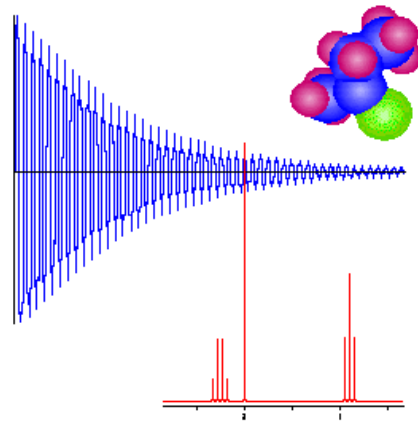


Detection of autosomal dominant polycystic kidney disease using NMR spectroscopic fingerprints of urine



Wolfram Gronwald

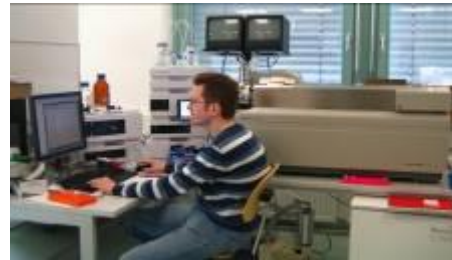
Munich Biomarker Conference 2011,
Munich

Institute of Functional Genomics



GC-APCI-TOF-MS

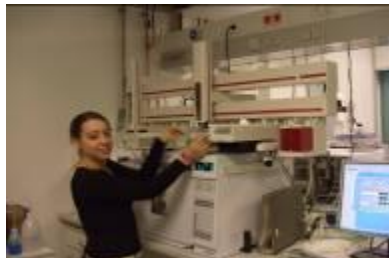
- Metabolic fingerprinting
- Metabolite identification



HPLC-MS/MS (4000 Q TRAP)

- Metabolite of central carbon metabolism by IP-LC-MS/MS
- Glucose flux analysis by IP-LC-MS/MS
- Intermediates of methionine and polyamine pathways
- Amino acids by iTRAQ-LC-MS/MS (45AAs)
- Analysis of tryptophan metabolites

Metabolomics at IFG



GC-MS with PrepStation

- Amino acids (30 AAs)
- Metabolic fingerprinting
- Small organic acids
- Amino acid enantiomers



GCxGC-TOF-MS

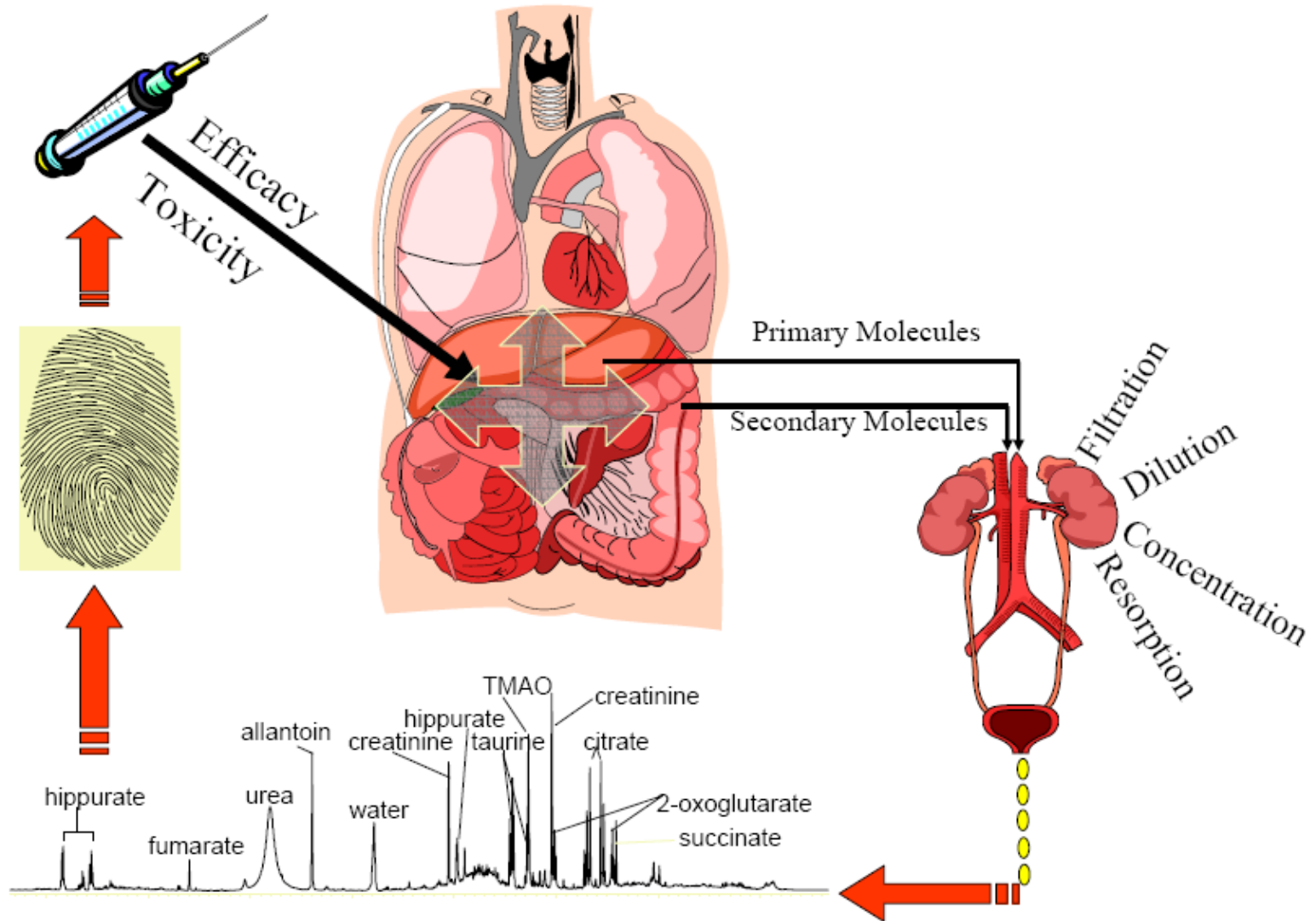
- Metabolic fingerprinting
- Small organic acids



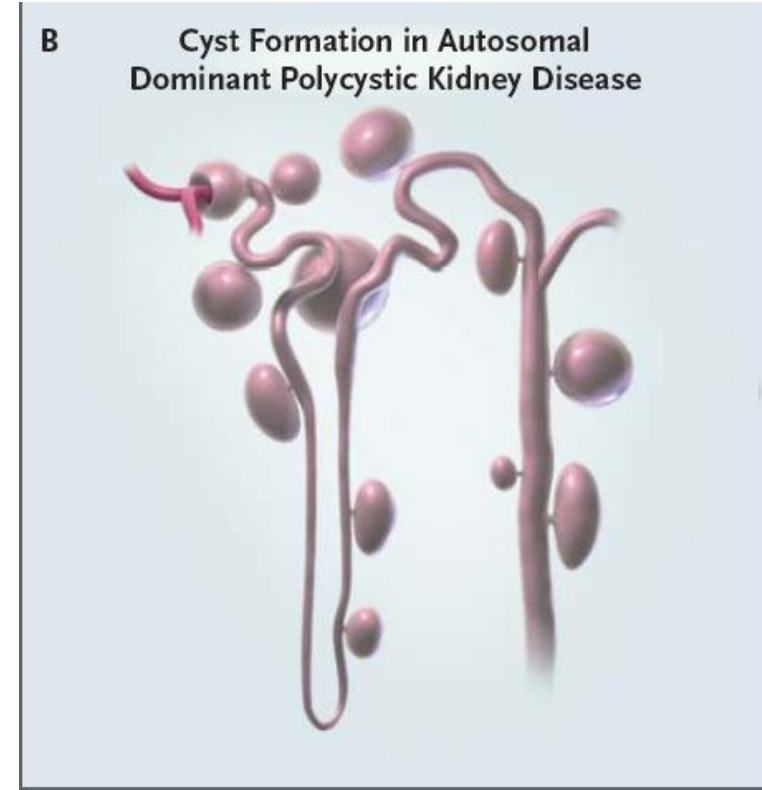
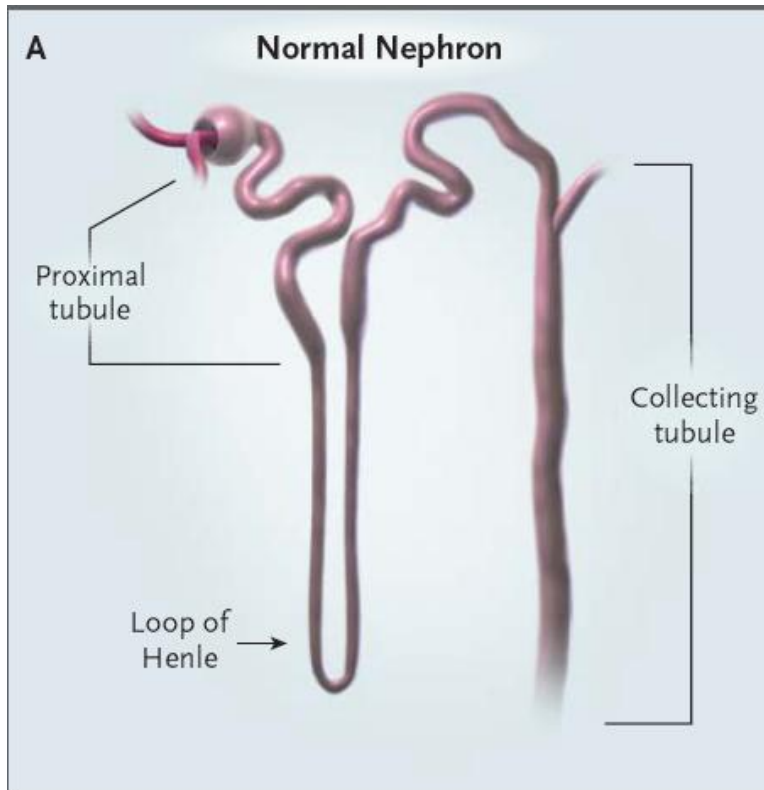
NMR

- Metabolic fingerprinting
- Metabolic profiling

The basic steps



ADPKD



Adapted from P. Wilson *N.Engl.J.Med*,**350**,151-164

- **ADPKD**

- Autosomal polycystic kidney disease (ADPKD)
- genetically inherited (1:400 - 1:1000)

The disease

- *Progression*
 - Polycystin 1 important for regulation of adhesion, migration of ureteric bud during renal development
 - Mutations disturb this process -> formation of cysts
 - Develops slowly, normally no symptoms until fourth decade of life and often diagnosed a lot later
 - Usually detected at a progressed state
 - At the end leads to renal failure -> organ transplant required

Motivation

- *Early detection*
 - Is it possible to discriminate between healthy and diseased patients at an relatively early state based on their urinary metabolic fingerprint?
 - Is it possible to discriminate autosomal polycystic kidney disease from other kidney complications?
- *Metabolites*
 - What are the relevant metabolites that allow discrimination?
 - Can we learn more about the biochemical paths involved in the disease?

The patients

- *Analysis of human urine samples by NMR*
 - 54 patients suffering from kidney cysts (age 36 - 53, mixed male/female)
 - 46 apparently healthy volunteers (age 20 - 50, mixed male/female)
 - 4 groups of patients suffering from other diseases

First steps in classification

- **Pretreatment of spectra**

- Buckets

- e.g. 0.01 ppm wide -> 700 buckets (features) correction of slight variations in signal positions across spectra

- Normalize

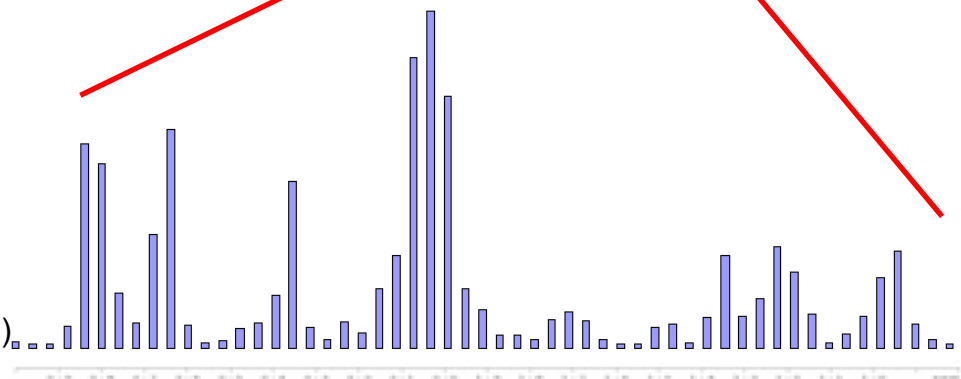
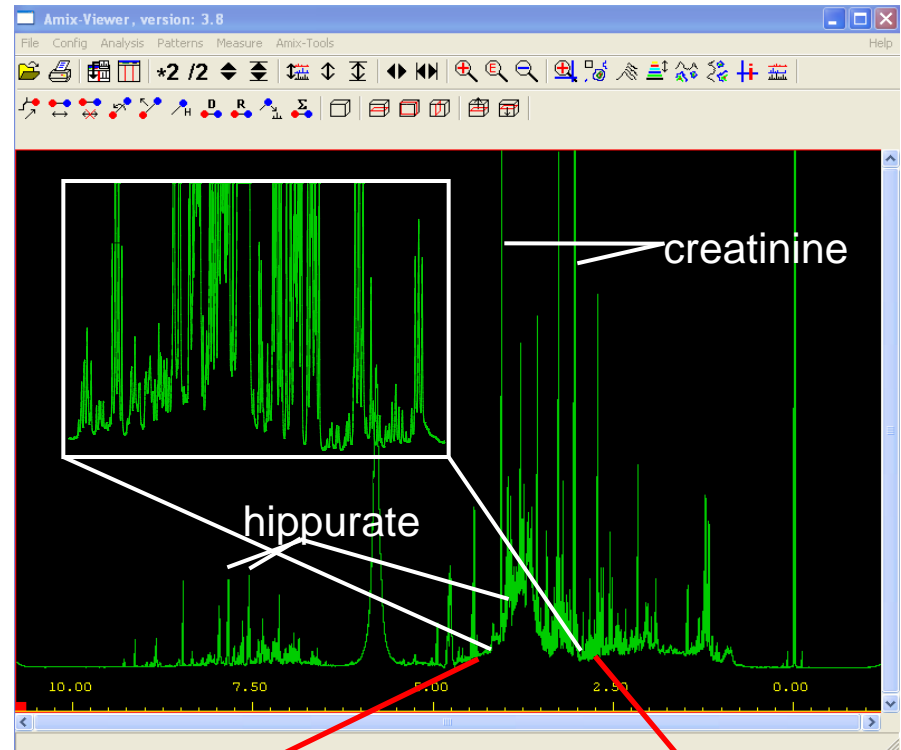
- e.g. to creatinine correction for varying fluid intake

- Scale

- using e.g. vsn to obtain uniform distribution of variance

$$I_S = \frac{1}{K} \sqrt{I_R + E}$$

Kohl, S.M., Klein, M.S., Hochrein, J., Oefner, P.J., Spang, R. & **Gronwald, W.** State-of-the Art Data Normalization Methods Improve NMR-Based Metabolomic Analysis. *Metabolomics*, advance online access, doi: 10.1007/s11306-011-0350-z (2011).

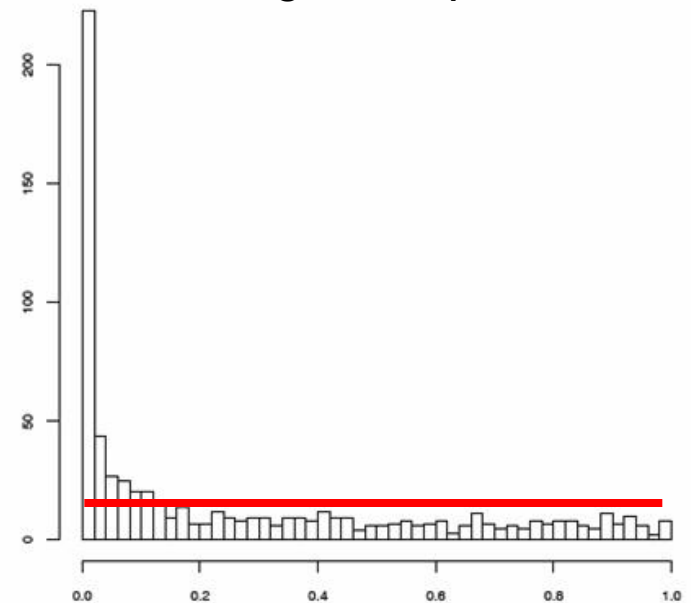


Relevant metabolites

- *Search* for differentially produced metabolites
 - Use of simple two-sided *t*-test for each bucket

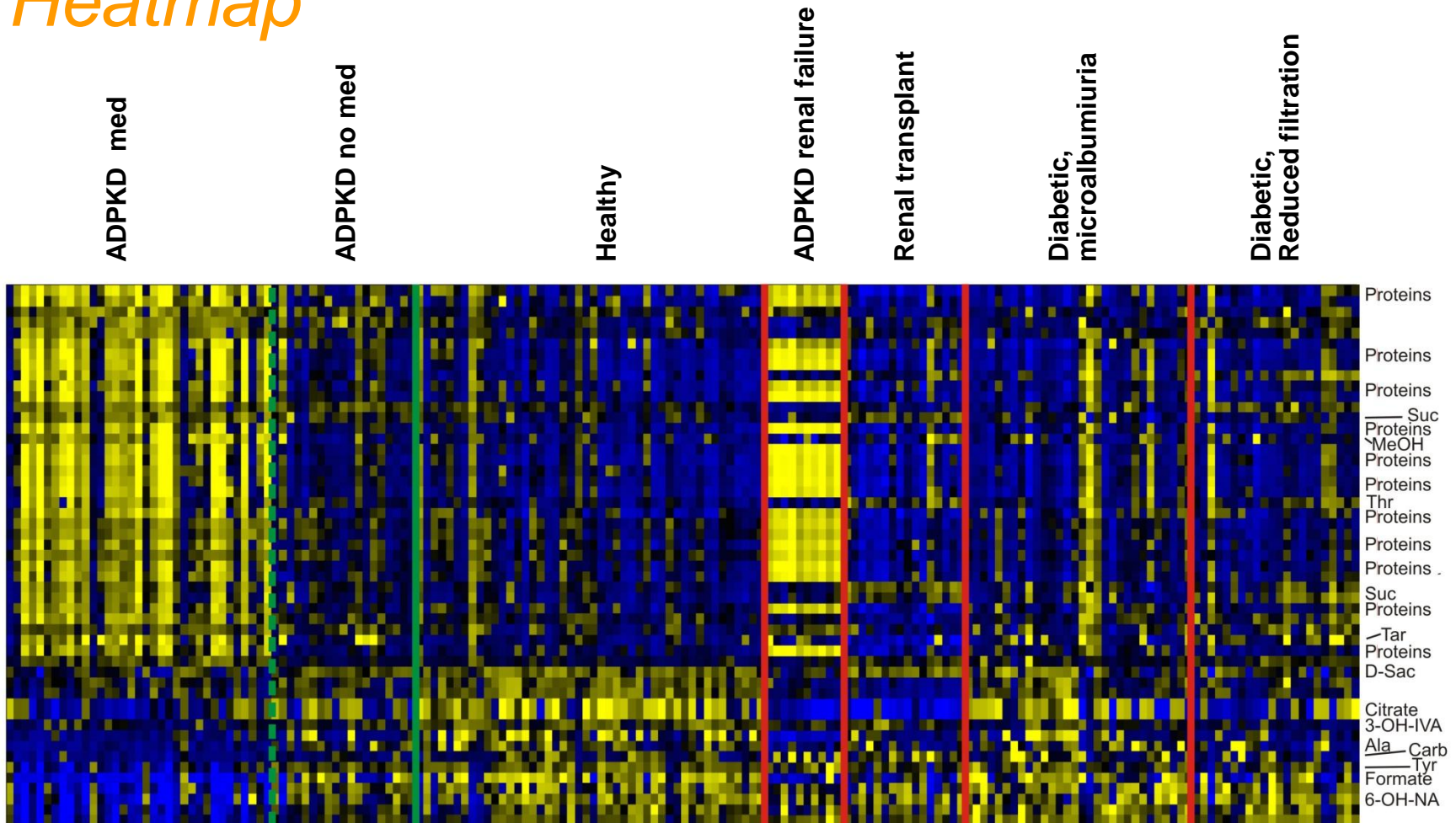
	healthy	diseased	...
Bucket 1	Int	Int	Int
Bucket 2	Int	Int	Int
Bucket 3	Int	Int	Int
⋮			

Histogram of *p*-values



Regulated metabolites

Heatmap



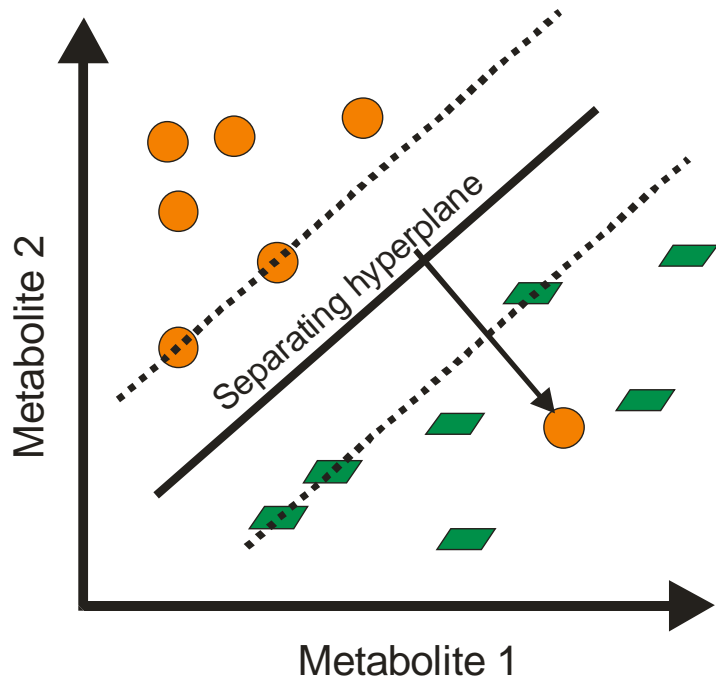
- Used abbreviations: 3-hydroxyisovaleric acid (3-OH-IVA), 6-hydroxynicotinic acid (6-OH-NA), alanine (Ala), carbohydrates (Carb), D-saccharic acid (D-Sac), methanol (MeOH), sucrose (Suc), tartaric acid (Tar), threonine (Thr), and tyrosine (Tyr).

Interpretation of results

- *Metabolites*

- Increased excretion of proteins in diseased patients hints at tubular problems in the kidneys
- Reduced excretion of citrate in diseased patients is a known complication that might lead to the development of kidney stones -> patients receive extra citrate
- Increased excretion of methanol something really new in context of kidney cysts -> generation of new hypotheses

Classification of samples



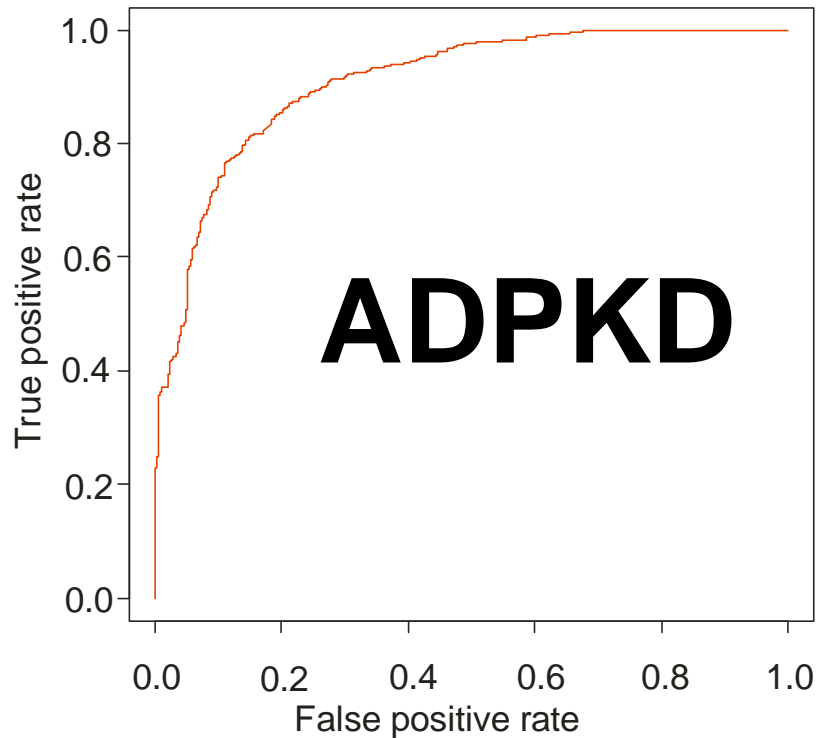
- Example for classification using **support vector machine (SVM)**
 - Optimal hyperplane (thick black line)
 - Red and green dots represent patients
 - Dot on dashed line called support vectors, define separating hyperplane
 - Each dimension represents one metabolite (bucket) 2 are shown here
 - Some misclassifications allowed (arrow) controlled by cost function

- Classification

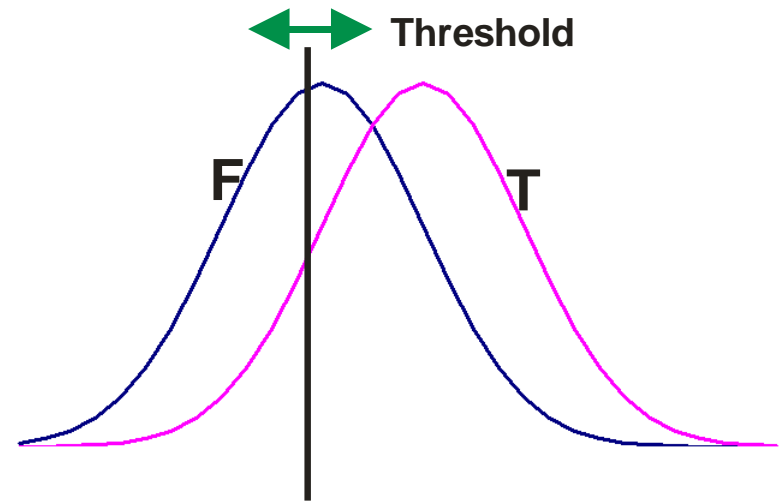
- **SVM** was trained to predict healthy/diseased patients
- Feature selection based on t-test
- Nested cross validation with inner loop parameter optimization.
- Number of features (buckets) optimized to e.g. 51

Performance of classification

- 54 ADPKD patients versus 46 healthy controls
- **Average prediction accuracy** of SVM $85.0 \pm 3.1\%$
- **Area under ROC curve** = 0.91



- SVM calculates for each patient probability for being diseased/healthy
- For classification different probability thresholds can be selected -> ROC analysis



- Classification results from nested cross validation on average 51 features

Summary

- *NMR in ADPKD*
 - Clear discrimination between diseased ADPKD patients and healthy controls
 - Identification of relevant metabolites
 - Distinction of ADPKD and other kidney diseases
 - Clear differences between ADPKD patients who received medication and those who did not
 - Increased methanol levels seem to be specific for ADPKD in a progressed state

Acknowledgment

University of Regensburg, Institute of Functional Genomics

- Matthias S. Klein
- Stefanie Kohl
- Markus Deutschman
- Carridad Louis
- Peter J. Oefner

University Clinic of Erlangen, Department of Nephrology

- Raoul Zeltner
- Bernd Schulze
- Kai-Uwe Eckardt

University Clinic of Regensburg Department for Internal Medicine II

- Stephan Reinhold
- Carsten A. Böger
- Bernhard Banas

Financial support

- Bavarian Genomic Network



- BMBF-Grant FUGATO-plus MeGA-M
- Reform-C

