



UNIVERSITA PALACKÉHO
OLOMOUC



FAKULTNÍ NEMOCNICE
OLOMOUC



NTNU

Norwegian University of
Science and Technology

Metabolomics – tools, data and interpretation

Žídková L., Friedecký D., Adam T.

Laboratory for Inherited Metabolic Disorders

Department of Clinical Chemistry, Palacky University and
Faculty Hospital, Olomouc

**IX. celostátní
sjezd**

České společnosti
klinické
biochemie

20. – 22. 9. 2009
Praha

Lecture overview

1. Definitions
2. Tools
3. Data analysis and interpretation
4. Examples
5. Where are we (in Olomouc)?

Simplified !!

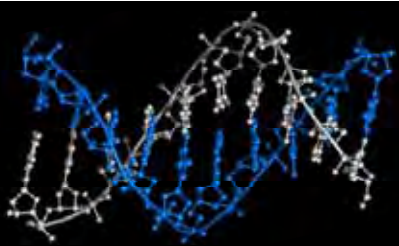
Definition of “Metabolomics”

Definition (one of many):

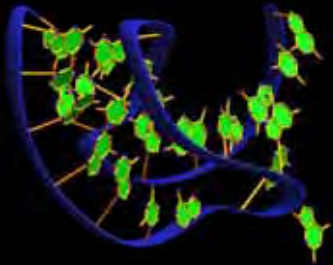
- **Metabolomics** is the measurement of „all“ low molecular weight metabolites within a biological sample under a defined temporal and physiological state.

Unsettled terminology

- Metabolomics, metabonomics, metabolic profiling,... targeted vs. Nontargeted



Genomics 25.000 genes



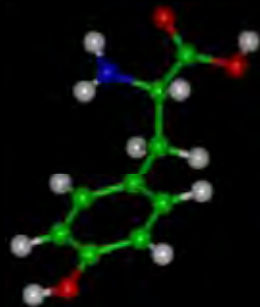
Transcriptomics 100.000 RNAs



Proteomics 1 mil. proteins



Metabolomics 2.400 compounds +



DYNAMIC in nature - Reflects phenotype - May be simpler than genome, proteome - less raw information than for other *omics, but more context.

What is a Metabolite?

- Any organic molecule detectable in the body with a MW < 1000 Da
- Includes sugars, nucleosides, organic acids, amino acids, lipids, steroids, peptides, oligonucleotides, alkaloids and drugs
- Includes human & microbial products
- Concentration > 1 μ M (NMR, excl. pollutants)

Medical Metabolomics – theor.

- Measure metabolite flux/pathway kinetics
- Track effects from toxins/drugs/surgery
- Identify functions of unknown genes
- Obtain a more “holistic” view of metabolism
- Diet strategies, drug testing (especially liver toxicity) in new drugs

Medical Metabolomics – practical

Genomics and proteomics tell you what might happen, but metabolomics tells you what actually did happen. Bill Lasley

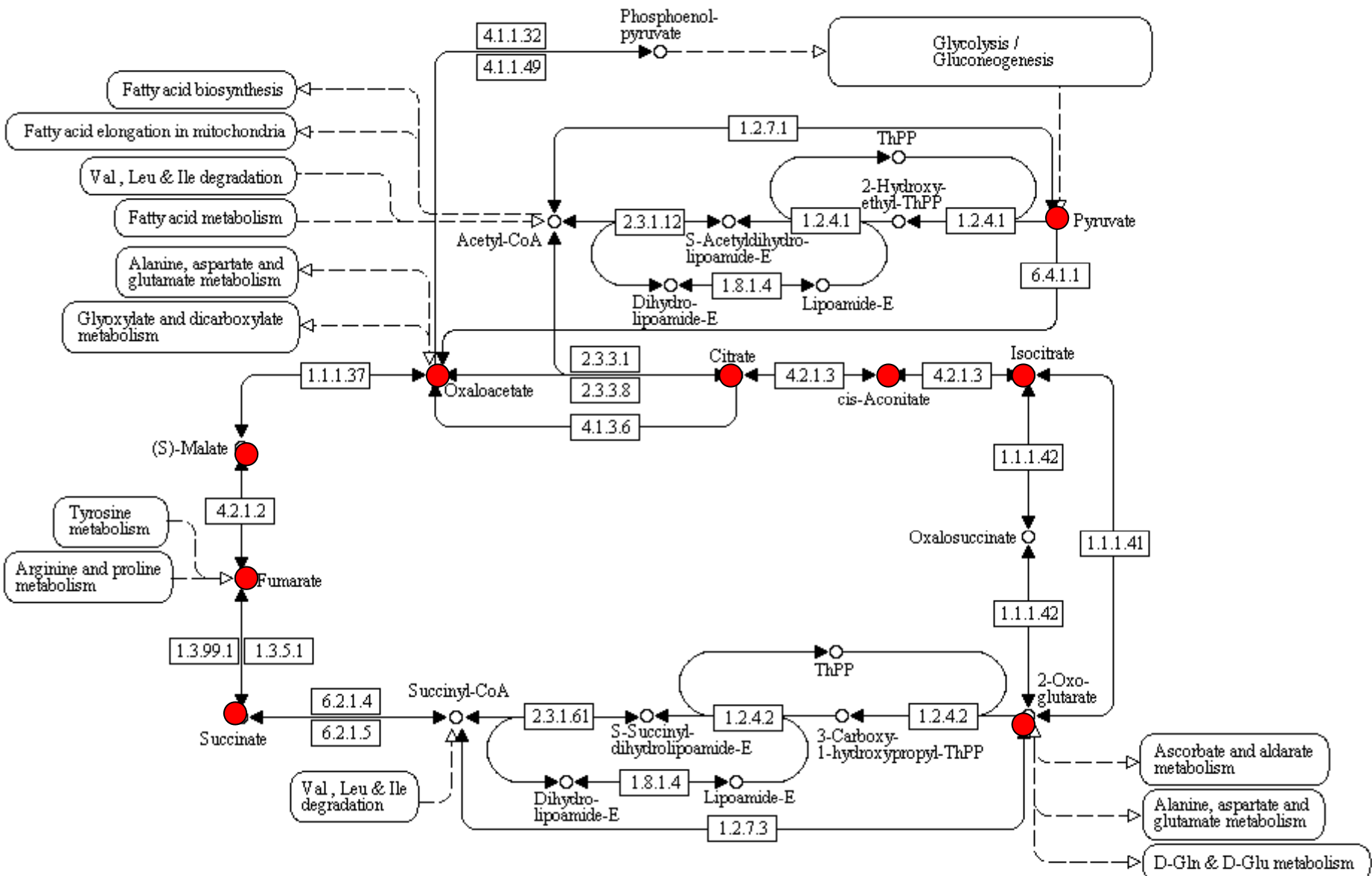
- Generate biomarkers for early detection of dis. eg necrosis, PC, Alzheimer's, infec/ inflam.
- Generate metabolic “signatures” for diseases

Coverage of the metabolome

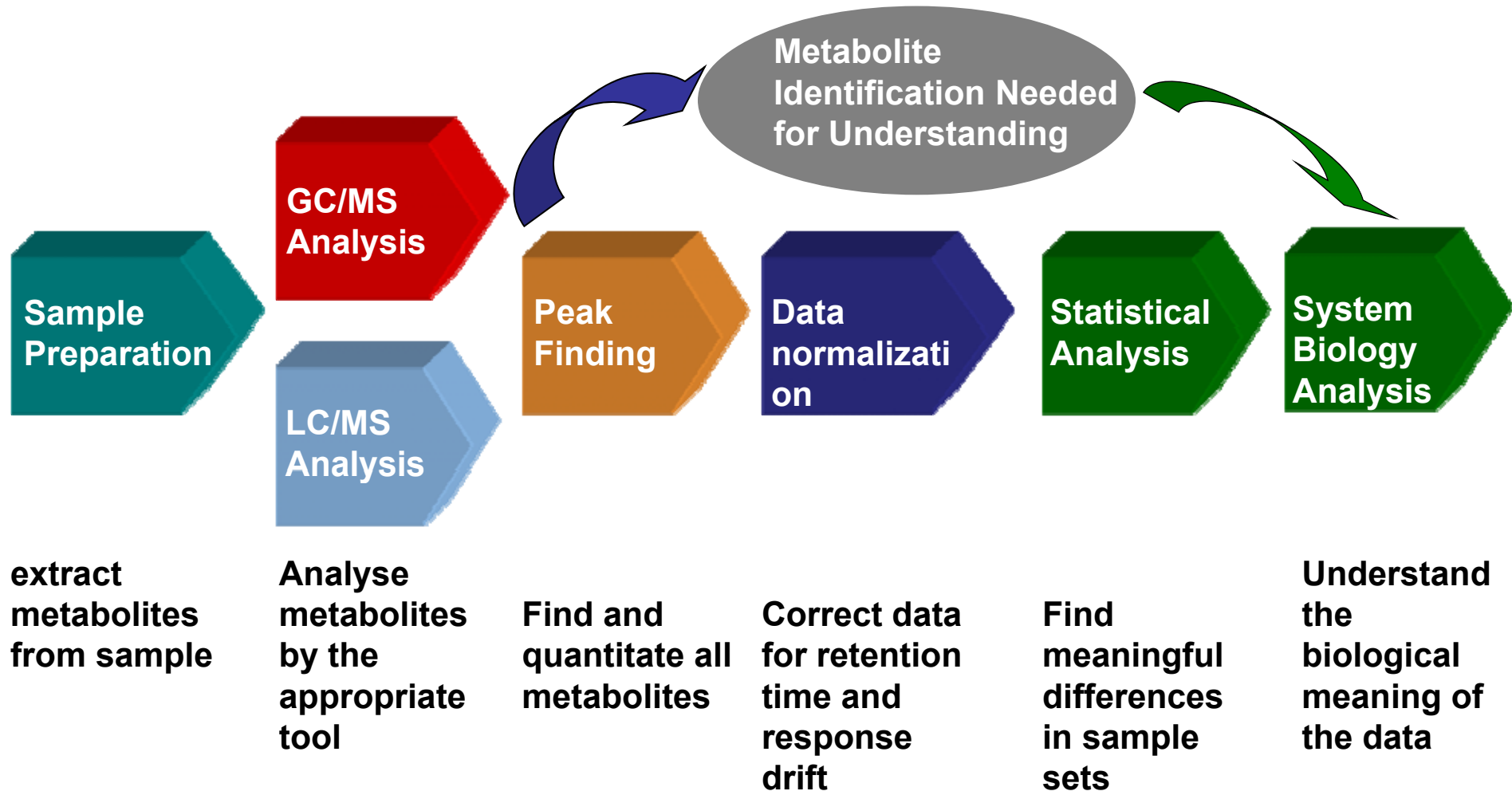
- available
- half life
- chemical instability
- detected by current methods

- Total number unknown
- Plants > 200,000
- Mammalian – can't be calculated from genome, 2.400 (>70,000 lipid permutations alone)

CITRATE CYCLE (TCA CYCLE)



Metabolomics Workflow



Tools

NMR, GCxGC-MS, CE-MS, LC-qTOF,...

GC/MS



LC/TMS



Olomouc - Metabolomics - Status quo

GC-MS

- Urine > 500 features (incl. Exometabolites)
- Cellular 43 compounds
- Plasma – under development

LC-TMS

- Urine > 110 features (incl.)
- Cellular 65 compounds
- Plasma – under development

Data analysis and interpretation

- ... a popravdě, literatura je absolutně plná úplných odpadků pocházejících z kombinace vysokého optimizmu, tváří v tvář zdánlivě pozitivnímu nálezu, statistické ignorance a strachu časopisů posoudit důkladně data, aby nenalezly něco nemilého...

Kell DB



Data analysis and interpretation

- Experiment design - Factorial designs
- Know your system
 - Understand your system's variances

Noise reduction & reproducibility !!

- QA/QC (Standards)
 - Recovery standards
 - Derivatization (or chemistry) standards
 - Injection standards

Std	Description	Purpose
MTRX	Large pool of human plasma	Assure all aspects of profiling process are within specifications
SMTRX	Pool created using aliquot from each Sample form one source	Assess effect of matrix on profiling process; distinguish biological- from process Variability
PRCS	Aliquot of ultra-pure water	Process blank to assess contribution to compound signals from process
SOLV	Aliquot of extraction solvents	Solvent blank used to segregate contamination sources in extraction
DS	Derivatization Standard	Assess variability of derivatization for GC/MS samples
IS	Internal Standard	Assess variability/performance of instrument
RS	Recovery Standard	Assess variability; verify performance of extraction/instrumentation

Deuterated standards

AA

glycine
alanine
valine
methionine
leucine
tyrosine
phenylalanine
Glutamic acid
glutamine
proline
cystine
ornithine
lysine
Homocysteine

Fatty a.

Methylmalonic acid
oxalic acid
Pipelicolic acid
glutaric acid
3,4-dihydroxyphenylacetic
methylsuccinic acid
docosanoic acid
tetracosanoic acid
hexacosanoic acid
phytanic acid
3,7-dihydroxycoprostanic acid
3,7,12-trihydroxycoprostanic acid

Acylcarnitines

L-carnitine.HCL
acetyl-L-carnitine.HCL
propionyl-L-carnitine.HCL
butyryl-L-carnitine.HCL + 10 othe

P&P

Orotic acid
Uric acid
Uracil
thymine
adenosine

Carboh.

glucose
galaktose
fruktose

Other

glycerol

Data Interpretation

- Standardization
- Normalization
 - Median centered
 - Z-transformation
 - $(\text{Obs} - \text{Mean}_c) / \text{StdDev}_c$
- Statistics vs. Data-mining
 - Statistics
 - Independent observations, False discovery
 - Data-mining
 - Random Forest, Non-negative factorization, Partition, Singular Value Decomposition

Data analysis and interpretation

- Data represented in a matrix

Chemometric Approach

- Principle Component Analysis (PCA)
- Soft Independent Modeling of Class Analogy (SIMCA)
- Partial Least-Squares (PLS) Method by Projections to Latent Structures
- Orthogonal PLS (OPLS)
- Targeted Profiling

Variables going across in different columns

Objects going down in different rows	X-var 1	X-var 2	X-var 3	Y-var 1	Y-var 2
Sample 1					
Sample 2...					

A propositional approach to describing and using metabolomics data (the x-data) for analyzing complex systems. These may have other specific properties (the y-data) which one may also wish 'explain' in terms of the x-data.

A1 \times Metabolism

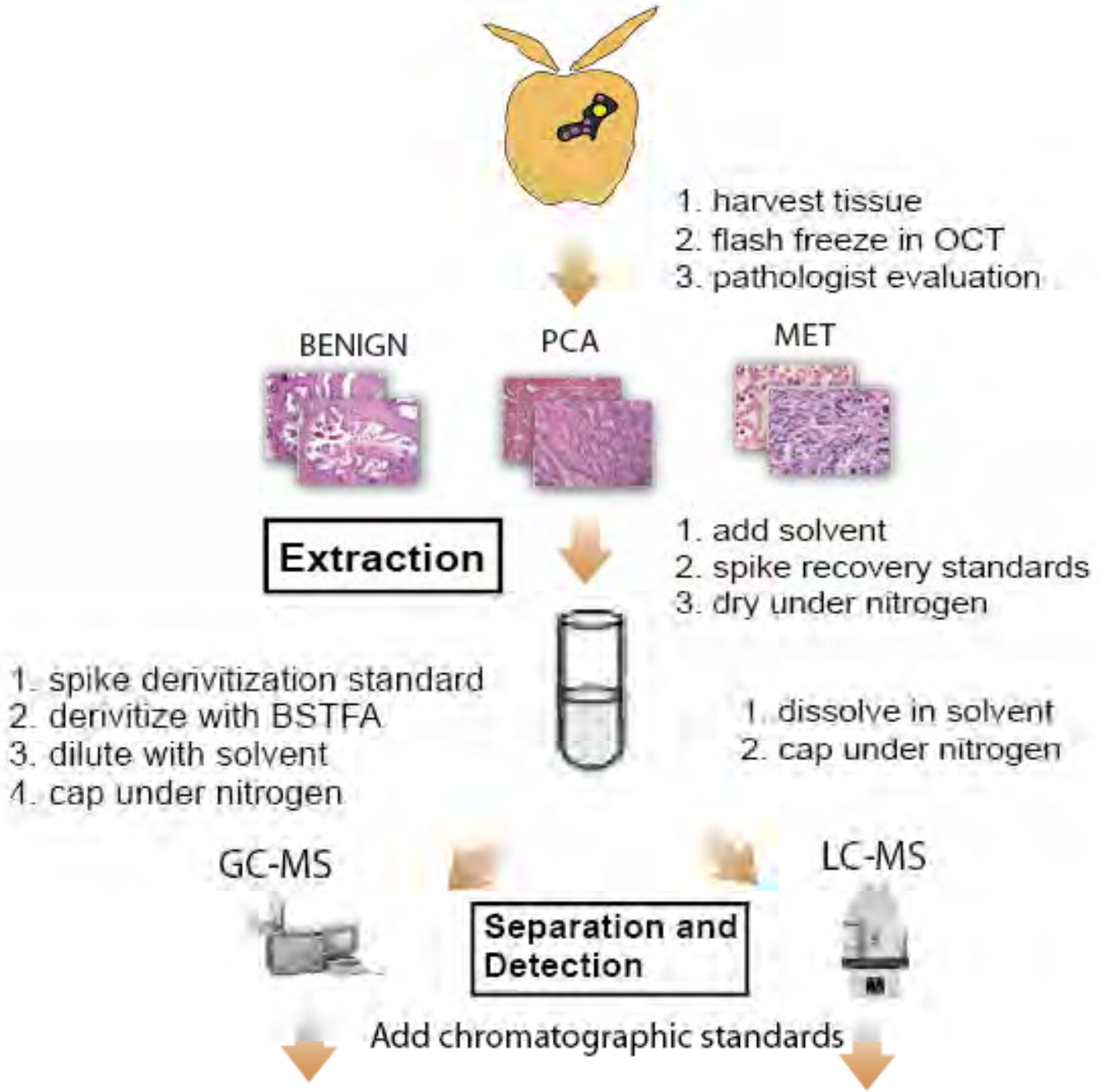
	A	B	C	D	E	F	G		
1	Metabolism	http://www.genome.jp/kegg/pathway.html#metabolism					Nature		
2	KEGG	Compound	Metabolism	Pathway	Humar	techniq	sampl	M	
3	C00489	Glutarate	Lipid Metabolism	Fatty acid metabolism	not	LC	T	G	
4	C00037	Glycine	Metabolism of Cofactors and	Porphyrin and chlorophyll me	H	GC	T U P	G	
5	C00387	Guanosine	Nucleotide Metabolism	Purine metabolism	H	LC	T U	G	
6	C05582	Homovanillate	Amino Acid Metabolism	Tyrosine metabolism	H			H	
7	C00262	Hypoxanthine	Nucleotide Metabolism	Purine metabolism	H			H	
8	C00294	Inosine	Nucleotide Metabolism	Purine metabolism	H			Ir	
9	C00788	L-Adrenaline	Amino Acid Metabolism	Tyrosine metabolism	H			A	
10	C00041	L-Alanine	Amino Acid Metabolism	Alanine and aspartate meta	H	GC	T U P	A	
11	C00049	L-Aspartate	Energy Metabolism	Nitrogen metabolism	H	GC	T	A	
12	C02984	L-Aspartyl-tRNA(Asp)	Amino Acid Metabolism	Alanine and aspartate meta	H				
13	C01832	Lauroyl-CoA	Lipid Metabolism	Fatty acid metabolism	H				
14	C00327	L-Citrulline	Amino Acid Metabolism	Arginine and proline metabo	H	GC	T U P	C	
15	C00097	L-Cysteine	Metabolism of Other Amino	Glutathione metabolism	H	GC	T U	C	
16	C03125	L-Cysteinyl-tRNA(Cys)	Amino Acid Metabolism	Cysteine metabolism	H				
17	C01693	L-Dopachrome	Amino Acid Metabolism	Tyrosine metabolism	H				
18	C00025	L-Glutamate	Metabolism of Cofactors and	Porphyrin and chlorophyll me	H	GC	T U P	G	
19	C01165	L-Glutamate 5-semialdehyde	Amino Acid Metabolism	Arginine and proline metabo	H				
20	C00064	L-Glutamine	Amino Acid Metabolism	Glutamate metabolism	H	GC	T U P	G	
21	C00155	L-Homocysteine	Amino Acid Metabolism	Methionine metabolism	H			H	
22	C05939	Linatine	Amino Acid Metabolism	Arginine and proline metabo	H				
23	C00407	L-Isoleucine	Amino Acid Metabolism	Valine, leucine and isoleucin	H			Is	
24	C03127	L-Isoleucyl-tRNA(Ile)	Amino Acid Metabolism	Valine, leucine and isoleucin	H				
25	C00328	L-Kynurenine	Amino Acid Metabolism	Tryptophan metabolism	H				
26	C00123	L-Leucine	Amino Acid Metabolism	Valine, leucine and isoleucin	H			L	
27	C00047	L-Leucyl-tRNA	Amino Acid Metabolism	Valine, leucine and isoleucin	H				

	D	E	F	G	H	I	J	K	L	M	N
Nature											
Way	Human	techniq	sample	Metabolite	Catalog n	Mr (g/mol)	Producer	Purity	Quantity	Group	
acid metabolism	not	LC	T	Glutaric acid	G340-7	132,12	Aldrich	99%	Malonic acid		
phytyl and chlorophyll me	H	GC	T U P	Glycine	10MG	75,1	NBC		> 50 mg	AA	
e metabolism	H	LC	T U	Guanosine	G-6752	283,2	Sigma	98%	> 50 mg	PP	
cine metabolism	H			Homovalinic acid	H-1252	182,2	Sigma		> 50 mg	Skříň - A	
e metabolism	H			Hypoxanthine	H-9377	136,1	Sigma		> 50 mg	PP	
e metabolism	H			Inosine	I-4125-1G	268,2	Sigma, S	99%	> 50 mg	PP	
cine metabolism	H			Adrenaline, L		183,21	Merck		> 50 mg	Skříň - C	
cine and aspartate metal	H	GC	T U P	Alanine	10MG	89,1	Calbiochem, NB		> 50 mg	AA	
gen metabolism	H	GC	T	Aspartic,L acid	10MG	133,1	NBC, Calbioche		> 50 mg	AA	
cine and aspartate metal	H										
acid metabolism	H										
cine and proline metabol	H	GC	T U P	Citrulline,L	10MG	175,2	LACHEMA		> 50 mg	AA	
thione metabolism	H	GC	T U	Cysteine, L	C-7755	121,2	Sigma	98%	> 50 mg	Skříň - A	
cine metabolism	H										
cine metabolism	H										
phytyl and chlorophyll me	H	GC	T U P	Glutamic,L acid	10MG	147,2	Calbiochem, NB		> 50 mg	AA	
cine and proline metabol	H										
amate metabolism	H	GC	T U P	Glutamine,L	10MG	146,2	Calbiochem		> 50 mg	AA	
ionine metabolism	H			Homocysteine		135,2	Loba		> 50 mg	AA	
cine and proline metabol	H										
e, leucine and isoleucin	H			Isoleucine,L	10MG	131,2	NBC, Calbioche		> 50 mg	AA	
e, leucine and isoleucin	H										
ophan metabolism	H										
e, leucine and isoleucin	H			Leucine,L	10MG	131,2	NBC		> 50 mg	AA	

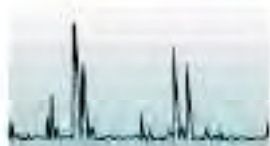
LETTERS

Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression

Arun Sreekumar^{1,2,3,4}, Laila M. Poisson^{5*}, Thekkelnayck M. Rajendiran^{1,3*}, Amjad P. Khan^{1,3*}, Qi Cao^{1,3}, Jindan Yu^{1,3}, Bharathi Laxman^{1,3}, Rohit Mehra^{1,3}, Robert J. Lonigro^{1,4}, Yong Li^{1,3}, Mukesh K. Nyati^{4,6}, Aarif Ahsan⁶, Shanker Kalyana-Sundaram^{1,3}, Bo Han^{1,3}, Xuhong Cao^{1,3}, Jaeman Byun⁷, Gilbert S. Omenn^{2,7,8}, Debashis Ghosh^{4,5,11}, Subramaniam Pennathur^{2,4,7}, Danny C. Alexander¹², Alvin Berger¹², Jeffrey R. Shuster¹², John T. Wei^{4,9}, Sooryanarayana Varambally^{1,3,4}, Christopher Beecher^{1,2,3} & Arul M. Chinnaiyan^{1,2,3,4,9,10}

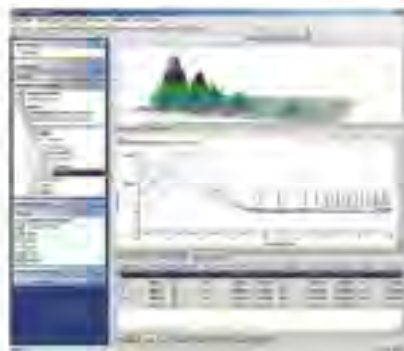


Add chromatographic standards



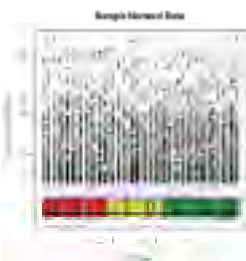
Mass spectra

Spectral Analysis



1. peak identification
2. peak deconvolution
3. peak alignment
4. database search
5. quantitation

Normalization



1. data imputation
2. median centering
3. scaling by IQR/Z normalization

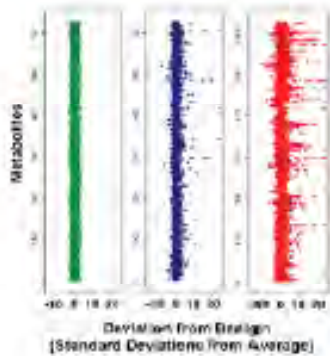


1. data imputation
2. median centering
3. scaling by IQR/Z normalization

Pathway mapping



Interpretation



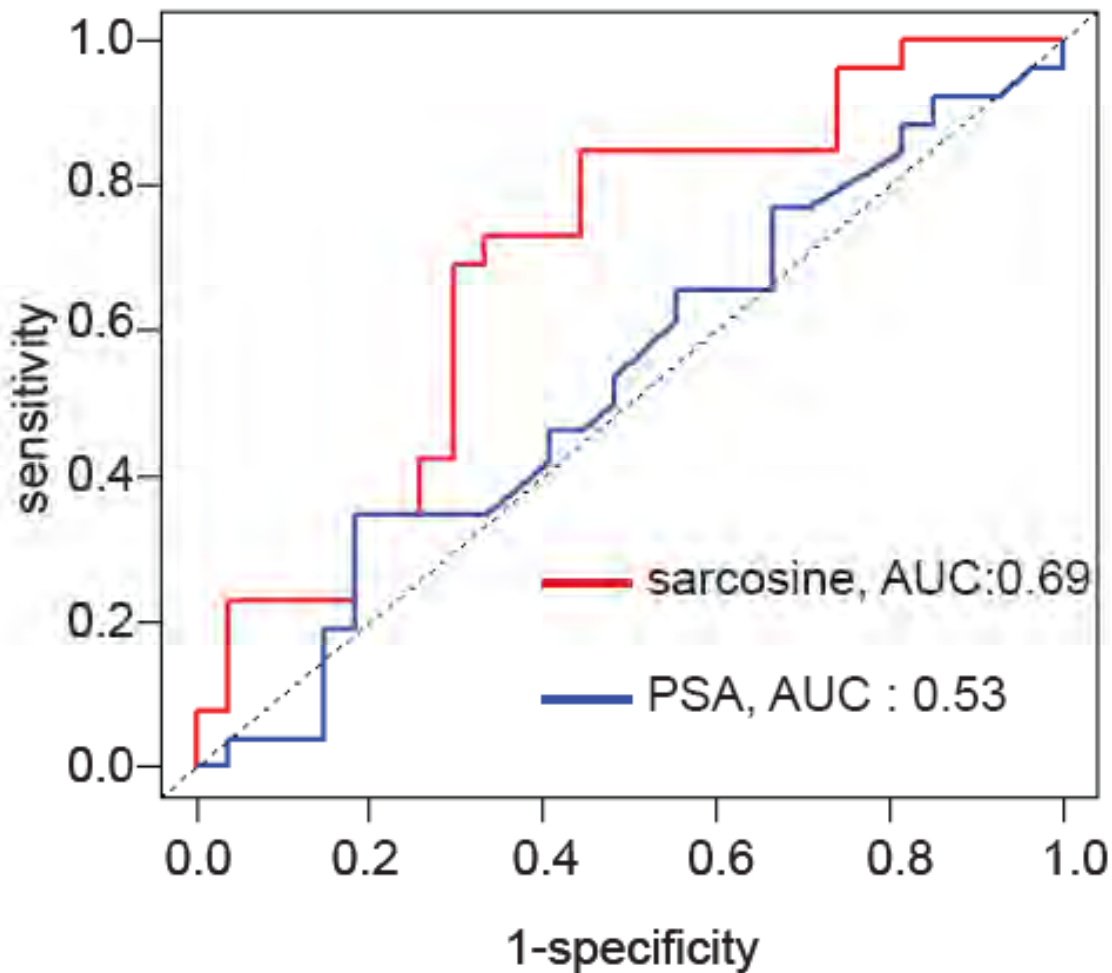
GC-MS



Validation

Class-specific metabolome

The AUC for sarcosine and PSA for patients having PSA levels in the clinically challenging range of 2-10 ng/ml



Supplementary Figure 15: Comparative Receiver Operator Curve (ROC) showing the relative discriminatory power of sarcosine assessed in urine sediments and serum PSA for the biopsy-proven cohort of patients.

Adopted from Sreekumar A, NATURE, 12 February 2009, p 910

Lecture summary

1. Definitions
2. Tools – many ways, reduce noise!
3. Data - 3D vs. 2D in CB, use more stats
4. Examples
5. Where are we (in Olomouc)?



UNIVERSITA PALACKÉHO
OLOMOUC



FAKULTNÍ NEMOCNICE
OLOMOUC



NTNU

Norwegian University of
Science and Technology



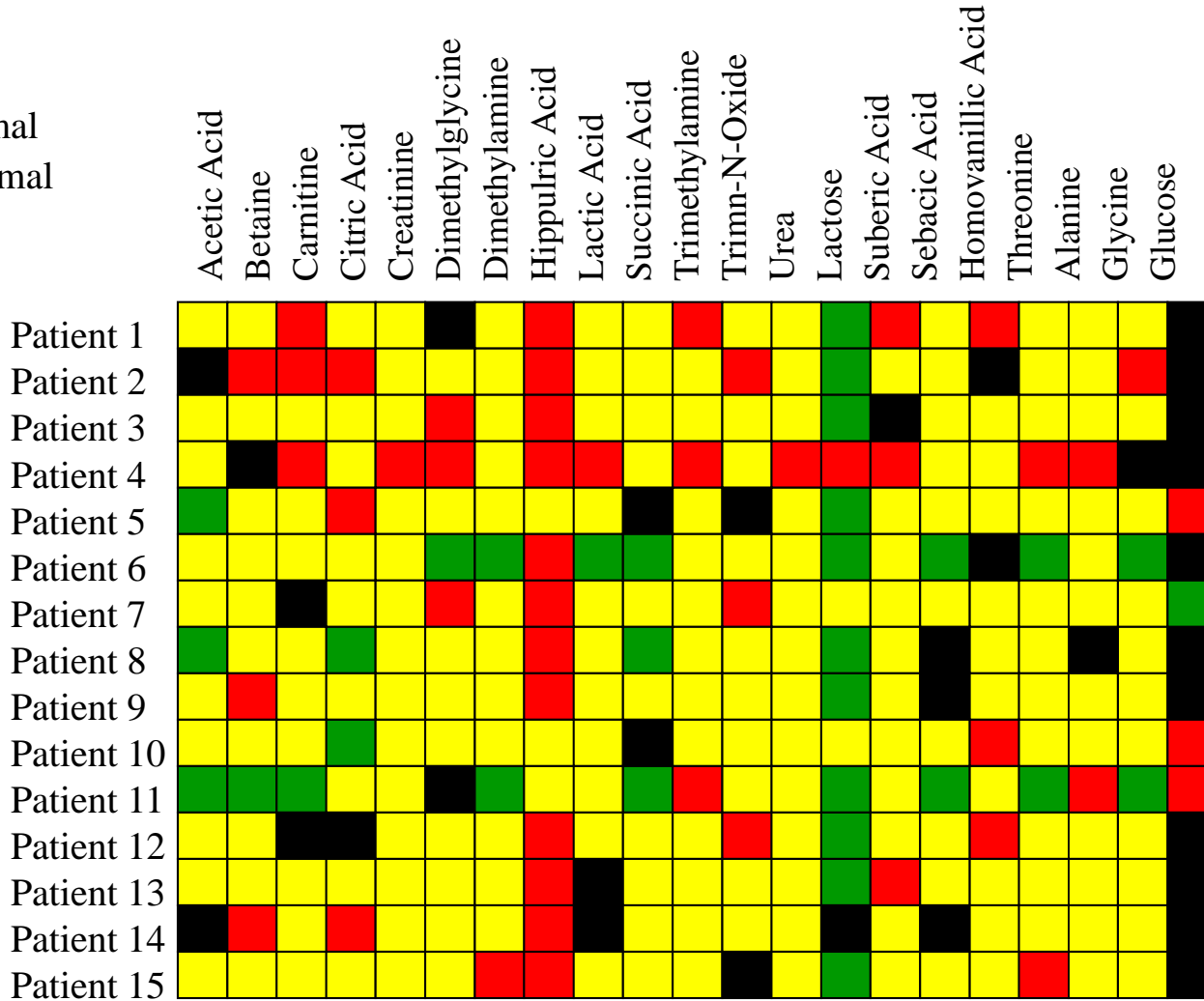
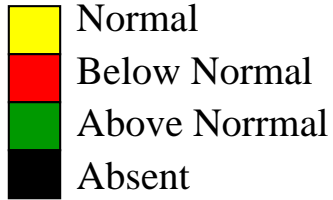
Grants:

MSM6198959205

A/CZ0046/2/0011

Děkuji za pozornost!

Metabolic Microarrays



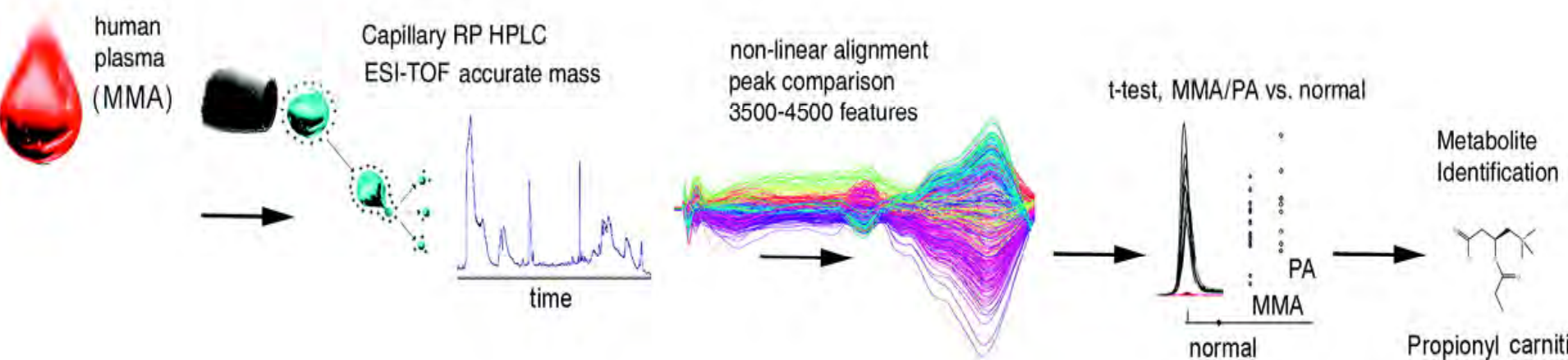
Data analysis and interpretation

- ... and frankly, the literature is absolutely full of complete rubbish resulting from a combination of over-optimism in the face of ostensibly positive findings, statistical ignorance and the fear of journals to scrutinize data too carefully lest they find something unpleasant...

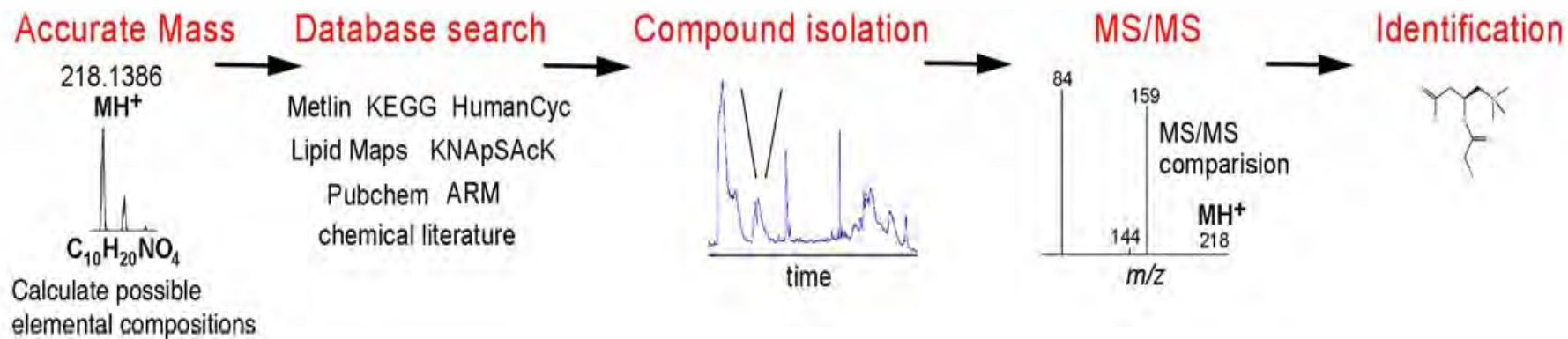
Kell DB



Finding Metabolomic Differences (untargeted)



Identification of Unknown Metabolites (targeted)



From Suizdak G, Clin Chem