Metabolomics in Systems Biology

Basil J. Nikolau
W.M. Keck Metabolomics Research Laboratory
Iowa State University

February 7, 2008

Outline

- What is metabolomics?
- Why is metabolomics important?
  - What are the issues for addressing this question?
- How are metabolomics studies conducted?
- Challenges of metabolomics
- Integrating metabolomics and enzymology in a system
  - What does metabolomics reveal about biology?

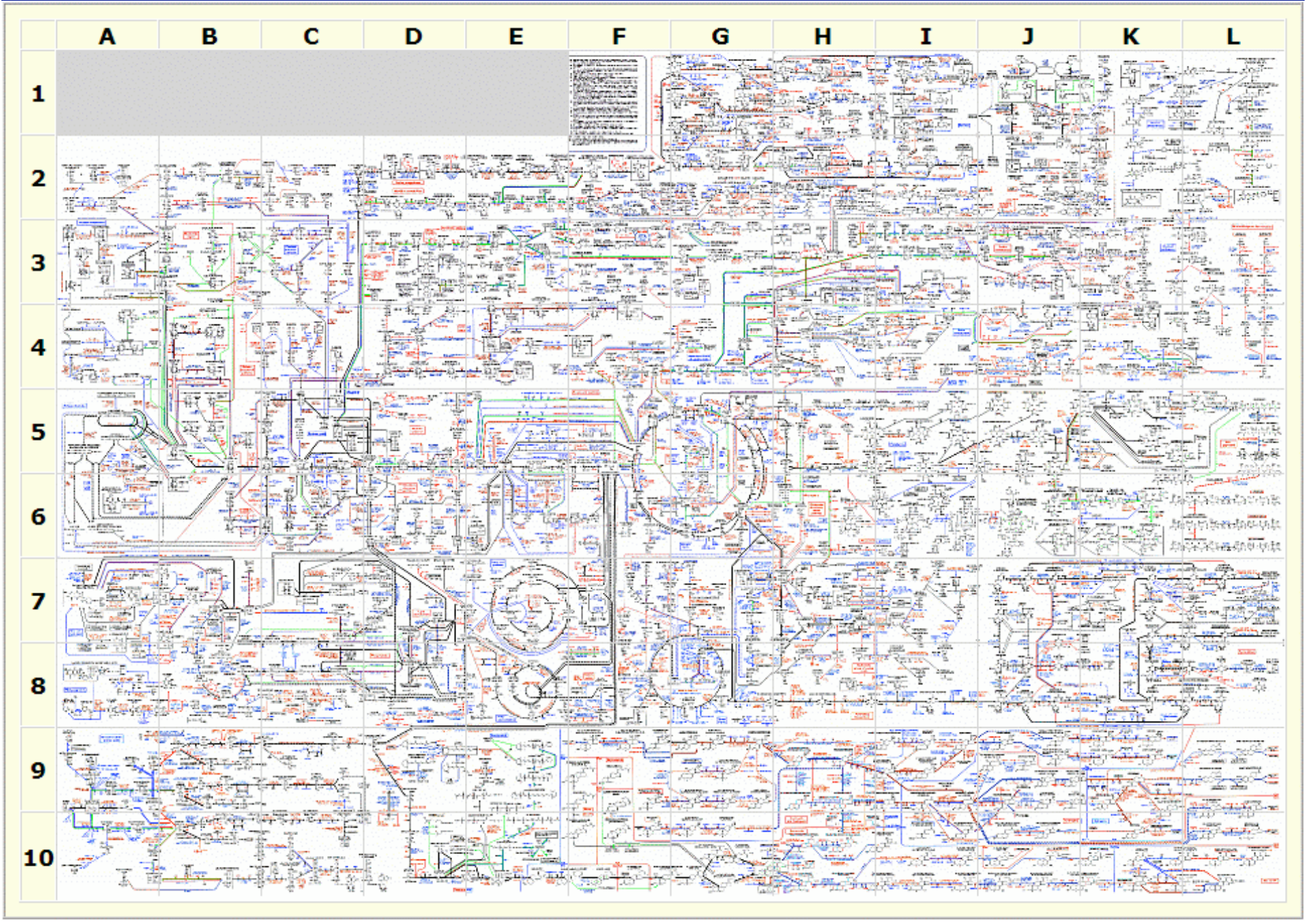
What is metabolomics?

Fusion of metabolism and genomics

= **metabolomics**

What is metabolism?

- **The biological chemical conversions**
  - Primary metabolism
    - Metabolism common to all life-forms (essential)
  - Secondary metabolism
    - Metabolism that is asymmetric in taxonomic distribution
      - Mostly non-essential
      - "Spice" of life
  - All biological chemical reactions are catalyzed by enzymes (gene encoded proteins)
  - Metabolism is conferred by the genome (developmental program) but in the context of environmental stimuli


A
B
C
D
E
F
G
H
I
J
K
L
1
2
3
4
5
6
7
8
9
10

What is metabolomics?

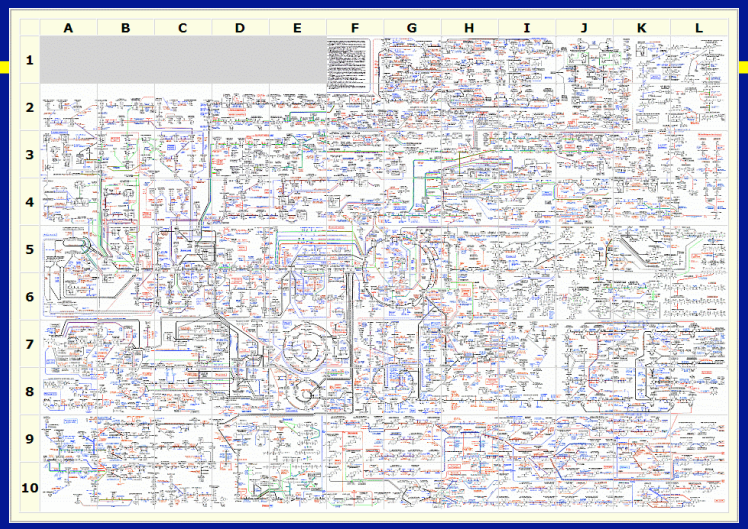
- **Goal #1: Metabolite profiling**
  - Chemically define a tissue at any time or place in development
  - Determine the chemical composition of the tissue in terms of "small" biomolecules
  - Identify "all" metabolites in a biological sample
- **Goal #2: The study of metabolism in the context of the entire genome**
  - Intellectually think of metabolism NOT in isolated blocks (pathways), but in the context of the entire genetic, developmental, physiological and environmental potential of the organism
  - Metabolic networks
    - Structure of network and regulation of the network

What does that mean?

Consider seeds

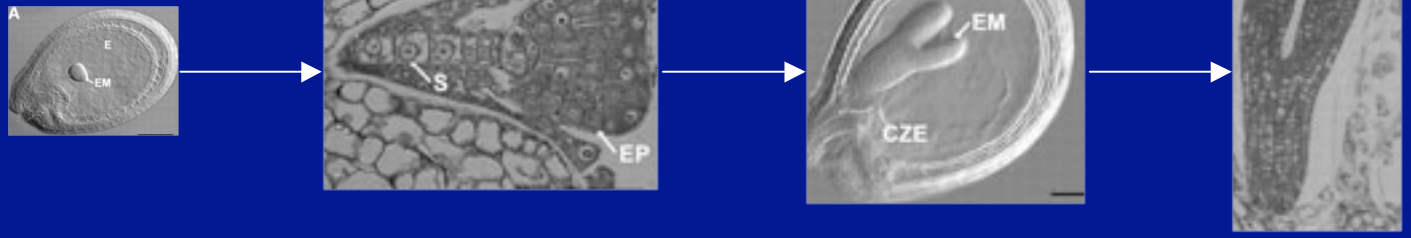
Inputs

sucrose
+
amino acids



Outputs

starch
+
protein
+
oil



Why is metabolomics important?

Metabolomics provides the chemical basis for phenotypes

Why is this important?

1. In the context of a fully defined genome, there is a need to define the biochemical and physiological functions of "all genes".

This is primarily being conducted by genetics

Metabolomics has a major role in interpreting the results of these experiments

Why is metabolomics important?

2. Defines biochemical differences associated with genetic mutants (forward genetics)

Provides functionality to undefined genes

3. Defines biochemical differences associated with natural variation

Provides access to allelic variation

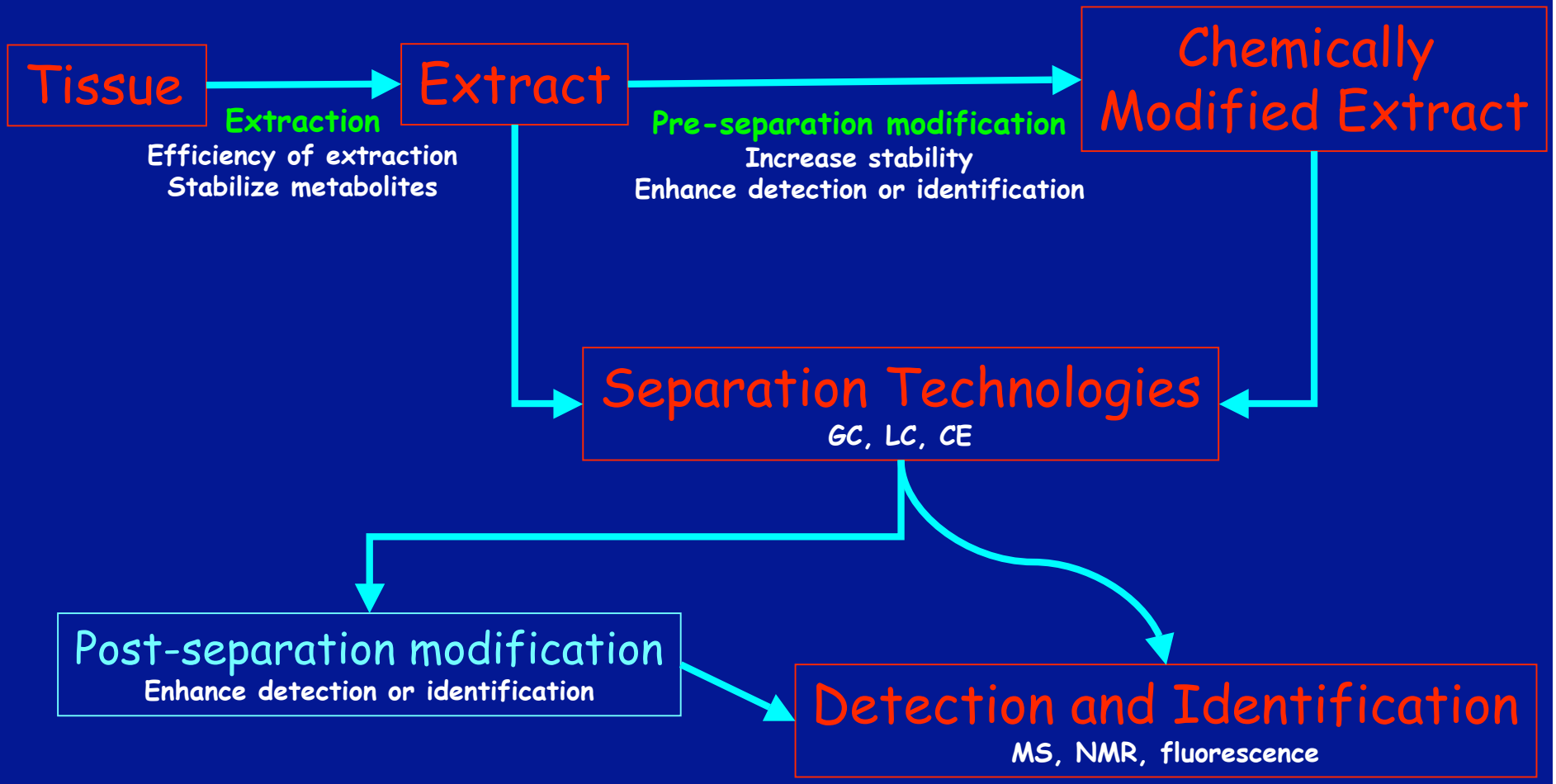
4. Defines biochemical difference associated with development or environmental stimuli

Provides insights into how metabolic networks are regulated

Metabolomics Technology

- Extraction
  - Efficiency
  - Stabilization
  - Chemical modification for separation and/or detection
- Separation
  - Chromatography
    - Gas chromatography (GC)
    - Liquid chromatography (LC)
  - Electrophoresis
- Identification/quantification
  - Mass spectrometry
  - Nuclear Magnetic Resonance (NMR)

Metabolite Profiling Strategies



Challenge of metabolomics

Challenge #1

What is the extent of the metabolome?

- Apply highly sensitive analytical technologies to identify and quantify metabolites in biological materials
- Identification primarily by mass-spectrometry or NMR
- How are "all metabolites" identified?
  - Targeted metabolomics vs non-targeted metabolomics

W.M. Keck Metabolomics Research Laboratory
Iowa State University

Chromatography

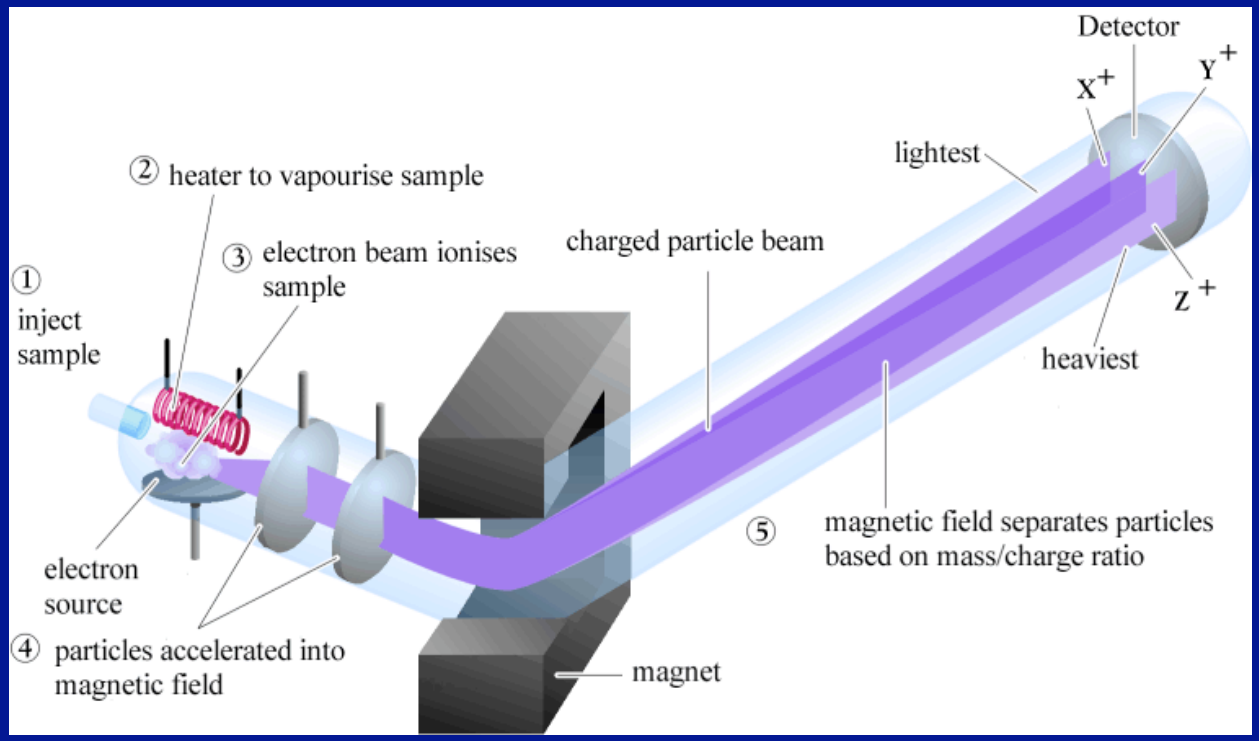
- Method of separation
- Stationary phase (analytes interact)
- Mobile phase (carrier of analytes)
- Gas-chromatography (GC)

Gas Chromatography (GC)

- http://ccl.northwestern.edu/netlogo/models/run.cgi?GasChromatography.789.585

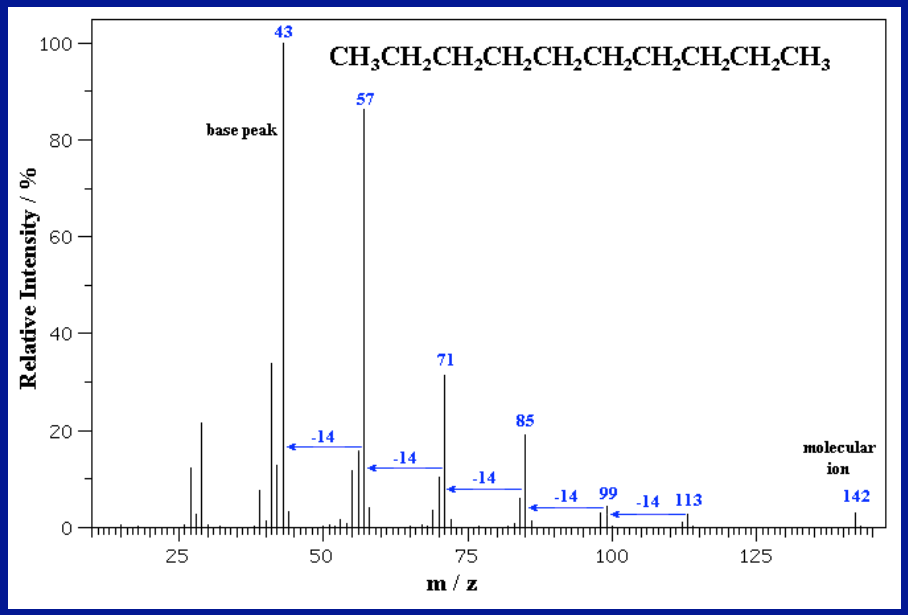
Detector to detect analytes

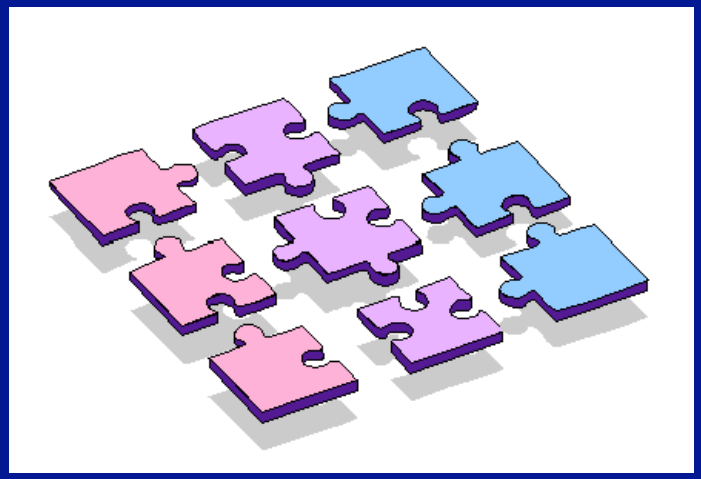
- Mass-spectroscopy
- Highly sensitive
- Provides structural information by fragmenting analytes and "weighing" the fragments



Mass spectroscopy outputs

- Chemical structure

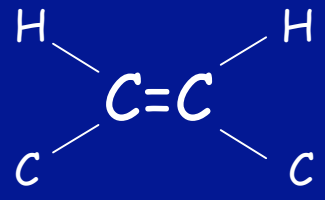


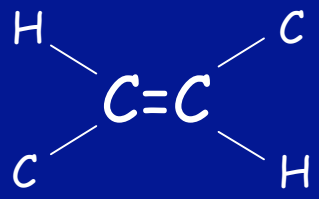


Analogous to jigsaw puzzle
Except instead of assessing shapes of fragments, you determine mass of fragments, and
All possible combinations are obtained

Mass spectroscopy challenge/limits

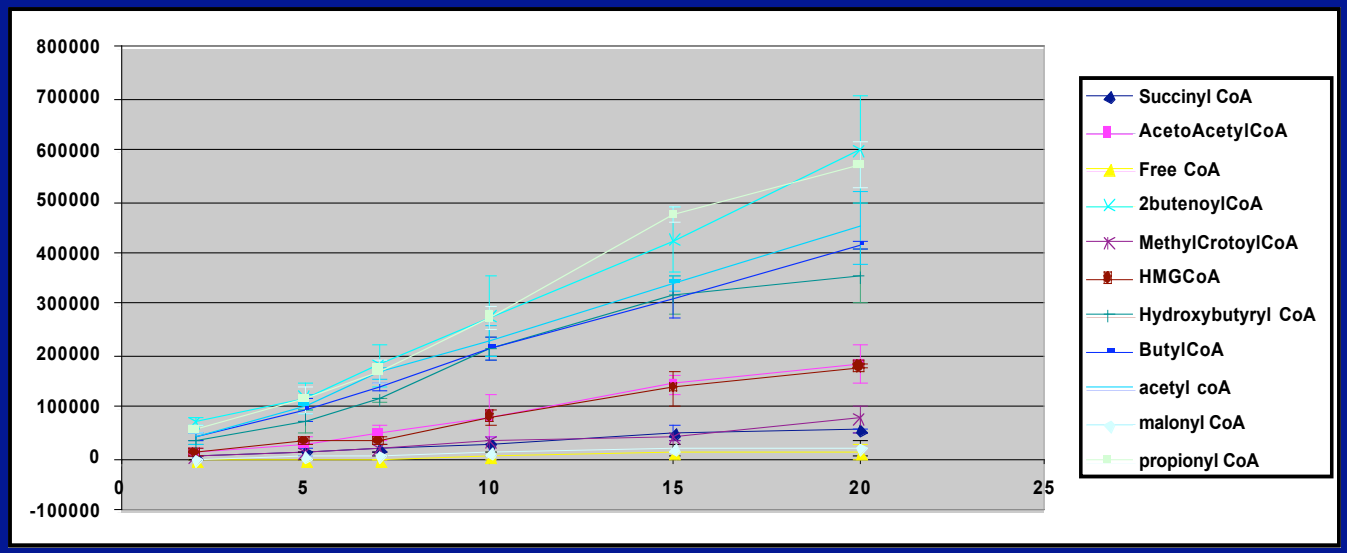
- Not all chemical structures can be determined by MS
- Limitation of instrument sensitivity
  - this is being overcome as more high-end MS instruments are being developed (TOF-MS)
- Limitation of chemistry - certain functional groups give same MS signature - but chemistry is different
  - e.g., isomers - cis/trans





Mass spectroscopy challenge/limits

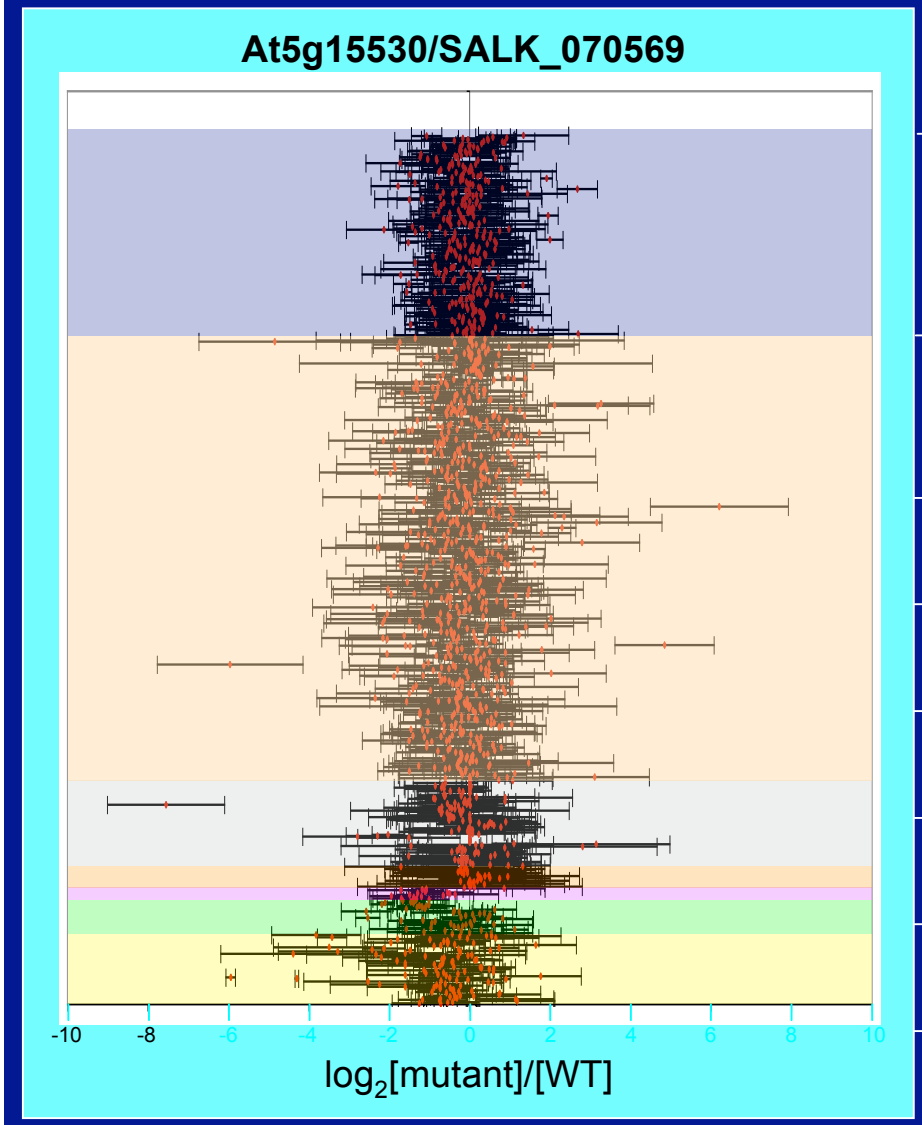
- Mass spectrometer response is different for each chemical
- Therefore, quantification is NOT absolute, unless you know the chemical and standardize the response of detector



From Suh-Yeon and Ann Perera

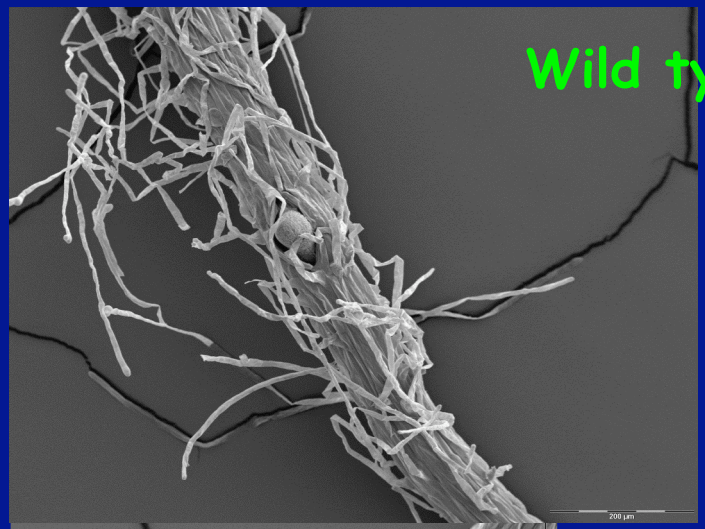
Solution:
Relative abundance
Ratio of abundance in two samples

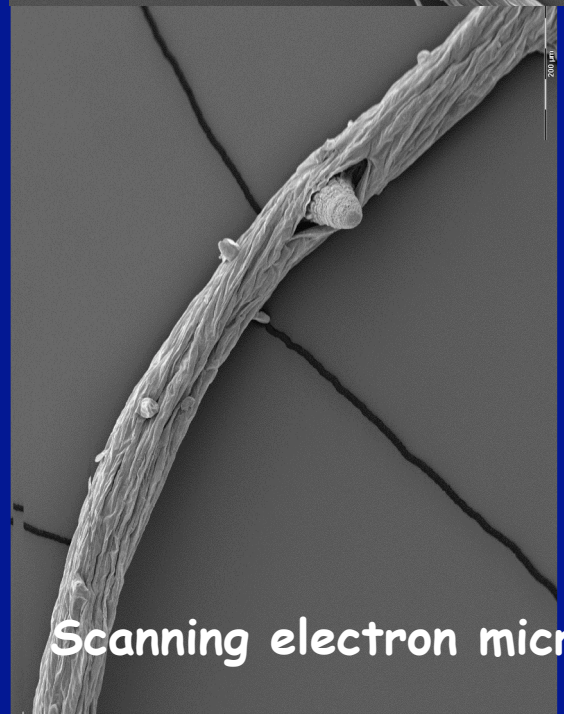
Metabolites

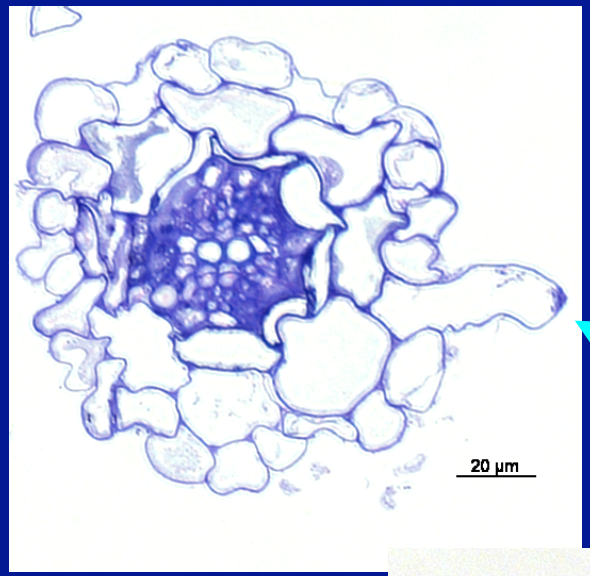


Analytical platform	Number of metabolites		
	Chemically defined	Chemically undefined	Total
Non-targeted GC-TOF (Fiehn)	102	197	299
Non-targeted UPLC-TOF (Sumner)	218	452	670
Lipidomics (Welti)	160	0	160
Isoprenoids (Lange)	15	7	22
Ceramides	20	0	20
Amines	25	7	32
Fatty acids	50	103	153
Total	590	766	1356

At4g29540 mutants lack root-hairs



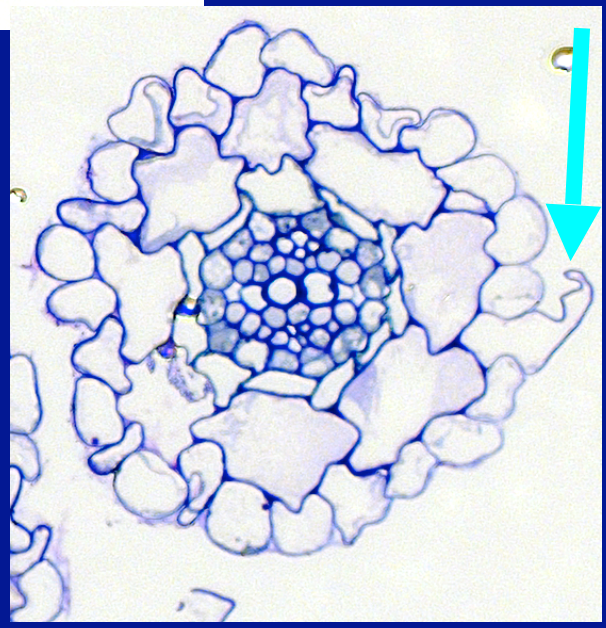




Root cross-section
light microscopy

Root-hair

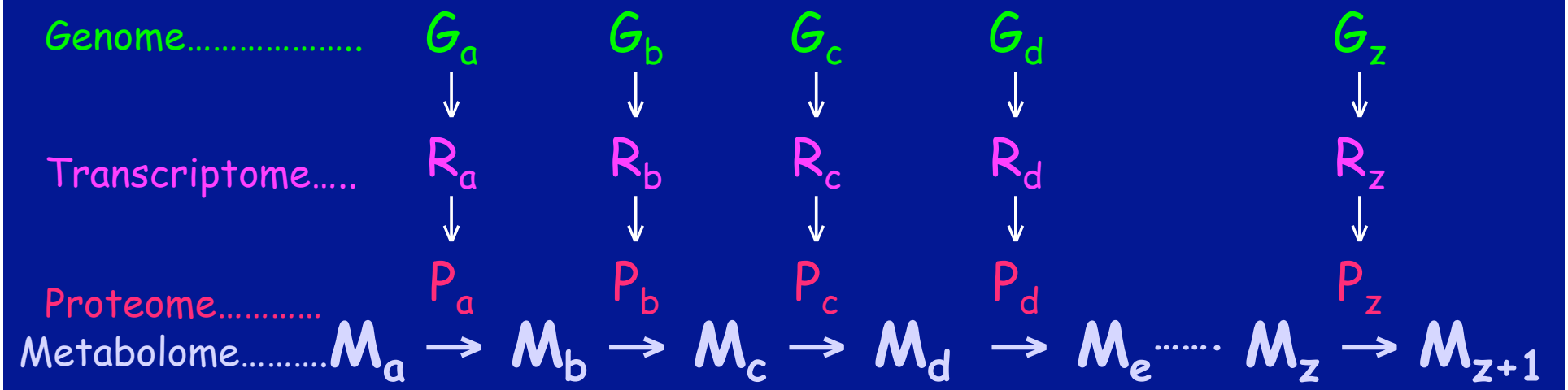
mutant



How to use metabolomics data?

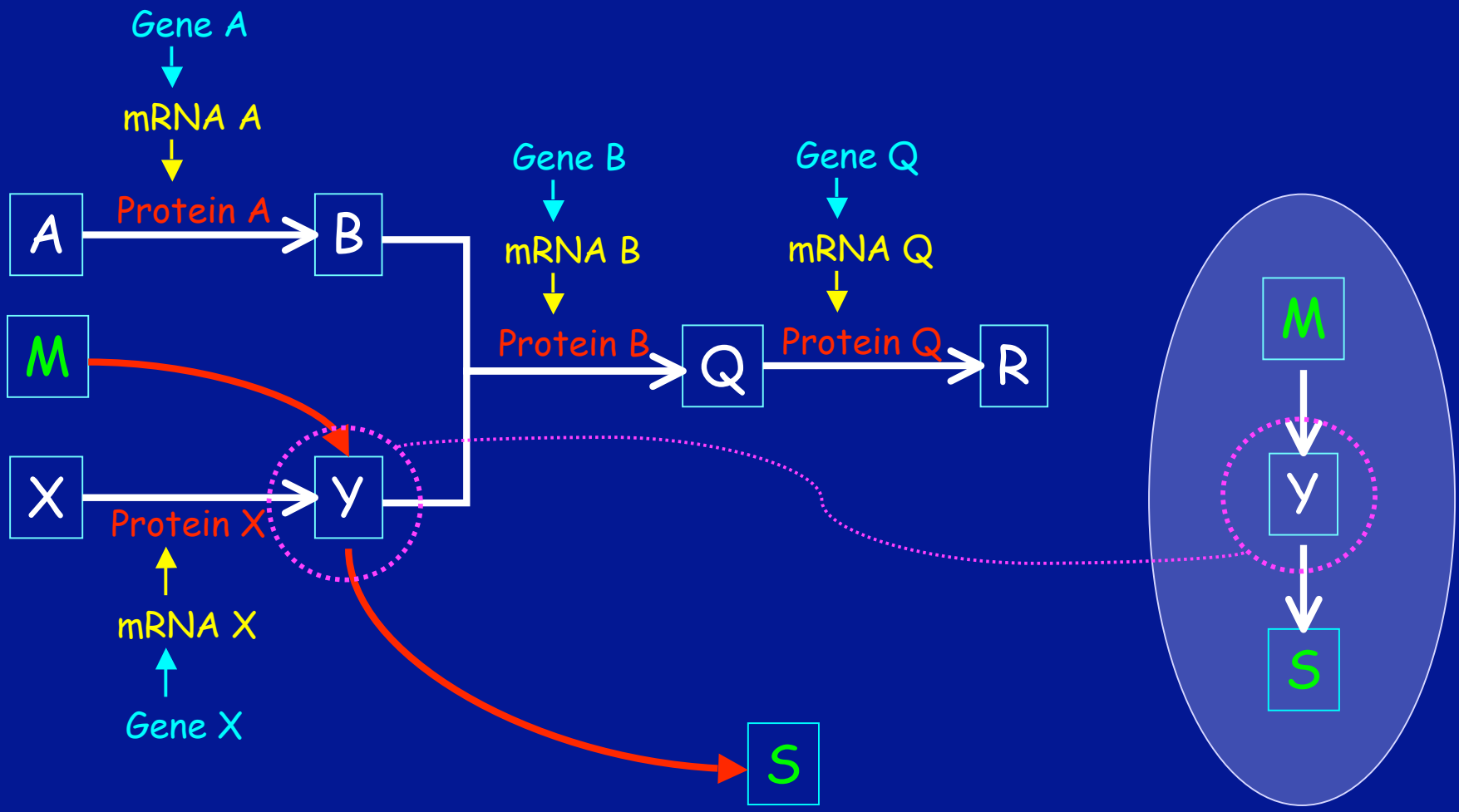
Challenge of metabolomics

Genome defines transcriptome and proteome BUT NOT the metabolome



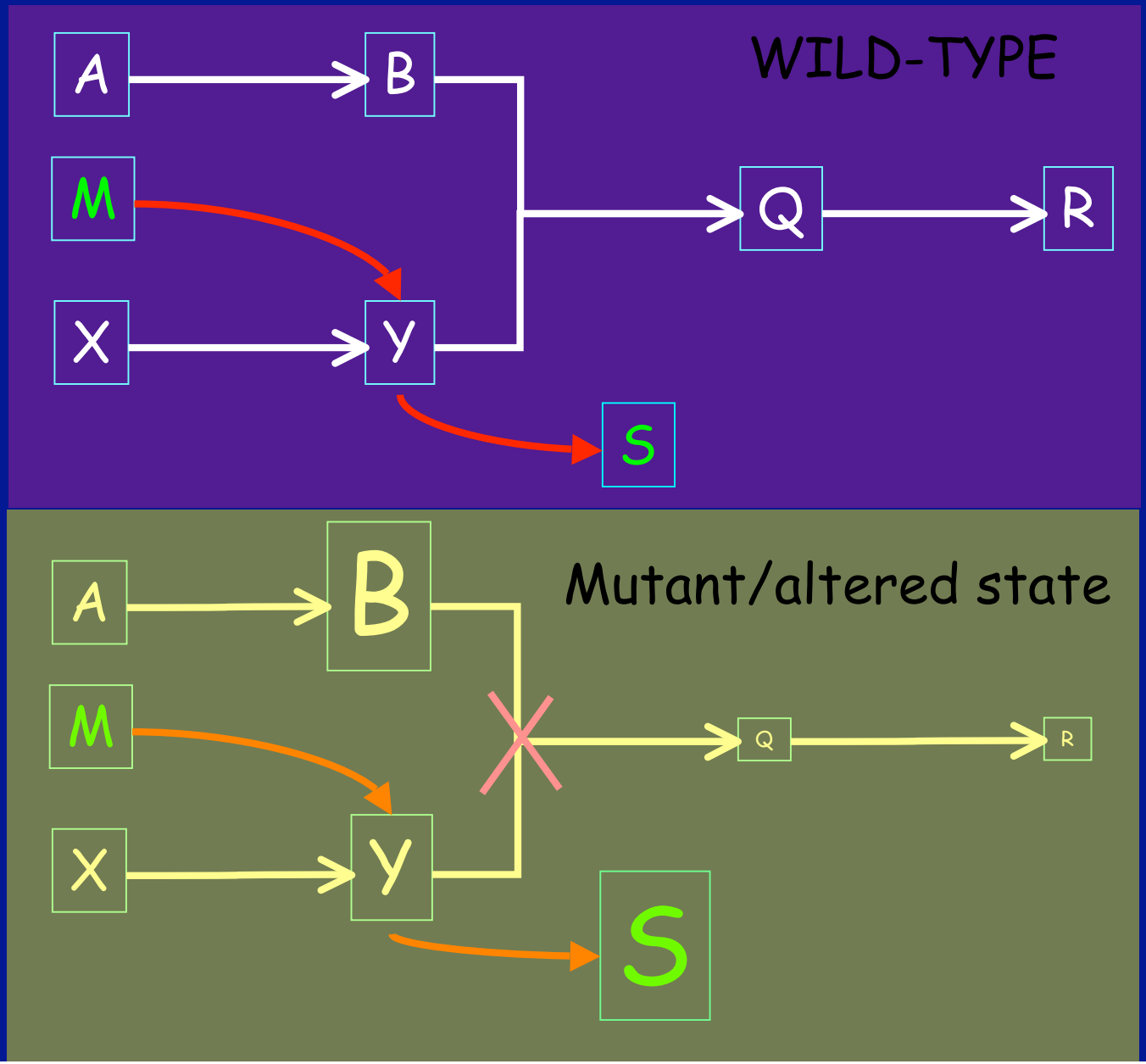
Yet, the metabolome represents the ultimate level at which the genome is expressed

Complex Metabolic Networks

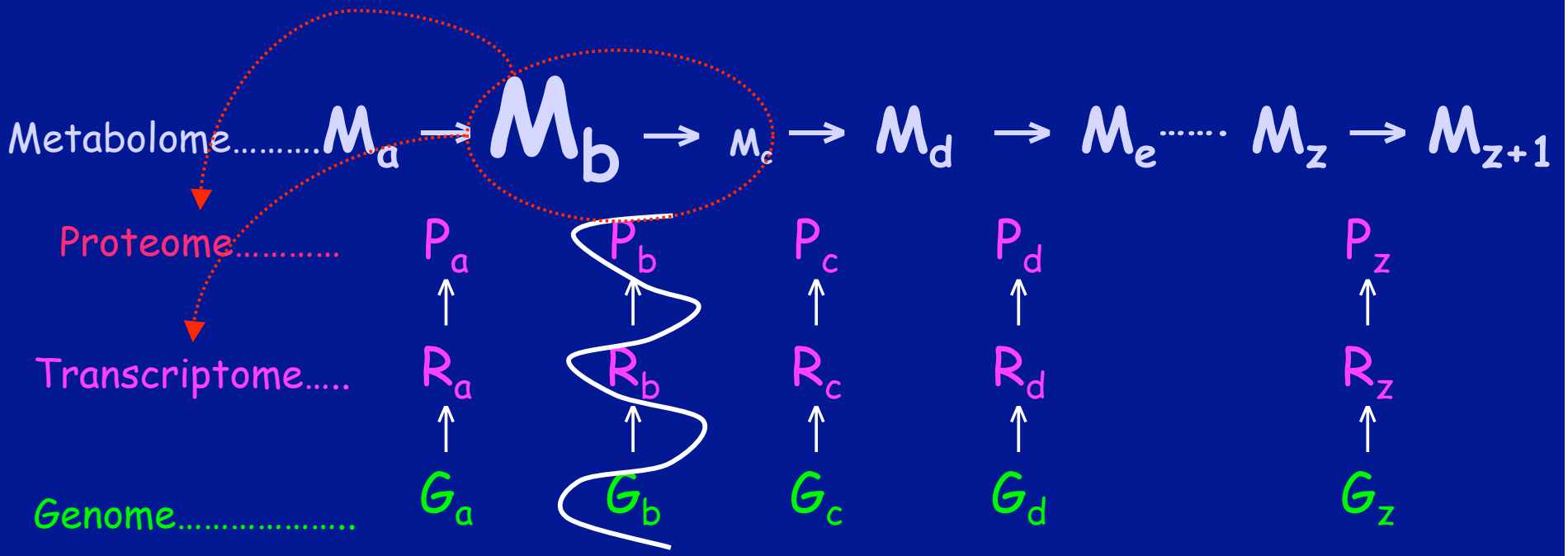


A single metabolite may be involved in multiple pathways

The role of metabolomics in functional genomics



Metabolomics can potentially reveals the primary breach of mutants



Effects on proteome or transcriptome is secondary to effect on metabolome

Omics data, like metabolomics is static data

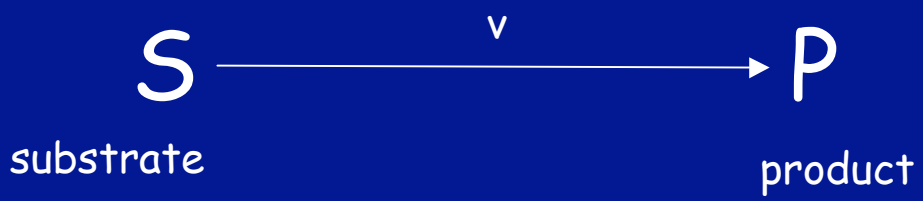
But biological systems are dynamic

Thus, what is the relationship between metabolomics (static data) and dynamic data

Metabolic FLUX (dynamic data)

Metabolic flux

The rate of chemical conversions



Enzyme catalyzed reaction

v = rate of conversion of S to P (mole/unit time)

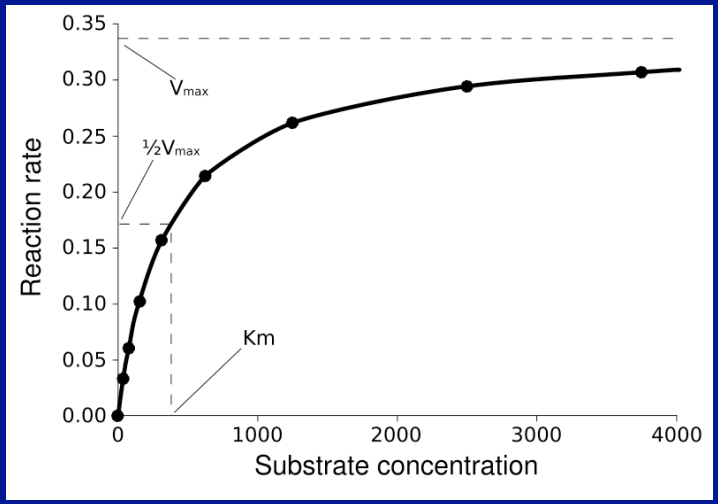
What is v dependent on?

ENZYMOLOGY

Enzymology

- v (enzyme velocity) is dependent on:
  - [S]
  - Enzyme concentration (direct correlation)
  - Enzyme property (K_m and V_{max})
- Simple math of system is Michaelis-Menten eqn. (hyperbolic curve)

$$v = \frac{V_{\max} \times [S]}{K_m + [S]}$$



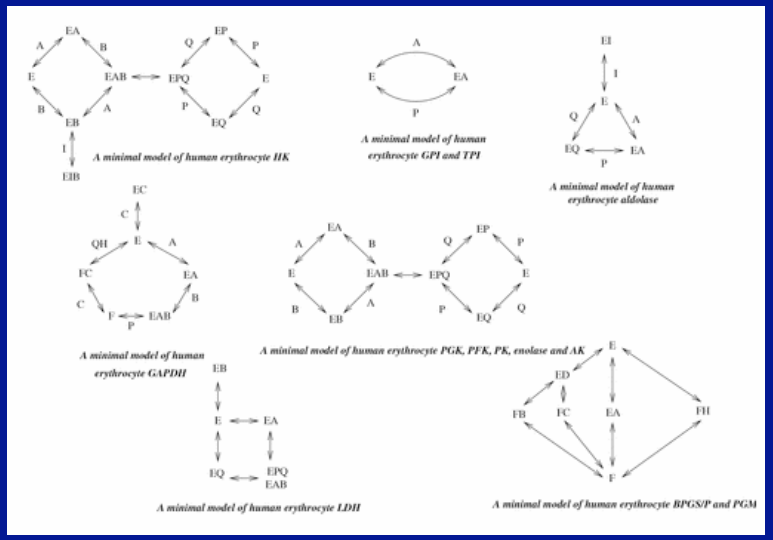
Enzymology - complexities

- Michaelis-Menten is simplification (only one substrate is considered
- Most enzymes have multiple substrates and products (math is much more complicated)

e.g., acetyl-CoA carboxylase

acetyl-CoA + Mg.ATP + HCO_3^- → malonyl-CoA + ADP + P_i

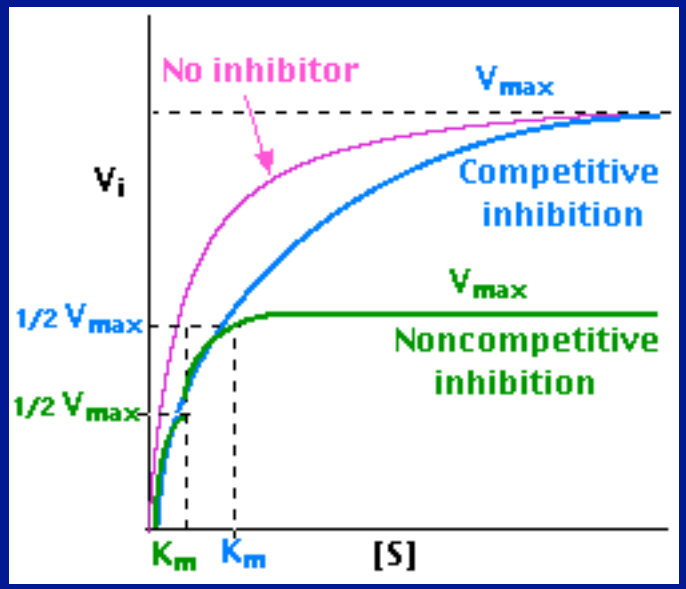
$$v = \frac{V_{max} \times [S]}{K_m + [S]}$$

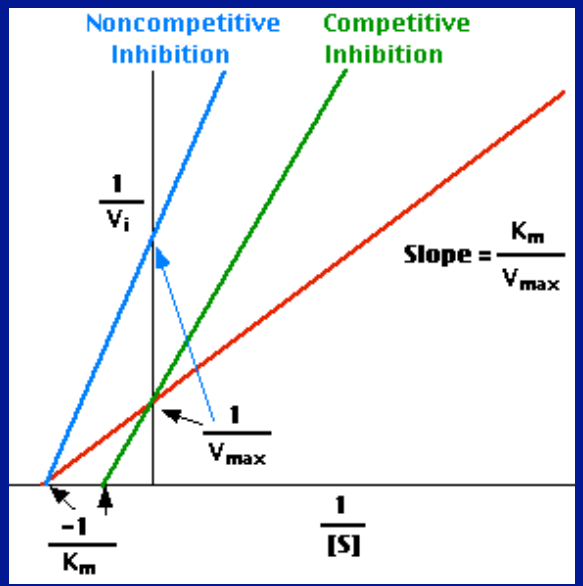


Enzymology - complexities

- Modulators of enzyme activity
- Inhibitors and activators
- Can affect binding of substrate or rate of catalysis - hyperbolic kinetics

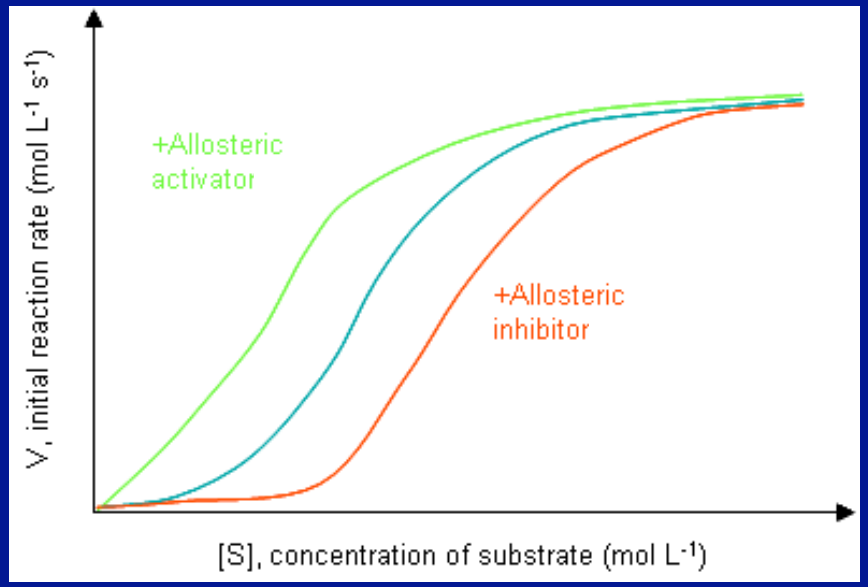
$$\frac{v = V_{\max} \times [S]}{K_m + [S]}$$





Enzymology - complexities

- Allosteric modulators of enzyme activity
- Inhibitors and activators
- These bind to enzyme NOT at active site, but affect catalysis - sigmoidal kinetics (cf., hyperbolic kinetics)



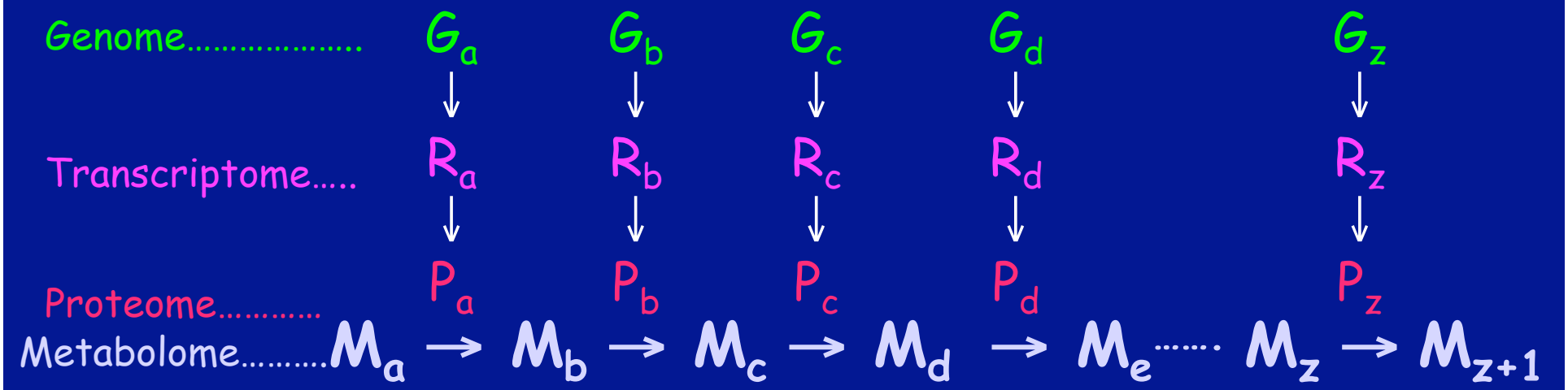
Enzymology - complexities

- Enzyme kinetic data is generated *in vitro*
  - Purified system
  - From initial rate conditions
    - [enzyme] << [S], and
    - [P] = 0

How do these data extrapolate to the *in vivo* situation?

This math applies to all arrows in this scheme

These are all enzyme catalyzed reactions
RNA polymerase (Gene → mRNA)
Ribosomes (mRNA → Protein)



The complexities and limitations we've discussed apply, and are the challenge of systems biology

Thank you