The Future of Metabolomics

David Wishart

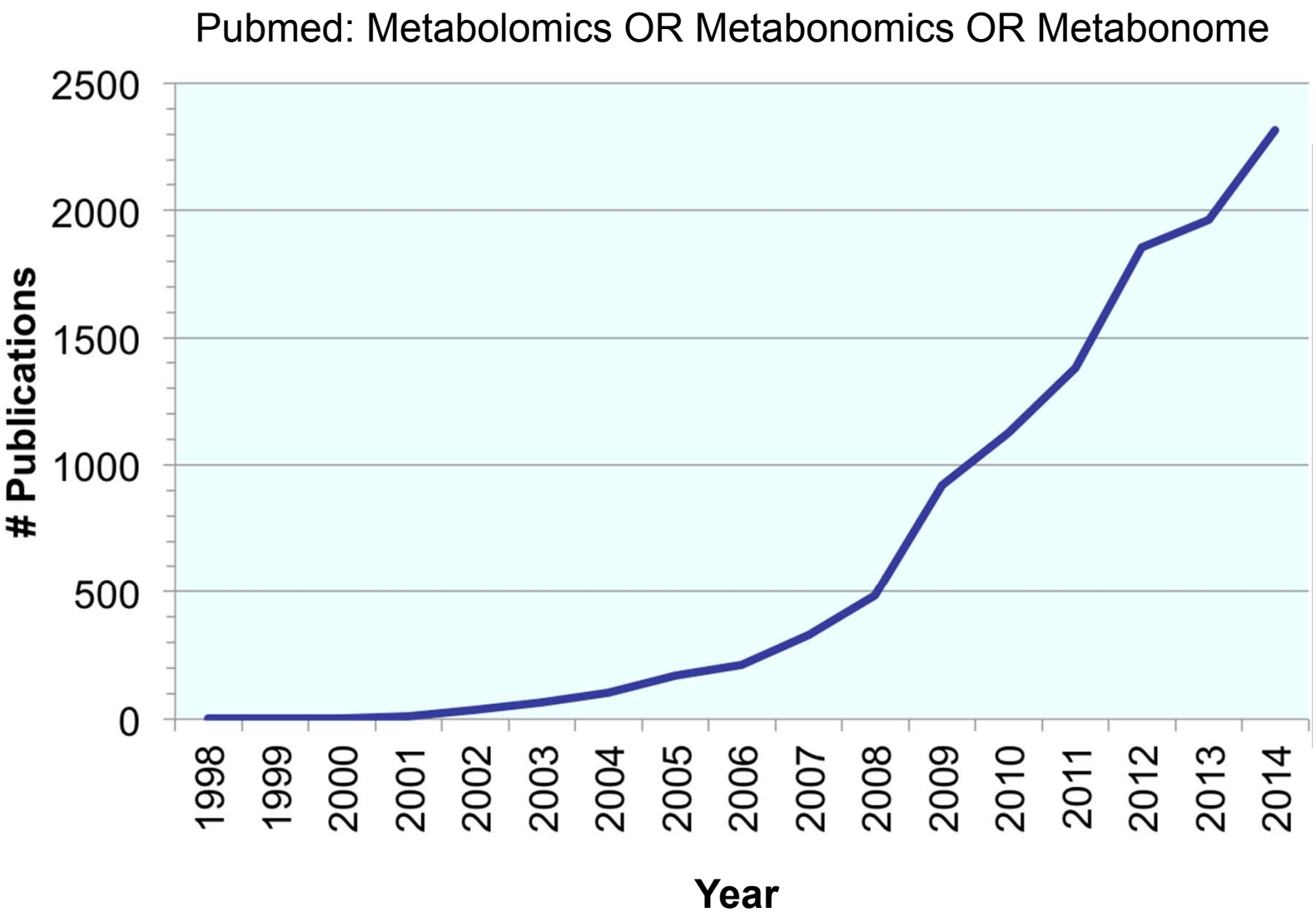
University of Alberta, Edmonton, AB, Canada

Birmingham, Alabama June 18, 2015

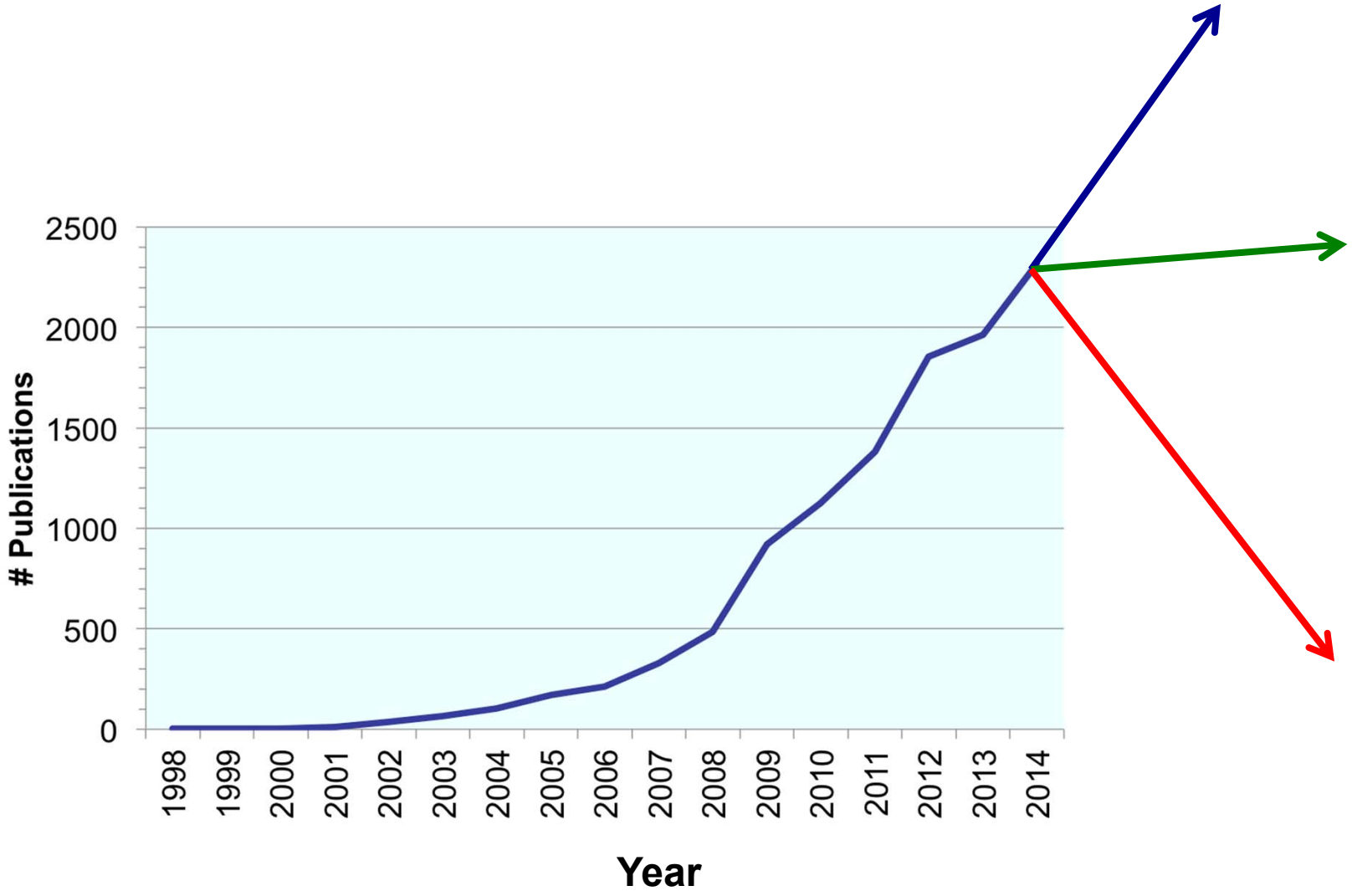
Why Small Molecules Matter

- >99% of food flavours and aromas come from small molecules
- >90% of common clinical tests measure small molecules
- 89% of all drugs are small molecules
- 83% of the most common diseases are due to the effects of small molecules
- 81% of all deaths in North America and Europe are due to the effects of small molecules
- >55% of drugs are derived via natural cmpds

Metabolomics Is Growing



But Where Is It Going?



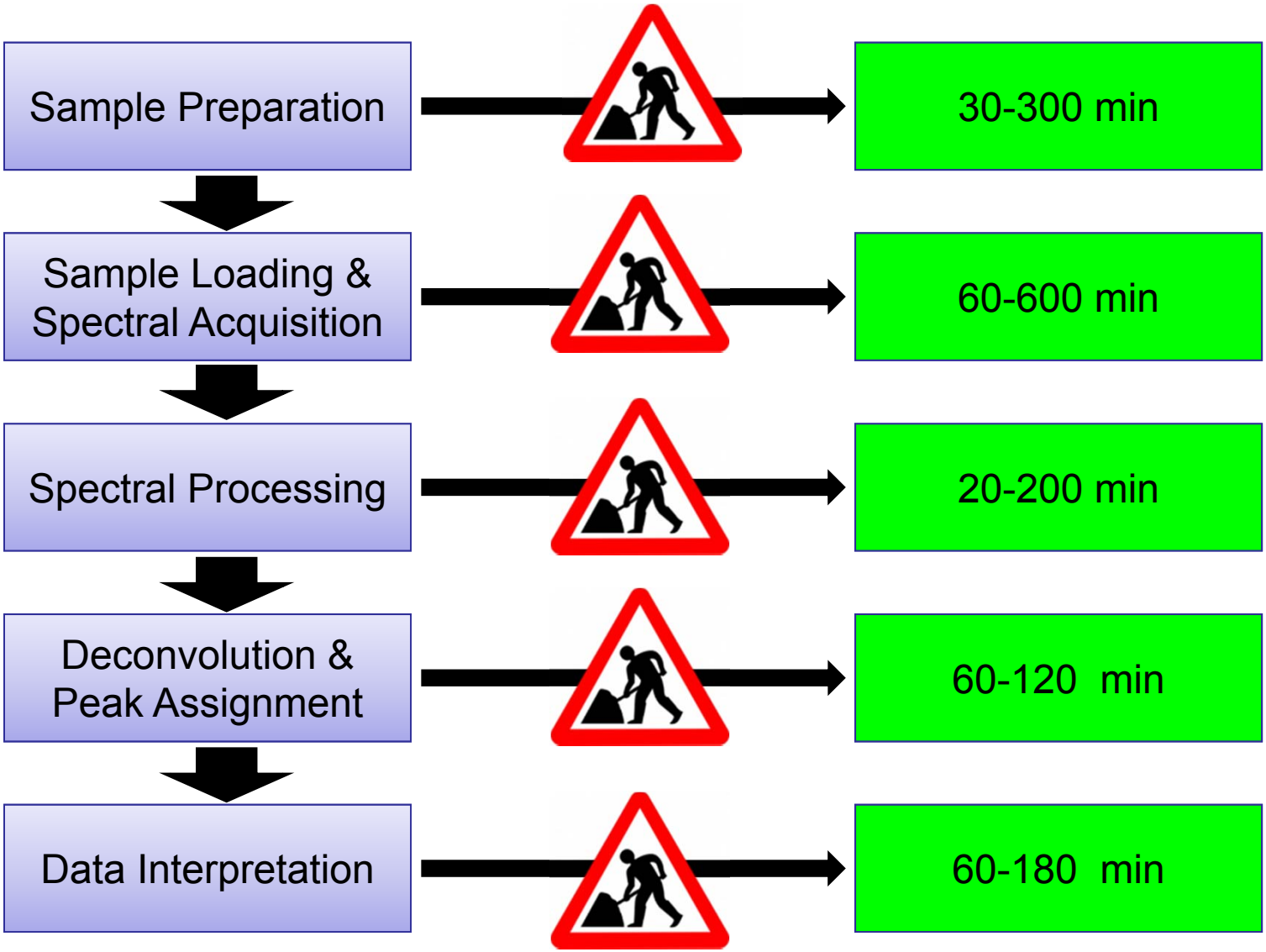
Key Bottlenecks in Metabolomics

- Lack of automation
- Incomplete metabolome coverage
- Expensive/large equipment
- Lack of quantification
- Inability to translate findings to the clinic
- Making metabolomics matter to drug companies

Key Trends in Metabolomics

- **Automated metabolomics**
- **Expanding metabolome coverage**
- **Making metabolomics portable**
- **Quantify, quantify, quantify…**
- **Moving metabolomics from the lab to the clinic**
- **Moving metabolomics (back) into drug development and discovery**

Metabolomics Workflow



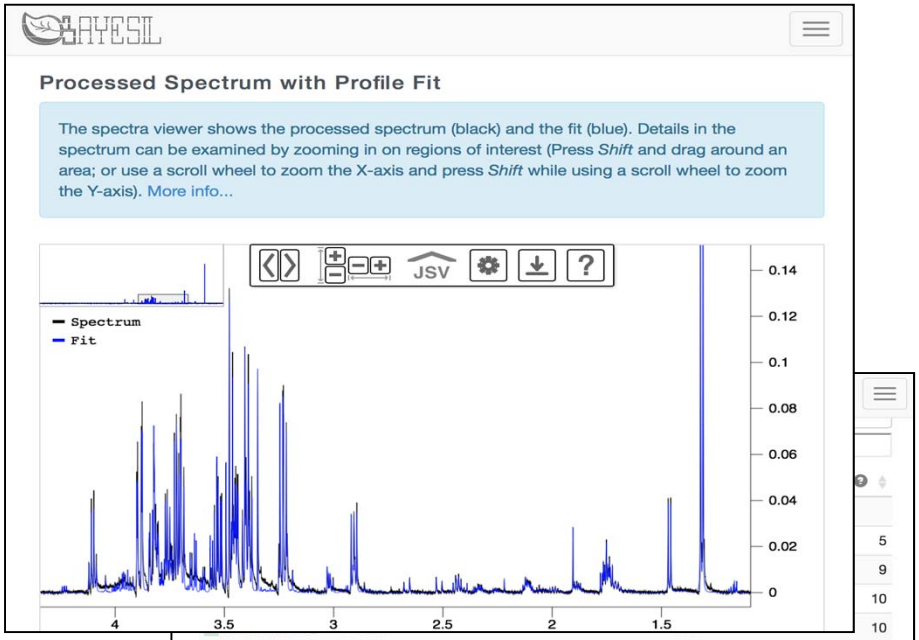
Automated Metabolomics

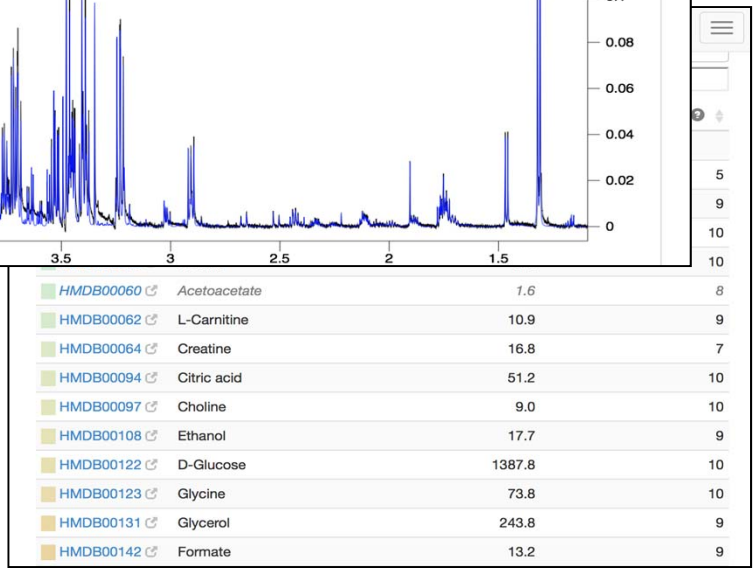
Bruker – Automated NMR

Sciex – Lipidyzer

Biocrates – Automated MS

Bayesil (Automated NMR)



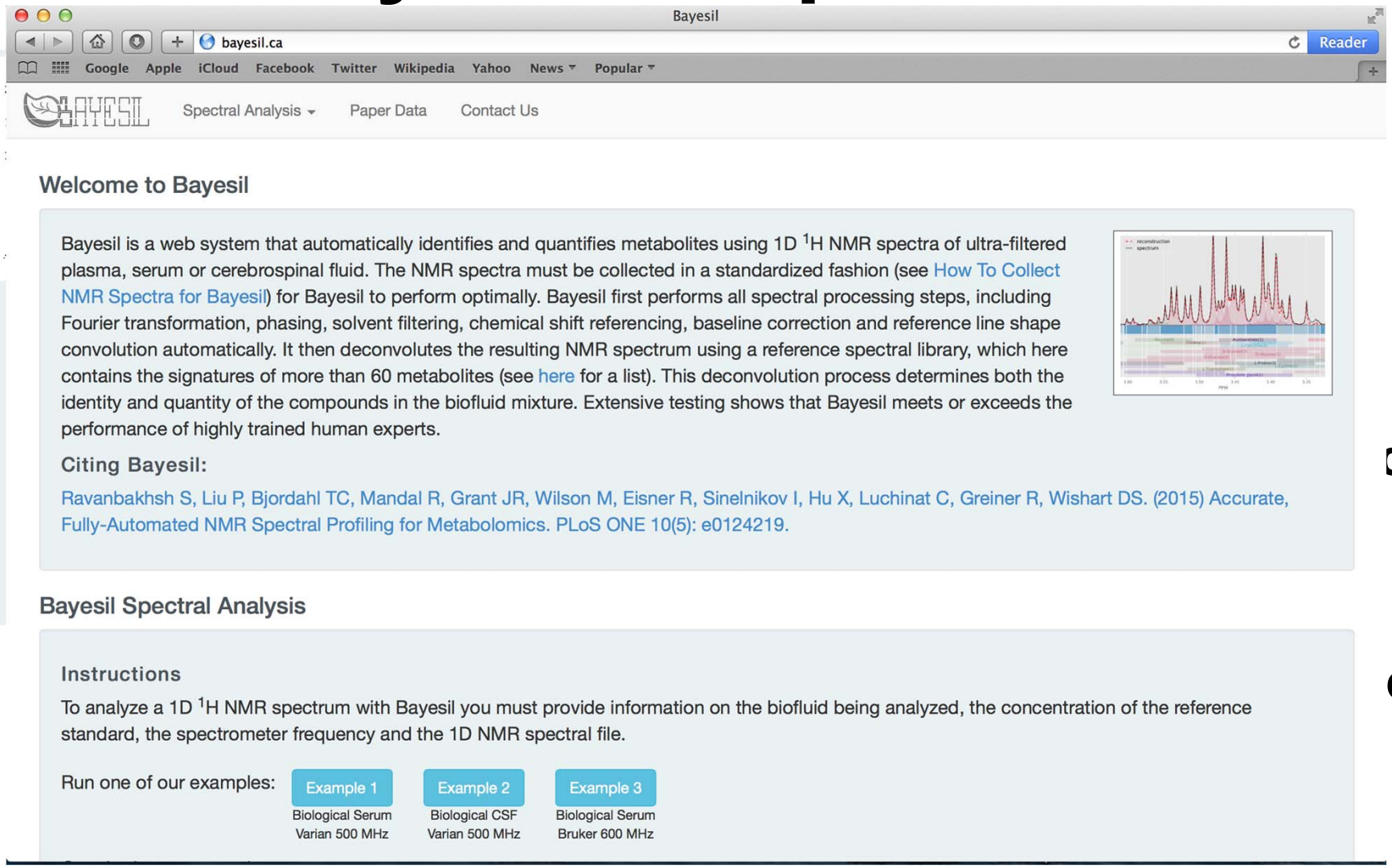


HMDB00060	Acetoacetate	1.6	8
HMDB00062	L-Carnitine	10.9	9
HMDB00064	Creatine	16.8	7
HMDB00094	Citric acid	51.2	10
HMDB00097	Choline	9.0	10
HMDB00108	Ethanol	17.7	9
HMDB00122	D-Glucose	1387.8	10
HMDB00123	Glycine	73.8	10
HMDB00131	Glycerol	243.8	9
HMDB00142	Formate	13.2	9

http://bayesil.ca

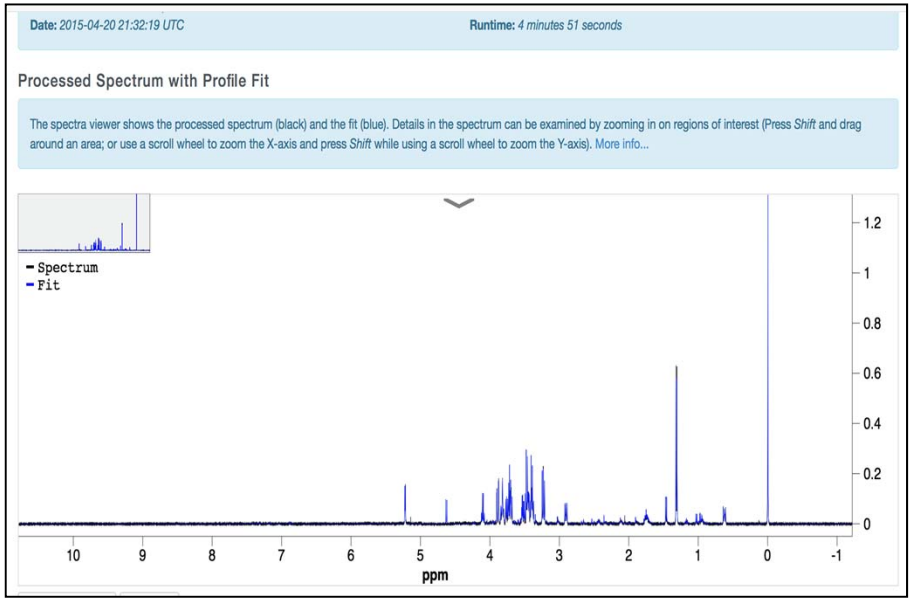
- Uses probabilistic graphical models (PGM) – similar to HMMs
- Fits shift & peak intensity similar to the way humans perform fitting and pattern finding
- Requires prior knowledge of probable biofluid composition
- Fully automated phasing, referencing, water removal, baseline correction, peak convolution, identification and quantification
- Free web server

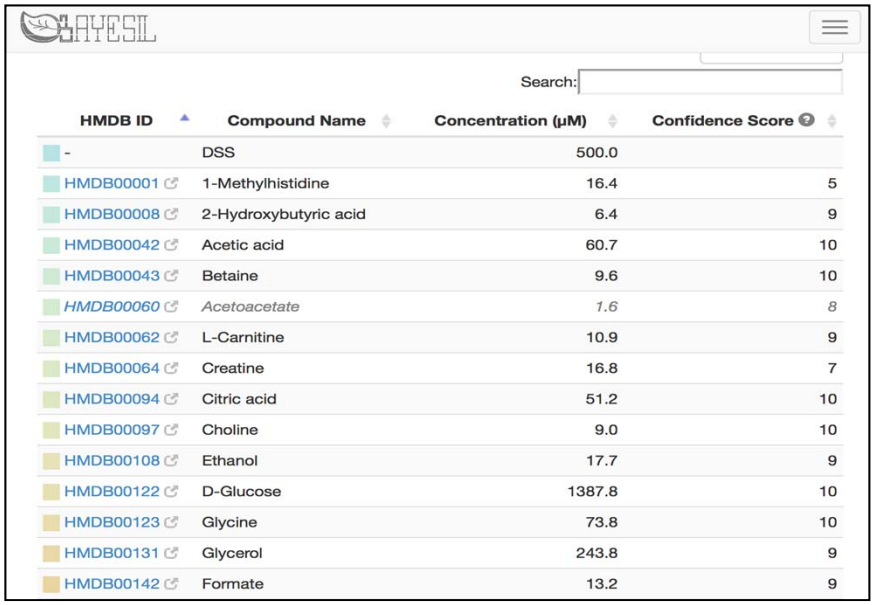
Bayesil in Operation



S. Ravanbakhsh, et al. (2015) PLoS One 10(5): e0124219

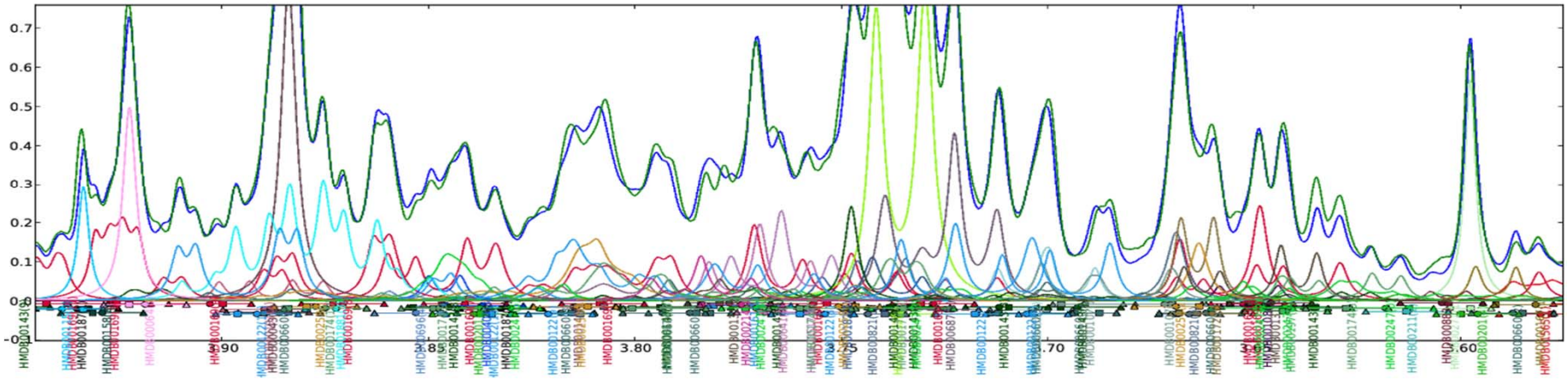
Bayesil Output





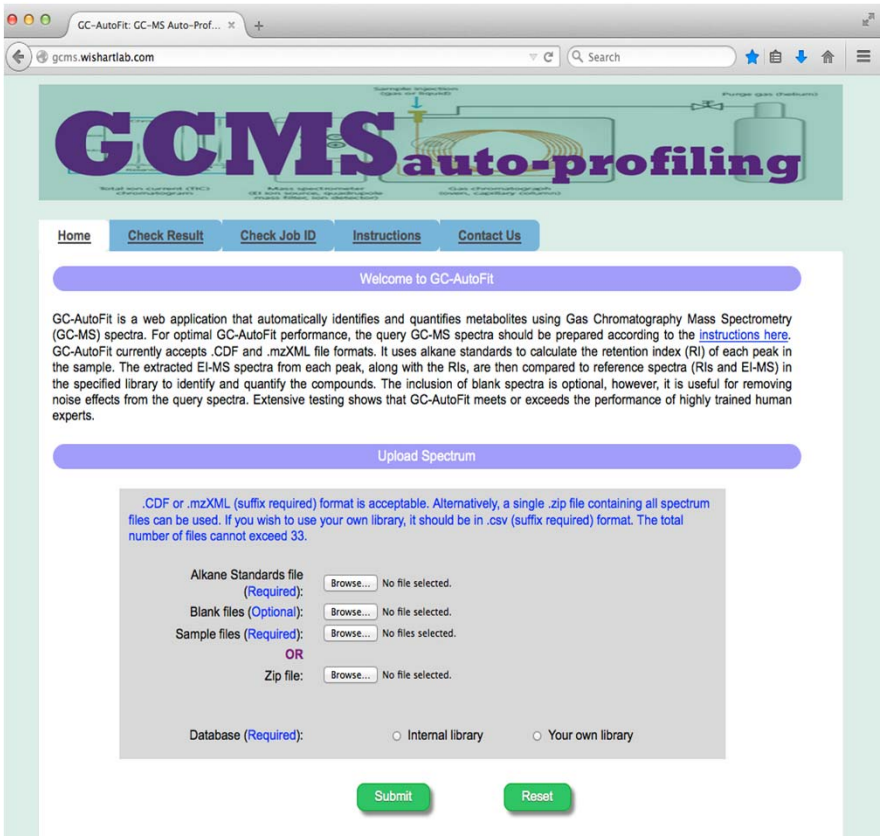
Search:

HMDB ID	Compound Name	Concentration (μM)	Confidence Score
-	DSS	500.0	
HMDB00001	1-Methylhistidine	16.4	5
HMDB00008	2-Hydroxybutyric acid	6.4	9
HMDB00042	Acetic acid	60.7	10
HMDB00043	Betaine	9.6	10
HMDB00060	*Acetoacetate*	*1.6*	*8*
HMDB00062	L-Carnitine	10.9	9
HMDB00064	Creatine	16.8	7
HMDB00094	Citric acid	51.2	10
HMDB00097	Choline	9.0	10
HMDB00108	Ethanol	17.7	9
HMDB00122	D-Glucose	1387.8	10
HMDB00123	Glycine	73.8	10
HMDB00131	Glycerol	243.8	9
HMDB00142	Formate	13.2	9



S. Ravanbakhsh, et al. (2015) PLoS One 10(5): e0124219

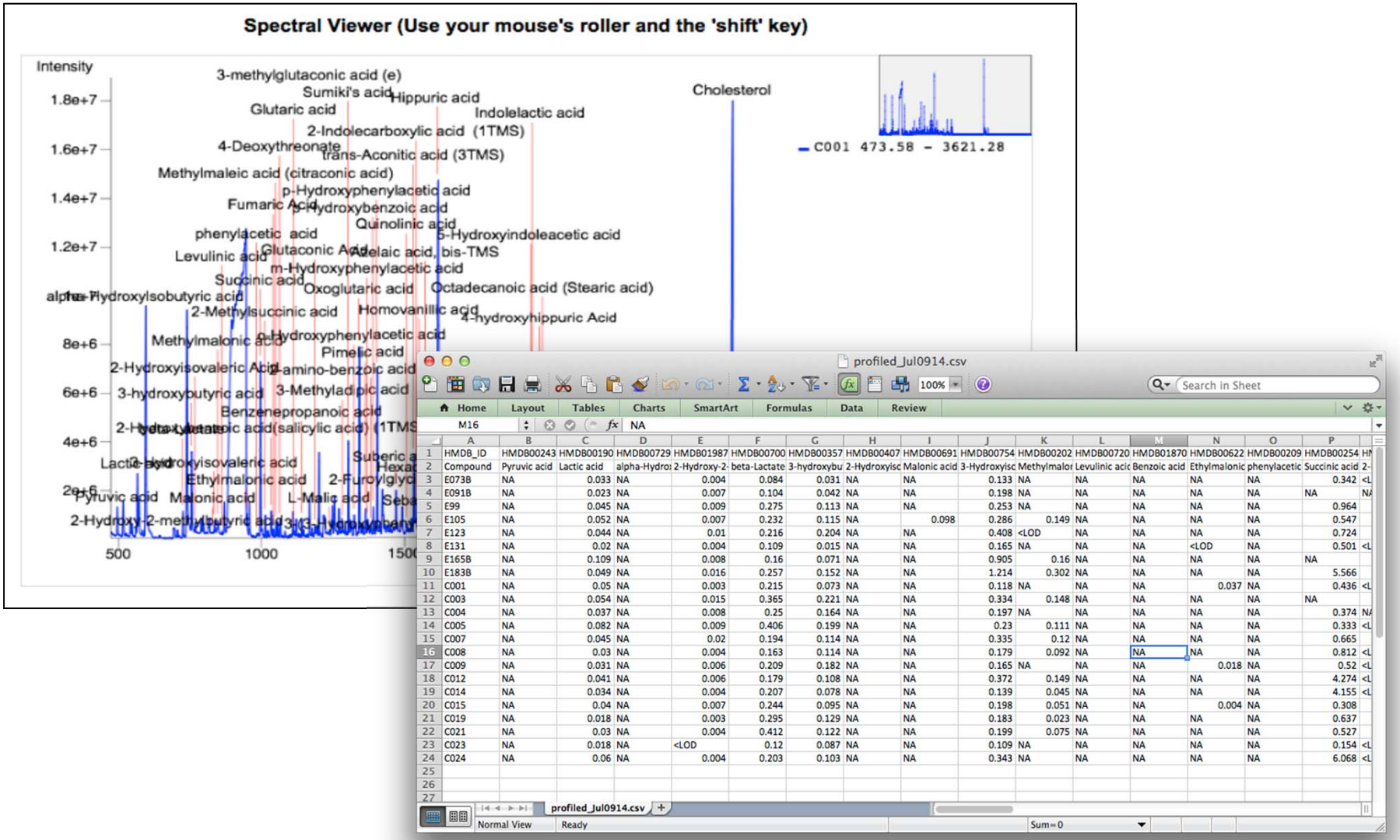
GC-AutoFit (Automated GC-MS)



http://gcms.wishartlab.com

- Requires 3 spectra (sample, blank, alkane standards)
- Performs auto-alignment, peak ID, peak integration and concentration calculation
- Accepts NetCDF or mzXML files
- 60 sec per spectrum
- 45-70 cmpds ID'd and quantified, 96% accuracy
- Optimized for blood, urine, saliva and CSF
- Still requires careful sample preparation & derivatization

GC-AutoFit Output

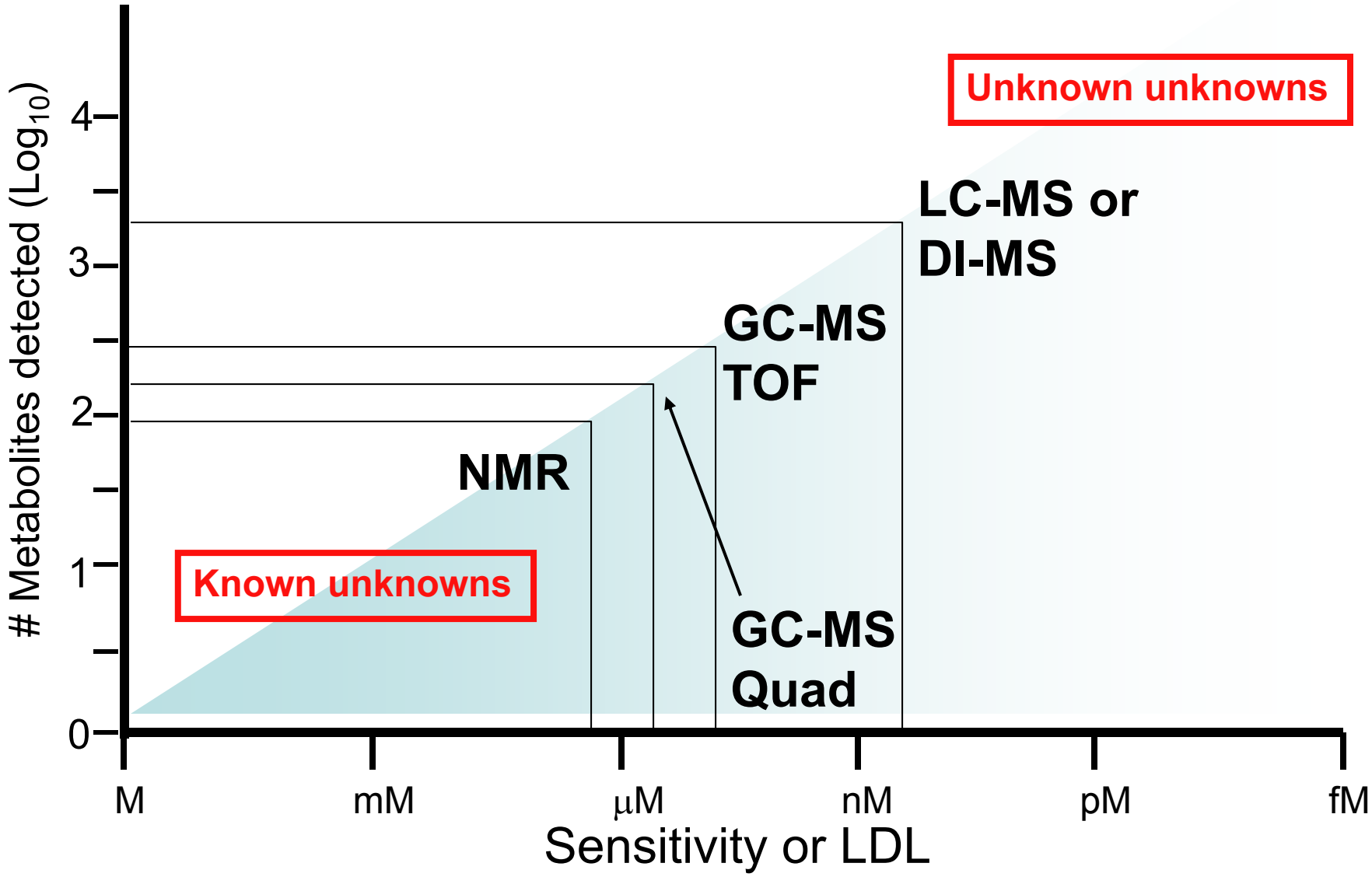


	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	HMDB_ID	HMDB00243	HMDB00190	HMDB00729	HMDB01987	HMDB00700	HMDB00357	HMDB00407	HMDB00691	HMDB00754	HMDB00202	HMDB00720	HMDB01870	HMDB00622	HMDB00209	HMDB00254
2	Compound	Pyruvic acid	Lactic acid	alpha-Hydro	2-Hydroxy-2-	beta-Lactate	3-hydroxybu	2-Hydroxyisc	Malonic acid	3-Hydroxyisc	Methylmalor	Levulinic acic	Benzoic acid	Ethylmalonic	phenylacetic	Succinic acid
3	E073B	NA	0.033	NA	0.004	0.084	0.031	NA	NA	0.133	NA	NA	NA	NA	NA	0.342
4	E091B	NA	0.023	NA	0.007	0.104	0.042	NA	NA	0.198	NA	NA	NA	NA	NA	NA
5	E99	NA	0.045	NA	0.009	0.275	0.113	NA	NA	0.253	NA	NA	NA	NA	NA	0.964
6	E105	NA	0.052	NA	0.007	0.232	0.115	NA	0.098	0.286	0.149	NA	NA	NA	NA	0.547
7	E123	NA	0.044	NA	0.01	0.216	0.204	NA	NA	0.408	<LOD	NA	NA	NA	NA	0.724
8	E131	NA	0.02	NA	0.004	0.109	0.015	NA	NA	0.165	NA	NA	NA	<LOD	NA	0.501
9	E165B	NA	0.109	NA	0.008	0.16	0.071	NA	NA	0.905	0.16	NA	NA	NA	NA	NA
10	E183B	NA	0.049	NA	0.016	0.257	0.152	NA	NA	1.214	0.302	NA	NA	NA	NA	5.566
11	C001	NA	0.05	NA	0.003	0.215	0.073	NA	NA	0.118	NA	NA	NA	0.037	NA	0.436
12	C003	NA	0.054	NA	0.015	0.365	0.221	NA	NA	0.334	0.148	NA	NA	NA	NA	NA
13	C004	NA	0.037	NA	0.008	0.25	0.164	NA	NA	0.197	NA	NA	NA	NA	NA	0.374
14	C005	NA	0.082	NA	0.009	0.406	0.199	NA	NA	0.23	0.111	NA	NA	NA	NA	0.333
15	C007	NA	0.045	NA	0.02	0.194	0.114	NA	NA	0.335	0.12	NA	NA	NA	NA	0.665
16	C008	NA	0.03	NA	0.004	0.163	0.114	NA	NA	0.179	0.092	NA	NA	NA	NA	0.812
17	C009	NA	0.031	NA	0.006	0.209	0.182	NA	NA	0.165	NA	NA	NA	0.018	NA	0.52
18	C012	NA	0.041	NA	0.006	0.179	0.108	NA	NA	0.372	0.149	NA	NA	NA	NA	4.274
19	C014	NA	0.034	NA	0.004	0.207	0.078	NA	NA	0.139	0.045	NA	NA	NA	NA	4.155
20	C015	NA	0.04	NA	0.007	0.244	0.095	NA	NA	0.198	0.051	NA	NA	0.004	NA	0.308
21	C019	NA	0.018	NA	0.003	0.295	0.129	NA	NA	0.183	0.023	NA	NA	NA	NA	0.637
22	C021	NA	0.03	NA	0.004	0.412	0.122	NA	NA	0.199	0.075	NA	NA	NA	NA	0.527
23	C023	NA	0.018	NA	<LOD	0.12	0.087	NA	NA	0.109	NA	NA	NA	NA	NA	0.154
24	C024	NA	0.06	NA	0.004	0.203	0.103	NA	NA	0.343	NA	NA	NA	NA	NA	6.068
25																
26																
27																

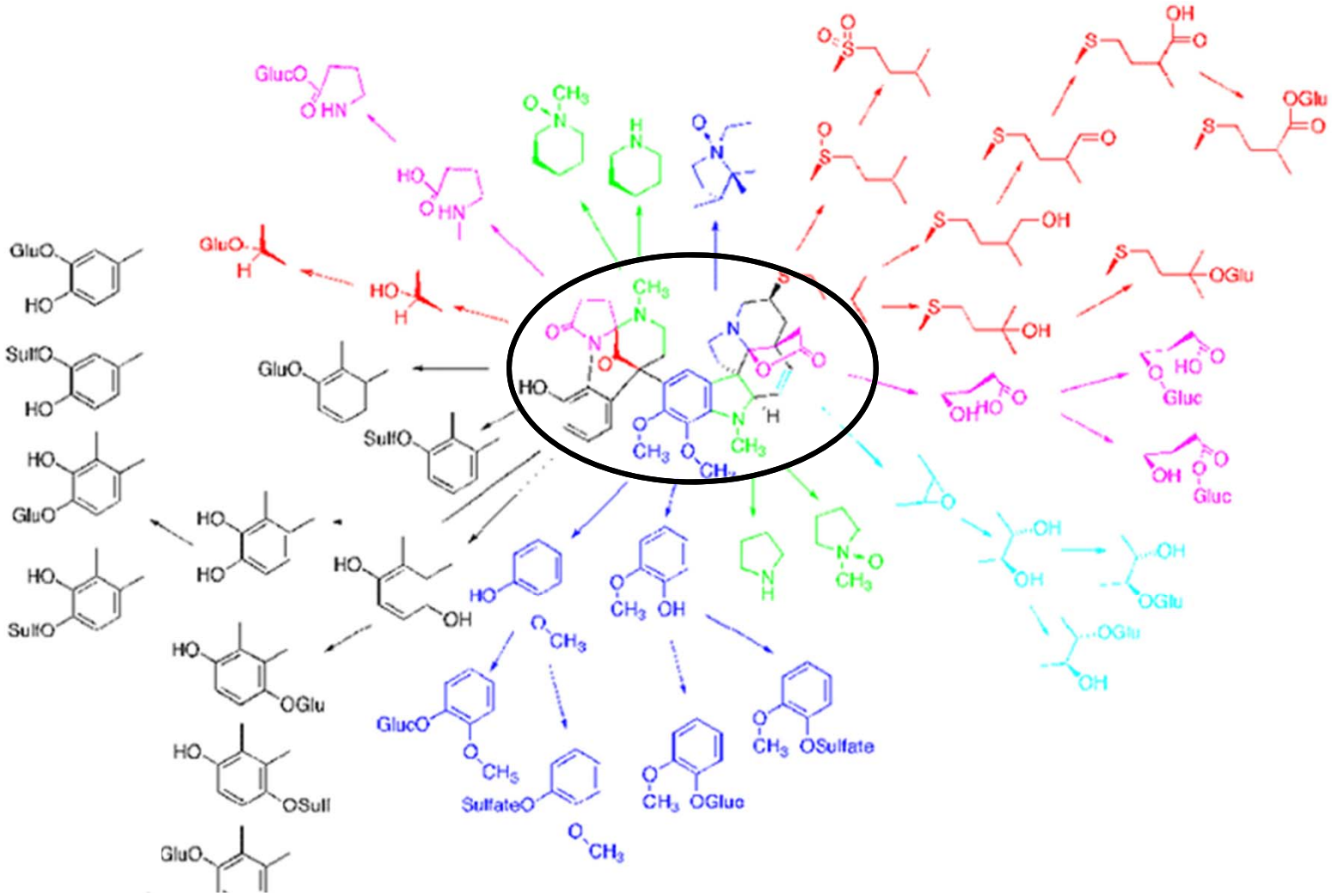
Key Trends in Metabolomics

- Automated metabolomics
- Expanding metabolome coverage
- Making metabolomics portable
- Quantify, quantify, quantify…
- Moving metabolomics from the lab to the clinic
- Moving metabolomics (back) into drug development and discovery

Technology & Sensitivity

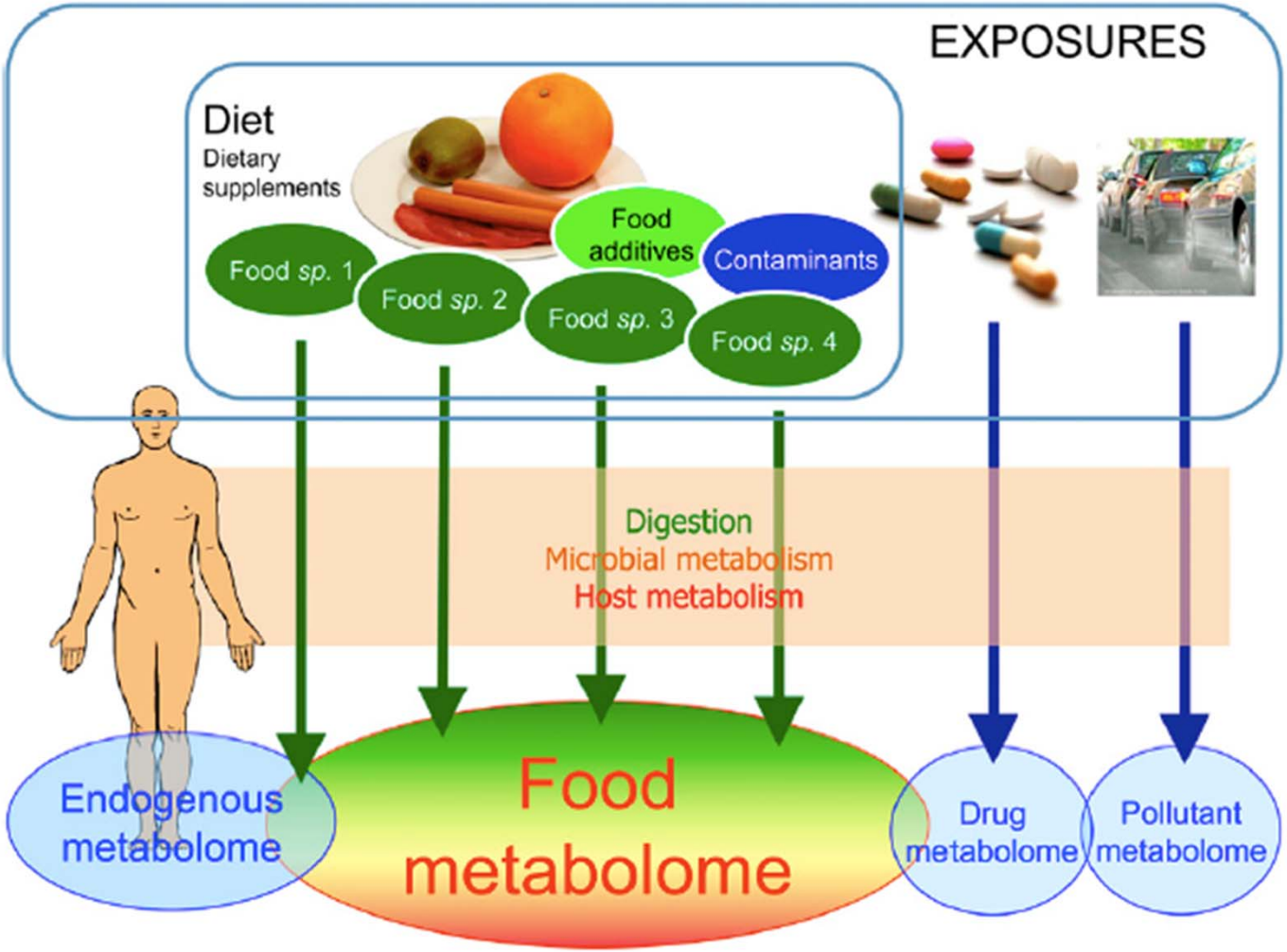


What Are The Unknown Unknowns?



Metabolites of Metabolites

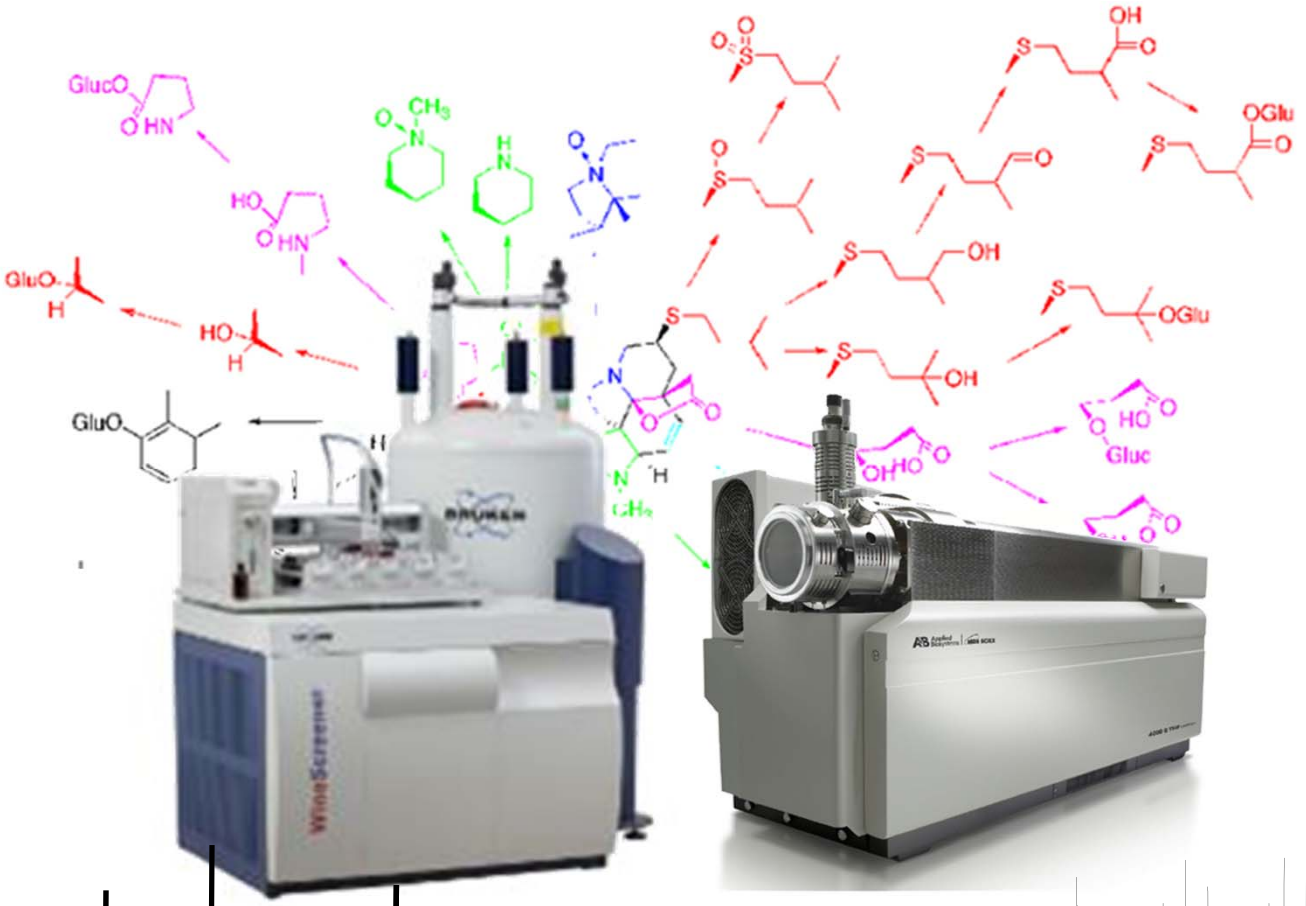
The Food Metabolome



Scalbert A. et al. (2014) Am. J. Clinical Nutr. 99(6):1286

Systematic Spectral Collection

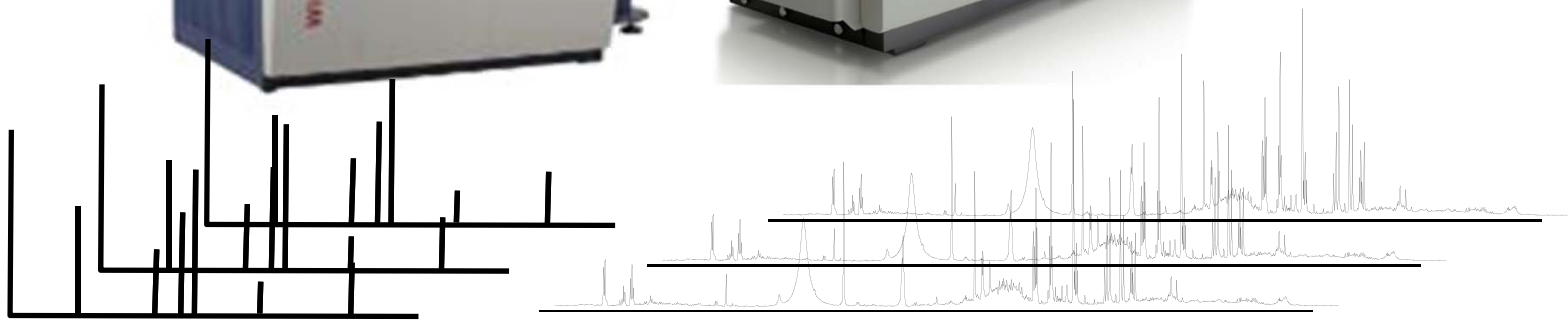
100,000+
Compounds



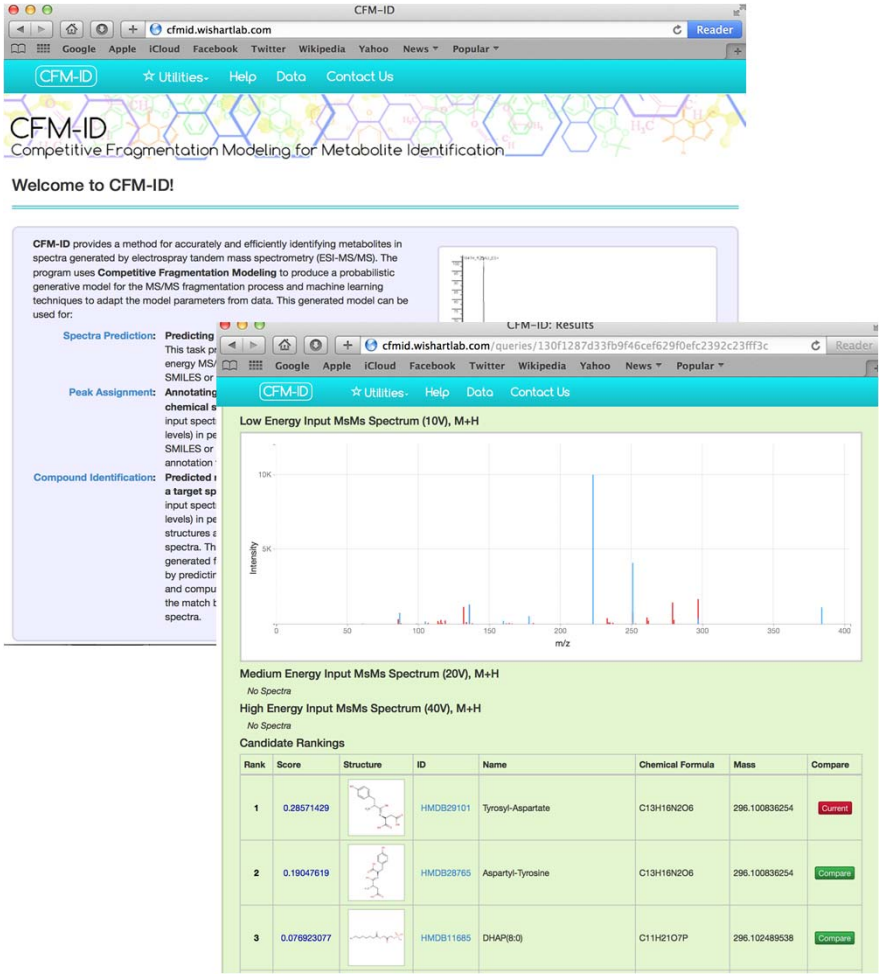
$$

$$

800,000+
NMR, MS/MS
Spectra



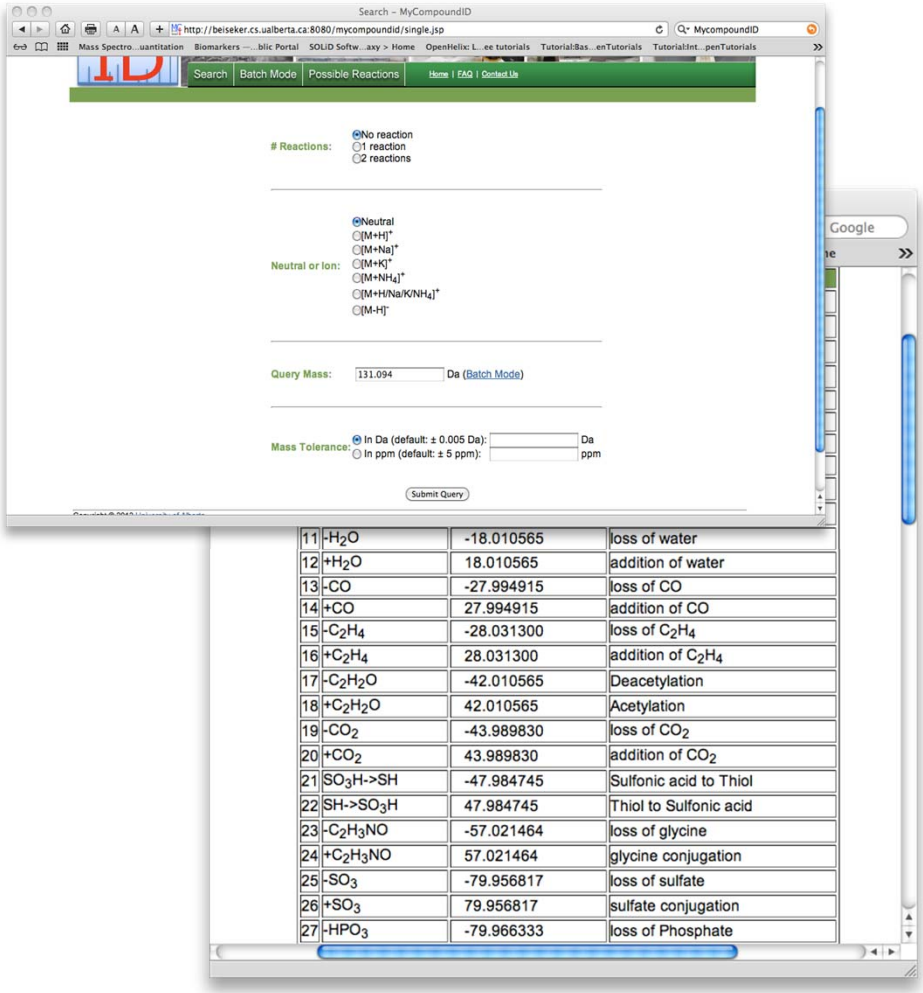
Systematic Spectral Prediction



http://cfmid.wishartlab.com

- Predicts MS/MS spectra from known compounds via advanced machine learning techniques
- 50% more accurate than other systems
- Matches predicted MS/MS spectra (from HMDB, KEGG or user choice) to input MS/MS spectra
- Permits rapid compound ID from MS/MS spectra

Predicting Metabolites of Metabolites



11	-H_2O	-18.010565	loss of water
12	+H_2O	18.010565	addition of water
13	-CO	-27.994915	loss of CO
14	+CO	27.994915	addition of CO
15	-C_2H_4	-28.031300	loss of C_2H_4
16	+C_2H_4	28.031300	addition of C_2H_4
17	-C_2H_2O	-42.010565	Deacetylation
18	+C_2H_2O	42.010565	Acetylation
19	-CO_2	-43.989830	loss of CO_2
20	+CO_2	43.989830	addition of CO_2
21	SO_3H->SH	-47.984745	Sulfonic acid to Thiol
22	SH->SO_3H	47.984745	Thiol to Sulfonic acid
23	-C_2H_3NO	-57.021464	loss of glycine
24	+C_2H_3NO	57.021464	glycine conjugation
25	-SO_3	-79.956817	loss of sulfate
26	+SO_3	79.956817	sulfate conjugation
27	-HPO_3	-79.966333	loss of Phosphate

http://mycompoundid.org

- Calculates the MW of metabolic transformations (+adducts, neutral loss fragments) from HMDB "parent" metabolites
- 375,809 compounds from one metabolic reaction and 10,583,901 from two reactions
- Number of putative compound hits (via mass matching) for MS-based metabolomic experiments increases 4-5X

Key Trends in Metabolomics

- Automated metabolomics
- Expanding metabolome coverage
- Making metabolomics portable
- Quantify, quantify, quantify…
- Moving metabolomics from the lab to the clinic
- Moving metabolomics (back) into drug development and discovery

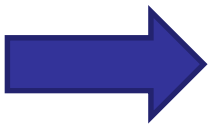
Personalized Medical Monitoring Devices

fitbit

iPhone

ACCU-CHEK®
Roche

Democratizing Metabolomics



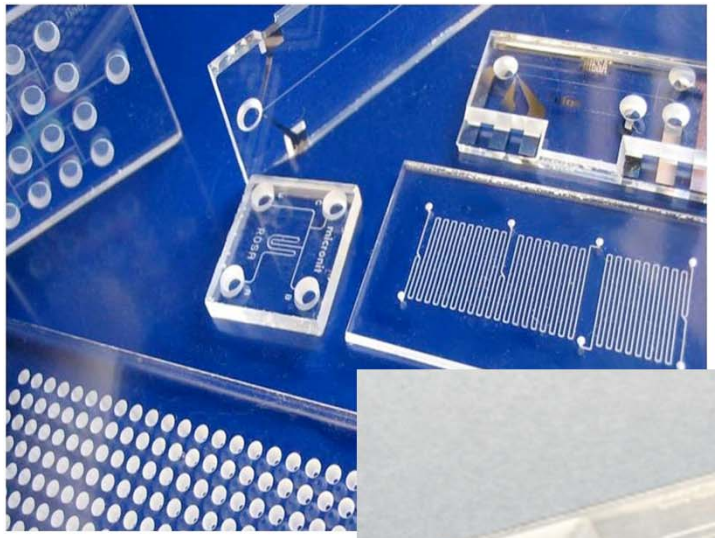
$10 million instrument, $200/test

$1000 instrument, $2/test

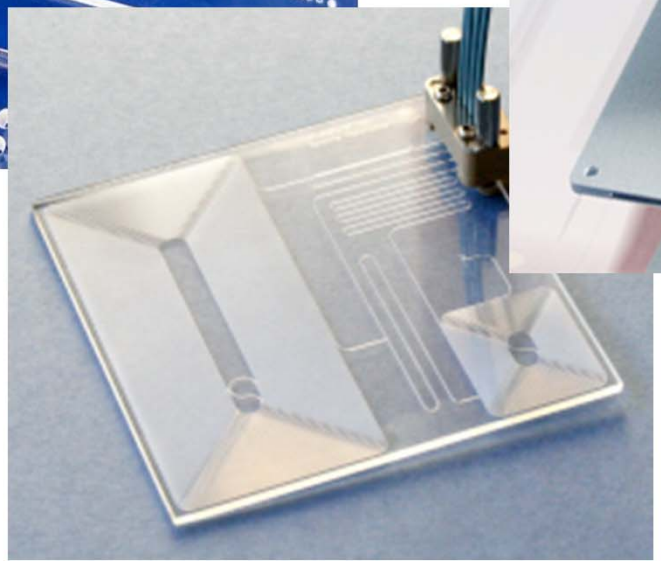
Not As Absurd As You Think

QUALCOMM TRICORDER XPRIZE

Miniaturization via Microfluidics & Nanotech



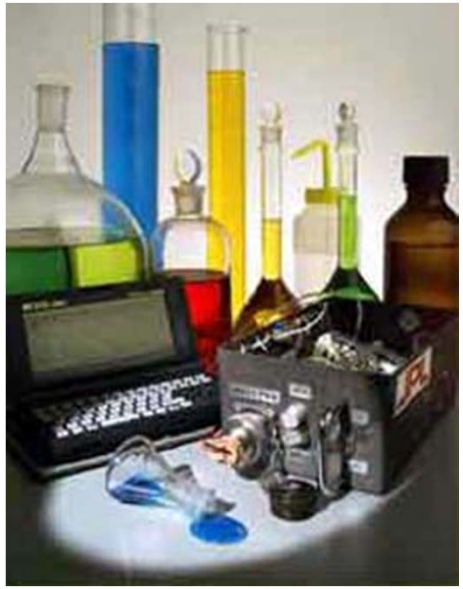
CE on a chip

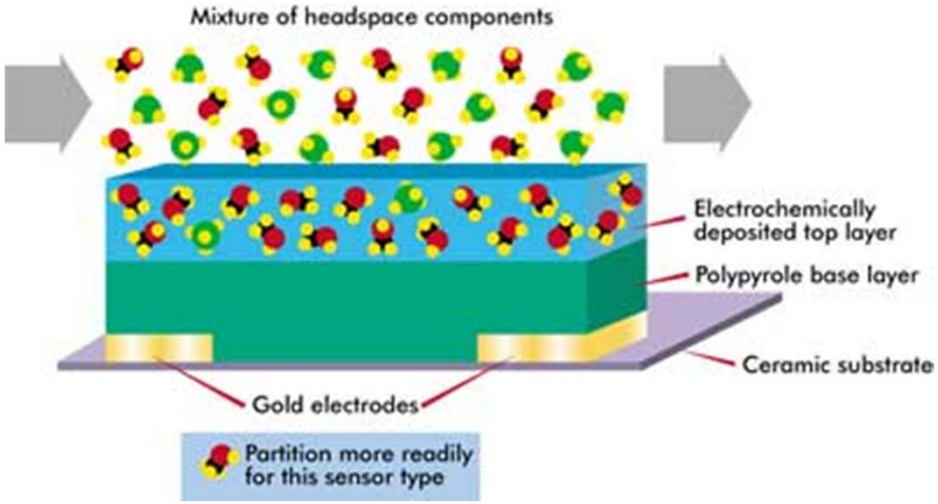


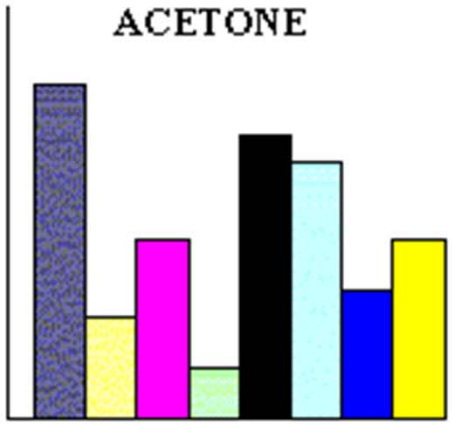
GC on a chip

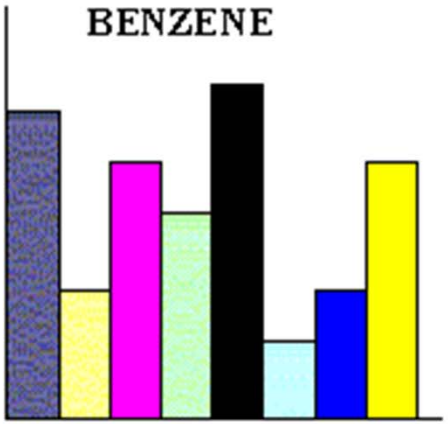
HPLC on a chip

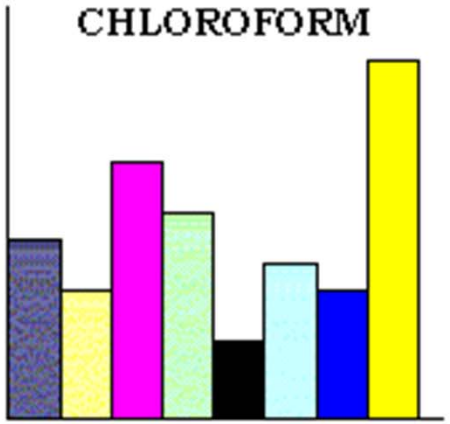
E-Nose for Volatile Metabolites



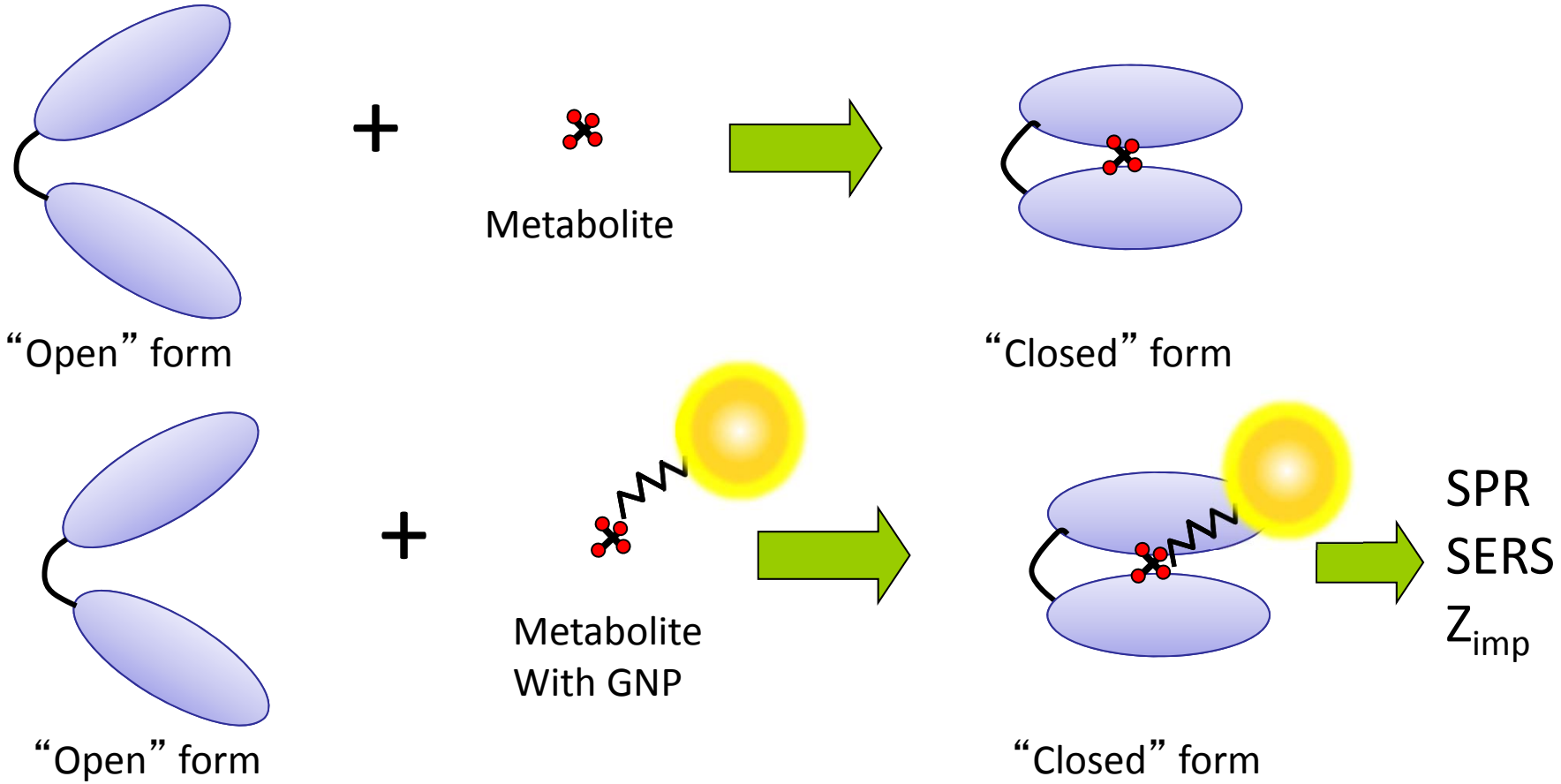




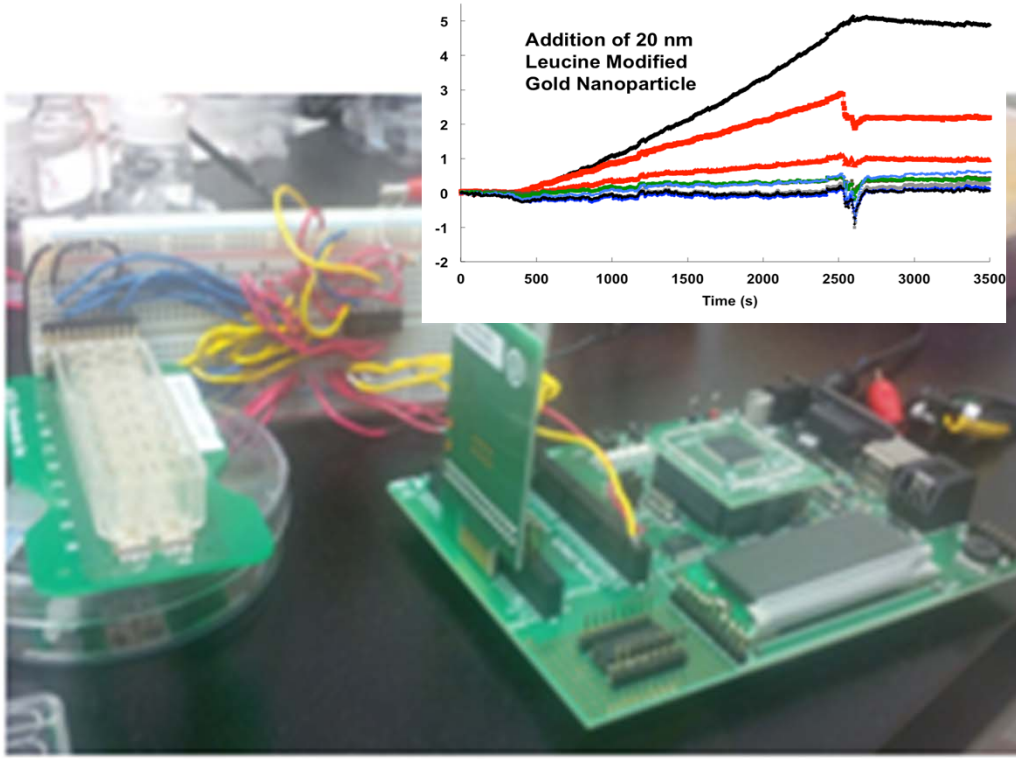




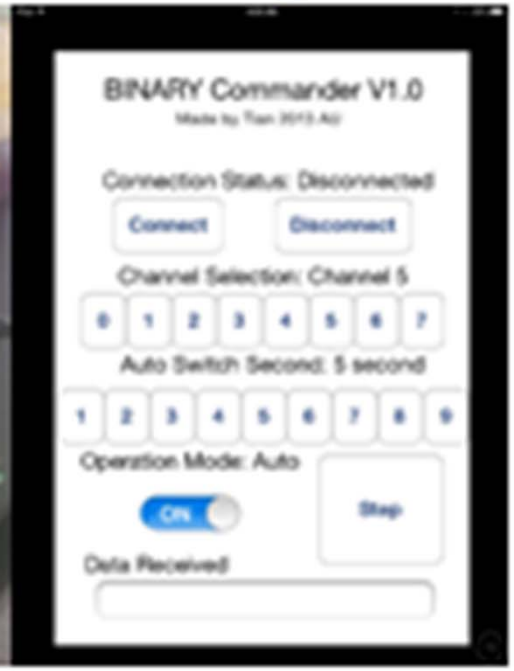
Protein and Aptamer-Mediated Metabolite Sensing



Working Prototype: Impedance-Based Metabolite Sensor



iPhone App



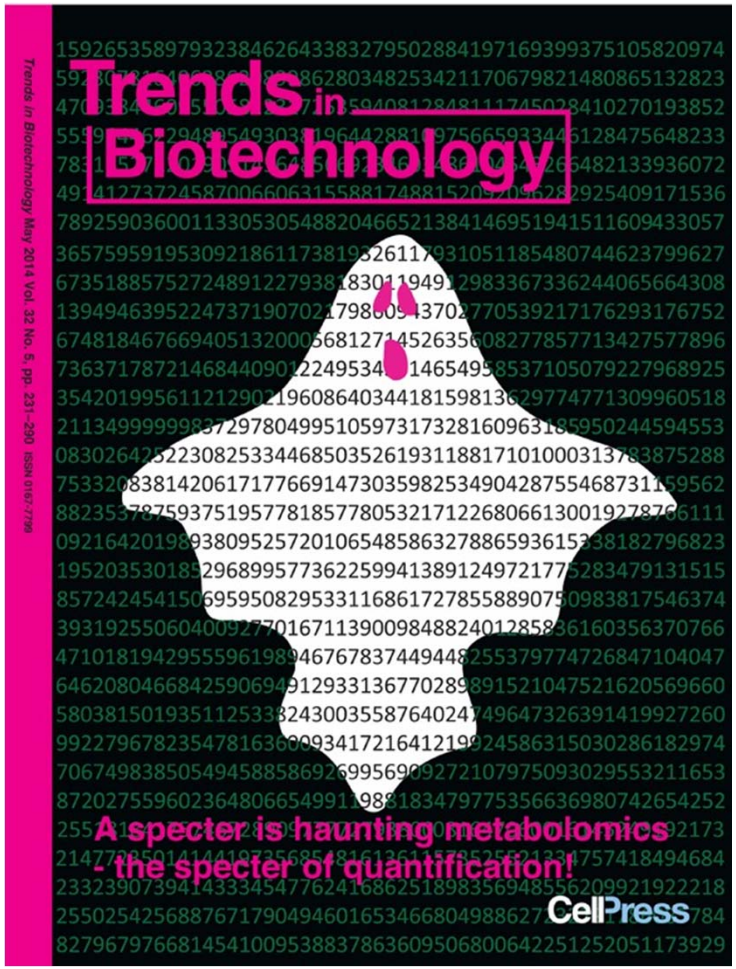
Developed by Dr. Jie Chen, University of Alberta

Key Trends in Metabolomics

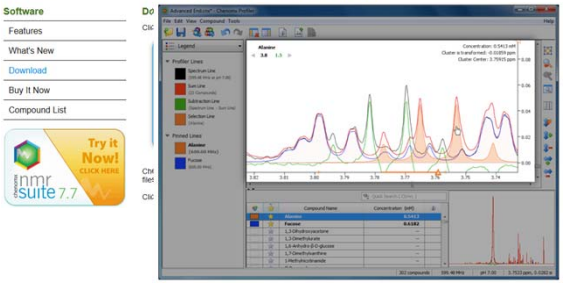
- Automated metabolomics
- Expanding metabolome coverage
- Making metabolomics portable
- Quantify, quantify, quantify…
- Moving metabolomics from the lab to the clinic
- Moving metabolomics (back) into drug development and discovery

Quantification & Metabolomics

- >90% of published metabolomics studies are semi-quantitative (relative peak areas, intensities)
- <10% of published metabolomics studies use absolute quantification
- The field MUST become more quantitative if findings are to be translated to practical applications



Quantitative Metabolomics (Commercial)

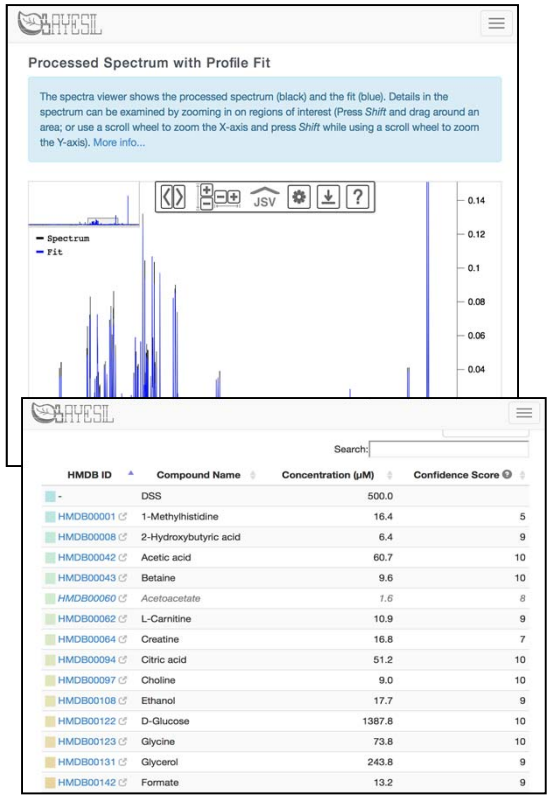


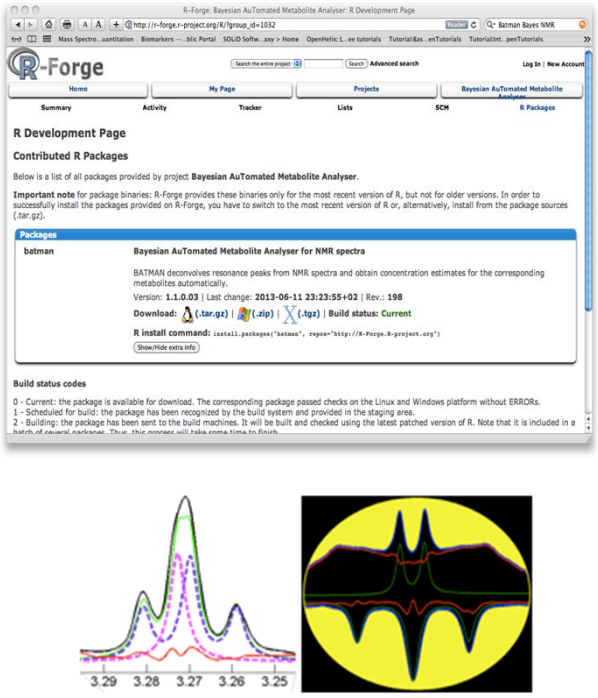
Chenomx– Automated NMR

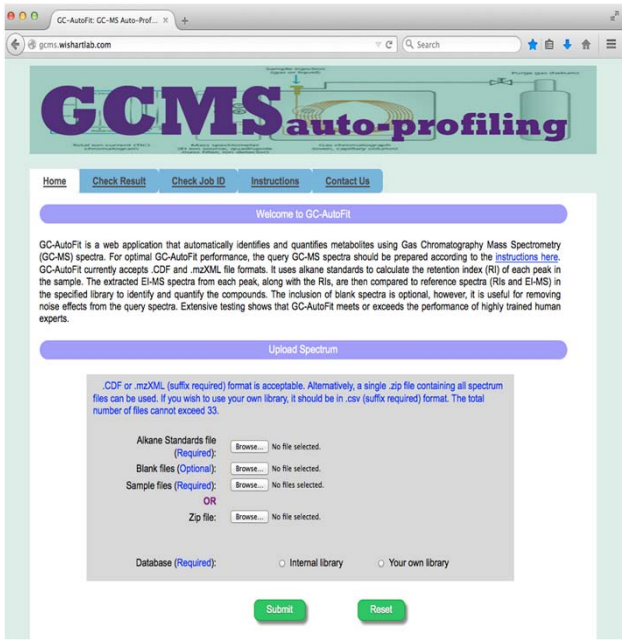
Bruker – Automated NMR

Biocrates – Automated MS

Quantitative Metabolomics (Academic)







Bayesil **Batman** **GC-Autofit**

Some Impressive Results…

Human Biofluid Omics "Records" for Absolute Quantification

	Metabolomics	Proteomics	Genomics (Transcripts)
Serum/Plasma	288 Identified & Quantified [1]	73 Identified & Quantified [4]	0
CSF	172 Identified & Quantified [2]	130 Identified & Quantified [5]	0
Urine	378 Identified & Quantified [3]	63 Identified & Quantified [6]	0

1. Psychogios N. et al. (2011) PLoS One 6(2): e16957
2. Mandal R. et al. (2012) Genome Med.;4(4):38.
3. Bouatra S. et al. (2013) PLoS One 8(9): e73076
4. MRM Proteomics Inc. (Victoria BC) reported in 2014
5. Percy AJ. et al. (2014) J. Proteome Res. (ePub Jun 9)
6. Chen YT. et al. (2012) J. Proteomics 75(12):3529

Key Trends in Metabolomics

- Automated metabolomics
- Expanding metabolome coverage
- Making metabolomics portable
- Quantify, quantify, quantify…
- Moving metabolomics from the lab to the clinic
- Moving metabolomics (back) into drug development and discovery

Some Grim Statistics

- Since 1970 > 700,000 biomarker papers published in PubMed
- Since 1970 <250 biomarkers have been approved for clinical use
- No markers approved (yet) using proteomics methods (lots use ELISA)
- 5 biomarker tests approved using transciptomics or gene chips

But Did You Know… Almost Everyone <25 Has Had A Metabolomic Test?

Newborn Screening

“Omics” Testing

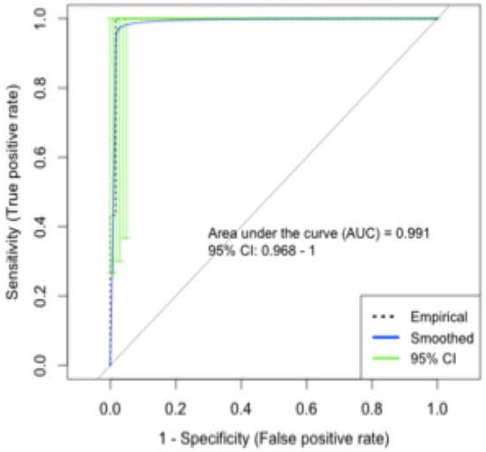
- Number of “approved” tests arising from Metabolomics/Clinical Chem. – 195
- Number of “approved” tests arising from or using Genomics – 100-110
- Number of “approved” single Protein tests (ELISA) – 60
- Number of “approved” tests arising from or using Transcriptomics – 5
- Number of “approved” tests arising from or using Proteomics - 0

How Does Metabolomics Do?

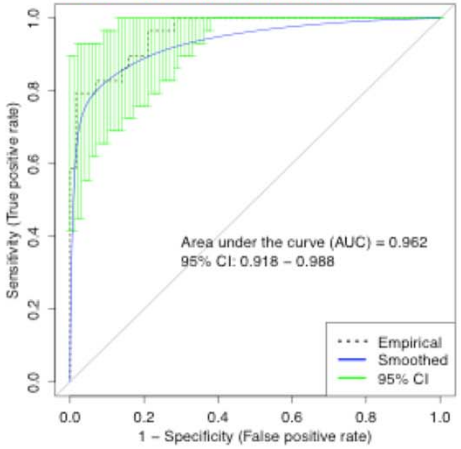
(Prediction & Diagnosis)

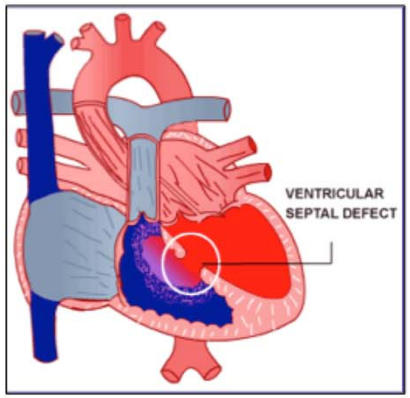
Predicting Diseases

Early Preeclampsia
First trimester maternal serum
AUC = 0.99 (2 metabolites)

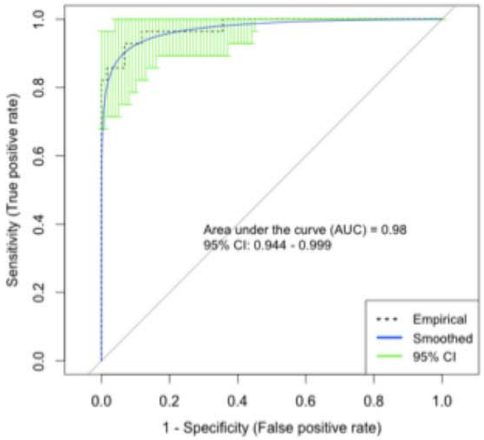


Late Preeclampsia
First trimester maternal serum
AUC=0.96 (8 metabolites)

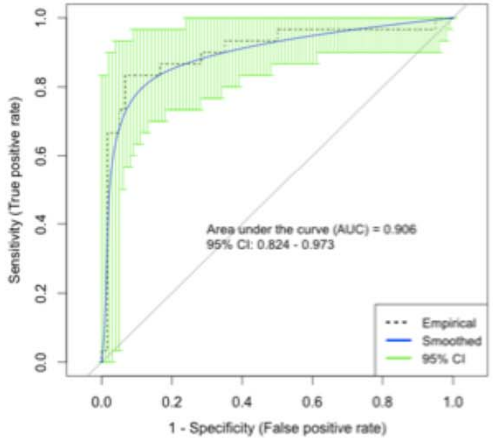




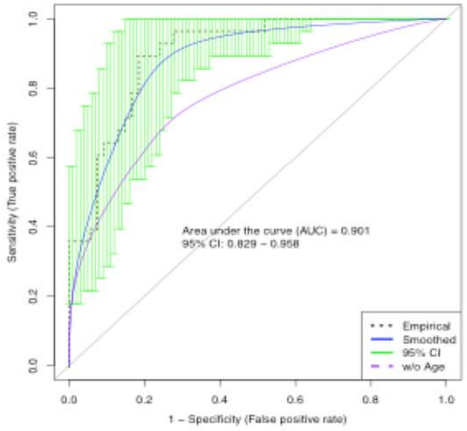
Congenital Heart Defects (CHD)
Maternal Serum
AUC=0.98 (3 metabolites)

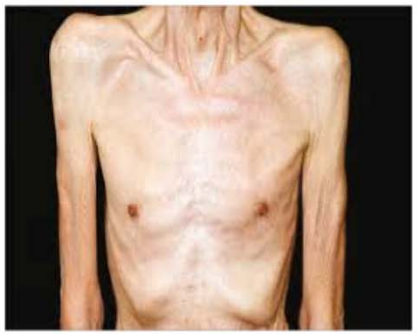


Trisomy 18
First trimester maternal serum
AUC=0.91 (7 metabolites)

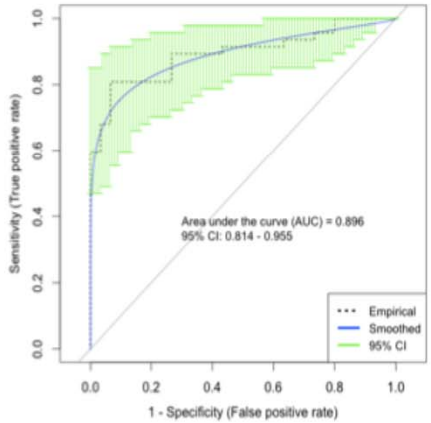


Trisomy 21
First trimester maternal serum
AUC=0.90 (3 metabolites + Age)



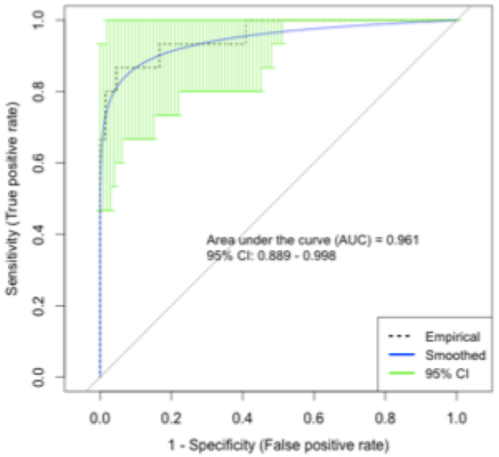


Cancer Cachexia
Adult Urine samples
AUC=0.90 (4 metabolites)

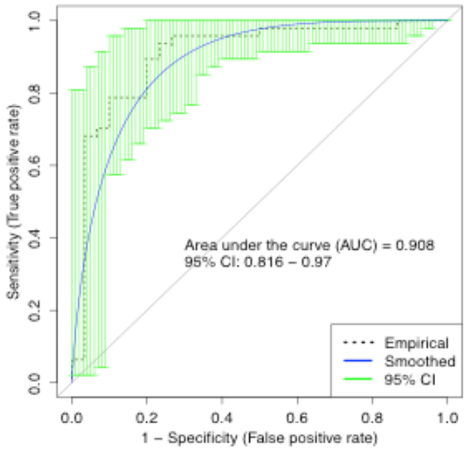


Diagnosing Diseases

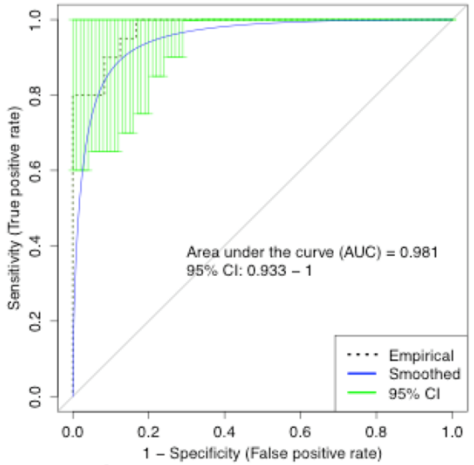
Adult kidney transplant rejection
Adult Urine samples
AUC=0.96 (9 metabolites)

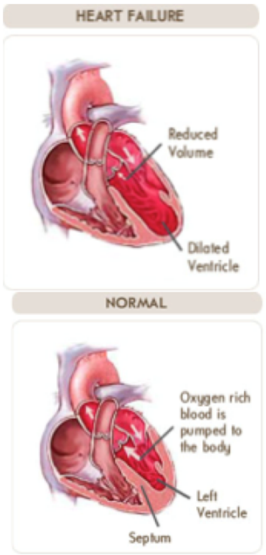


Pediatric kidney transplant rejection
Pediatric Urine samples
AUC=0.91 (4 metabolites)

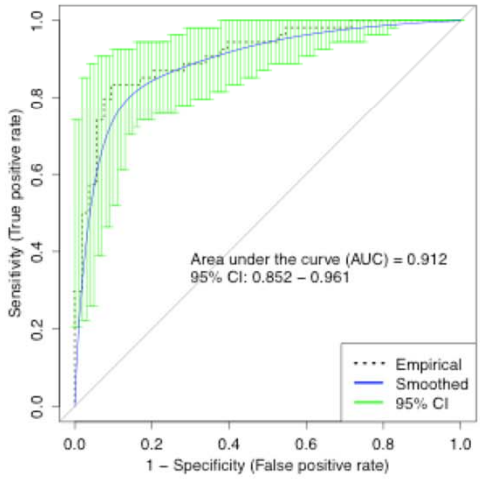


Systolic Heart Failure vs. Diastolic HF
Adult Serum samples
AUC=0.98 (4 metabolites)

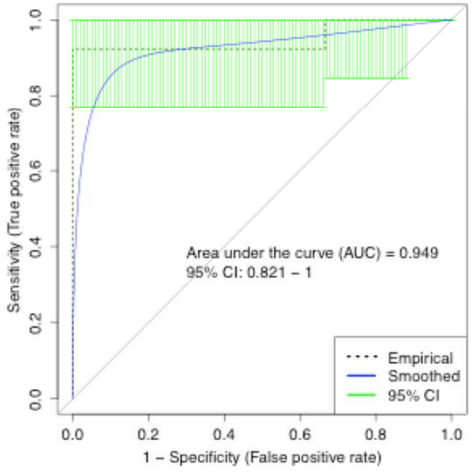




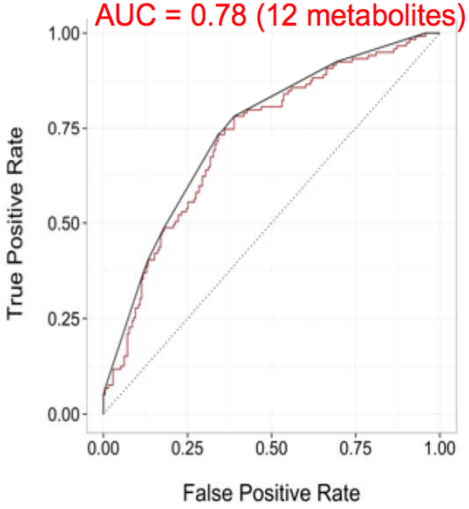
Chronic Fatigue Syndrome
Adult Serum samples
AUC=0.91 (6 metabolites)

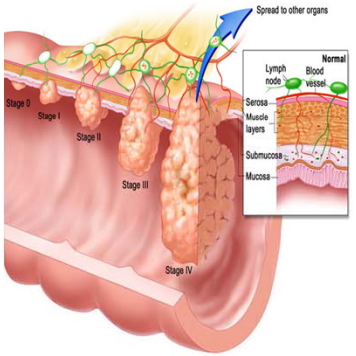


Eosinophilic Esophagitis (EoE)
Pediatric Urine Samples
AUC=0.95 (3 metabolites)



Colonic Polyps
Adult Urine Samples
AUC = 0.78 (12 metabolites)



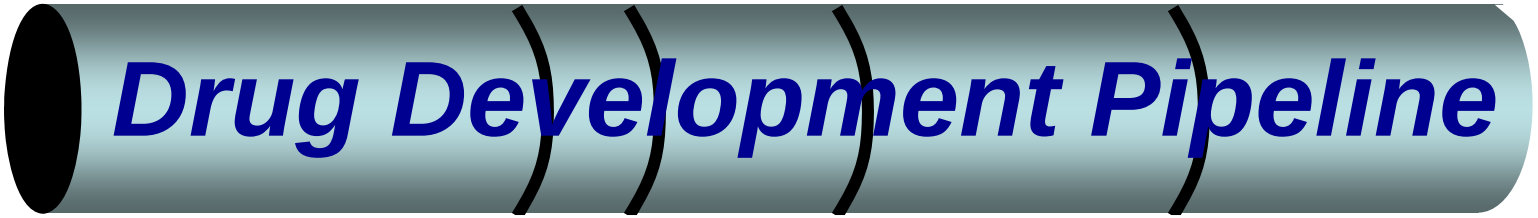


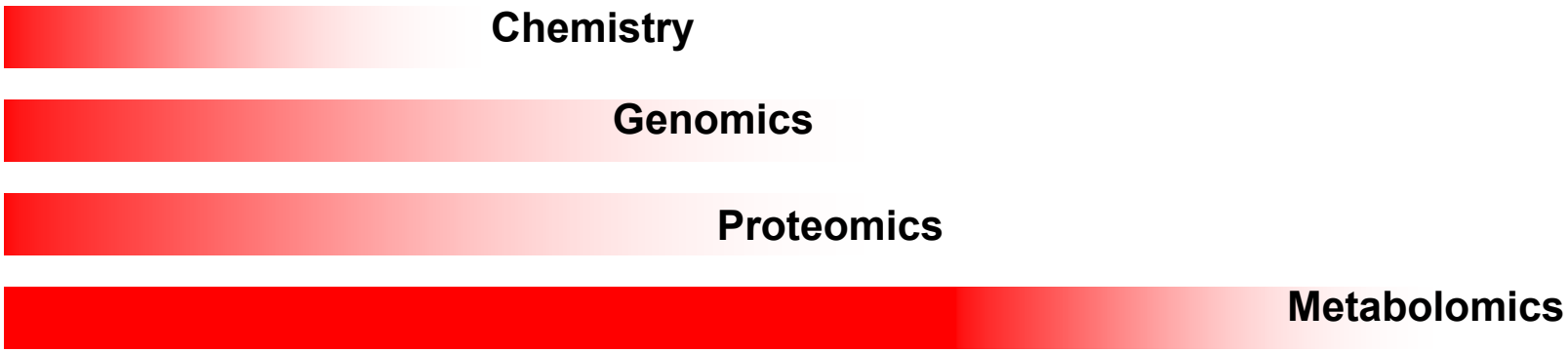
Key Trends in Metabolomics

- Automated metabolomics
- Expanding metabolome coverage
- Making metabolomics portable
- Quantify, quantify, quantify…
- Moving metabolomics from the lab to the clinic
- Moving metabolomics (back) into drug development and discovery

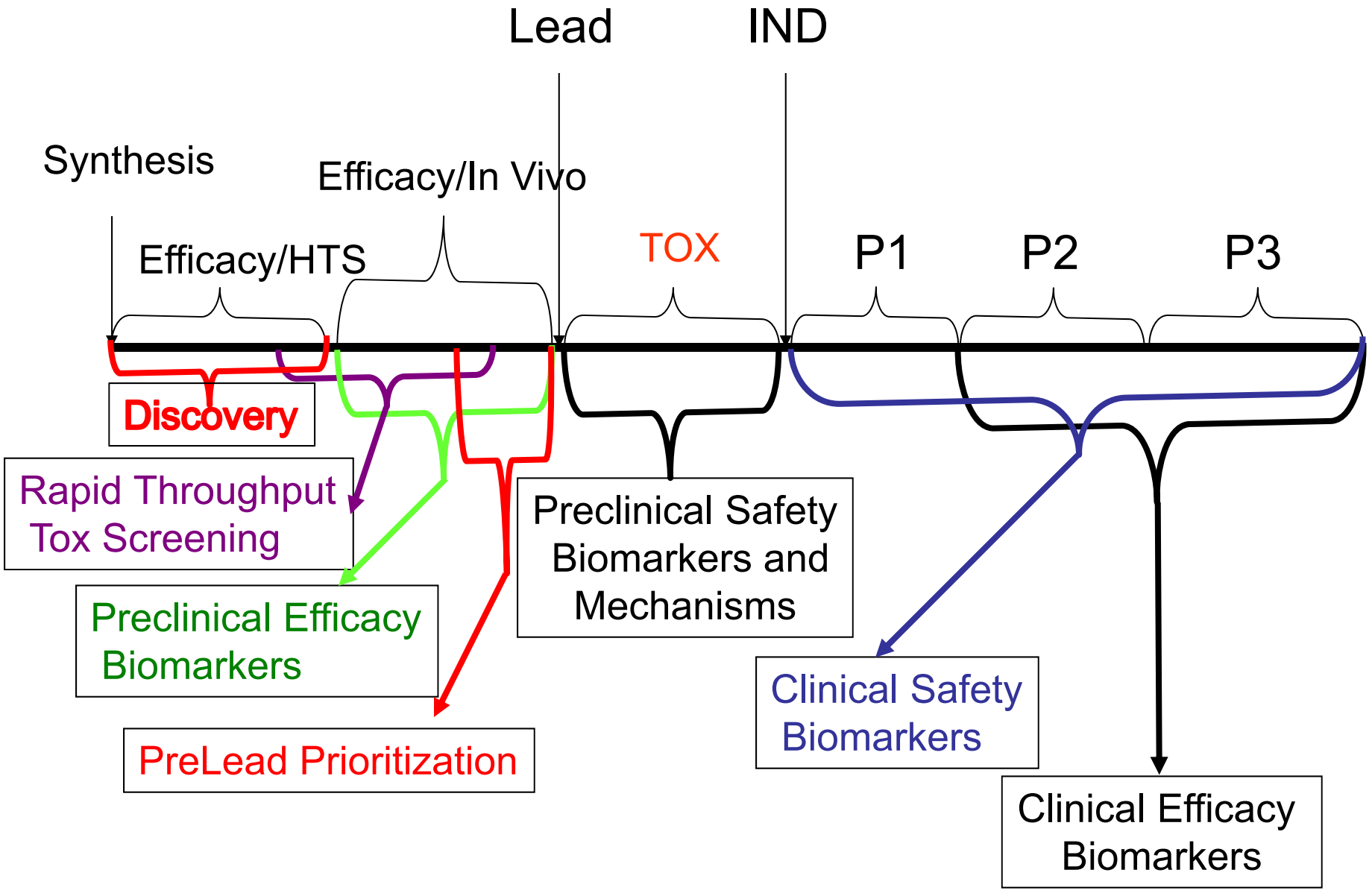
Metabolomics & The Drug Industry

$280	$60	$90	$250	$120 million
3.5 yrs	1 yr	2 yrs	3 yrs	2.5 yrs
Discovery	Phase I	Phase II	Phase III	FDA Approval

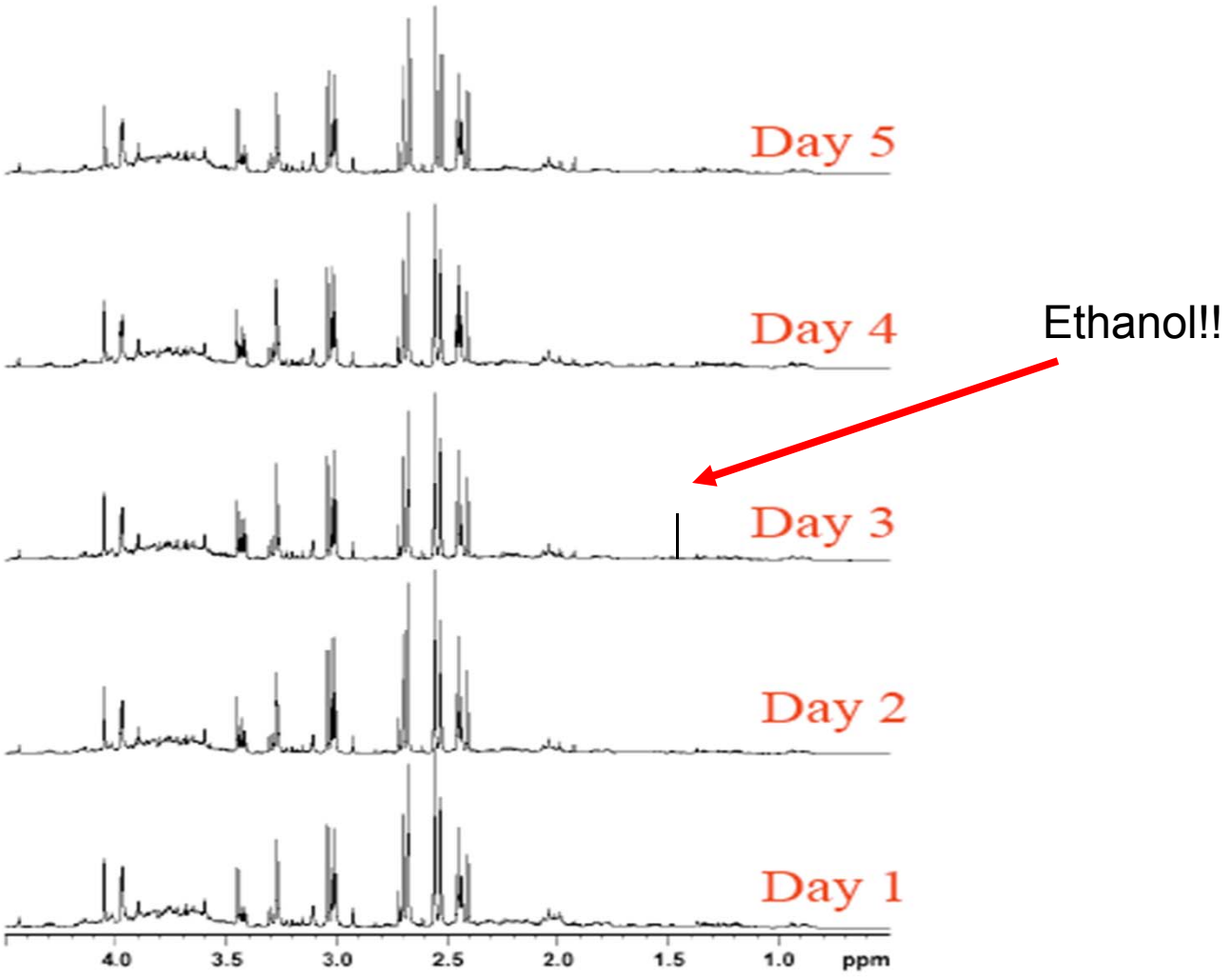




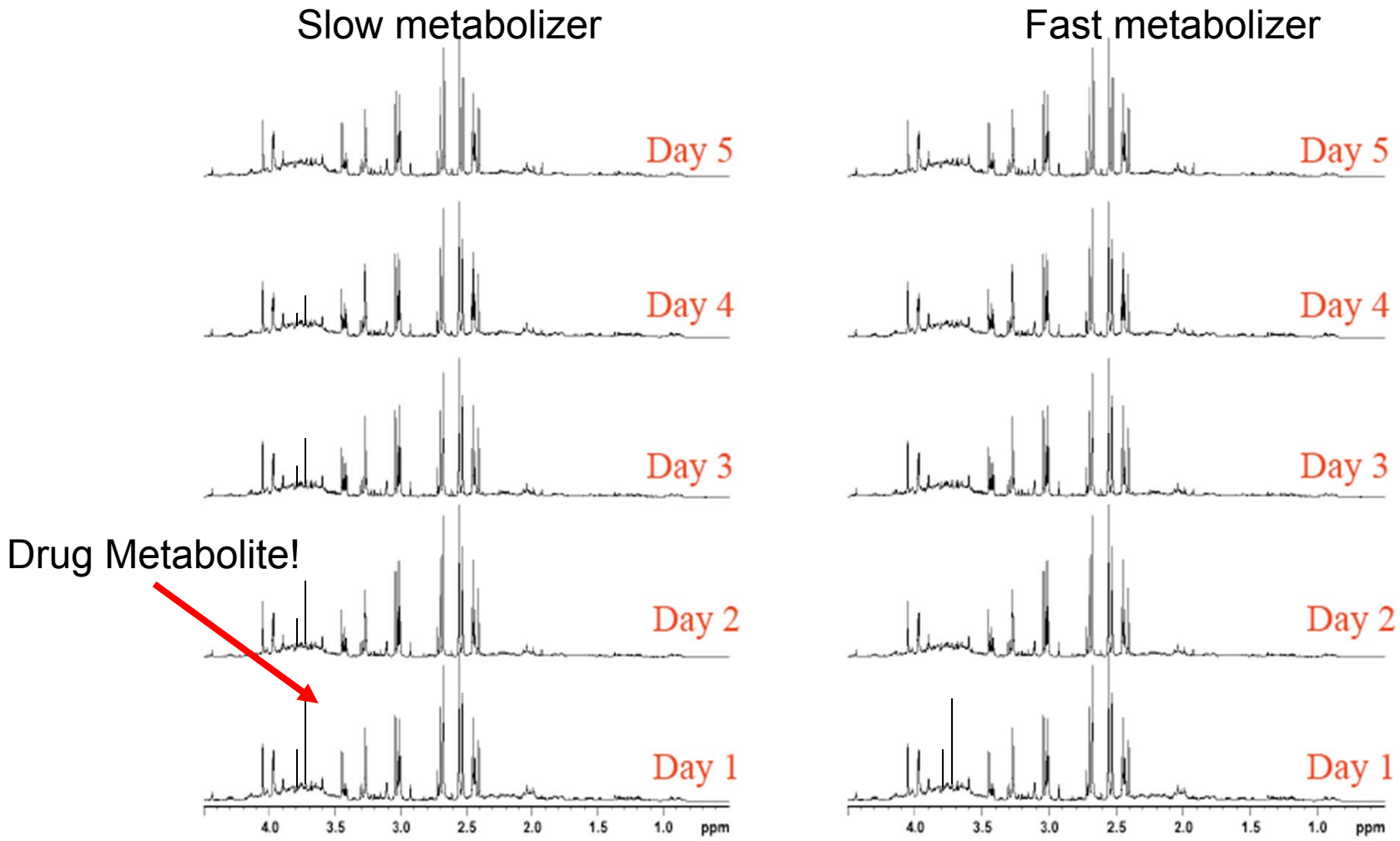
Metabolomics in Drug Development



Applications in Patient Compliance (P2/P3 Trials)

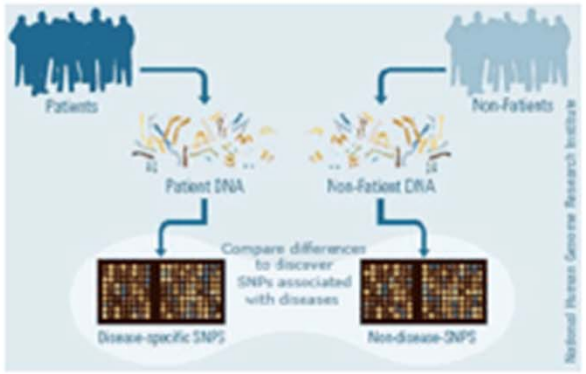


Applications in Drug Monitoring/Customization

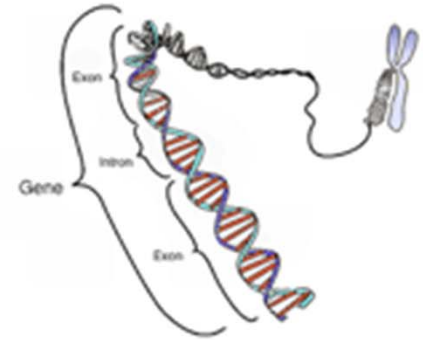


Traditional Drug Discovery

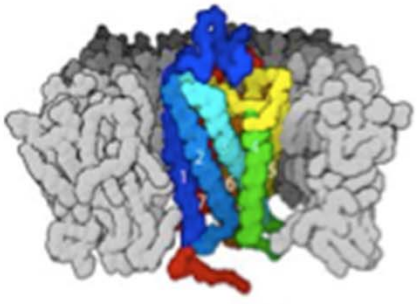
1:5 success $2-10 million, 2-4 yrs



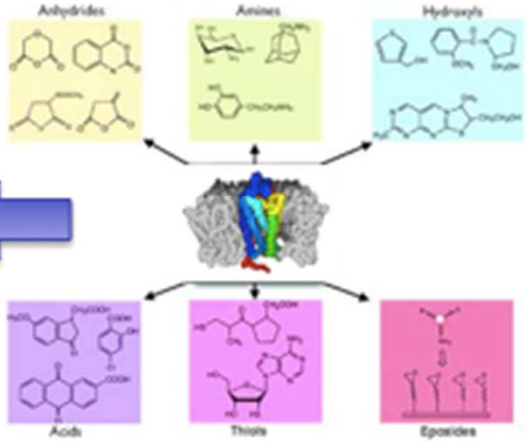
1:2 success



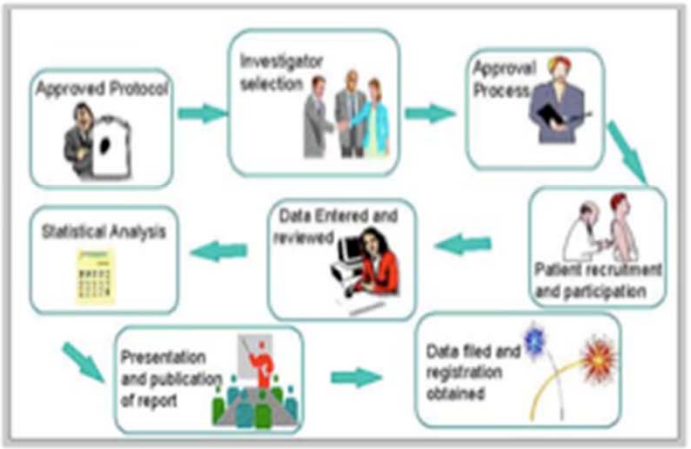
1:2 success



0.001% Success Rate, 20+ years, >$1 billion



1:5 success, 1-5 yrs

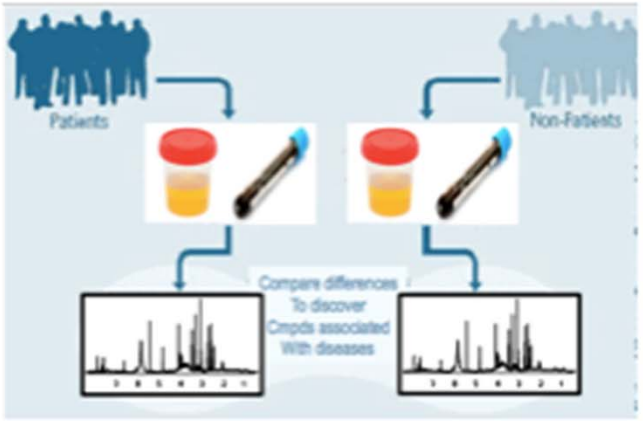


1:500 success, $1 billion, 15 yrs

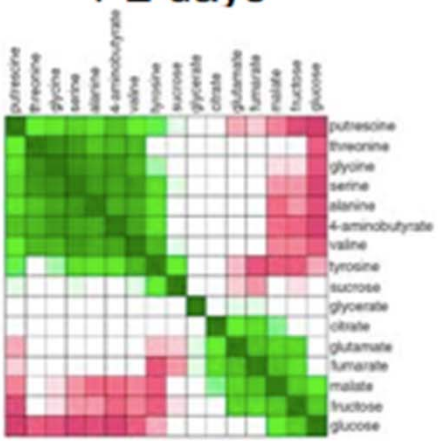
1:2 success

Metabolite-Based Drug Discovery

1:2 Success, $200K, 1 yr



1-2 days

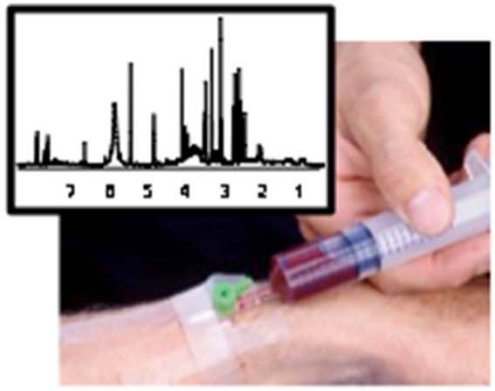


1-2 days

15% Success Rate, 1+ years, <$250,000

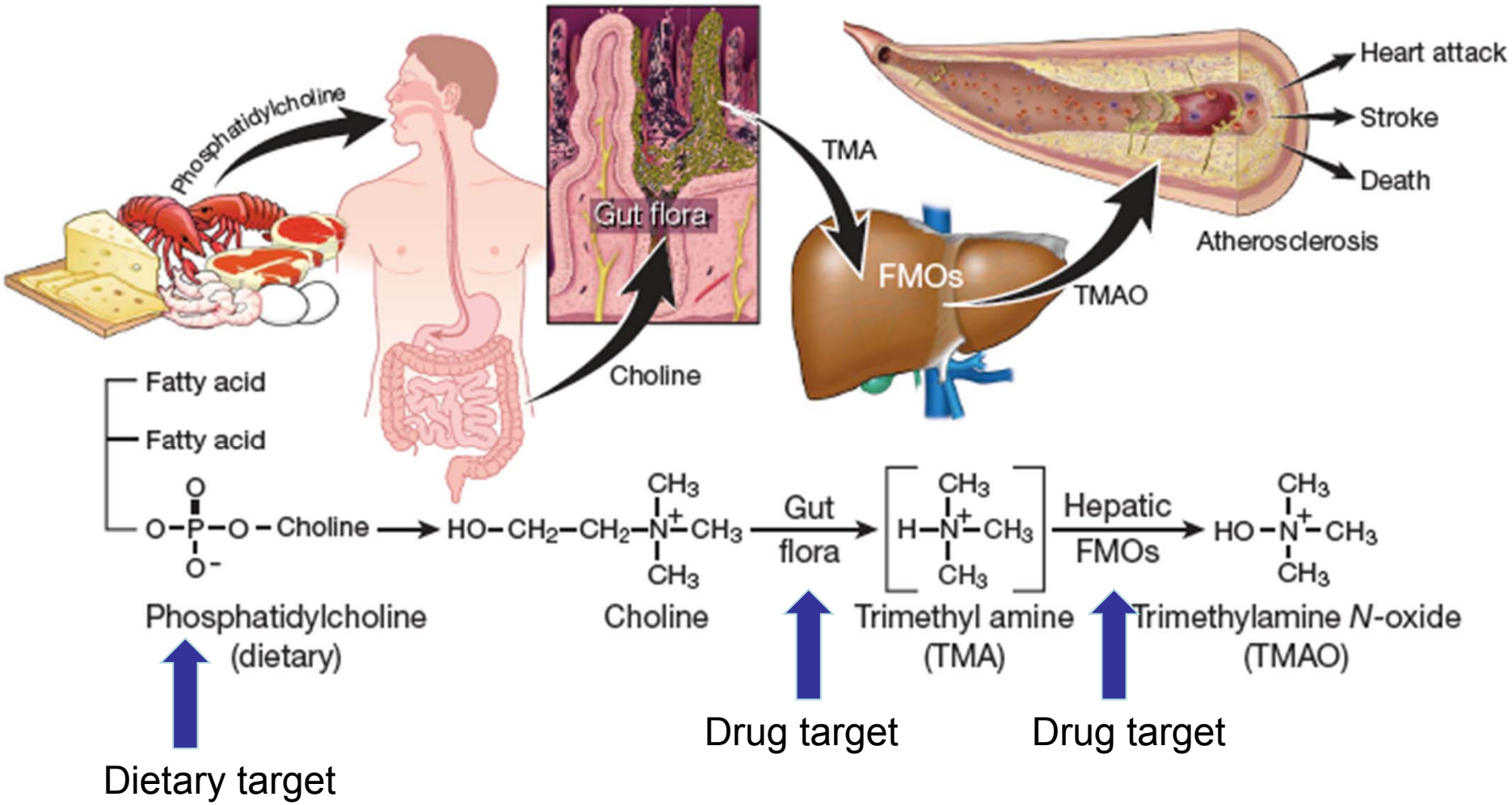
1:2 Success, 1-2 hours

1-2 weeks (cmpds/diet)
5-10 yrs (enzymes, MAbs)



1:2 Success, $200/yr

Metabolomics, CVD & Therapy



Summary – The Future of Metabolomics

- Automated metabolomics
- Expanding metabolome coverage
- Making metabolomics portable
- Quantify, quantify, quantify…
- Moving metabolomics from the lab to the clinic
- Moving metabolomics (back) into drug development and discovery

Acknowledgements

- Augustin Scalbert
- Russ Greiner
- Jie Chen
- Khalid Aziat
- Felicity Allen
- Richard Fedorak
- Liang Li
- Rupa Mandal
- Jennifer Reid
- Tamara Lim
- Ray Bahado-Singh
- Jason Dyck
- David Rush
- Tom Blydt-Hansen
- Siamak Ravanbaksh
- Beomsoo Han
- Jeff Xia
- Tim Jewison
- Allison Pon
- Craig Knox

GenomeAlberta

GenomeCanada

TMIC The Metabolomics Innovation Centre

CIHR IRSC
Canadian Institutes of Health Research
Instituts de recherche en santé du Canada