Knowledge that will change your world

3rd UAB Metabolomics Workshop
June 14-18, 2015

Introduction to metabolomics research

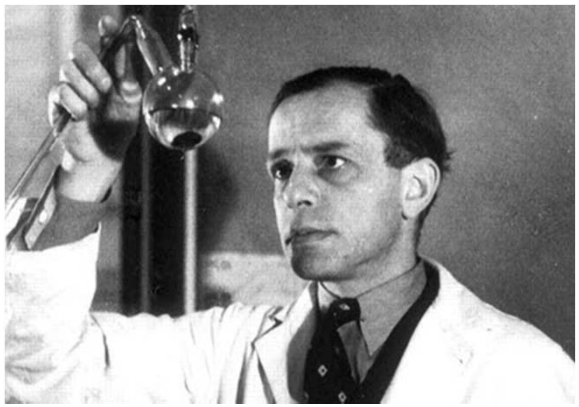
Stephen Barnes, PhD

Director, Targeted Metabolomics and Proteomics Laboratory

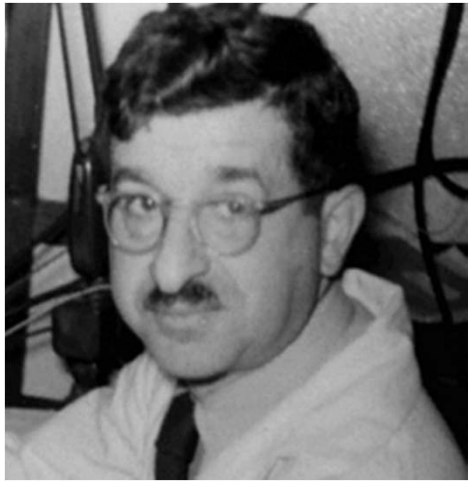
Where did metabolomics come from?

Are metabolomics and metabonomics different?

From nuclear weapons 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 (^{2}H), and later carbon (^{13}C) and nitrogen (^{15}N), to study biochemical pathways**
- **Understanding the pathways of metabolism was born**

Direction of NIH Research 1950-2015

Metabolomics

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

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

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

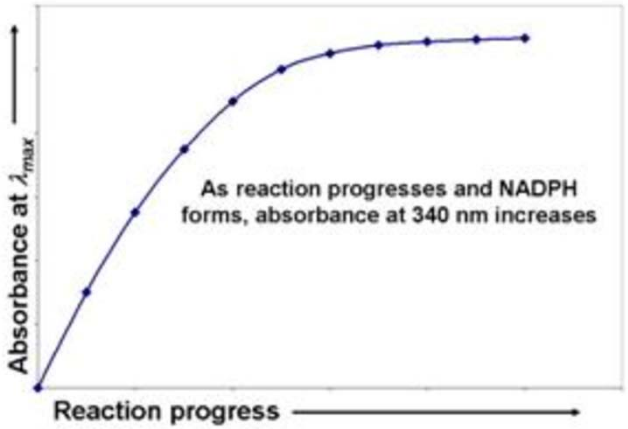
2012 Human genome ENCODE project reveals the extent of DNA expression and roles for "junk" DNA, as well as intergenic proteins

2014 –"deep" proteomics reveals the presence of 400+ proteins that are not encoded by the genome

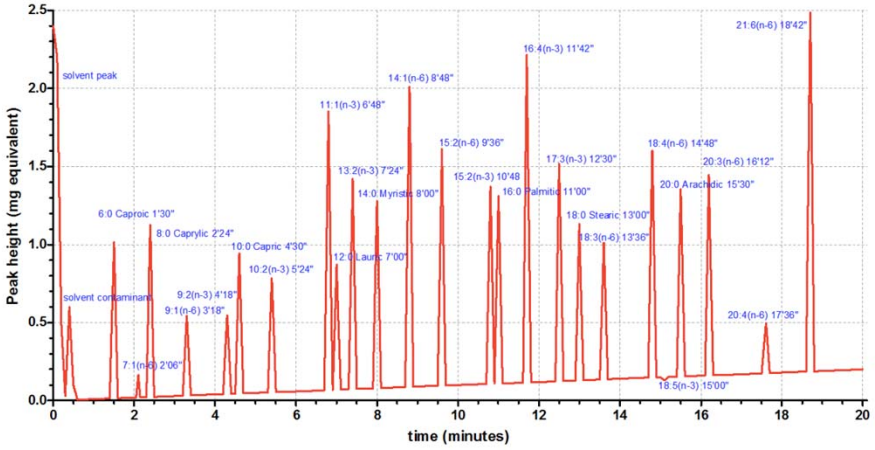
Bloch Lynen Krebs

Metabolism to metabolomics

- **Measured with enzymes – NAD(P)H absorbance/fluorescence**
  - **Studies of glycolytic and the TCA cycle intermediates one at a time**



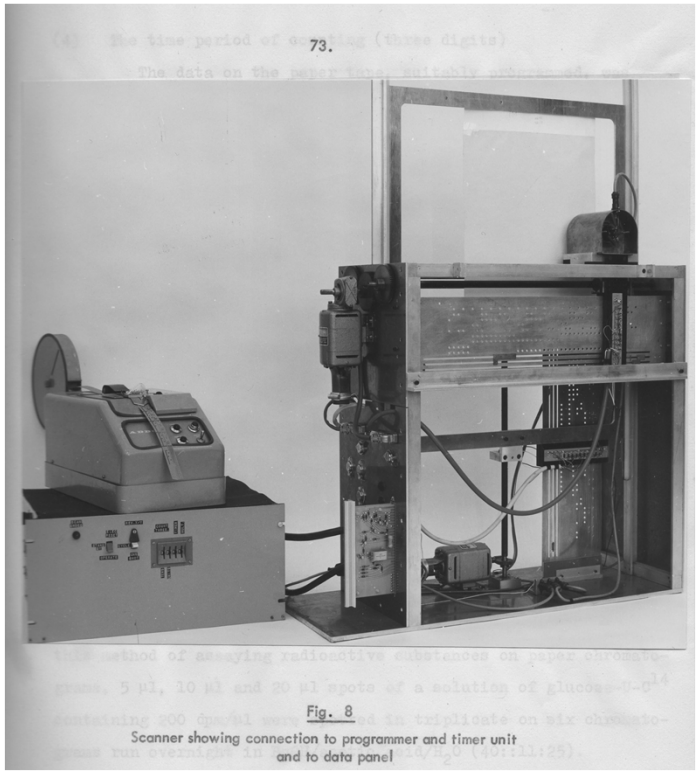
- **Organic acids, fatty acids and amino acids by GC**
  - **Volatile derivatives, Flame Ionization Detection**
  - **GC-MS started in mid-70s**
  - **Open tubular capillary GC gave far higher chromatographic resolution than the packed ¼" ID columns (1975/6)**



https://d1u1p2xjjiahg3.cloudfront.net/bcb5de3b-6cee-4c92-bb76-aa0c96198806.jpg

Origins of practical metabolomics

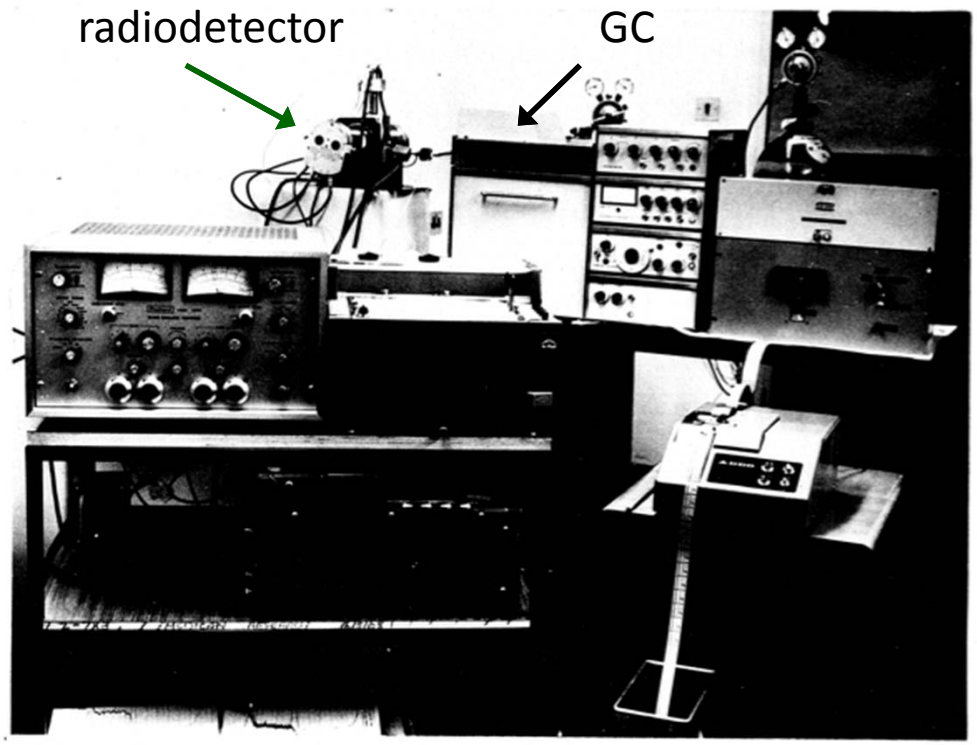
Imperial College 1967-1970



Radio 2D-paper chromatography scanner with digitization of collected data

The room had 20 of these scanners – data analyzed by a central computer (in 1968)

Courtesy of K.R. Mansford, PhD



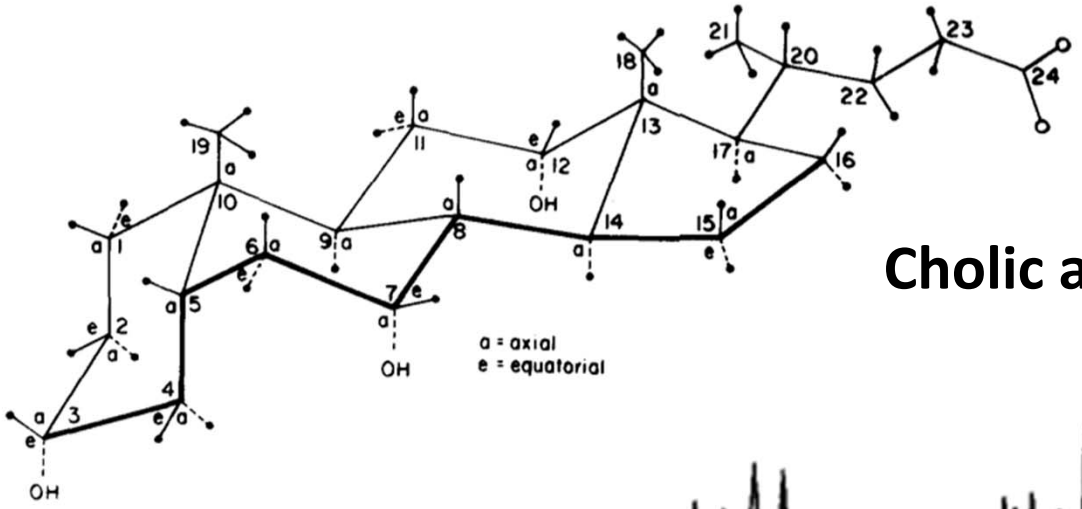
Radio gas-liquid chromatography with digitization of collected data

Developed this for my PhD work (1967-1970) to study glucose metabolism in acellular slime moulds

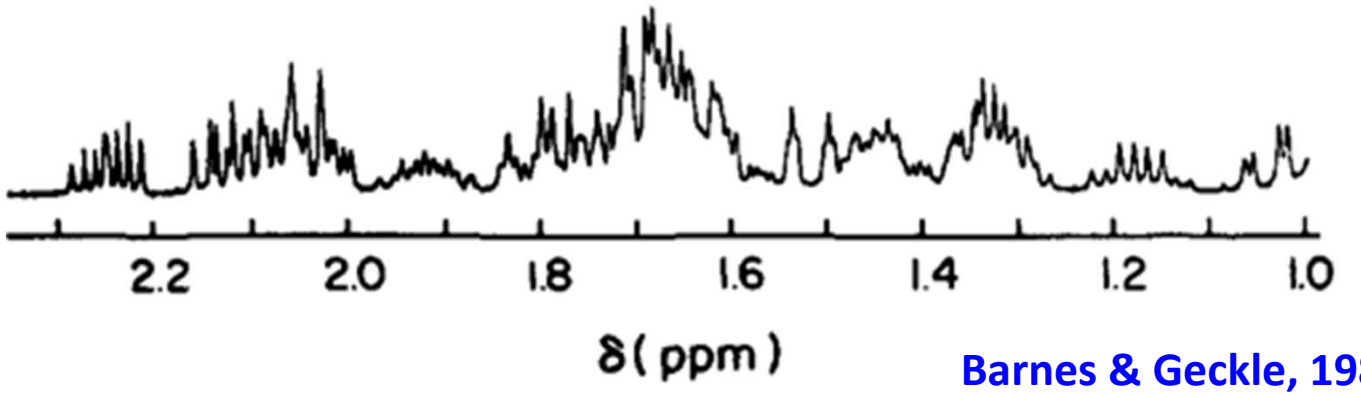
Metabonomics is a term coined by those pioneering NMR metabolomics

How NMR became a player

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

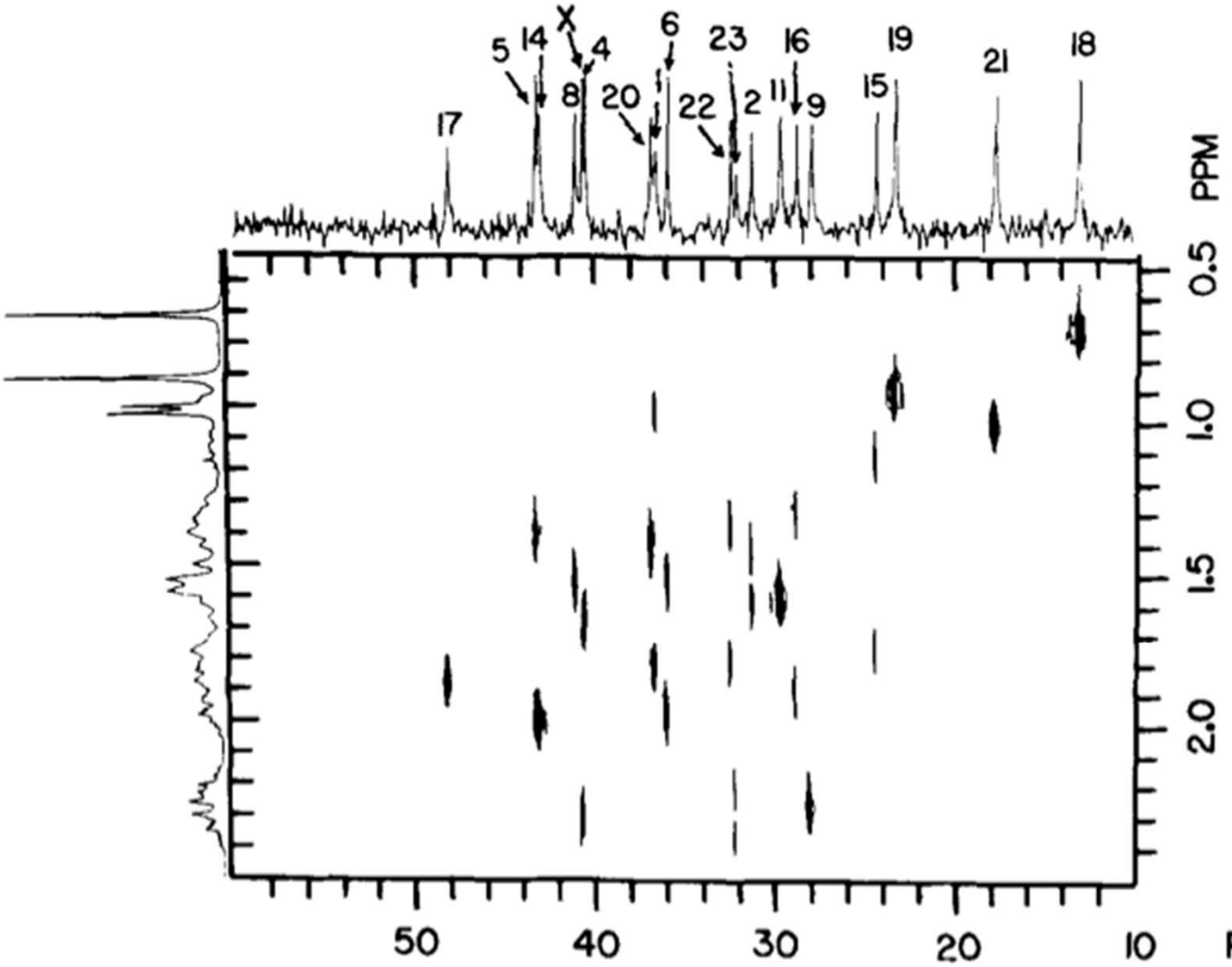


Cholic acid



Barnes & Geckle, 1982

Pulse sequences in NMR (HetCor)



Cholic acid

Waterhous et al. (1985)

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**

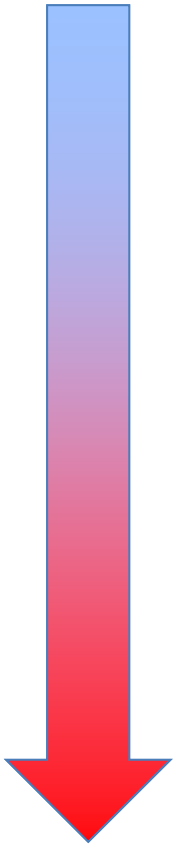
Types of LC-MS analysis

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

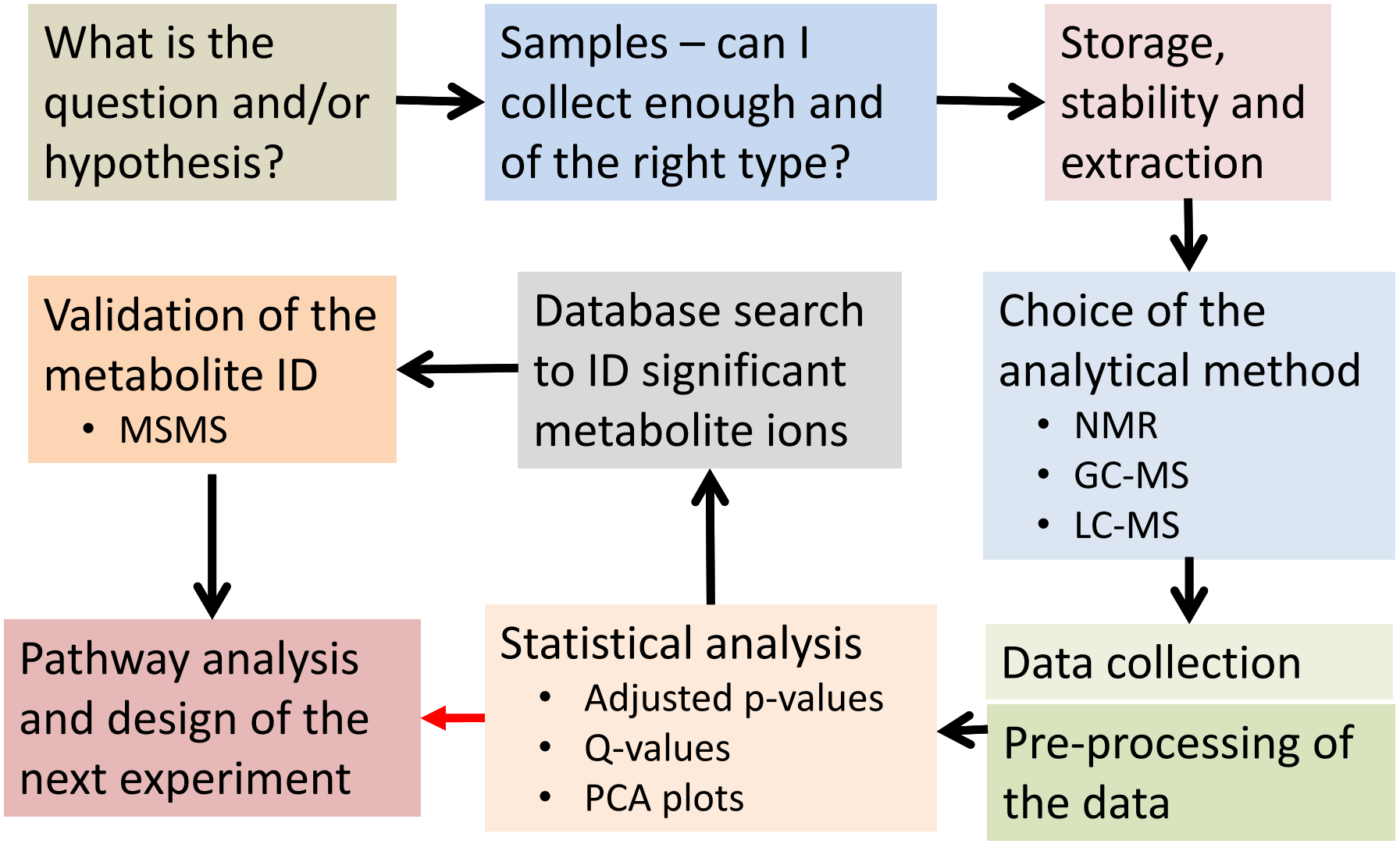
TripleTOF

Ion Mobility

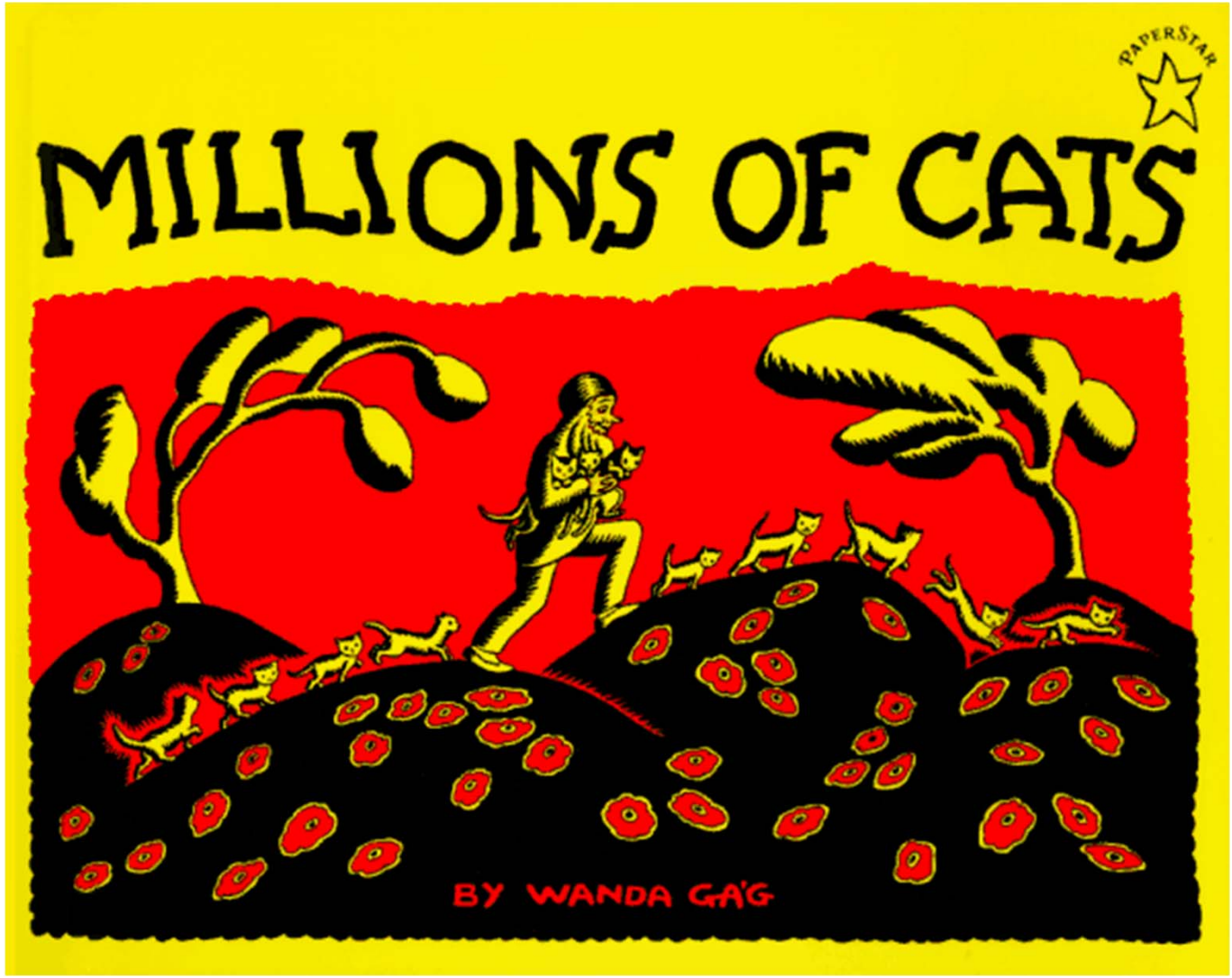
World without gas!

PHOTO: BRENDAN SMIALOWSKI/AFP/GETTY IMAGES

Metabolomics workflow

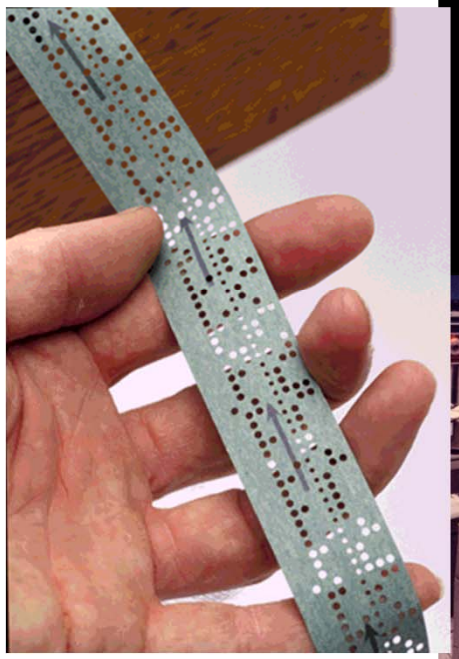


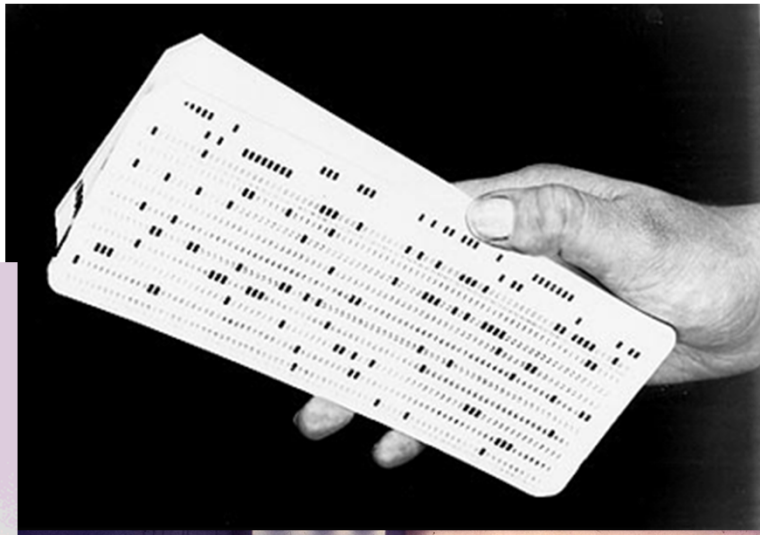
Data explosion



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 2015

- 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 cost ~$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

Does the “cloud” present a viable option?

Yes, if we can transfer the data to other computers with greater processing power and cheaper long-term storage, but……..

Is this really a “safe” solution?

The Cloud and computing in 2015

- The manufacturers are turning to putting software and your data into the Cloud (assuming you can overcome HIPPA constraints)
- In proteomics, they are putting their programs there
  - SCIEX is using BASESPACE (with Illumina)
  - You upload your data to an Amazon server
  - The programs are downloadable Apps
- For now, metabolomics uses XCMS
  - Either online or as a server-based software
  - Cloud next?

See 2013 and 2014 workshop discussions on computing by Sean Wilkinson

http://www.uab.edu/proteomics/metabolomics/workshop/workshop_july_2013.php

http://www.uab.edu/proteomics/metabolomics/workshop/workshop_june_2014.php

The metabolome is very complex!

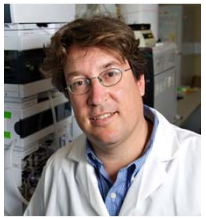
Great challenges in metabolomics

- 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 substantial MSMS database (but it's coming)
- Storing and processing TBs/PBs of data
- Standards and standard operating procedures
- Being able to do the analyses in "real time"

I want to start metabolomics, but who with?

Charles Burandt
MRC2
U. Michigan

Art Edison
SECIM
U. Florida



Oliver Fiehn
UC-Davis

NIH Common Fund
Regional
Comprehensive
Metabolomics
Research Centers

Sreekumaran Rao
Mayo Clinic

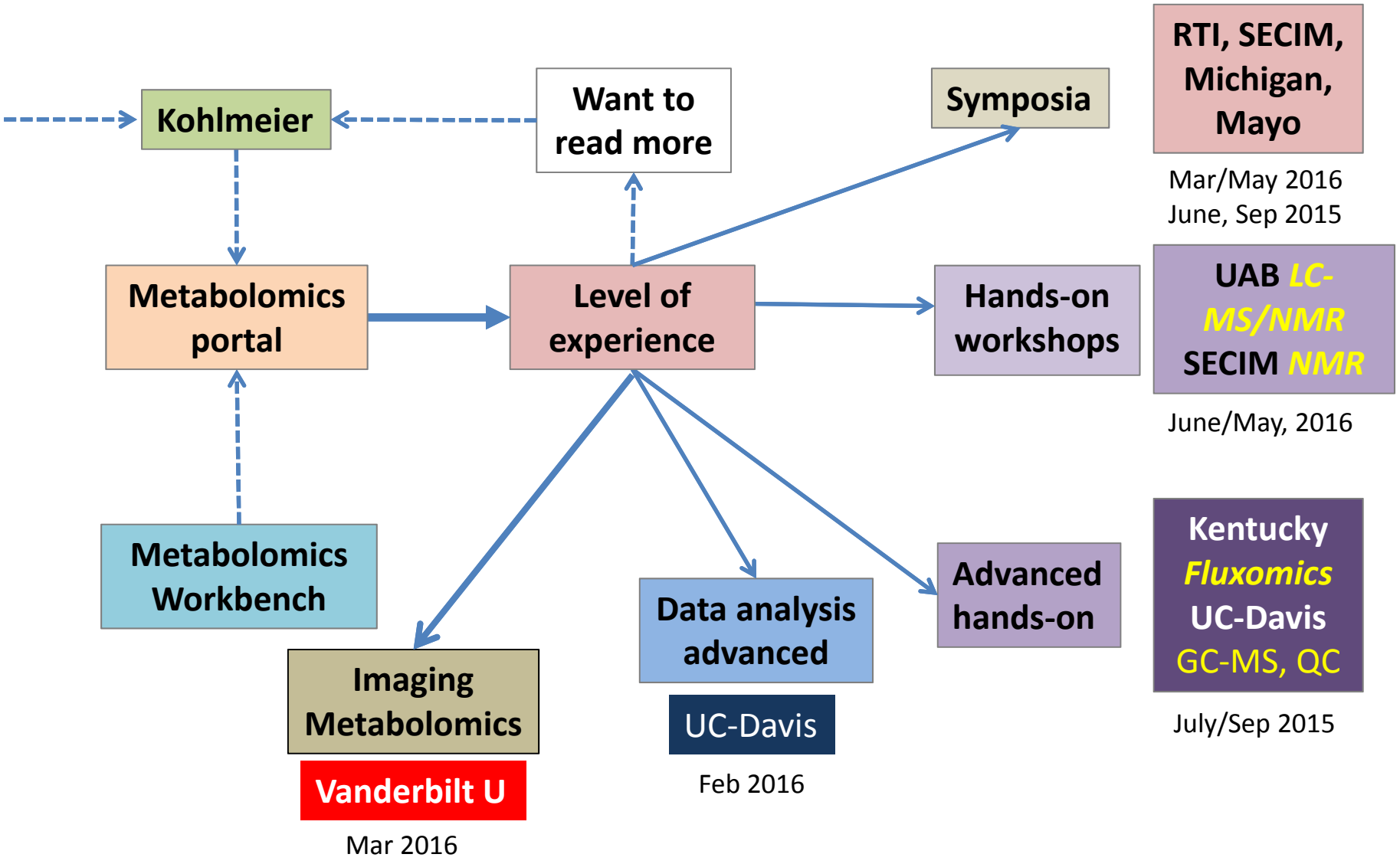
Susan Sumner
RTI International

Rick Higashi
U. Kentucky

http://www.metabolomicsworkbench.org

Each of these regional centers has a pilot program, typically up to $50k with annual deadlines in mid-February (last one in 2016)

Workflow for metabolomics training



Structure of the workshop

- Introduction to experimental design
  - Optimal planning and sample collection
- Sample processing/extraction
  - Primary data collection by NMR, LC-MS and imaging
- Introduction to data processing and statistical analysis
- Electives:
  - Advanced data processing; pathway analysis
  - Advanced sample processing for imaging; ion mobility analysis; MSMS interpretation
- Integration of metabolomics and its future

Terrific speakers

- **Richard Caprioli (Vanderbilt)**
  - **Director of the National Imaging Mass Spectrometry Center**
- **Art Edison (U Florida)**
  - **Director of SECIM**
- **David Wishart (U Alberta)**
  - **Pioneer in the development of metabolomics**

Returning speakers

Kathleen Stringer,
PharmD, U Michigan

Wimal Pathmasiri, PhD
RTI Intl

Xiuxia Du, PhD
UNC-Charlotte

Rodney Snyder, MS
RTI Intl

Paul Benton, PhD
Scripps Res Inst

Shuzhao Li, PhD
Emory U

UAB trainers

Janusz Kabarowski, PhD

Matthew Renfrow, PhD

Jeevan Prasain, PhD

N. Rama Krishna, PhD

Landon Wilson

Ali Arabshahi

D. Ray Moore II

Haley Albright

Ronald Shin, PhD

Industry Speakers

Rob Mohney
Metabolon

Jeremiah Tipton
SCIEX

Roy Martin and Tom Beaty
Waters

Thank you – questions?