Exploring metabolomic data from designed experiments using ANOVA Multiblock Orthogonal Partial Least Squares

Julien Boccard
School of Pharmaceutical Sciences
University of Geneva, University of Lausanne
Experimental design in metabolomics

Several factors can be evaluated simultaneously, e.g., dose, time, genotype.

Each factor has different levels, e.g., quantitative, qualitative, ordinal.

Full factorial designs allow the systematic evaluation of:
- main effects
- interactions between factors

Many experimental setups generate multivariate data, e.g., spectra, omics, multi-components response.
ANOVA and multivariate data analysis

- MANOVA is not able to handle underdetermined systems (k>n)
- PCA mixes the different sources of variation

How to account for the study design and covariances between variables?

- Existing strategies associate ANOVA decomposition of the experimental matrix with projection methods

\[ X = X_\mu + X_\alpha + X_\beta + X_{\alpha\beta} + X_{\text{Res}} \]

Existing strategies:
- ASCA
- ANOVA-PCA
- ANOVA-PLS
- ANOVA-TP
- AComDim
ANOVA Multiblock OPLS workflow

Experimental matrix (n x k)

ANOVA decomposition (n x k)

For each main effect and interaction term:

→ Extraction of levels barycentres from *pure effect submatrices* by Singular Value Decomposition

ASCA
ANOVA Multiblock OPLS workflow

Experimental matrix
(n x k)

ANOVA decomposition
(n x k)

For each main effect and interaction term:
→ Computation of experimental submatrices
Pure effect submatrices + Residuals

Hypothesis:
The structure of a significant effect will emerge from noise
ANOVA Multiblock OPLS workflow

Experimental matrix
(n x k)

ANOVA decomposition
(n x k)

Joint analysis of the submatrices

Prediction of level barycentres based on experimental submatrices → multiblock OPLS
AMOPLS model outputs

Supervised multiblock OPLS model
Joint decomposition based on predictive/orthogonal component(s)

Observations scores
→ Check for sample groupings

Variables loadings
→ Highlight relevant biomarkers

Balance between block saliences
→ Assign each latent structure to a factor

Assess statistical relevance by comparison with ANOVA residuals
→ Permutation tests (effect-to-residuals ratio)
Paraquat Neurotoxicity - Dataset

Dataset
✓ UHPLC-TOF/MS metabolic profiles of 3D aggregating rat brain cells
✓ 36 observations x 1’397 variables (m/z @ RetTime)

2 Factors design
(i) Maturation (2 levels): Immature / Mature
(ii) Paraquat (3 levels): Control / Paraquat 0.5 µM / Paraquat 1 µM
Paraquat Neurotoxicity - PCA

- The principal components are not easily interpretable.
- The factors under study affect the metabolic profiles.
- The effects of Maturation and Paraquat are mixed up.
- The principal components are not easily interpretable.
- The groups are not well-separated.
Paraquat Neurotoxicity - AMOPLS

4 terms ANOVA decomposition:

→ 6 components AMOPLS model
5 predictive (2 main effects + 1 interaction) + 1 orthogonal

Goodness of fit
R² 0.903

Permutation tests
✓ Global model (p<5%)
✓ Maturation (p<1%)
✓ Paraquat (p<1%)
✗ Interaction (p>5%)

Effect-to-residuals ratio
Paraquat Neurotoxicity - Maturation effect

$t_{p1}$: Scores related to Maturation

Maturation has a strong impact on metabolic profiles

Mature cells exhibit more elaborated phenotypes

Shift from tissue growth to neuronal cells differentiation

- Gangliosides (GA1, GM2)
- Arachidonic acid metabolism
- Glutamate dipeptides
- Precursors of serotonin
- Sphingomyelin

- Acylcarnitines
- Tyrosine dipeptides
- Neurotensin
- Hydroxyvitamin D
Paraquat Neurotoxicity - Treatment effect

Paraquat treatment induces a dose-related response

Intoxicated cells show altered metabolic profiles

Shift from neurotransmission to oxidative stress
Conclusions

AMOPLS combines ANOVA decomposition of the sources of variation, OPLS interpretability and multiblock modelling.

- Supervised analysis of ANOVA submatrices
- Joint analysis of the effects
- Easy interpretation
- Contribution of each ANOVA submatrix
- Objective evaluation of the effects
- Broad field of application (biomarker discovery, method development, etc.)
Acknowledgements

Prof. Serge Rudaz
Dr. Fabienne Jeanneret
Dr. David Tonoli

Prof. Florianne Tschudi-Monnet
Dr. Jenny Sandström von Tobel