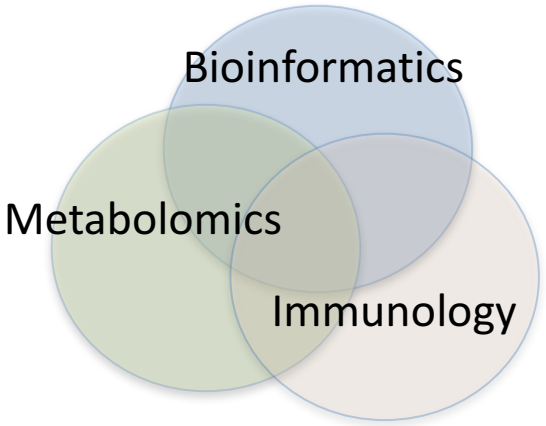
EMORY UNIVERSITY

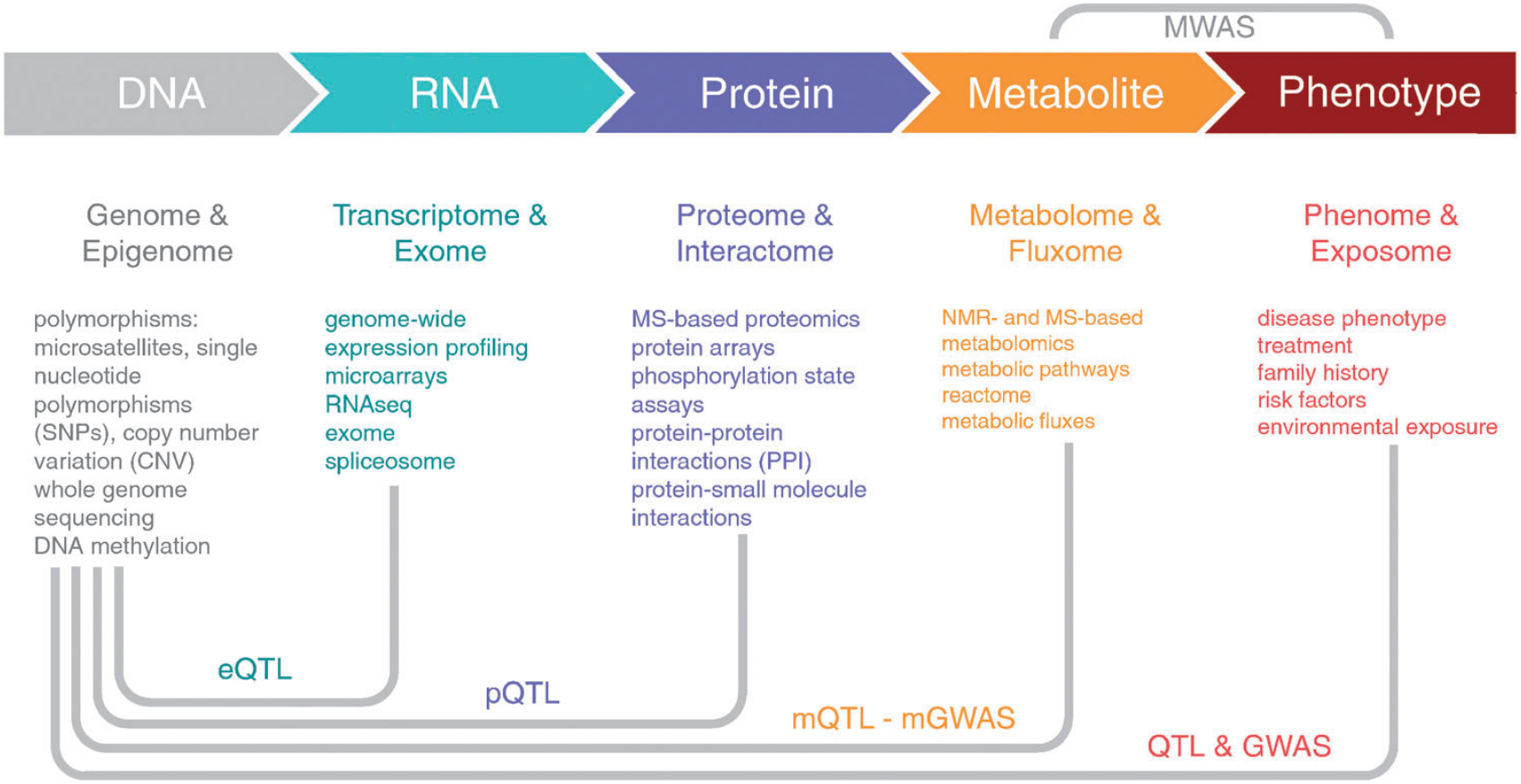
Overview of Metabolomics Workflow



Shuzhao Li, Ph.D
Assistant Professor
Department of Medicine
Emory University
August 26, 2019

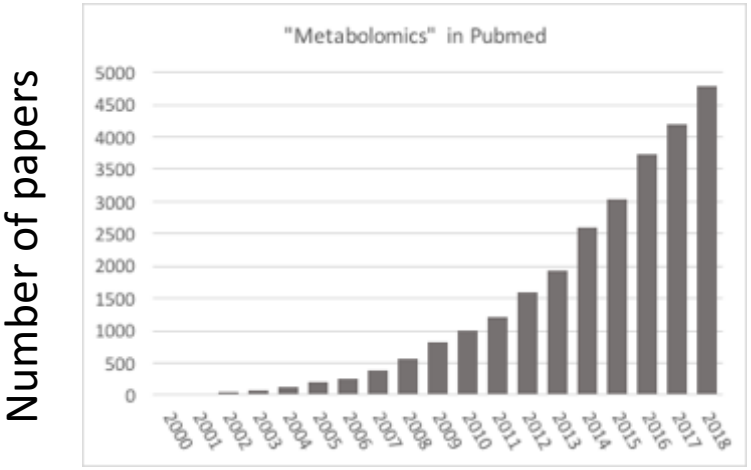
shuzhao.li@gmail.com shuzhao-li@GitHub

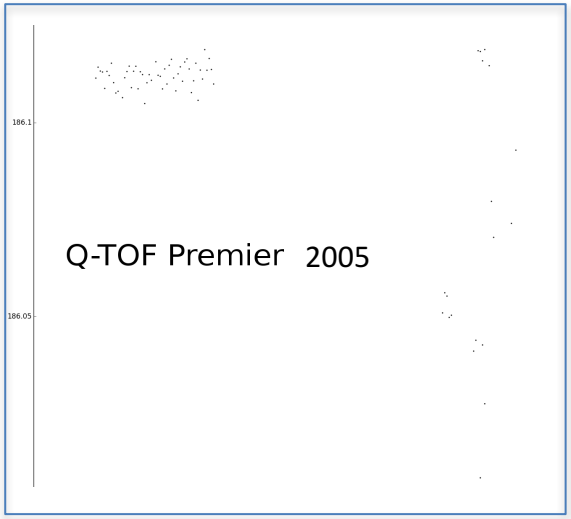
Biomedical –omics data

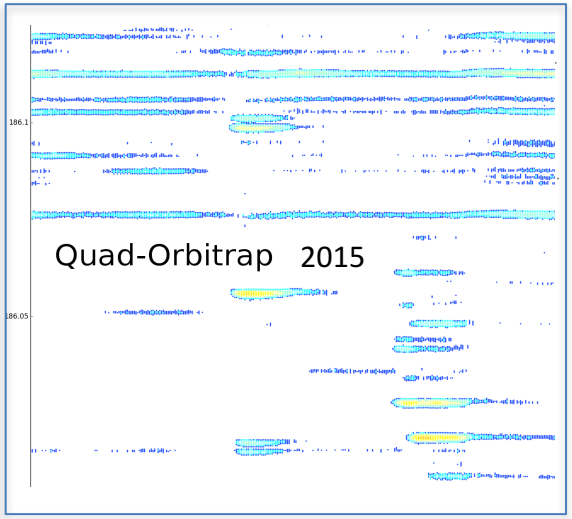


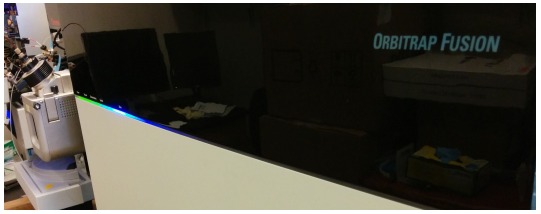
Dumas, 2012. Mol. BioSyst. 8:2494

Advance of Metabolomics

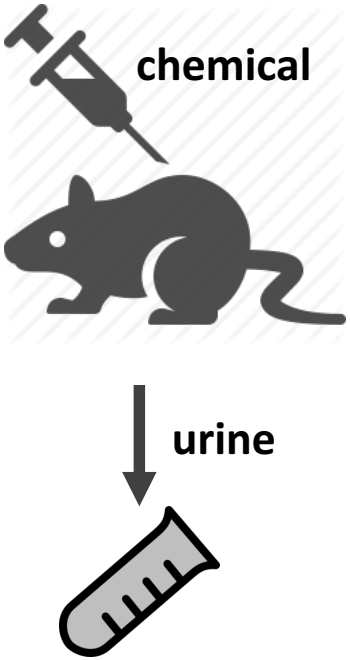


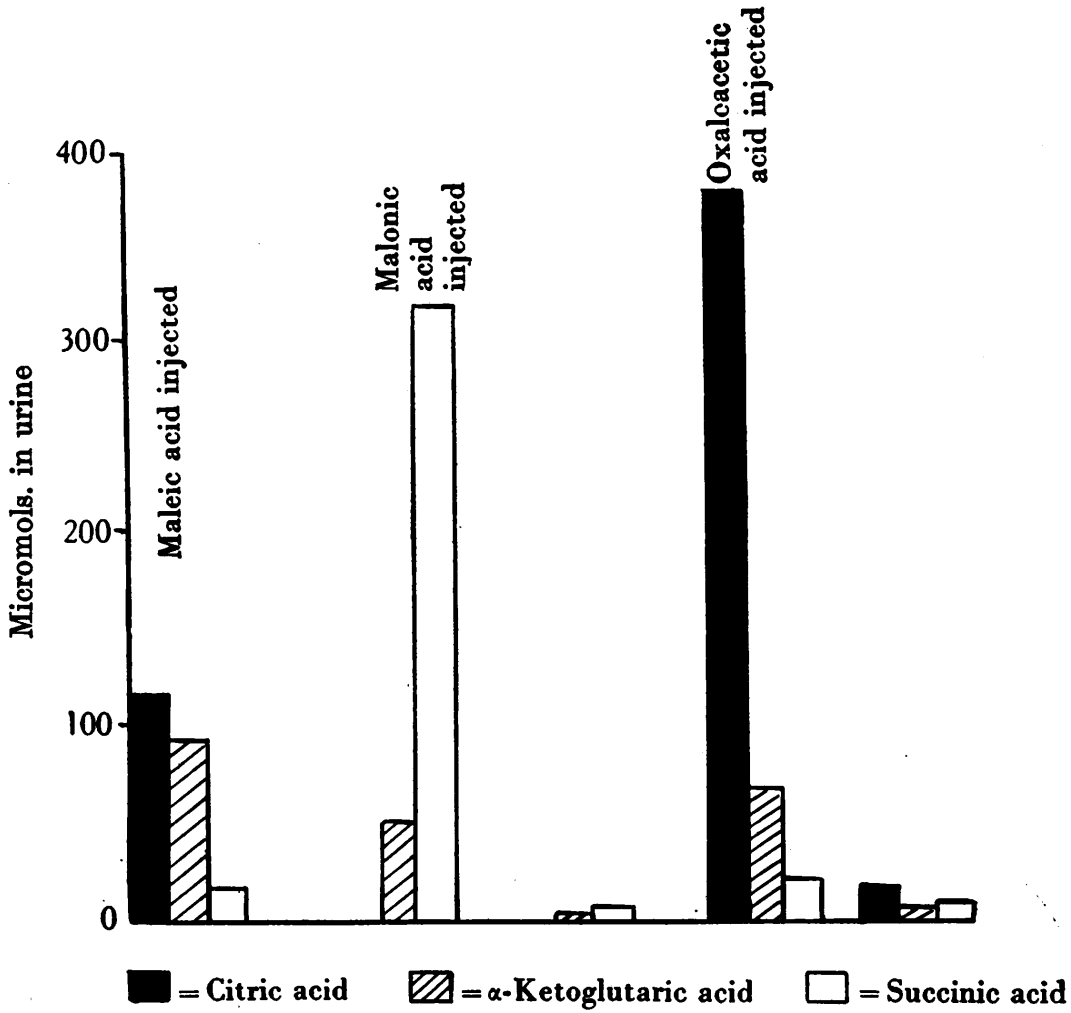




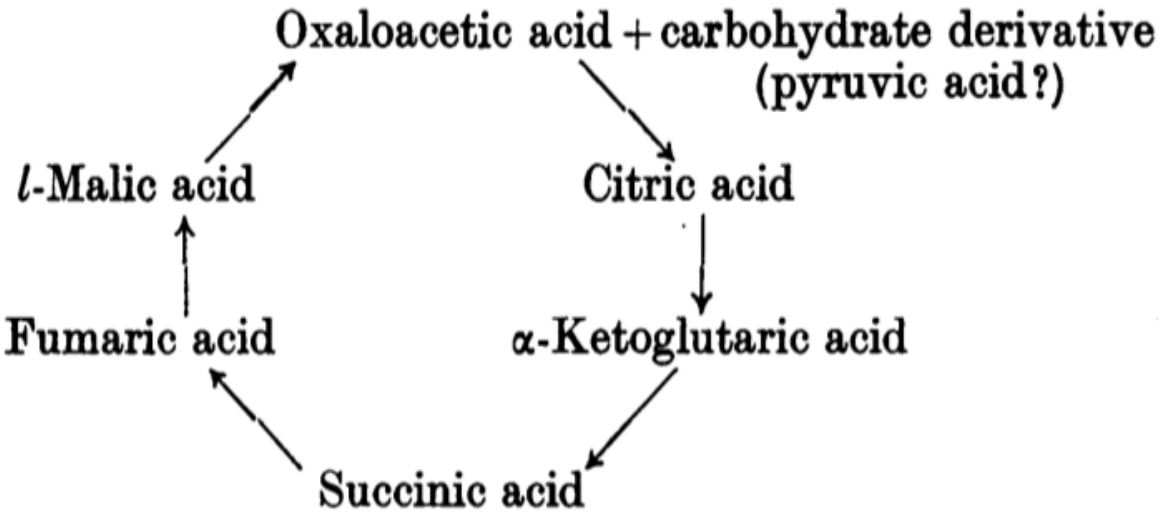


Krebs et al, 1937





Krebs Cycle, 1937



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

Beadle 1941

Genetic Control of Biochemical Reactions in Neurospora

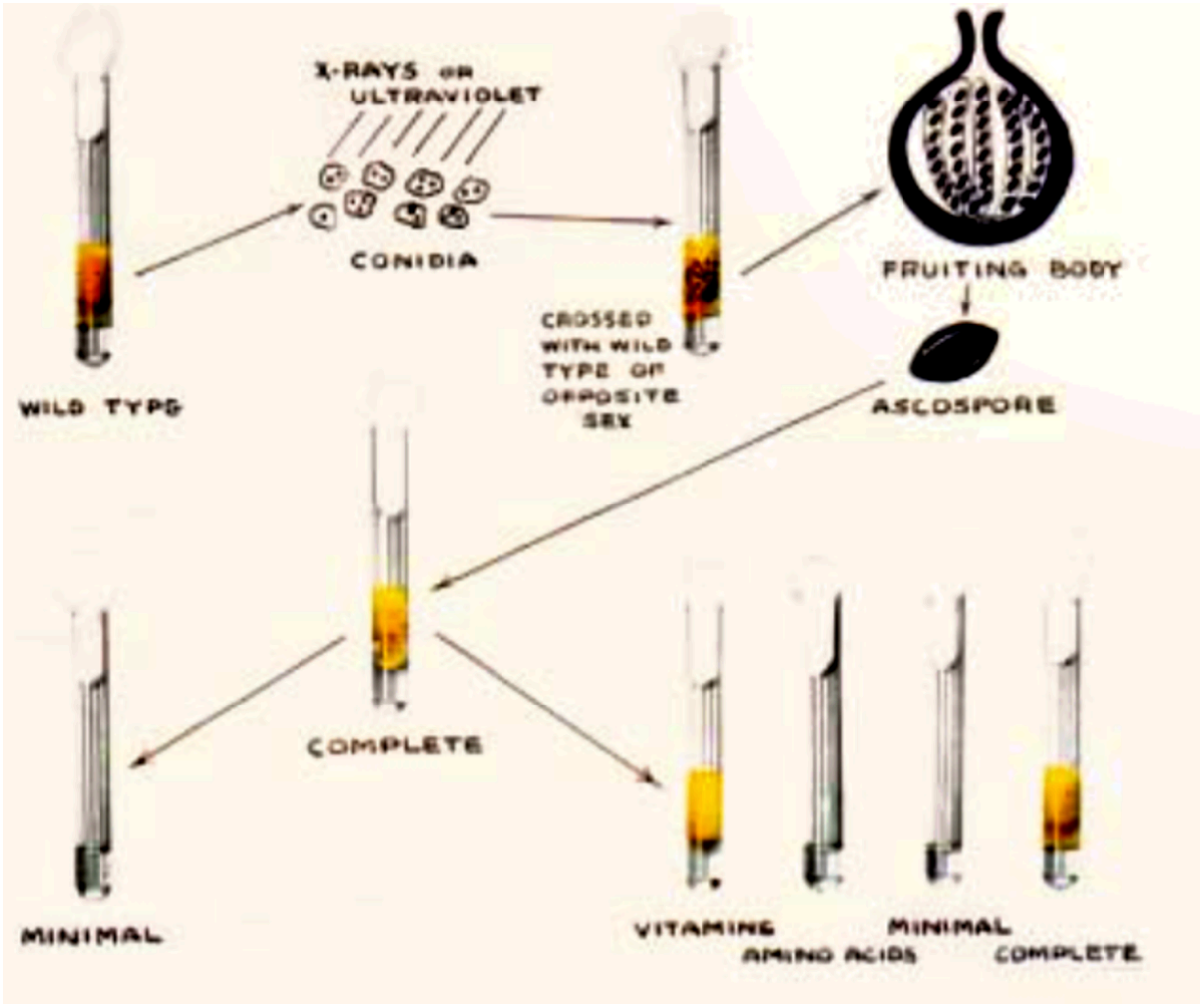
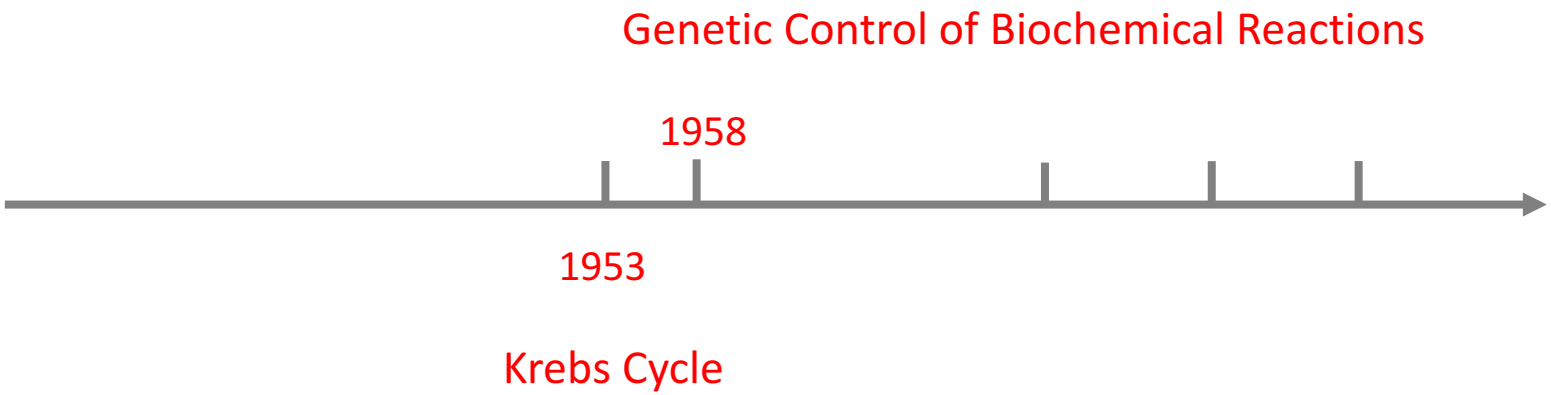
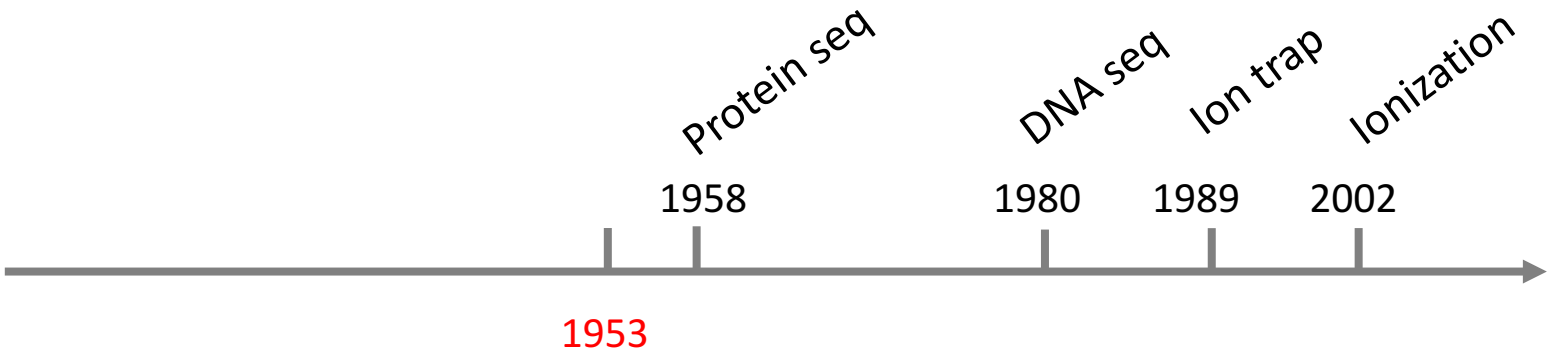


Figure from Nature Reviews Genetics **5**, 949-954, 2004

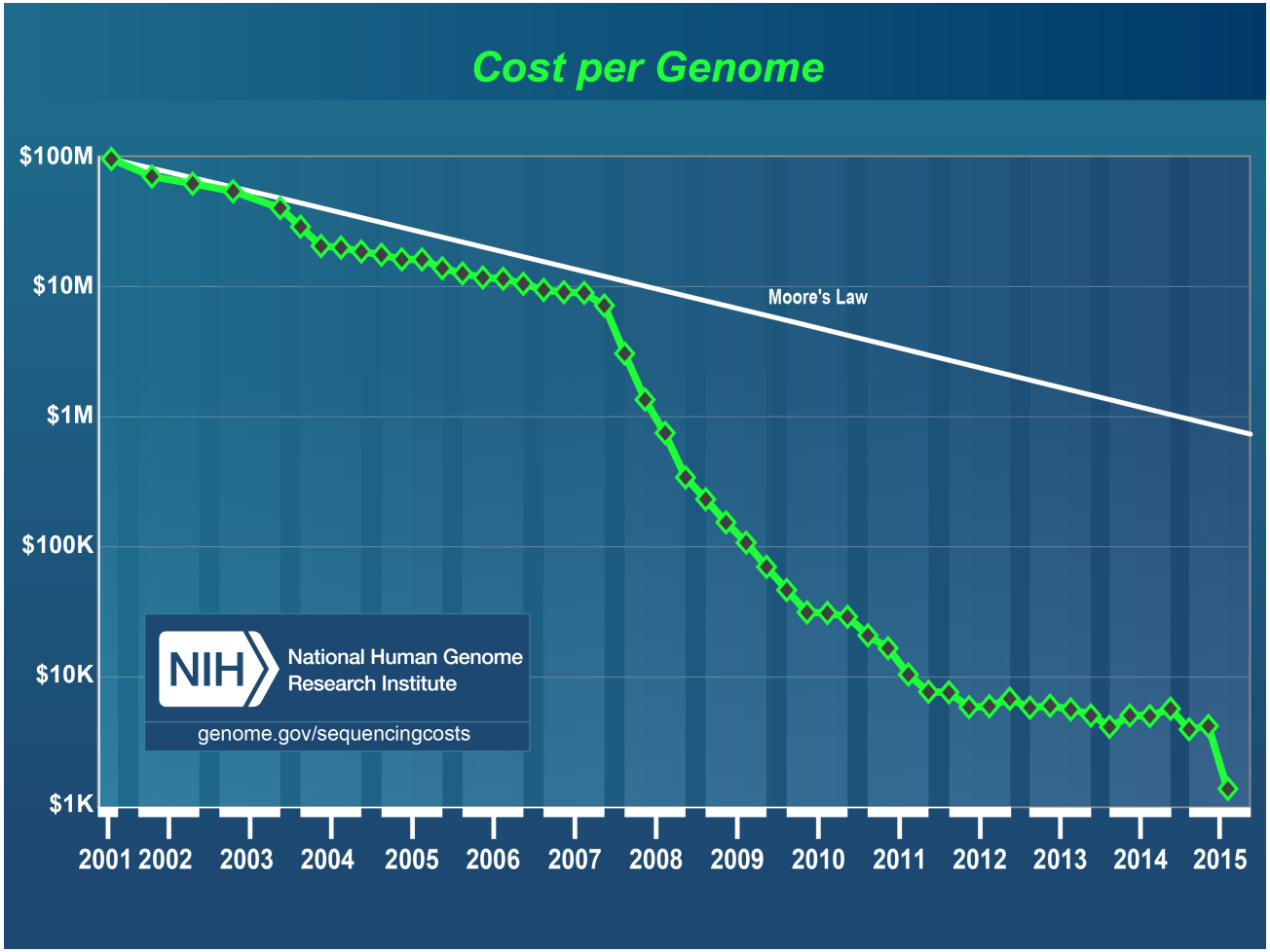
Related Nobel Prizes



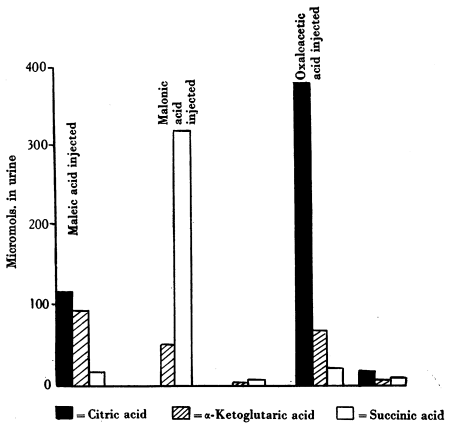
Related Nobel Prizes

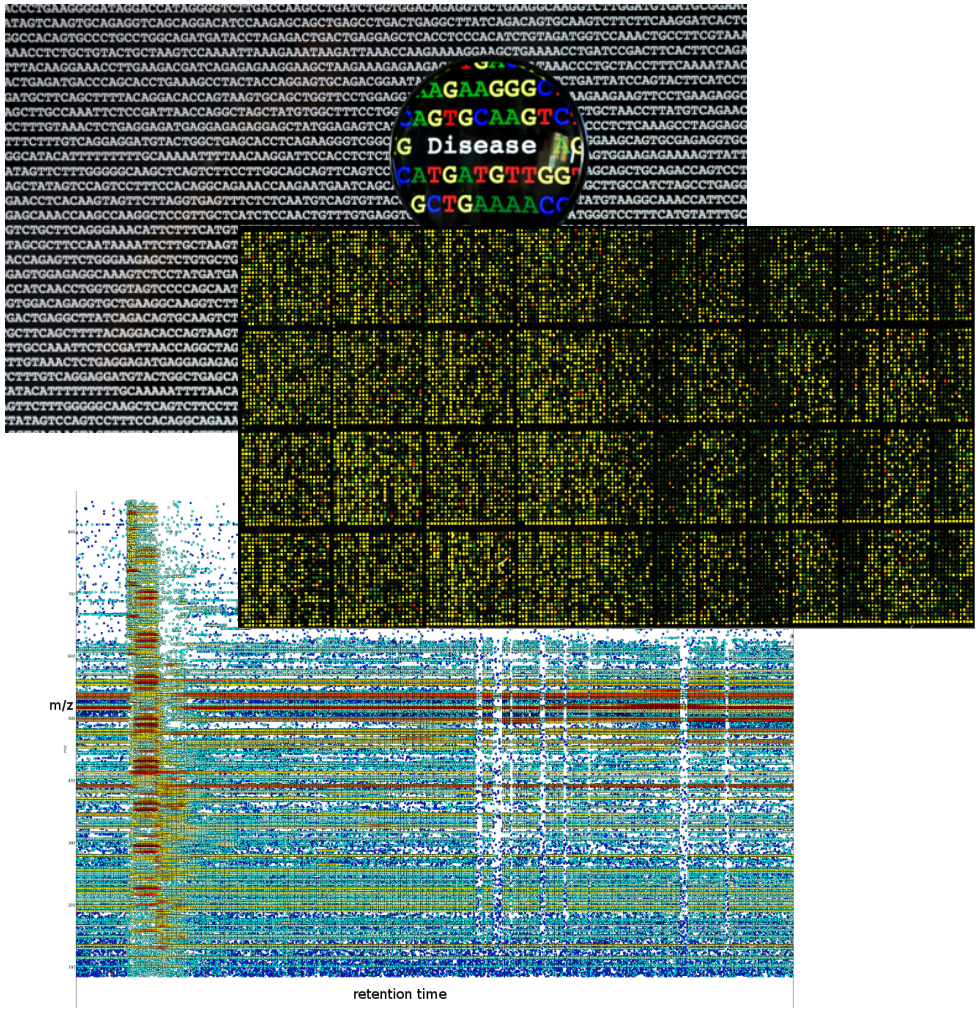


Economy of data generation (genomics)

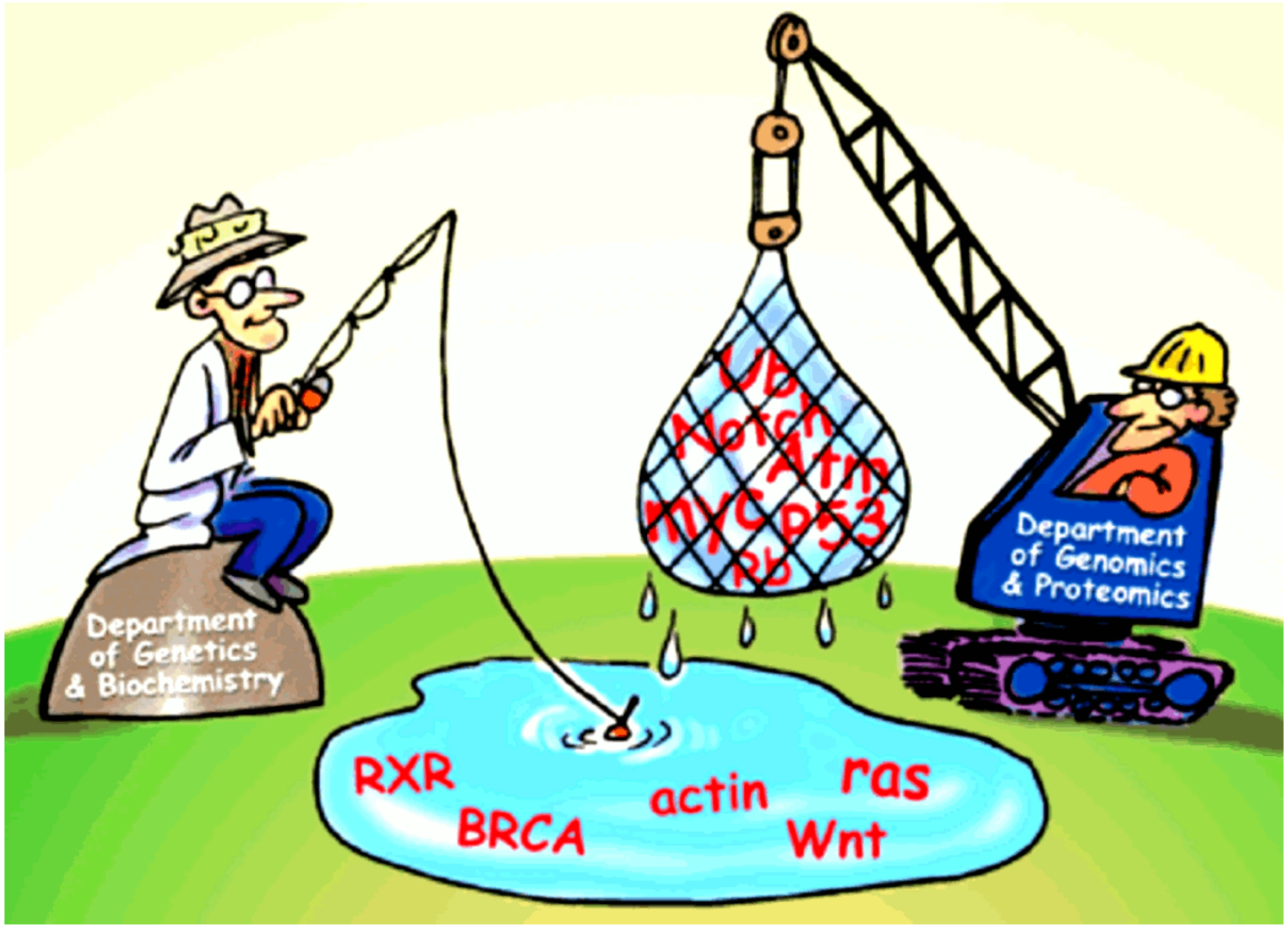


The data deluge





When science goes -omics

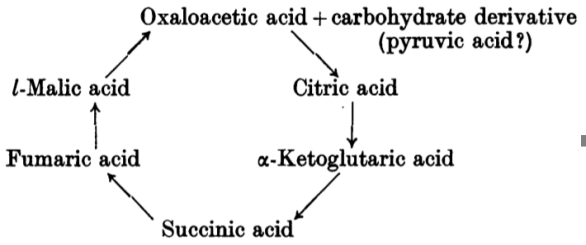


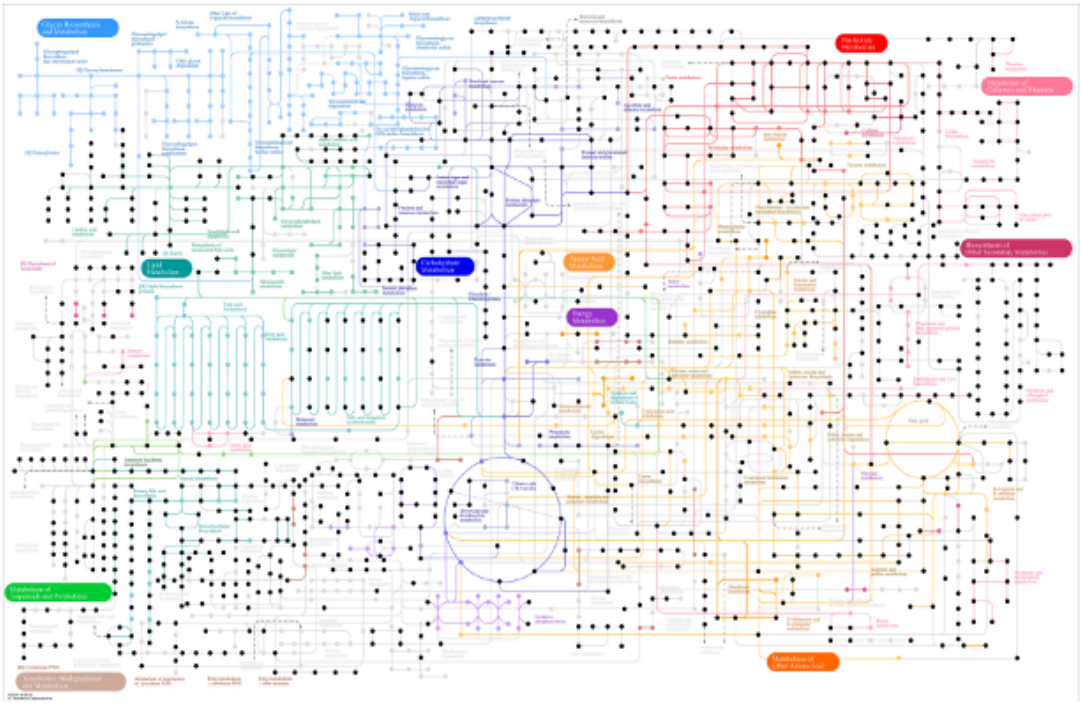
Science (2001) 291:1221

Working with big data

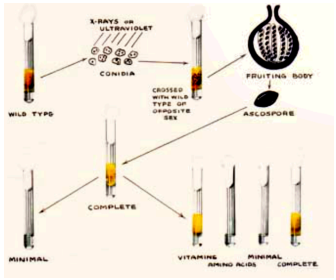
- Global profiling may support many competing hypotheses
- False discovery rate
- Latent structures in data
- Statistical analysis and inference, machine learning
- Combining prior knowledge and domain knowledge
- Integration of different types of data and evidence
- Simulation

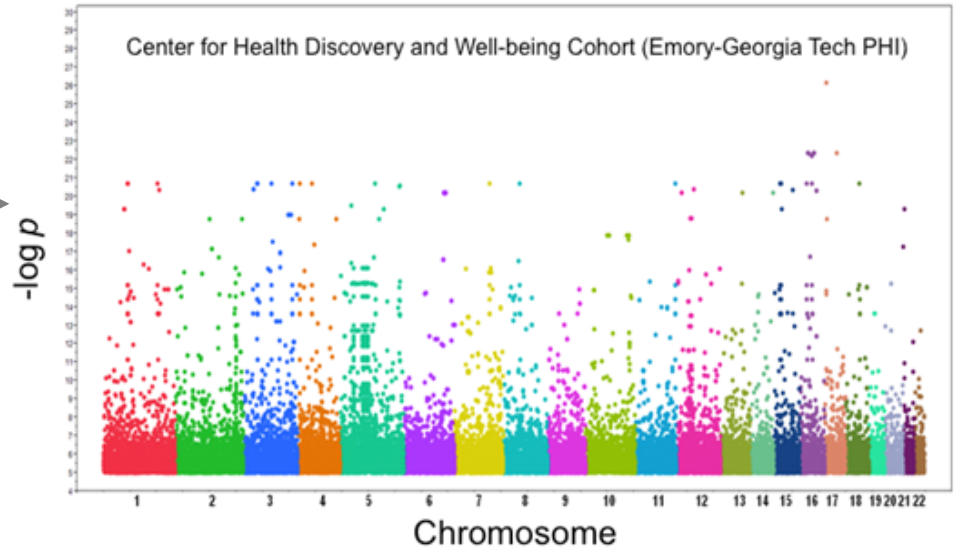
Newer metabolomics cover majority of metabolic pathways



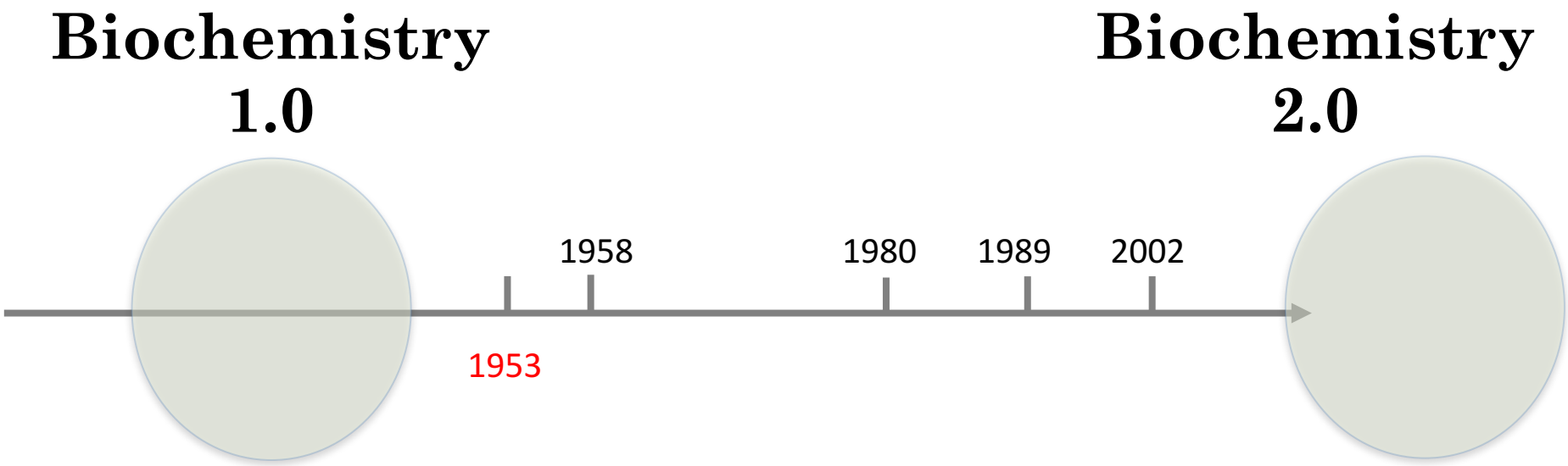


G × M, Metabolome wide association of genetic variations

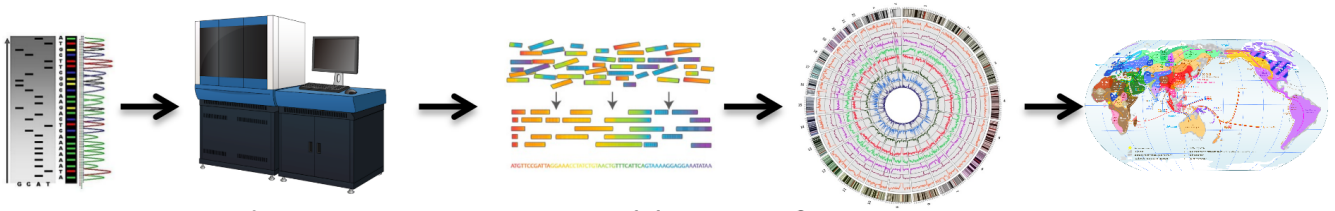




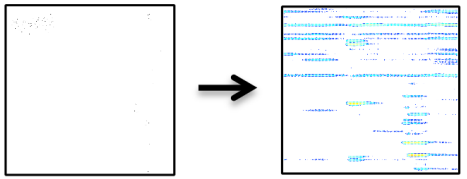
Courtesy: Dean P. Jones and Greg Gibson


Biochemistry 1.0
Biochemistry 2.0
1953
1958
1980
1989
2002

Challenges and Opportunities of Metabolomics

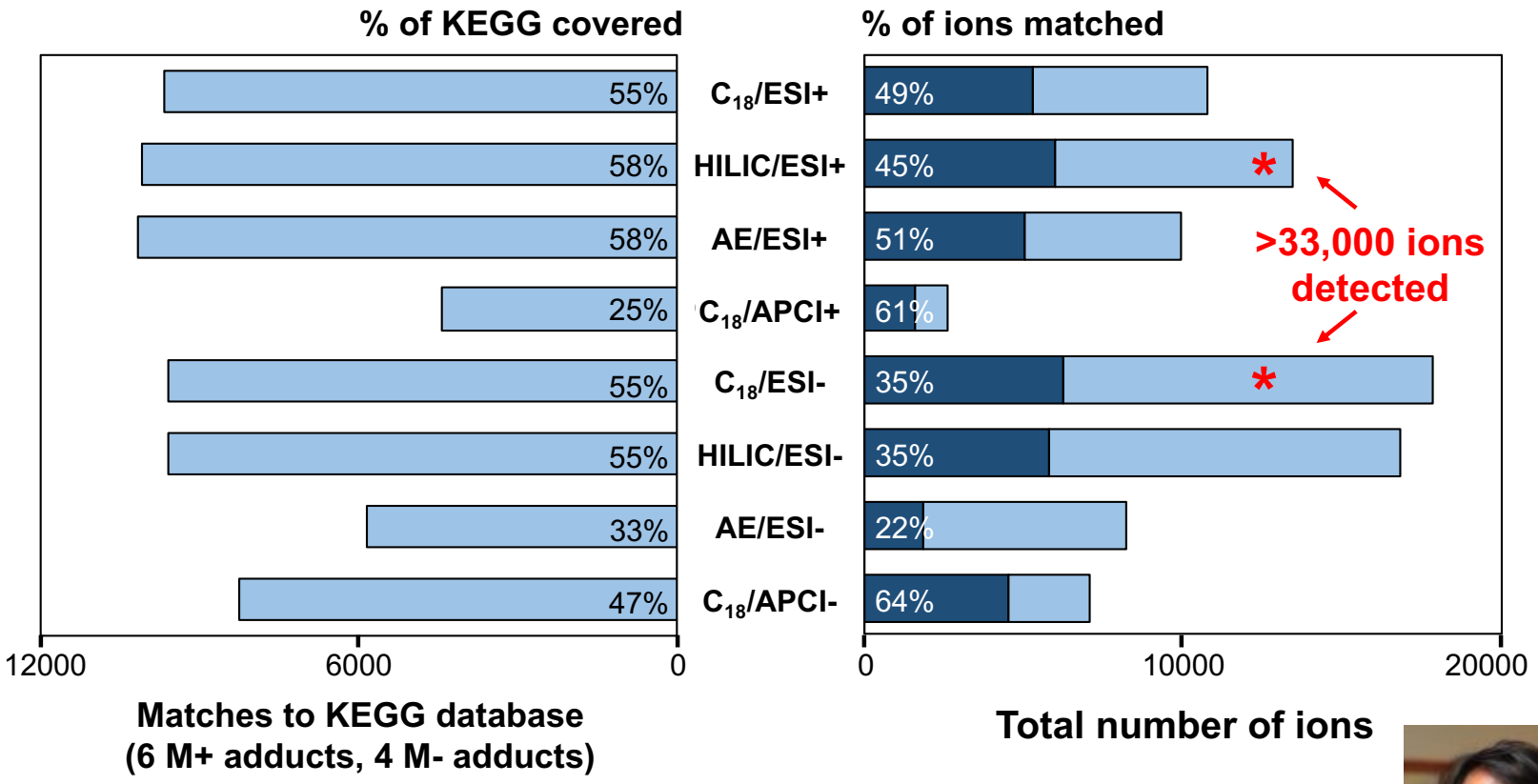


Automated sequencer Assembly Reference genome Variations Genomics



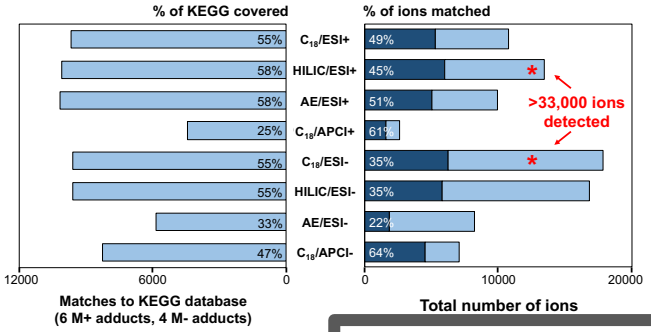
? → ? → ? Metabolomics

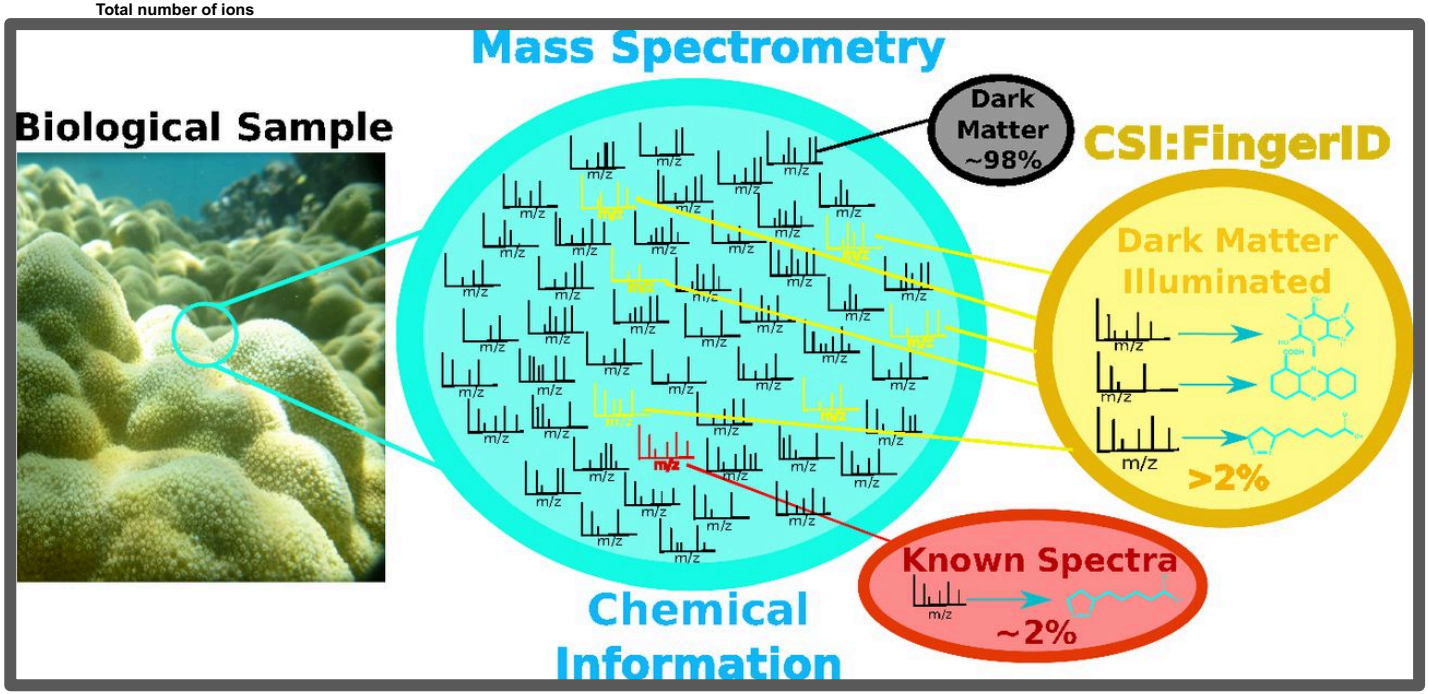
Most metabolomic peaks are unidentified



Ken Liu et al (2016) JOEM 58: S53-S61

Dark matters of Metabolomics





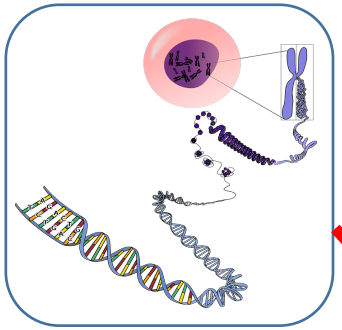
Ricardo R. da Silva et al. PNAS 2015;112:41:12549

Metabolomics is Biochemistry 2.0

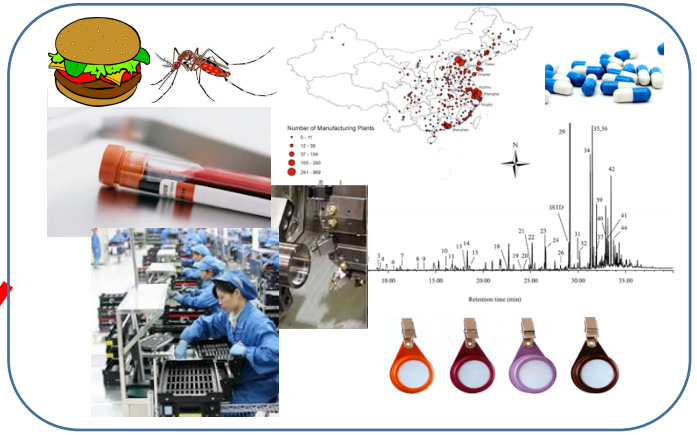
- Catalog unidentified chemicals in humans
- Update of metabolic models, reactions and pathways (new maps)
- Signaling and regulatory roles of metabolites and chemicals
- Linking genomes and environmental exposures (exposome), including microbial products, diet and drugs
- Enabling precision medicine with digital health
- Data and computation intensive

G × M × E

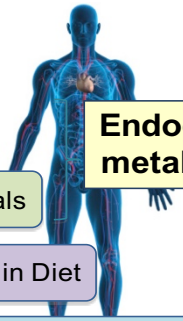
Genome



Environment



Gene function



- Core Biological Metabolome
- Microbiome-related Chemicals
- Non-nutritive Chemicals in Diet
- Supplements and Pharmaceuticals

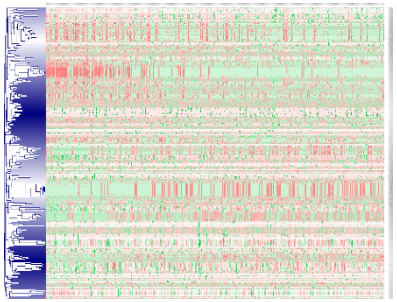
Endogenous metabolome

Environmental metabolome

- Commercial Products
- Environmental Chemicals

Body burden

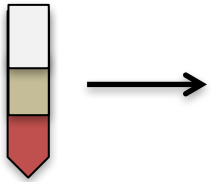
Molecular response



Metabolome

Figure by Doug Walker, Dean Jones & Shuzhao Li

Metabolomics workflow (simplified)

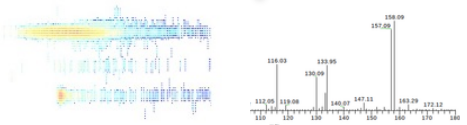


Sample preparation

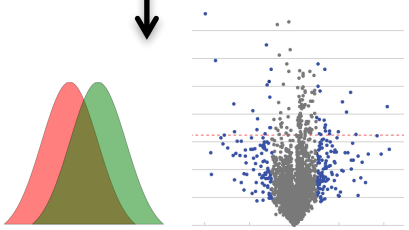
chromatography

Mass spectrometry

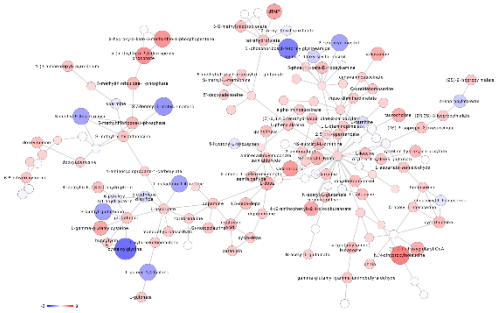
Biological question



Feature extraction



Statistical analysis



Metabolic pathways
Data integration

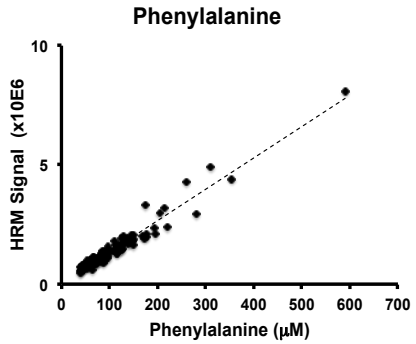
Li et al. 2016.
Biol Res Nurs. 18:12

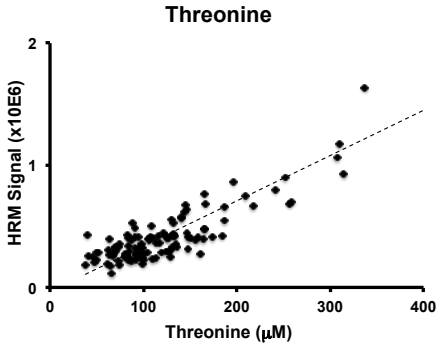
Metabolomics technologies / platforms

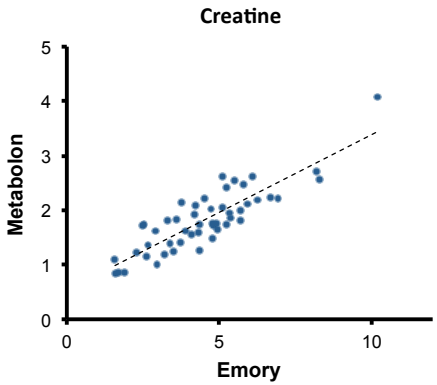
- NMR vs Mass spectrometry
- Targeted vs untargeted Metabolomics
- LC vs GC
- MS vs MS^n
- IM, EC, …
- Isotope tracing

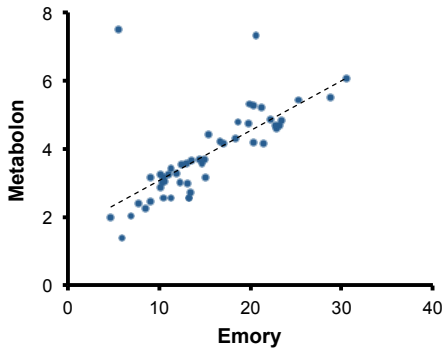
Accurate quantification in high-throughput mode shown by cross-laboratory comparisons

HERCULES
Exposome Research Center







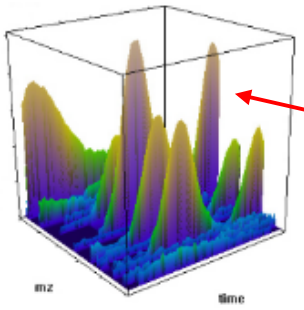


HRM metabolite quantification in 30 orphan samples

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

Slide courtesy of Prof. Dean Jones

Metabolomics data (LC-MS)



mass/charge (m/z)

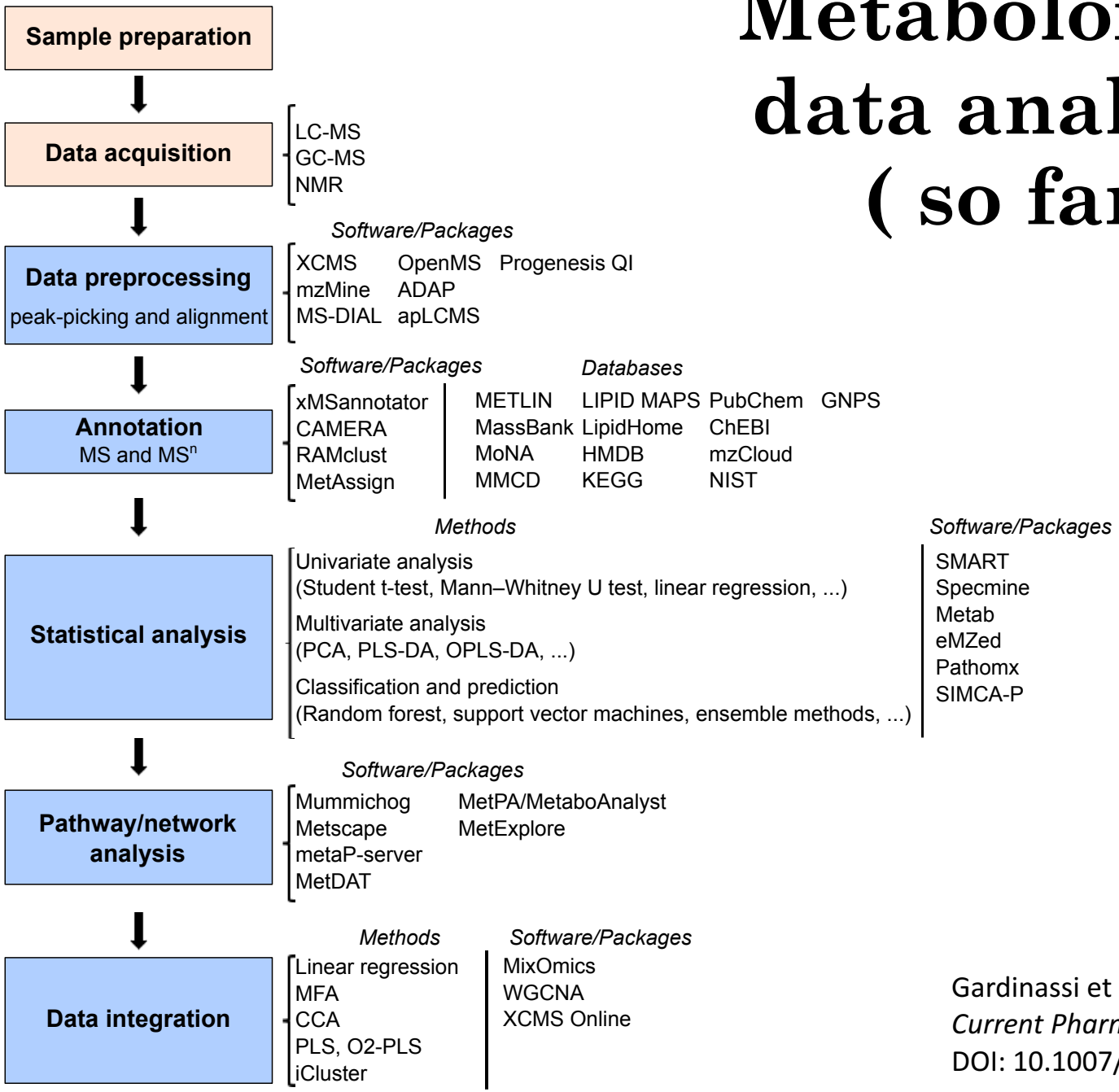
retention time

intensity

feature

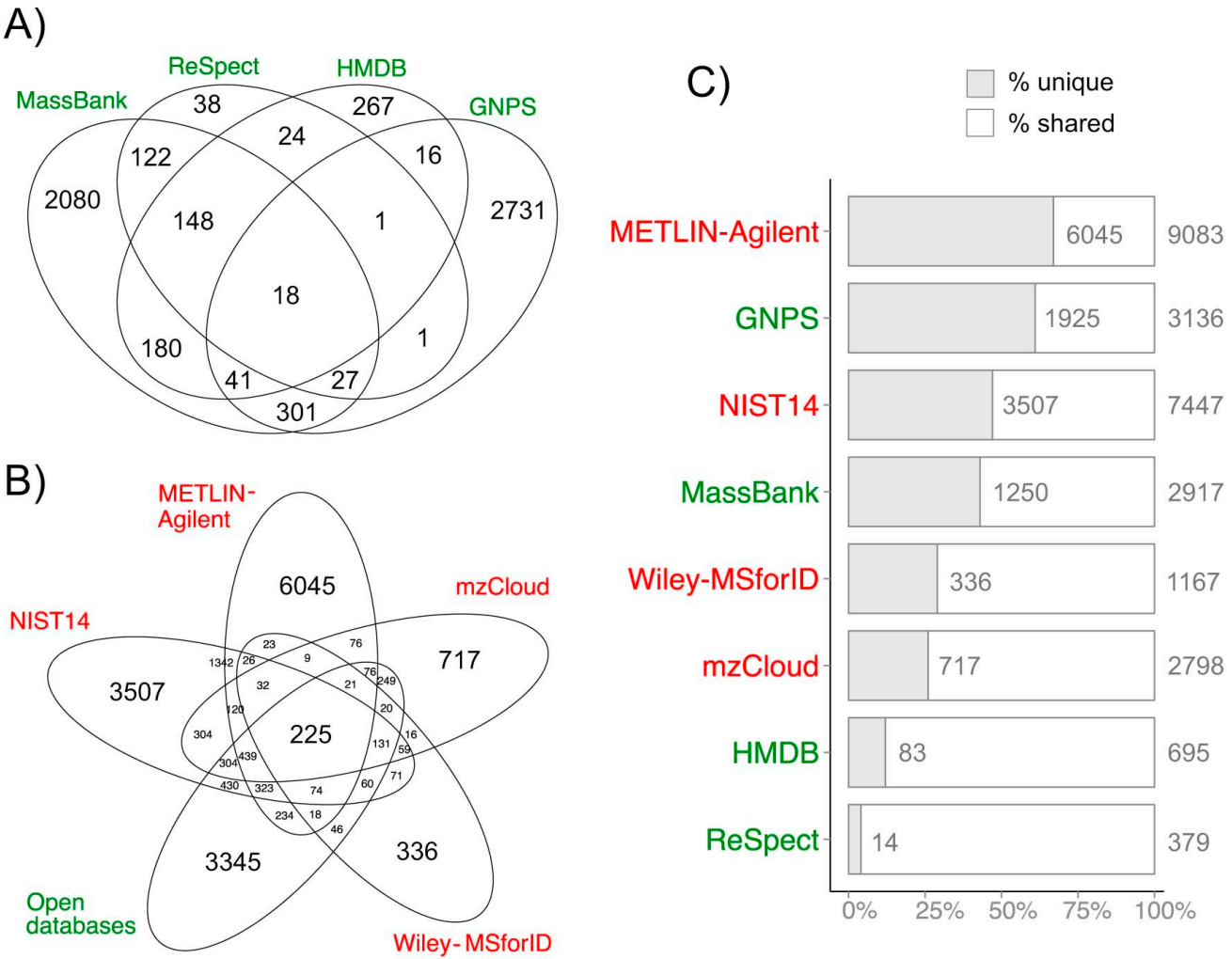
m/z	Retention time	Sample_1	Sample_2	Sample_3
702.86649413	70.4811138781	56905.1184820213	78201.4457063353	24695.6510487344
715.4970731384	56.55139775	83341.1725499997	1085.6171033898	0
715.6080685464	57.1626518527	93484.5395499998	0	0
722.705208561	65.2698532639	55396.5458522493	57091.7796614066	110944.335033457
730.7388823899	66.8110957481	60206.9068679558	62566.9945317202	88473.0389390921
732.7358893181	67.2539834098	95115.557321815	121927.272992316	125077.730263205
734.7314651732	69.0473678869	18862.86035	18500.1329562712	35179.9382
740.7679844539	66.5585431864	104536.742757076	143943.802149299	91120.9347911557
742.7651707472	67.4809438574	103257.45203488	134695.193645035	104576.322126369
750.7969865893	67.5681700041	117046.427960669	107939.940388836	73107.5173642583
752.793714331	67.3189767434	60630.5403168073	63435.6029427167	34751.816765529
760.8266370588	68.7380703739	41461.8047407238	54465.3458564024	28357.2944064846
102.0663370112	95.4437104918	19665.4401722973	101522.479348442	102707.039800486
112.3625162102	324.2816800135	151985.843922093	17555.644	57954.9573157841
121.072097577	88.0562910071	38408.6990506757	81408.9040789102	96405.1045547668
136.0486704167	74.5368317344	55806.64625	14052.7744	0

Metabolomics data analysis (so far)



Gardinassi et al (2017)
Current Pharmacology Reports,
DOI: 10.1007/s40495-017-0107-0.

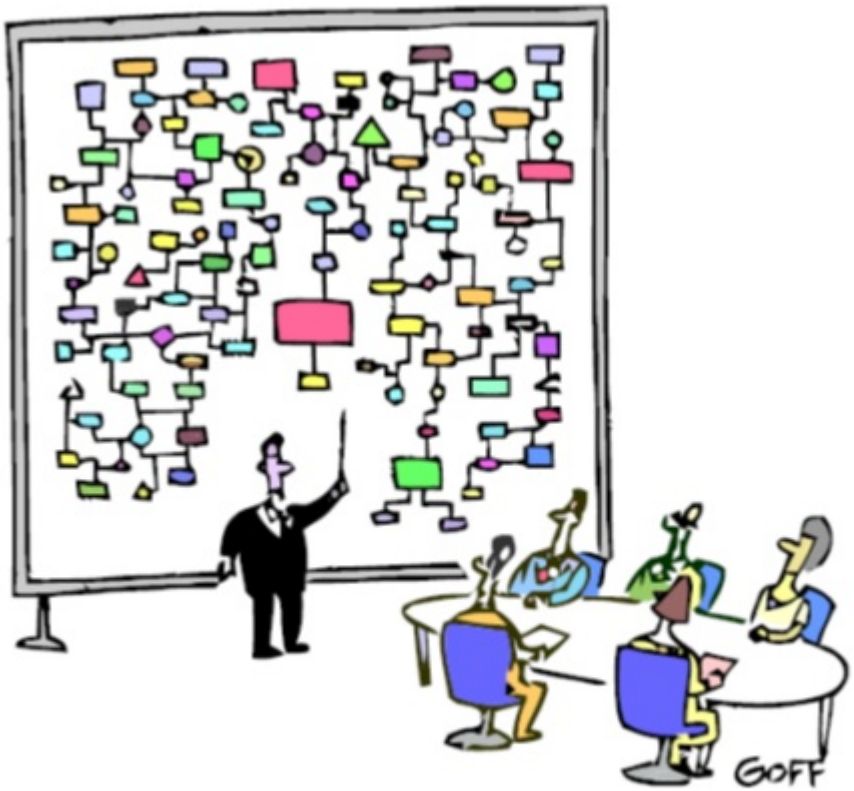
Mass Databases (2016)



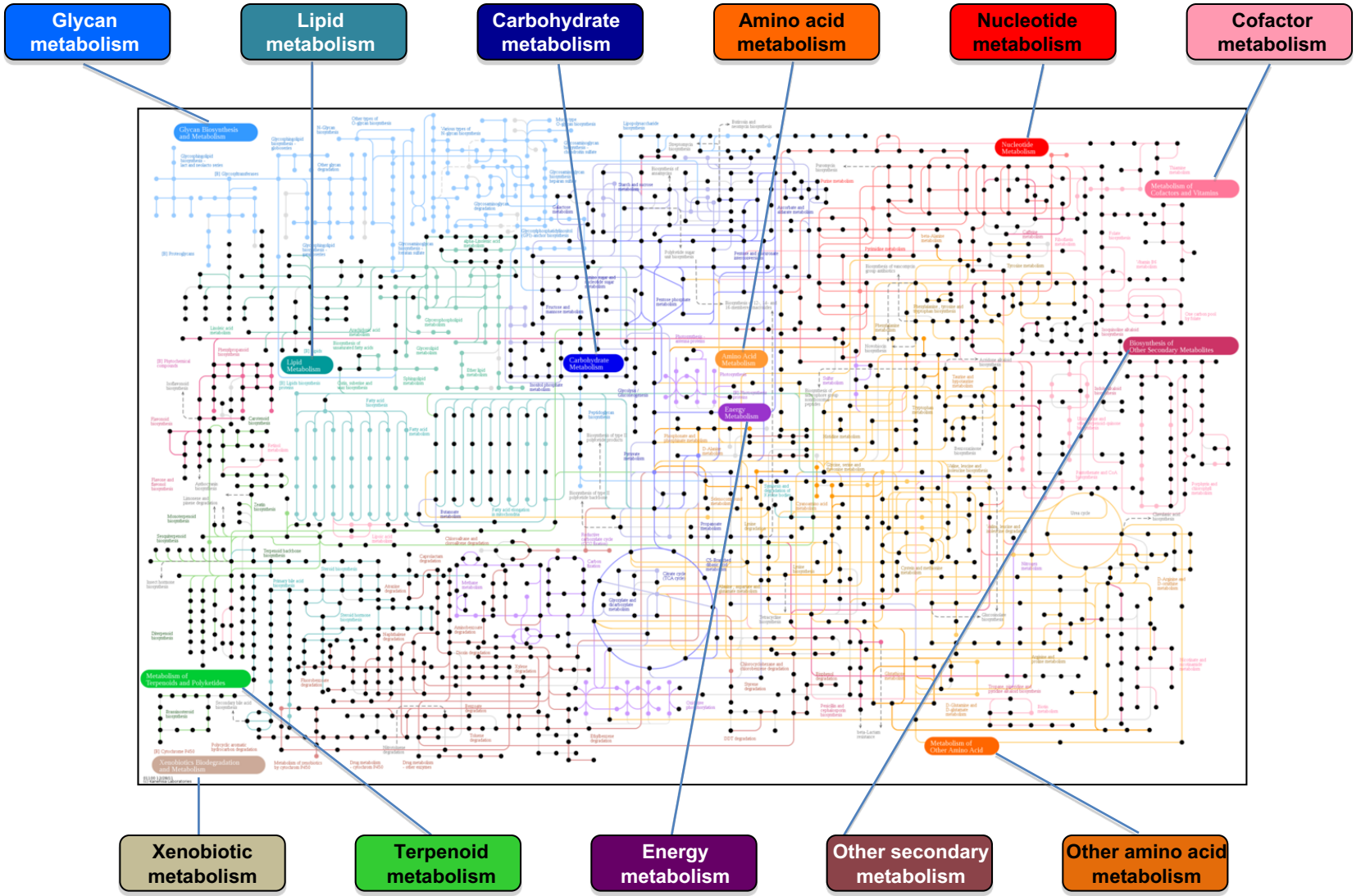
(A) Venn diagram showing the overlap between open mass spectral databases (HMDB, MassBank, GNPS, and ReSpect). (B) Venn diagram showing the overlap between five commercial databases (Agilent METLIN PMD, mzCloud, NIST 14, and Wiley MS) and open databases described in A. (C) Number and percentage of unique and shared compounds (i.e., InChIKey) with MS^n ($n \geq 2$) data in each database in relation to all eight resources.

Vinaixa, M., Schymanski, E.L., Neumann, S., Navarro, M., Salek, R.M. and Yanes, O., 2016. Mass spectral databases for LC/MS-and GC/MS-based metabolomics: state of the field and future prospects. *TrAC Trends in Analytical Chemistry*, *78*, pp.23-35.

Lot of pathways are known



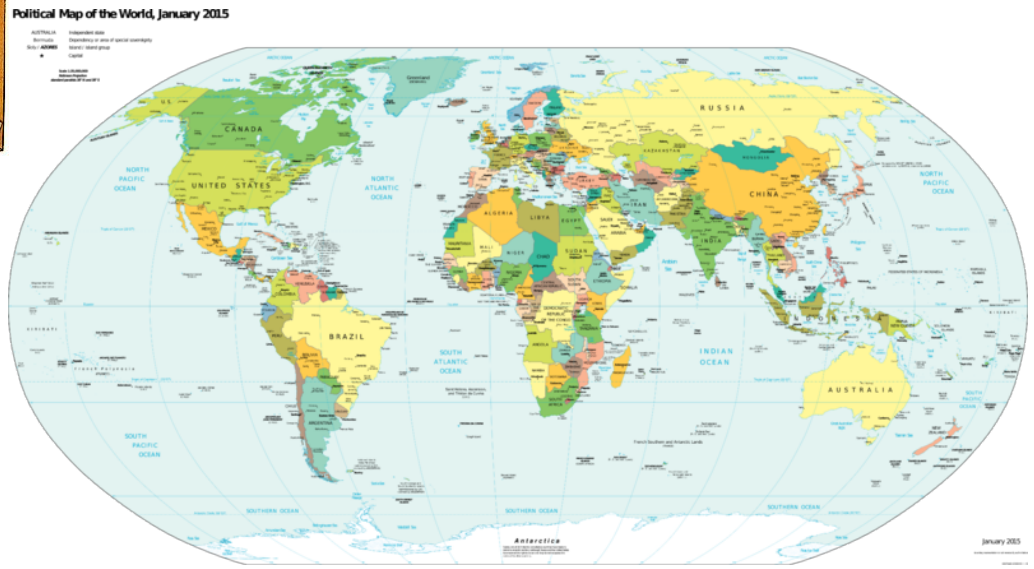
"And that's why we need a computer."



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

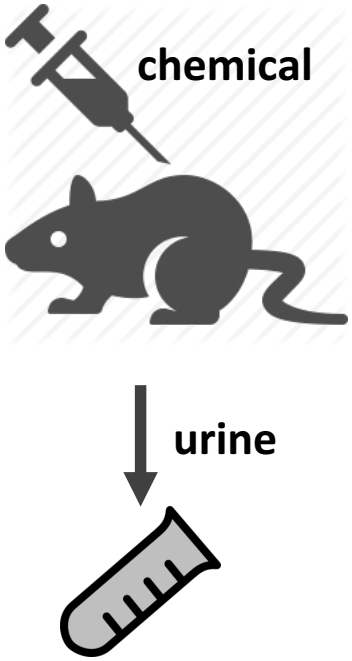
Towards new metabolic maps

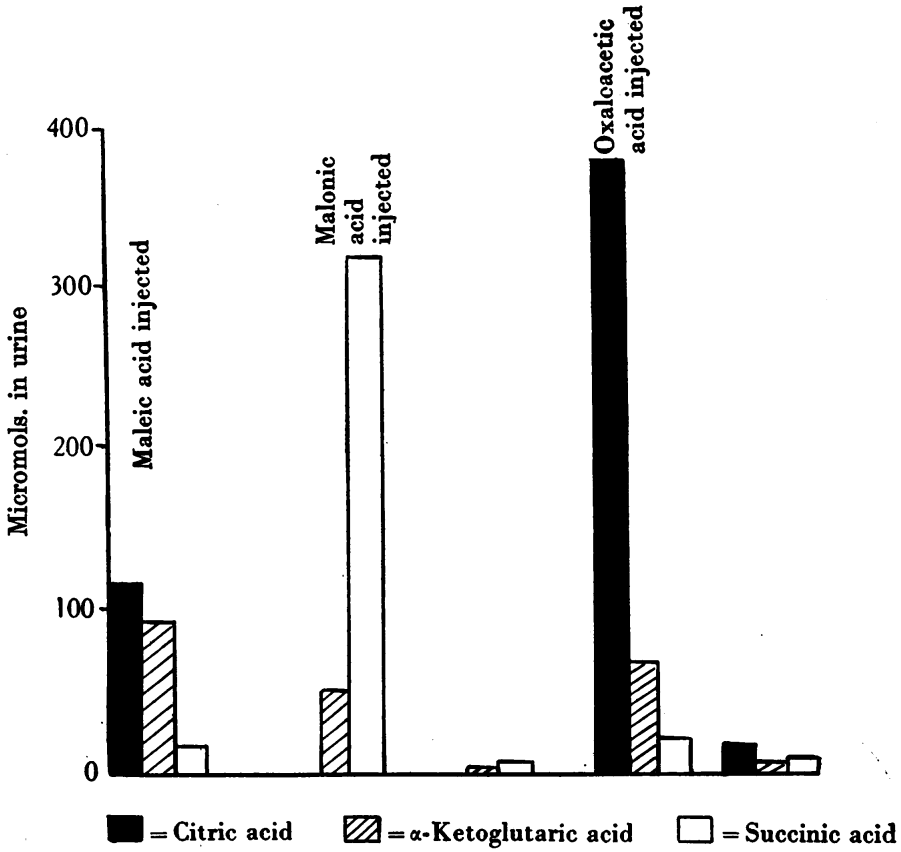
1482

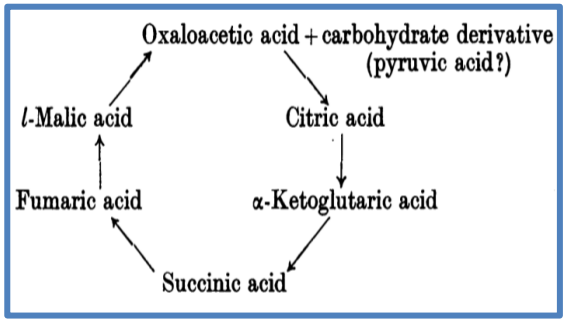


2015

Charting metabolic pathways

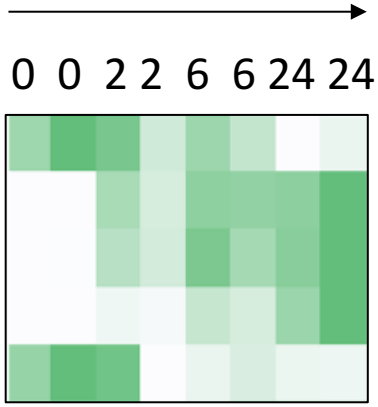




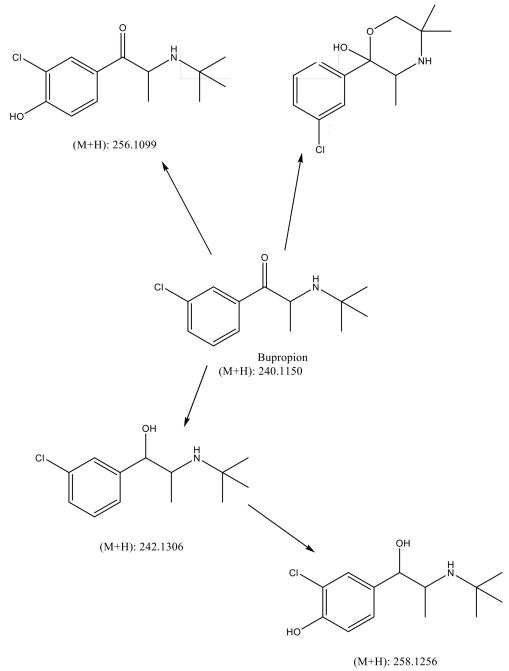


Krebs et al. 1938.
Biochem Journal. 32:113

Charting metabolic pathways (2019)

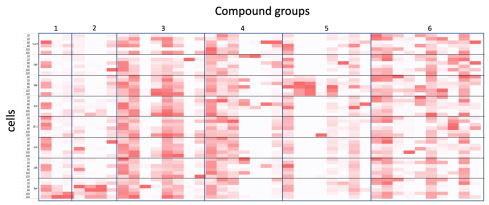


Bupropion
Hydroxybupropion
Desaturated bupropion
Erythro/threohydrobupropion
?



Jones/Li/Morgan/Miller
NIEHS U2C ES030163

Charting metabolic pathways



- Capturing unknown; tandem mass spec; isotope tracing in cells and whole organisms
- Combination with genetic tools
- Domain applications, tumors, microbiome, immune cells
- Azimuth database; Interfacing with computational reconstruction of biochemical networks

Objectives

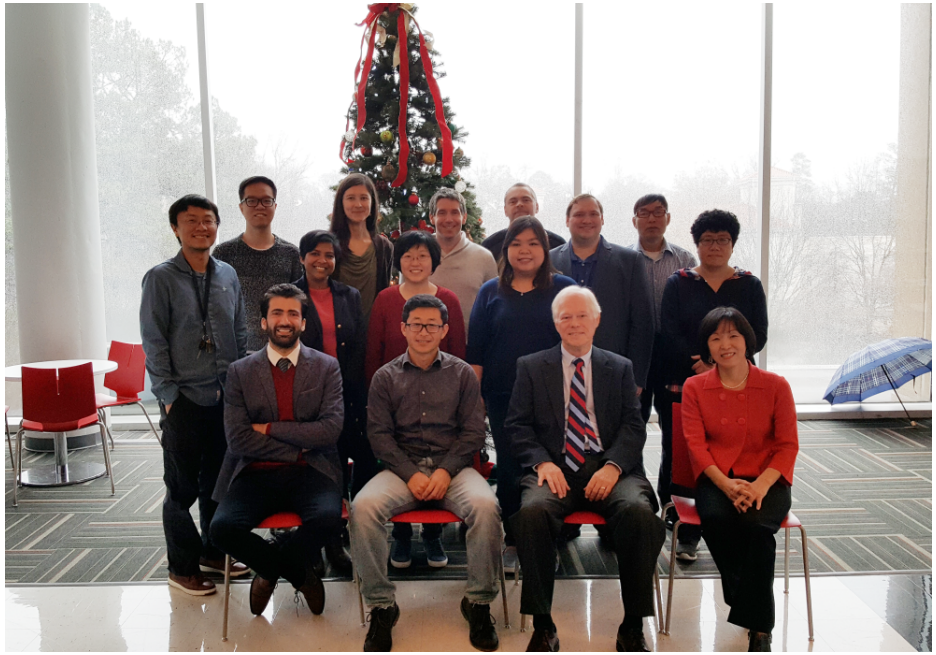
- Understanding of experimental design and statistics
- Using Jupyter notebooks for data analysis
- Web tools such as XCMS Online, MetaboAnalyst and Mummichog server
- Metabolite reporting standards
- Discussion of research projects

Acknowledgement

CHDS
Barbara Cohn
Piera Cirrilo
Nickilou Krigbaum
Stanford University
Bali Pulendran
University of Colorado
Myron Levin
Adriana Weinberg

Emory University
Rafi Ahmed
Mark Mulligan (NYU)
Nadine Rouphael
Mary Galinski
Judith Fridovich-Keil
Taylor Fischer
Miriam Vos

Dean P. Jones
Young-Mi Go
Douglas Walker (MSSM)
ViLinh Tran
Bill Liang
Luiz Gardinassi
Ken Liu
M. Ryan Smith
Tianwei Yu
Edward Morgan
Choon Myung Lee
Grant Singer



Funding support from US National Institutes of Health U19 AI090023, HHSN272201300018I, UH2 AI132345, U01 CA235493, U2C ES030163, P30ES019776, U2C ES026560, P50 ES026071, R01 DK107900, R01 AI121252, R01 GM124061, and EPA (83615301)