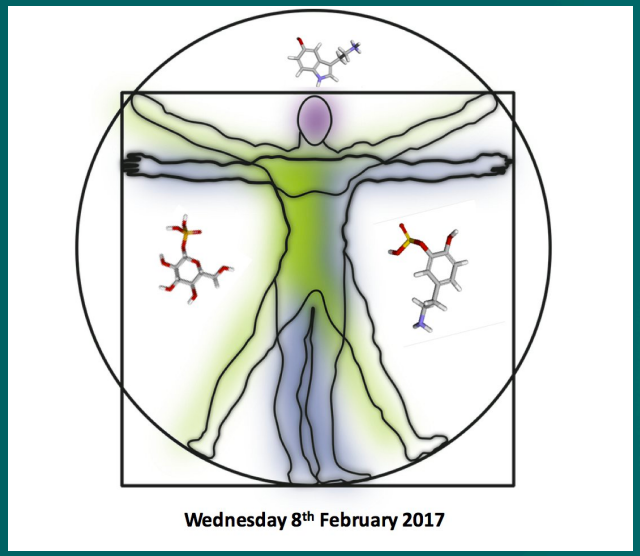
Metabolomics: molecules of life

An introduction



Reza Salek PhD

Metabolism and Molecular Informatics

The European Bioinformatics Institute (EMBL-EBI)

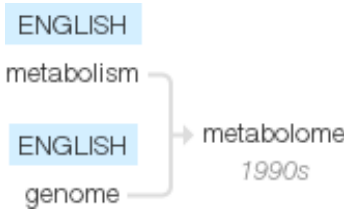
Email: Reza.salek@ebi.ac.uk

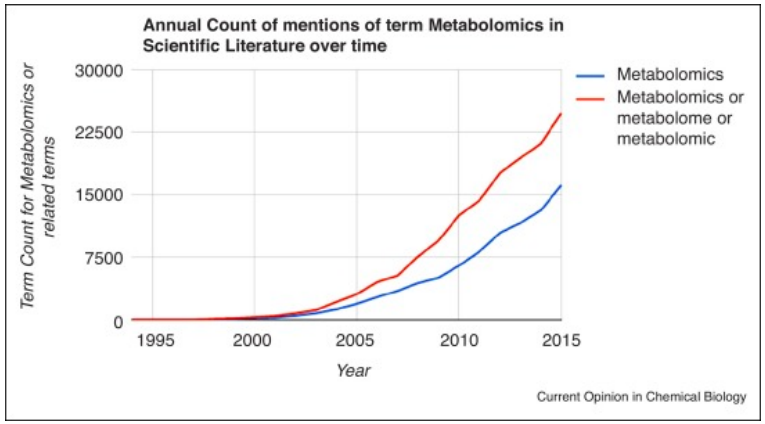
EMBL-EBI

Some Definitions

metabolome

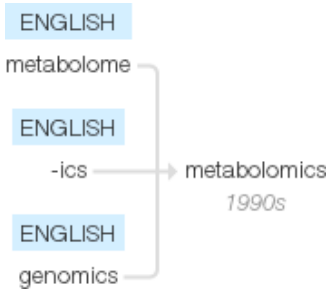
/mɪˈtabələʊm/
noun
BIOCHEMISTRY
noun: **metabolome; plural noun: metabolomes**
the total number of metabolites present within an organism, cell, or tissue.

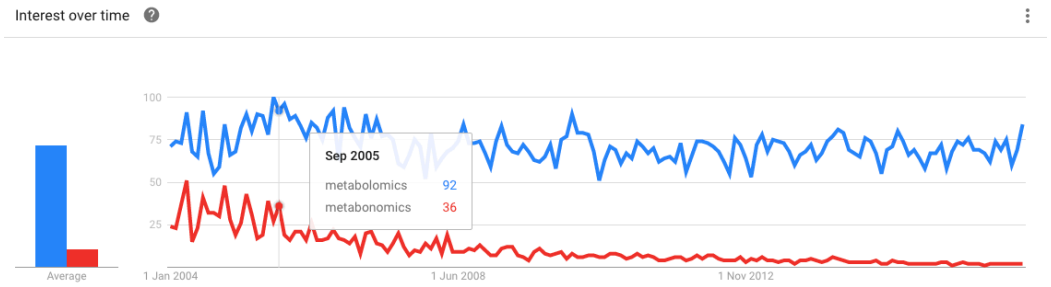




metabolomics

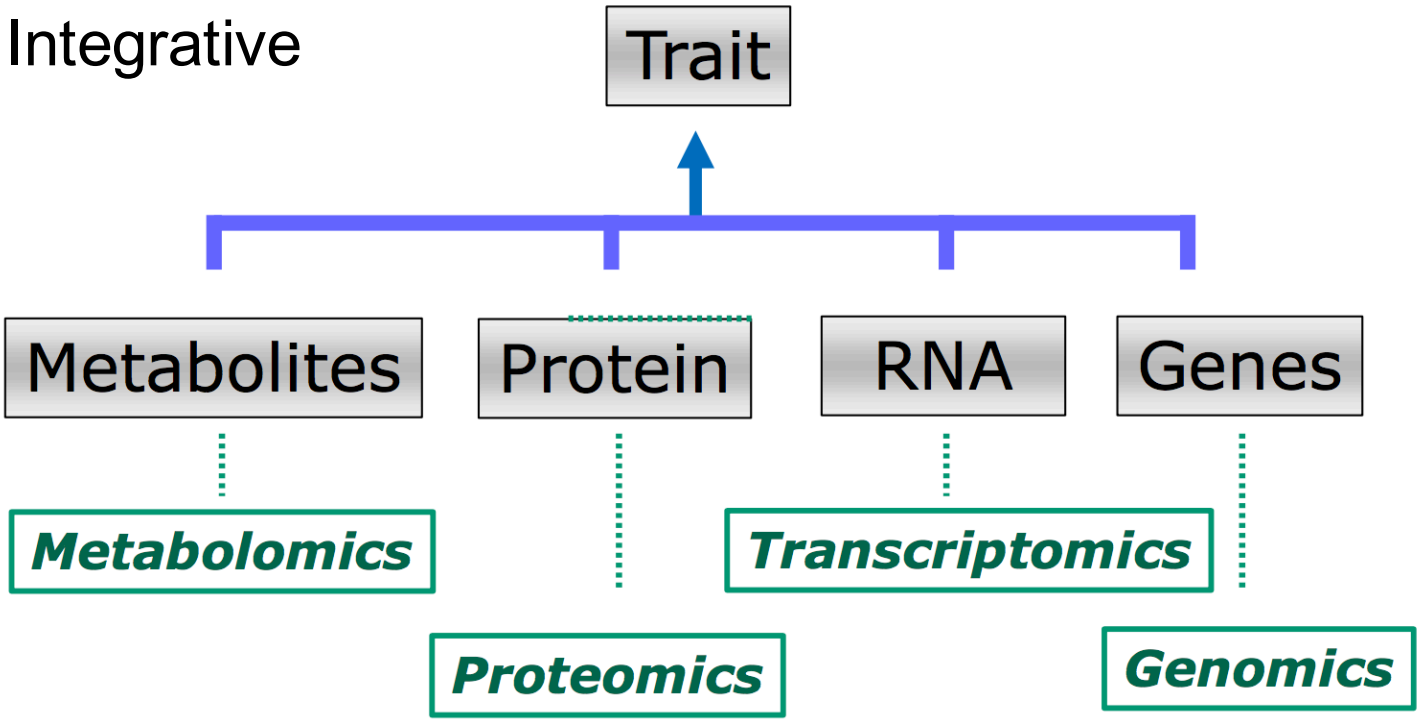
mɪˈtabəlɒmɪks/
Noun
BIOCHEMISTRY
noun: **metabolomics**
the scientific study of the set of metabolites present within an organism, cell, or tissue, often by measuring simultaneously (100s -10000s), many of which are not identified (features or analyts,)



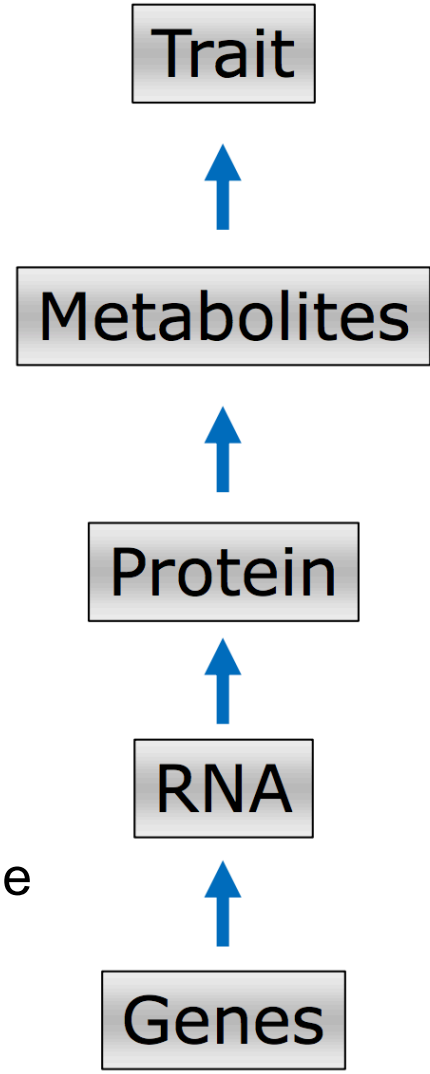


Central dogma in Biology

Integrative



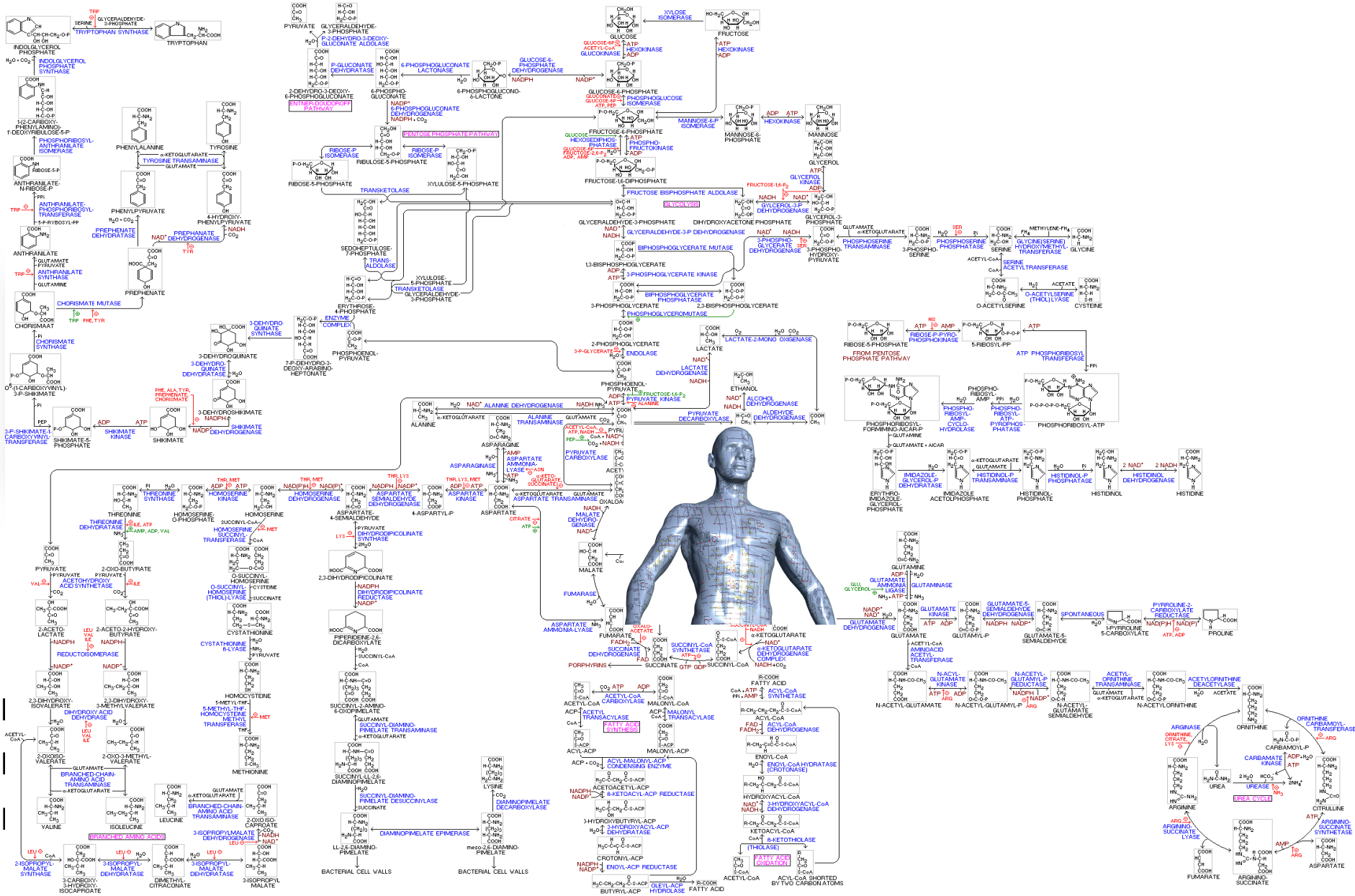
Reductionist



- 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

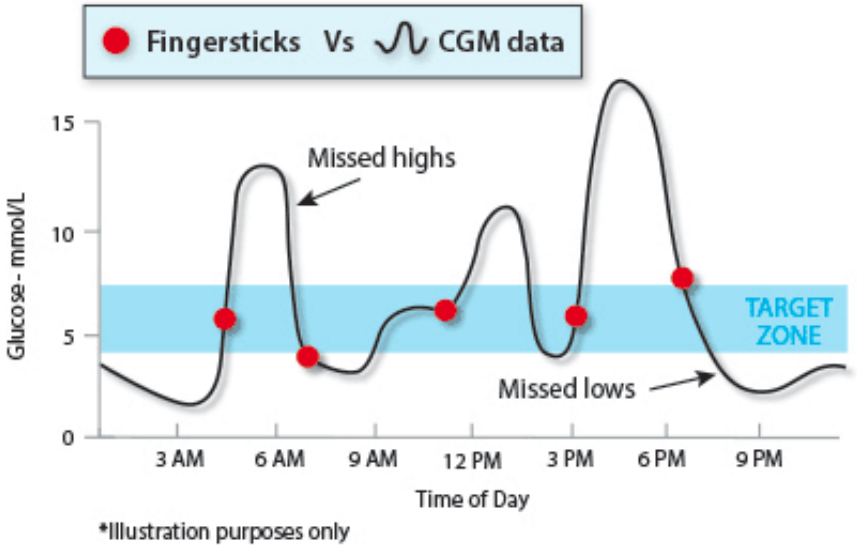


Uroscopy – Early metabolomics

- 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.

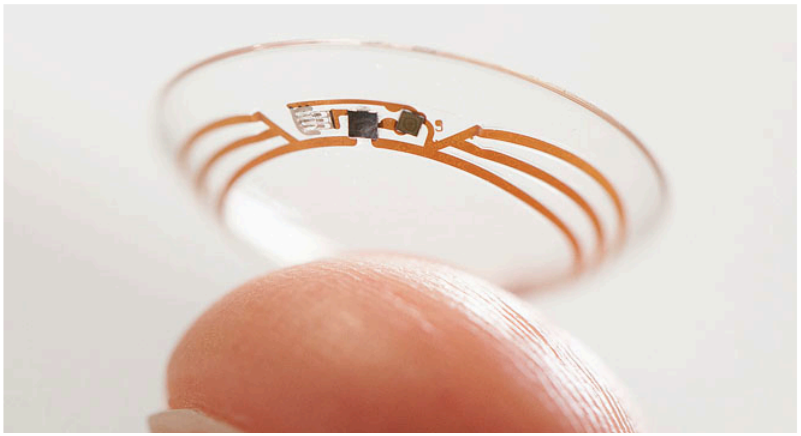
The Urine Wheel for diagnosing metabolic diseases, from *Epiphanie Medicorum* by Ullrich Pinder in 1506

Bio Marker Diagnosis and Monitoring



Google announces 'smart' contact lenses that monitor glucose levels

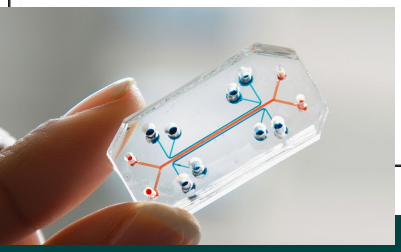
Published January 16, 2014 / FoxNews.com



Images from Google

EMBL-EBI

Metabolomics on a chip



J. H. Sung, Shuler lab, Cornell University.

EMBL-EBI

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

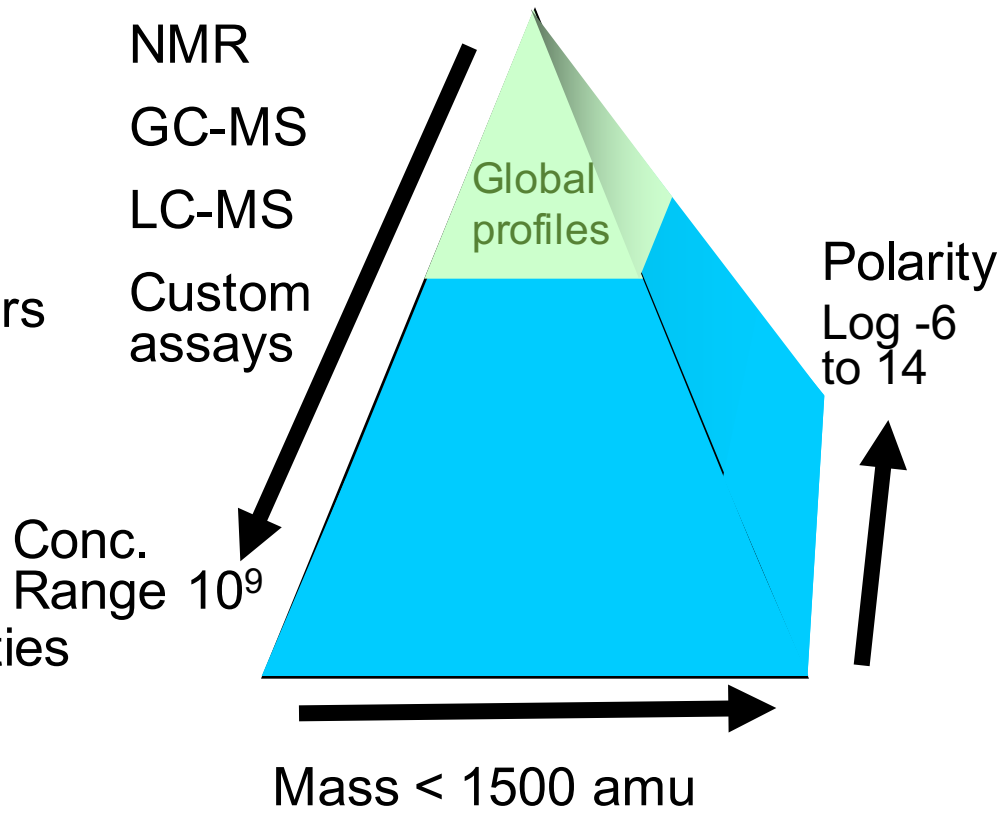
Cell size vs Metabolite – by Lane et al,

- 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 3
- 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

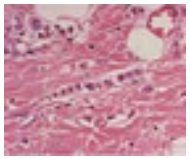
- Concentrations depend on **pool sizes and compartmen**tation: lymphocytes cell nucleus occupies >50% of internal volume, epithelial cells <10%.
- **Mitochondria numbers vary** from 1-1000 each ca. 1 fL, occupying a substantial fraction of the cell volume

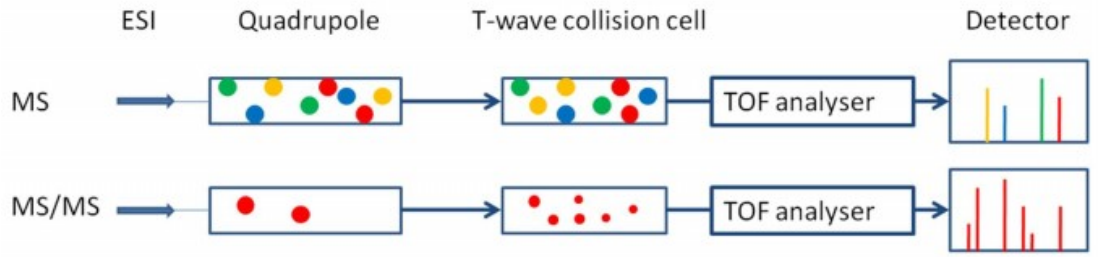
Dynamic range of metabolome

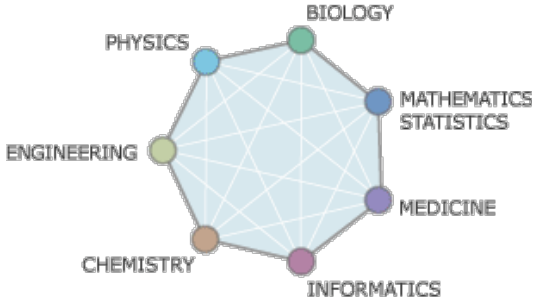
- Size varies “only” 2 orders of magnitude (~150 to 1500 Da)
- Abundance varies hugely (many orders of magnitude)
- Large number of structural isomers
  - i.e Lipids (TAG)
    - 40 common fatty acids
    - 3 fatty acids per TAG → 64000 species
- Hugely different chemical properties
  - Hydrophobicity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



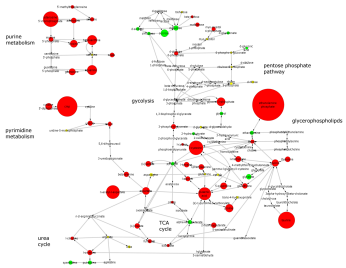
Snapshot metabolomics

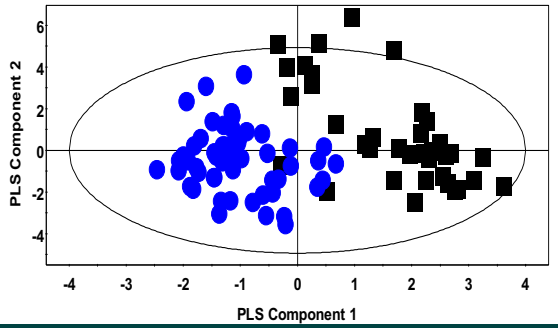


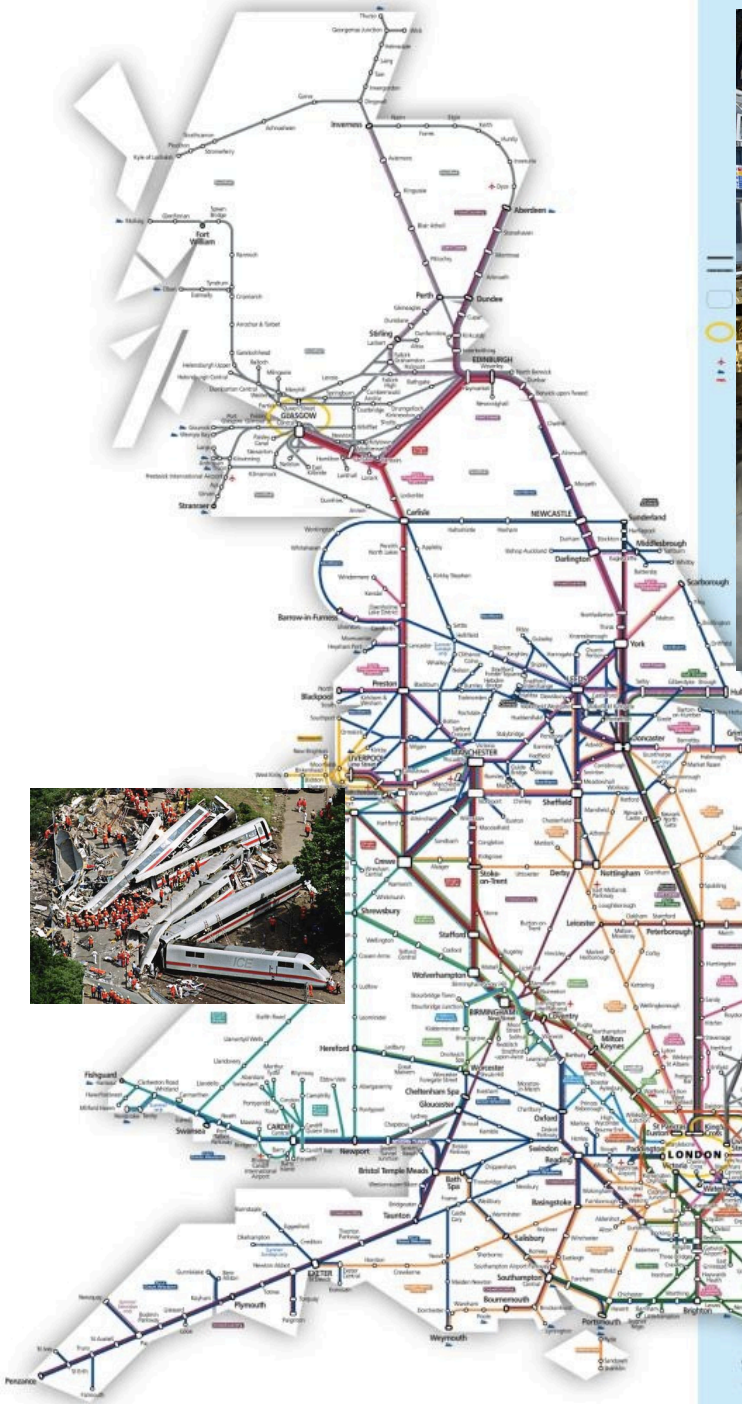




Oltvai-Barabasi, Science, Oct 02



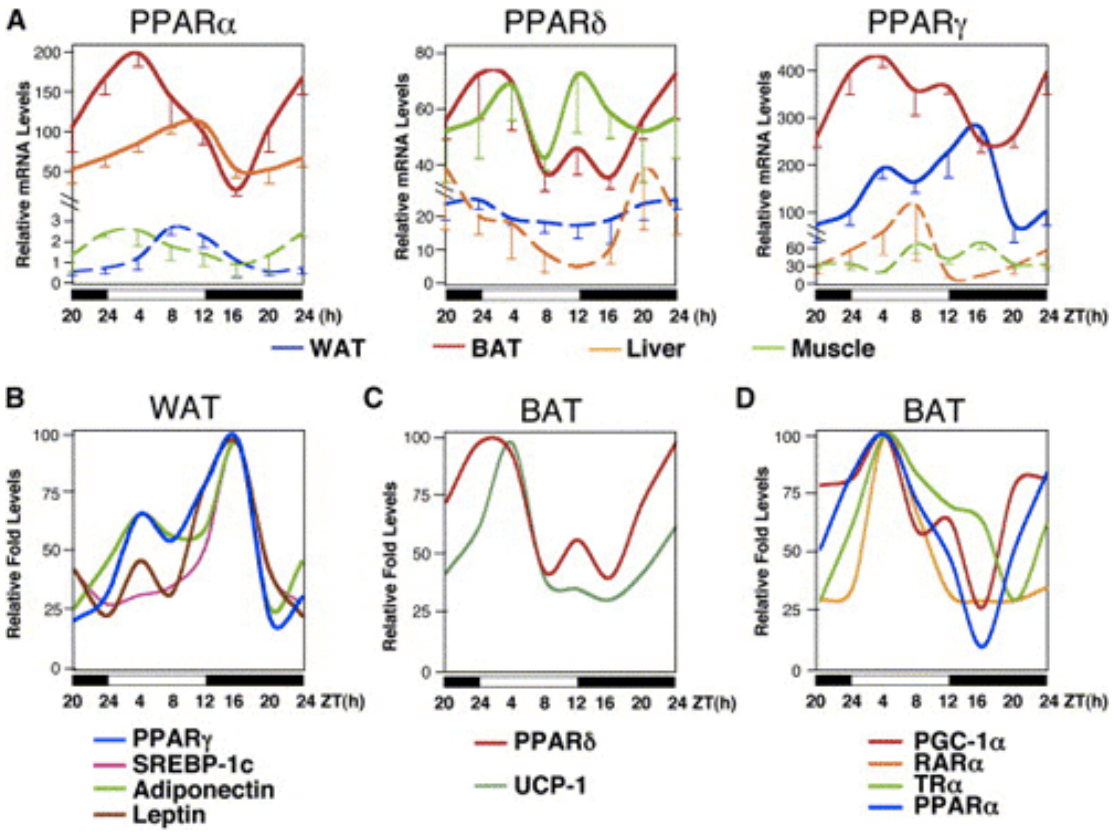



Inverness
Aberdeen
Perth
Dundee
Stirling
EDINBURGH
GLASGOW
Carlisle
NEWCASTLE
Sunderland
Middlesbrough
Darlington
Scarborough
Barrow-in-Furness
York
Preston
Blackpool
Hull
Grimsby
LIVERPOOL
MANCHESTER
Sheffield
Doncaster
Crewe
Stoke-on-Trent
Derby
Nottingham
Leicester
Peterborough
Norwich
Shrewsbury
Stafford
Wolverhampton
BIRMINGHAM
Coventry
Milton Keynes
Ipswich
Cambridge
Colchester
Harwich
Hereford
Worcester
Fishguard
Cheltenham Spa
Gloucester
Oxford
CARDIFF
Swansea
Newport
Swindon
Reading
LONDON
Bristol Temple Meads
Bath Spa
Basingstoke
Taunton
Salisbury
Ashford
Dover
Exeter
Southampton
Bournemouth
Portsmouth
Brighton
Eastbourne
Plymouth
Weymouth
Penzance
Continental Europe
© 2012 Andrew Smithers


Departures
Liverpool Street

EMBL-EBI

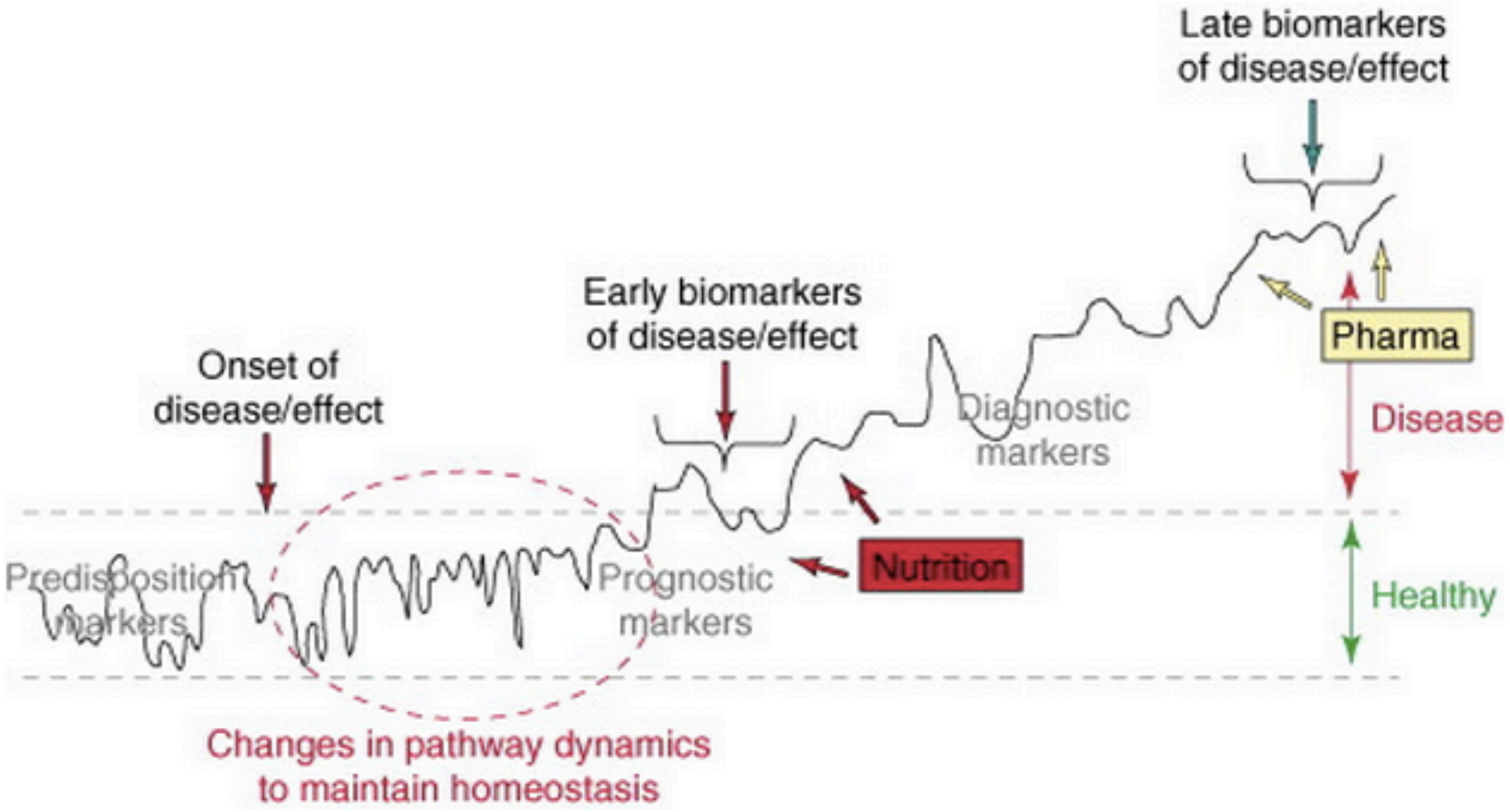
Diurnal Rhythm



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

DOI: 10.1371/journal.pgen.1002143

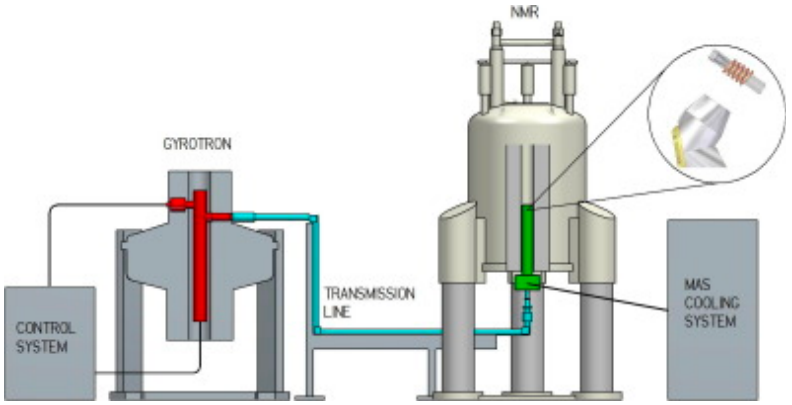
Why is it important

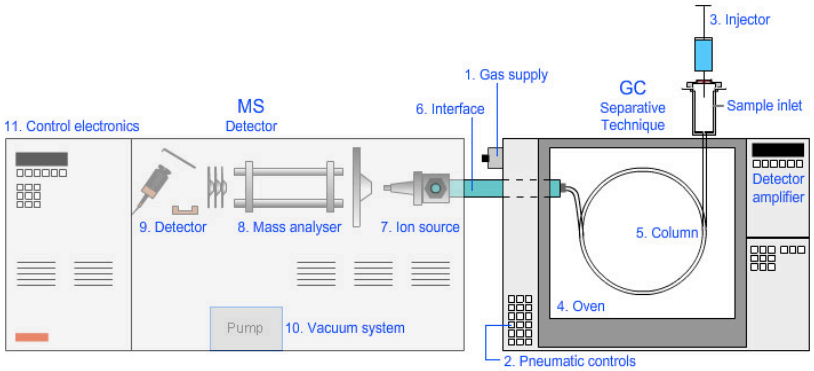


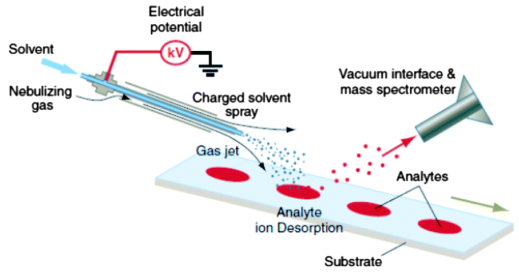
Jan van der Greef, Paul Stroobant, Rob van der Heijden, The role of analytical sciences in medical systems biology, Current Opinion in Chemical Biology, Volume 8, Issue 5, October 2004, Pages 559-565, ISSN 1367-5931, http://dx.doi.org/10.1016/j.cbpa.2004.08.013.

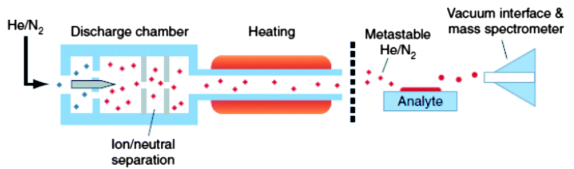
Global profiling tool

- 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

Exploiting High Mass Accuracy to ID Compounds

Type	Mass Accuracy
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)

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

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

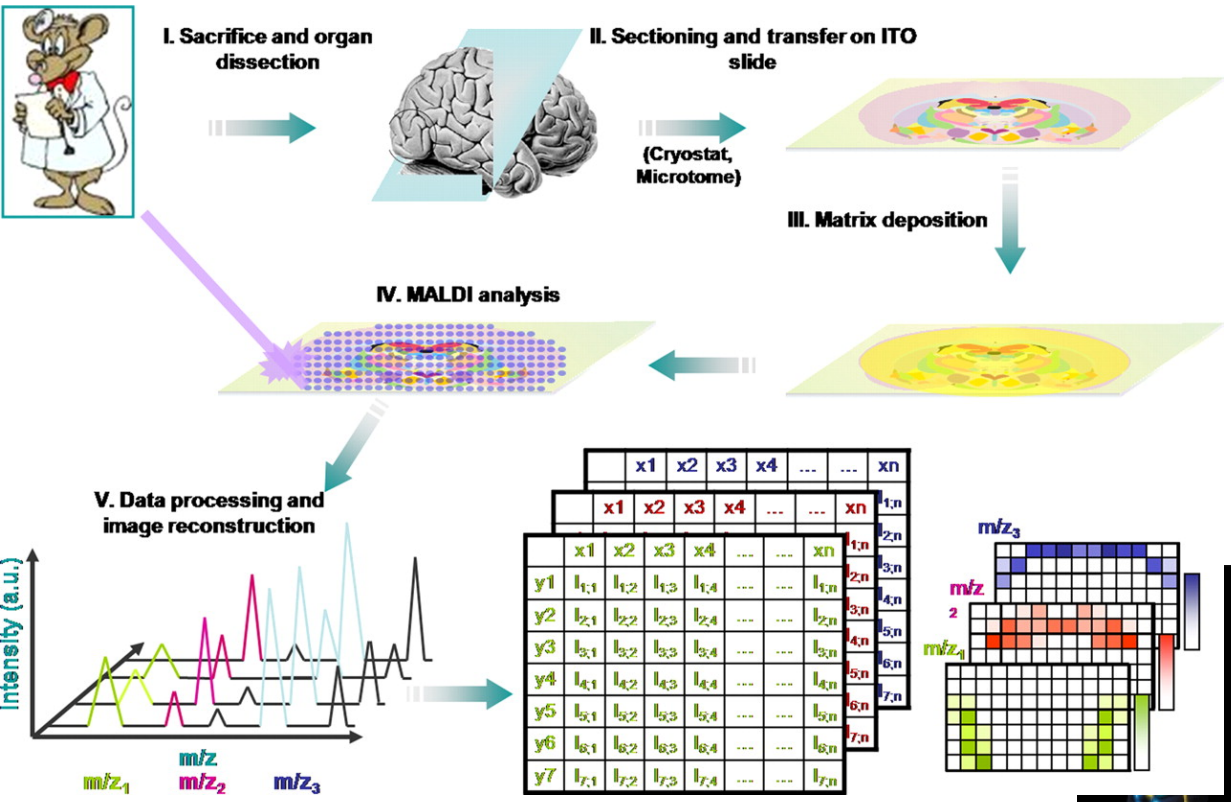
Advantages of MS

- It is much more sensitive than NMR. This mean 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

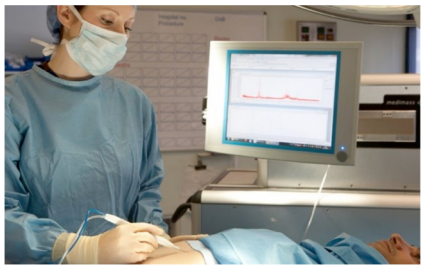
- Similarity of isomers can make identification difficult. For LC-MS libraries and not complete
- Samples must be derivatised before analysis which increases the preparation time.
- Sample analysis time is longer than NMR and are less reproducible
- Some large molecules cannot be measured by this technique.

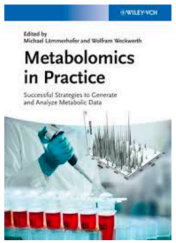
Imaging Mass spectrometry

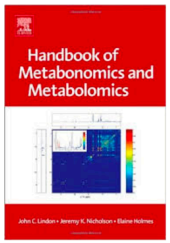


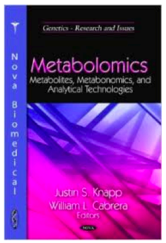
METASPACE EMBL

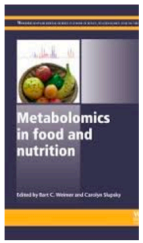
Applications of metabolomics is growing

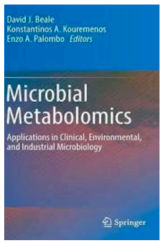


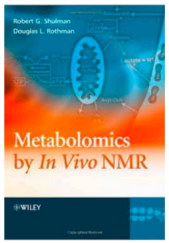


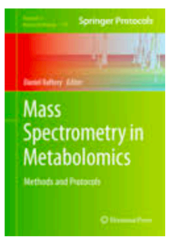


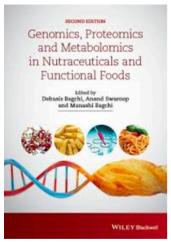


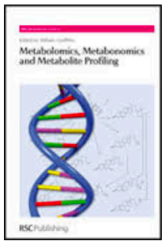


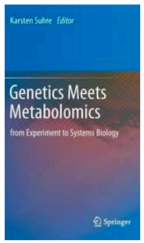


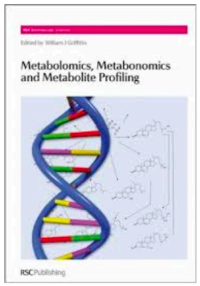


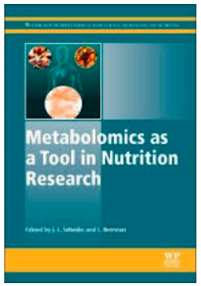


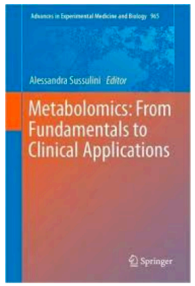


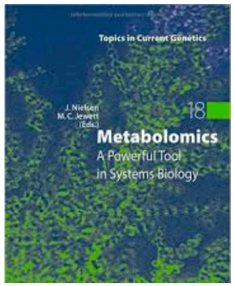


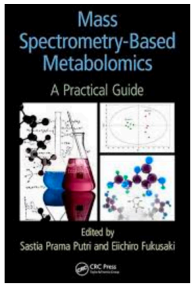


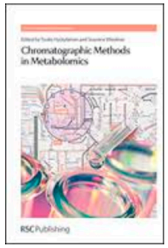


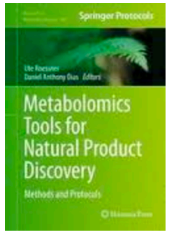


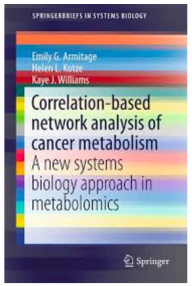


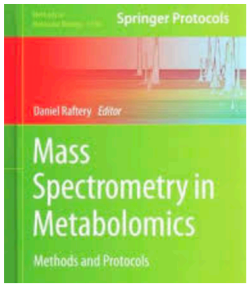


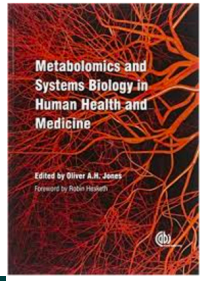


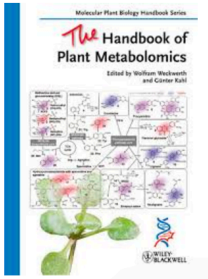


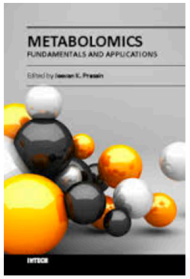


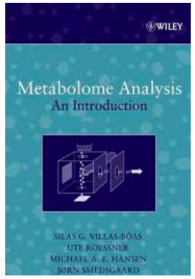


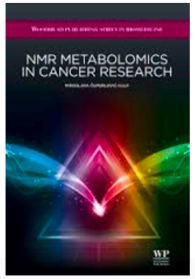


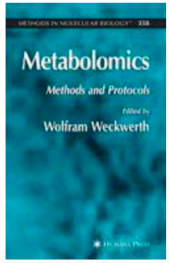


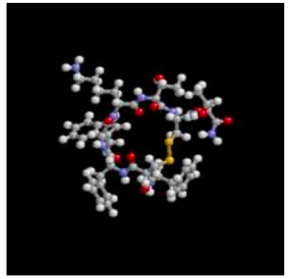






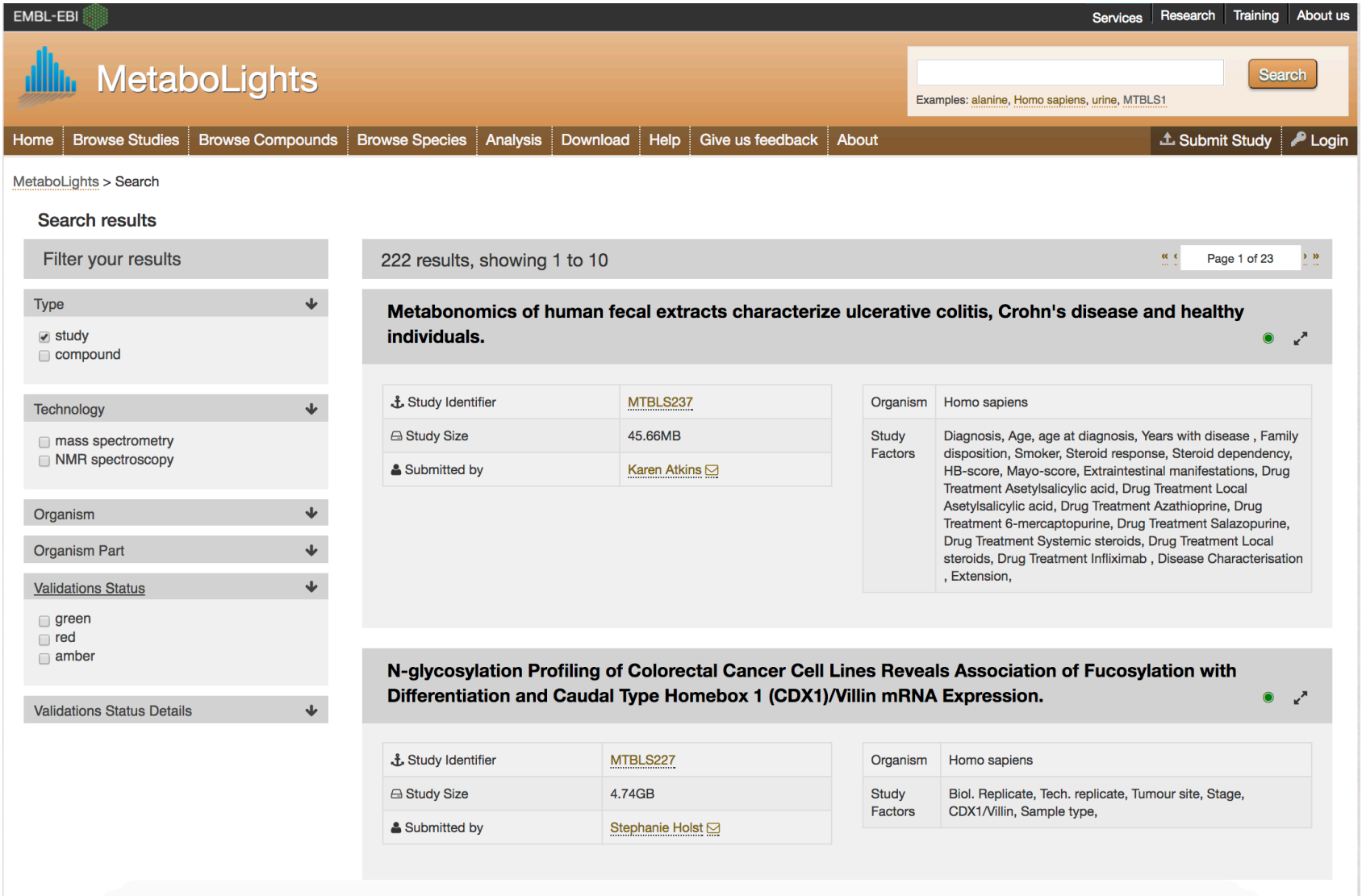






EMBL-EBI

MetaboLights – metabolomics data sharing

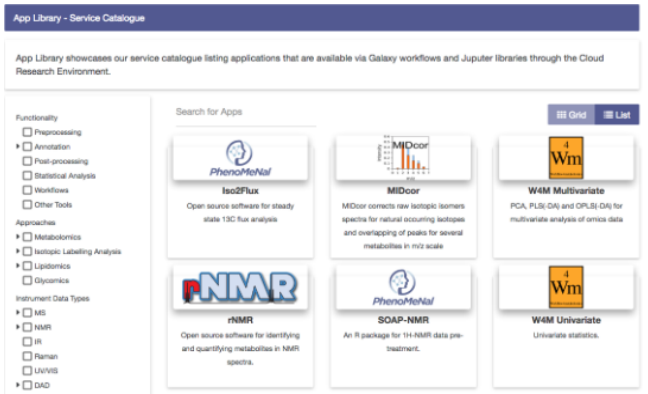


Large scale computing for medical metabolomics

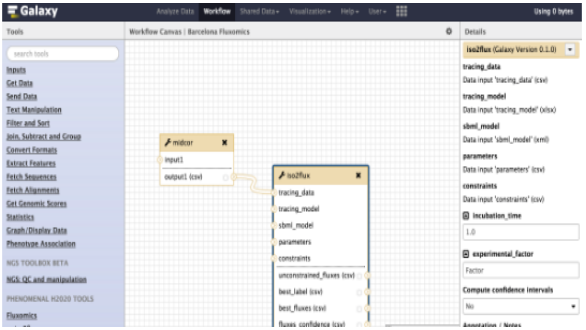
All libraries in the App Library run Within the CRE on the cloud. This simplifies running and maintaining applications

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

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



The open-source App Library



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

PhenoMeNal
Google Cloud Platform
egi
openstack
amazon web services

EMBL-EBI

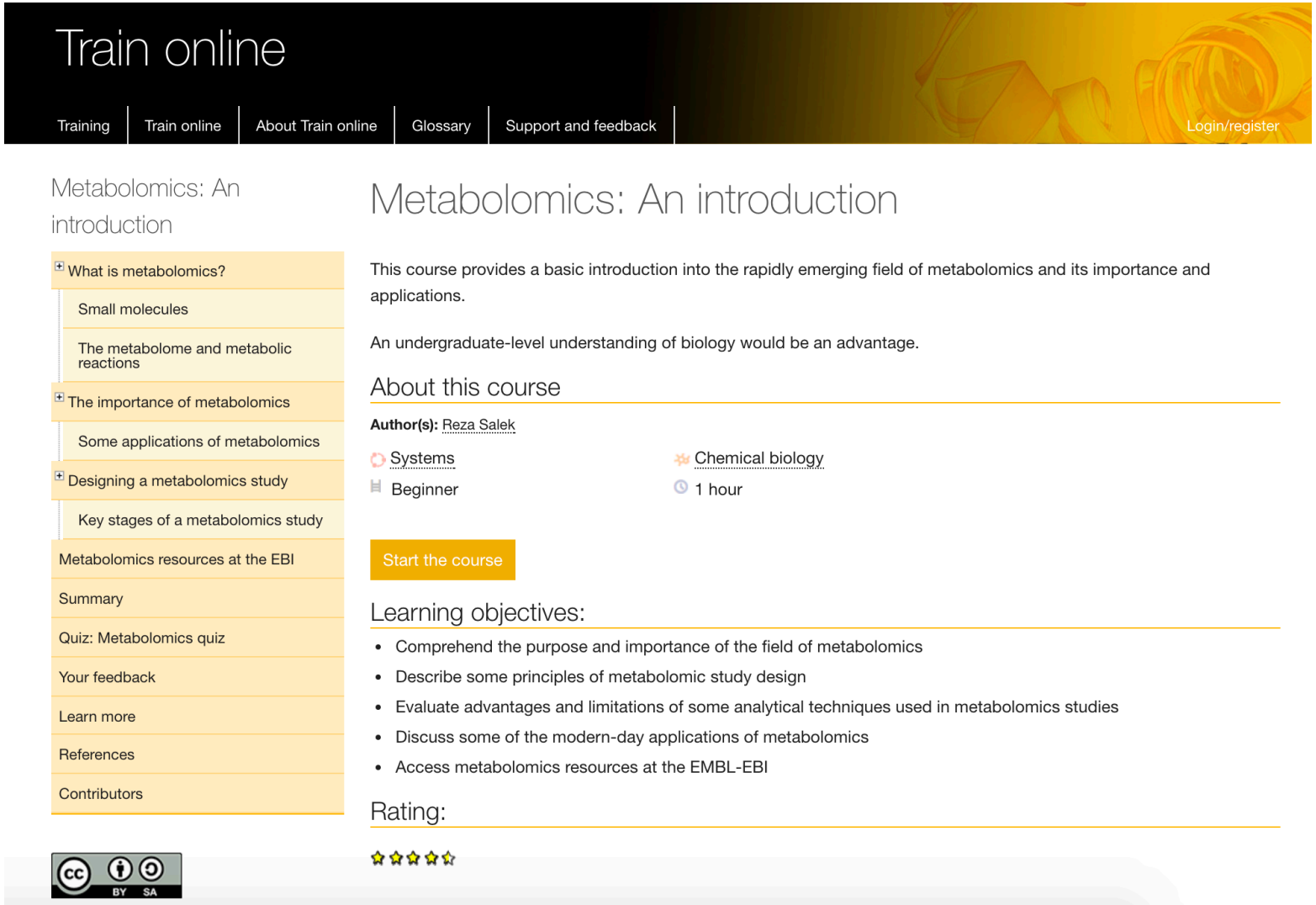
Some interesting references

- 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
http://www.ebi.ac.uk/training/online/course/introduction-metabolomics/what-metabolomics
http://www.metabonews.ca/archive.html

Online course

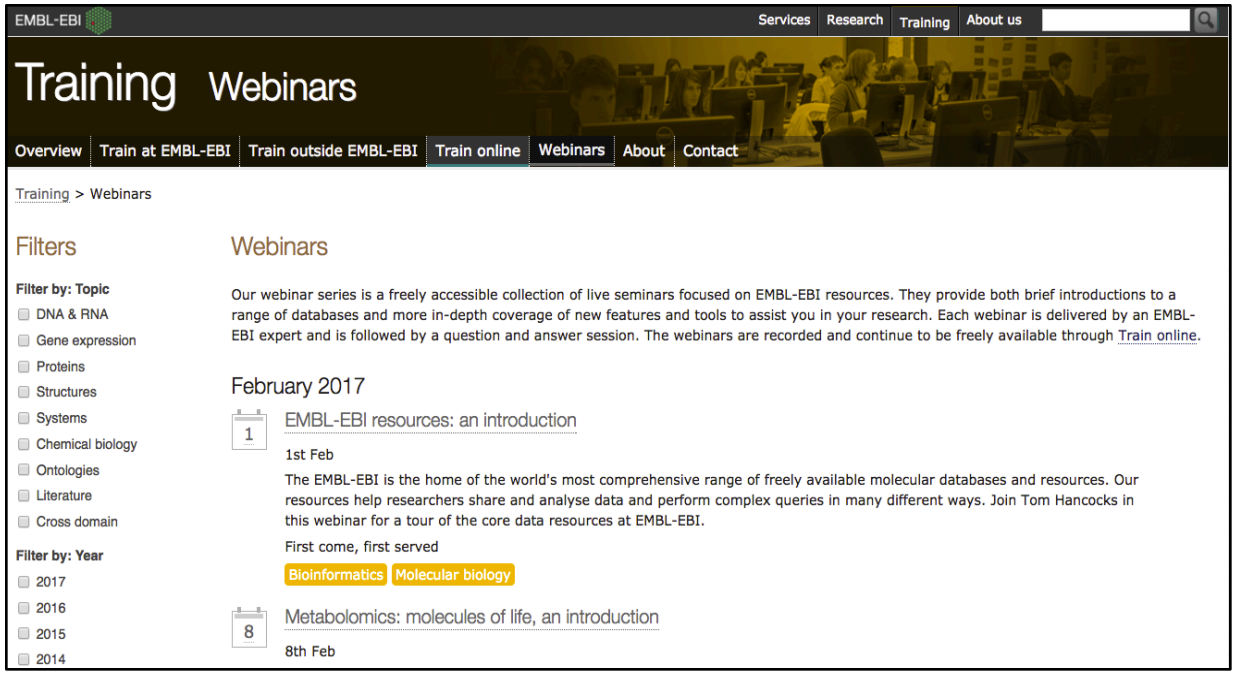


http://www.ebi.ac.uk/training/online/course/introduction-metabolomics

EMBL-EBI

Upcoming webinars

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



Feedback

Tell us what you think