Taiwan Biobank Metabolomics Workshop

2025.06.12 | Taiwan Biobank

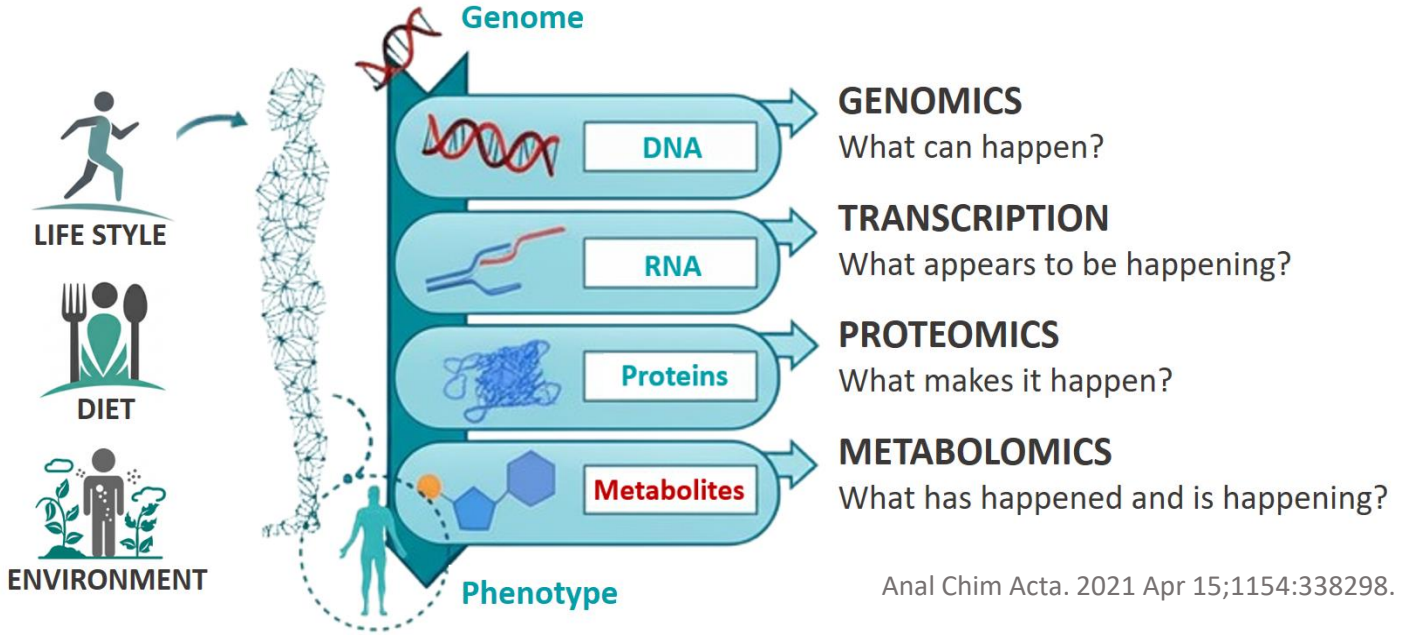
Outline

- Introduction to Metabolomics
- NMR Technology in Metabolomics
- Overview of the Taiwan Biobank Metabolomics Dataset
- Data Analysis Workflow of TWB NMR Metabolomics Dataset
- Demonstration / Hands-on

Introduction to Metabolomics

What is Metabolomics?

- **Metabolomics** is the comprehensive study of the small molecules that reflect the physiological state of a biological system.
- It provides a **metabolic snapshot** that links genetic and environmental changes to observable traits, thereby bridging the gap between **genotype and phenotype.**



Anal Chim Acta. 2021 Apr 15;1154:338298.

Introduction to Metabolomics

The Evolution of Metabolomics: A Chronological Journey

300 B.C. | 1614 | 1913 | 1946 | 1971 | 1984 | 1990s | 2000s | 2010s | 2020s

Mass spectrometry
Thomson J.J. and Anderson F.

Quantitative basis of metabolism
Sanctorius S.

NMR
Bloch F. and Purcell E.

NMR detects diabetes
Nicholson J.K. et al.

- **"Metabolome" coined**
  Oliver S.G. et al.
- KEGG

Body fluids to predict disease

- 2000–1500 BC: Ant test for urine glucose in ancient China
- 300 BC: Urine taste used for disease diagnosis in Egypt & Greece

Early era of metabolomics
Pauling L. et al.

Metabolic profile
Horning E.C. and Horning M.G.

- 2003 METLIN
- 2004 The Metabolomics Society
  The Human Metabolome Project
- 2005 HumanCyc and KaPPA-view
- 2007 HMDB
- 2009 ProMeTra

- 2010 MetPA
- 2011 Reactome, iPath, Pathos, Paintomics
- 2012 MetaMapp, MetScape
- 2013 Bruker IVDr
- 2014 SMPDB
- 2015 Real-time metabolome profiling
- 2016 BioCyc

- Integration with multi-omics and AI
- Percision medicine

Pre-metabolomics era
Instrumental and methodological development

Post-metabolomics era
Databases, tools and high-throughput techniques development

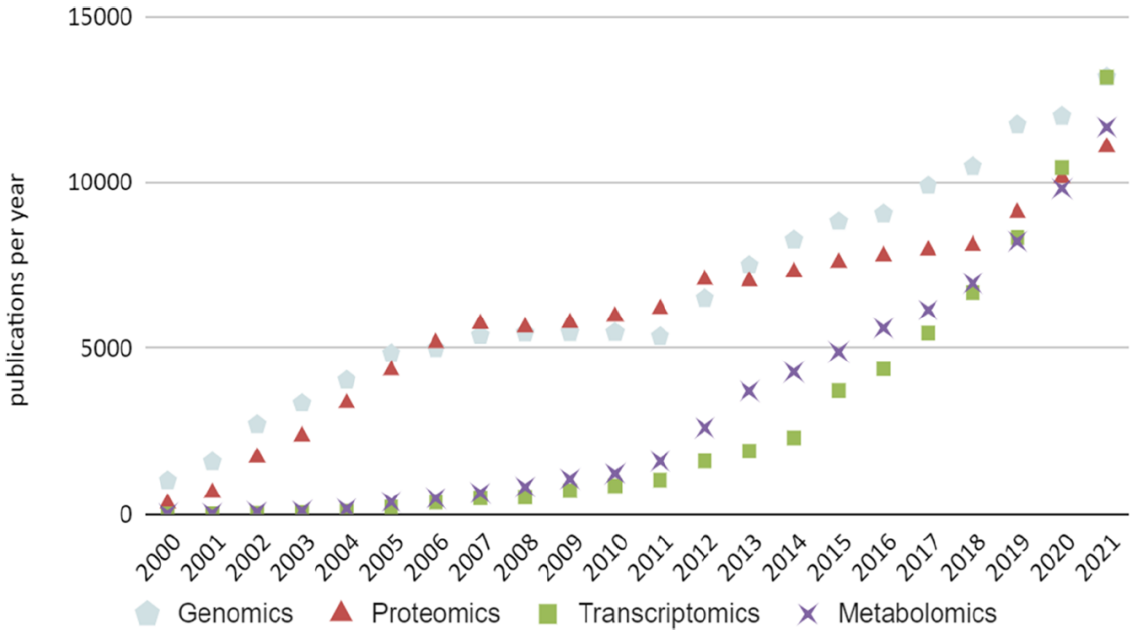
Introduction to Metabolomics

Why Metabolomics Matters?

- Captures the **end-point of phenotype** and real-time physiological state
- Sensitive to **early disease detection**, **drug response**, and **environmental influences**
- Enables **biomarker discovery** and advances **precision medicine**
- Integrates with other omics for **systems biology**

Challenges and Future Directions

- **High complexity and variability**, with **limited metabolite annotation**
- Requires **standardization**, **advanced analytical tools**, and **AI/ML integration**
- Expanded applications
  Public health, nutrition, sports science, mental health, aging, toxicology



Anal Chem. 2023 Feb 28;95(8):3909-3916.

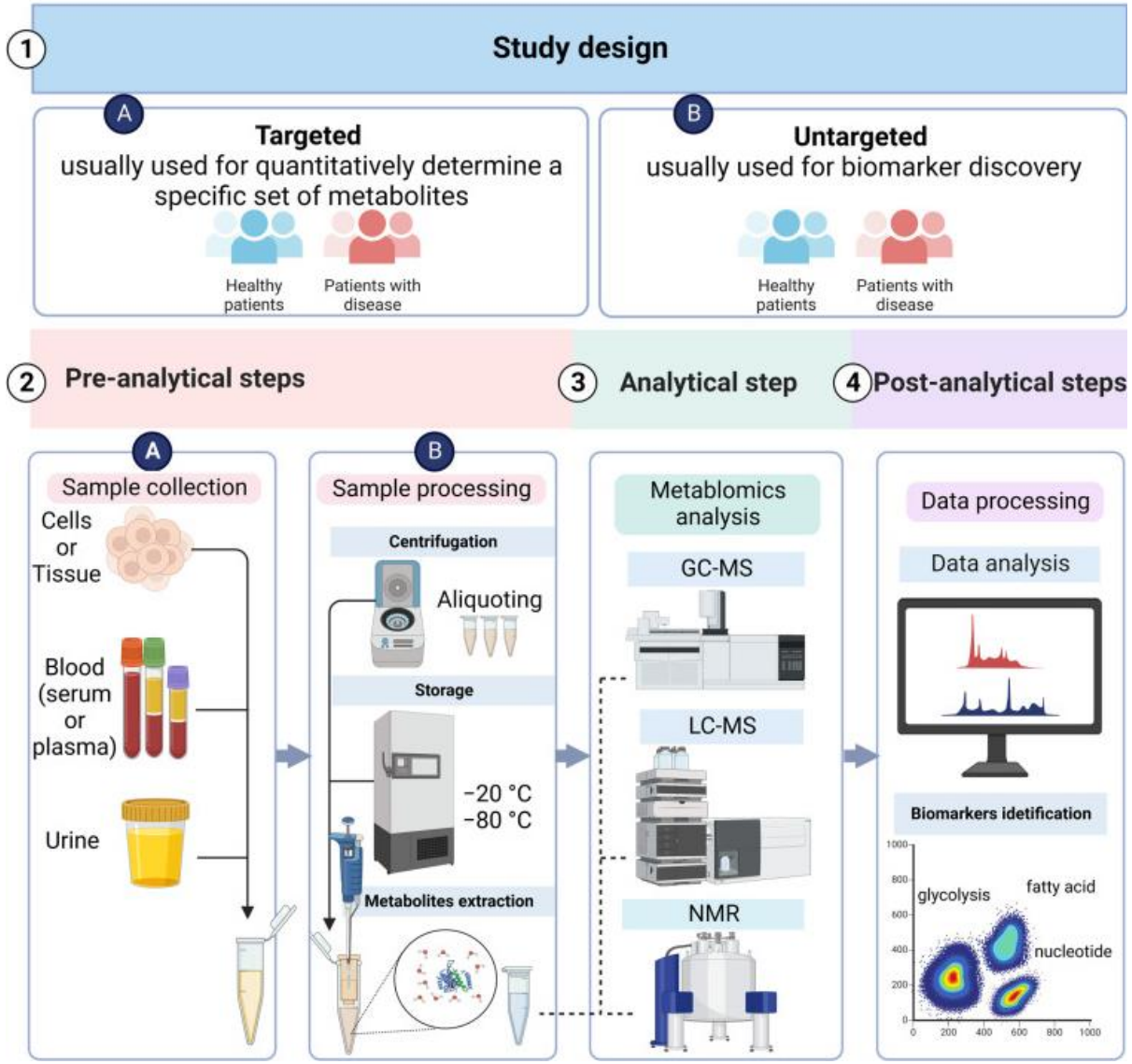
Introduction to Metabolomics

Types of Metabolomics

- Targeted
- Untargeted

Metabolomics Workflow Steps

- Study design
- Sample collection and processing
- Platform choice
- Data processing, results interpretation, and biomarker identification



Diagnostics (Basel). 2023 Feb 23;13(5):861

Comparison of NMR, GC-MS, and LC-MS in Metabolomics

Feature	NMR	LC-MS	GC-MS
Sensitivity	Low (10^{-6} mol/L)	High (10^{-15} mol/L)	High (10^{-12} mol/L)
Reproducibility	High	Moderate	Low
Quantitative analysis	Quantitative	Not very quantitative	Quantitative
Metabolite Coverage	Limited (< 200)	More	Few
Sample Preparation	Simple (buffer addition only)	Moderate (requires extraction, desalting, etc.)	Complex (requires volatilization; often needs derivatization)
Sample Destruction	Non-destructive	Destructive	Destructive
Throughput	High (fixed analysis time, simple workflow)	Moderate to high (depends on automation and setup)	Moderate to high (fast analysis but needs preprocessing)
Instrument Cost	High	Moderate	Moderate
Application	• Absolute quantification • High-abundant metabolites • Structural elucidation	• Non-targeted/targeted metabolomics • Biomarker discovery	• Volatile metabolite • Plant metabolomics • Environmental studies

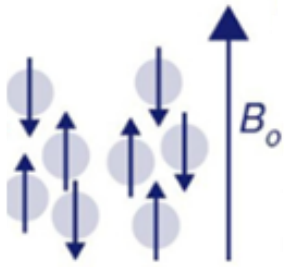
Diagnostics (Basel). 2023 Feb 23;13(5):861

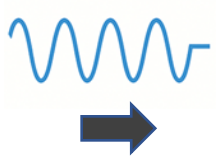
NMR Technology in Metabolomics

Basic Principle of NMR (Nuclear Magnetic Resonance)

Nuclear

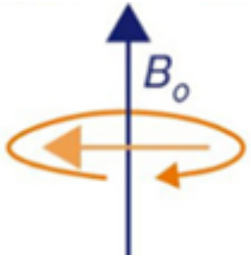
Magnetic

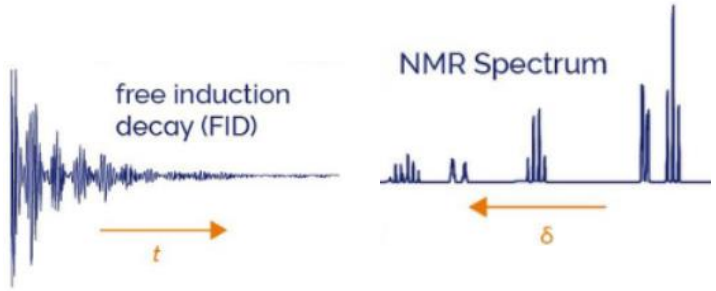




RF Pulse Excitation

Resonance





Before Magnetic Field

Inside atoms, tiny particles called nuclei spin in all directions. They are not affected by any magnetic force.

In Magnetic Field

When placed in a strong magnetic field, some of the spins line up with the field. This creates a tiny magnetic signal we can measure.

Signal Excitation and Relaxation

A short burst of energy makes the spins tip out of alignment. As they relax back, they release signals that we can detect.

From Signal to Spectrum

These signals are like echoes. A computer turns them into a chart showing what kinds of molecules are in the sample.

https://nmr.oxinst.com/what-is-benchtop-nmr

Bruker IVDr Platform

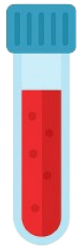
What is IVDr (in vitro diagnostic research)?

- A standardized, high-throughput **NMR-based metabolomics platform**
- Designed for **in vitro diagnostic research** across multiple biological fluids (e.g., plasma, serum, urine)
- Provides **quantitative** profiling of metabolites and lipoproteins with high **reproducibility** and strong **inter-laboratory consistency**.
- Automated sample preparation, data acquisition, and analysis using SOPs
- Globally harmonized system across labs using the same hardware/software

Applications & Benefits

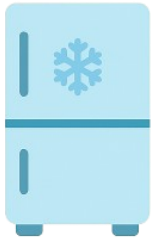
- Clinical research, Epidemiology and population studies, Longitudinal cohort monitoring and Early disease biomarker discovery
- Ready-to-use interpretation modules

Overview of the Taiwan Biobank Metabolomics Dataset



Sample type and collection

- Plasma was collected in sodium citrate tubes
- Standardized phlebotomy and processing protocols are applied across all sites
- Tecan automatic liquid handling platform



Sample storage

- Plasma aliquots stored at −80 ˚C
- Storage duration: up to 16 years (range: 0–16 years)

NMR platform

- Bruker IVDr system based on the UltraShield Plus 600 MHz NMR spectrometer
- Analysis modules

  B.I.-QUANT-PS™ for molecule quantification, and B.I.-LISA™ for lipoprotein

Data characteristics

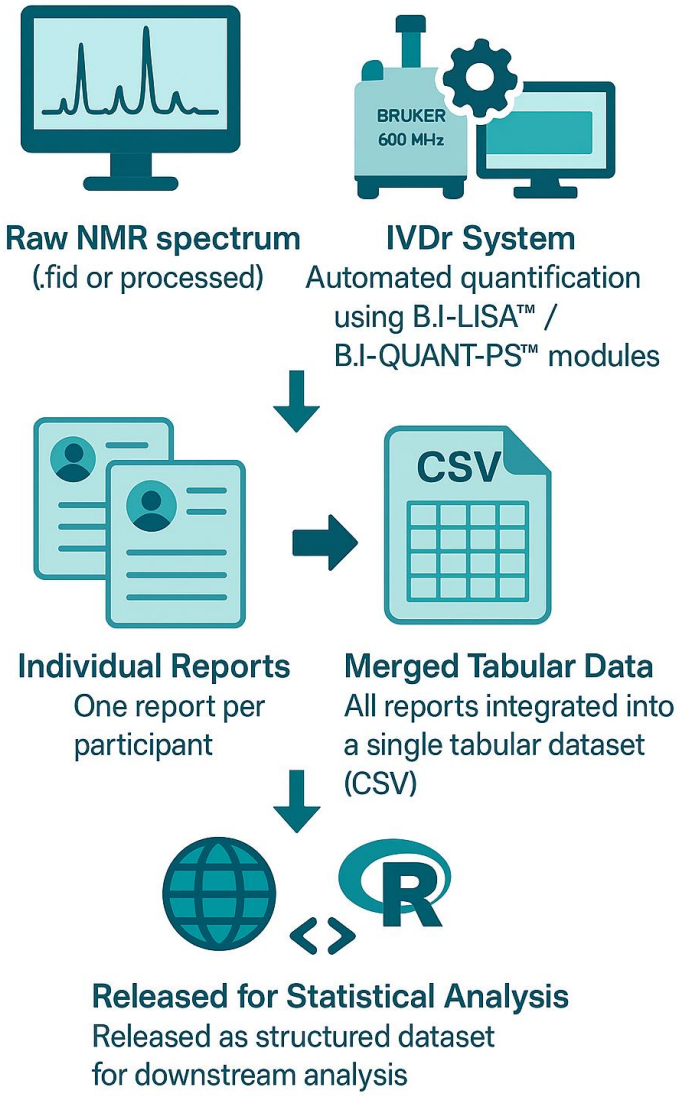
Overview of the Taiwan Biobank Metabolomics Dataset

Two formats of data provided

- Raw NMR spectra (.fid, .1i, .1r) for custom processing
  Bruker TopSpin software
- Analysis-ready matrix (.csv) processed via Bruker IVDr system

Sample and metabolite information

- Total sample size: 4,210
- Metabolites quantified: ~150
- Visit https://taiwanview.twbiobank.org.tw/ for details



Taiwan Biobank Metabolomics Dataset: Metabolites Coverage

41 Metabolite Concentrations

- Alcohols and derivatives
- Amines and derivatives
- Amino acids and derivatives
- Carboxylic acids
- Essential nutrient
- Keto acids and derivatives
- Sugars and derivatives
- Sulfones
- Technical additives

Analysis Report

Bruker **I**VDr **Quant**ification in **P**lasma/**S**erum B.I.Quant-PS™

Sample ID: APGK_expno10.100000.11r

Measuring Date: 14-Feb-2023 11:25:35
Reporting Date: 14-Feb-2023 12:53:51, 7 page(s), Version 2.0.0
Quantification Method Version: Quant-PS 2.0.0

3 Amino acids and derivatives

Compound	Conc. mmol/L	LOD mmol/L	r mmol/L	ρ %	Δ mmol/L	95% Range mmol/L	Graphics (*)
2-Aminobutyric acid	< 0.05	0.05	0.000	0 ○	0.894	⩽ 0.10	
Alanine	0.47	0.02	0.474	100 ●	0.007	0.29 - 0.64	
Asparagine	< 0.05	0.05	0.000	0 ○	5.494	⩽ 0.08	
Creatine	< 0.01	0.01	0.003	99 ●	0.003	⩽ 0.07	
Creatinine	0.09	0.01	0.090	99 ●	0.003	0.06 - 0.14	
Glutamic acid	< 0.05	0.05	0.000	0 ○	1.881	⩽ 0.24	
Glutamine	1.00	0.02	1.038	98 ●	0.043	0.30 - 0.83	
Glycine	0.26	0.01	0.256	100 ●	0.005	0.17 - 0.44	
Histidine	0.09	0.02	0.085	99 ●	0.002	0.07 - 0.16	
Isoleucine	0.08	0.03	0.077	97 ●	0.009	0.03 - 0.11	
Leucine	0.15	0.01	0.152	97 ●	0.013	0.07 - 0.20	
Lysine	0.26	0.04	0.259	70 ○	0.067	⩽ 0.29	
Methionine	0.07	0.05	0.065	74 ○	0.013	0.05 - 0.13	
N,N-Dimethylglycine	< 0.01	0.01	0.007	97 ●	0.001	⩽ 0.01	
Ornithine	0.03	0.02	0.034	52 ○	0.032	⩽ 0.16	
Phenylalanine	0.04	0.03	0.042	98 ●	0.003	⩽ 0.07	
Proline	0.86	0.05	0.862	80 ○	0.231	⩽ 0.59	
Sarcosine	< 0.01	0.01	0.006	78 ○	0.001	⩽ 0.01	
Threonine	0.29	0.04	0.293	54 ○	0.372	⩽ 0.24	
Tyrosine	0.06	0.03	0.061	97 ●	0.005	⩽ 0.08	
Valine	0.33	0.03	0.330	100 ●	0.009	0.15 - 0.35	

(*) Gray horizontal boxes represent 95% concentration range, black vertical lines represent sample value.

Taiwan Biobank Metabolomics Dataset: Metabolites Coverage

112 Lipoprotein parameters

Major classes

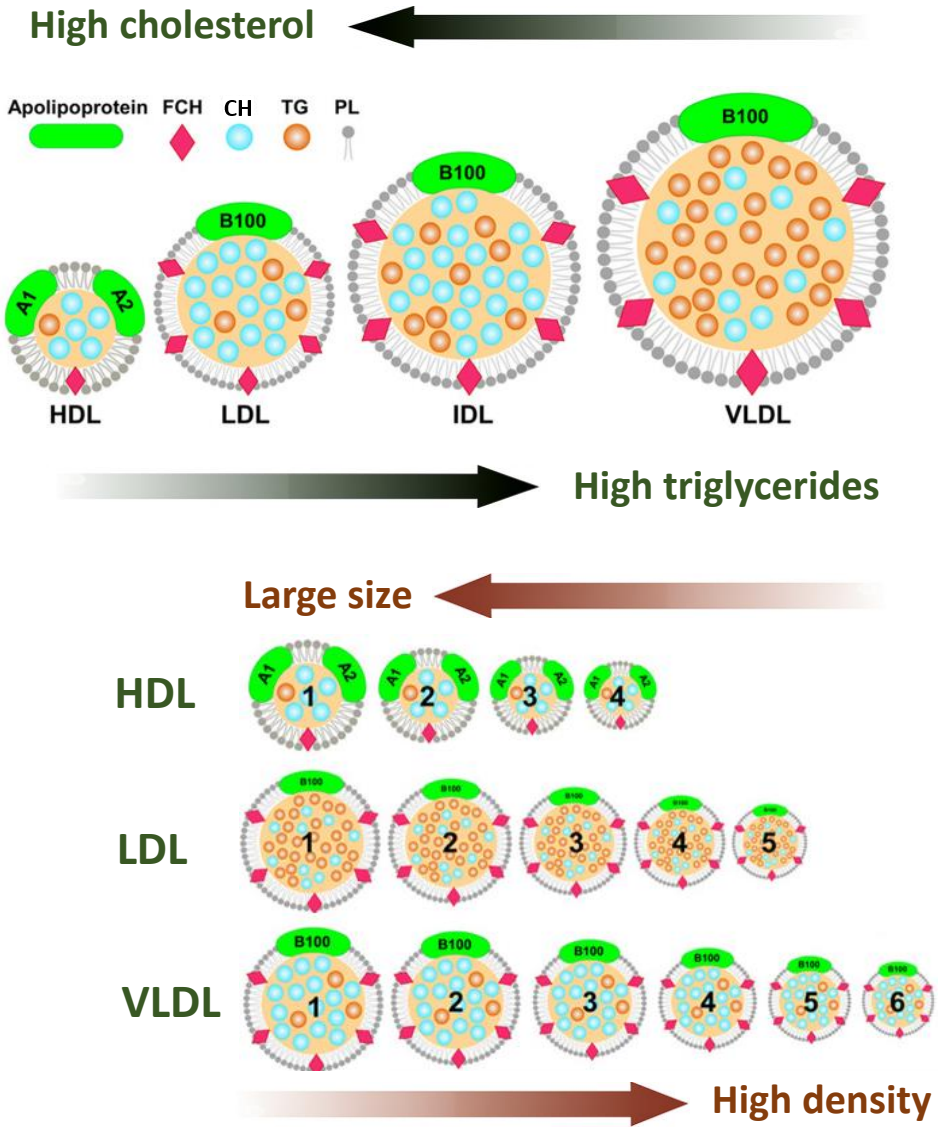
- VLDL, IDL, LDL, and HDL

Subclasses

- LDL-1 to LDL-6,…, HDL-1 to HDL-4

For each class and subclass

- Particle number (PN)
- Phospholipids (PL)
- Free cholesterol (FC)
- Esterified cholesterol (CH)
- Triglycerides (TG)
- Apolipoproteins (A1, A2, AB)



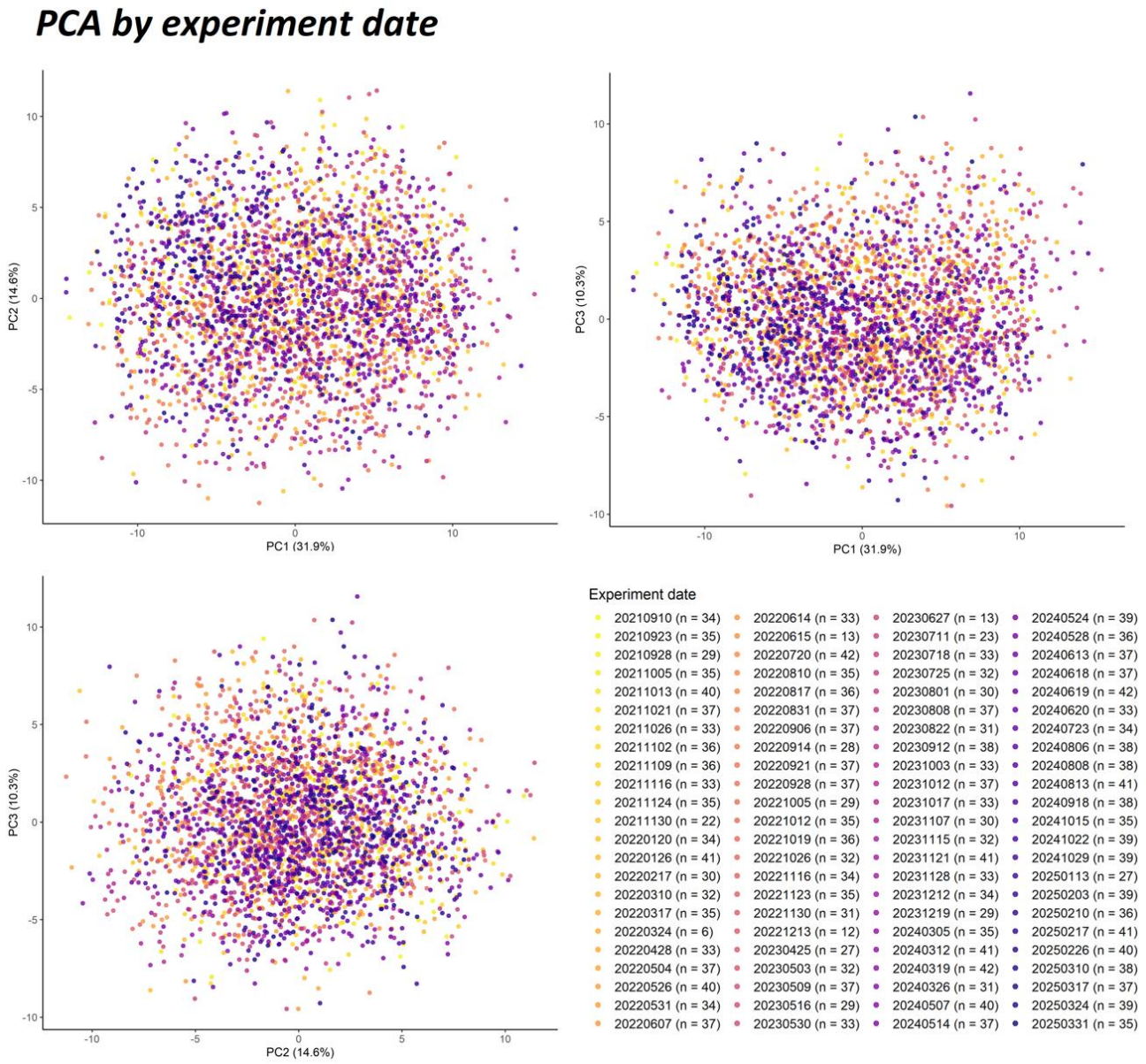
J Proteome Res. 2022 Apr 1;21(4):1061-1072.

Taiwan Biobank Metabolomics Dataset: Data Properties and Validation

	TWB Metabolomics Cohort at baseline		
	Male n = 1,350	Female n = 2,177	Overall n = 3,527
Age	51.3 (11.1)	51.2 (10.7)	51.3 (10.8)
Fasting time (hours)	10.7 (3.4)	10.6 (3.6)	10.5 (3.6)
Alcohol consumption	90 (6.7)	34 (1.6)	124 (3.5)
Smoking status	182 (13.5)	28 (1.3)	210 (6.0)
Betel nut chewing	40 (3.0)	2 (0.1)	42 (1.2)
BMI (kg/m^2)	25.6 (4.1)	23.9 (3.9)	24.5 (4.1)
Experiment year			
2021	281 (20.8)	276 (12.7)	557 (15.8)
2022	116 (8.6)	522 (24.0)	638 (18.1)
2023	464 (34.4)	492 (22.6)	956 (27.1)
2024	338 (25.0)	608 (27.9)	946 (26.8)
2025	151 (11.2)	279 (12.8)	430 (12.2)
Blood storage time(years)	7.1 (2.6)	7.4 (2.7)	7.2 (2.7)

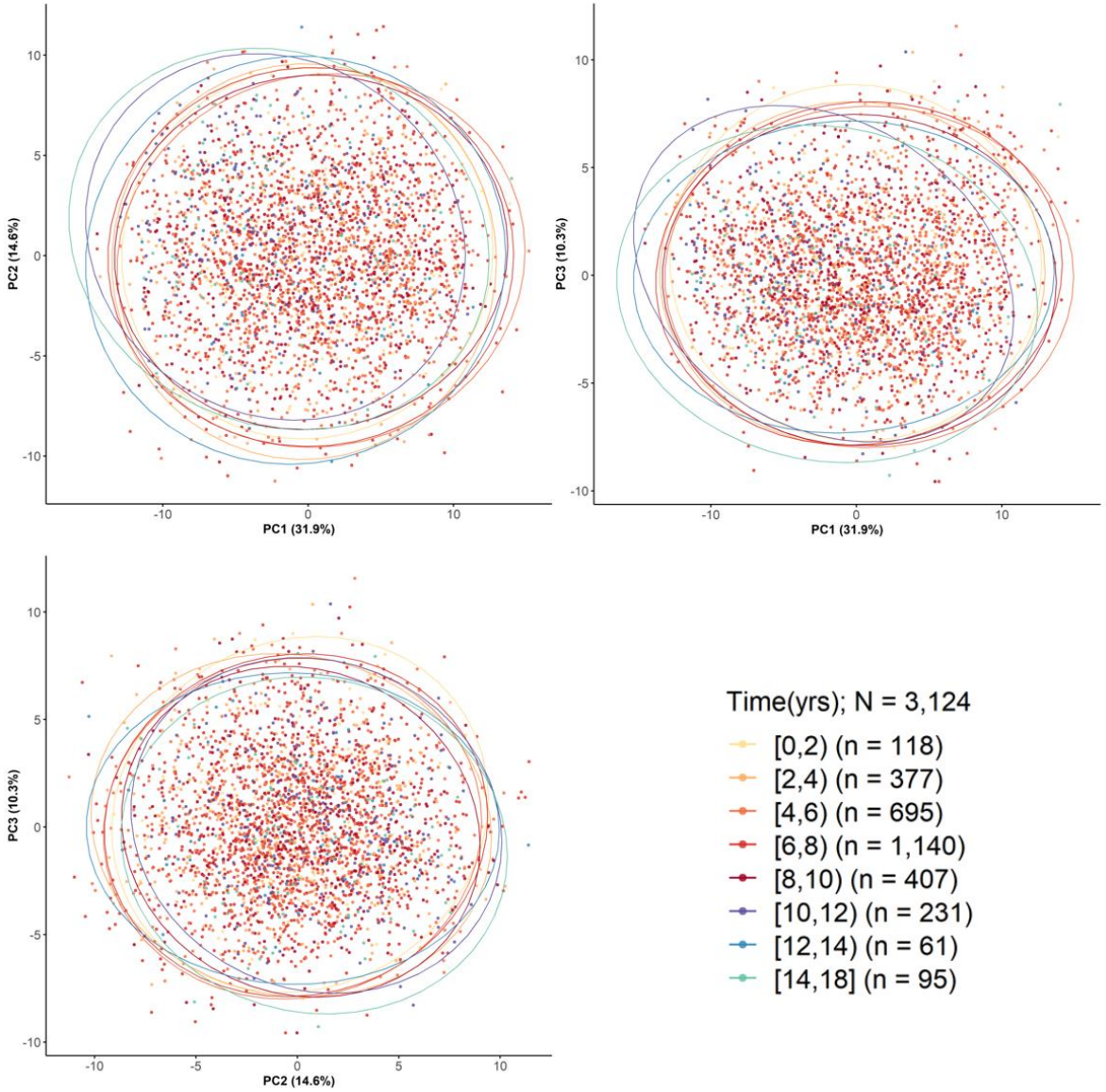
1. Continuous variables are presented as mean (S.*D.*); categorical variables are shown as n (%).
2. This table represents baseline participants. The TWB cohort additionally includes 683 individuals with follow-up data.

Taiwan Biobank Metabolomics Dataset: Data Properties and Validation

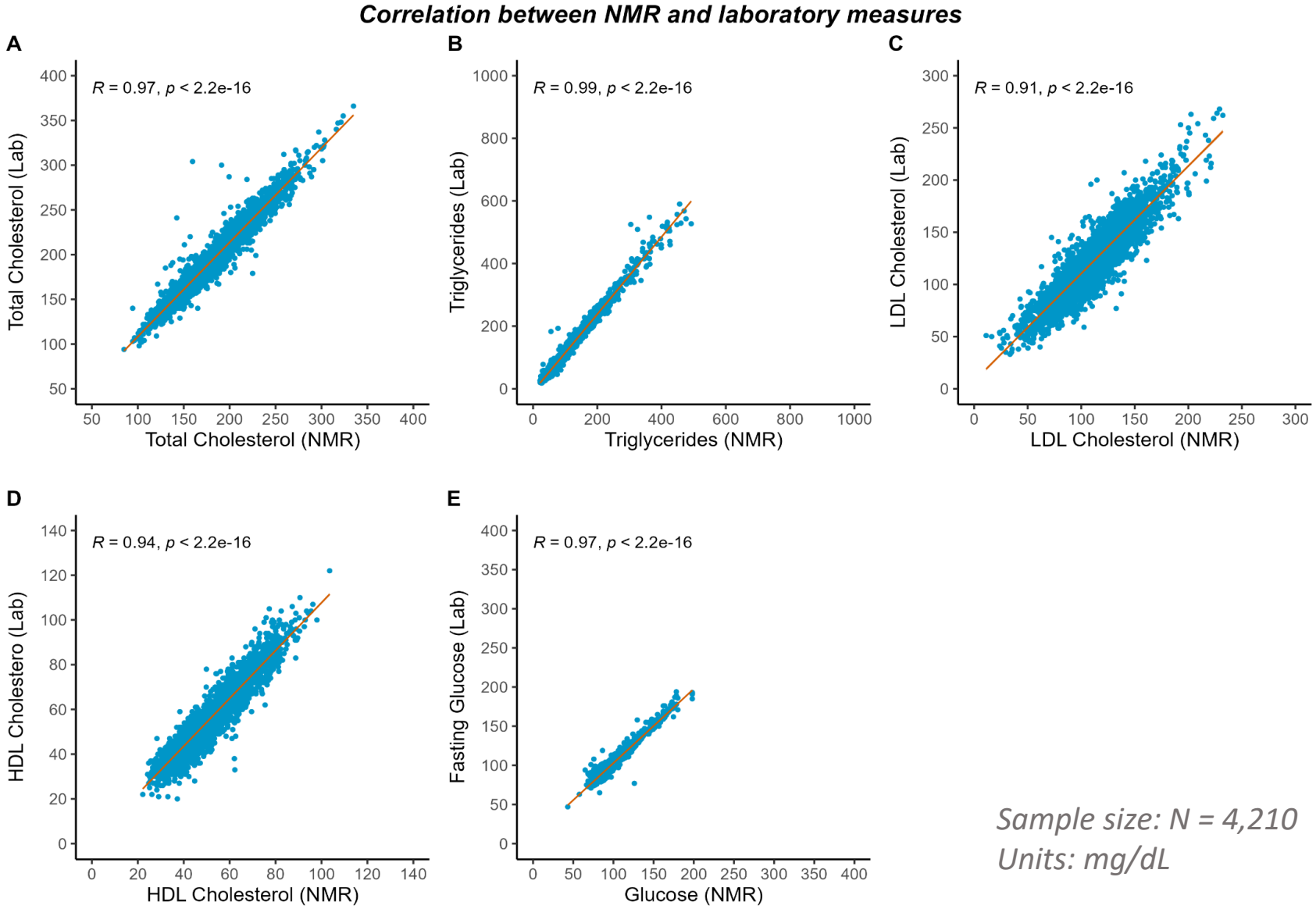


Taiwan Biobank Metabolomics Dataset: Data Properties and Validation

Blood storage duration before metabolomics experiment

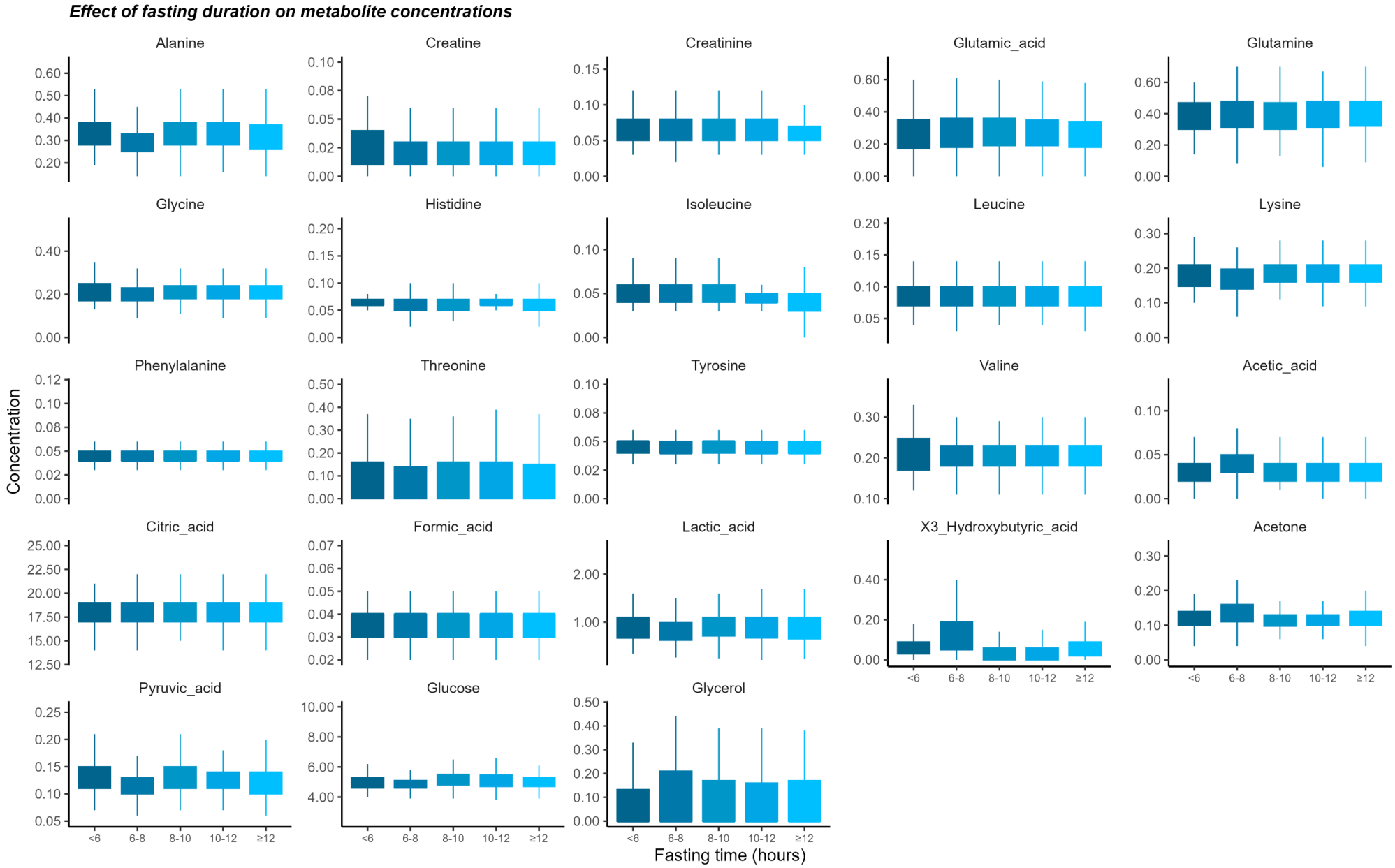


Taiwan Biobank Metabolomics Dataset: Data Properties and Validation

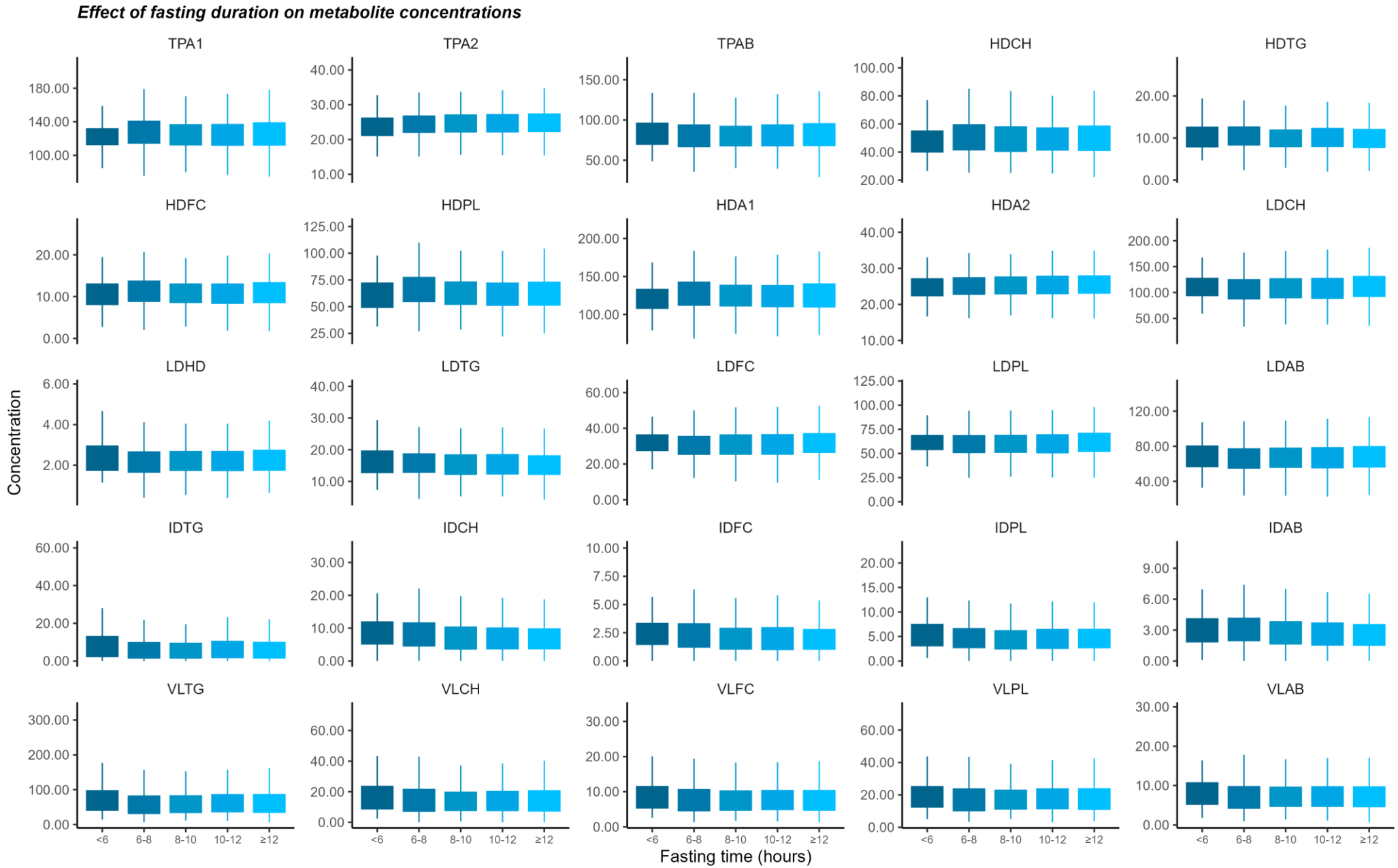


Sample size: N = 4,210
Units: mg/dL

Taiwan Biobank Metabolomics Dataset: Data Properties and Validation

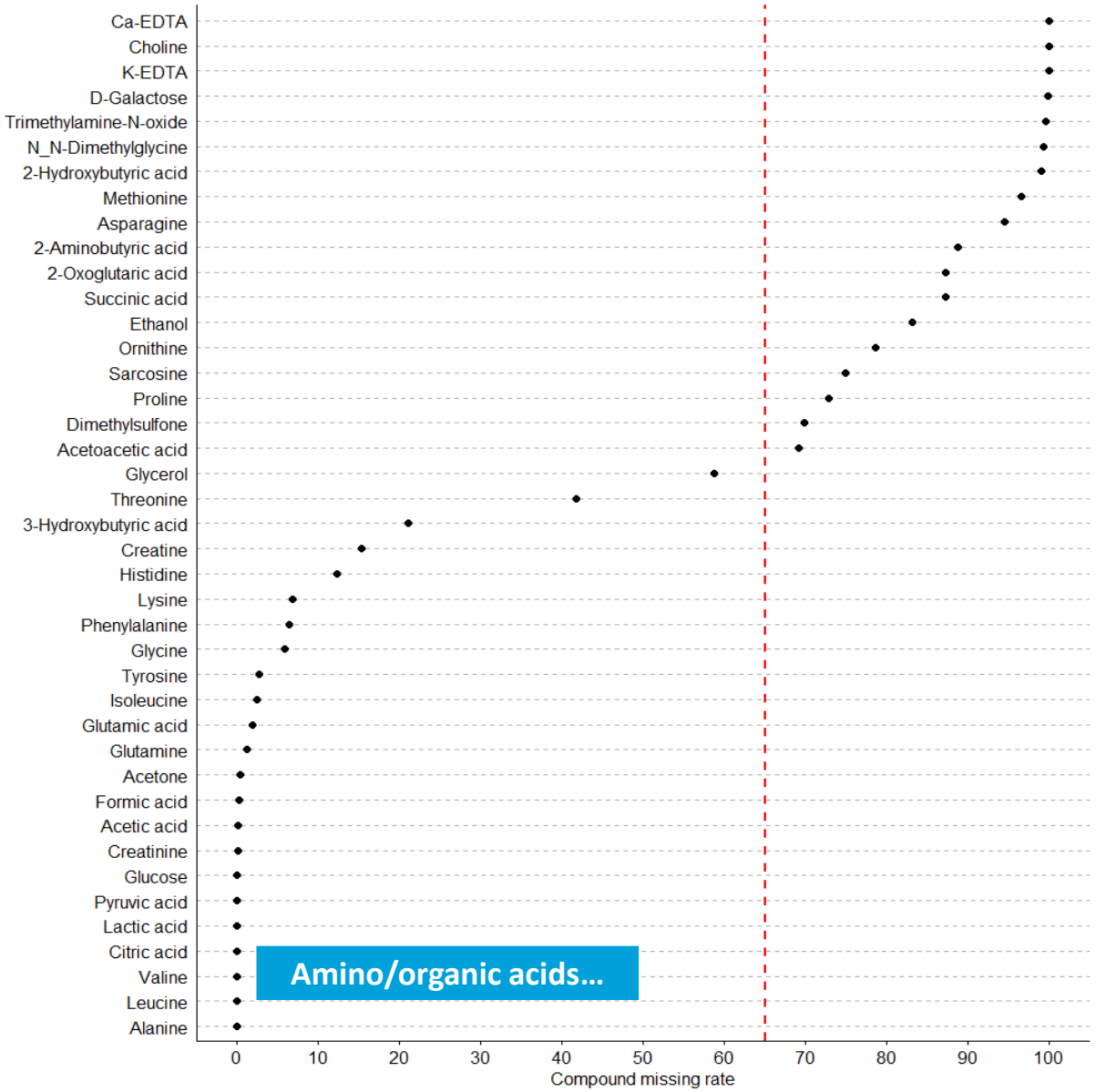


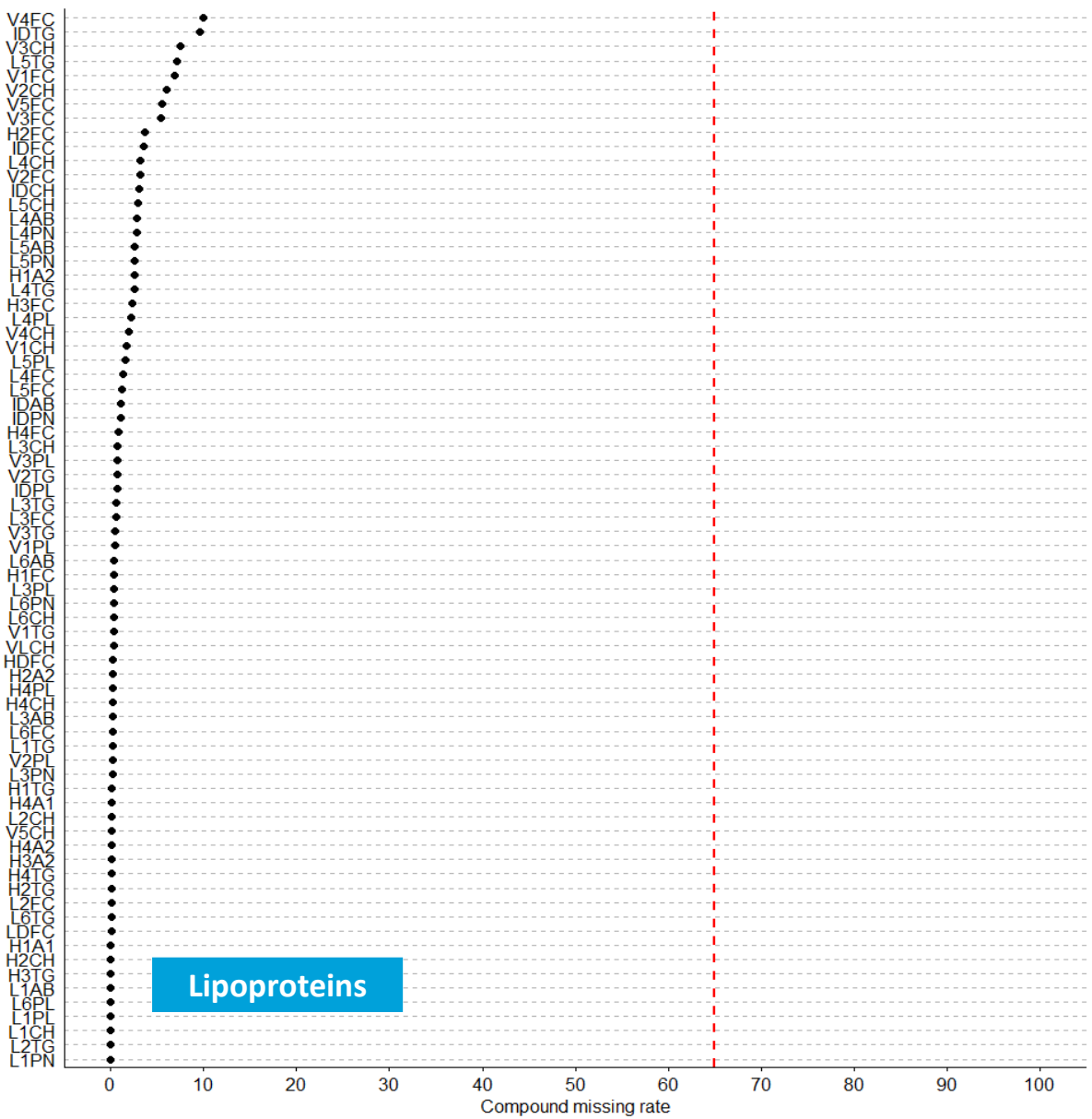
Taiwan Biobank Metabolomics Dataset: Data Properties and Validation



Taiwan Biobank Metabolomics Dataset: Data Properties and Validation

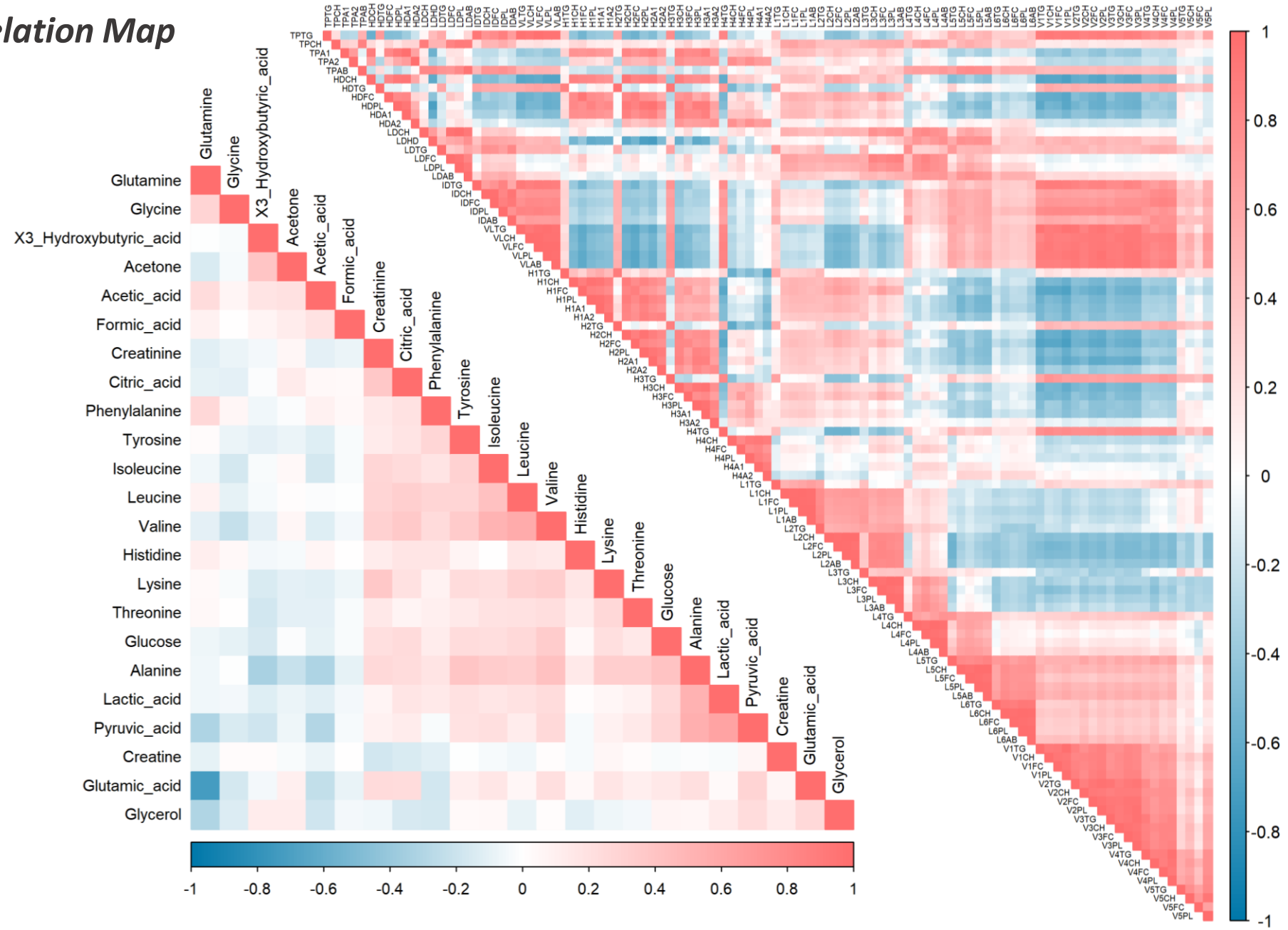
The missing rate for all 153 compounds



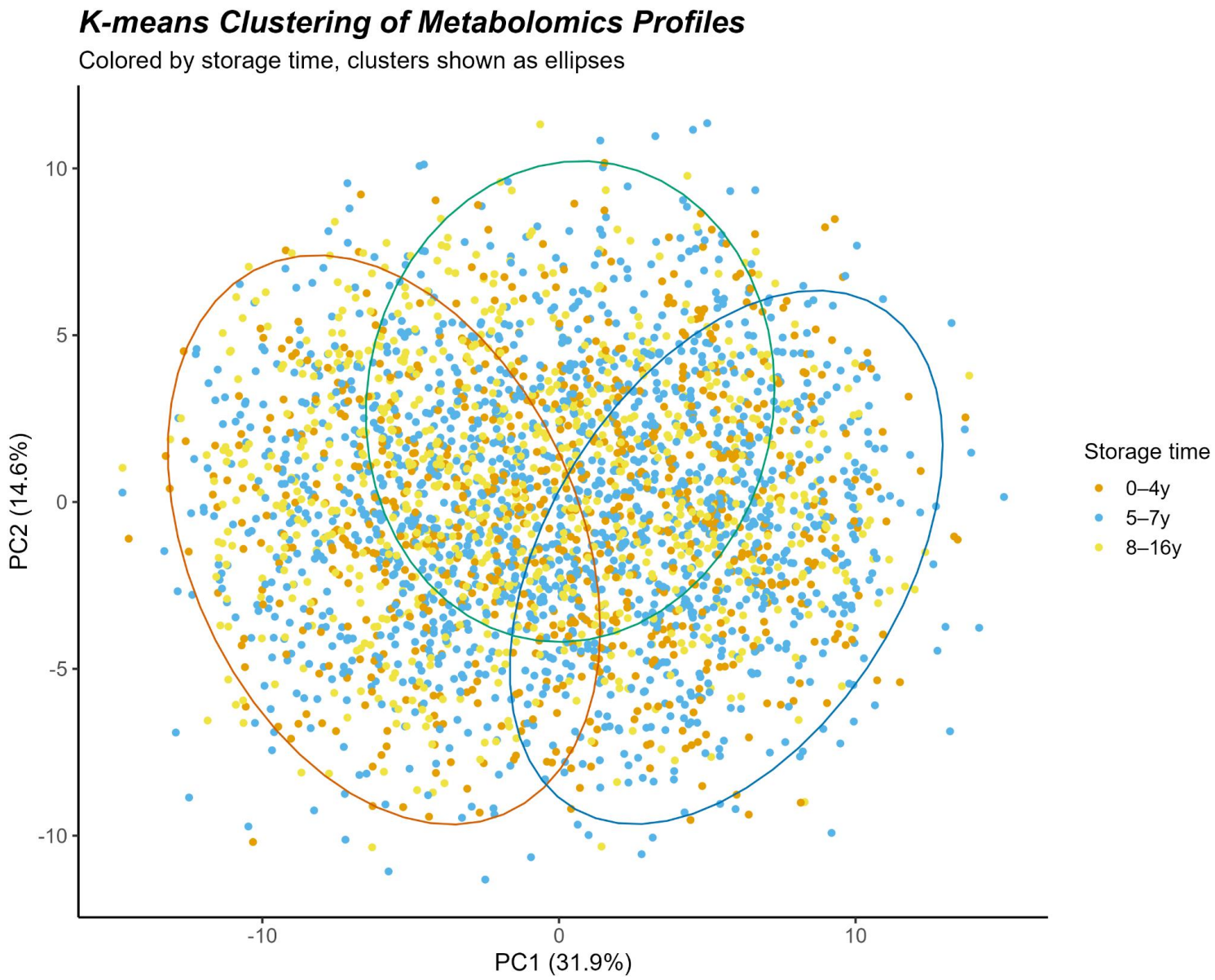


Taiwan Biobank Metabolomics Dataset: Data Properties and Validation

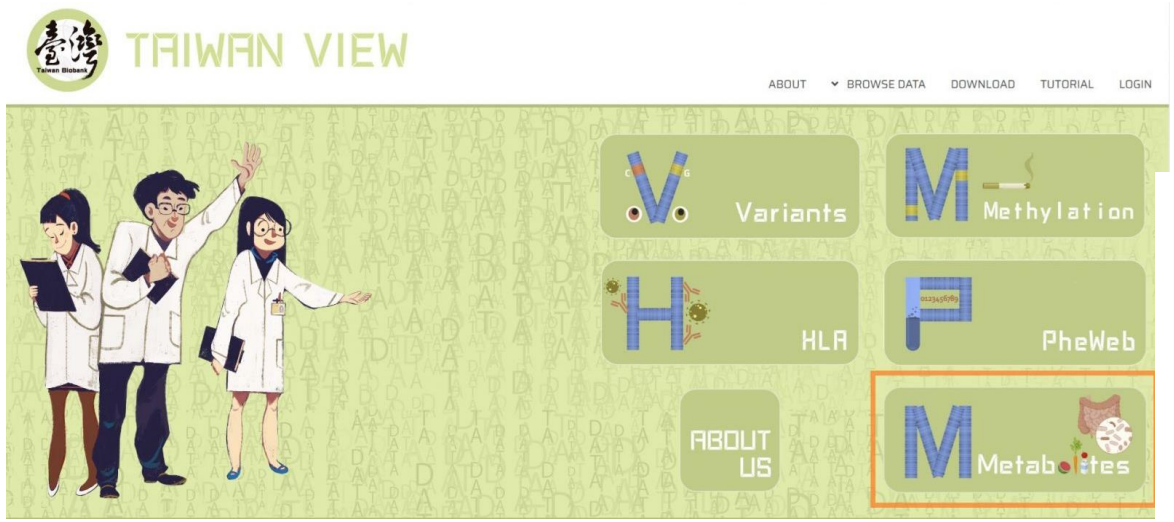
Metabolite Correlation Map

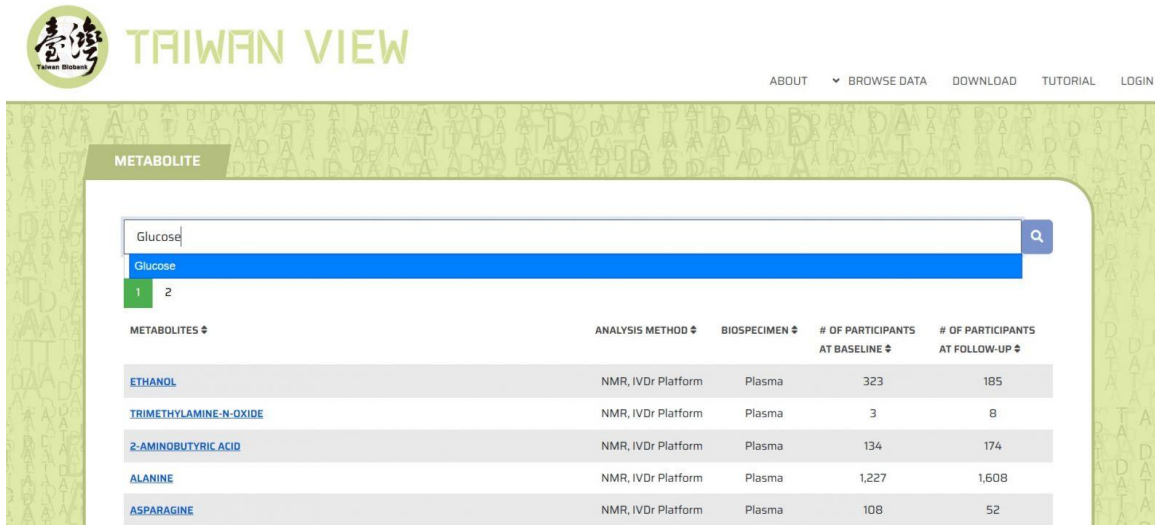


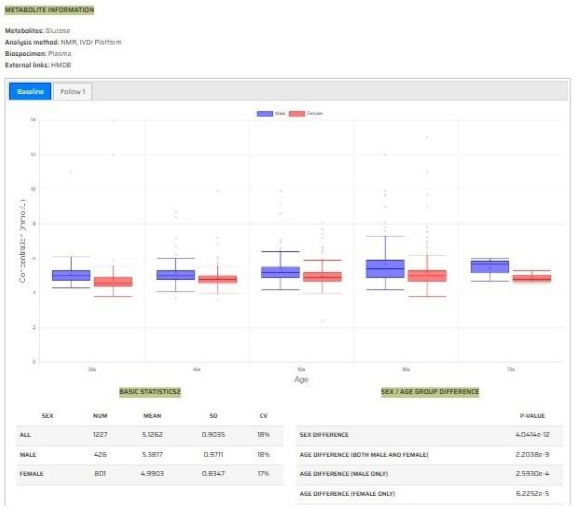
Taiwan Biobank Metabolomics Dataset: Data Properties and Validation

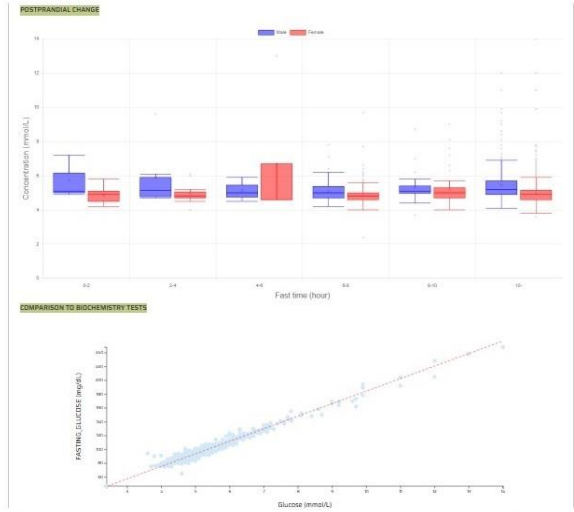


Web-Based Interactive Summary of TWB Metabolomics Data: Taiwan View









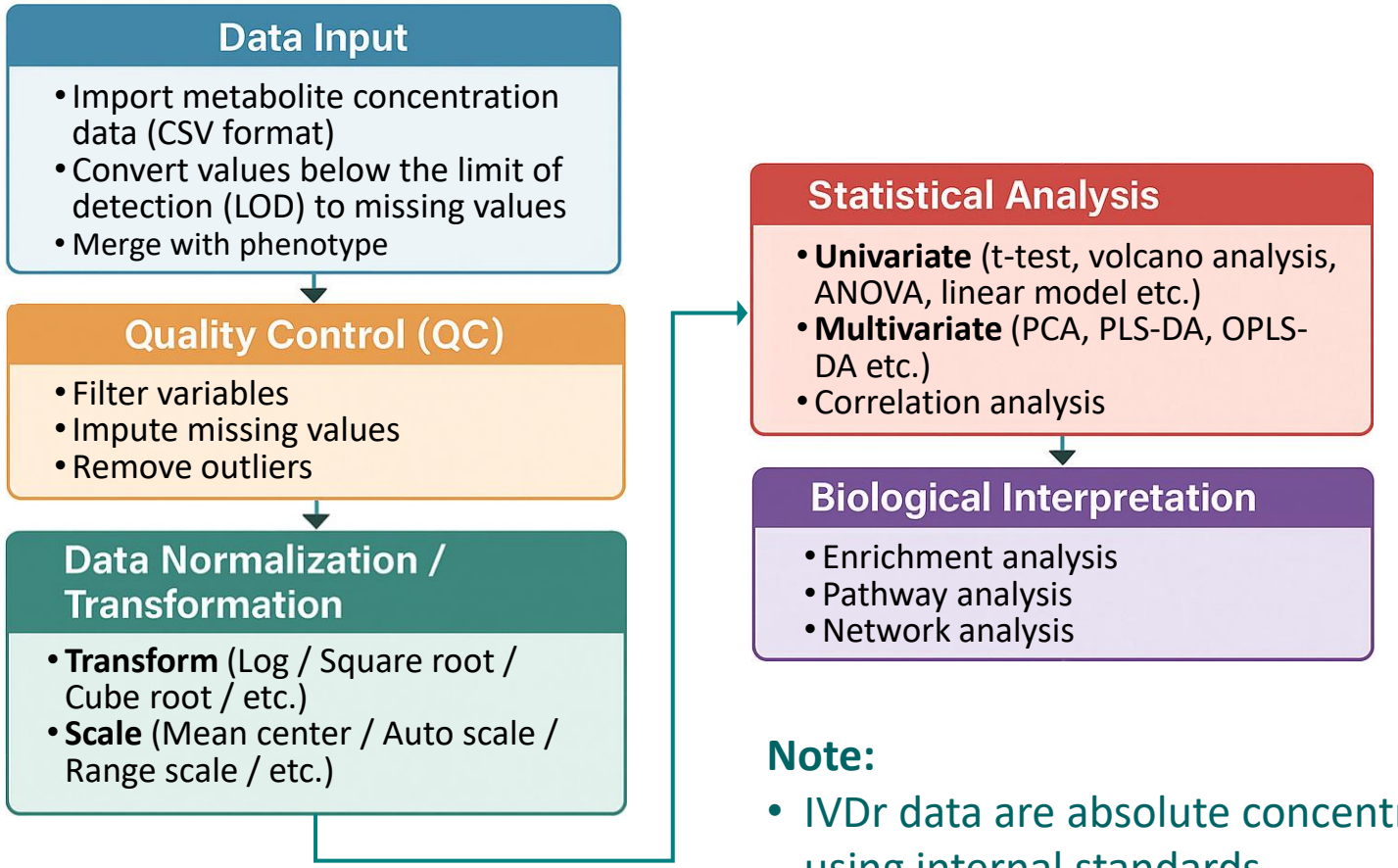
CORRELATION BETWEEN OTHER METABOLITES

ALL | Male | Female

METABOLITES	SPEARMAN'S RANK CORRELATION COEFFICIENT	P-VALUE
2-Aminobutyric acid	1.7328e-1	4.5264e-2
3-Hydroxybutyric acid	-1.6063e-1	4.7349e-7
Acetoacetic acid	-1.3634e-1	1.9690e-3
Alanine	3.8935e-1	1.0850e-45
Calculated Figures, Apo-A1 / Apo-B100, Apc-B100/Apo-A1	2.0597e-1	3.1991e-13
Calculated Figures, IDL Particle Number, IDL Particle Number	1.1596e-1	4.6709e-5
Calculated Figures, LDL Cholesterol / HDL Cholesterol, LDL-Chol/HDL-Chol	1.7381e-1	8.8630e-10
Calculated Figures, LDL-1 Particle Number, LDL-1 Particle Number	-2.4799e-1	1.1853e-18

https://www.biobank.org.tw/

Data Analysis Workflow of TWB NMR Metabolomics Dataset



Note:

- IVDr data are absolute concentrations, already corrected using internal standards.
- Common row-wise normalization (e.g., total area) may not be required and may distort values.