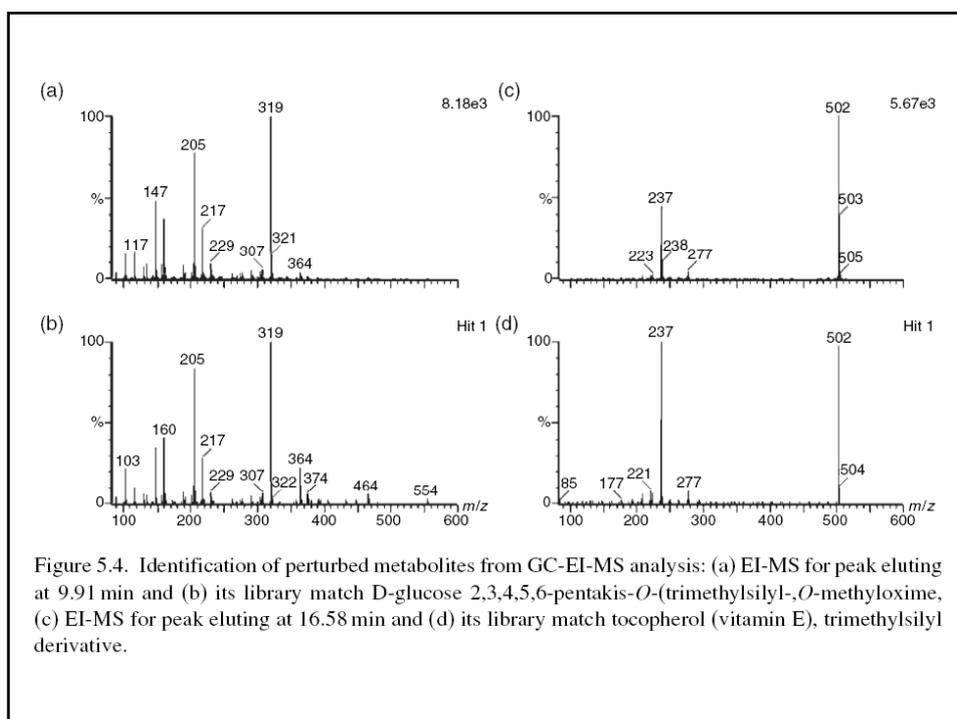
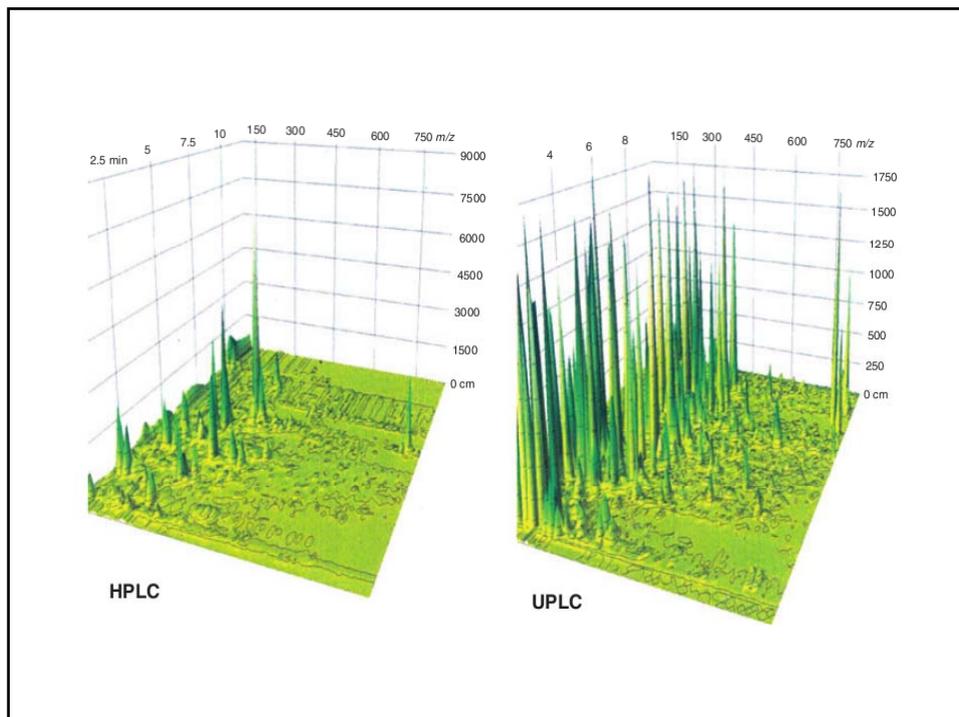
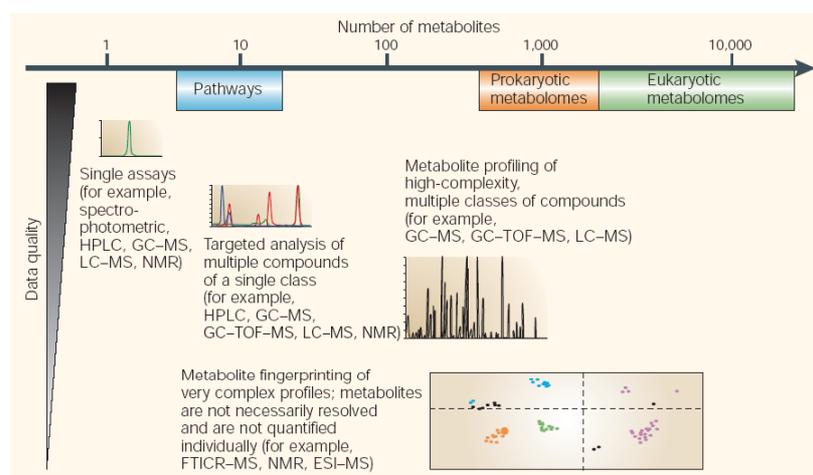


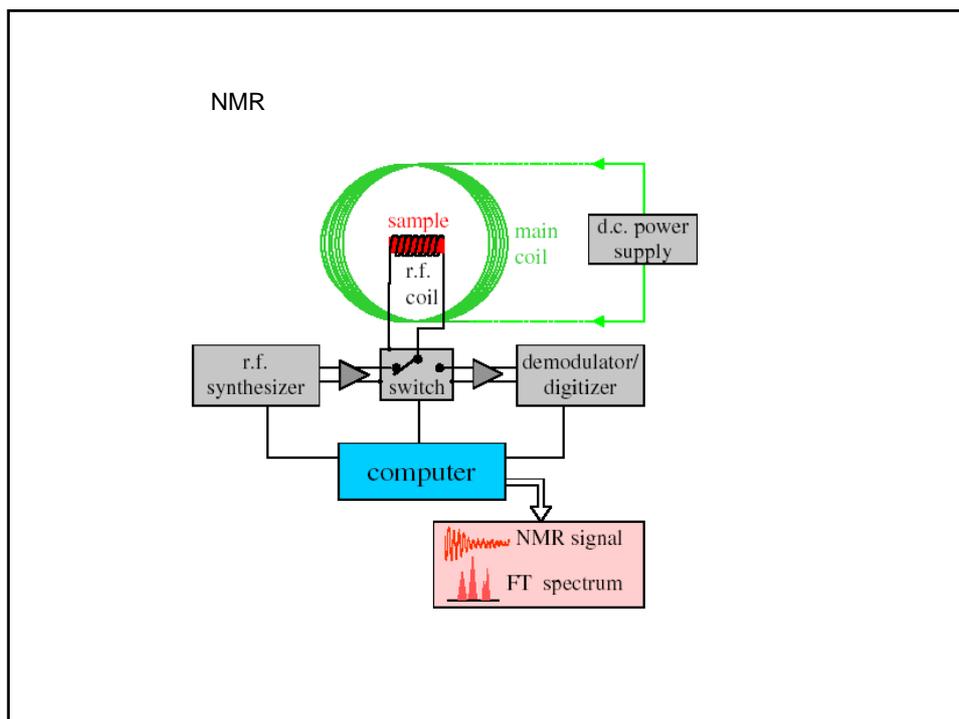
Figure 5.4. Identification of perturbed metabolites from GC-EI-MS analysis: (a) EI-MS for peak eluting at 9.91 min and (b) its library match D-glucose 2,3,4,5,6-pentakis-*O*-(trimethylsilyl)-*O*-methylxime, (c) EI-MS for peak eluting at 16.58 min and (d) its library match tocopherol (vitamin E), trimethylsilyl derivative.



## Quality vs metabolic coverage



The trade-off between metabolic coverage and the quality of metabolic analysis



## NMR-based metabolomics

### What kind of samples can be analysed by NMR?

- All type of biological liquids (urine, plasma, cerebrospinal fluid, amniotic fluid, sperm, synovial fluid, saliva) or cellular or organ extracts
- All kind of biological samples such as biopsies of organs and cell cultures

### Why is NMR competitive?

- NMR offers a direct biochemical window into a living system in a holistic way (no a priori selection)
- NMR is fully quantitative
- There is no need for special sample preparation (fractionation, derivatization, ...)
- NMR is non-destructive and allows to completely recover the samples
- NMR has emerged into a high throughput analysis system with minimal sample preparation (cost effective)
- Nearly all metabolic intermediates have unique NMR signatures

### Factors underlying the enormous challenges of metabolomics

Radius of a typical eukaryotic cell (meter)	$5 \times 10^{-6}$
Volume of one cell (liter)	$5 \times 10^{-13}$
Maximum number of cells in 10 g tissue	$2 \times 10^{10}$
Maximum quantity (mole) of a metabolite with one copy/cell recoverable from 10 g cells/tissue	$3 \times 10^{-15}$
Detection limit for MS (mole)	$1 \times 10^{-18}$
Dynamic range limit for MS (factor)	$1 \times 10^6$
Detection limit for $^1\text{H}$ NMR (mole)	$1 \times 10^{-9}$
Dynamic range limit for NMR (factor)	$1 \times 10^6$

### NMR vs MS

**NMR and MS dominate metabolomics research. "There is no one magic tool that can capture the diversity of composition and concentration which is present in a single sample," says Aram Adourian, senior director of advanced technologies at Beyond Genomics. "It really depends on the question you are asking. You need to have an array of tools available."**

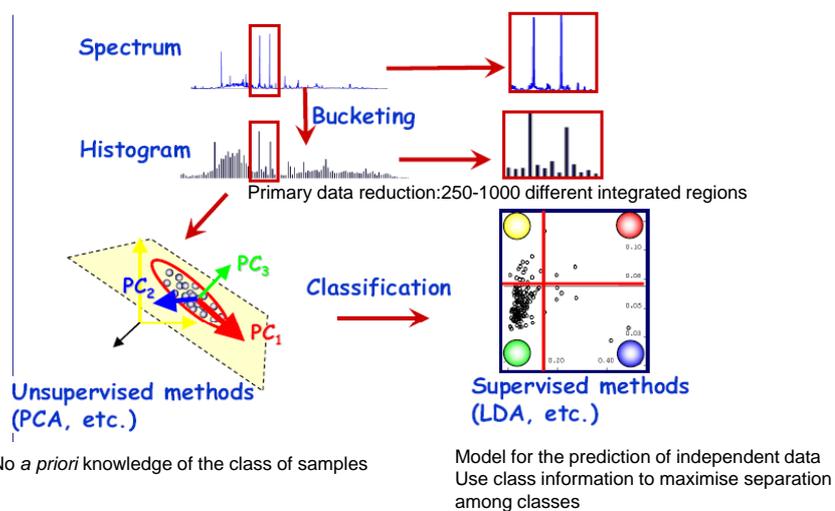
[www.the-scientist.com](http://www.the-scientist.com)  
 Amy Adams, "Metabolomics",  
 Volume 17 | Issue 8 | 38 | Apr. 21, 2003

## Applications and examples of NMR-based metabonomics

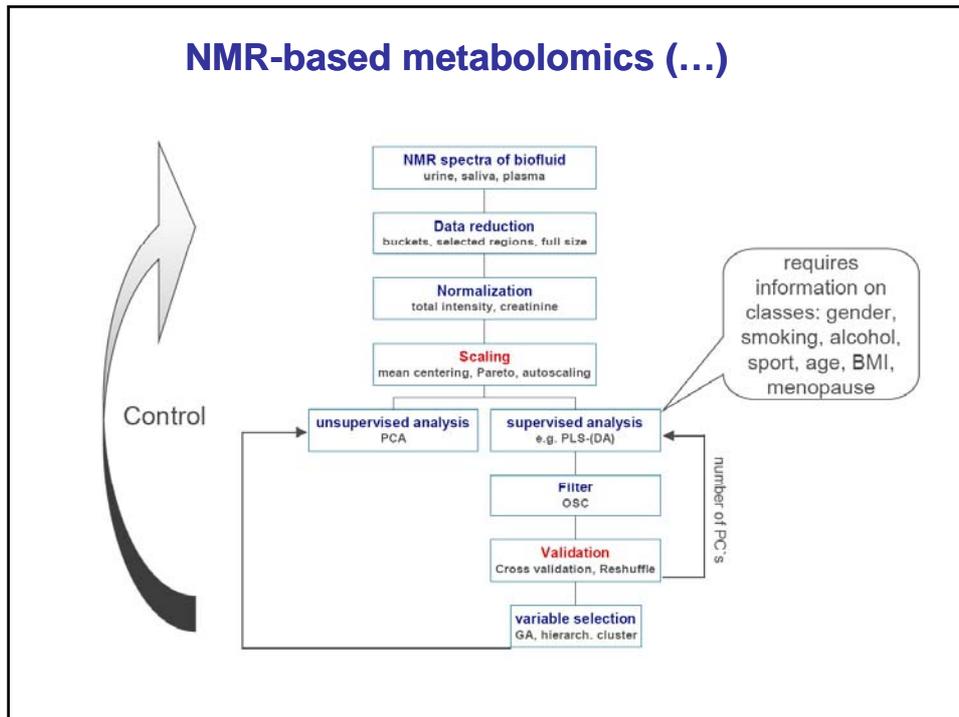
APPLICATION	EXAMPLES
Classification of toxicity	Nephrotoxicity Hepatotoxicity Phospholipidosis Testicular toxicity Mitochondrial
Classification of disease	Inborn errors of metabolism Cancer (prostatic, brain, renal etc) Renal disease Diabetes Muscular Dystrophy
Investigation of physiological status	Diurnal variation Hormonal variation
Monitoring efficacy of therapeutic intervention	Renal transplantation (cyclosporin)
Functional genomics	Assessment of strain differences in animal models Evaluation of transgenic models
Characterisation of natural products	Batch to batch variation in commercial Feverfew

Antti et al.  
<http://www.acc.umu.se/~tnkjtg/Chemometrics/Editorial>

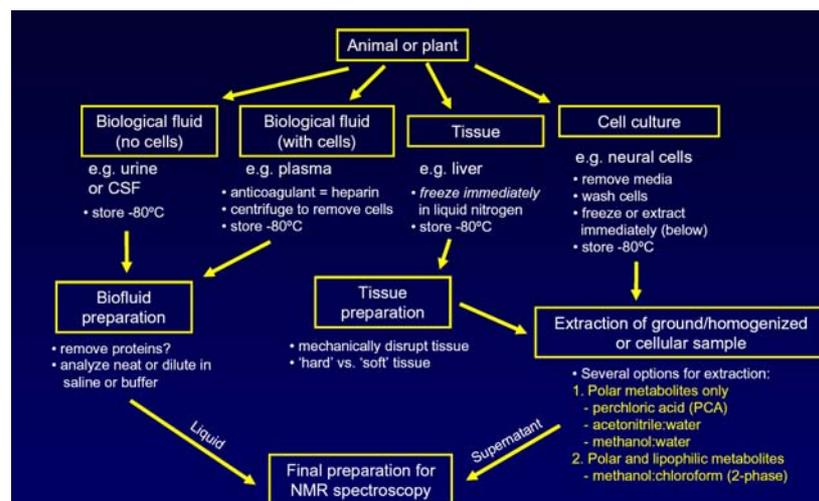
## NMR-based metabolomics: the concept

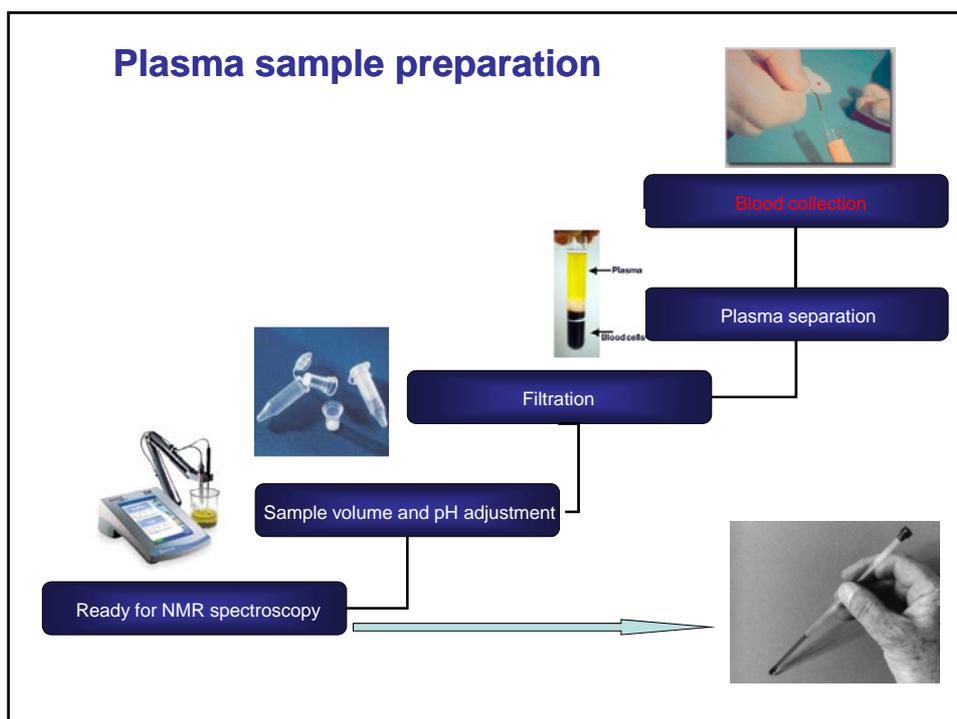


## NMR-based metabolomics (...)



## Sample preparation for NMR-based metabolomics





### Human Metabolome Database

Location: <http://www.hmdb.ca/>

Description: The Human Metabolome Database (HMDB) is a freely available electronic database containing detailed information about small molecule metabolites found in the human body. It is intended to be used for applications in metabolomics, clinical chemistry, biomarker discovery and general education. The database is designed to contain or link three kinds of data: 1) chemical data, 2) clinical data, and 3) molecular biology/biochemistry data. The database currently contains nearly 2500 metabolite entries including both water-soluble and lipid soluble metabolites as well as metabolites that would be regarded as either abundant ( $> 1 \mu\text{M}$ ) or relatively rare ( $< 1 \text{ nM}$ ). Additionally, approximately 5500 protein (and DNA) sequences are linked to these metabolite entries. Each MetaboCard entry contains more than 90 data fields with half of the information being devoted to chemical/clinical data and the other half devoted to enzymatic or biochemical data. Many data fields are hyperlinked to other databases (KEGG, PubChem, MetaCyc, ChEBI, PDB, Swiss-Prot, and GenBank) and a variety of structure and pathway viewing applets. The HMDB database supports extensive text, sequence, chemical structure and relational query searches.

### ChemSpider

Location: <http://www.chemspider.com/>

Description: ChemSpider is a free access service providing a structure centric community for chemists. Providing access to millions of chemical structures and integration to a multitude of other online services ChemSpider is the richest single source of structure-based chemistry information.

**KEGG: Kyoto Encyclopedia of Genes and Genomes**

Location: <http://www.genome.jp/kegg/>

Description: The goal of this website is to build a bioinformatics resource as complete computer representation of the cell, the organism, and the biosphere, which will enable computational prediction of higher-level complexity of cellular processes and organism behaviors from genomic and molecular information. For metabolites especially <http://www.genome.ad.jp/kegg/liqand.html>, which includes possibilities to search for chemical formula, name, exact mass, and pathway.

**METLIN Metabolite Database**

Location: <http://metlin.scripps.edu/>

Description: METLIN is a metabolite database for metabolomics containing over 15,000 structures, it is also represents a data management system designed to assist in a broad array of metabolite research and metabolite identification by providing public access to its repository of current and comprehensive mass spectral metabolite data.

**ChEBI**

Location: <http://www.ebi.ac.uk/chebi/>

Description: Chemical Entities of Biological Interest is a freely available dictionary of molecular entities focused on 'small' chemical compounds. The term 'molecular entity' refers to any constitutionally or isotopically distinct atom, molecule, ion, ion pair, radical, radical ion, complex, conformer, etc., identifiable as a separately distinguishable entity. The molecular entities in question are either products of nature or synthetic products used to intervene in the processes of living organisms.

**PubChem**

Location: <http://pubchem.ncbi.nlm.nih.gov/>

Description: PubChem provides information on the biological activities of small molecules.

Including:

PubChem Compound: Search unique chemical structures using names, synonyms or keywords. Links to available biological property information are provided for each compound.

PubChem Substance: Search deposited chemical substance records using names, synonyms or keywords. Links to biological property information and depositor web sites are provided.

PubChem BioAssay: Search bioassay records using terms from the bioassay description, for example "cancer cell line". Links to active compounds and bioassay results are provided.

Structure Search: Search PubChem's Compound database using a chemical structure as the query. Structures may be sketched or specified by SMILES, MOL files, or other formats.

**Comparative Toxicogenomics Database (CTD)**

Location: <http://ctd.mdibl.org/>

Description: The Comparative Toxicogenomics Database (CTD) elucidates molecular mechanisms by which environmental chemicals affect human disease.

Chemical-gene/protein interactions and chemical- and gene-disease relationships are curated from the published literature, and integrated with diverse data (chemicals, genes/proteins, human diseases, references, sequences, vertebrate and invertebrate organisms, and the Gene Ontology) to facilitate environmental health research.

**Lipid Maps**

Location: <http://www.lipidmaps.org/>

Description: This website is focussed on the lipid section of the metabolome by developing an integrated metabolomic system capable of characterizing the global changes in lipid metabolites ("lipidomics"). All enzymes and other proteins involved in lipid metabolism will be mapped and integrated into the networks of lipidomics.

**Lipid Library**

Location: <http://www.lipidlibrary.co.uk/>

Description: Includes: definitions, structures, composition, occurrence, biochemistry and functions of most types of fatty acids and lipids.

Comments: This website is regularly updated also with the most recent literature on the analysis of lipids and many practical tips can be found. The site is really a must for lipid scientists.

**PRIMe**

Location: <http://prime.psc.riken.jp/>

Description: RIKEN Metabolomics platform, is a Web-based service for metabolomics and transcriptomics as systems for understanding life. This project measures standard metabolites by means of multi-dimensional NMR spectroscopy, GC/MS, LC/MS, and CE/MS. We also provide unique tools for metabolomics, transcriptomics, and integrated analysis of a range of other "-omics" data.

**MassBank**

Location: [www.massbank.jp](http://www.massbank.jp)

Description: This site presents the database of comprehensive, high-resolution mass spectra of metabolites. Supported by the [JST-BIRD](#) project, it offers various query methods for standard spectra from [Keio Univ.](#), [RIKEN PSC](#), and others.