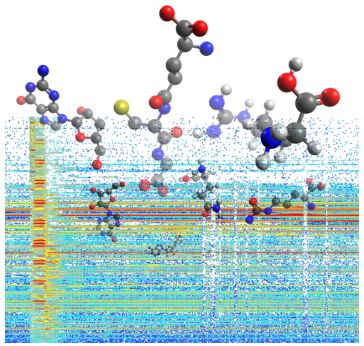
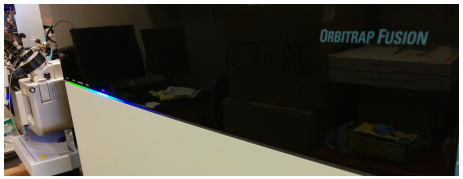
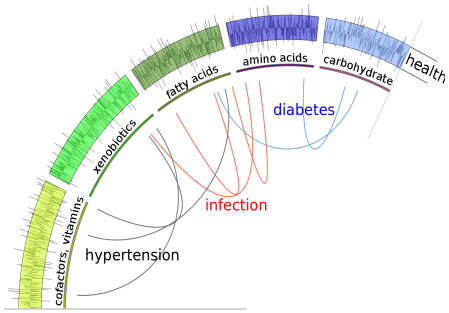




High-resolution Metabolomics for Systems biology and medicine

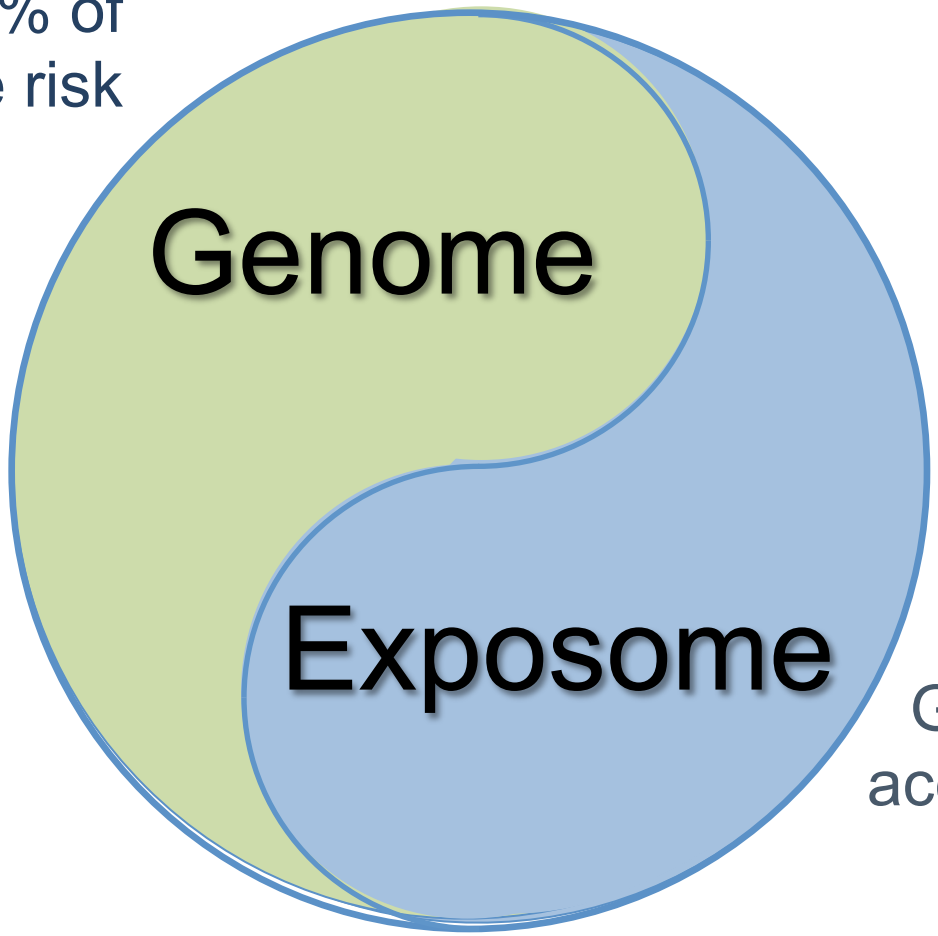
Shuzhao Li, Ph.D and Dean P. Jones, Ph.D
Department of Medicine, Emory University

June 18, 2015

EMORY
SCHOOL OF
MEDICINE

Disease risk is a combination of genomics and exposures

10-20% of disease risk



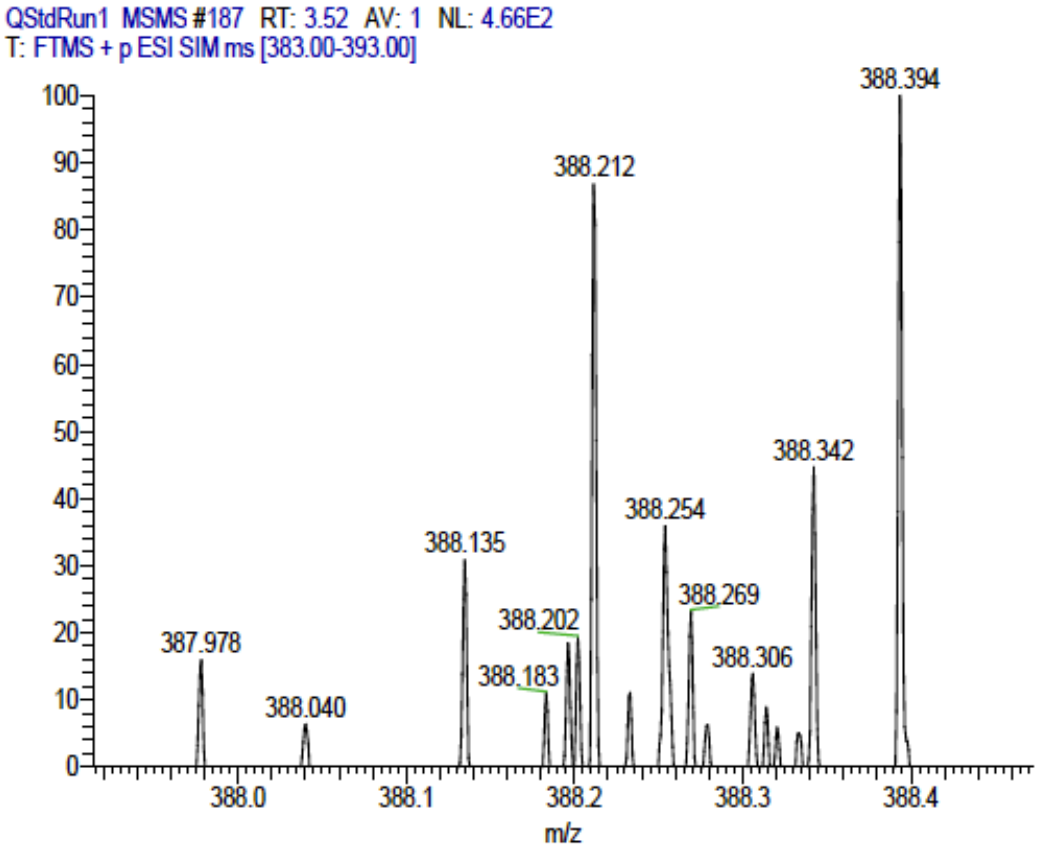
GxE interactions account for 80-90% of disease risk

Jones et al 2012 Annu Rev Nutr 32:183-202

Outline

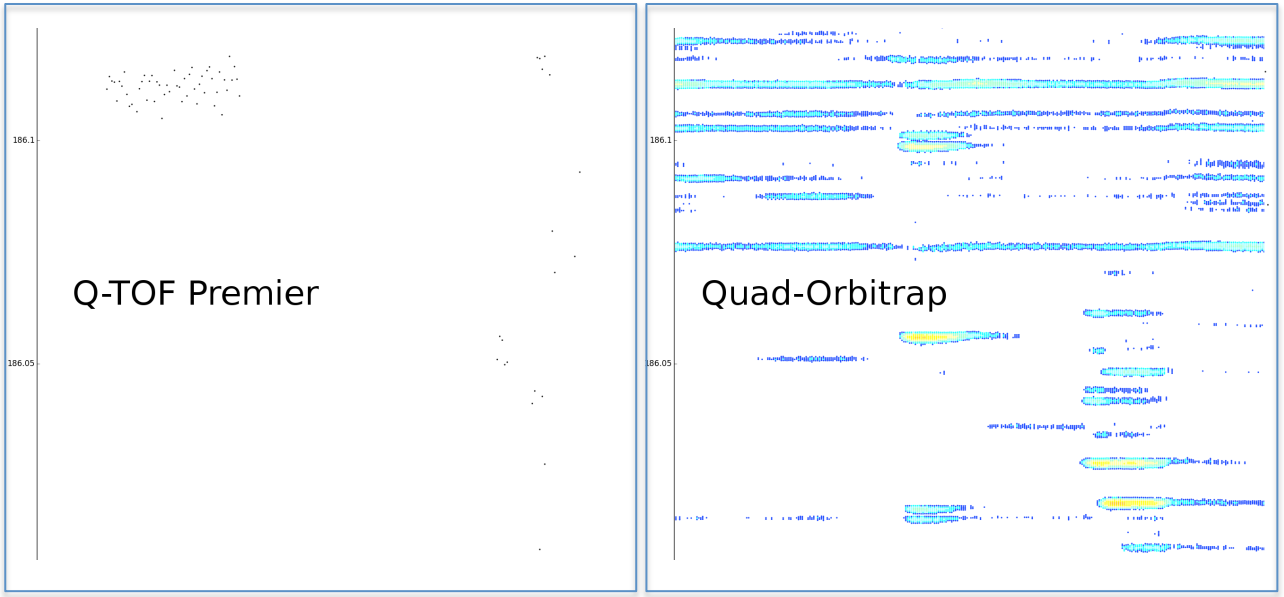
- The development of high-resolution metabolomics
- Metabolic networks and *mummichog*
- Broader impacts of metabolomics

Resolution and mass accuracy of high-resolution mass spectrometers allow unprecedented detection of low-abundance ions in human plasma and other complex mixtures



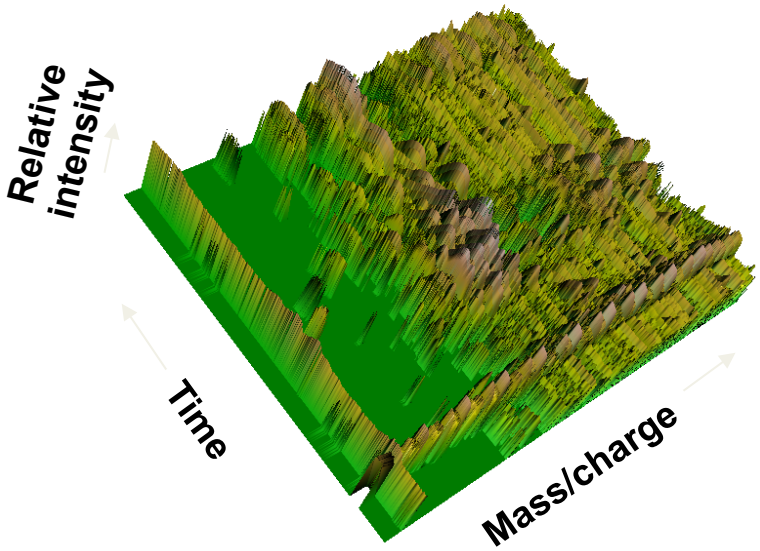
Uppal et al BMC Bioinformatics 2013

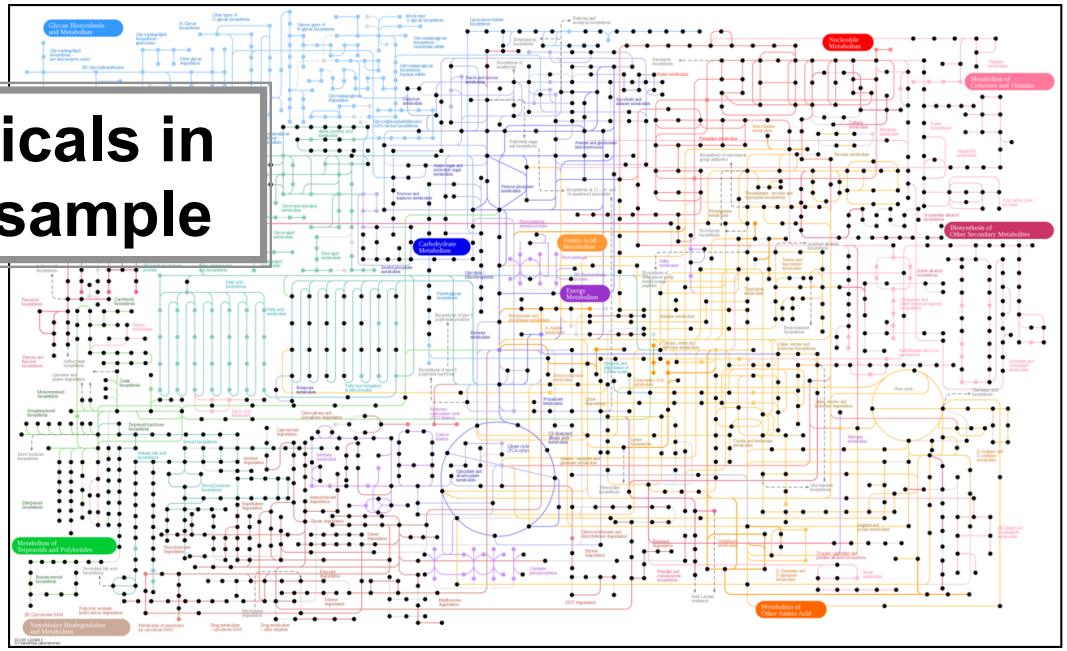
Resolution + sensitivity



High-resolution metabolomics developed at Emory measures individual biochemistry with resolution approaching that for genomics

Measure 20,000 chemicals in an individual plasma sample

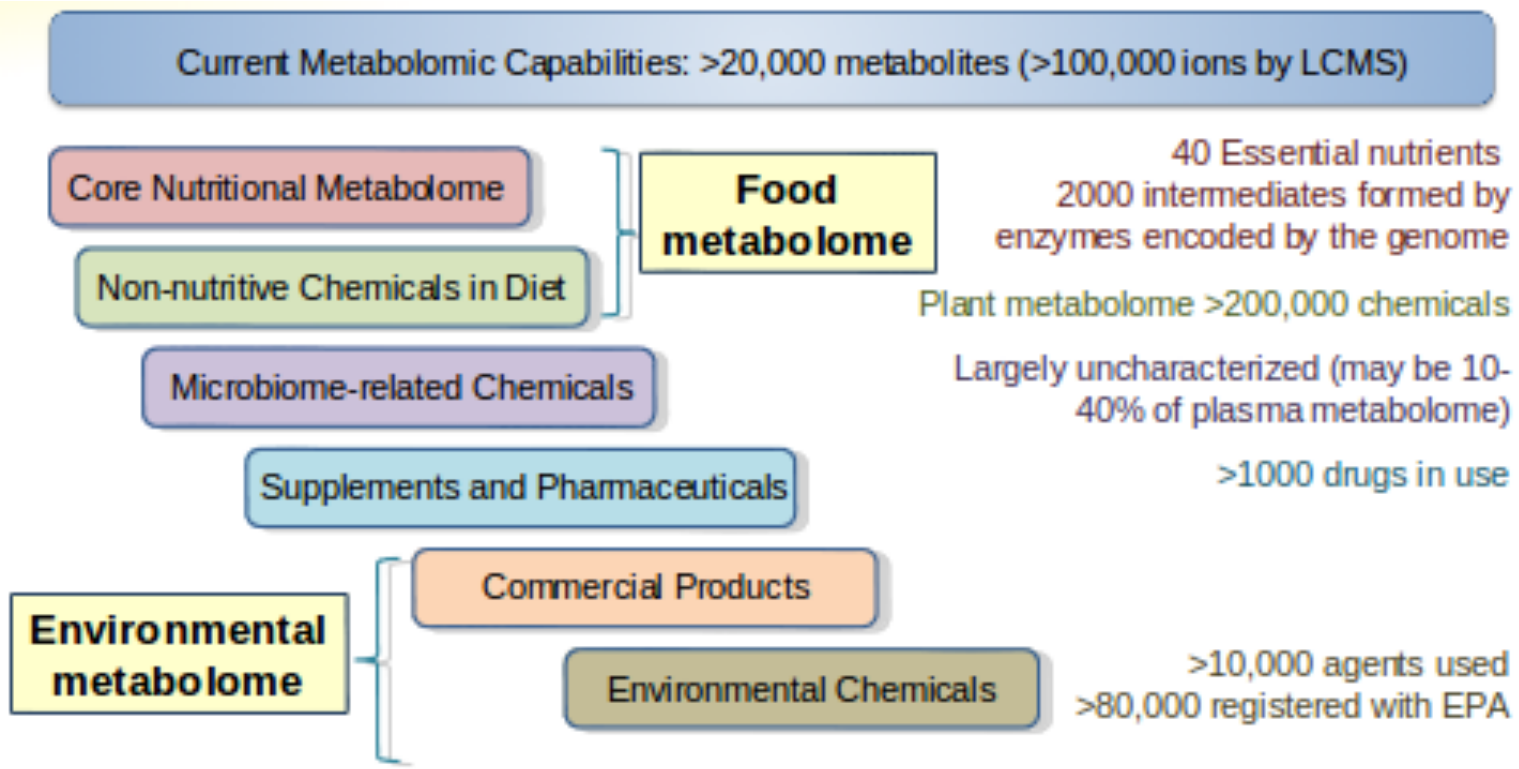




Technologies appear to be good enough to measure 1 million chemicals

Jones et al 2012 Annu Rev Nutr 32:183-202

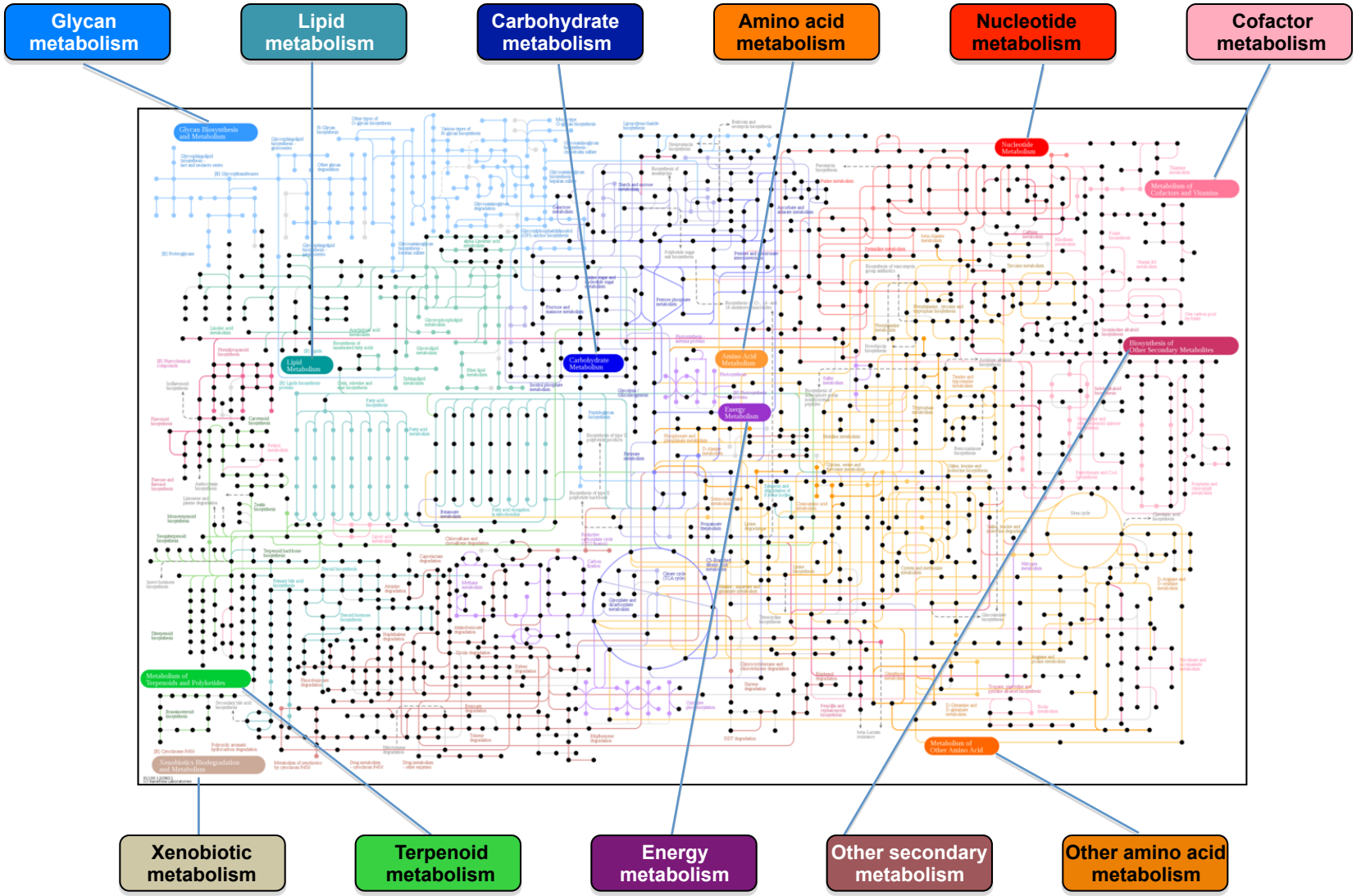
What's in a metabolome?



*Metabolome refers to chemicals associated with life

Jones, Park Ziegler Annu Rev Nutr 2012

Accurate mass *m/z* match more than half of metabolites in KEGG human metabolic pathways (shown in black); 146 of 154 pathways are represented



Jones 2012 Annu Rev Nutr **32**:183-202

xMSanalyzer data for 174 serum samples (K Uppal et al BMC:Bioinformatics 2013)

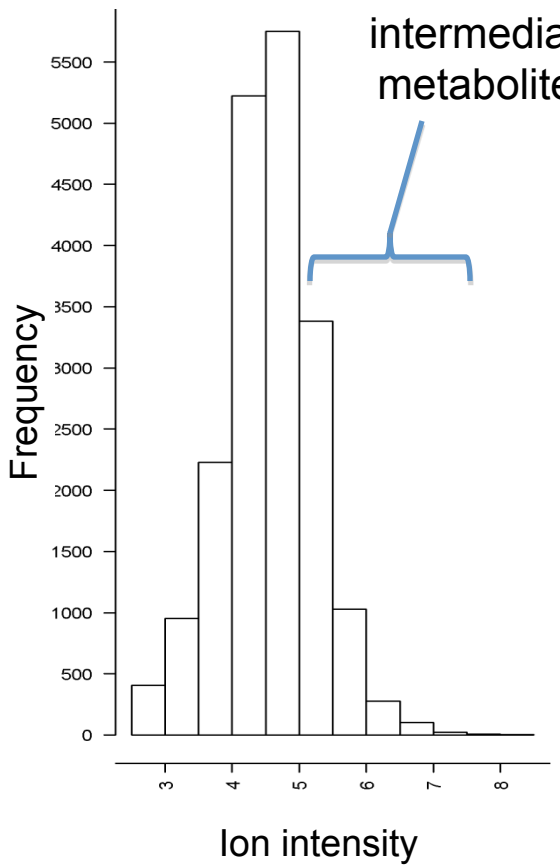
Improved data extraction over most approaches: 34,768 ions, triplicate analyses

Summary for C18: 19,383 ions Range of detection over 5 orders of magnitude of intensity

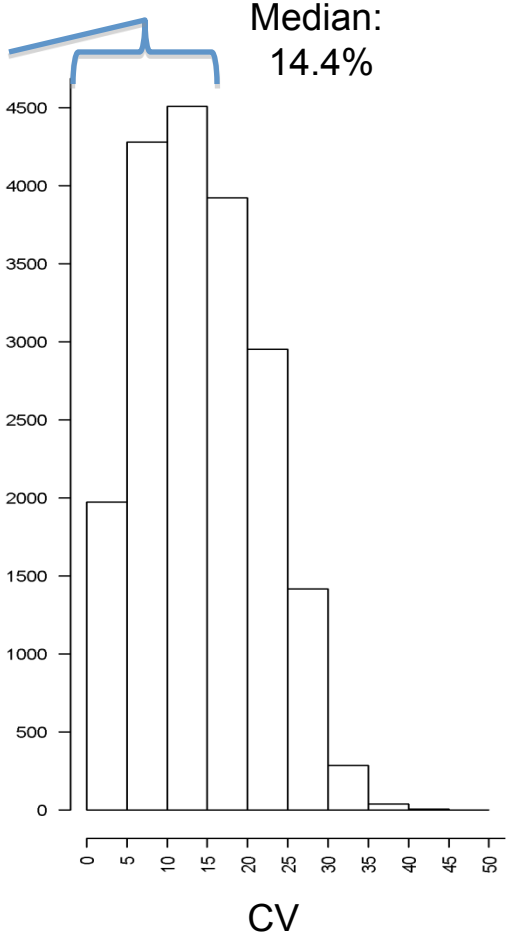
With triplicate analyses, CV is obtained for each metabolite in each sample:
6,247 had median CV < 10% Mean intensity of ions with CV <10%: 3.0 x 10^5

Intensity
Mean: 1.2 x 10^5
Median: 2.0 x 10^4

Enriched in intermediary metabolites

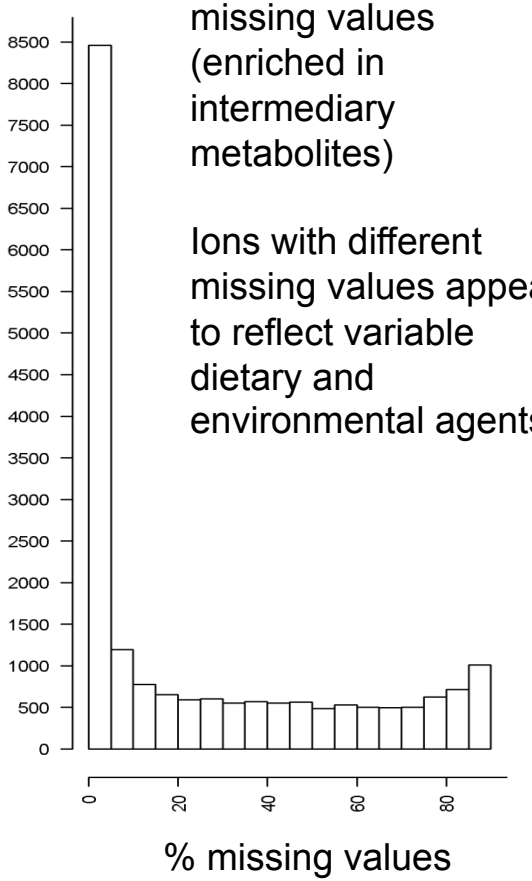


CV
Median:
14.4%



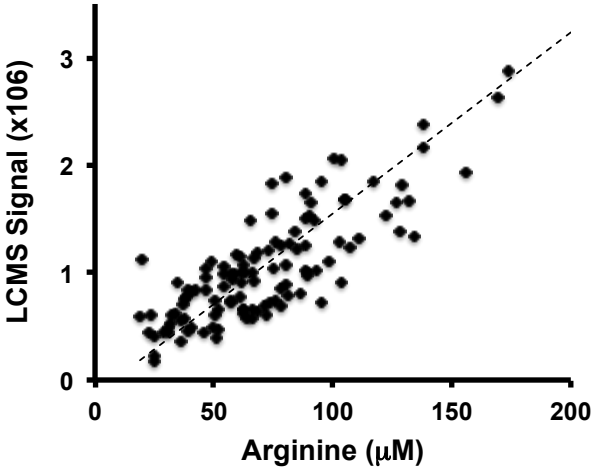
Missing values
>8000 had <5% missing values (enriched in intermediary metabolites)

Ions with different missing values appear to reflect variable dietary and environmental agents

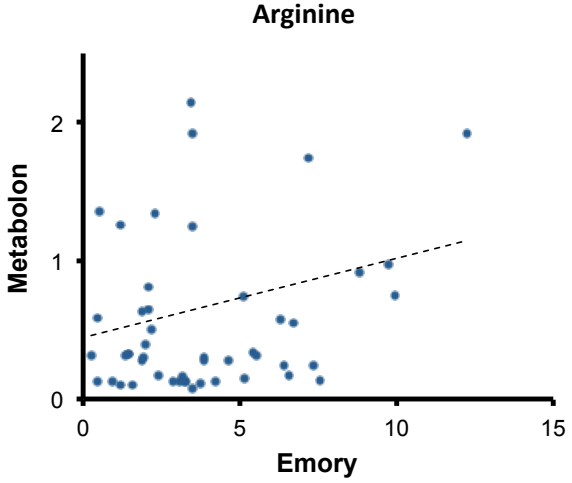


Cross-platform validation

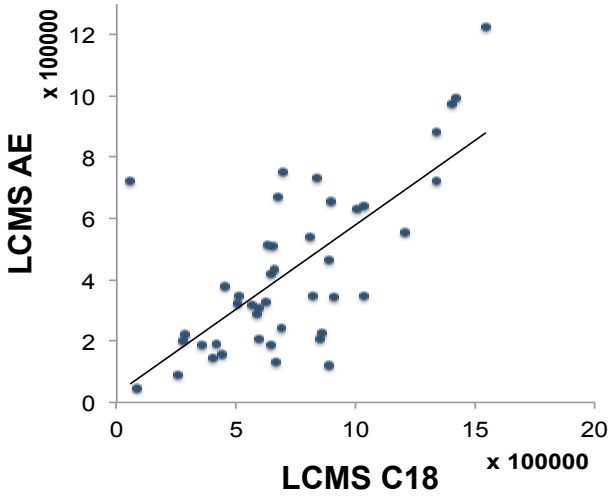
Comparison C18-LCMS to automated amino acid analysis



Comparison C18-LCMS to Metabolon GCMS



Comparison C18-LCMS to Anion Exchange LCMS



HRM metabolite quantification in 30 orphan samples

Identity	Mean ± SD (µM)	HMDB (µM)
Arginine	148 ± 39	60 to 140
Glycine	280 ± 62	212-329
Histidine	100 ± 12	75 to 143
Ornithine	83 ± 28	54 to 94
Phenylalanine	131 ± 18	48 to 88
Threonine	136 ± 22	102 to 260
Tryptophan	56 ± 7	44 to 78
Tyrosine	84 ± 23	54 to 143
Glucose	4310 ± 1153	3900 to 6100
Kynurenine	2.0 ± 0.4	1.4 to 2.4
Carnitine	52 ± 9	30 to 57
Creatinine	93 ± 13	59 to 109
Creatine	16 ± 8	8.4 to 65

Alternate Workflows

Targeted Metabolomics

Select analytic target to test hypothesis

↓

Select and test analytic method

↓

Perform power calculation; design experiment

↓

Conduct experiment

↓

Analyze samples and perform statistical analysis

High-resolution metabolomics

Pose scientific question (with or without hypothesis)

↓

Select relevant samples

↓

Analyze samples by high-resolution MS with advanced data extraction algorithms

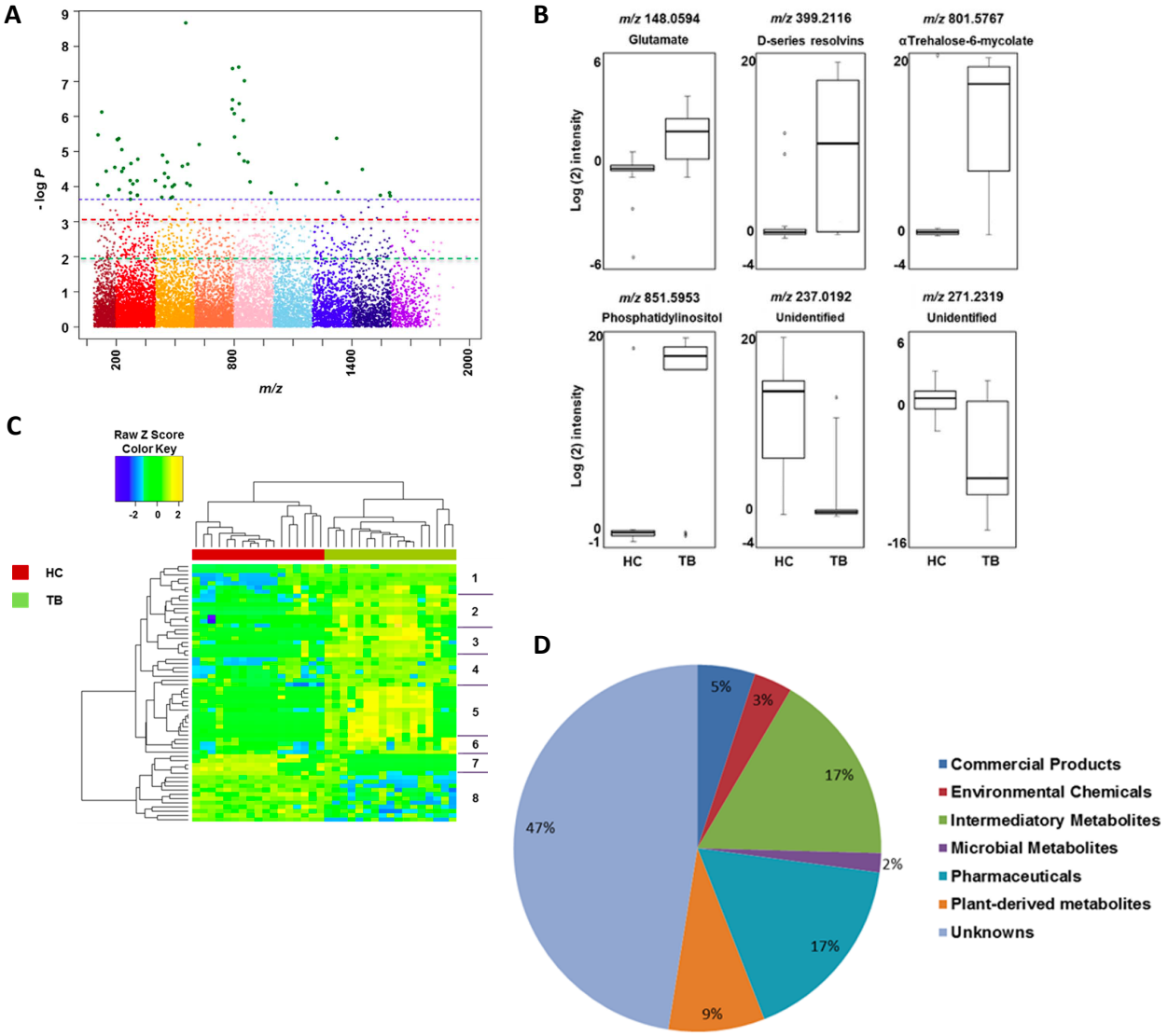
↓

Use bioinformatic methods and database tools to obtain significant metabolites and pathways

↓

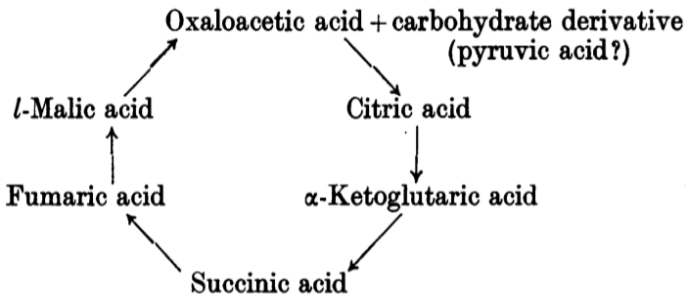
Perform MS/MS and co-elution studies to verify metabolites

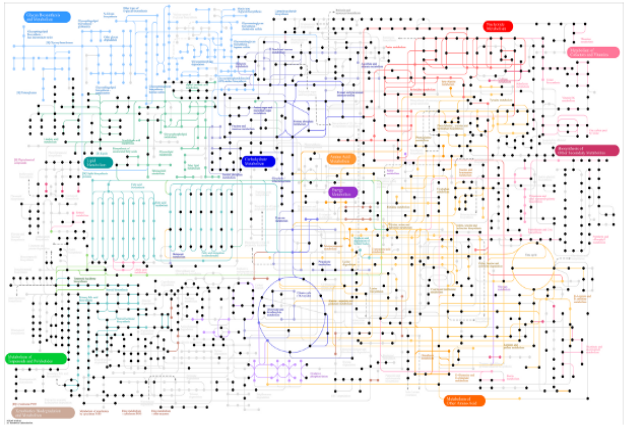
Pilot study of pulmonary Tuberculosis



Frediani, Jennifer K., et al. "Plasma Metabolomics in Human Pulmonary Tuberculosis Disease: A Pilot Study." *PloS one* 9.10 (2014): e108854.

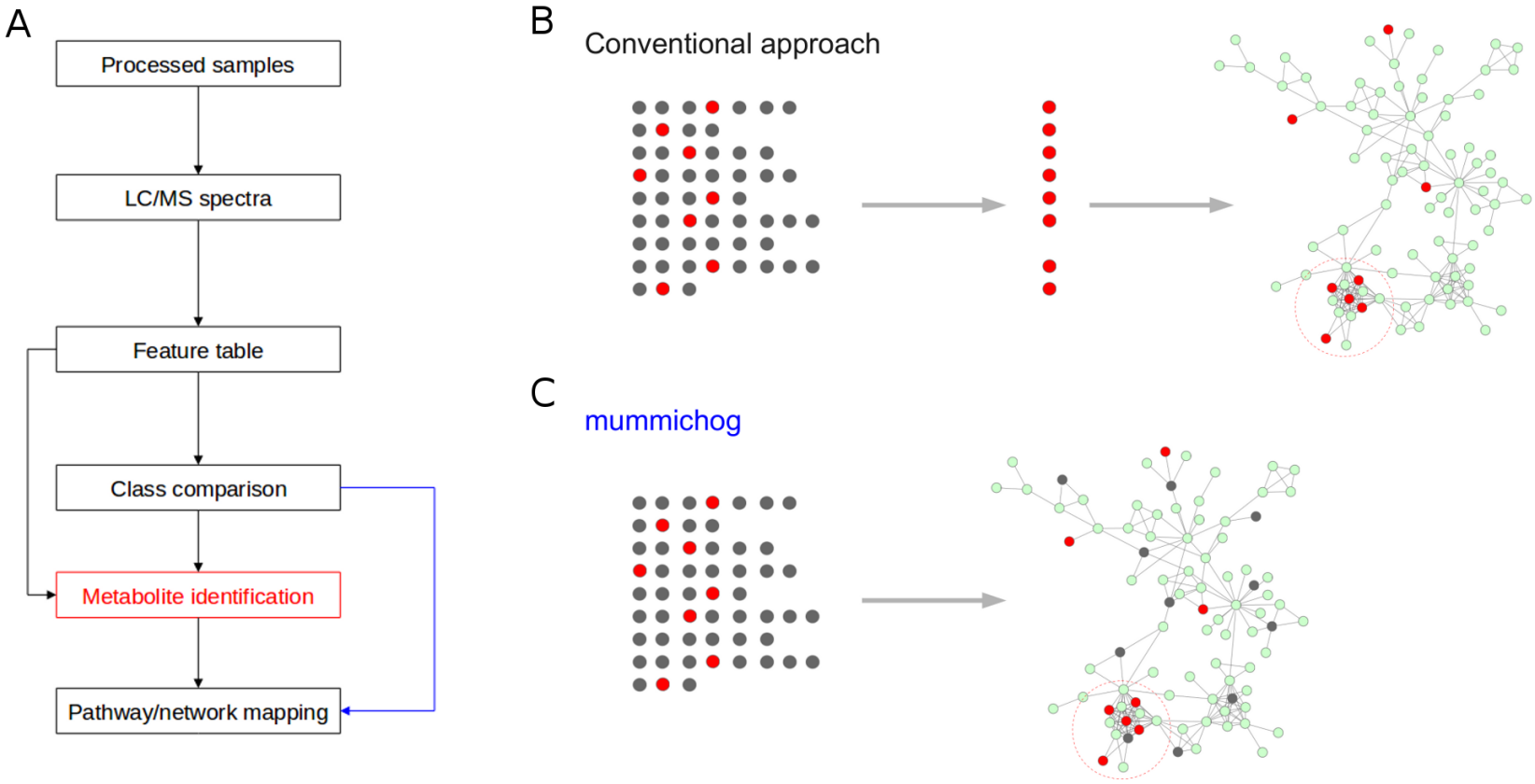
Connecting HRM with metabolic pathways





Krebs et al. 1938. Biochem Journal. 32:113

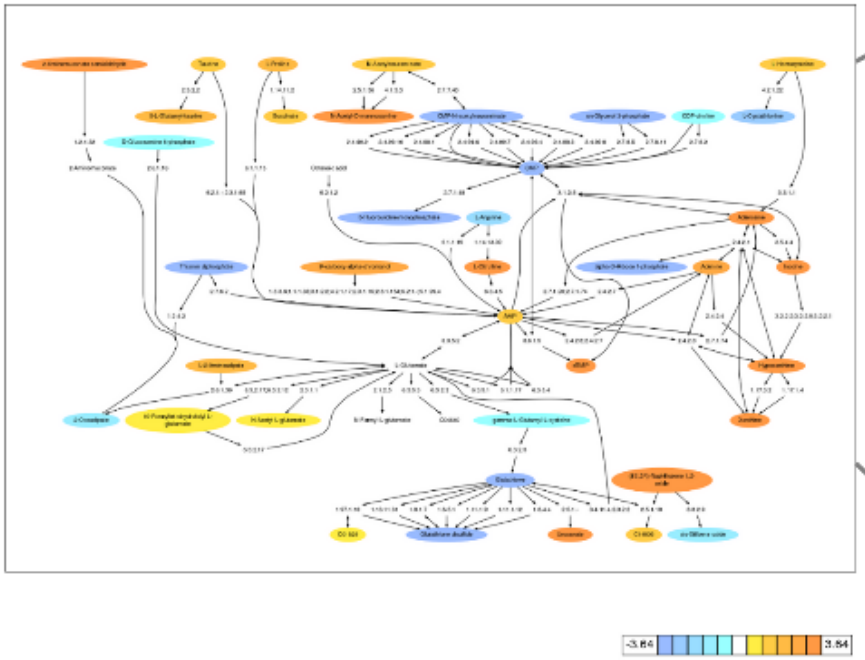
Mummichog interpretation of metabolomics data



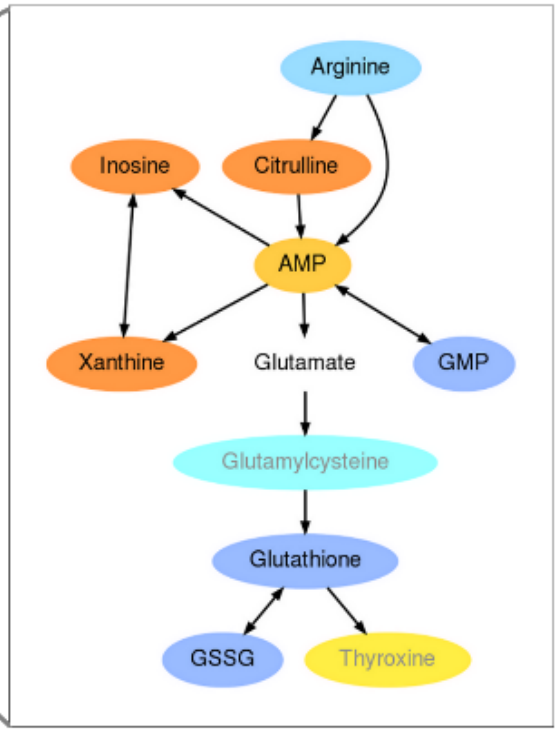
Li et al. 2013. PLoS Computational Biology. 9:e10031323

Metabolite network after viral activation

A

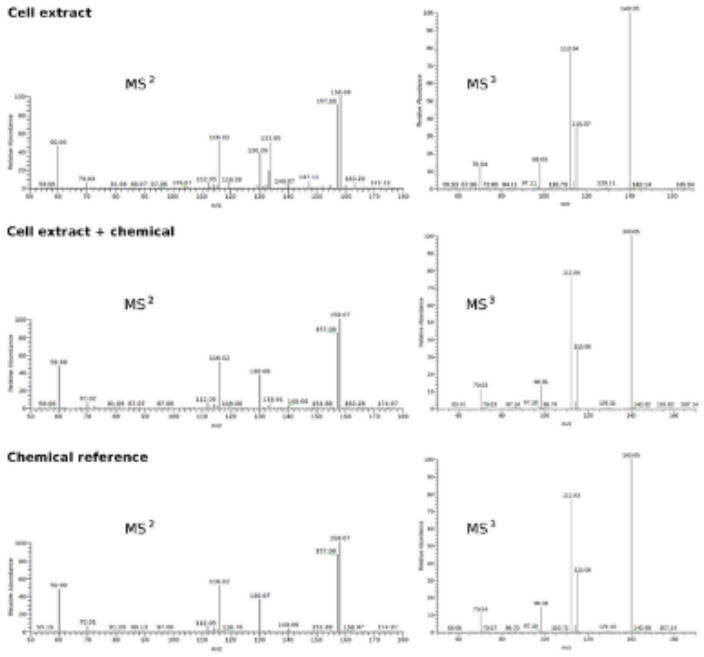


B

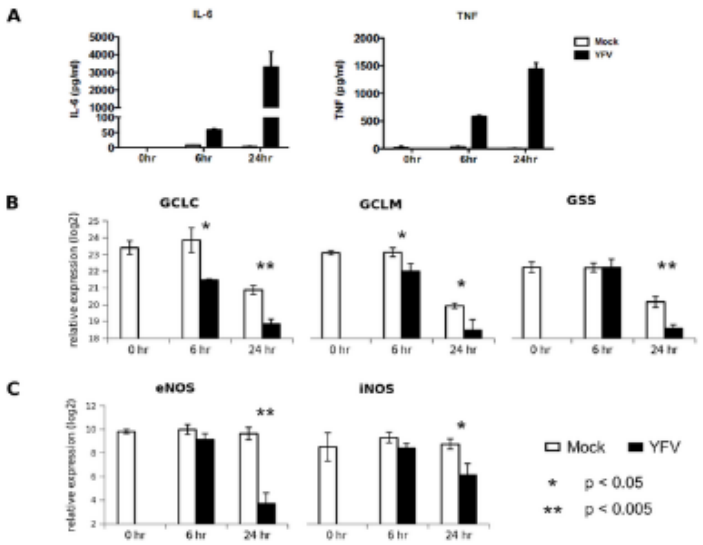


Li et al. 2013. PLoS Computational Biology. 9:e10031323

Experimental validation of *mummichog* prediction



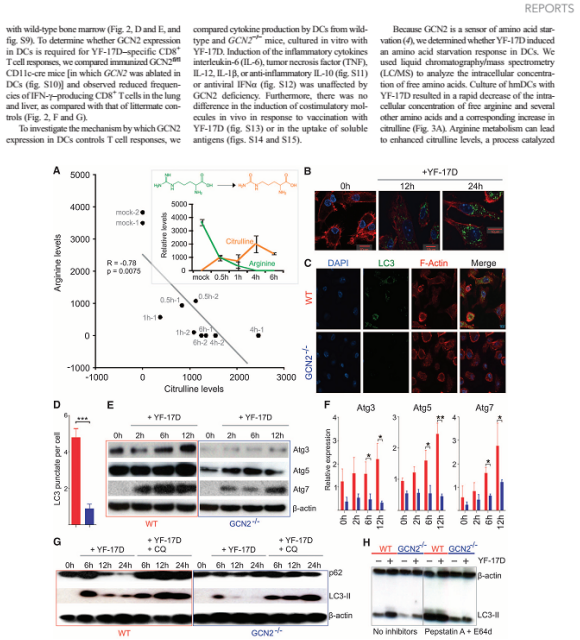
Tandem mass spectrometry confirmed 9/11 metabolites



Gene expression supported GSH/GSSG depletion and Arg/Cit conversion

Li et al. 2013. PLoS Computational Biology. 9:e10031323

Mummichog application cases



REPORTS

with wild-type bone marrow (Fig. 2, D and E, and fig. S9). To determine whether GCN2 expression in DCs is required for YF-17D–specific $CD8^+$ T cell responses, we compared immunized $GCN2^{fl/fl}$ $CD11c$-cre mice [in which *GCN2* was ablated in DCs (fig. S10)] and observed reduced frequencies of IFN-γ–producing $CD8^+$ T cells in the lung and liver, as compared with that of littermate controls (Fig. 2, F and G).

To investigate the mechanism by which GCN2 expression in DCs controls T cell responses, we compared cytokine production by DCs from wild-type and $GCN2^{-/-}$ mice, cultured in vitro with YF-17D. Induction of the inflammatory cytokines interleukin-6 (IL-6), tumor necrosis factor (TNF), IL-12, IL-1β, or anti-inflammatory IL-10 (fig. S11) or antiviral IFNα (fig. S12) was unaffected by GCN2 deficiency. Furthermore, there was no difference in the induction of costimulatory molecules in vivo in response to vaccination with YF-17D (fig. S13) or in the uptake of soluble antigens (figs. S14 and S15).

Because GCN2 is a sensor of amino acid starvation (*4*), we determined whether YF-17D induced an amino acid starvation response in DCs. We used liquid chromatography/mass spectrometry (*LC/MS*) to analyze the intracellular concentration of free amino acids. Culture of hmDCs with YF-17D resulted in a rapid decrease of the intracellular concentration of free arginine and several other amino acids and a corresponding increase in citrulline (Fig. 3A). Arginine metabolism can lead to enhanced citrulline levels, a process catalyzed

Fig. 3. YF-17D induces autophagy in dendritic cells via a mechanism dependent on GCN2. (**A**) Inverse correlation between the concentrations of free arginine and citrulline in hmDCs stimulated with YF-17D. Mock-treated DCs at 6 hours (without virus) are shown as controls. (Inset) Mean relative abundance ± SD of arginine and citrulline. (**B**) Culture of hmDCs with YF-17D induces autophagy as visualized by means of confocal microscopy. (**C**) Comparison of autophagy (LC3 punctate staining) in mBMDCs from wild-type or $GCN2^{-/-}$ mice, cultured in vitro with YF-17D for 6 hours. (**D**) Counts of LC3 granules per cell. (**E**) Comparison of the autophagy proteins in BMDC from wild-type or $GCN2^{-/-}$ mice cultured in vitro with YF-17D. 0h represents 30 min after DCs are cultured in low volume of low fetal bovine serum medium with YF-17D. Basal levels of Atg5 and Atg7 in freshly isolated DC were similar in wild-type and $GCN2^{-/-}$ mice (fig. S20). (**F**) Densitometric analysis of Western blots from three independent experiments. (**G**) Autophagy flux experiments depicting p62 and LC3II accumulation by chloroquine after culture with YF-17D. (**H**) Autophagy flux showing accumulation of LC3II 6 hours after YF-17D culture after treatment with lysosomal inhibitors (pepstatin and E64D). Data are representative of three independent experiments. $^*P < 0.05$; $^{**}P < 0.005$, Student's *t* test. Error bars indicate mean ± SEM

Ravindran et al. 2014.
Science 343:313

Aging Cell (2014) **13**, pp596–604 Doi: 10.1111/acel.12215

Effects of age, sex, and genotype on high-sensitivity metabolomic profiles in the fruit fly, *Drosophila melanogaster*

Jessica M. Hoffman,[1†] Quinlyn A. Soltow,[2,3,4†] Shuzhao Li,[2] Alfire Sidik,[1,5] Dean P. Jones[2,3,5*] and Daniel E. L. Promislow[1,7*]

[1]Department of Genetics, University of Georgia, Athens, GA 30602, USA
[2]Division of Pulmonary Allergy & Critical Care Medicine, Department of Medicine, Emory University, Atlanta, GA 30322, USA
[3]Department of Medicine, Clinical Biomarkers Laboratory, Emory University, Atlanta, GA 30322, USA
[4]ClinMet Inc., 3210 Merryfield Row, San Diego, CA 92121, USA
[5]Center for Health Discovery & Well Being, Emory University, Atlanta, GA 30322, USA

Summary

Researchers have used whole-genome sequencing and gene expression profiling to identify genes associated with age, in the hope of understanding the underlying mechanisms of senescence. But there is a substantial gap from variation in gene sequences and expression levels to variation in age or life expectancy. In an attempt to bridge this gap, here we describe the effects of age, sex, genotype, and their interactions on high-sensitivity metabolomic profiles in the fruit fly, *Drosophila melanogaster*. Among the 6800 features analyzed, we found that over one-quarter of all metabolites were significantly associated with age, sex, genotype, or their interactions, and multivariate analysis shows that individual metabolomic profiles are highly predictive of these traits. Using a metabolomic equivalent of gene set enrichment analysis, we identified numerous metabolic pathways that were enriched among metabolites associated with age, sex, and genotype, including pathways involving sugar and glycerophospholipid metabolism, neurotransmitters, amino acids, and the carnitine shuttle. Our results suggest that high-sensitivity metabolomic studies have excellent potential not only to reveal mechanisms that lead to senescence, but also to help us understand differences in patterns of aging among genotypes and between males and females.

Key words: age; aging; *Drosophila melanogaster*; genetic variation; genotype; heritability; metabolomics; sex; systems biology.

Introduction

Lifespan is a highly heritable trait. Over the past 20 years, researchers working on lab-adapted organisms have been able to identify evolutionarily conserved genetic pathways which, when knocked down or overexpressed, are able to dramatically increase lifespan. These successes underscore two critical questions: first, at the molecular level, what are the underlying mechanisms by which these genes affect longevity; second, at the population level, do these same genes account for standing variation in longevity in natural populations?

These questions are complicated by the fact that the age at which an individual dies depends not only on its genotype, but also on a lifetime of effects accumulated through environmental exposure, the environment-specific response of genes, and the downstream physiological consequences of these complex factors. Fortunately, whole-genome sequencing and genome-wide association (GWA) studies now make it possible to identify segregating alleles that affect complex phenotypes such as body height, diabetes, schizophrenia, and even longevity (Jeck *et al.*, 2012), but GWA studies suffer from numerous challenges, and these are further compounded in analyses of lifespan. First, alleles identified in GWA studies typically explain just 0.1–1.0% of the variation in complex traits (Park *et al.*, 2010). Second, the genetic basis of lifespan appears, at least in part, to differ between the sexes (Burger & Promislow, 2004). Third, lifespan includes a substantial degree of stochasticity, varying dramatically even among genetically identical individuals raised in a constant and identical environment (Kirkwood *et al.*, 2005). Finally, and perhaps most importantly, lifespan is a highly composite trait potentially influenced by the functional decline of many underlying processes. To fully understand the genetics of lifespan, we need to understand the genetics not simply of age at death, but rather of the underlying causes of death.

Here, we suggest that many of the challenges that we face in our attempts to define the pathways that account for age-related declines in function, and for genetic variation in these declines, can be resolved through the use of high-resolution metabolomics (Mishur & Rea, 2012). If we can decompose the physiological processes that influence morbidity and mortality to their constituent components (i.e., the metabolome), we will be an important step closer to bridging the gap between genotype and phenotype (Fig. 1). The metabolome is effectively a functional intermediate between genotype and phenotype.

Previous work illustrates how the metabolome can serve as a strong bridge between genotype and phenotype. While allelic variation typically explains only a small fraction of the variation in complex phenotypes, GWA studies of the metabolome have found genetic variants capable of explaining up to 60% of the variance in the concentration of individual metabolites (Suhre *et al.*, 2011). The metabolome also appears to be a sensitive indicator of age-related physiological changes both in invertebrates and vertebrates (e.g., Sarup *et al.*, 2012; Yu *et al.*, 2012). Moreover, different mutants that extend longevity share common metabolomic signatures (e.g., *Caenorhabditis elegans*: Fuchs *et al.*, 2010; *Mus musculus*: Wijeyesekera *et al.*, 2012).

While we have learned much from these initial studies, most have been limited by the use of relatively low-sensitivity metabolomics technology and by limited genetic information. Studies of age and the

Correspondence
Daniel Promislow, Department of Pathology and Department of Biology, University of Washington, Box 357705, 1959 NE Pacific Street, Seattle, WA 98195, USA. Tel.: +206 616 6994; fax: +206 616 8271; e-mail: promislo@uw.edu
*Co-senior authors, †Co-first authors.
‡Present address: Department of Molecular Biosciences, University of Texas, Austin, TX 78712, USA
†Present address: Department of Pathology and Department of Biology, University of Washington, Seattle, WA 98195, USA

Accepted for publication *25 January 2014*

Hoffman et al, 2014.
Aging Cell 13: 596-604

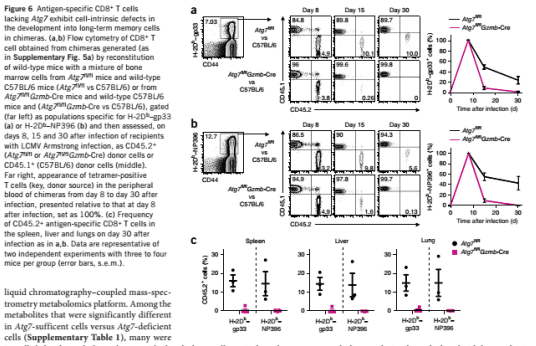


Figure 6 Antigen-specific $CD8^+$ T cells lacking *Atg7* exhibit cell-intrinsic defects in the development into long-term memory cells in chimeras. (**a**,**b**) Flow cytometry of $CD8^+$ T cell obtained from chimeras generated (as in **Supplementary Fig. 5a**) by reconstitution of wild-type mice with a mixture of bone marrow cells from $Atg7^{fl/fl}$ mice and wild-type C57BL/6 mice ($Atg7^{fl/fl}$ vs C57BL/6) or from $Atg7^{fl/fl}$*Gzmb*-Cre mice and wild-type C57BL/6 mice and ($Atg7^{fl/fl}$*Gzmb*-Cre vs C57BL/6), gated (far left) as populations specific for H-$2D^b$–gp33 (**a**) or H-$2D^b$–NP396 (**b**) and then assessed, on days 8, 15 and 30 after infection of recipients with LCMV Armstrong infection, as $CD45.2^+$ ($Atg7^{fl/fl}$ or $Atg7^{fl/fl}$*Gzmb*-Cre) donor cells or $CD45.1^+$ (C57BL/6) donor cells (middle). Far right, appearance of tetramer-positive T cells (key, donor source) in the peripheral blood of chimeras from day 8 to day 30 after infection, presented relative to that at day 8 after infection, set as 100%. (**c**) Frequency of $CD45.2^+$ antigen-specific $CD8^+$ T cells in the spleen, liver and lungs on day 30 after infection as in **a**,**b**. Data are representative of two independent experiments with three to four mice per group (error bars, s.e.m.).

liquid chromatography–coupled mass-spectrometry metabolomics platform. Among the metabolites that were significantly different in *Atg7*-sufficient cells versus *Atg7*-deficient cells (**Supplementary Table 1**), many were in well-defined metabolic pathways with clear links to cell survival, and some have been linked to T cell differentiation (**Supplementary Fig. 6**). For example, among the pathways that were significantly different in the genotypes were the carnitine shuttle and di-unsaturated fatty acid β-oxidation, both of which are part of fatty acid metabolism in the mitochondria (**Fig. 7a**). T cell–intrinsic mitochondrial fatty acid oxidation (FAO) has been shown to be a critical metabolic pathway for the stable development of long-lived memory T cells[26,27]. These data might indicate that autophagy contributes to the production of the lipid substrates for mitochondrial FAO and the fueling of oxidative phosphorylation in these cells[28]. Consistent with a role for autophagy in lipid metabolism in T cells during the transition to the memory phase, our metabolite analysis showed that lipid biosynthetic pathways were dysregulated in the absence of Atg7 (**Fig. 7a**). Also of note, the glucosamine 6-phosphate and glycan-biosynthesis pathways were perturbed in *Atg7*-deficient T cells (**Fig. 7a** and **Supplementary Fig. 6**). The hexosamine pathway provides substrates for the glycosylation of receptors for growth factor cytokines, which leads to receptor stability and the maintenance of survival signals[29]. Given the importance of growth factor cytokines in supporting the development of memory T cells[2], it is possible that T cells lacking autophagy machinery cannot maintain the proper growth factor signals that support memory T cell development, linking autophagy to glycosylation and the maintenance of cellular metabolism.

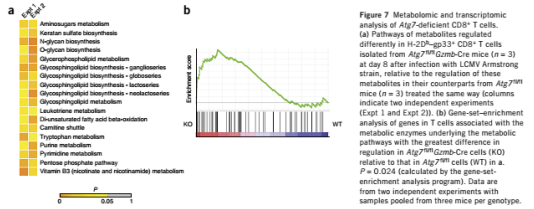
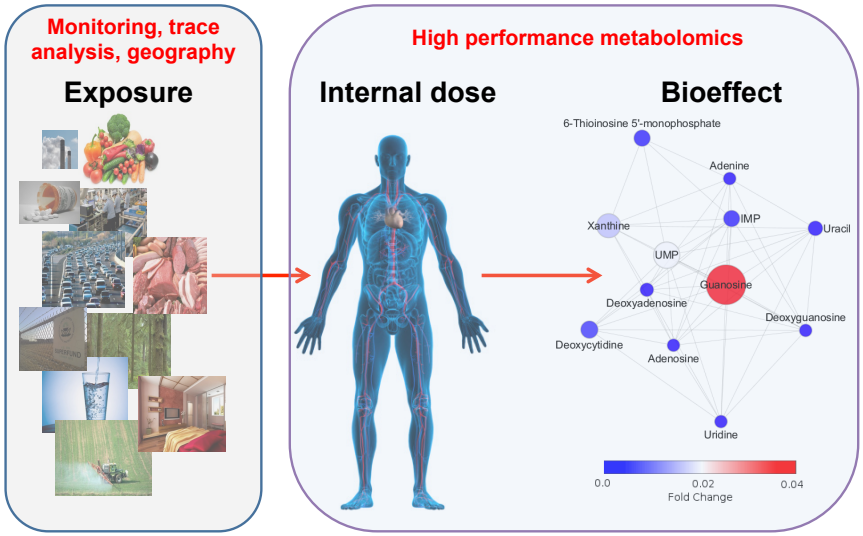
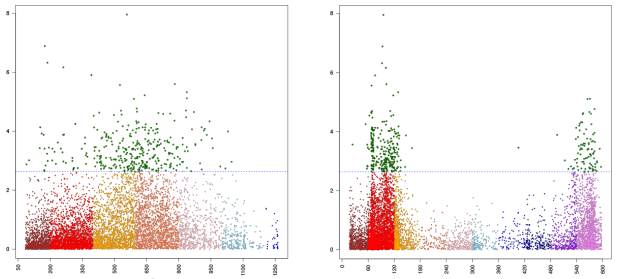


Figure 7 Metabolomic and transcriptomic analysis of *Atg7*-deficient $CD8^+$ T cells. (**a**) Pathways of metabolites regulated differently in H-$2D^b$–gp33$^+$ $CD8^+$ T cells isolated from $Atg7^{fl/fl}$*Gzmb*-Cre mice ($n = 3$) at day 8 after infection with LCMV Armstrong strain, relative to the regulation of these metabolites in their counterparts from $Atg7^{fl/fl}$ mice ($n = 3$) treated the same way (columns indicate two independent experiments (Expt 1 and Expt 2)). (**b**) Gene-set–enrichment analysis of genes in T cells associated with the metabolic enzymes underlying the metabolic pathways with the greatest difference in regulation in $Atg7^{fl/fl}$*Gzmb*-Cre cells (KO) relative to that in $Atg7^{fl/fl}$ cells (WT) in **a**. $P = 0.024$ (calculated by the gene-set-enrichment analysis program). Data are from two independent experiments with samples pooled from three mice per genotype.

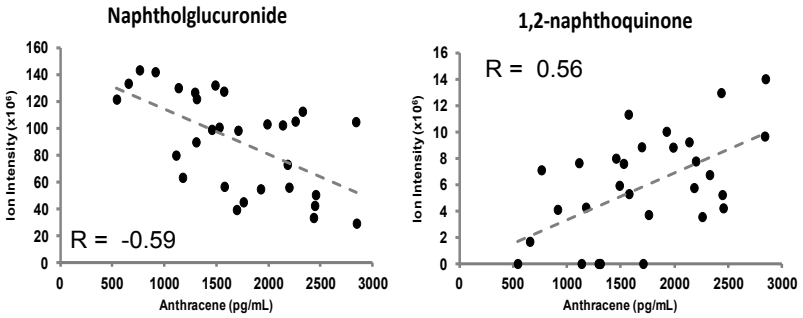
Xu et al, 2014.
Nature Immunology. 15:1152

Gauging environmental exposure and bioeffect simultaneously

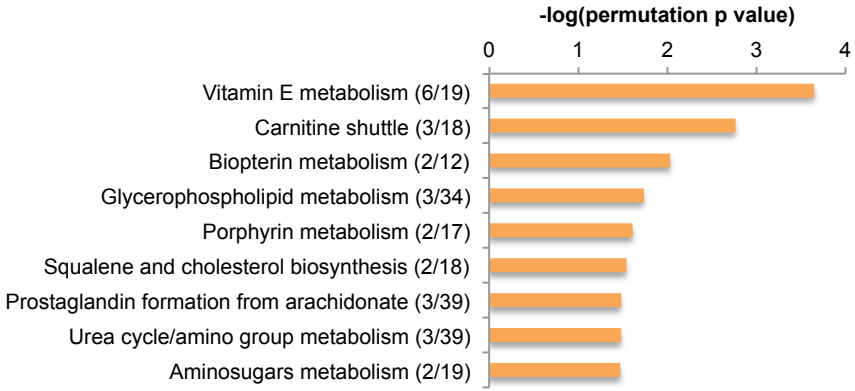




Significant metabolite features via ANOVA



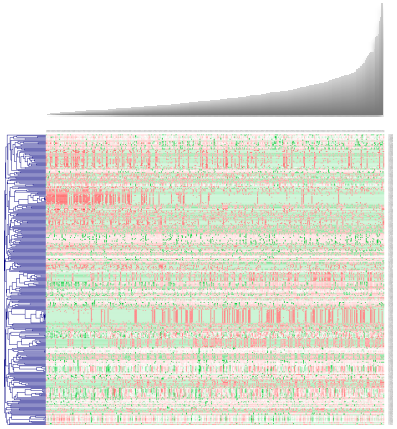
PAH metabolites were correlated with targeted analysis of plasma PAHs

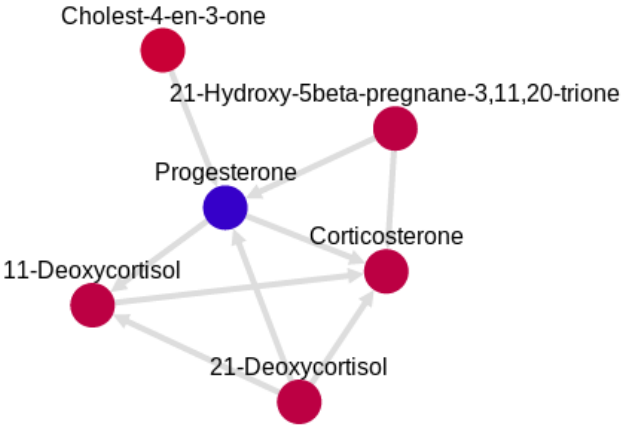


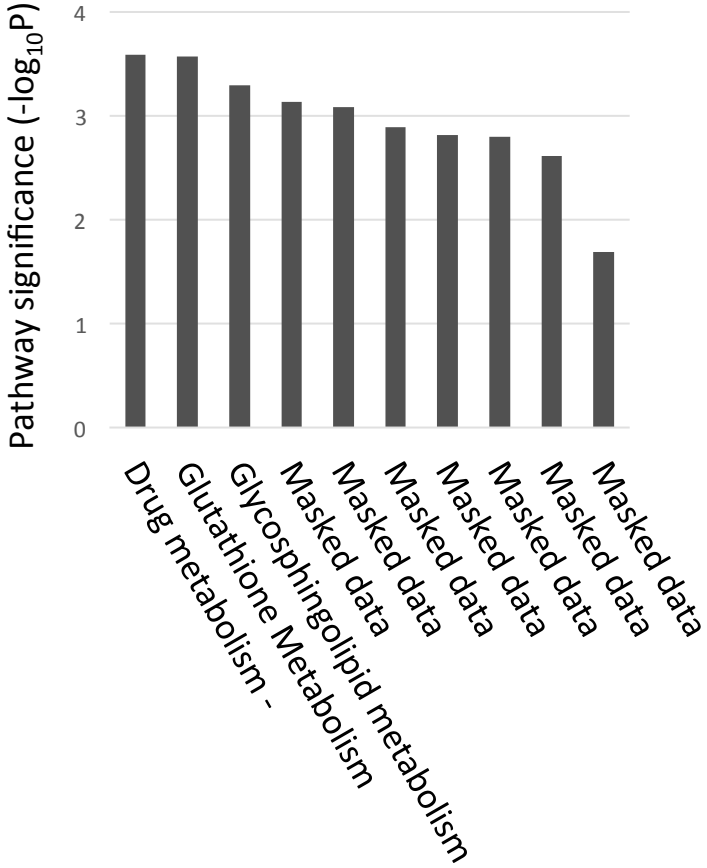
Significant pathways via *mummichog*

Courtesy: Doug Walker

Metabolome-wide association study (MWAS) of LDL cholesterols



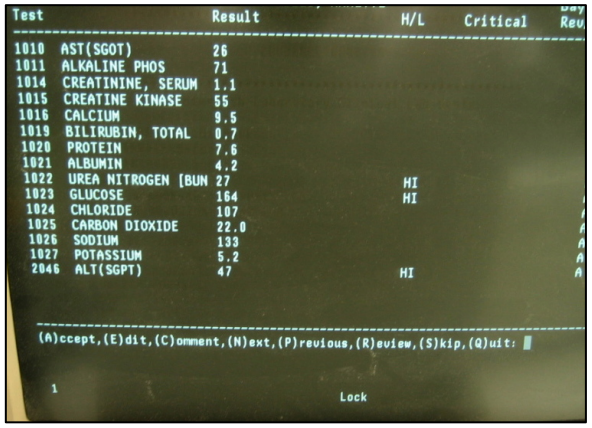


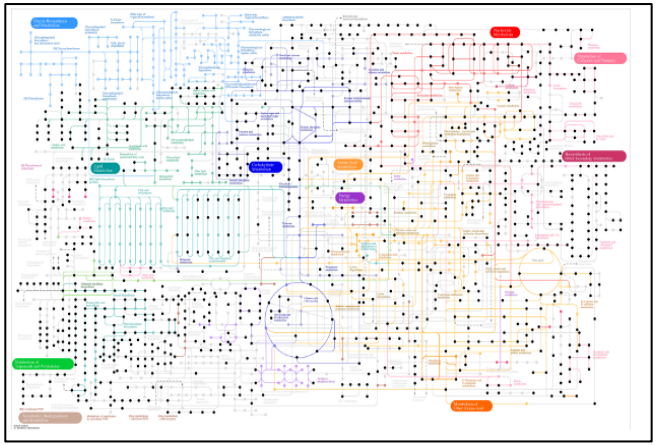


Integration of metabolomics and transcriptomics

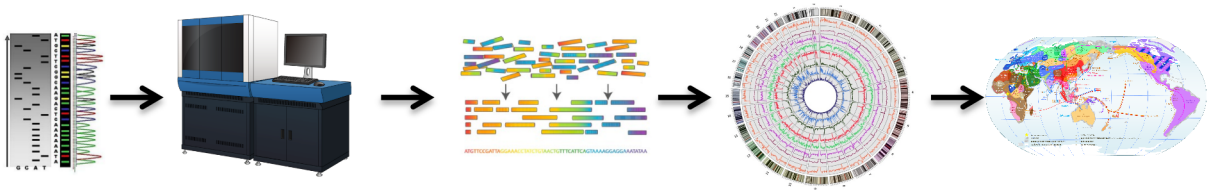
Unpublished data deleted

Towards Universal health screen

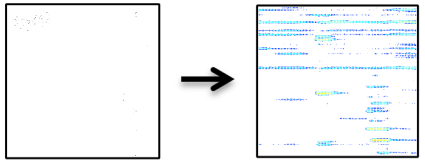




Computational metabolomics in the making



Automated sequencer Assembly Reference genome Variations Genomics



? ? ? Metabolomics

Summary

- High-resolution metabolomics leads to new work flow
- *Mummichog* is an effective tool to bridge genomics and metabolomics
- Impacts to environmental sciences, epidemiology, systems biology, medicine…
- Many challenges ahead

EMORY UNIVERSITY
DEPARTMENT OF MEDICINE

CLINICAL BIOMARKERS LABORATORY

Thanks, Team!

High-Resolution Metabolomics Team

Shuzhao Li

Biology

Human population, Intervention

Animal model

Cell culture

Foodstuff, Environmental

Tom Ziegler
Young-Mi Go,
Many collaborators

Analytical

LTQ-FT

LTQ-Velos Orbitrap

Q-Exactive-HF

Fusion

Bill Liang
ViLinh Tran

Computational

apLCMS

xMSanalyzer

xMSannotator

xmsPANDA

Karan Uppal
Tianwei Yu
Sophia Banton
Andrei Todor

Pathway analysis

Mummichog, other tools

Metabolite identification

MS^n

Doug Walker
Collaborators

EMORY
SCHOOL OF MEDICINE

Jones et al 2012 Annu Rev Nutr 32:183-202
T Yu et al 2009 Bioinformatics 25:1930-6
JM Johnson et al 2008 Clin Chim Acta 396:43-48
JM Johnson et al 2010 Analyst 135: 2864-2870
K Uppal et al 2013 BMC Bioinformatics 14:15
Q Soltow et al 2013 Metabolomics 9:S132–S143
T Yu et al 2013 J Proteome Res 12:1419-27
S Li et al 2013 Plos Comp Biol 9:e1003123