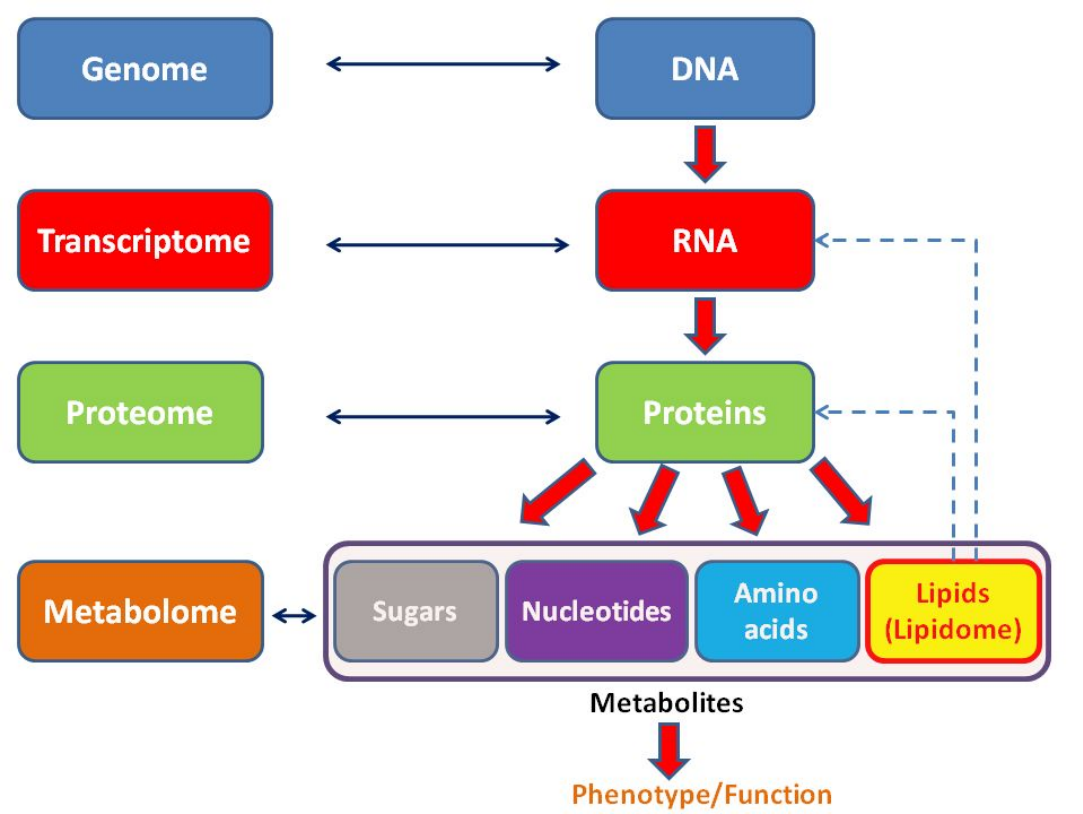
Metabolomics

The Mechanism of Life

The Machinery

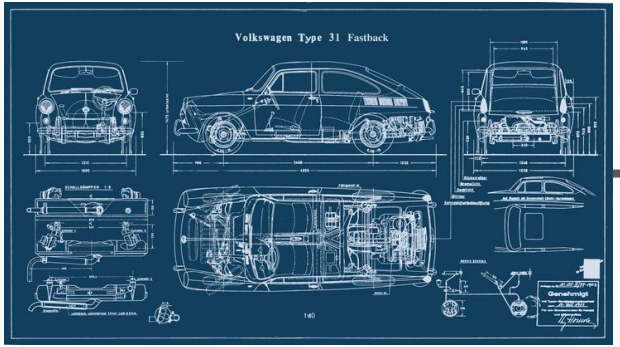
The Fuel/Exhaust



https://en.wikipedia.org/wiki/Metabolome

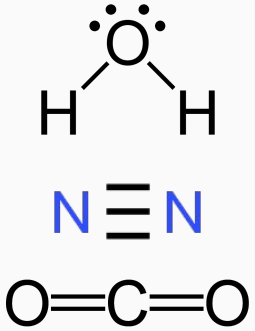
Central Dogma Plus

Genome



Transcriptome

Proteome

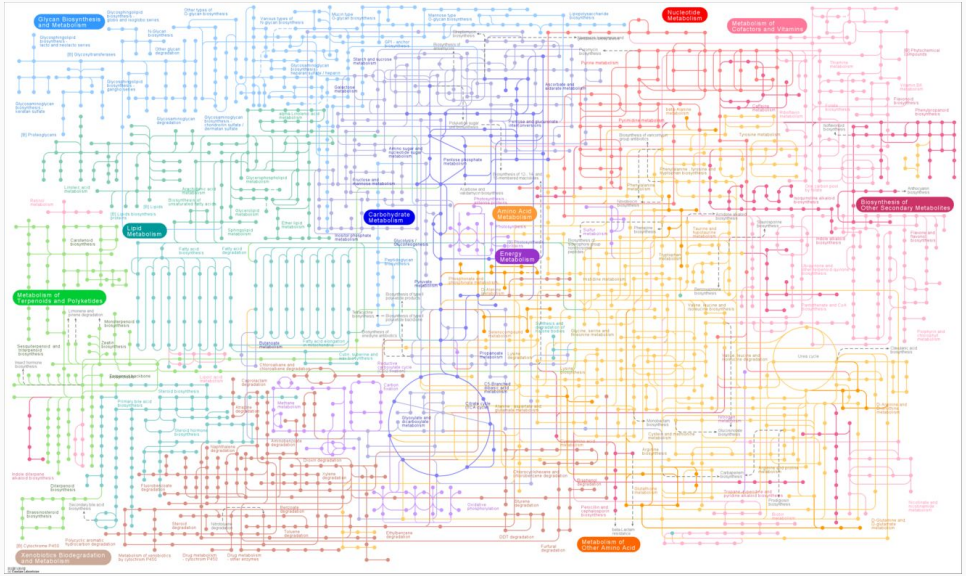


Metabolome

Terminology

Metabolism: Network of chemical reactions

- How a cell extracts energy from its environment
- How a cell synthesizes its “building blocks”

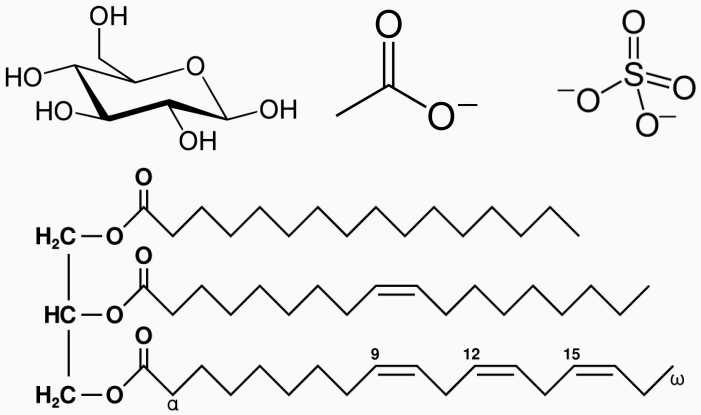


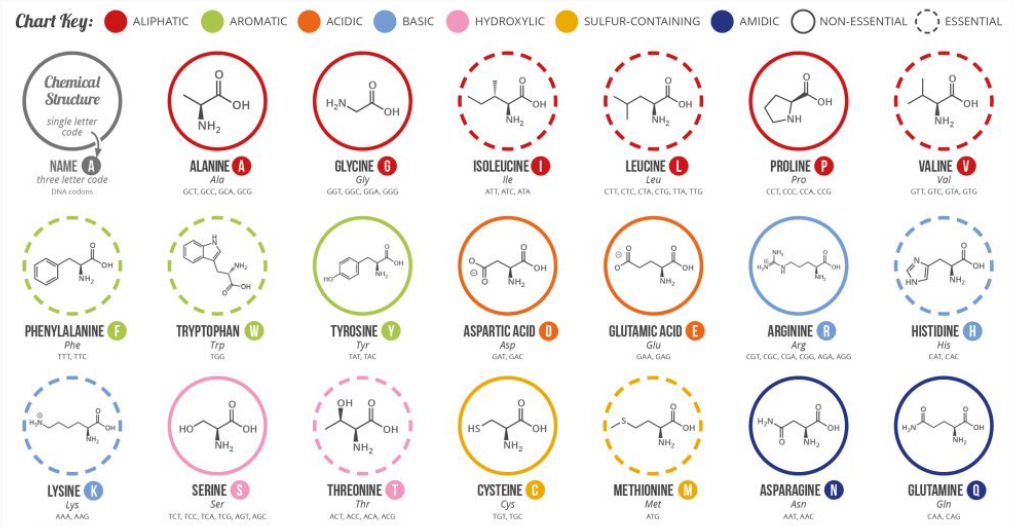
Terminology

Metabolite: Chemical compounds used in metabolism

- Amino acids, glucose, acetate, lactate, *etc.*

Metabolome: Metabolites produced/used by cells





The Human Metabolome

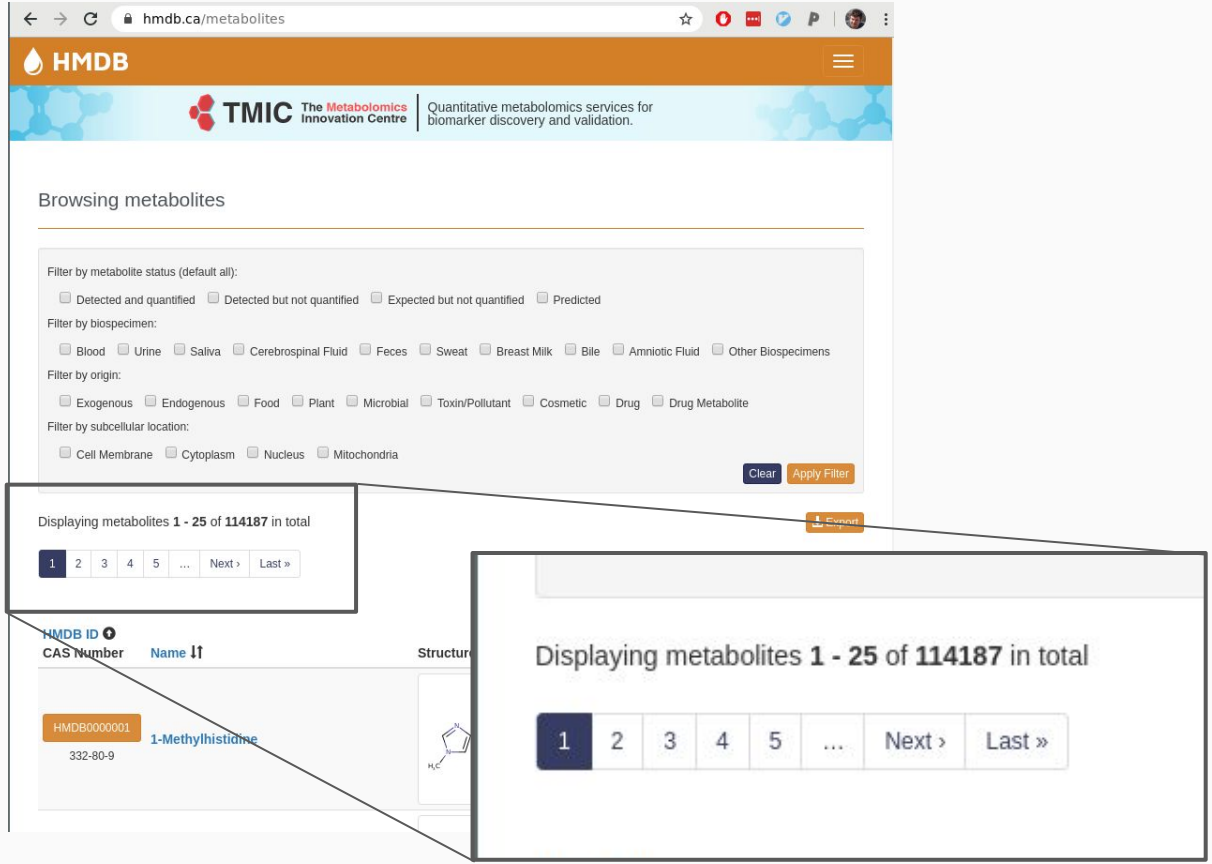
Human Metabolome DB: hmdb.ca

100k+ metabolites

Various locations in human body/fluids

Various origins

- Endogenous - created by cells
- Exogenous - ingested (i.e. food)



Workflow

1. **Collect samples:** fluids, cells, media, etc
2. **Separate molecules:** chromatography
3. **Detect using one of:**
    a. Mass spectrometry
    b. NMR (Nuclear Magnetic Resonance)
    c. UV (Ultraviolet-Visible) or IR (Infrared) Spectroscopy
    d. Flame Ionization
4. **Analyze**: bioinformatics + statistics

Collect: Intra- vs. Extra-cellular

Can perform intra-cellular metabolomics, but it's simpler to perform extra-cellular metabolomics (*exometabolomics*)

Extracellular sources - no cells:

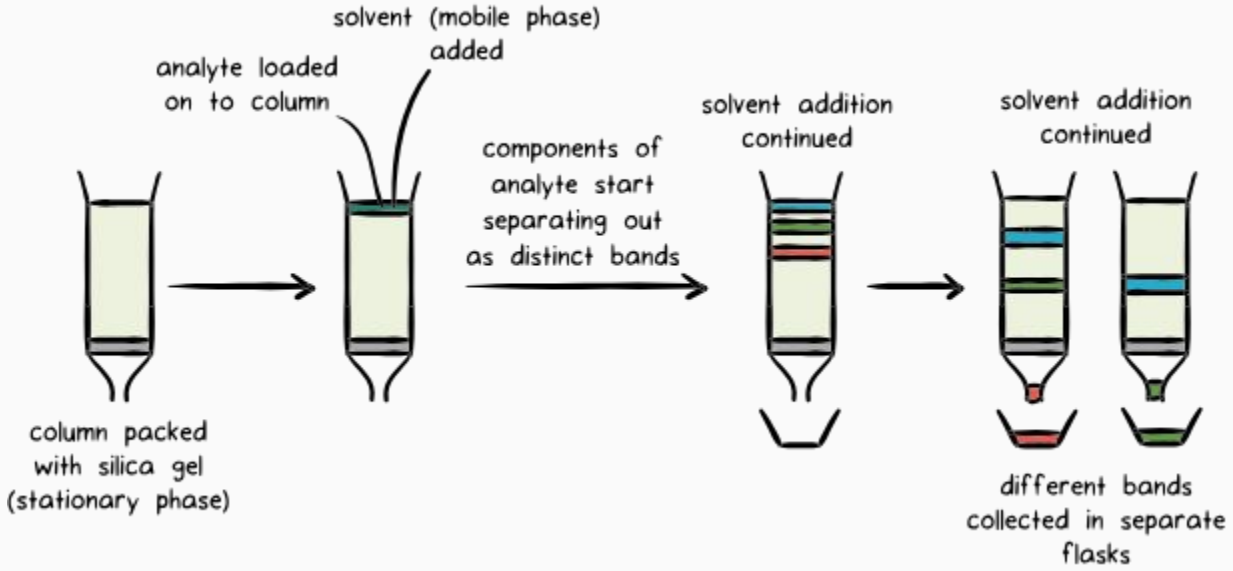
- Tissue: blood serum, cerebrospinal fluid (CSF), saliva, sweat etc
- Microbial: Just look at the (spent) medium! What did the microbes secrete/take up?

Intracellular is very complicated to do properly

- Must isolate cells from media
- Complex experimental setup may be required: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4559092/

Separate: Chromatography

Can separate by size, charge



GC vs LC: Gas vs Liquid Chromatography

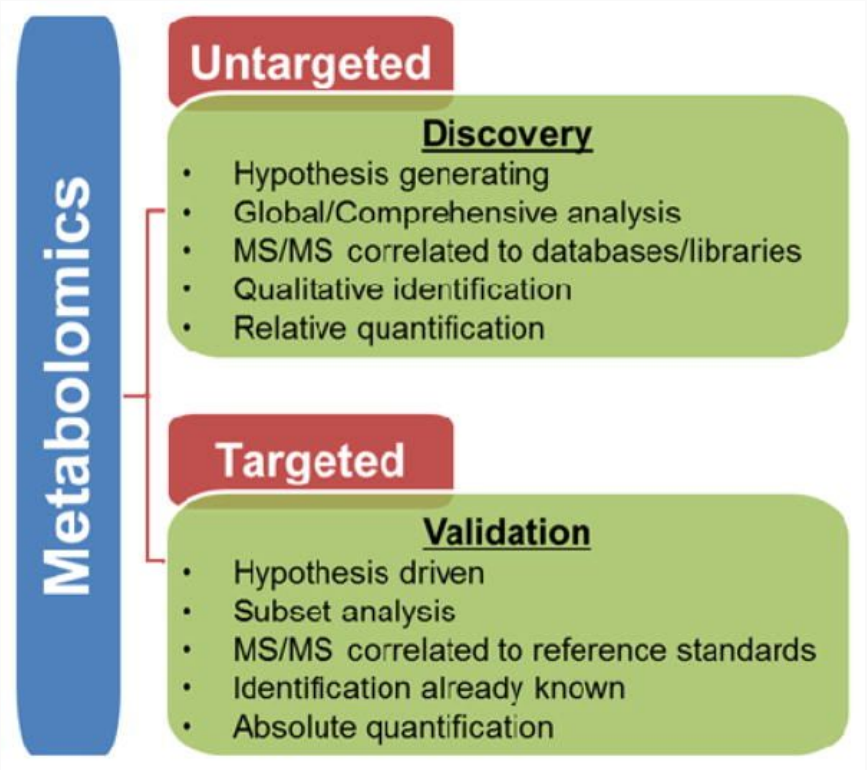
Design Choice: Targeted vs. Untargeted

Untargeted:

- What metabolites are there?
- Qualitative
- Measure “all” of the metabolites in the sample

Targeted:

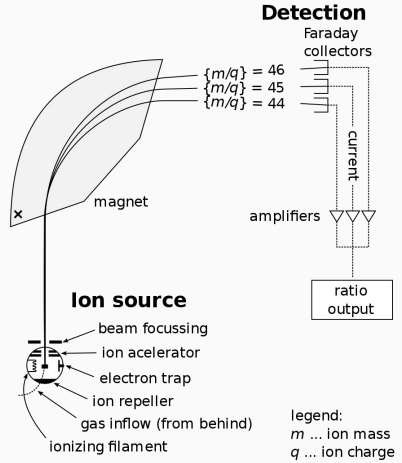
- Measure a specific set of metabolites
- Quantitative

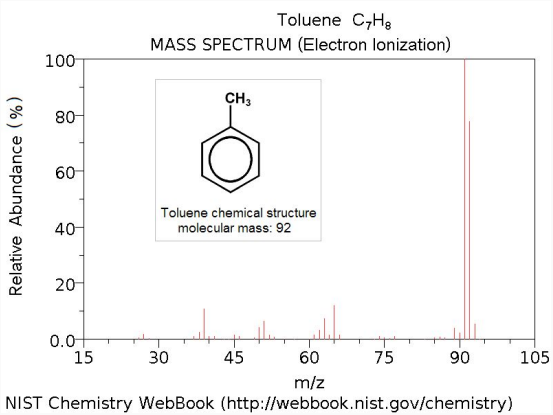


Schrimpe-Rutledge, Alexandra C., Simona G. Codreanu, Stacy D. Sherrod, and John A. McLean. 2016. “Untargeted Metabolomics Strategies-Challenges and Emerging Directions.” *Journal of the American Society for Mass Spectrometry* 27 (12): 1897–1905.

Detect: Mass Spectrometry

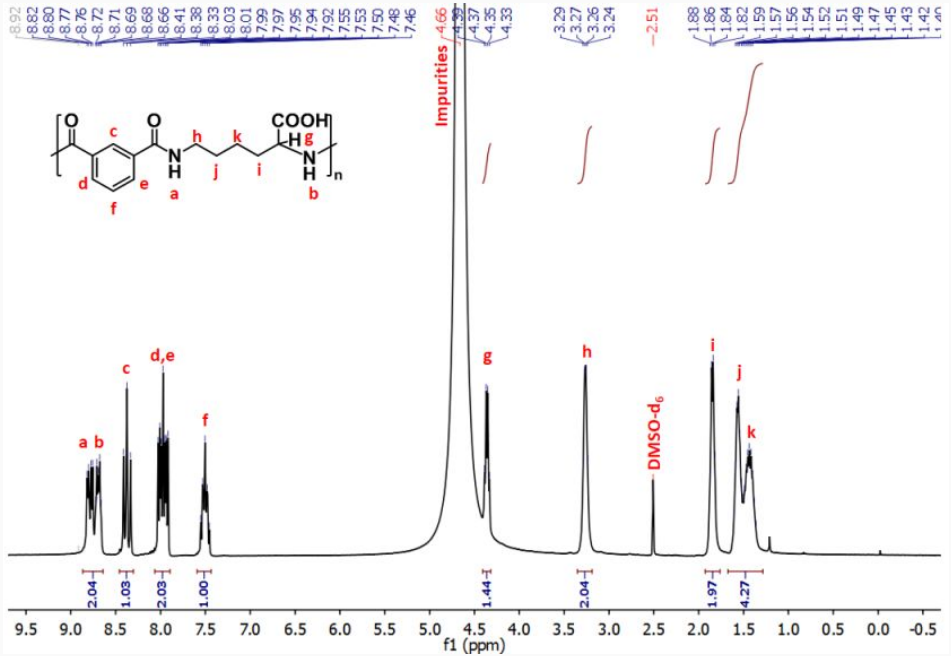
- Identify mass/charge (m/z) of particles
- Vaporizes compounds from chromatograph
- Each molecule has a (hopefully unique) set of m/z spectra





Detect: NMR Spectrometry

- Measures magnetic resonance between atoms in a structure
- Each atomic bond configuration has specific signature
- Analyze spectra to identify content



Detection Methods

Need to carefully select your detection platform(s)

- Many complementary, analytical trade-offs

Experimental reagents, sample manipulation, *etc.* impact your results

Some media are less amenable to metabolomics

- High salt concentrations mess up chromatography columns

Isotopic tracers/labels

Analyze: Metabolomics Data

- Quantities of metabolites for many samples
- **Forms a matrix**
- Pattern detection analysis
  - Unsupervised: PCA, Hierarchical Clustering, etc
  - Supervised: Regression/Partial Least Squares
- Pathway enrichment
  - Map metabolites → pathways, identify enrichment
  - Input to Flux Balance Analysis

Analysis Considerations

- No single accepted metabolomics analysis approach
- Metabolite abundances thought to be “truncated log-normal” distributed
  - Truncated because abundances can’t be negative
- Check for log-normality yourself! (e.g. Wald test)
- If log-normality holds, can use any statistical method that assumes normality
  - e.g. Linear regression, PCA, etc...

Example Study

Original Article | Open Access | Published: 23 March 2020

A population-based resource for intergenerational metabolomics analyses in pregnant women and their children: the Generation R Study

Ellis Voerman, Vincent W. V. Jaddoe, Olaf Uhl, Engy Shokry, Jeannie Horak, Janine F. Felix, Berthold Koletzko & Romy Gaillard

Metabolomics **16**, Article number: 43 (2020) | Cite this article

- Prospective study of 994 mother-child pairs
- Either mother, infant, or 10y/o
- Sampled blood serum or umbilical cord blood
- Liquid chromatography followed by mass spec (LC-MS) to target:
  - Amino acids
  - Non-esterified fatty acids
  - Phospholipids
  - Carnitines
- Measured ~200 metabolites in each sample

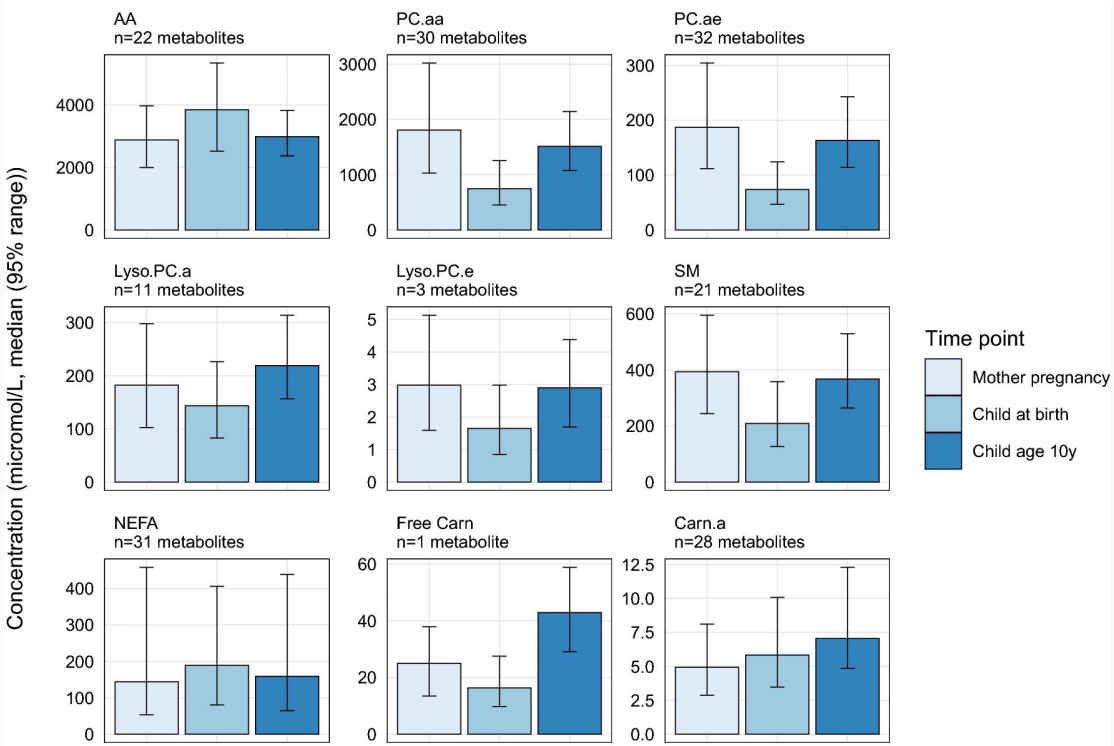
Example Study: Differential Concentration

Metabolites groups by structure:

- AA - amino acid
- NEFA - non-esterified fatty acids
- PC - phospholipid
- Carn - carnitines
- Etc

Plotted absolute concentration divided by group

“Differential concentration” analysis



Example Study: PCA

Principal Component Analysis of all metabolite concentrations

More PCs required to explain variance in 10y/o compared with other sample types

Time point	Number of metabolites	Number of PCs				
		50%	75%	85%	95%	99.5%
Mother early pregnancy	195	3	15	35	88	163
Child at birth	194	4	21	46	101	169
Child age 10 years	181	6	27	50	98	157

% explained variance

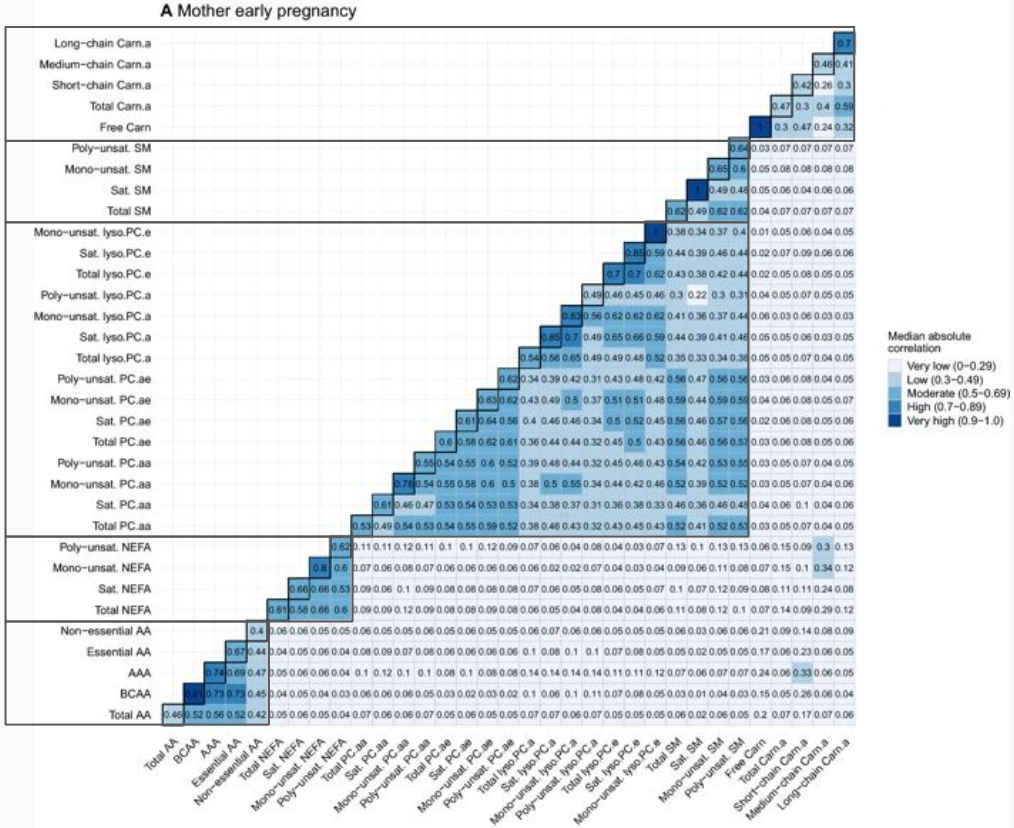
# of components required

Example Study: Correlation w/in Individual

Metabolite concentration correlated by structure group

Plotted median absolute correlation (most are +) of all pairwise metabolites from each grouping

Correlation structure of mother in pregnancy similar to infant (not shown here, see study)



Common Challenges

Metabolomics analyzes a **broad range of features**

- Small, hydrophilic carbohydrates (*e.g.* glucose) – Large, hydrophobic lipids (*e.g.* triacylglycerides) – Complex, natural compounds (*e.g.* antibiotics)
  - Different sizes, charge, composition
  - Different biological roles and implications

Many metabolites are not in databases: our picture of cellular metabolism is incomplete

Identification of novel metabolites is difficult

Approaches for Microbiome Analysis

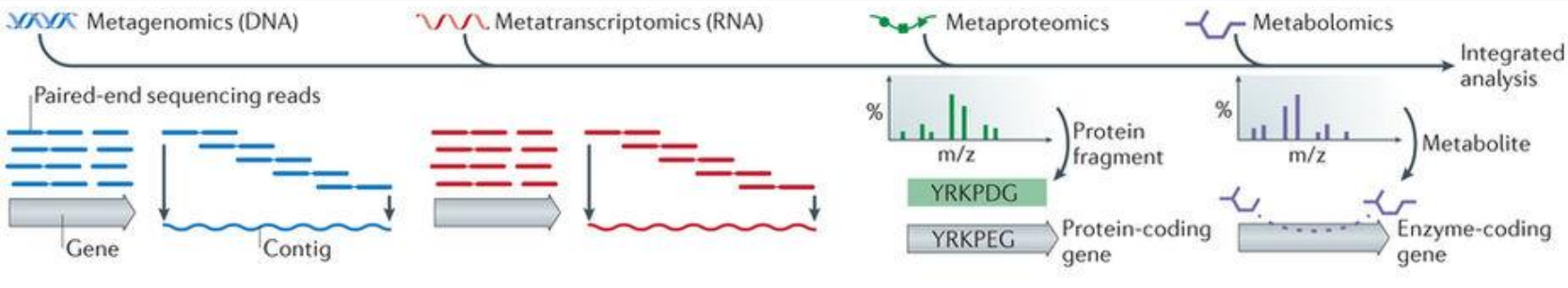
“Metataxonomics”: What is the composition?

- Use marker genes (*e.g.* 16S rRNA)

Metagenomics: What is the composition and functional potential?

Metatranscriptomics: What genes are collectively expressed?

Metabolomics: What metabolic byproducts are used/produced?



Detection Methods

LC/MS	NMR
Separation required (sometimes multiple chromatographic methods)	No separation required
Can resolve and quantify individual metabolites in complex mixtures	Limited ability to resolve complex mixtures
Biased against volatile metabolites	
High sensitivity and dynamic range	Low(er) sensitivity
Limited capacity for quantification	Absolute quantification of metabolites

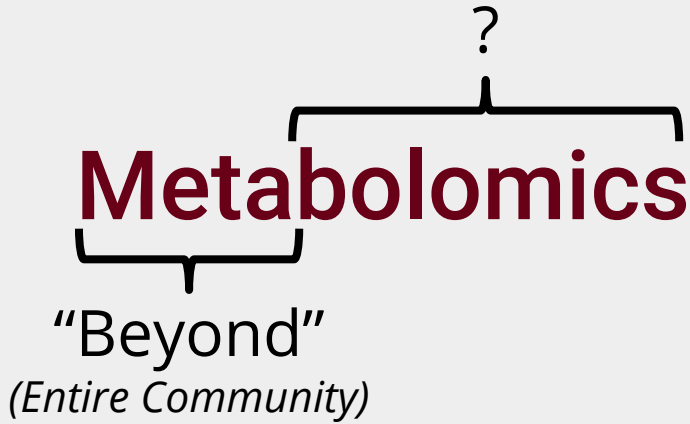
Detection Methods

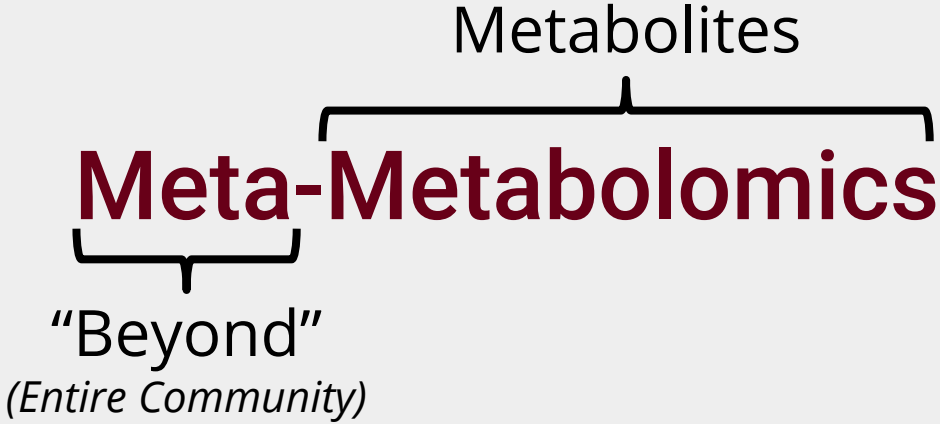
LC/MS	GC/MS
Separation required (sometimes multiple chromatographic methods)	Separation required
Can resolve and quantify individual metabolites in complex mixtures	
Biased against volatile metabolites	Can quantify volatile and uncharged metabolites, isomeric compounds (*e.g.* sugars, lipids)
High sensitivity and dynamic range	
Limited capacity for quantification	Absolute quantification of metabolites

Why Metabolomics?

Metagenomics enables the prediction metabolites (based on what enzymes are present)
Metabolomics enables the quantification/identification of metabolites through direct measurements

Provides a good measure of **phenotype** and **biochemical activity**


?
Metabolomics
“Beyond”
(Entire Community)


Metabolites
Meta-Metabolomics
“Beyond”
(Entire Community)