



Knowledge that will change your world

Introduction to metabolomics research

Stephen Barnes, PhD
4-7117; sbarnes@uab.edu

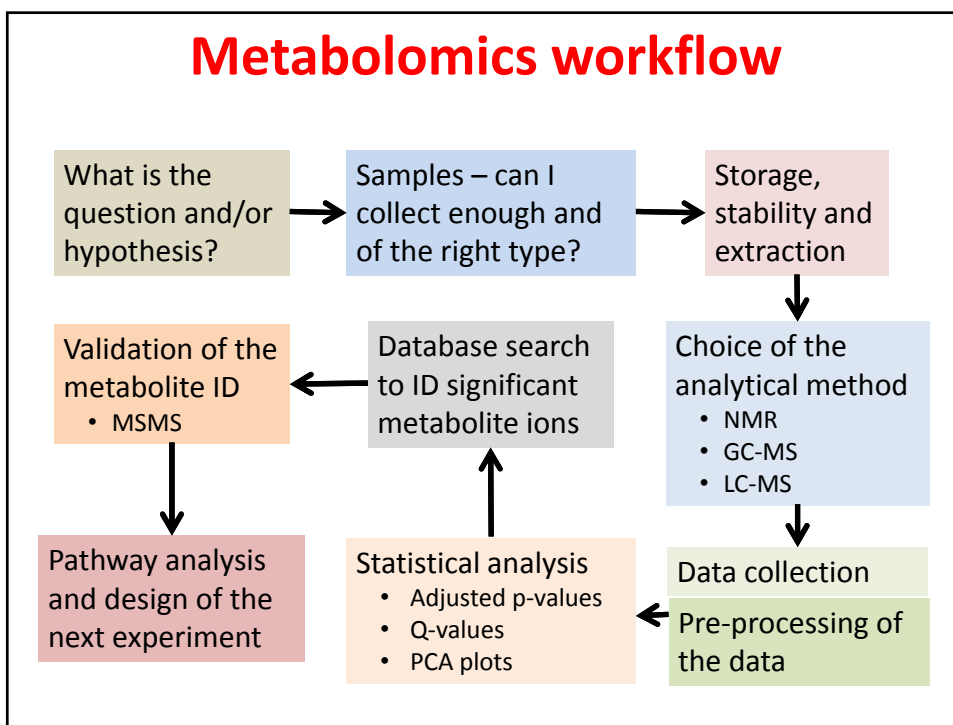
What is “Metabolomics”?

- **Metabolomics is like other types of –omics analysis (microarray, proteomics, etc.)**
 - **Offers a “comprehensive” view of all detectable chemicals (not just metabolites)**
 - **Can be applied to body fluids**
 - Plasma/sera, urine, saliva, tears, etc.
 - **Also to tissues**
 - Liver, lung, heart, kidney, etc.
 - **And to single cells**
 - Human, rodent, yeast, bacteria, etc.

World without gas!



Metabolomics workflow



Course goals

1. To understand the **vital** roles of metabolites
 - To provide energy for the chemical and enzymatic processes of life
 - To provide the building blocks for the macromolecules (DNA, RNA, proteins, carbohydrates, lipids)
 - As co-factors
 - As signalling molecules

Course goals

2. To understand the **origins** of metabolites
 - Produced by (human) cells
 - Produced by **the things that we eat**
 - Plants (wheat, corn)
 - Fruits (apples, oranges, strawberries)
 - Vegetables (rice, potatoes, broccoli, peas)
 - Dairy products, including fermented forms
 - Meat from other animals
 - Xenobiotics
 - Produced by **microorganisms** in our bodies

Complex metabolism of daidzein

Soybean – malonylglucoside of daidzein

Tofu/soy milk – β -glucoside of daidzein

Gut wall – phlorizin hydrolase - daidzein

Dihydrodaidzein
O-Desmethylanglensin

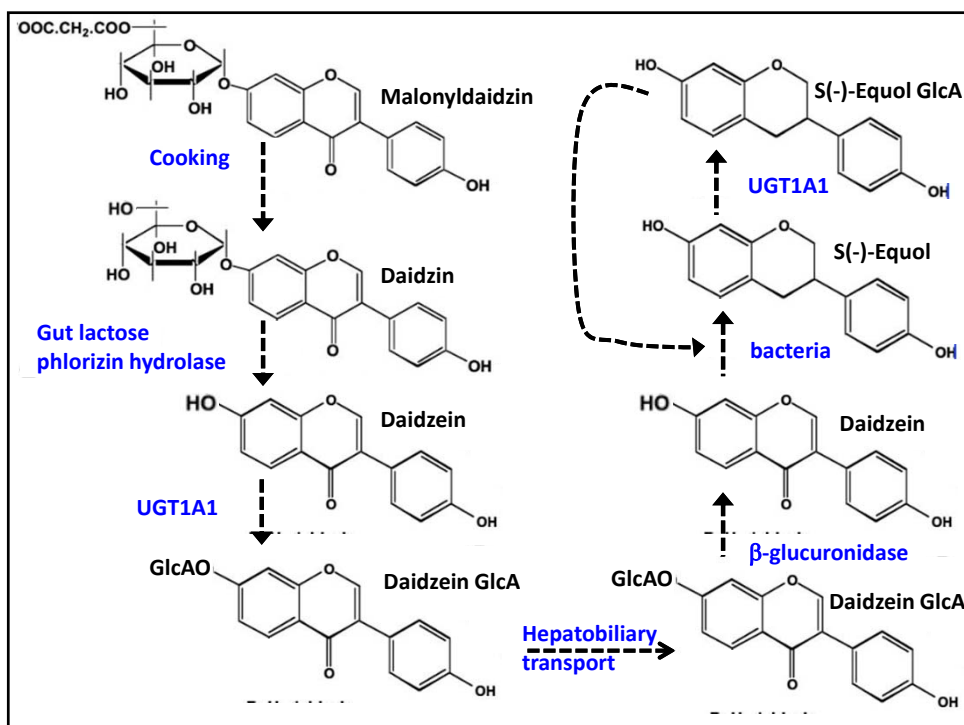
Enterocyte – daidzein β -glucuronide

Liver – hepatobiliary transport

Bacterial hydrolase - daidzein

Bacterial reduction – S(-)-equol

Enterocyte (?) – liver - S(-)-equol β -glucuronide



Course goals

3. To understand that a metabolomics experiment is **high dimensional**
 - i.e., it compares the intensities of hundreds, if not thousands, of distinct species
 - Very important statistical consequences
 - Cannot afford to do a robust experiment that fully satisfies theoretical statistical principles
 - Very important to sit down with a statistician prior to executing an experiment

Course goals

4. To select the appropriate method for extracting/recovering metabolites
 - **Metabolites encompass an enormous range of chemistries**
 - Gaseous (H_2 , H_2S)
 - Volatile (butyric acid, acetone, skatole)
 - Hydrophilic (glucose)
 - Charged-positive/negative (amino acids, nucleotides, organic acids, amines)
 - Hydrophobic (lipids, steroids, hydrocarbons)
 - **No single method suitable for all metabolites**

Course goals

5. Selecting the analytical approach

– *In situ* analysis

- Laser ablation of frozen tissue
- Other desorption methods
- Magic angle spinning NMR
- Other spectroscopic methods

– Extracted samples

- NMR
- GC-MS (1- and 2D chromatography and MSMS)
- LC-MS (1- and 2D chromatography and MSMS)

– Targeted vs untargeted analysis

Course goals

6. Analysis of the data

– Data alignment

- NMR methods
- LC-MS and GC-MS methods (XCMS)

– Statistical evaluation

- Univariate and multivariate analysis
- XCMSonline
- Metabosearch

Course goals

7. Identifying the “interesting” metabolites

- Use of MS (absolute mass)
 - METLIN
 - MetaboSearch
- MSMS (fragmentation spectra)
 - METLIN
- Metabolite standards
- Importance of retention time
 - Multiple column conditions

Course goals

8. Pathways and applications

- MetaboSearch
- KEGG pathway mapping
- Applications to:
 - Adverse cardiovascular risk
 - Diabetes
 - Lens and kidney diseases
 - Cancer

A brief history of metabolomics

Nuclear physics moves to biology

- 1897 JJ Thomson discovers the electron (cathode rays)

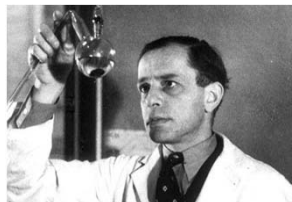


- 1919 Aston using a mass spectrograph shows that Neon with a non-integer MW (20.2 Da) is composed of two isotopes, ^{20}Ne and ^{22}Ne



<http://www.asms.org/Publications/Historical/HistoryofMassSpectrometry/tabid/94/Default.aspx>

Transition to biology



Ralf Schoenheimer



David Rittenberg

- While the politicians, tyrants, dictators and despots were salivating at the thought of developing nuclear weapons from unstable isotopes in the early part of the 20th Century, two scientists began the pursuit of the peaceful use of stable isotopes, initially deuterium (^2H), and later carbon (^{13}C) and nitrogen (^{15}N), to study biochemical pathways
- Understanding the pathways of metabolism was born

Direction of NIH Research 1950-2012

Metabolomics



1995-present Omics galore
Transcriptomics (microarray, deep sequencing, RNA seq)
Proteomics

2012 Human genome ENCODE
project reveals the extent of DNA expression and roles for "junk" DNA, as well as intergenic proteins

2006 First ENCODE project on 1% of the human genome reveals RNAs coming from more than one gene

1950s-60s emphasis on determining metabolic pathways – 20+ Nobel prizes



Bloch Lynen Krebs

1950s-early 1980s
Identification and purification of proteins

1980-1988 Sequencing of genes – cDNA libraries – *orthogonal research*

1988-2000 Sequencing of the human genome – *period of non-orthogonal research* – where did all the genes go? junk DNA?

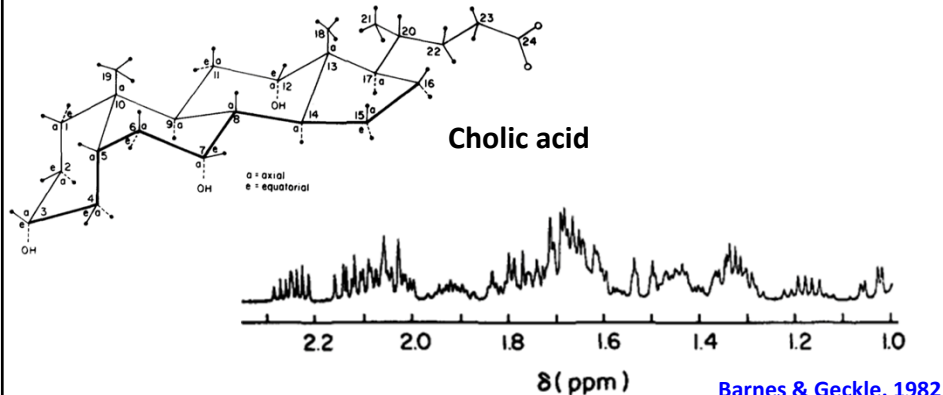
2004 Tiling arrays reveal that most of the genome is expressed

Metabolism to metabolomics

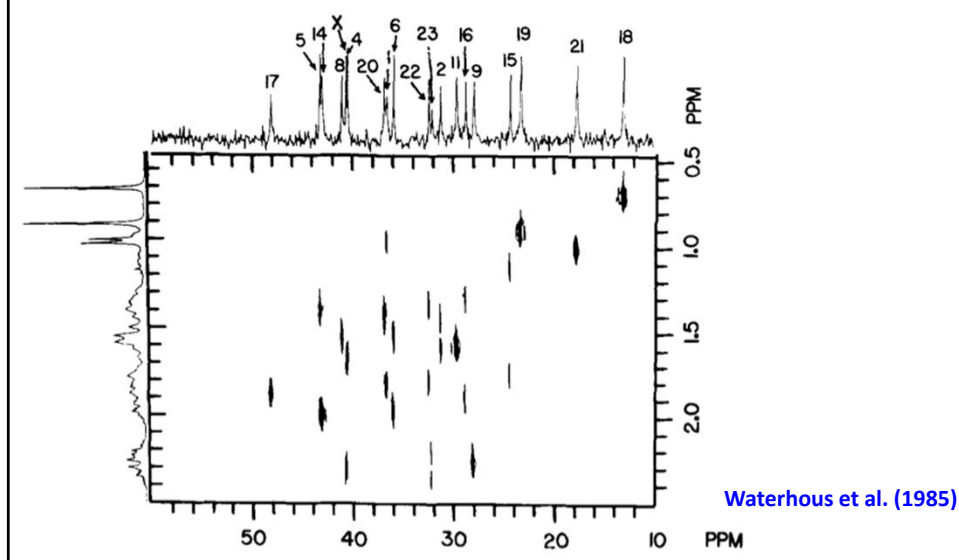
- Measured with enzymes – NAD(P)H absorbance/fluorescence
 - Studies of glycolytic and the TCA cycle intermediates one at a time
- Organic acids, fatty acids and amino acids by GC
 - Volatile derivatives, Flame Ionization Detection
 - GC-MS started in mid-70s
 - Capillary GC gave far higher chromatographic resolution than the packed $\frac{1}{4}$ " ID columns (1975/6)

How NMR became a player

- Mid 60s – introduction of Fourier transform analysis
- Late 70s – introduction of superconducting magnets
- Early 80s - pulse sequences



Pulse sequences in NMR (HetCor)



Gas chromatography

- **Built on critical steps**
 - 1908 Twsett introduces the concept of chromatographic separation (of plant pigments)
 - 1941 Martin and Consden conceptualize the rules of partition chromatography (get the Nobel Prize)
 - 1950 James and Martin describe gas chromatography of volatile fatty acids
 - 1975 (Finally) open tubular, capillary gas chromatography becomes commercially available

Progress in LC-MS

- Commercial HPLC appeared in the early 1970s to separate thermally stable and unstable molecules
- The challenge remained to find a way to get the unstable compounds into the gas phase
 - Applied to macromolecules (peptides, proteins) as well as metabolites
- Thermospray had some initial success
- **Electrospray ionization** and **chemical ionization** radically changed analysis, allowing compounds to go into the gas phase at atmospheric pressure and room temperature

LC-MS

- Suddenly, there were what appeared to be no limits (or very few) to what could be analyzed
- Unheard of, robust mass spectrometers came into play
 - “A reliable mass spectrometer” was considered in 1990 to be an oxymoron

Types of LC-MS analysis

Single quadrupole
LC-MS analysis

LC-time-of-flight
(TOF)-MS

FT-ICR MS

Orbi-trap

Triple quadrupole
LC-MS analysis

Multiple reaction
monitoring (MRM)

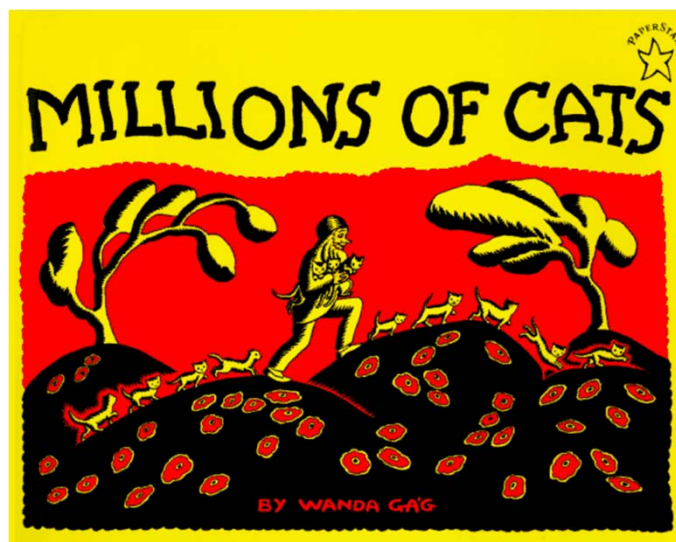
Q-TOF

TripleTOF



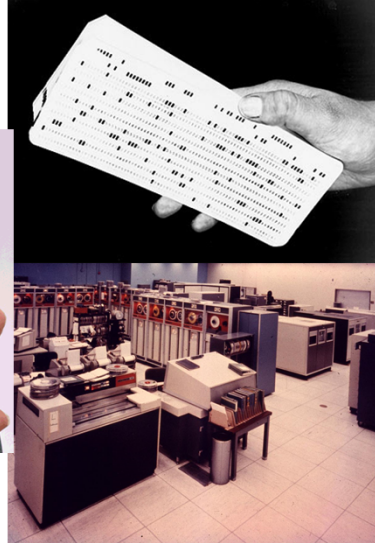
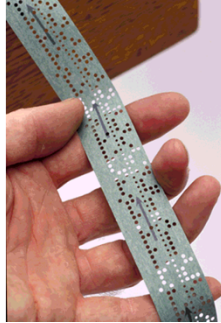
Ion Mobility

Data explosion



Changing times in Computing

- 1950 The Cambridge colleagues of Watson and Crick calculated the structure of DNA by putting data onto punched cards and taking them by train to London for analysis – and to the fog – the “cloud” in 1950s
- 1964 Seymour Cray develops the CDC 6600 (1 Mflops)
- 1967 I used paper tape to collect data from a radio gas chromatograph and then submitted them via a terminal reader to the CDC 6600 at the University of London



Today in Computing



On my desk in 2013

- The Apple MacBook Air with 2 quad core Intel i7 processors
 - Operates at 2.0 GHz
 - Memory of 8 GB
 - Access 1.333 GHz
 - 512 GB Flash memory storage
 - 10 Gbs Thunderbolt I/O
- Also costs ~\$2,000



IBM Blue-Gene

- Parallel processing with 2,048 700 MHz computers operating at 4.733 Tflops
- Replaced by Cheaha, in its current configuration it has 48 compute nodes with two 2.66GHz 6-core Intel CPUs per node (576 cores total)
- It operates at 6.125 Tflops

Great challenges in metabolomics

- **The extent of the metabolome**
 - From gaseous hydrogen to earwax
- **Having complete databases**
 - METLIN has 60,000+ metabolite records, but your problem always creates a need to have more
 - Current lack of a MSMS database
- **Storing and processing TBs of data**
- **Standards and standard operating procedures**
- **Being able to do the analyses in real time**

NIH Common Fund Metabolomics Program

- **Metabolomics Workbench:**
<http://www.metabolomicsworkbench.org/>
- **Regional Comprehensive Metabolomics Research Centers**
 - University of Michigan: <http://mrc2.umich.edu/index.php>
 - UC Davis Metabolomics Center: <http://metabolomics.ucdavis.edu/>
 - RTI International: <http://www.rti.org/page.cfm?objectid=3BC41B11-068E-1405-9A6F79D91D8D69EC>
 - SE Center for Integrated Metabolomics: <http://secim.ufl.edu/>
 - Resource Center for Stable Isotope Metabolomics:
<http://bioinformatics.cesb.uky.edu/bin/view/RCSIRM/>
 - Mayo Clinic Metabolomics Resource: <http://www.mayo.edu/research/core-resources/metabolomics-resource-core/overview>