# Data integration and network construction with muscle metabolome and meat quality data in pigs

Julia Welzenbach, Christine Große-Brinkhaus, Christiane Neuhoff,

Karl Schellander, Christian Looft, Ernst Tholen

Institute of Animal Science
Animal Breeding and Husbandry / Genetics group
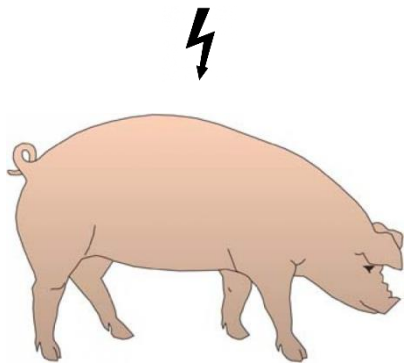University of Bonn, Germany

universität**bonn**

reduction in meat quality[1]
by impairment of…

e.g. stress conditions
before slaugthering

high drip loss

1) technological meat
characteristics

➢ reduced acceptance of
processing industries

2) sensory meat
characteristics

➢ reduced consumer
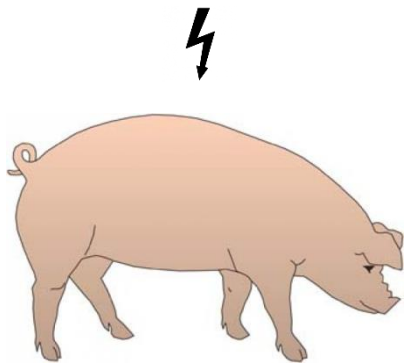acceptance

[1]Fischer et al. 2007

land-
wirtschaftliche
fakultät

universität**bonn**

e.g. stress conditions before slaugthering

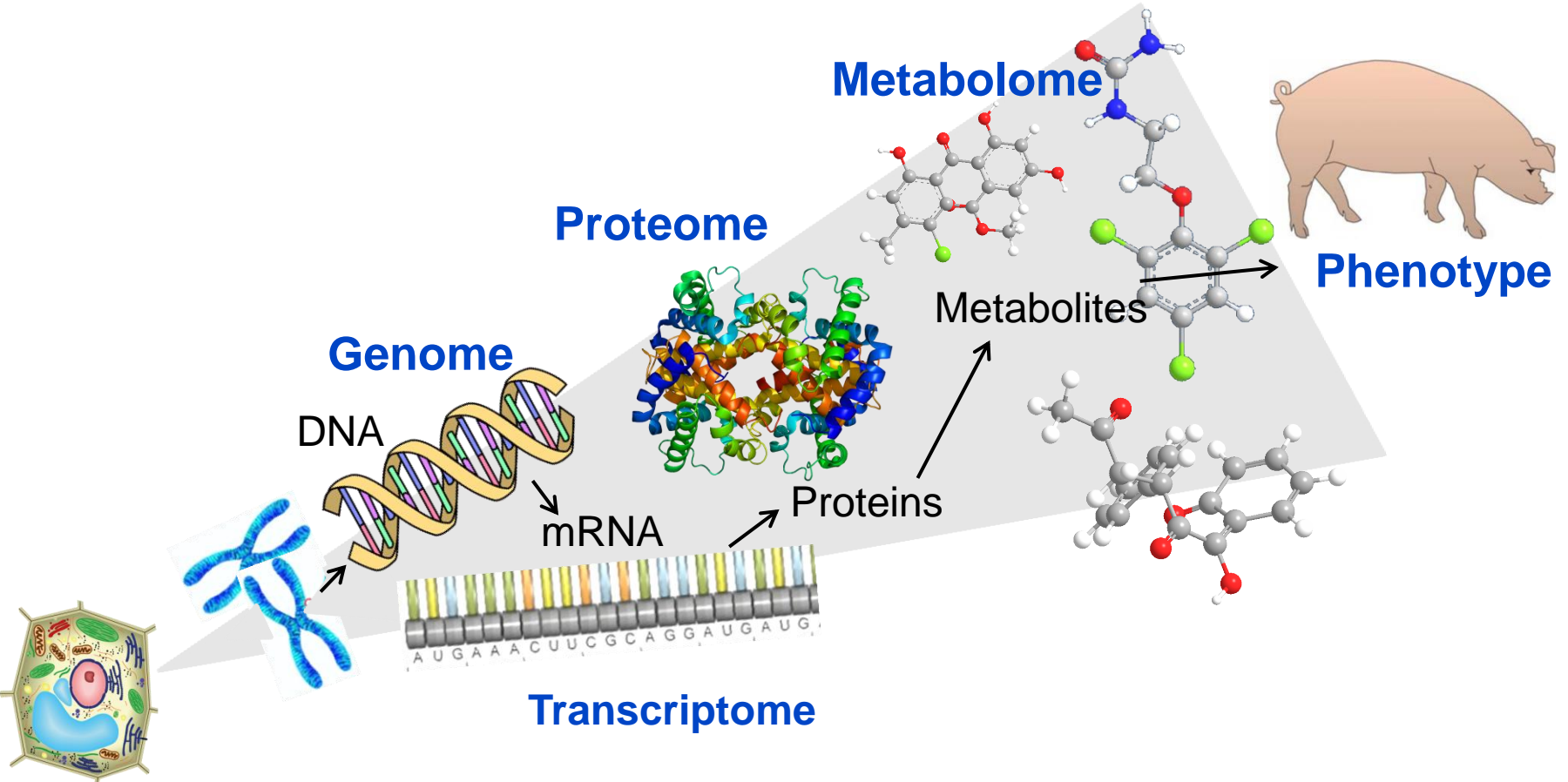high drip loss

reduction in meat quality[1] by impairment of…

1) technological meat characteristics

➢ reduced acceptance of processing industries

2) sensory meat characteristics

➢ reduced consumer acceptance

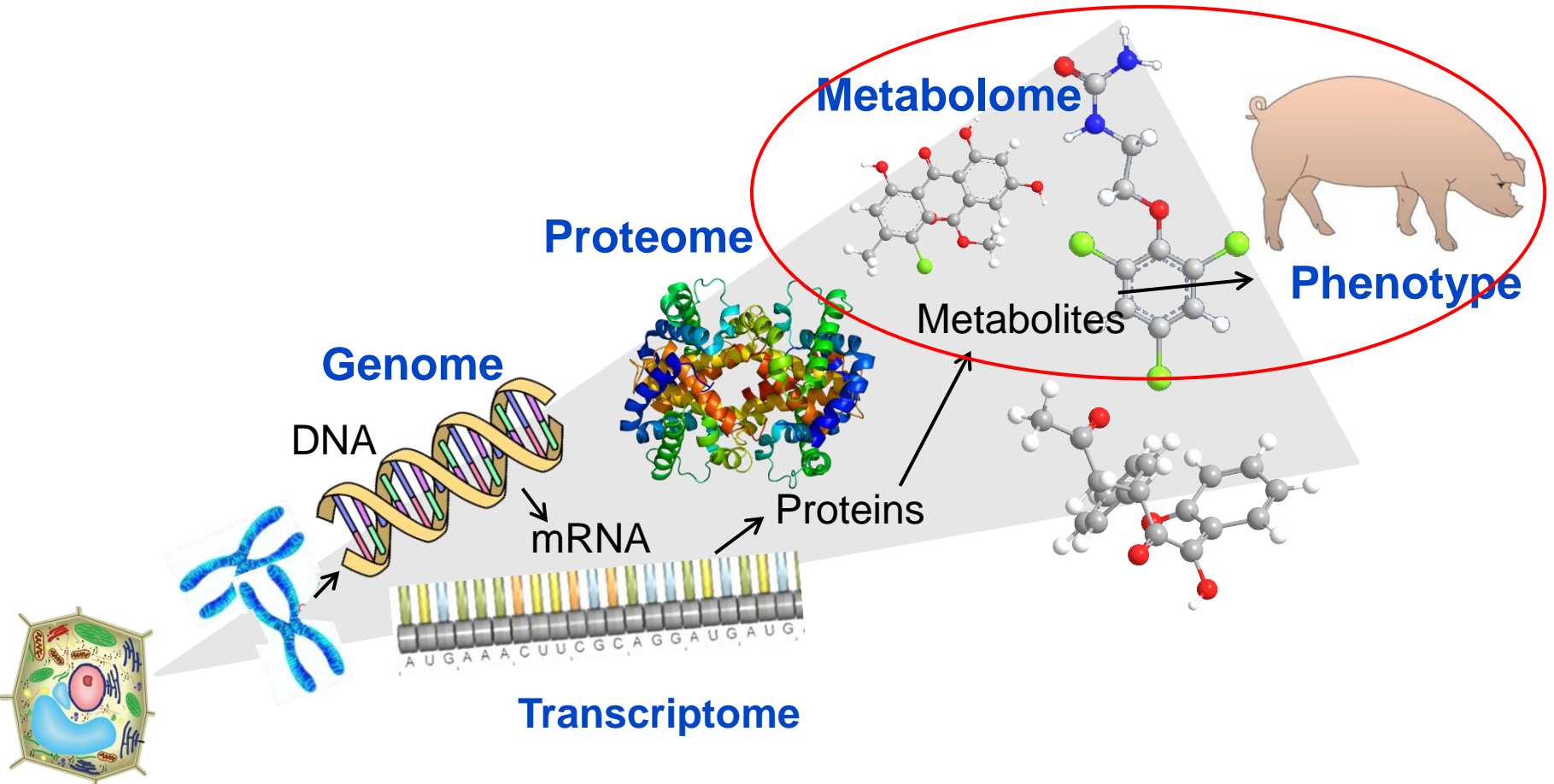➡ **How to select for low drip loss despite low heritabilty!?**

land-wirtschaftliche fakultät

# Introduction
# - the way from genome to phenotype



Metabolome

Proteome

Genome

Phenotype

DNA

Metabolites

mRNA

Proteins

Transcriptome

[1]Bino et al., 2004

# Introduction
## - the way from genome to phenotype



**Metabolome**

**Proteome**

**Genome**

**Phenotype**

DNA

Metabolites

Proteins

mRNA

Transcriptome

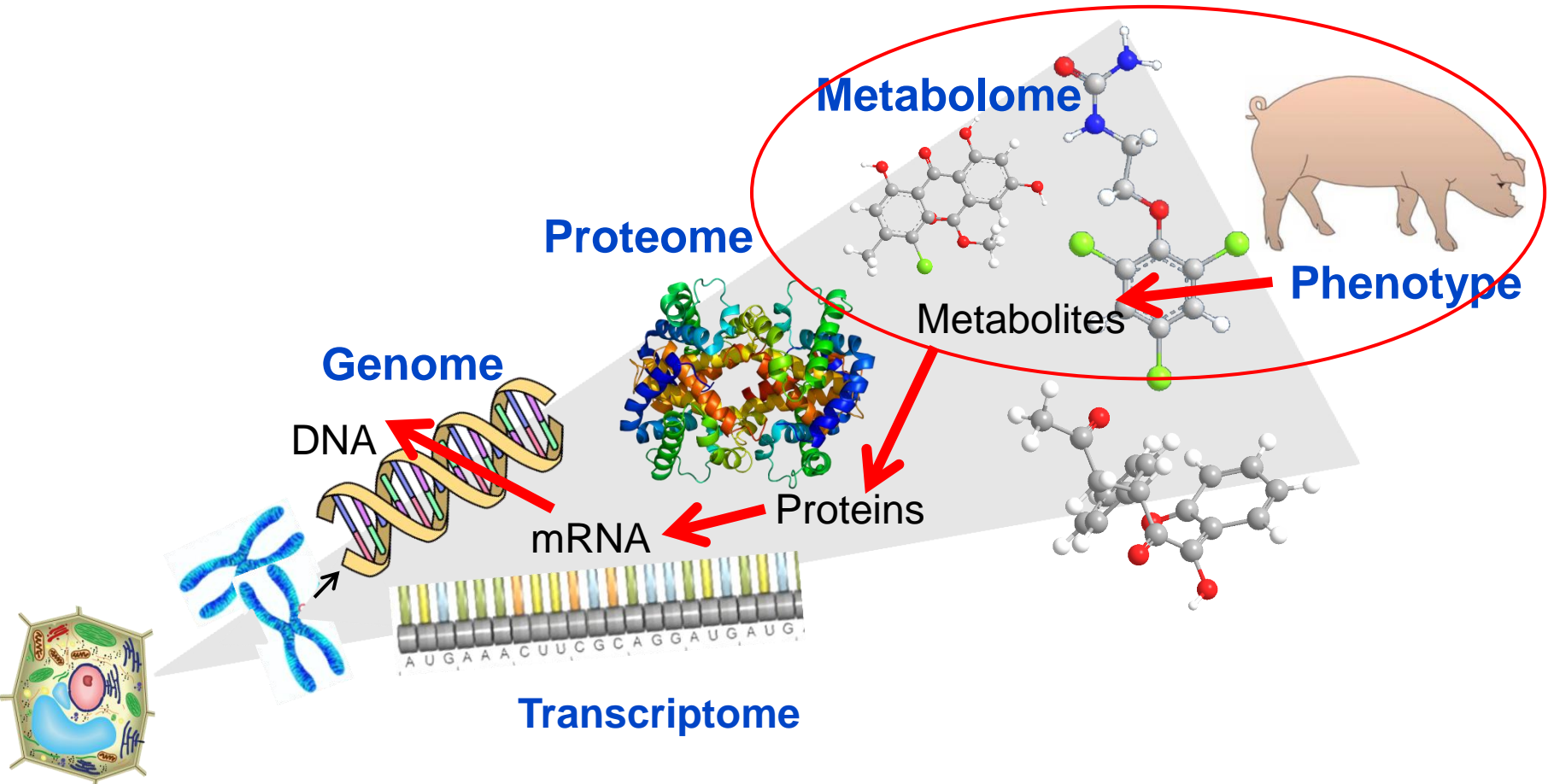→ Metabolites as more reliable phenotype[1]

[1]Bino et al., 2004

3

# Introduction
# - the way from genome to phenotype

Metabolome

Phenotype

Proteome

Metabolites

Genome

DNA

Proteins

mRNA

Transcriptome

→ Metabolites as more reliable phenotype[1]

→ Fill in the gap between phenotype and genome

[1]Bino et al., 2004

# Objective of the study

Investigation of the **relationship** between

**metabolite profiles** and **drip loss**

➢ get new insights in the complex biochemical processes of drip loss

➢ test different statistical approaches to reveal the most important metabolites

➢ detect metabolites as potential biomarkers

# Animals and drip loss phenotyping

- 97 $F_2$ pigs of Duroc × Pietrain resource population

- Meat samples of *M. longissimus dorsi*

- Record of drip loss 24h pm

  ➢ Bag-Methode of *Honikel* (1986)

  ➢ drip loss range: 0,4 – 5,3%

# Quantification and annotation of metabolites

o by gas – and liquid chromatography with mass spectrometry (GC-MS, LC-MS)

o „untargeted" metabolomics profiling

- Detection of the whole metabolome

➢ 1993 metabolites detected

o functional annotation

- Human Metabolome Database, Lipid Maps, METLIN
- 400 of 1993 metabolites annotated
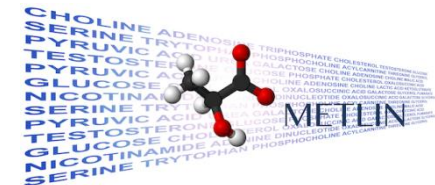
# Statistical analysis I

o Pre-correction of phenotypes and metabolite profiles

- generalized lineare model (GLM) with slaugther weight and season

o Statistical approaches to analyse the drip – metabolite associations

1. Correlation analysis

2. Principal component analysis (PCA)

3. Weighted network analysis (WNA)
   - using Ⓡ *WGCNA*[1]

4. Random forest regression[2] (RFR)
   - using Ⓡ *party*[3]

[1]Langfelder & Horvath 2008, [2]Breimann 2001, [3]Strobl et al. 2007

# Statistical analysis I

o Pre-correction of phenotypes and metabolite profiles

- generalized lineare model (GLM) with slaugther weight and season

o Statistical approaches to analyse the drip – metabolite associations

1. Correlation analysis
2. Principal component analysis (PCA)
3. Weighted network analysis (WNA)
   - using Ⓡ *WGCNA*[1]
4. Random forest regression[2] (RFR)
   - using Ⓡ *party*[3]

**handle the „Large p, small n"-problem (Overfitting)**

[1]Langfelder & Horvath 2008, [2]Breimann 2001, [3]Strobl et al. 2007

# Statistical analysis II

## 2. PCA

- Condenses the metabolite profiles into representative, uncorrelated principle components (PCs)

- Metabolites are quantified by their corresponding loadings
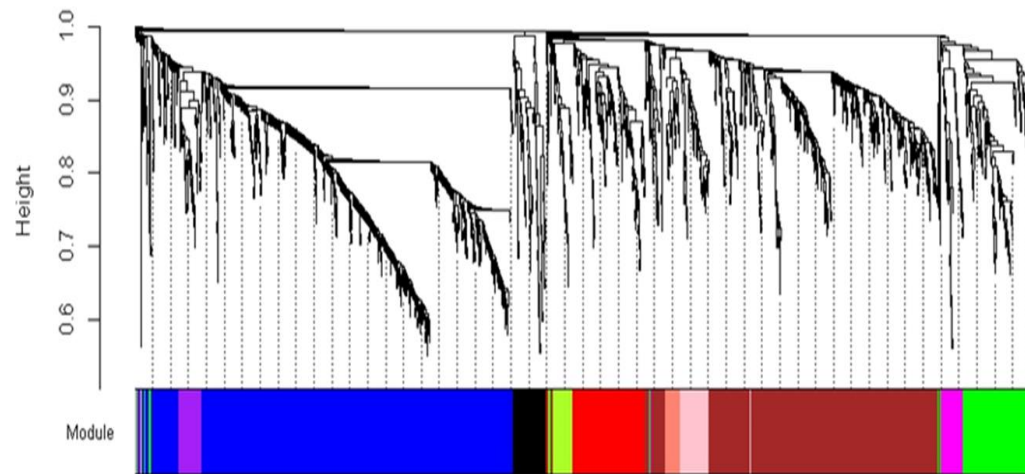
## 3. Weighted network analysis (WNA)

- Generates biological interpretable modules based on a hierarchical clustering dendrogram

- Metabolites are characterised by

➢ **module membership (MM)**
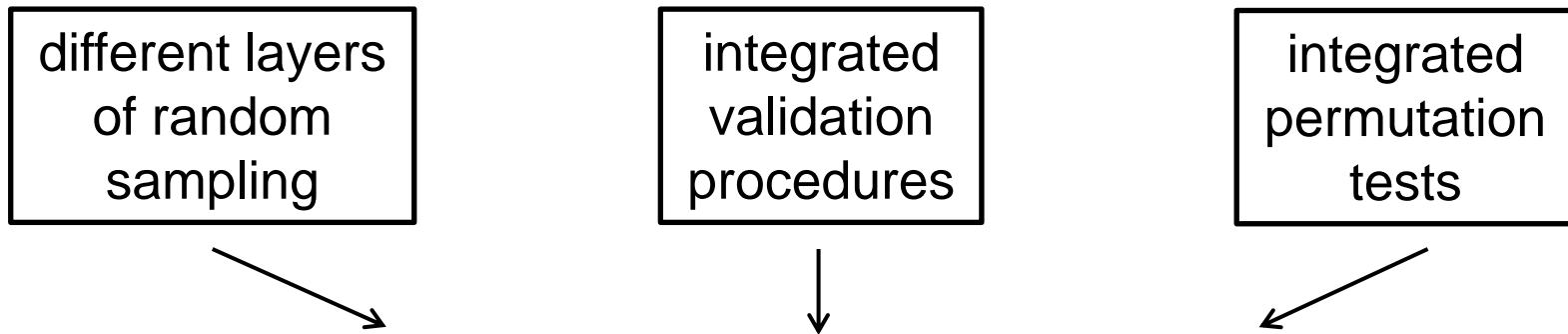  - Connectivity of the metabolites within a module

➢ **metabolite significance (MS)**
  - Based on the trait ↔ metabolite correlation

# Statistical analysis III

## 4. Random Forest Regression (RFR)

- supervised learning tool using tree-based methods

- key characteristics of RFR:

| | | |
|---|---|---|
| different layers of random sampling | integrated validation procedures | integrated permutation tests |

➢ RFR is able to handle datasets with complex interaction structures and highly correlated variables

- RFR calculates *variable importance (VI)* values, that are based on prediction accuracy[1]

[1]Strobl et al. 2007

# Results
# – Identification of biomarkers

## 1. Correlation analysis

- 71 (5) metabolites positive (negative) correlated (p ≤ 0,05)

- Range: 0.24 to 0.28 respectively -0.24 to -0.23

## 2. PCA

- First 3 PCs specify 46.9 % of metabolite expression variance

- Loadings are very weak (range: -0.1 to 0.1) → not significant[1]

➤ PCA not used for biomarker identification

➤ But: PCA was used to reduce the data set for RFR

[1]DiLeo et al., 2011

## 3. Weighted Network analysis (WNA)

- Clusters the metabolites into 10 modules
  - two modules significantly associated with drip loss

| trait | module | cor. | p-value | number metabolites |
|-------|--------|------|---------|--------------------|
| drip loss | ‚purple' | + 0.21 | p ≤ 0.04 | 52 |
| | ‚green' | + 0.21 | p ≤ 0.04 | 49 |

## 4. Random Forest Regression (RFR)

- 293 metabolites with significant (p ≤ 0,05) *variable importance* (VI) for drip loss
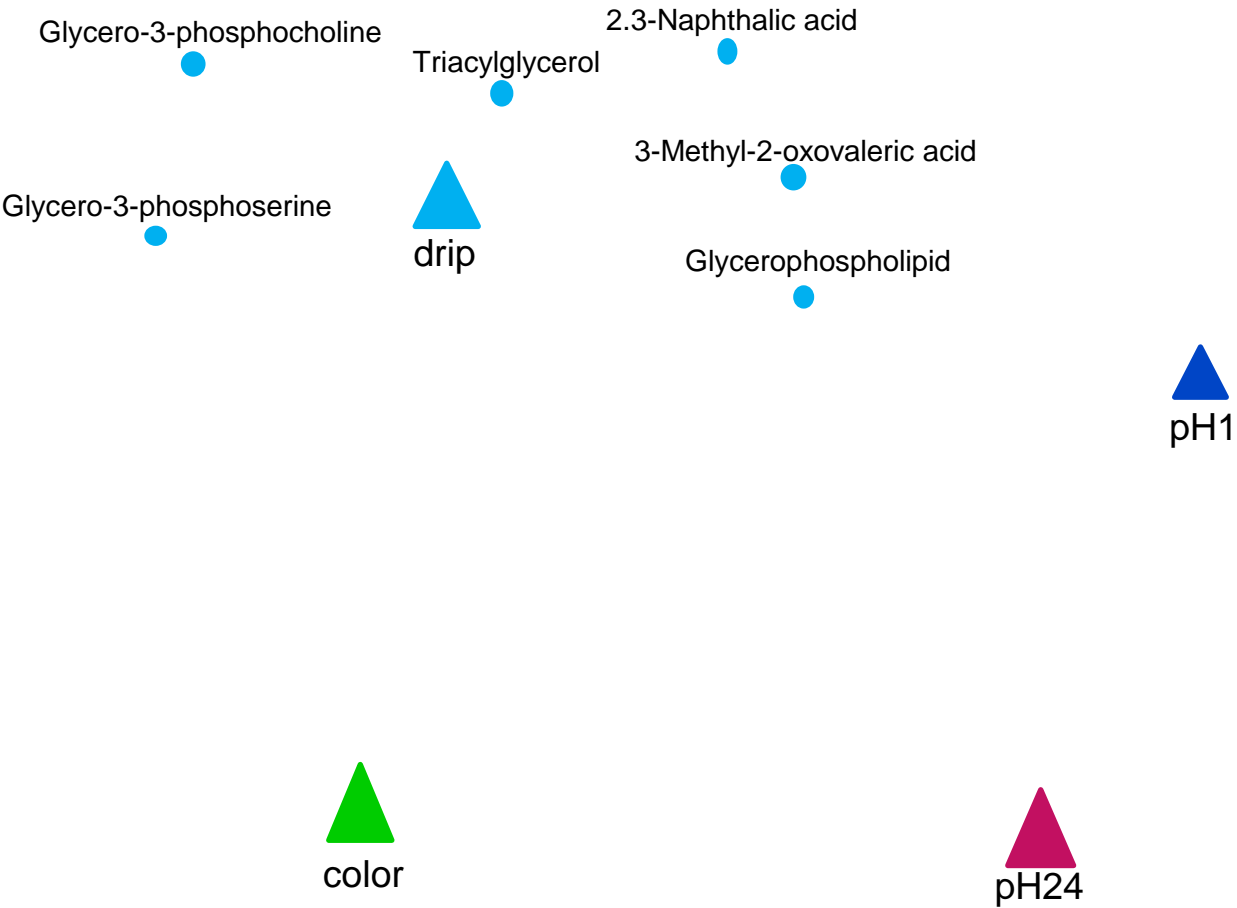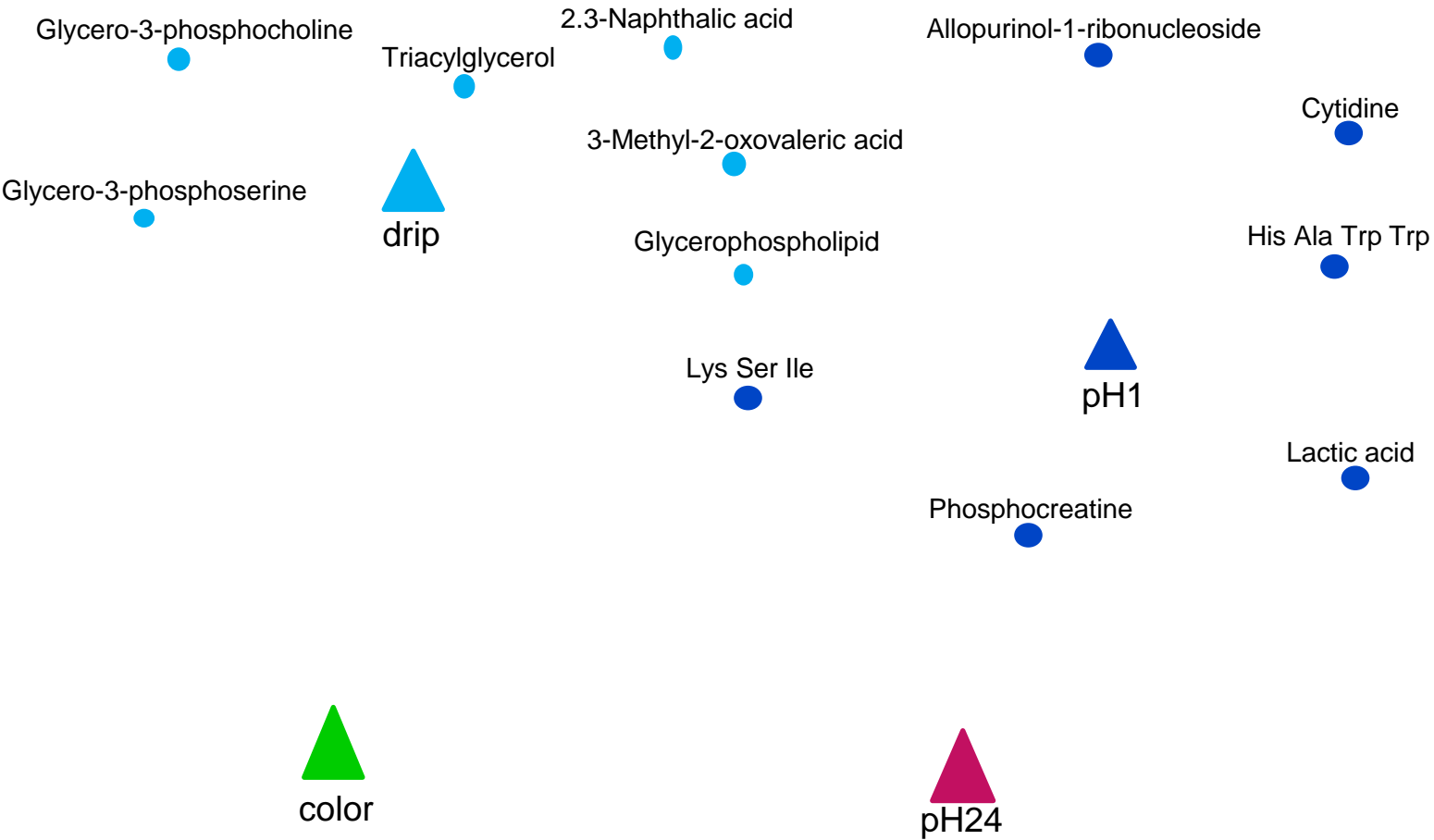
# Results – network construction

# Results – network construction
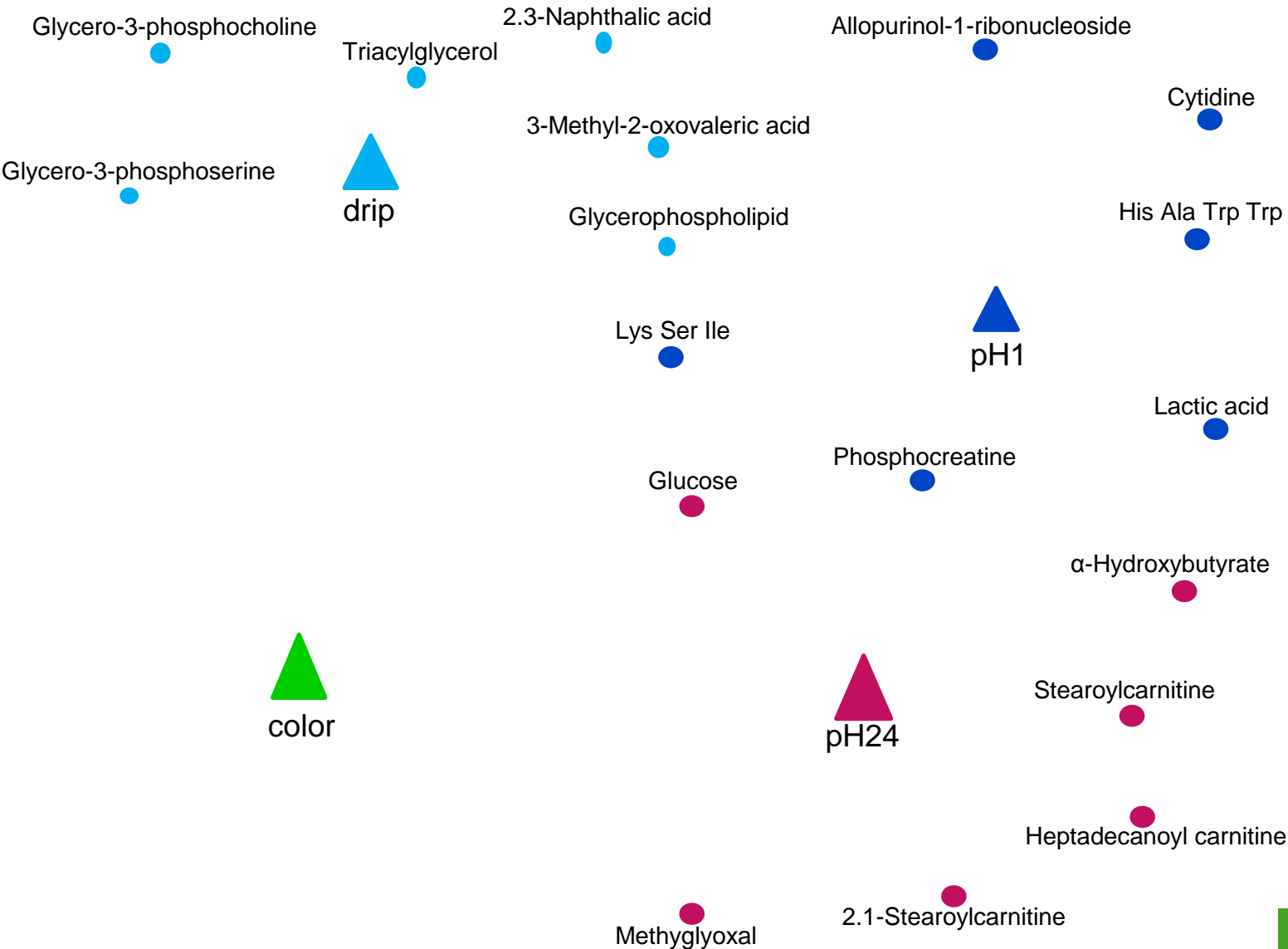


Glycero-3-phosphocholine

2.3-Naphthalic acid

Allopurinol-1-ribonucleoside

Triacylglycerol

Cytidine

**Dotted lines:**
*connectivity* between metabolites

3-Methyl-2-oxovaleric acid

Glycero-3-phosphoserine

drip

Glycerophospholipid

His Ala Trp Trp

Nicotinamide adenine dinucleotide

Lys Ser Ile

pH1

Lactic acid

Phosphocreatine

Glucose

Glucose-6-phosphate

Fructose-6-phosphate

α-Hydroxybutyrate

color

Stearoylcarnitine

pH24

Phosphoglycolic acid

Inosine-5-monophosphate

Heptadecanoyl carnitine

Octulose-1.8-bisphosphate

Methyglyoxal

2.1-Stearoylcarnitine

# Results – network construction

Glycero-3-phosphocholine

Triacylglycerol

2.3-Naphthalic acid

Allopurinol-1-ribonucleoside

Cytidine

3-Methyl-2-oxovaleric acid

Glycero-3-phosphoserine

drip

Glycerophospholipid

His Ala Trp Trp

pH1

Lys Ser Ile

Nicotinamide adenine dinucleotide

Lactic acid

Phosphocreatine

Glucose

Fructose-6-phosphate

Glucose-6-phosphate

α-Hydroxybutyrate

color

pH24

Stearoylcarnitine

Phosphoglycolic acid

Inosine-5-monophosphate

Heptadecanoyl carnitine

Octulose-1.8-bisphosphate

Methyglyoxal

2.1-Stearoylcarnitine

**Dotted lines:**
*connectivity* between metabolites

Arrows:
*directed connections* metabolite → trait

land-wirtschaftliche fakultät

# Results – network construction



**Dotted lines:** *connectivity* between metabolites

Arrows: *directed connections* metabolite → trait

# Results – network construction
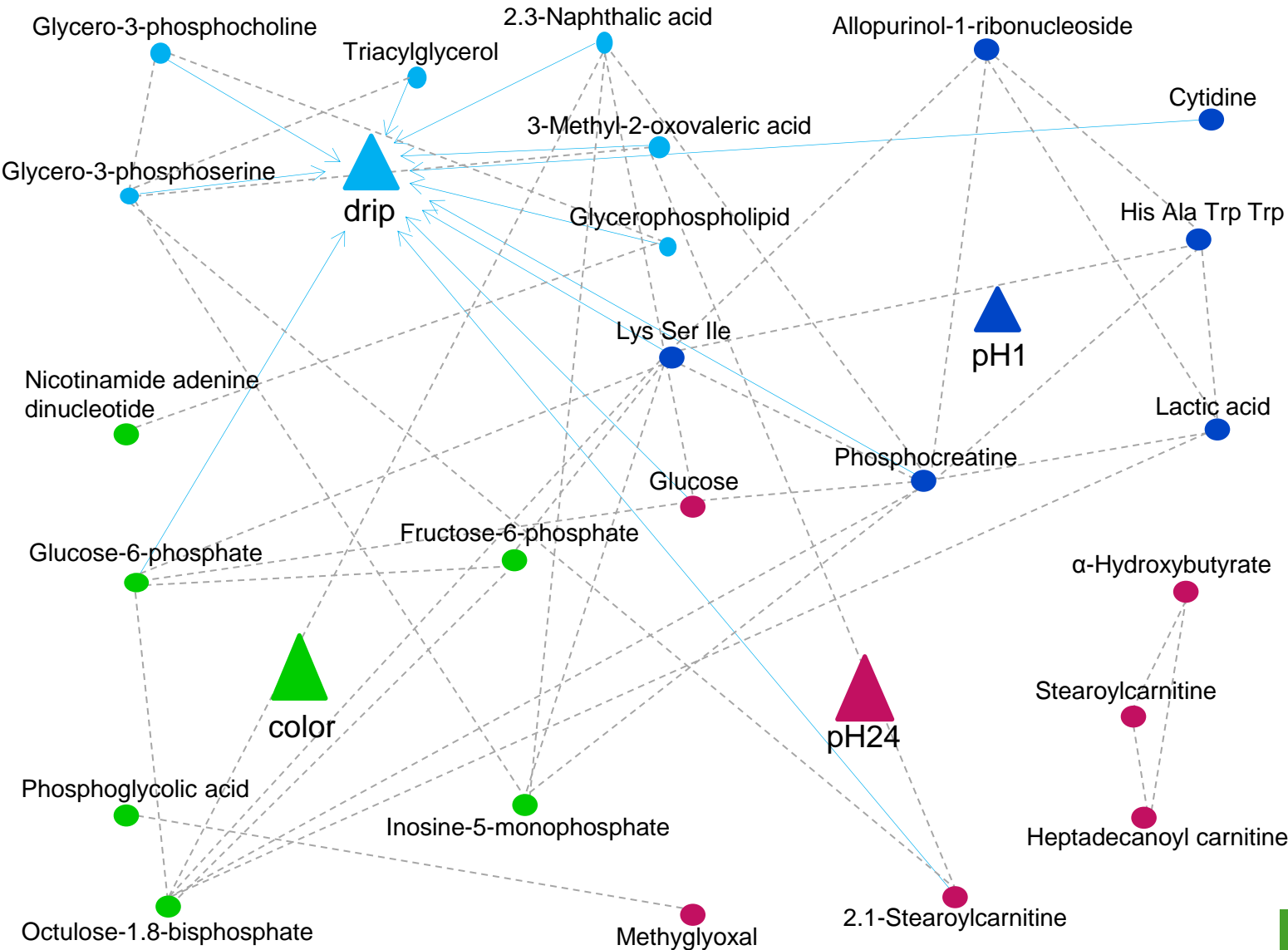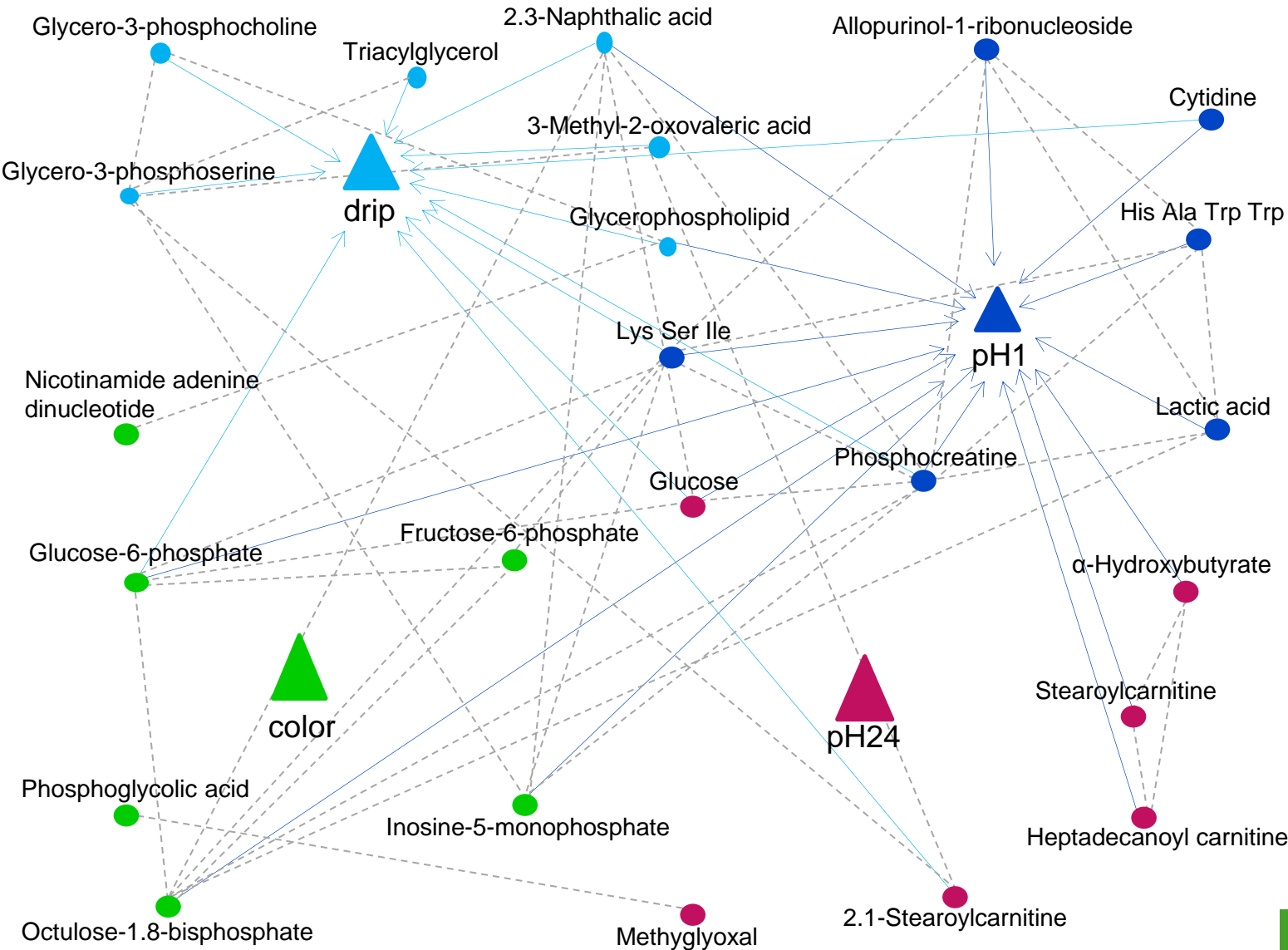


**Dotted lines:**
*connectivity* between metabolites

Arrows:
*directed connections* metabolite → trait

12

# Results – network construction

**Dotted lines:** *connectivity* between metabolites

Arrows: *directed connections* metabolite → trait

Dotted lines: *connectivity* between metabolites

Arrows: *directed connections* metabolite → trait

# Results – network construction



Dotted lines: *connectivity* between metabolites
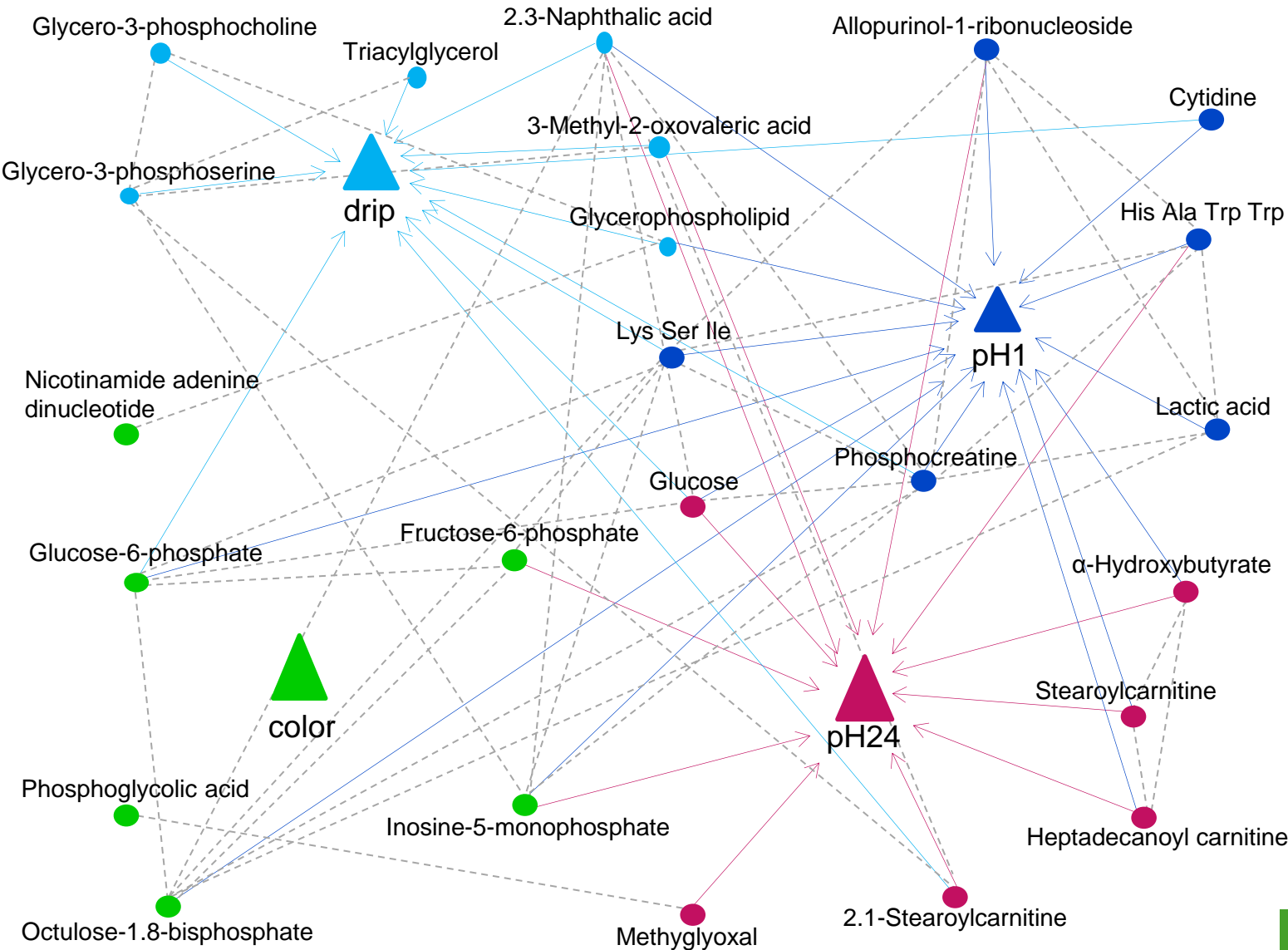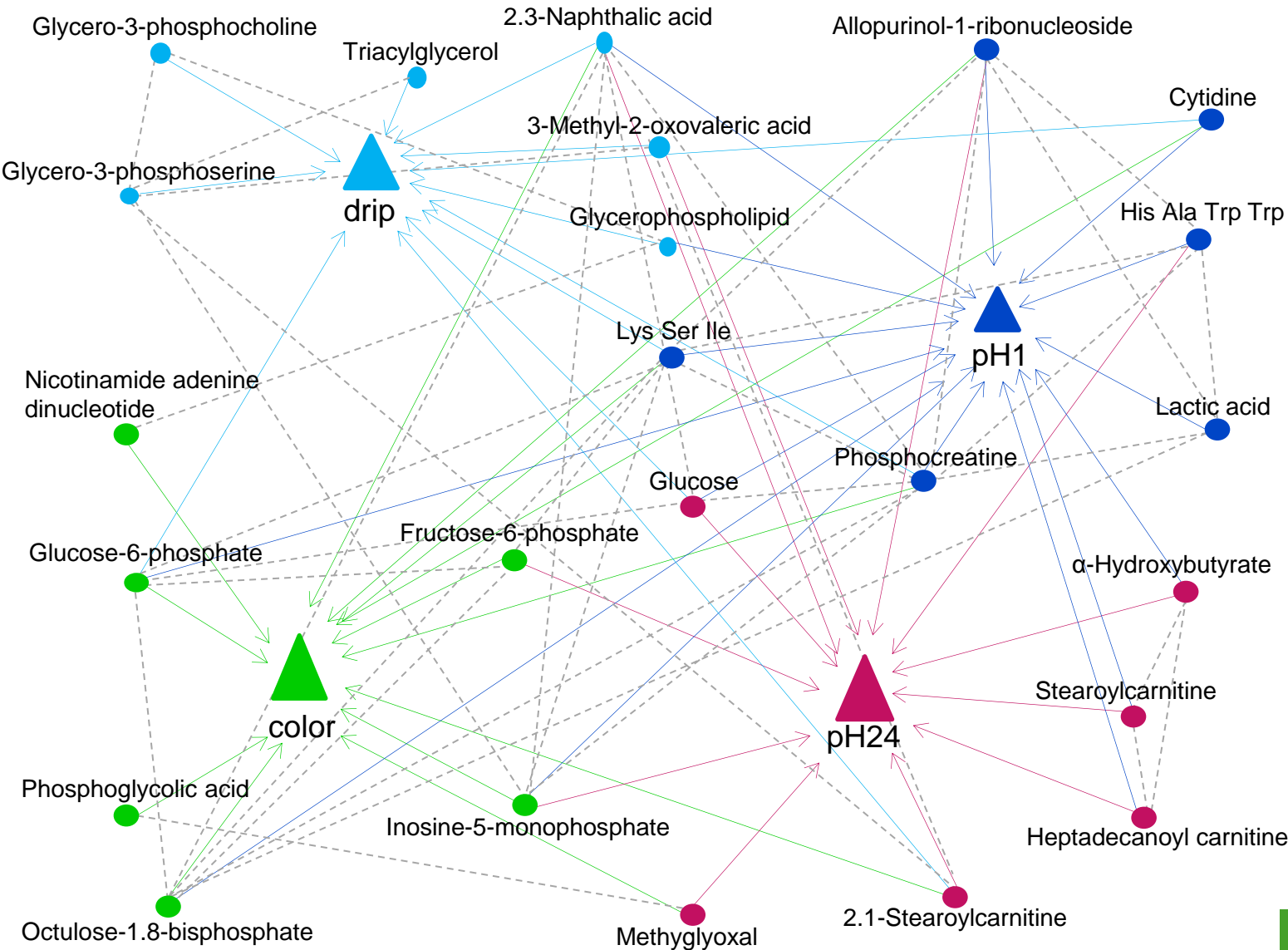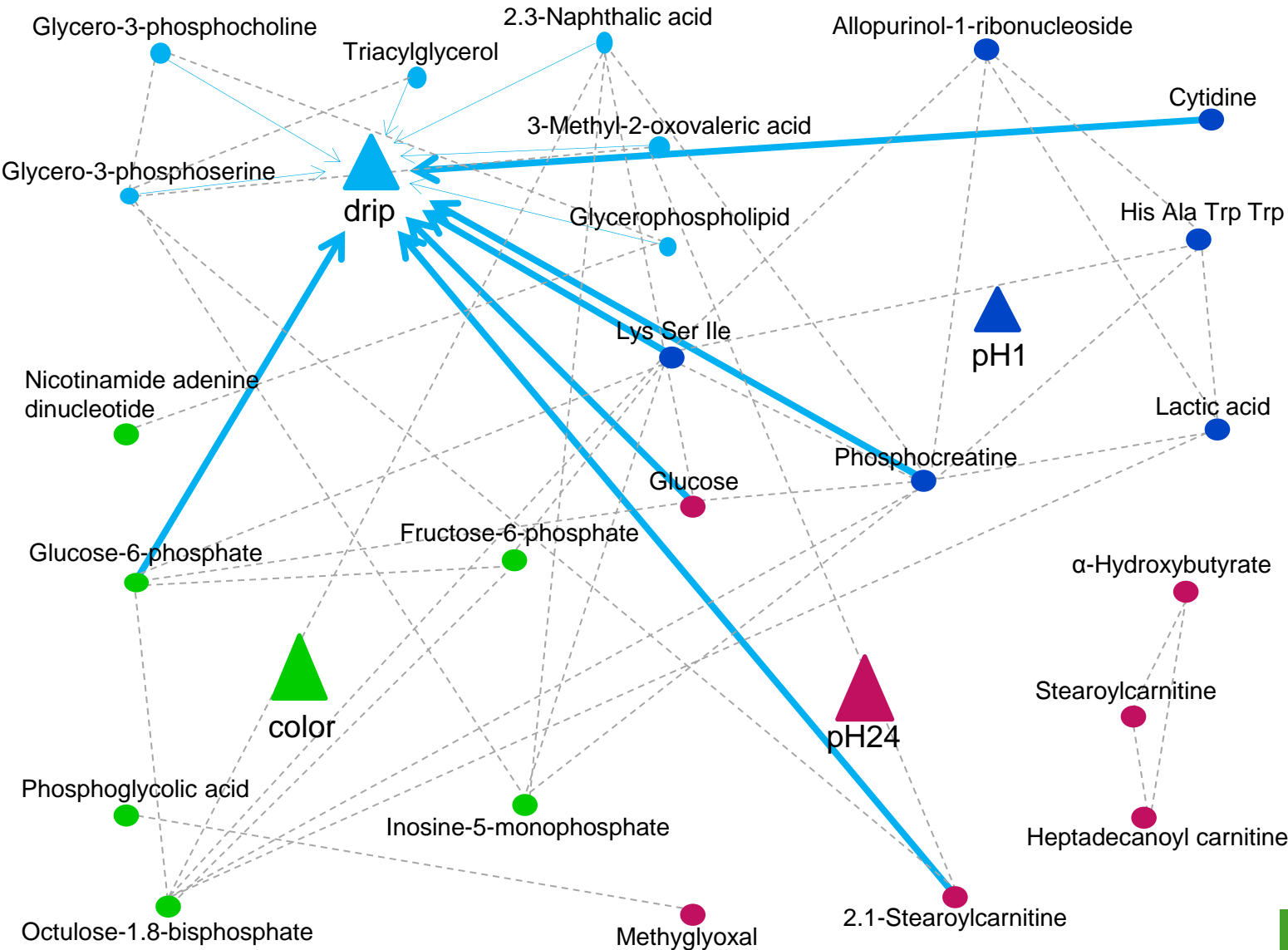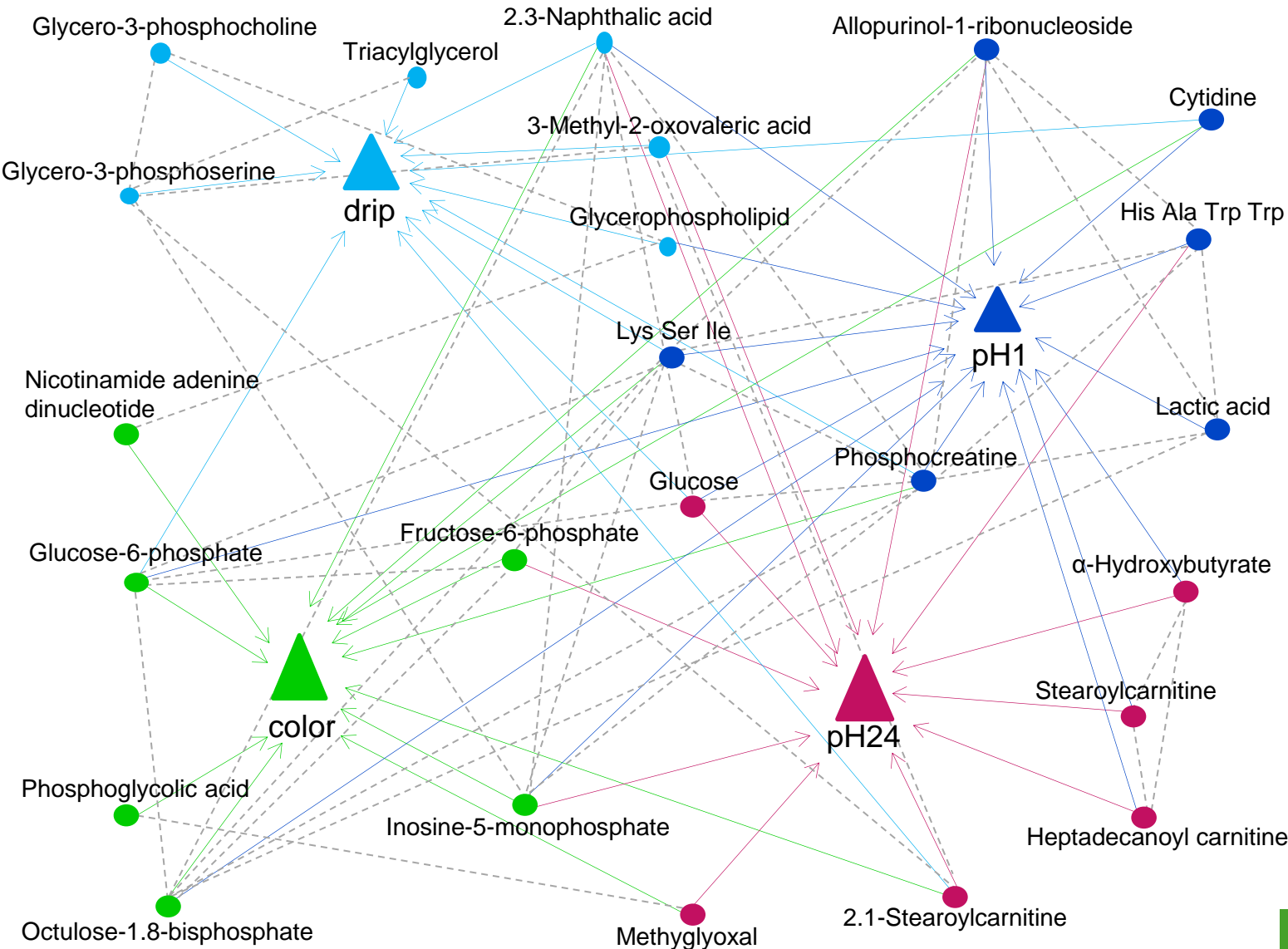
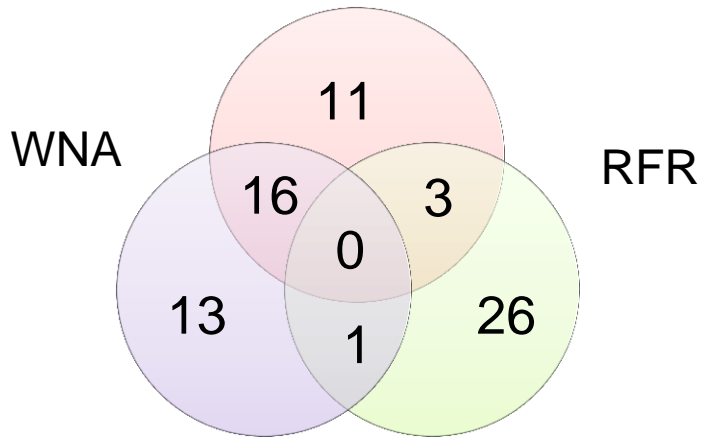Arrows: *directed connections* metabolite → trait

Dotted lines: *connectivity* between metabolites

Arrows: *directed connections* metabolite → trait

# Results - prediction accuracy of selected metabolites

- overlap in "Top 30" significant metabolites for drip loss

correlation analysis



WNA

11

16    3

0

13    1    26

RFR

- 20 metabolites identified by more than one method

- multiple $R^2$: 32.73 %

- stepwise Regression of the 20 metabolites

- resulted in 5 important metabolites

| biomarker for drip loss | regression coefficient |
|---|---|
| Glycerophosphocholine | 45.82* |
| C24:1 Sphingomyelin | -50.53* |
| Ubiquinon (Prenol Lipid) | 532.20* |
| C26:1 Sphingomyelin | 1870.00** |
| not annotated | 398.50*** |

- multiple $R^2$: 26.61 %

land-wirtschaftliche fakultät

# Conclusion and perspective

➢ Correlation analysis, WNA and RFR are suitable in identification of predictors

➢ Glycerosphospho- und Sphingolipids are the most promising biomarkers

➢ Selected set of metabolites has moderate prediction accuracy and also an effect on other meat quality traits

➢ Requirements of the development of reliable metabolite biomarkers
   ↳ enhanced abilities of metabolite quantification and annotation

➢ Future perspective: usage of combined omics-profiles as more exact phenotype and in GWAS of identifiy ‚real' powerful SNPs

Genome ⟷ Transkriptome ⟷ Proteome ⟷ Metabolome → Phenotype

14
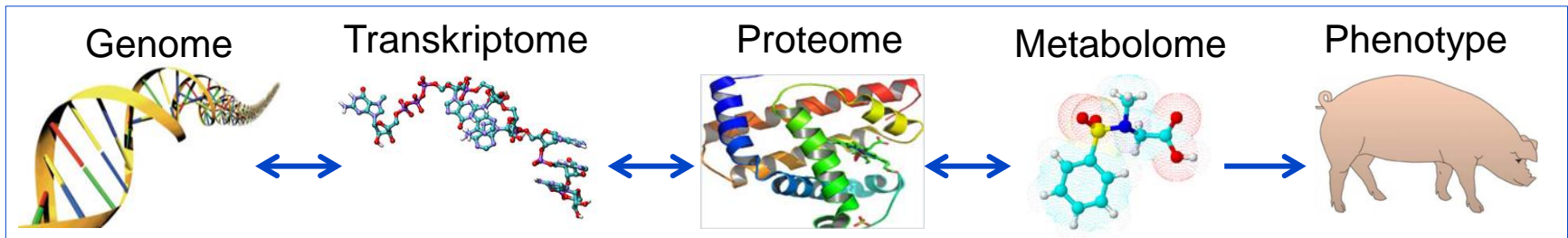
# Thank you for your attention

# Literature

- Bertram H. C.,Schafer A., Rosenvold K., Andersen H. J. (2004) Physical changes of significance for early post mortem water distribution in porcine M. longissimus. Meat Sci. **66**: 915-924.
- Bino R. J., Hall R. D., Fiehn O., Kopka J., Saito K. (2004) Potential of metabolomics as a functional genomics tool. Trends in Plant Science. **9**: 418–425.
- Breiman L. (2001) Random forests. Mach. Learn. **1**: 5–32.
- Fischer K., Lindner J. P., Freudenreich P., Spindler M., Schüssler G. (2007) Schnellanalytische Bestimmung des Wasserbindungsvermögens und anderer Merkmale des PSE-Status von Schweinefleisch mit Hilfe der VIS/NIR-Spektroskopie. Mitteilungsblatt der Fleischforschung Kulmbach. **46**: 209-215.
- Honikel K. O., Kim C. J., Hamm R., Roncales P (1986) Sarcomere shortening of pre-rigor muscles and ist influences on drip loss. Meat Sci. **16**: 267-282.
- Lambert I. H., Nielsen J. H., Andersen H. J., Ortenblad, N. (2001) Cellular model for induction of drip loss in meat. J. Agric. Food Chem. **49**: 4876-4883.
- Ortenblad N., Young J. F., Oksbjerg N., Nielsen J. H., Lambert I. H. (2003) Reactive oxygen species are important mediators of taurine release from skeletal muscle cells. Am. J. Physiol. Cell. Physiol. **284**: 1362-1373.
- Piironen V., Lampi A. M. (2004) Occurrence and levels of phytosterols in foods; in Dutta PC (ed): Phytosterols as Functional Food Components and Nutraceuticals. New York, Marcel Dekker,1–32.
- Zhang B. und Horvath S. (2005) A general framework for weighted gene coexpression network analysis. Statistical Applications in Genetics and Molecular Biology Vol. 4: Iss. 1, Article 17.

# Statistical analysis

## Random Forest Regression (RFR)

- supervised learning tool using tree-based methods with integrated permutation tests

- RFR procedure:
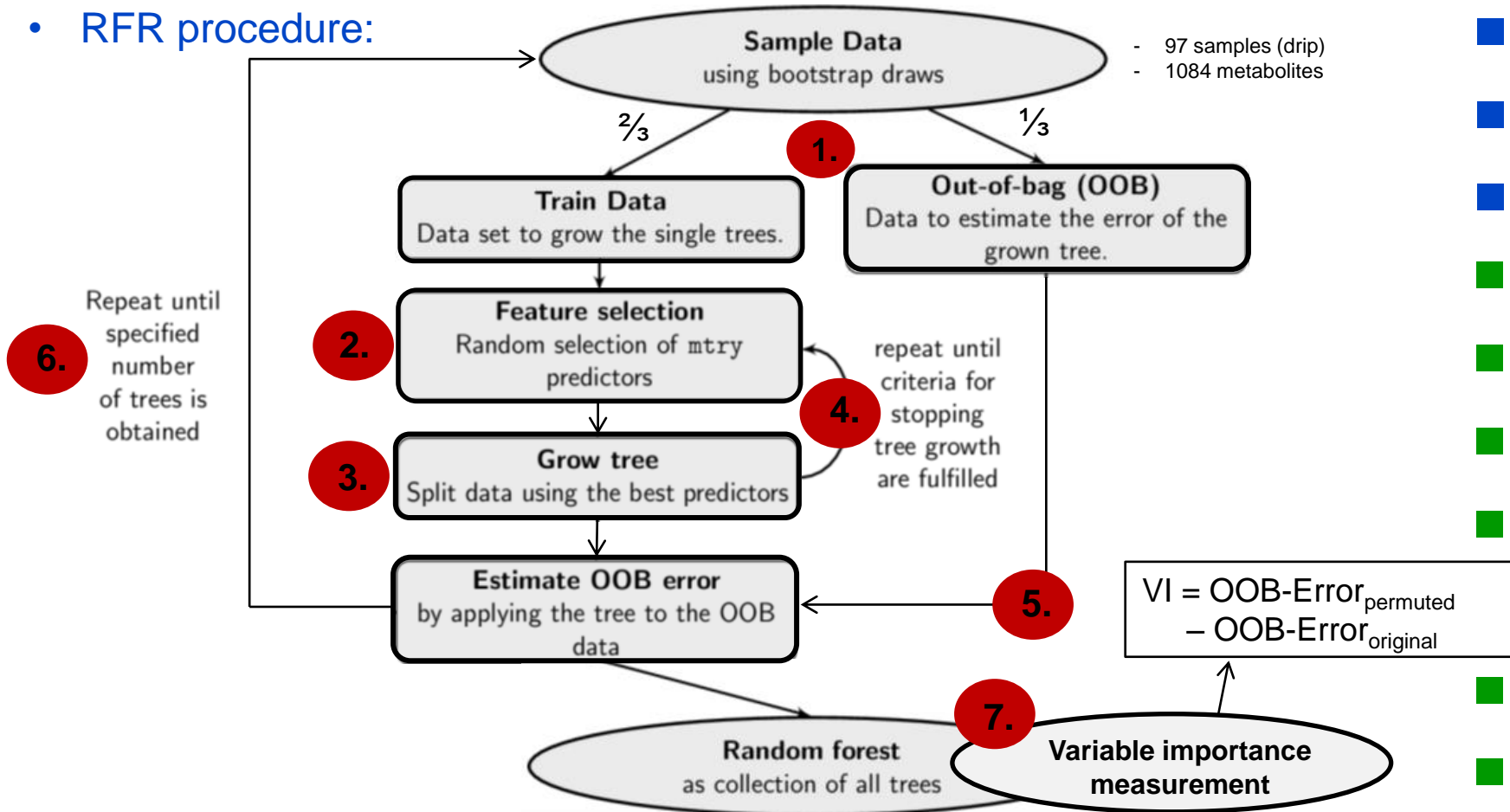
# Results - important biomarkers

- Ranking of metabolites in „top 30" of correlation analysis, WNA and RFR

| drip loss | Cor | MS | VI | pH1 | Cor | MS | VI |
|---|---|---|---|---|---|---|---|
| 2.3-Naphthalic acid | 23. | × | 10. | His Ala Trp Trp | 5. | 4. | 2. |
| Glycero-3-phosphocholine | 8. | × | 7. | Cytidine | 25. | 8. | 12. |
| Glycero-3-phosphoserine | × | 28. | 23. | Allopurinol-1-ribonucleoside | × | 9. | 25. |
| Glycerophosholipid | 22. | 14. | × | Lactic acid | 24. | × | 10. |
| Triacylglycerol | 19. | 12. | × | Lys Ser Ile | 19. | × | 6. |
| 3-Methyl-2-oxovaleric acid | 21. | 13. | × | Phosphocreatine | 26. | × | 21. |

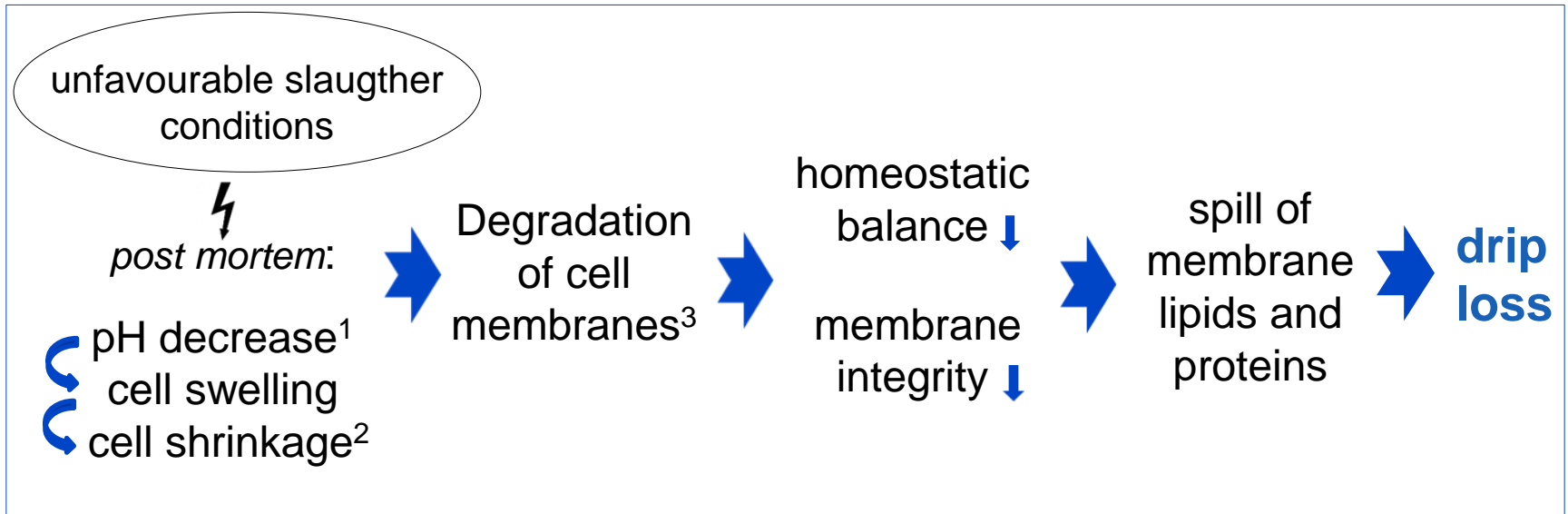| pH24 | Cor | MS | VI | color | Cor | MS | VI |
|---|---|---|---|---|---|---|---|
| α-Hydroxybutyrate | 1. | 1. | × | Octulose-1.8-bisphosphate | 7. | 1. | × |
| Heptadecanoyl carnitine | 2. | 2. | × | Fructose-6-phosphate | 27. | 9. | × |
| Stearoylcarnitine | 3. | 4. | × | Glucose-6-phosphate | 23. | 7. | × |
| Gle-cholesterol | × | × | 2. | Inosine-5-monophosphate | 28. | 10. | × |
| Methyglyoxal | × | × | 9. | Phosphoglycolic acid | 11. | 12. | × |
| Glucose | × | × | 11. | Nicotinamide adenine dinucleotide | 4. | × | 2. |

# Results – potential biomarkers
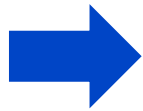
consistent results of statistical approaches:

Glycerosphospholipids (GPLs), Sphingolipids (SLs)



membrane lipid

unfavourable slaugther conditions

*post mortem*:

pH decrease[1]
cell swelling
cell shrinkage[2]

→ Degradation of cell membranes[3] → homeostatic balance ↓ / membrane integrity ↓ → spill of membrane lipids and proteins → **drip loss**
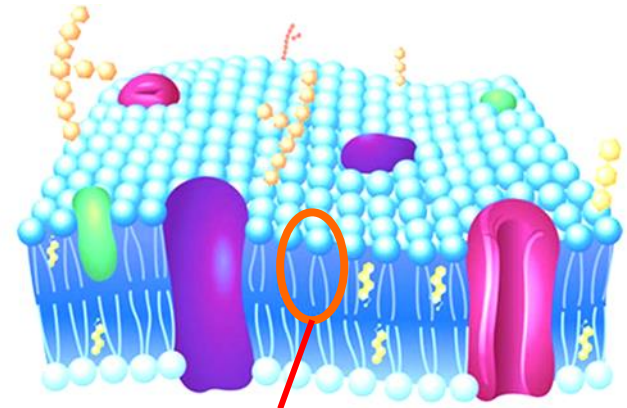
[1]Ortenblad et al., 2003;  [2]Bertram et al., 2004;  [3]Lambert et al., 2001