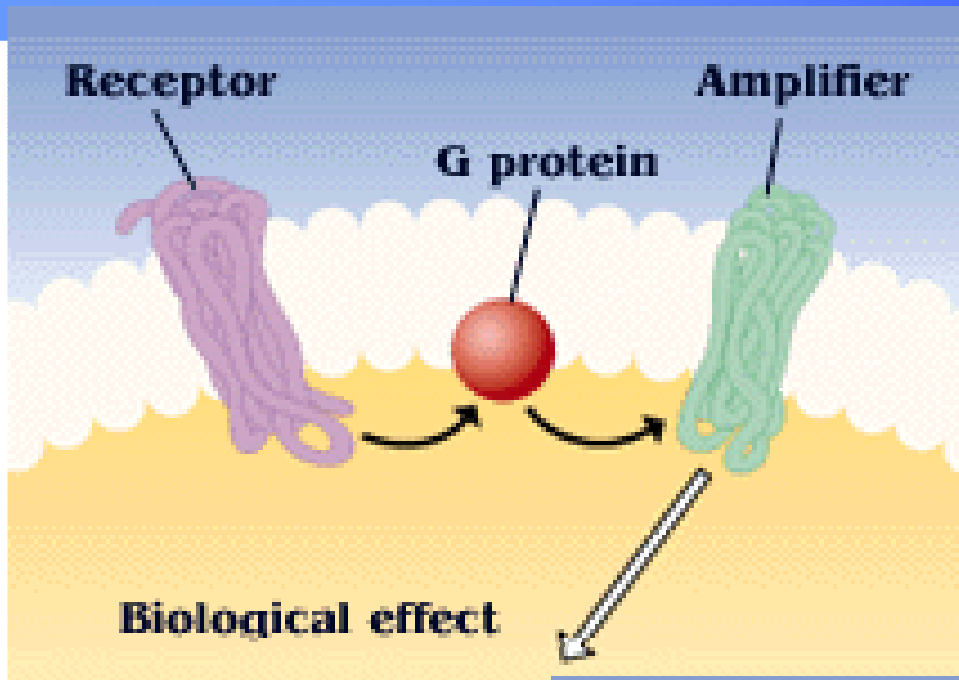


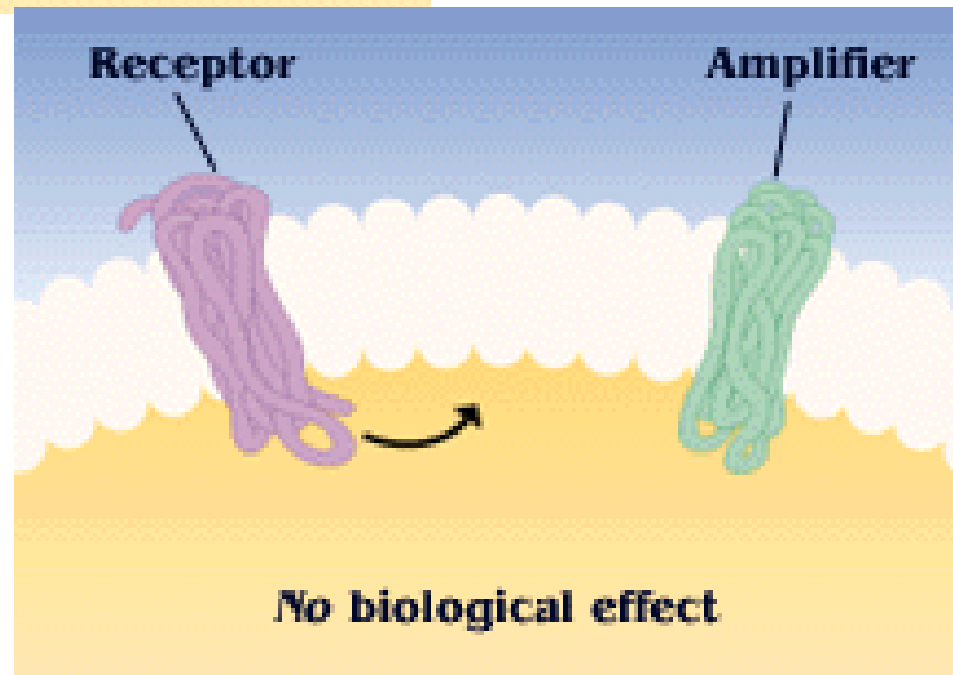
Bioinformatics, Metabolomics and personalized medicine

A yellow scroll graphic with a black outline, featuring a vertical strip on the left side and a small circular detail at the top right corner. The text is centered within the scroll.

Torbjörn Lundstedt
KI 080609

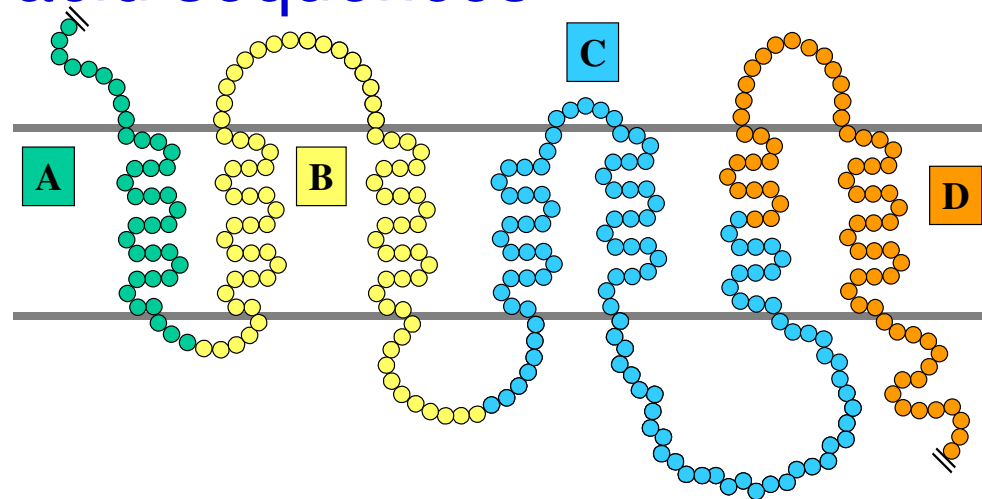


Multivariate analysis of G-protein coupled receptors



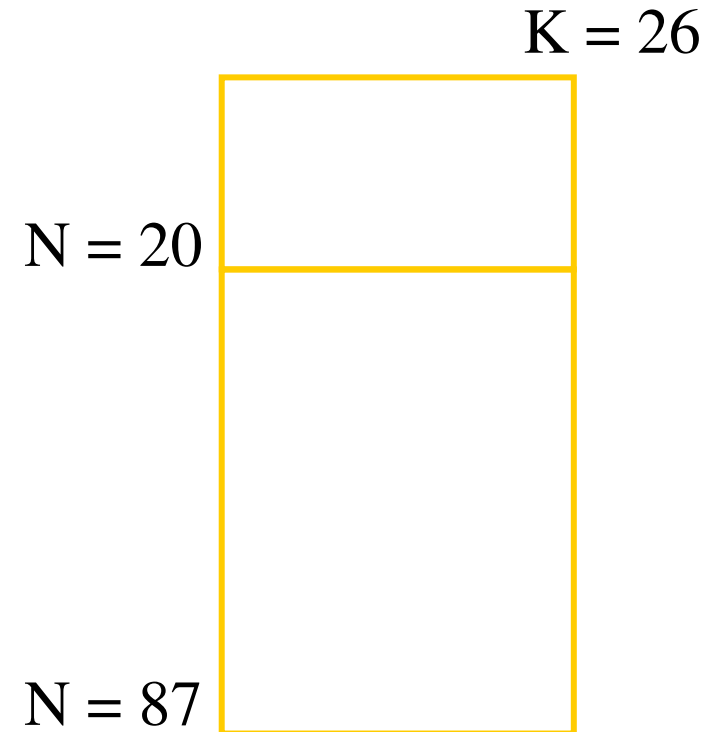
Bioinformatic examples

- G-proteincoupled receptors 2D structure
GPCR
- 7 trans-membrane regions 7TM
- Amino acid sequences



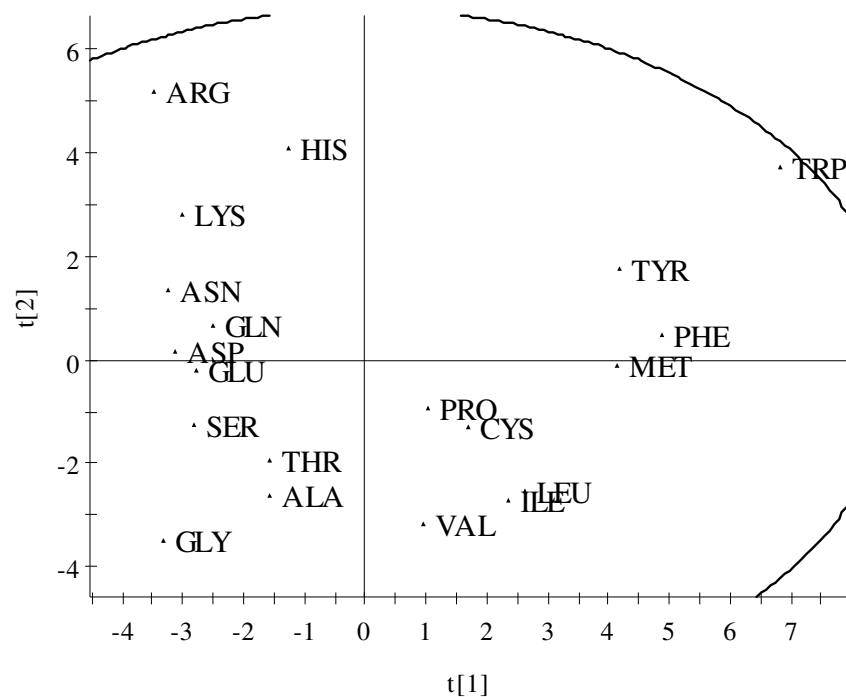
20 natural amino acids

- 20 natural amino acids
- Characterised with both experimental and calculated variables
- Observations = 20
- Variables = 26
(in this example)
- Have been expanded to 87 amino acids
(Maria Sandberg et al.)

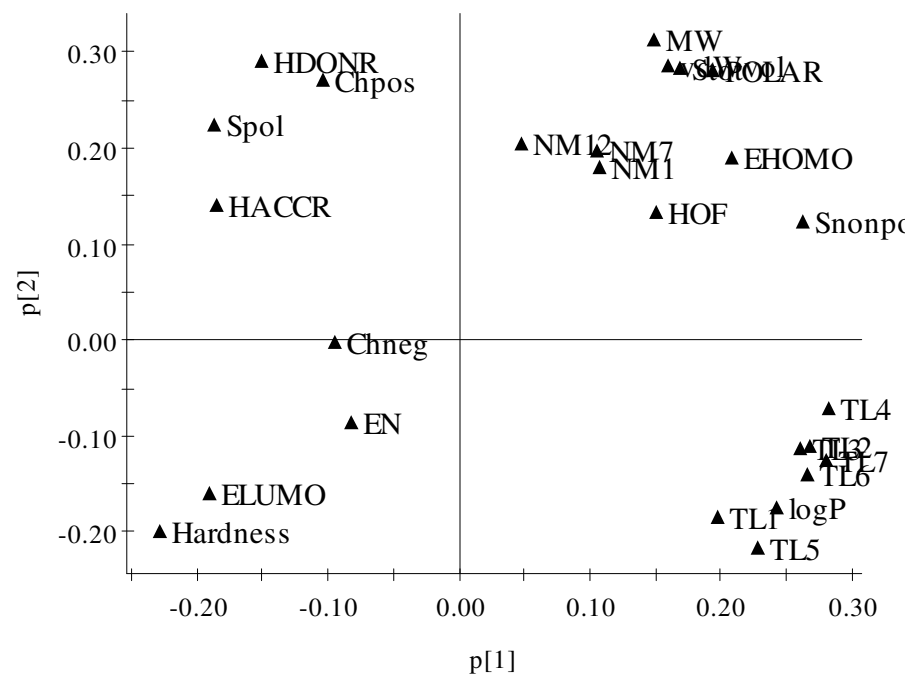


20 natural amino acids

87mia_all-020522.M2 (PCA-X), 20 natural mias data
t[Comp. 1]/t[Comp. 2]

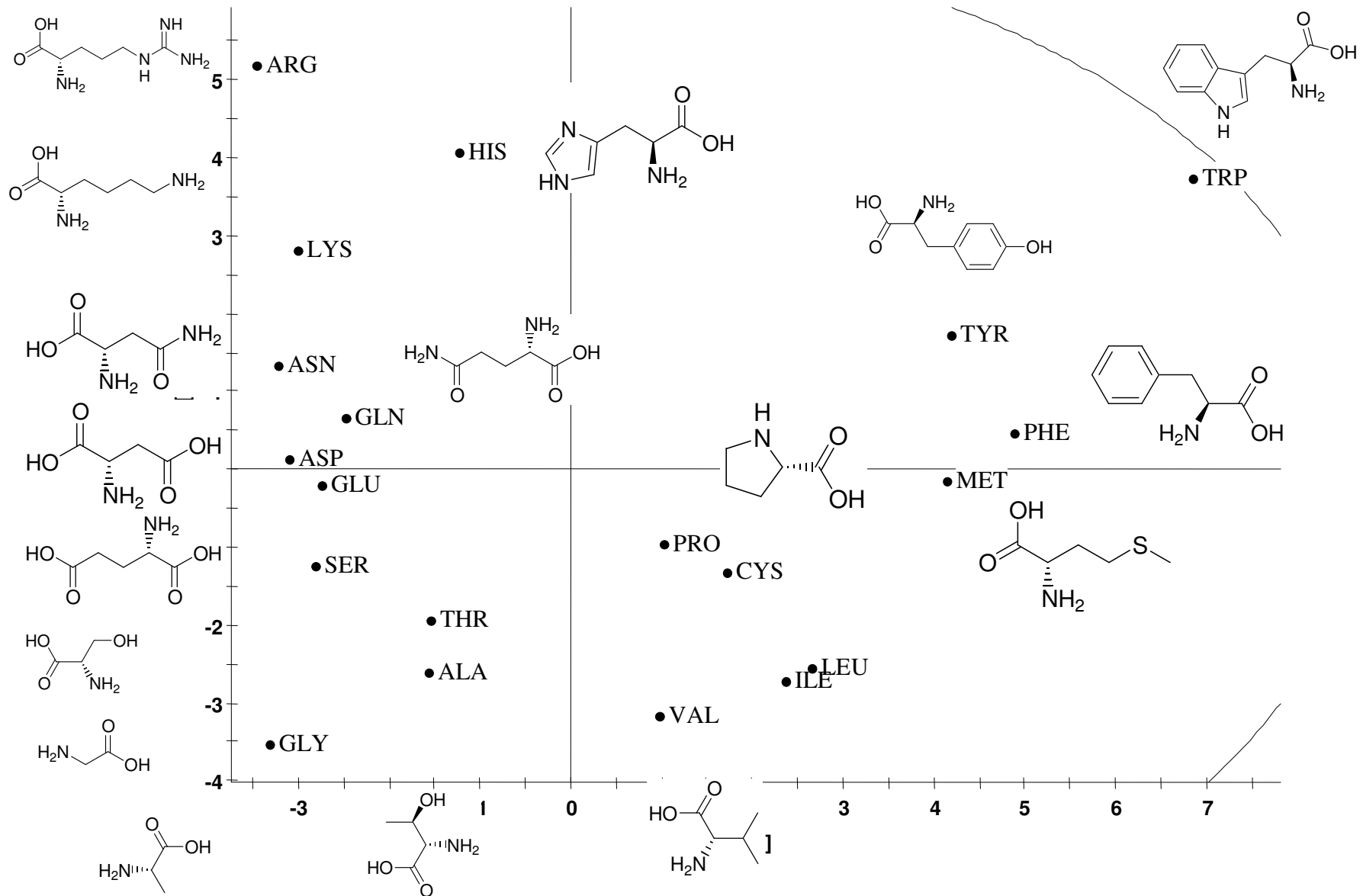


87mia_all-020522.M2 (PCA-X), 20 natural mias data
p[Comp. 1]/p[Comp. 2]



- Two first principal components t_1 vs. t_2 \rightarrow compare with p_1 and p_2 for comparison

20 natural amino acids



Principal properties

— natural amino acids —

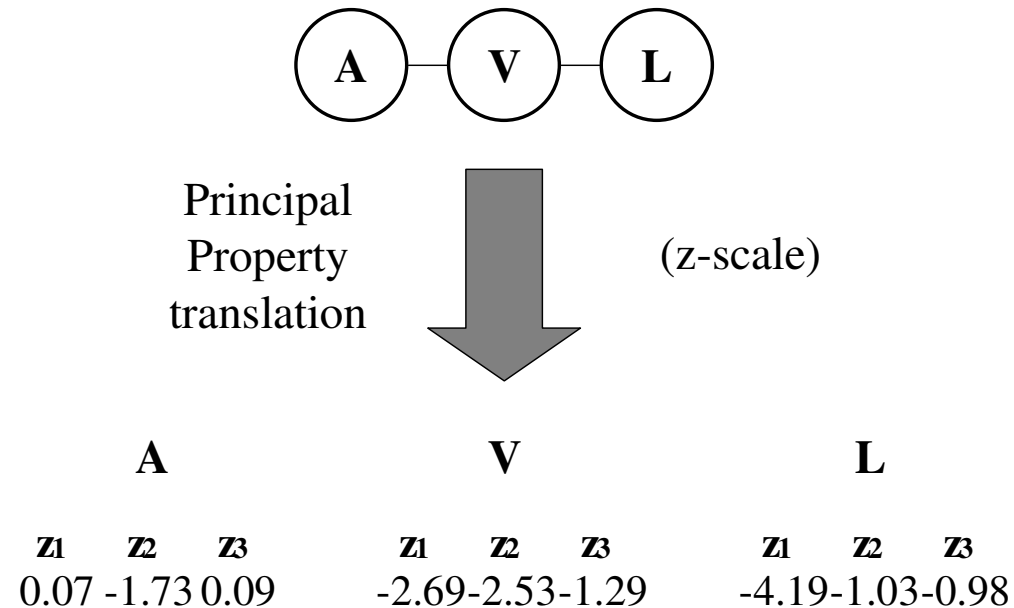
No.	Name	Name	One letter code	z_1	z_2	z_3
1	Alanine	ALA	A	0.07	-1.73	0.09
2	Valine	VAL	V	-2.69	-2.53	-1.29
3	Leucine	LEU	L	-4.19	-1.03	-0.98
4	Isoleucine	ILE	I	-4.44	-1.68	-1.03
5	Proline	PRO	P	-1.22	0.88	2.23
6	Phenylalanine	PHE	F	-4.92	1.30	0.45
7	Tryptophan	TRP	W	-4.75	3.65	0.85
8	Methionine	MET	M	-2.49	-0.27	-0.41
9	Lysine	LYS	K	2.84	1.41	-3.14
10	Arginine	ARG	R	2.88	2.52	-3.44
11	Histidine	HIS	H	2.41	1.74	1.11
12	Glycine	GLY	G	2.23	-5.36	0.30
13	Serine	SER	S	1.96	-1.63	0.57
14	Threonine	THR	T	0.92	-2.09	-1.40
15	Cysteine	CYS	C	0.71	-0.97	4.13
16	Tyrosine	TYR	Y	-1.39	2.32	0.01
17	Asparagine	ASN	N	3.22	1.45	0.84
18	Glutamine	GLN	Q	2.18	0.53	-1.14
19	Aspartic acid	ASP	D	3.64	1.13	2.36
20	Glutamic acid	GLU	E	3.08	0.39	-0.07

- Amino acids characterised with experimental and calculated variables
- PCA generates three descriptors for each amino acid
 - z_1 high lipophilic, low hydrofobic
 - z_2 high large, low small
 - z_3 elektronik properties

Hellberg *et al.*, 1987 (20 aa 's)
 Jonsson *et al.* 1989 (55 aa 's)
 Sandberg *et al.* 1998 (87 aa 's)

Principal properties

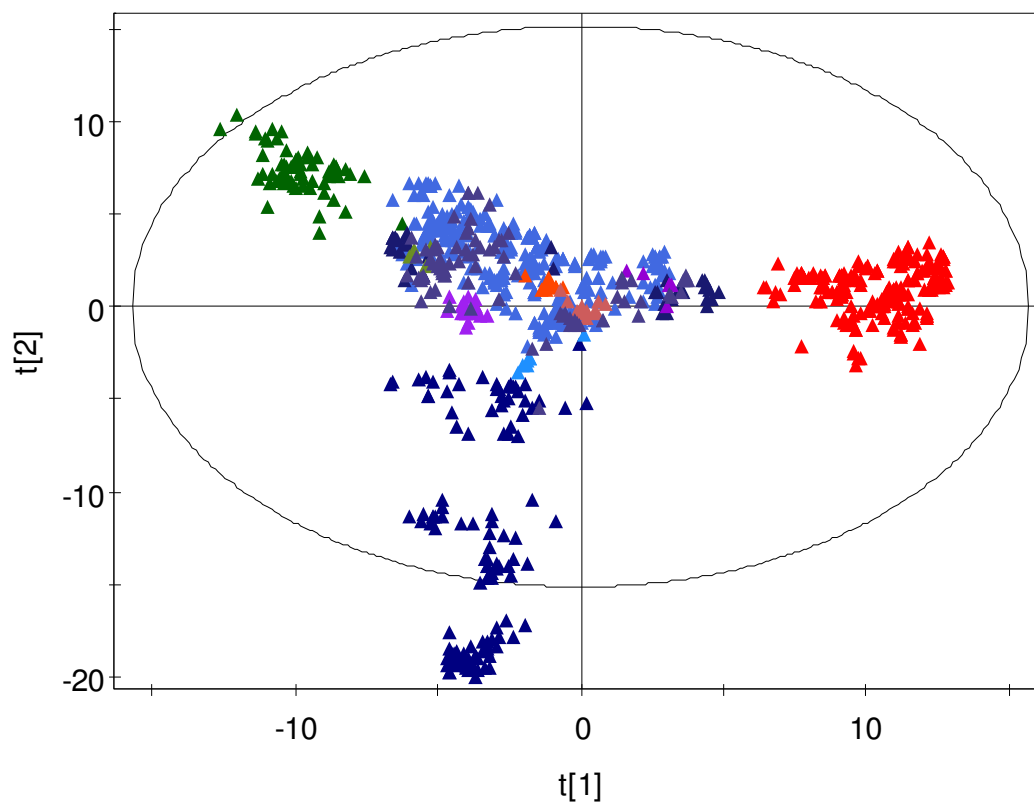
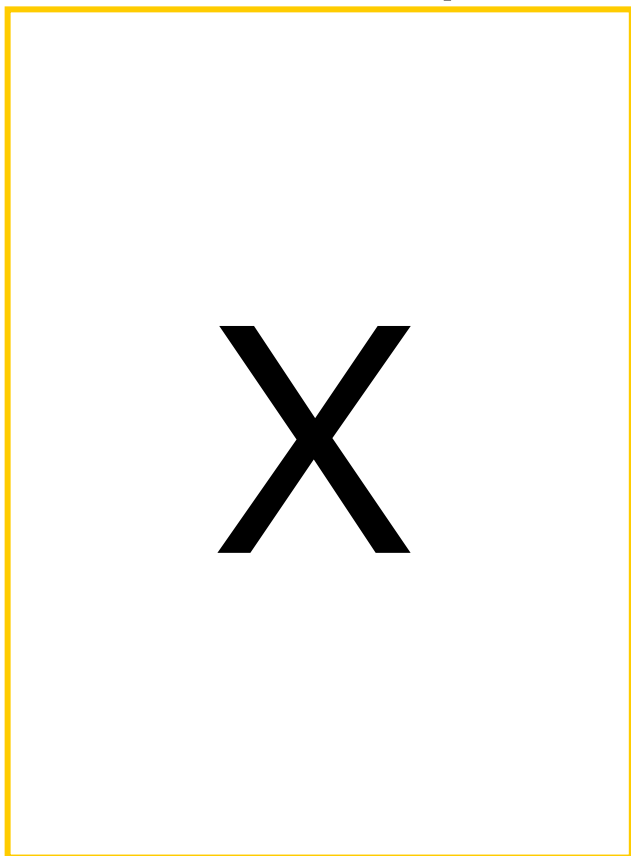
— natural amino acids —



Used to characterise amino acid sequences!

7TM GPRCs analysis, t1/t2

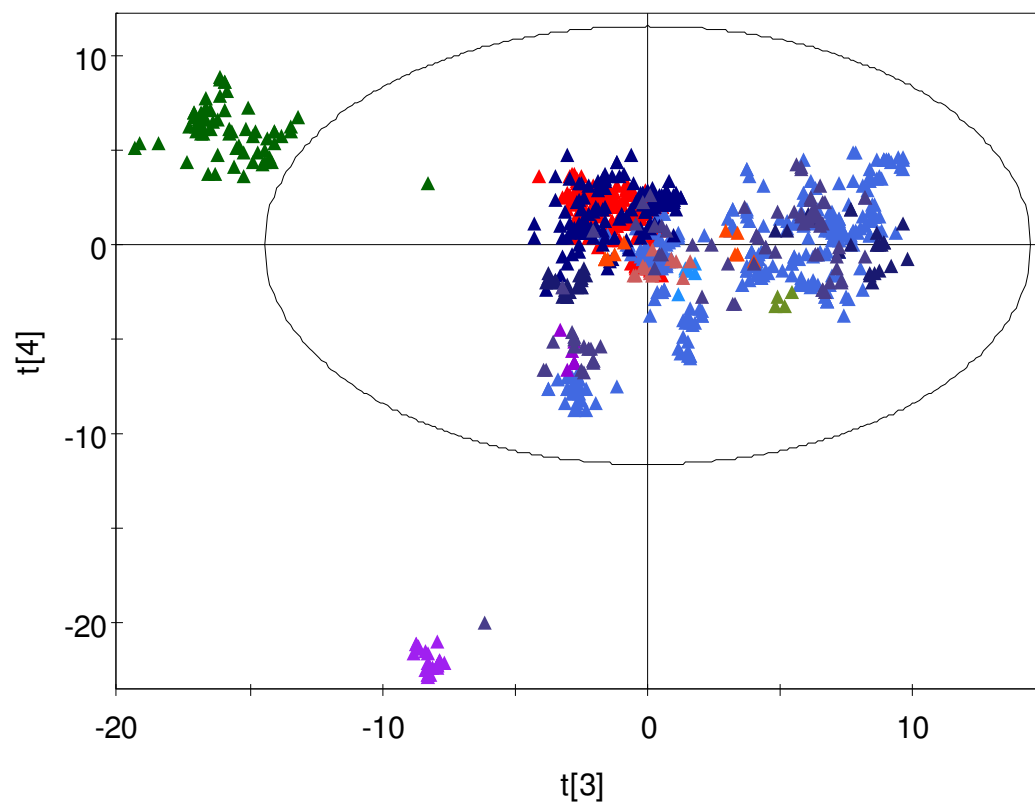
675 variables
(135×5)



897 sequences

7TM GPRCs analysis, t3/t4

gpa-s.M15 (PC), PCA all färg, Work set
Scores: t[3]/t[4]



Ellipse: Hotelling T2 (0,05)
Simca-P8.0 by Umetrics AB 2001-09-06 12:58

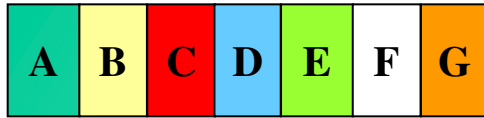
Hierarchical PCA

- Dividing the data in different levels depending on origin
- More than one PCA model to examine one problem
- Simplify interpretation

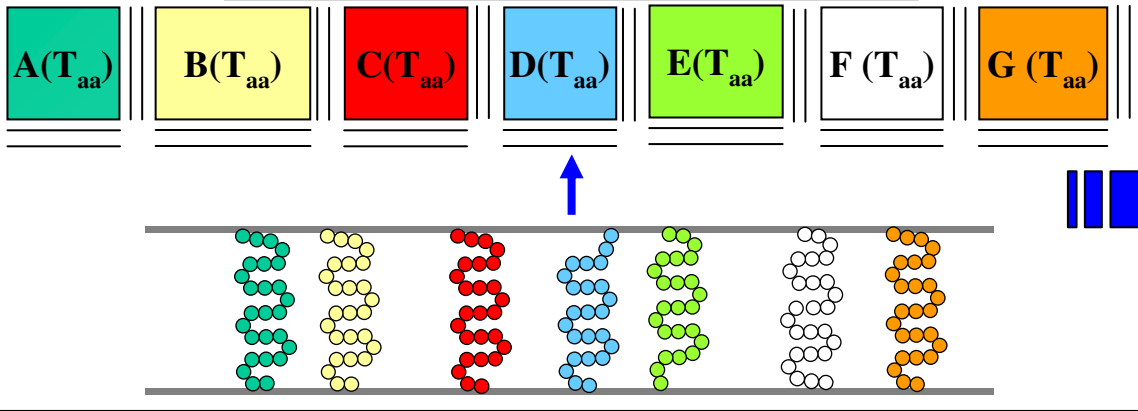
Characterisation

Level 1 – Interpretation

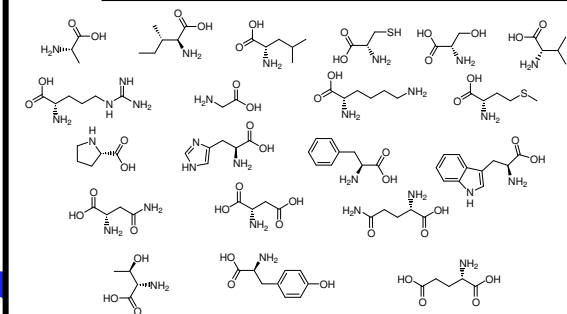
Receptor characterisation



Trans-membrane characterisation



Amino acid characterisation

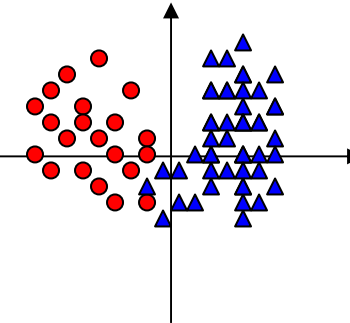


PCA

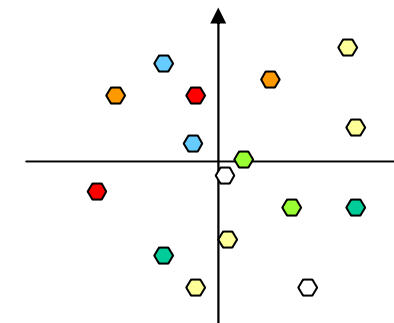


One receptor level

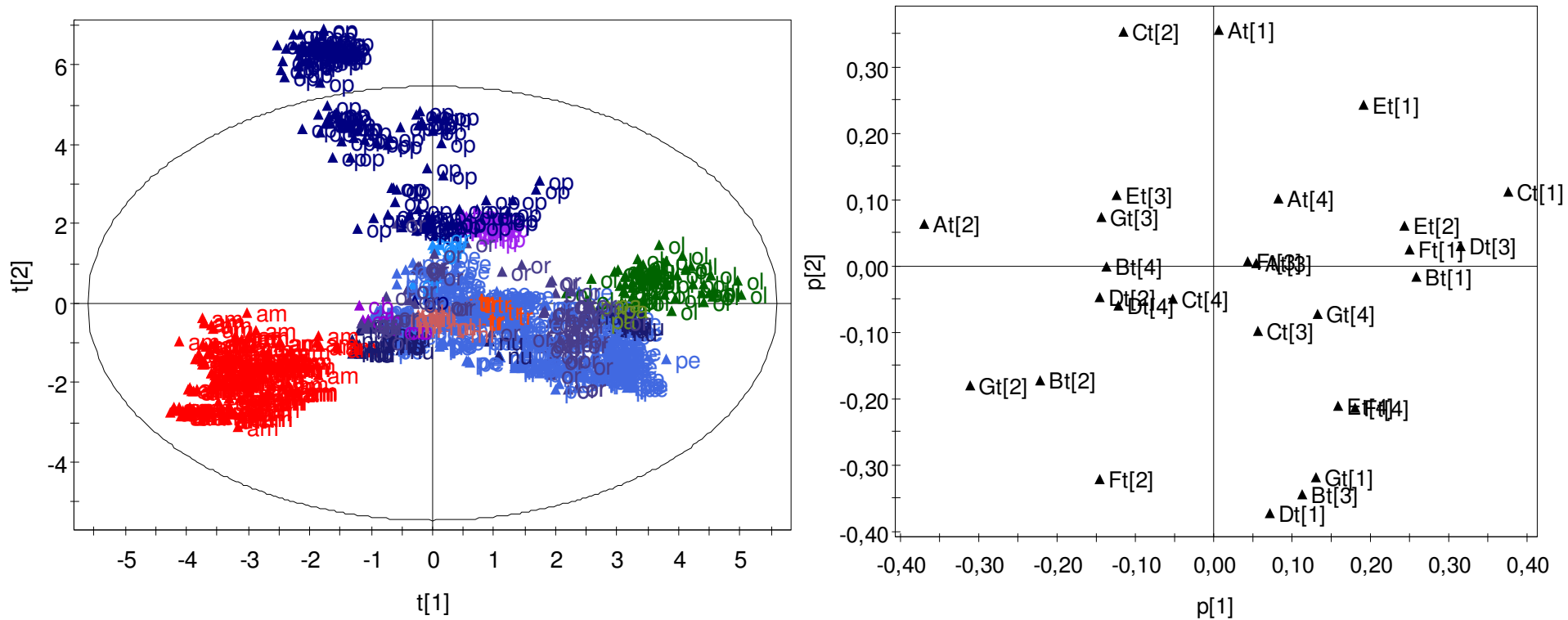
Principal scores



Principal loadings

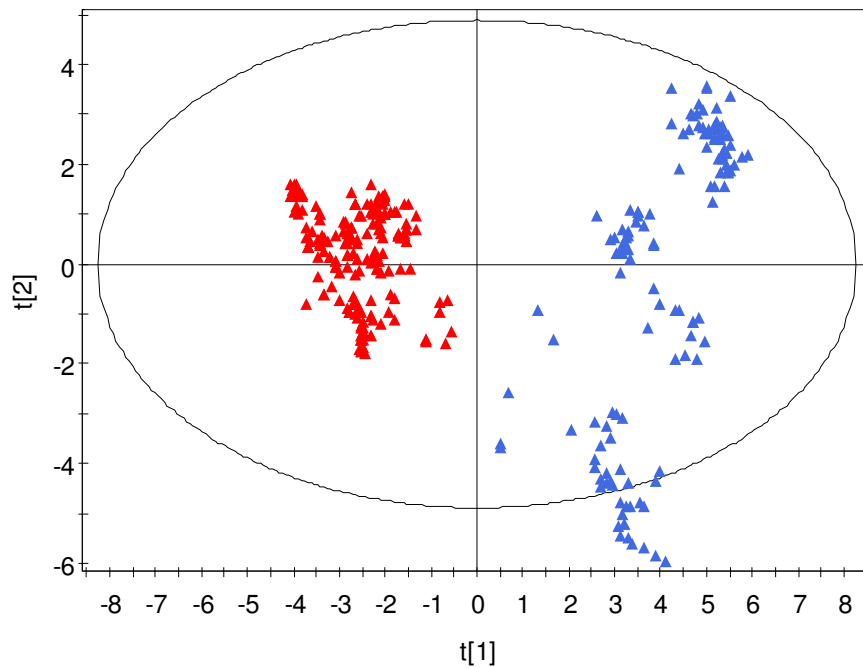


Hierarchical PCA



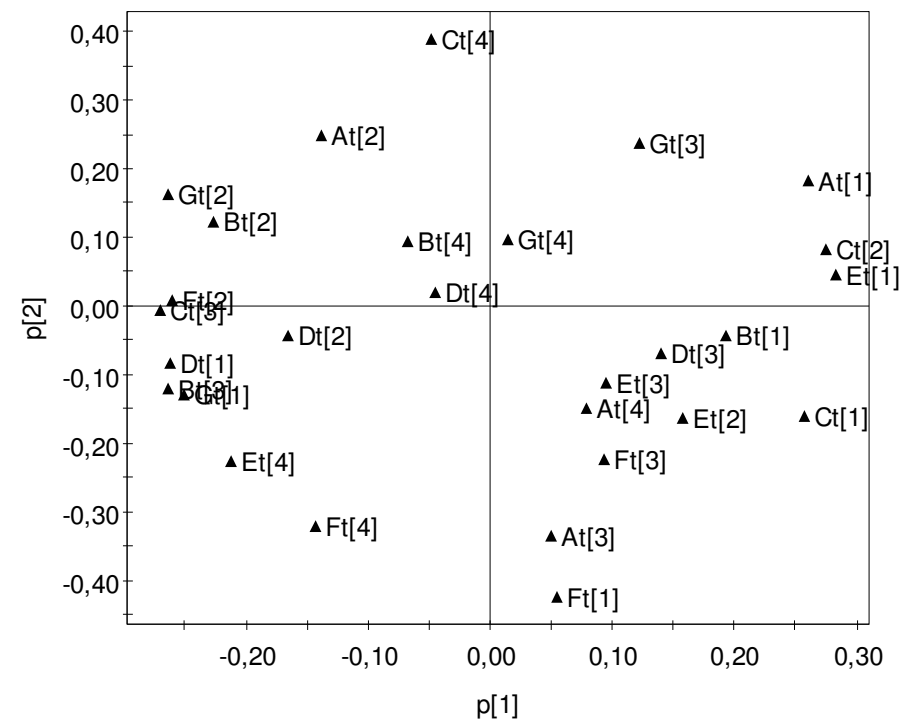
Analysis of sub-groups

s_hierar.M5 (PC), fyra komp., Work set
Scores: t[1]/t[2]



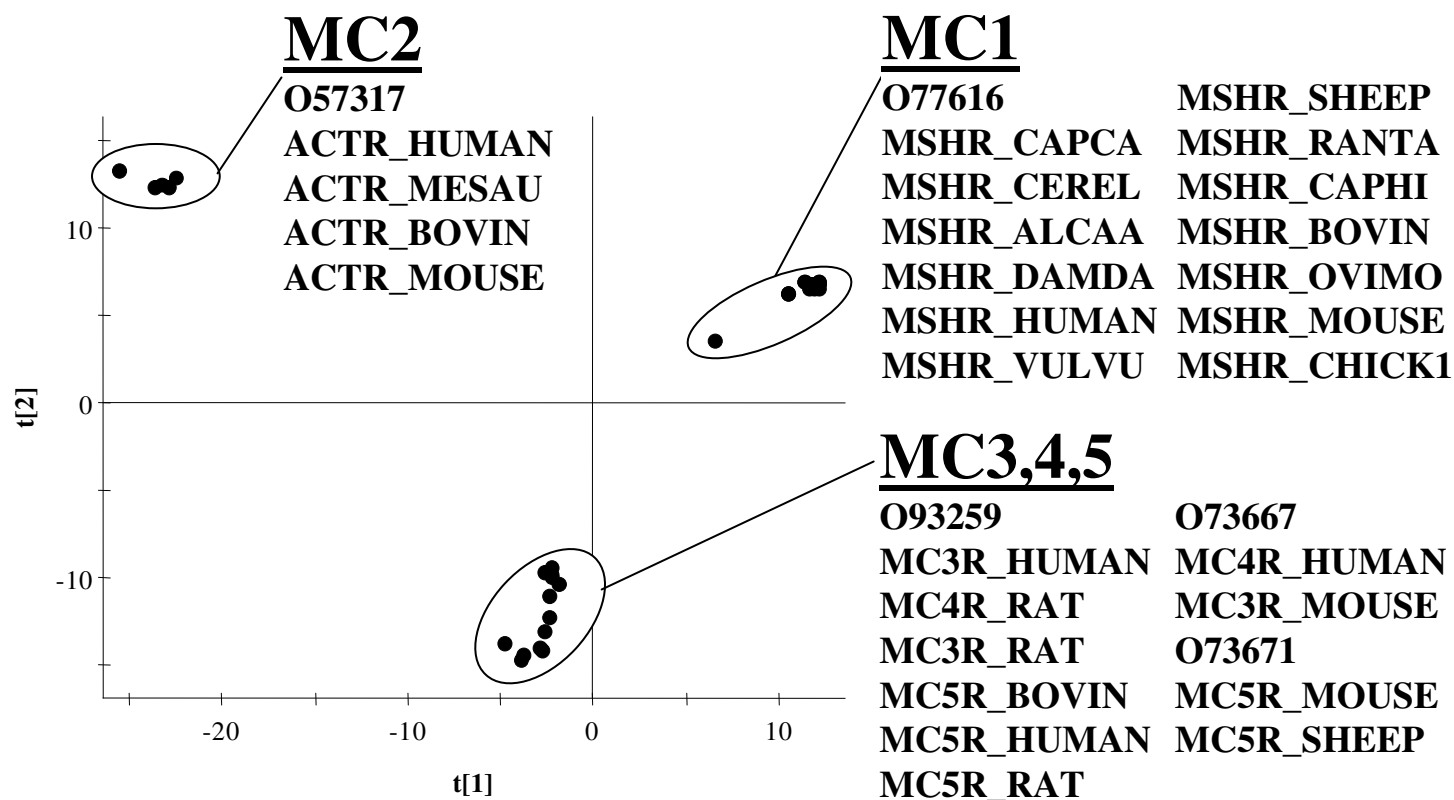
Ellipse: Hotelling T2 (0,05)
Simca-P8.0 by Umetrics AB 2001-09-10 09:24

s_hierar.M5 (PC), fyra komp., Work set
Loadings: p[1]/p[2]



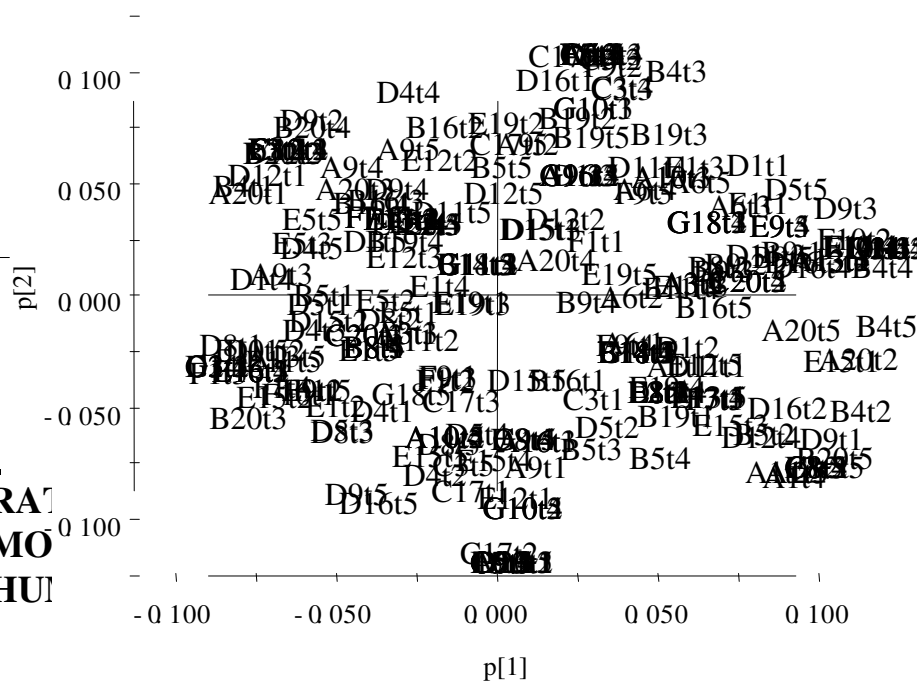
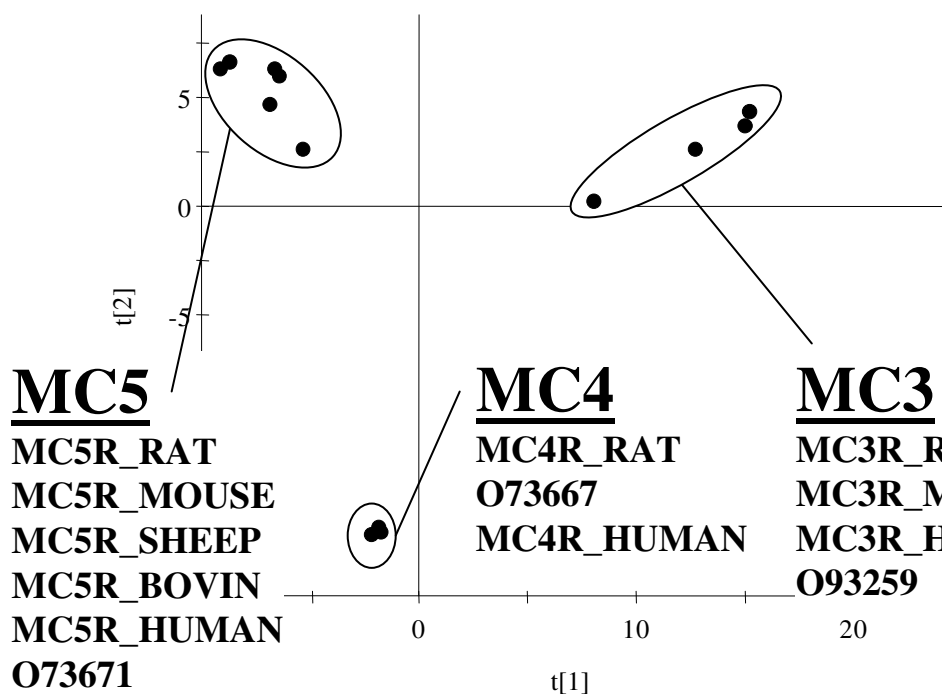
Classification – GPCR

Analysis of the melanocortin receptors; MC1, MC2, MC3, MC4 and MC5. MC1 and MC2 are quite different compared to MC3-MC5.

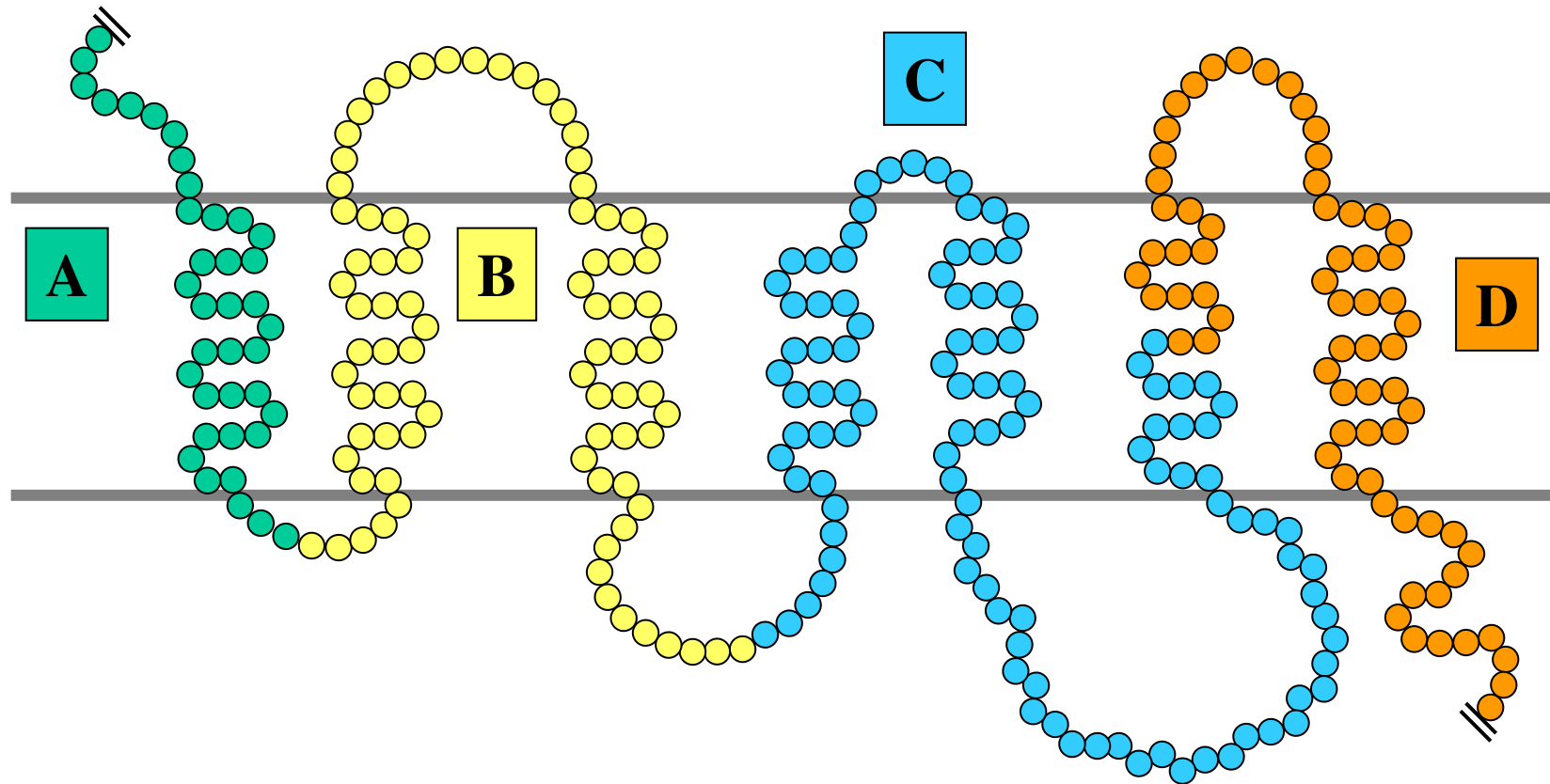


Classification – GPCR

7tm.M34 (PCA-X), PCA mc3,4,5
p[Comp. 1]/p[Comp. 2]



Schematic Overview of MC1 and MC3 Receptor Chimeras



Divided in four parts – characterised with binary coding

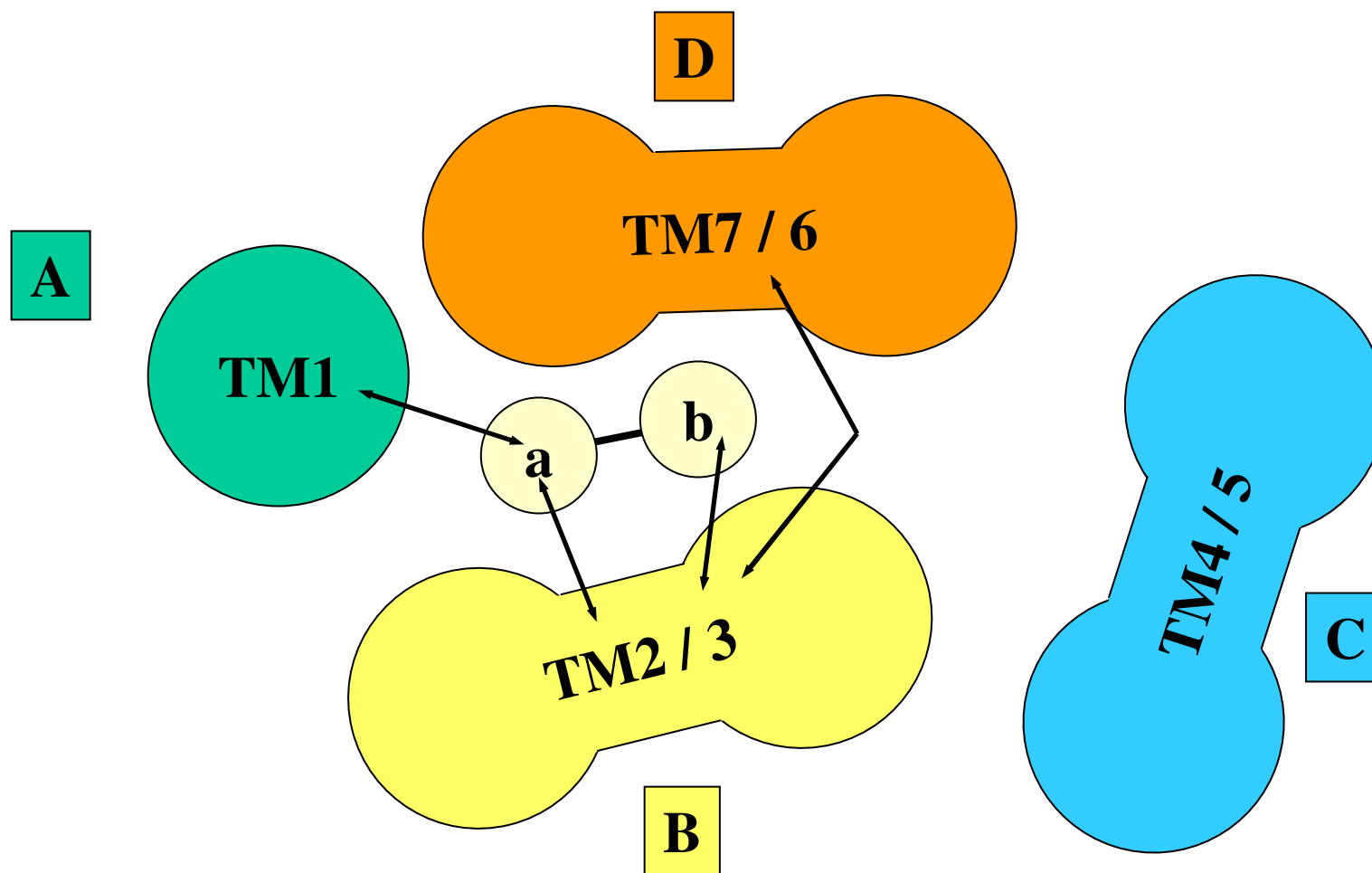
Sequences of Melanocortin Receptor Active Peptides

Position	1	2	3	4	5	6	7	8	9	10	11	12	13
Seq. Part	a										b		
Peptide													
MSH	Ser	Tyr	Ser	Met	Glu	His	Phe	Arg	Trp	Gly	Lys	Pro	Val
MS04	Ser	Ser	Ile	Ile	Ser	His	Phe	Arg	Trp	Gly	Leu	Cys	Asp
MS05	Ser	Ser	Ile	Ile	Ser	His	Phe	Arg	Trp	Gly	Lys	Pro	Val
MS06	Ser	Tyr	Ser	Met	Glu	His	Phe	Arg	Trp	Gly	Leu	Cys	Asp

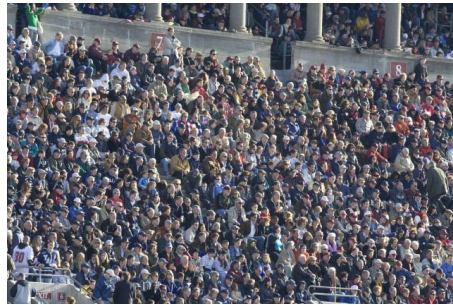
Receptor-Ligand-Response Matrix

Molecules		Descriptors X Y				Response		
Z	X	Z				X		Y
Receptor	Peptide	A	B	C	D	a	b	Log(K _i)
A ₁ B ₁ C ₁ D ₁	a ₁ b ₁	0	0	0	0	0	0	0.21
A ₁ B ₁ C ₁ D ₁	a ₂ b ₂	0	0	0	0	1	1	0.82
A ₁ B ₁ C ₁ D ₁	a ₂ b ₁	0	0	0	0	1	0	-0.07
A ₁ B ₁ C ₁ D ₁	a ₁ b ₂	0	0	0	0	0	1	1.65
A ₁ B ₂ C ₂ D ₂	a ₁ b ₁	0	1	1	1	0	0	1.37
A ₁ B ₂ C ₂ D ₂	a ₂ b ₂	0	1	1	1	1	1	3.82
A ₁ B ₂ C ₂ D ₂	a ₂ b ₁	0	1	1	1	1	0	2.03
A ₁ B ₂ C ₂ D ₂	a ₁ b ₂	0	1	1	1	0	1	2.72
A ₁ B ₁ C ₁ D ₂	a ₁ b ₁	0	0	0	1	0	0	1.37
A ₁ B ₁ C ₁ D ₂	a ₂ b ₂	0	0	0	1	1	1	3.36
A ₁ B ₁ C ₁ D ₂	a ₂ b ₁	0	0	0	1	1	0	2.3
A ₁ B ₁ C ₁ D ₂	a ₁ b ₂	0	0	0	1	0	1	2.18
A ₁ B ₂ C ₂ D ₁	a ₁ b ₁	0	1	1	0	0	0	1.56
A ₁ B ₂ C ₂ D ₁	a ₂ b ₂	0	1	1	0	1	1	4.14
A ₁ B ₂ C ₂ D ₁	a ₂ b ₁	0	1	1	0	1	0	2.31
A ₁ B ₂ C ₂ D ₁	a ₁ b ₂	0	1	1	0	0	1	2.83
A ₁ B ₂ C ₁ D ₁	a ₁ b ₁	0	1	0	0	0	0	1.85
A ₁ B ₂ C ₁ D ₁	a ₂ b ₂	0	1	0	0	1	1	4.19
A ₁ B ₂ C ₁ D ₁	a ₂ b ₁	0	1	0	0	1	0	2.57
A ₁ B ₂ C ₁ D ₁	a ₁ b ₂	0	1	0	0	0	1	3.7
A ₁ B ₁ C ₂ D ₂	a ₁ b ₁	0	0	1	1	0	0	1.55
A ₁ B ₁ C ₂ D ₂	a ₂ b ₂	0	0	1	1	1	1	3.11
A ₁ B ₁ C ₂ D ₂	a ₂ b ₁	0	0	1	1	1	0	2.41
A ₁ B ₁ C ₂ D ₂	a ₁ b ₂	0	0	1	1	0	1	2.4
A ₁ B ₁ C ₂ D ₁	a ₁ b ₁	0	0	1	0	0	0	0.54
A ₁ B ₁ C ₂ D ₁	a ₂ b ₂	0	0	1	0	1	1	1.84
A ₁ B ₁ C ₂ D ₁	a ₂ b ₁	0	0	1	0	1	0	0.86
A ₁ B ₁ C ₂ D ₁	a ₁ b ₂	0	0	1	0	0	1	1.57
A ₂ B ₂ C ₁ D ₁	a ₁ b ₁	1	1	0	0	0	0	1.77
A ₂ B ₂ C ₁ D ₁	a ₂ b ₂	1	1	0	0	1	1	3.08
A ₂ B ₂ C ₁ D ₁	a ₂ b ₁	1	1	0	0	1	0	2.76
A ₂ B ₂ C ₁ D ₁	a ₁ b ₂	1	1	0	0	0	1	2.88
A ₂ B ₂ C ₂ D ₁	a ₁ b ₁	1	1	1	0	0	0	1.84
A ₂ B ₂ C ₂ D ₁	a ₂ b ₂	1	1	1	0	1	1	3.67
A ₂ B ₂ C ₂ D ₁	a ₂ b ₁	1	1	1	0	1	0	2.63
A ₂ B ₂ C ₂ D ₁	a ₁ b ₂	1	1	1	0	0	1	2.86
A ₂ B ₂ C ₂ D ₂	a ₁ b ₁	1	1	1	1	0	0	1.97
A ₂ B ₂ C ₂ D ₂	a ₂ b ₂	1	1	1	1	1	1	4.49
A ₂ B ₂ C ₂ D ₂	a ₂ b ₁	1	1	1	1	1	0	3.02
A ₂ B ₂ C ₂ D ₂	a ₁ b ₂	1	1	1	1	0	1	3.17

Interaction Between Peptide and MCR



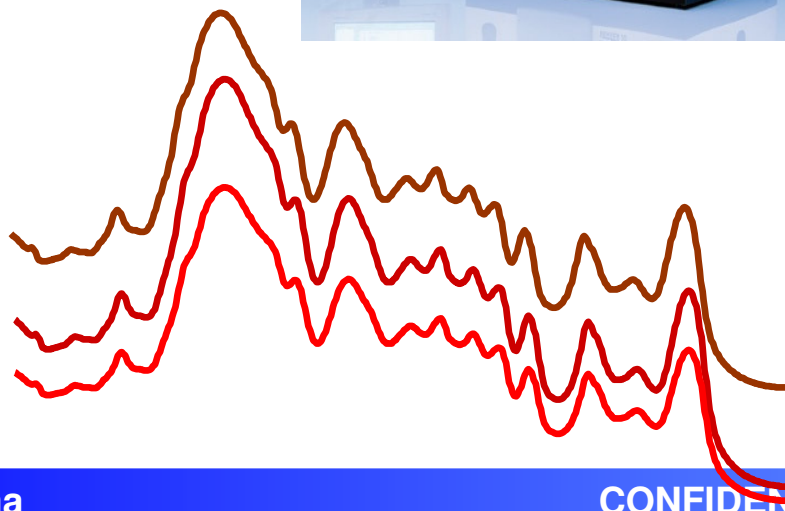
Metabonomics – Simplified view



Biological samples



Biochemical analysis of endogenous metabolites

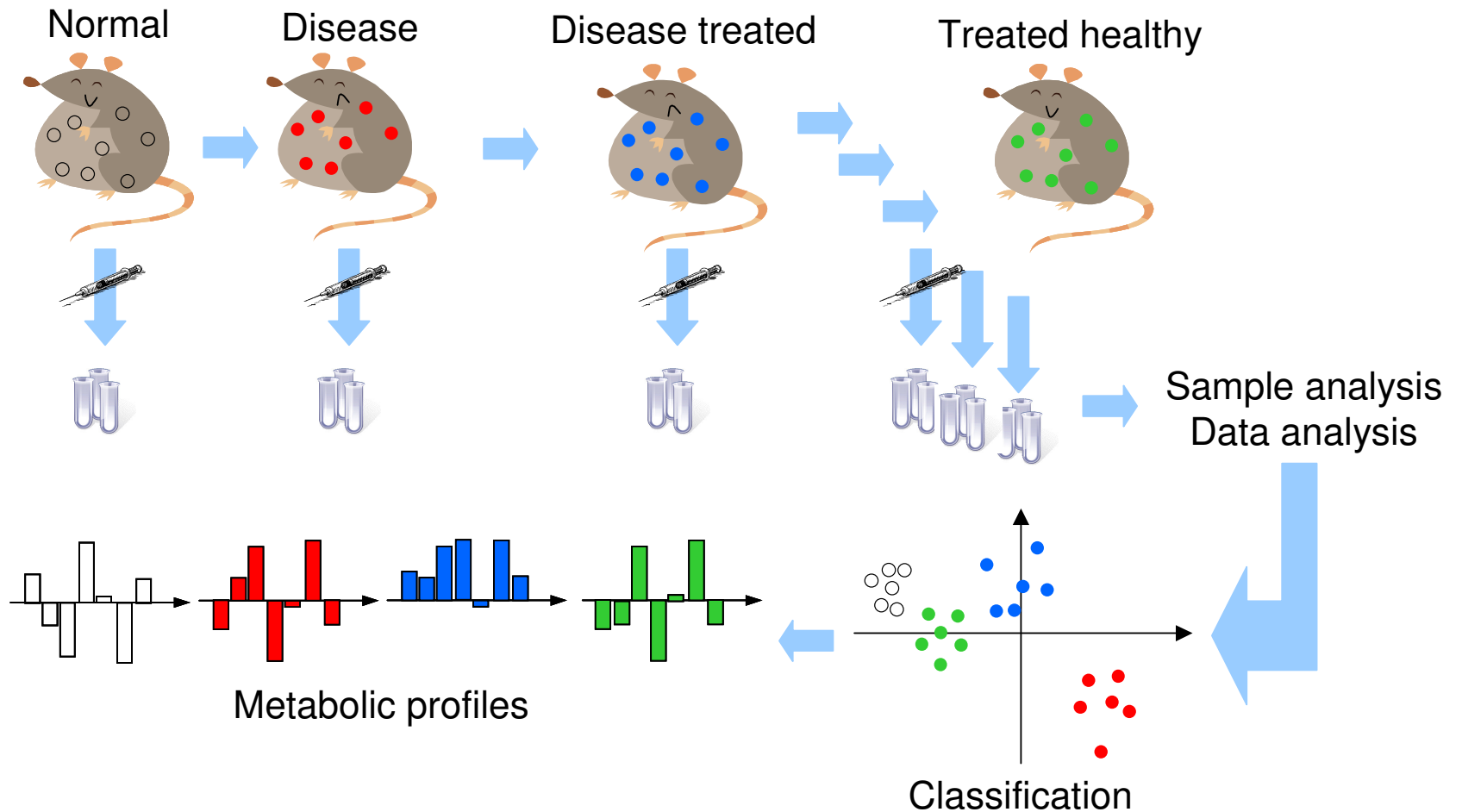


Information?

The challenge in modern biology is not in data collection but rather in maximizing information in data -

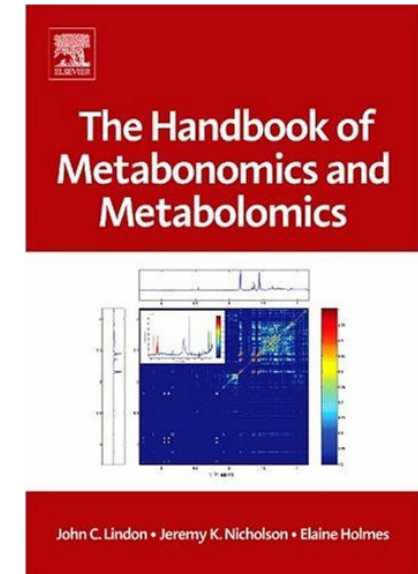
How to do it!

Metabonomics at a glance



Strategy

1. Formalize the aim
 - What do we want?
2. Selection of objects
 - Design of Experiments (DOE)
 - Multivariate design (MVD)
3. Sample preparation and profiling of human and animal samples
 - *In vivo*, *in vitro* samples
 - Blood, Urine, Cerebral Spinal Fluid (CSF)
 - Synovial fluid (joint), Bowel fluids, Feces, Tissues
4. Integration and evaluation of collected data
 - Exploratory analysis, Interpretation & Visualization
 - Prediction models
 - Patterns
 - Target identification
5. Identify and define IPR opportunities and strategies



Pigs

- Heart Ich/rep
- Microdialysis

”Data filtering”

- Large data – difficult to handle and overview
- Starting point – ask specific question and filter the data, use prior information
- Depending on the question the analysis can be focused at finding different patterns
- Information identified in data subsets can thereafter be applied to whole dataset
- Combine metabonomics data with questionnaire → connection genetics vs. lifestyle and environment

Parameters ?

- Example of filtering questions
- Compound classification – prior information
 - ❑ Mechanism of action
 - ❑ Classes of compound
 - ❑ Indication
 - ❑ Side effects
 - ❑ Etc
- Use structure characterisation and PCA to generate principal properties → make representative/diverse selection to base first model on (e.g. sub-set)

Multivariate design

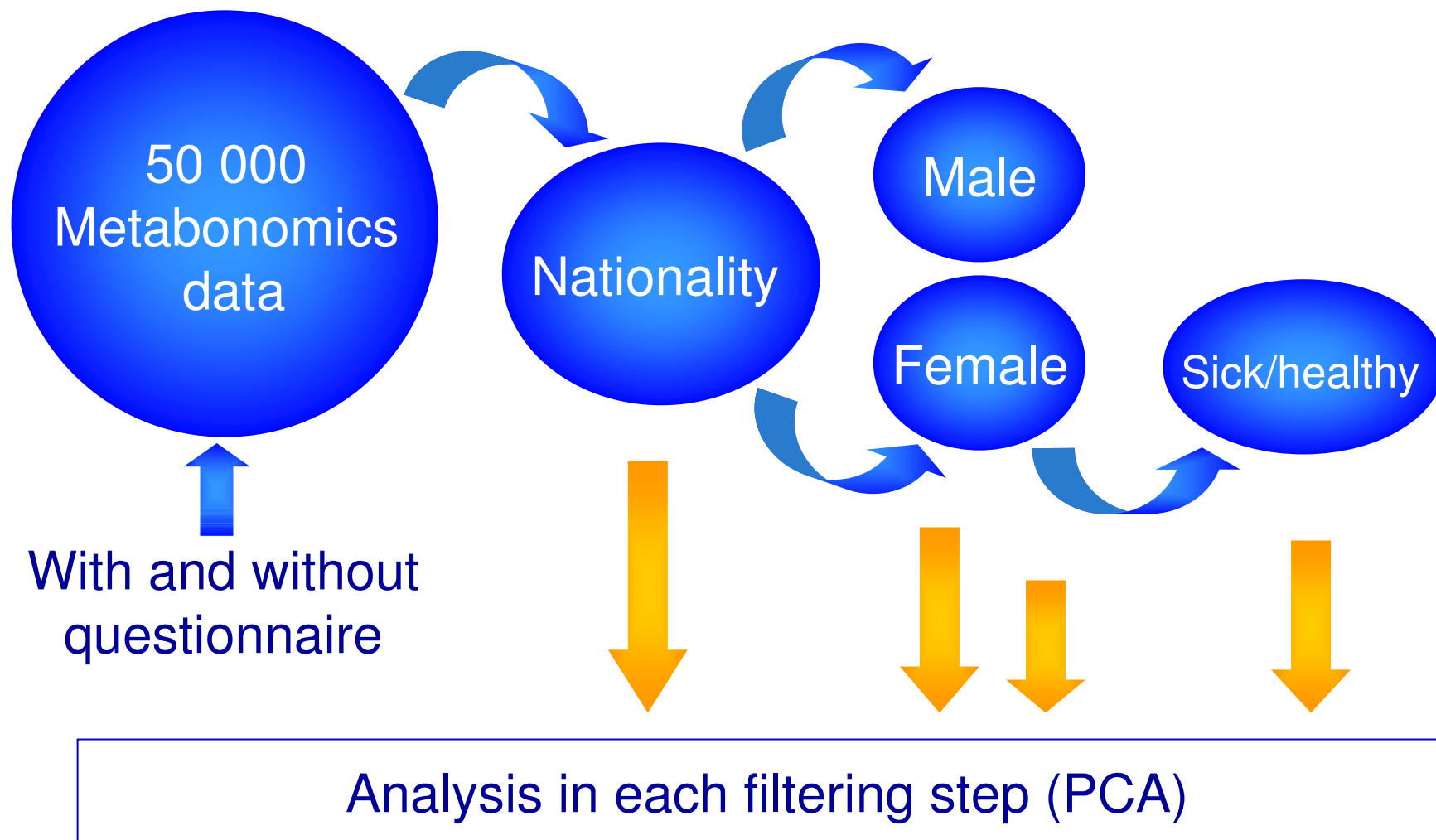
- Make PCA on reduced data set
- Make multivariate design(s) in the generated scores (principal properties) → cluster based design
- Select training set and test set according to a multivariate design – if possible
 - ❑ Validate the model
 - ❑ Obtain a more robust model
 - ❑ Classify
 - ❑ Identify biomarkers

Example of questions for design

- Gender
 - ❑ male/female
Looking for genetic patterns and lifestyle patterns/markers separating the sexes
- Genetic profile
 - ❑ Asian, European, American
Looking for genetic patterns/markers
- Nationality
 - ❑ China, UK or Asian, European
Looking for lifestyle and environment patterns/markers
- Anamnesis
 - ❑ "healthy" vs. un-healthy
Looking for disease pattern/markers

And
combinations
there of

Data layout

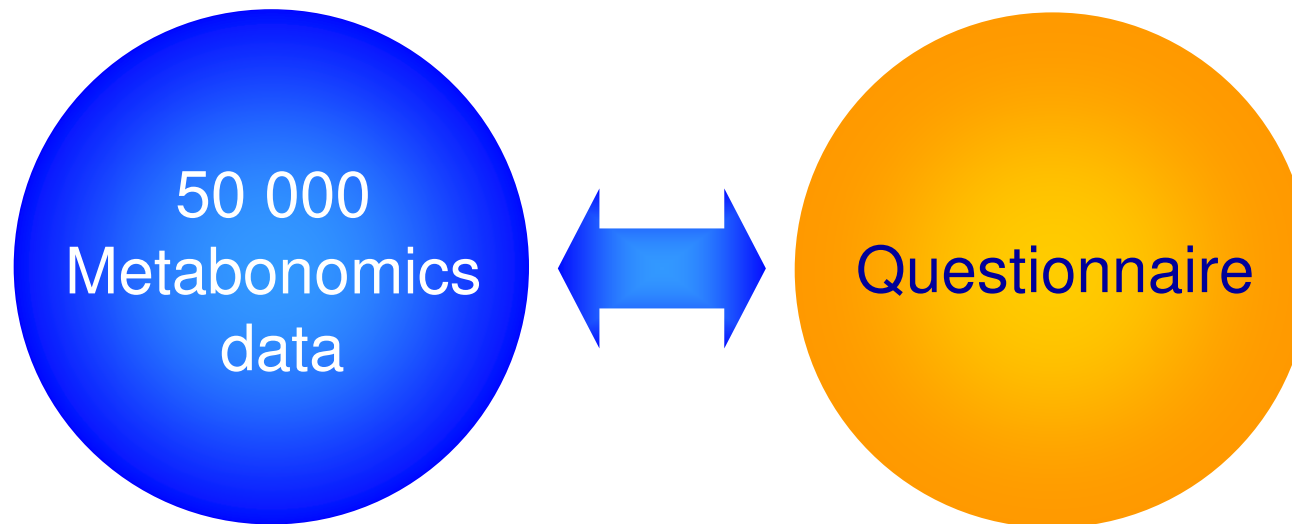


Data layout



- Pregnant (folic acid, hormones)
- Diabetes (insulin)
- Gastric/stomach ulcer (Losec/Nexium)
- Head ace (paracetamole)
- Smoker/non-smoker
- Age < 50 <
- BMI

Lifestyle vs. metabonomics



- Identifying patterns in lifestyle by relating metabonomics to questionnaire (PLS)

Parameters ?

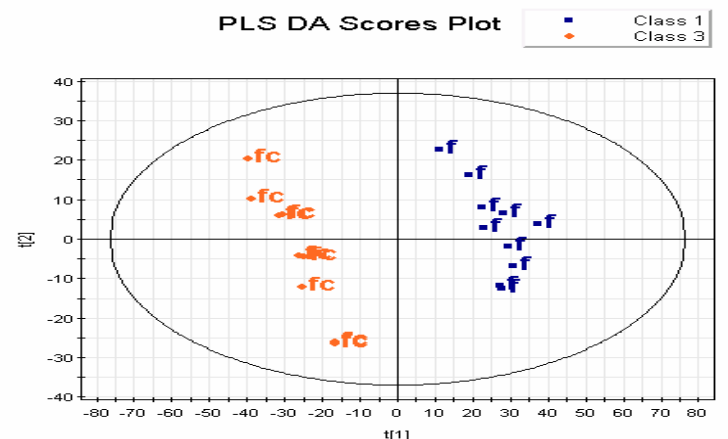
- Example of filtering questions
- Compound classification – prior information
 - ❑ Mechanism of action
 - ❑ Classes of compound
 - ❑ Indication
 - ❑ Side effects
 - ❑ Etc
- Use structure characterisation and PCA to generate principal properties → make representative/diverse selection to base first model on (e.g. sub-set)

Possible problems

- Selection of control group *in vivo*?
- Selection of reference structures?
- Selection of colon's (LC/GC)?
- Sample handling/work up?
- Training set, test set – model validation, classification, biomarker
- Fast/Slow responders

Class specific models 2 - PLS-DA

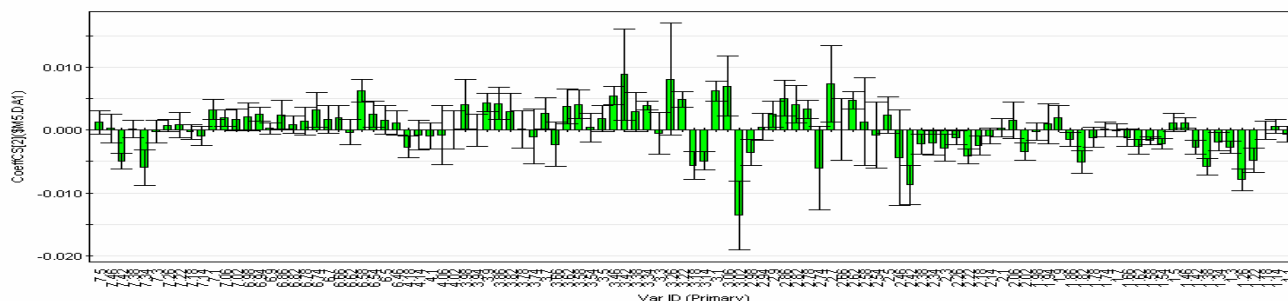
- PLS-DA Maximum separation projection
- PLS model built on group membership (1 or 0)
- Coefficients show as a whole how do the variables change when going from one group to another
- Here controls (f) and treated (fc) are separated



SIMCA-P+ 10.0 - 12/06/2003 14:54:48

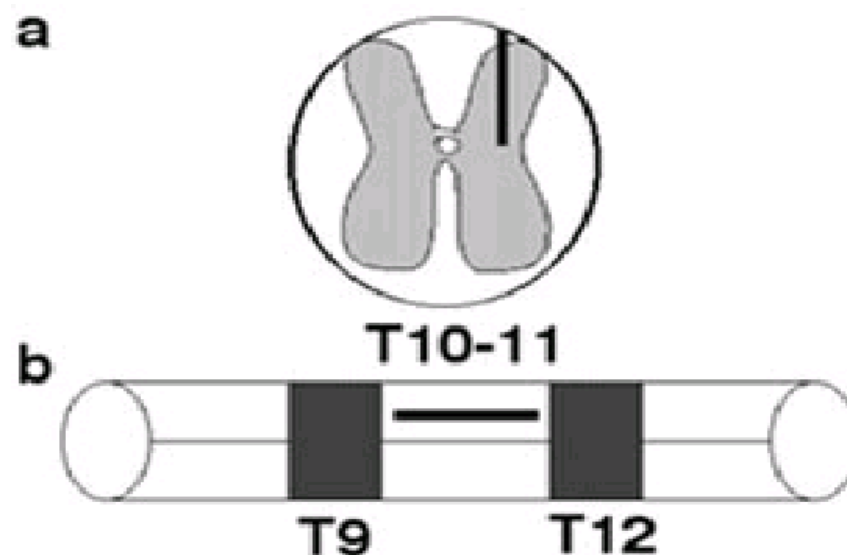
PLS-DA between “f” and “fc”

- Coefficients show as a whole how do the variables change when going from one group to another
- Chemical shift regions or masses influential for the separation of the two classes



Incision model short description

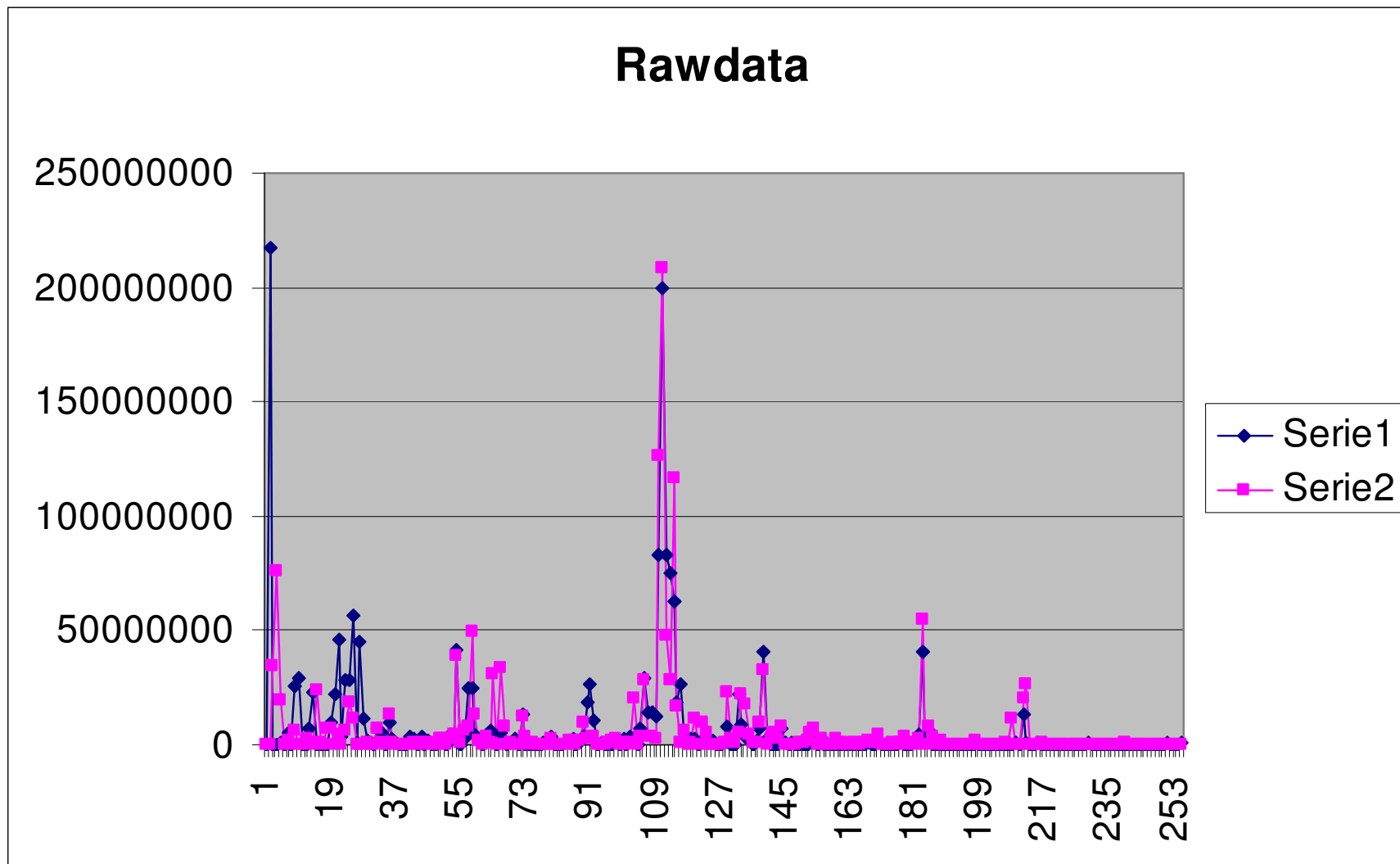
- Surgical intervention in rats. The SCI was induced by making a longitudinal incision into the right dorsal horn of the T10-11 segment
- AP173 and AP713 was given post-injury 5 minutes after injury with a topical dose of 5 μg in saline solution
- Tarlov scale was tested during 1 to 7 days
- The other responses was estimated day 7



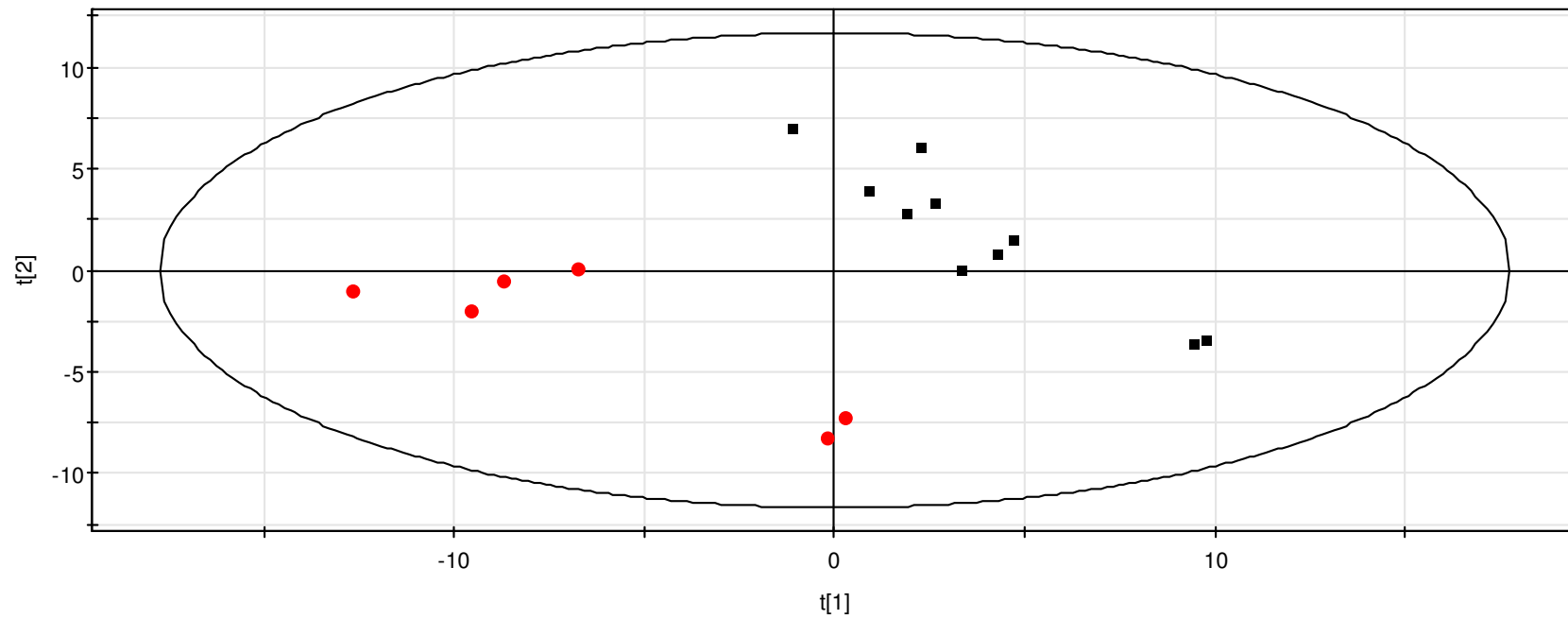
Status

- CSF, Plasma and three segments of spinal cord from normal rats
- CSF, plasma and three segments of spinal cord 5h after injury
- GC-MS and solid phase NMR

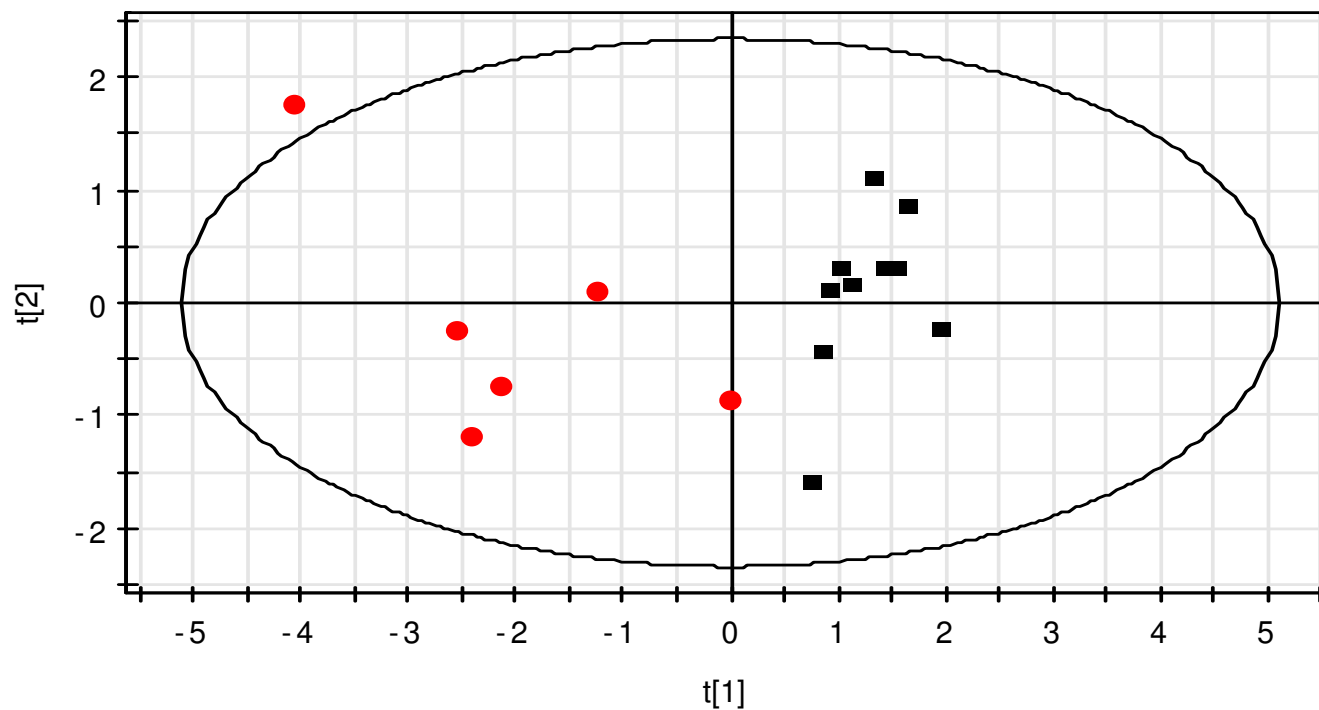
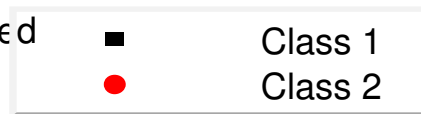
GC/MS PROFILE



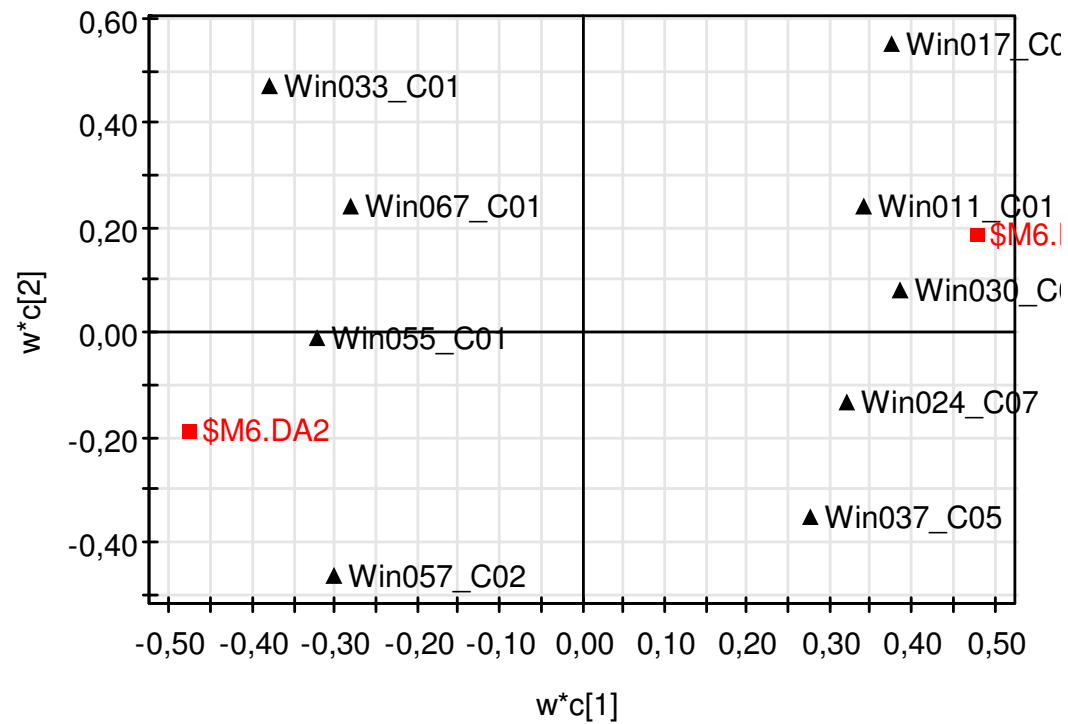
CSF_070130_Reg_process.M3 (PLS-DA), Untitled
t[Comp. 1]/t[Comp. 2]
Colored according to classes in M3



CSF_070130_Reg_process.M6 (PLS-DA), Untitled
t[Comp. 1]/t[Comp. 2]
Colored according to classes in M6



CSF_070130_Reg_process.M6 (PLS-DA), Untitled X
 $w^*c[Comp. 1]/w^*c[Comp. 2]$ Y



Arthritis – Diagnosis

- Early diagnosis critical
 - ❑ Reduce symptoms of wear
 - ❑ More successful treatment with early medication*
- Rheumatoid arthritis
 - ❑ Physical examination, antibodies (today not specific for RA), X-ray
- Osteoarthritis
 - ❑ Physical examination, X-ray
- New diagnostic tools are needed...

* <http://www.reumatikerforbundet.org/start.asp?sida=3955>

Case study 1: Sample information

Analysed by gas chromatography/time of flight mass spectrometry (GC/TOFMS)

Deconvolution (MCR v.1.13)

- ❑ 59 Samples
- ❑ 181 Spectral Profiles

Outliers (possible)

- ❑ Not primary detected

Included Isotope labeled mixture

- ❑ 11 Internal standards + Methyl Stearate

Main Groups

❑ 3 Health Groups

- Control
- Ra
- Oa

Rheumatoid Arthritis – brief background

- Worldwide prevalence of approximately 1%*
- Autoimmune disease, the body attacks itself, aetiology largely unknown**
- Treatment; irreversible disease, no known cure, medication to maintain mobility and ease pain
- Early diagnosis critical
 - ❑ Reduce symptoms of wear
 - ❑ More successful treatment with early medication*
- Diagnosis for rheumatoid arthritis
 - ❑ Physical examination, antibodies (today not specific for RA), X-ray, MRI
- New diagnostic tools are needed...

* Feldmann, M. et al., Cell. 85: 307-310 (1996)

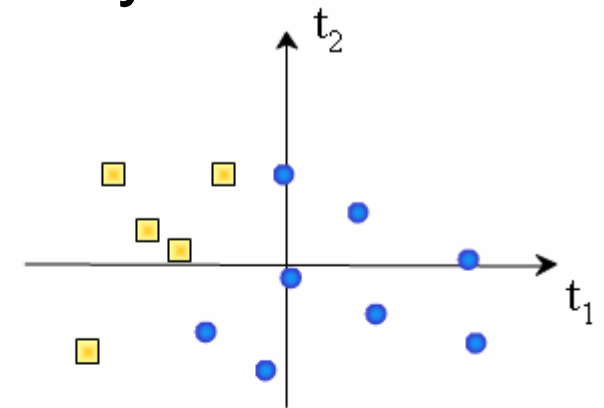
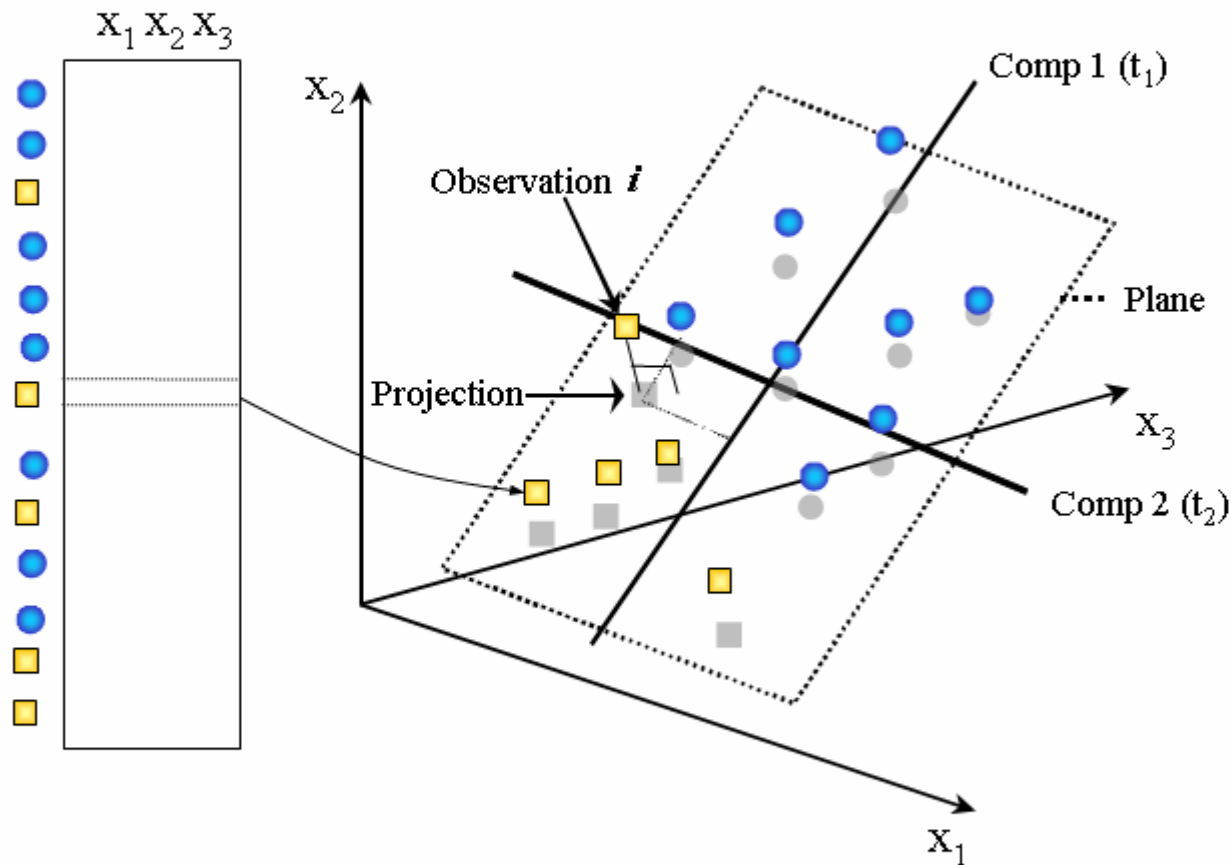
** Krishnan, E., Joint Bone Spine. 70: 496-502 (2003)

Isotope Labeled Internal Standards

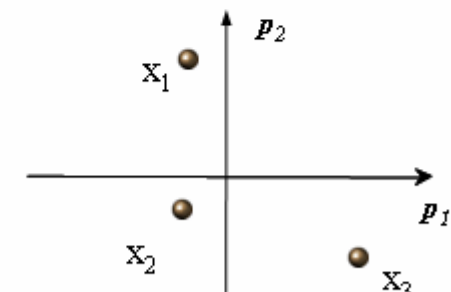
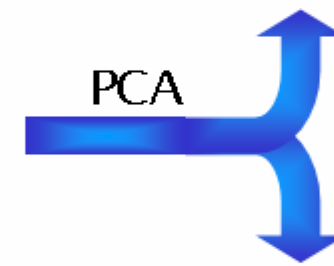
L-Proline-13C5-TMS (221 sec)	Win04_C02
Succinic Acid-D4-2TMS (225 sec)	Win04_C07
Salicylic Acid-D6-2TMS (265 sec)	Win08_C01
Di-Na-alfa-Ketogluarat-13C4-2TMS (277 sec)	Win09_C03
L-Glumatic Acid-13C5-15N (285 sec)	Win10_C02
Putrescine-D4-4TMS (305 sec)	Win11_C07
Myristic Acid-13C3-TMS (324 sec)	Win13_C03
D-Glucose-13C6-5TMS (331 sec)	Win14_C01
Hexadecanoic Acid-13C4-TMS (354 sec)	Win16_C02
Methyl stearate (365 sec)	Win17_C03
Sucrose-13C12-8TMS (430 sec)	Win24_C01
Cholesterol-D7-TMS (483 sec)	Win29_C02

Principal Component Analysis

PCA



Scores (observations)

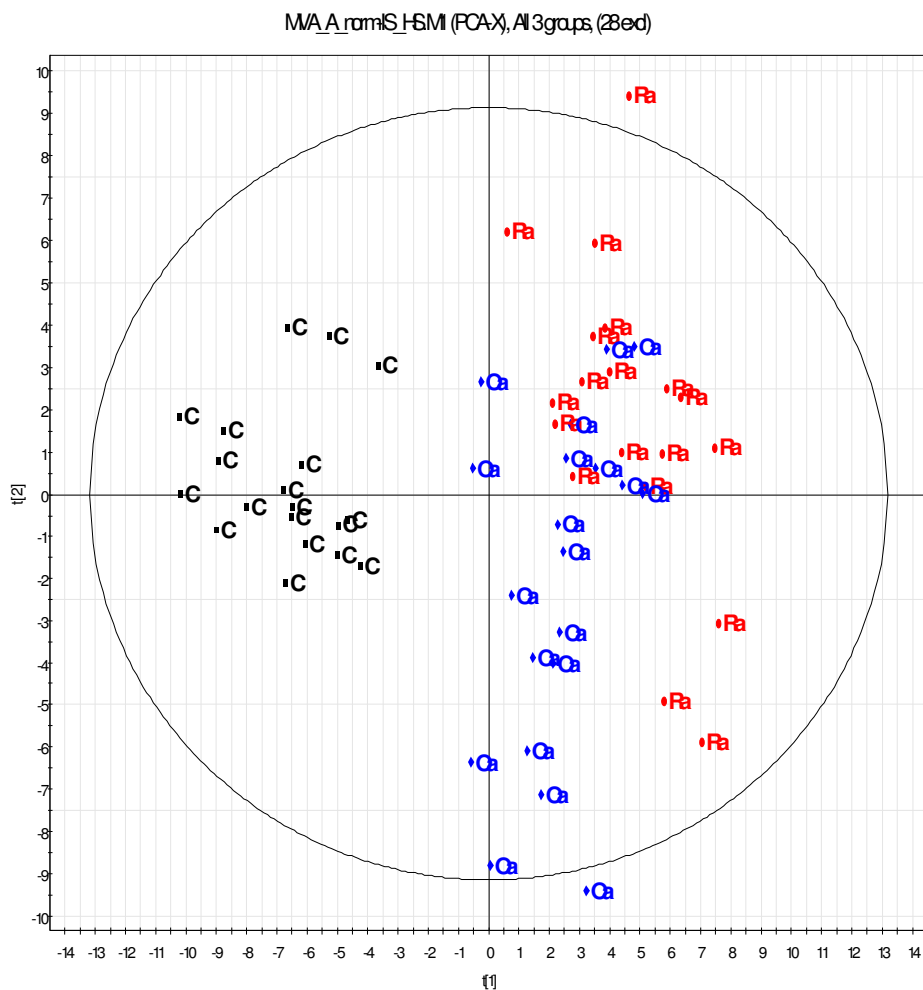


Loadings (variables)

Data Overview

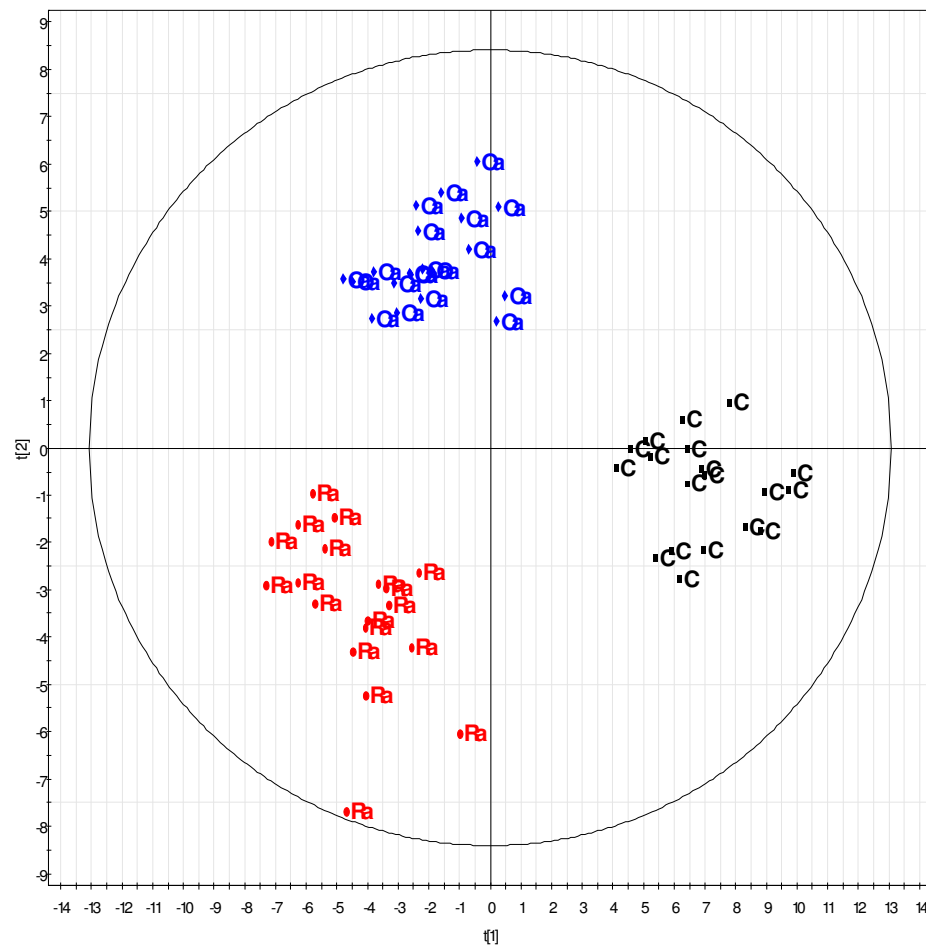
PCA – 181 variables

PLS-DA – 181 variables



R2X[1] = 0,149112 R2X[2] = 0,0715346
Ellipse: Hotelling T2 (0,95)

SMCAP11-20160410 13:41:52

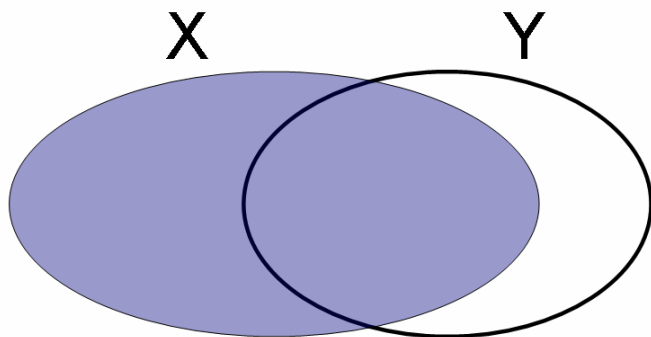


R2X[1] = 0,148724 R2X[2] = 0,0635685
Ellipse: Hotelling T2 (0,95)

SMCAP11-20160410 13:42:46

The OPLS method

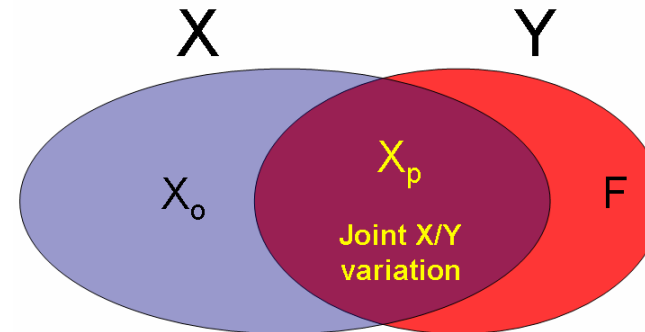
PLS, MLR, PCR, RR etc...



Mixes Y-orthogonal and Y-predictive variation
Uni-directional, Models Y from X

OPLS

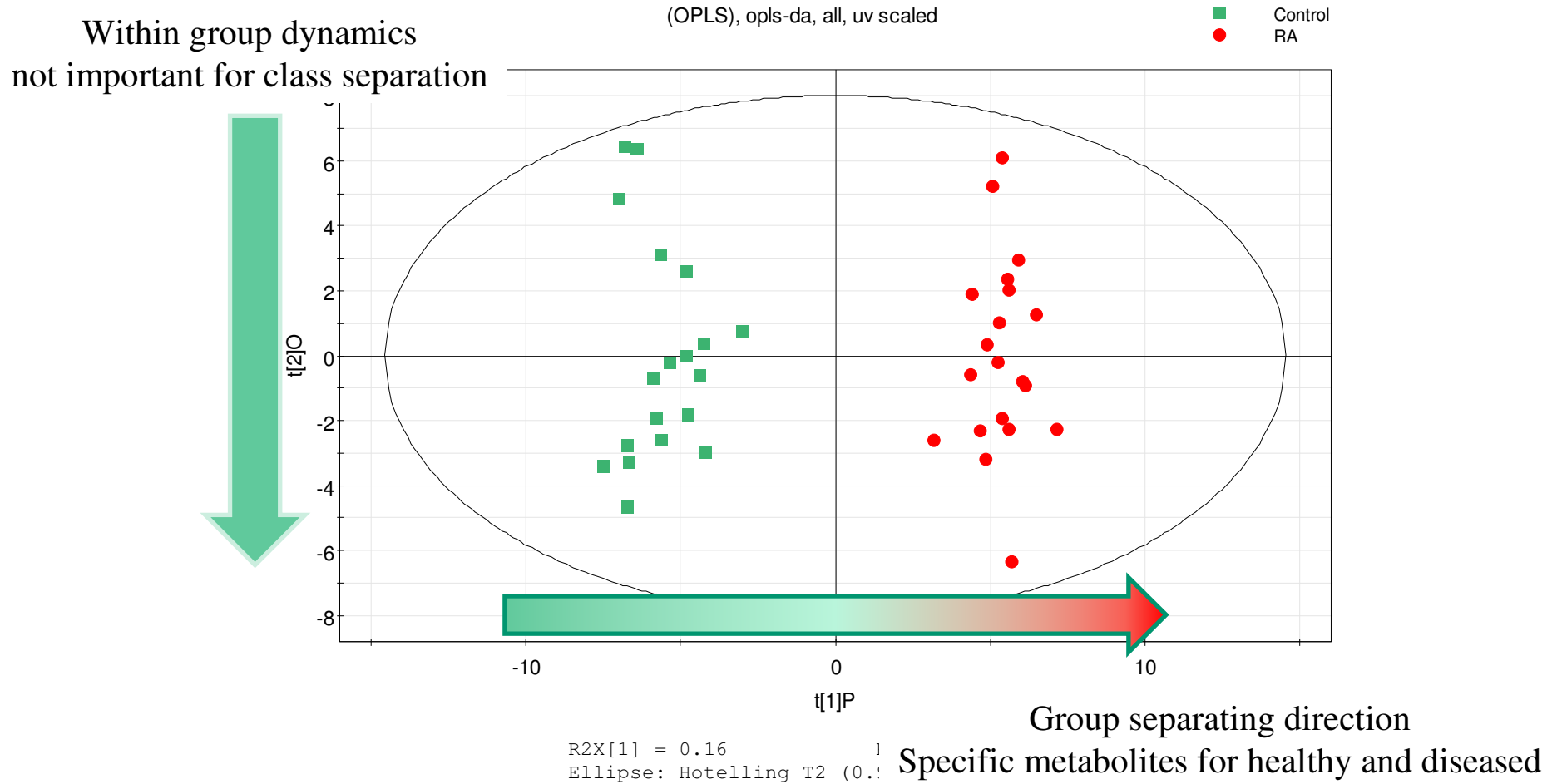
Orthogonal Projections of Latent Structures



Models X AND Y
Separates Orthogonal and Predictive variation
-e.g. 'between block' from 'within block'
Bi-directional
Only uses predictive variation for modeling Y

Rheumatoid arthritis: Control vs. RA

OPLS-DA* - 204 putative biomarkers



*Bylesjö, M.; Rantalainen, M.; Cloarec, O.; Nicholson, J. K.; Holmes, E.; Trygg, J.,
OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification.

RA: Comparison of the human case and animal models

- Great overlap of metabolites between humans and animals
 - ❑ Different metabolites show overlap in different animal models
 - ❑ Allows for identification of relevant animal models
 - ❑ Selection of model system for treatment studies

EM	Human Rheumatoid Arthritis	Mouse Collagen Induced Arthritis	Rat Adjuvant Induced Arthritis
EC001	↑	na	Na
EC002	↑	?	?
EC003	↑	↓	↓
EC004	↑	0/↓	↓
EC005	↓	na	na
EC006	↓	↓	↓
EC007	↓	↓	↓
EC008	↓	↓	↑
EC009	↓	↓	↓
EC010	↓	↑	↑
EC011	↓	0/↓	↓
EC012	↓	na	na
EC013	↓	↓	↓
EC014	↓	↓	?
EC015	↓	↓	↓
EC016	↓	?	↓
EC017	0	↓	↓
EC018	↑	↑	↓ / ?
EC019	↓	↓	↓
EC020	↓	↓ / ?	↓
EC021	?	↑ / ?	↑
EC022	↓	↓	↓
EC023	0	↓	↓
EC024	↑	↓	0/↑
EC025	↑	↓	↓
EC026	0/↑	↓	↑

RA: Comparison of therapies in animal model

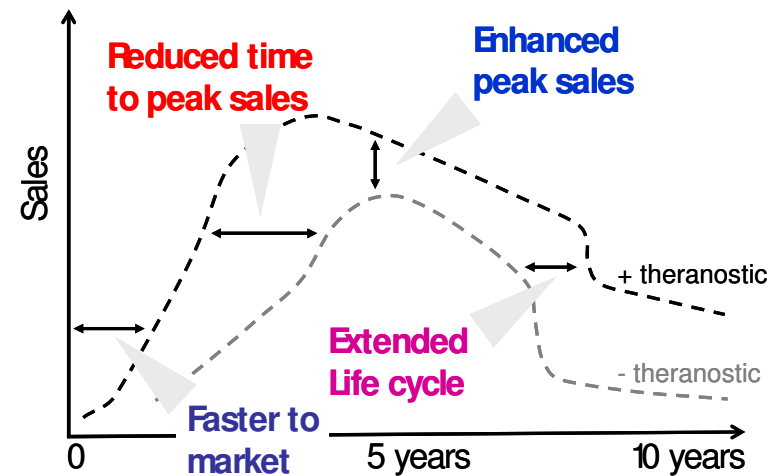
- Metabolites levels are affected by administered therapeutics
 - ❑ New drug (X) restore levels in more metabolites compared to MTX*
 - ❑ Useful in development of novel drugs
 - ❑ Tool in clinical studies to verify therapeutic effect in clinical studies
 - ❑ Concomitant development of novel drug and diagnostic test, theranostics?

	Vehicle	MTX	X 1mg	X 3mg	X 10mg
EC004	0/↑	↓	↓	0/↓	↓
EC006	0/↑ / ?	0/ ?	0	↑	↑
EC007	↓	0/↑	0/↑	0/↓	↑
EC009	0	↑	↓	↑	↑
EC010	↑	↑	↑	↑	↑
EC011	0	0/↓	↓	0/↓	↑
EC012	0/↓	↑	0/↓	↑	↑
EC013	↑	0/↑	0/↓	↑	0/↑
EC014	↑	0/ ?	↑	↑	↑
EC015	0/↑	↑	0/↓	↑	↑
EC016	0	↓	↑	↑	↓
EC017	↓	↓	↓	↓	↓
EC018	↓	↓	↓	0/↑	0/↑
EC019	↓	↓	↓	↓	↓
EC022	↑	↑	0/↑	↑	↑
EC023	↓	0/↓	↓	↓	↓
EC024	↓	↓	↓	↓	↓
EC025	↓	↓	↓	↓	↓
EC026	↑	↑	↑	↑	↑

*MTX, methotrexate

Theranostics

- Theranostic approaches aim to target **therapies** to those patients most likely to benefit by combining with **diagnostic** or **prognostic** biomarkers
- Greatest benefit is realised with therapies that are effective, but in unknown patient sub-populations – HA therapy in OA is an excellent example
- An effective theranostic can drive sales of therapies as illustrated
- Regulation can be achieved within current pivotal clinical trial protocols
- Potential for the approach to track therapeutic efficacy



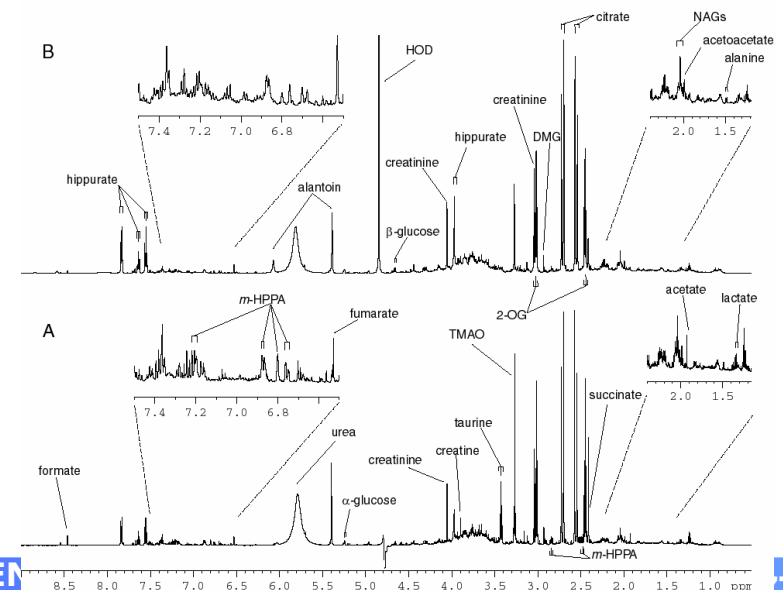
Adapted from Gilham (2002) Theranostics: an emerging tool in drug discovery and commercialisation. *Drug Discovery World*

Benefits

- Identification of responder/non responders populations
- Targeted patient recruitment to pivotal clinical trials, resulting in shorter, lower risk studies
- Stronger claims through mechanism of action knowledge and potential for data on therapeutic efficacy
- Improved HE evidence and addressing payer “entitlement” concern will increase payer uptake
- Much higher adoption in smaller market size – 50% penetration of 50% market size is better than 1-5% of the full market

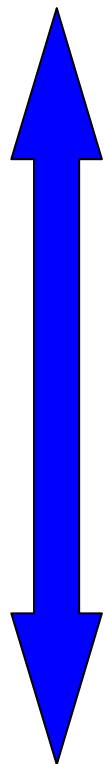
Case study 2: Functional foods study

- NMR spectroscopy to detect metabolism changes due to food supplement
- N = 47 biofluid samples
 - ❑ 9 healthy individuals
 - ❑ Given prepared foodstuff for consumption
 - ❑ Multiple visits – document effect over time
- K = 32 768 variables
 - ❑ NMR shifts
- Goal: Effect of food supplement?

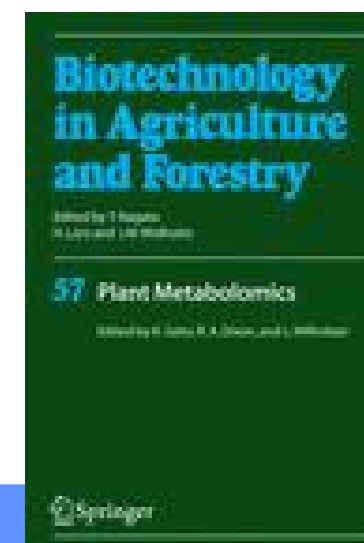
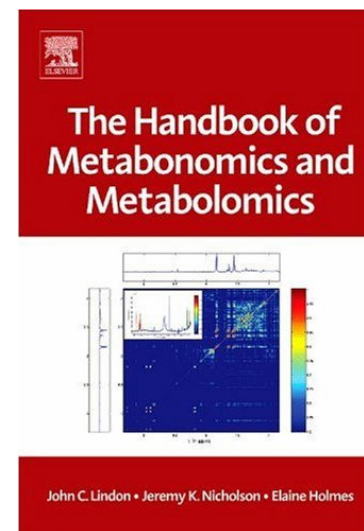


Chemometrics

- the information aspect for studying complex systems



1. Define the aim
 - What do we want?
2. Selection of objects (e.g. samples, time points, internode, experiments)
 - Design of Experiments (DOE)
 - Multivariate design (MVD)
3. Sample preparation and characterisation
 - Experimental protocol (e.g. FTIR, GCMS, Microarray)
 - Data processing (e.g. normalisation of FTIR, microarrays)
4. Evaluation/Validation of collected data
 - Exploratory analysis, Interpretation & Visualization
 - Dynamic study



2008-06-11

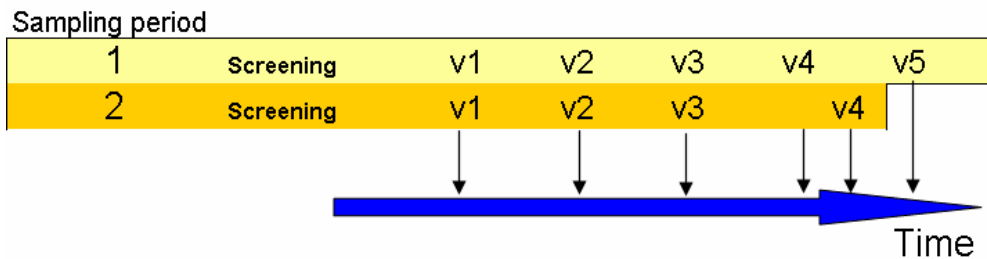
Johan Trygg, Research Group for
Chemometrics, Umeå University

Example:

Functional foods study

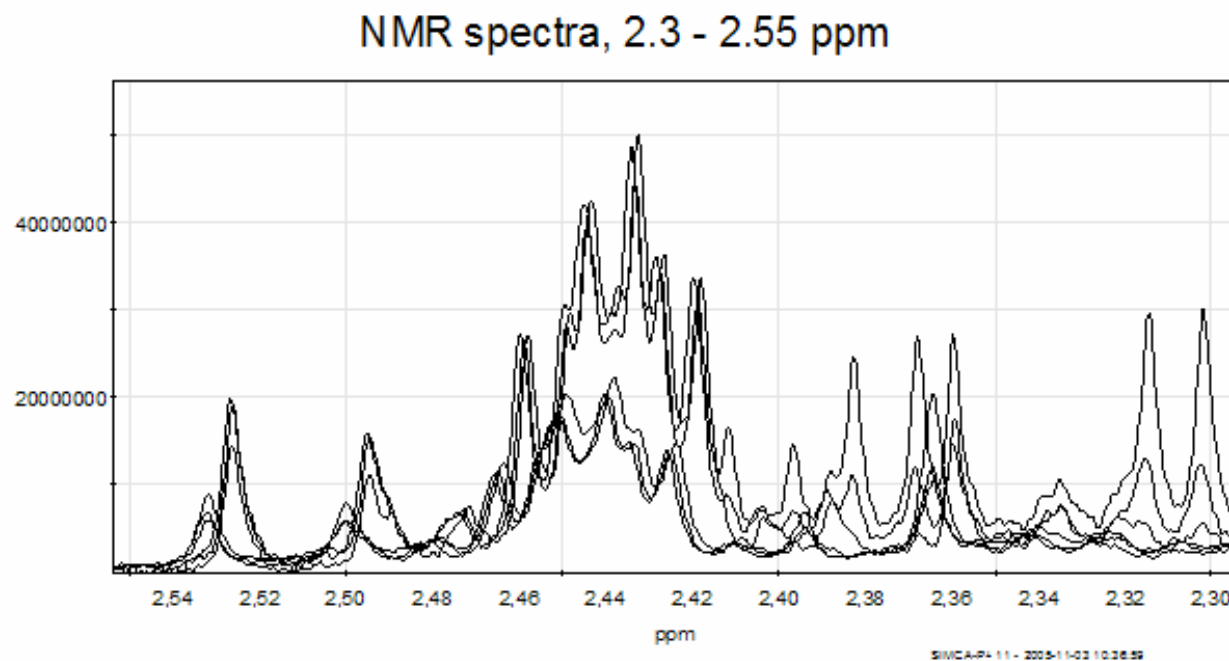
- Functional foods: Foodstuffs with a documented health-promoting effect – besides energy addition
- Centre for Human Studies of Foodstuffs, Sweden
 - ❑ Inclusion/exclusion criteria
 - ❑ 9 individuals given prepared food
 - ❑ Multiple visits – document effect

Study design



2008-06-11

Case study 2: Overview of all individuals/samples



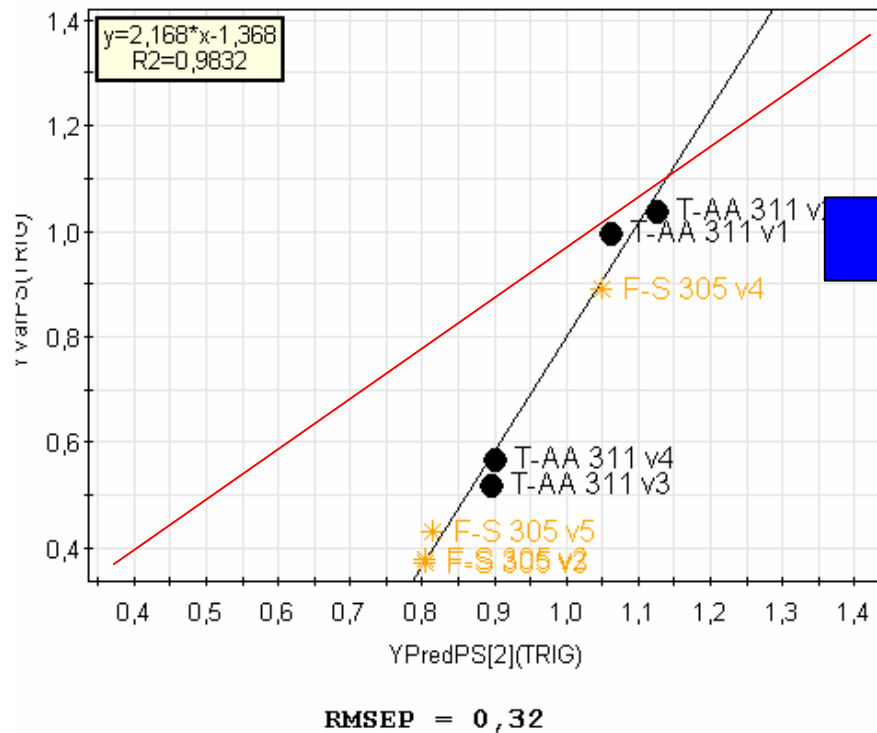
Clear separation.

But ☹ ... due to different sampling periods

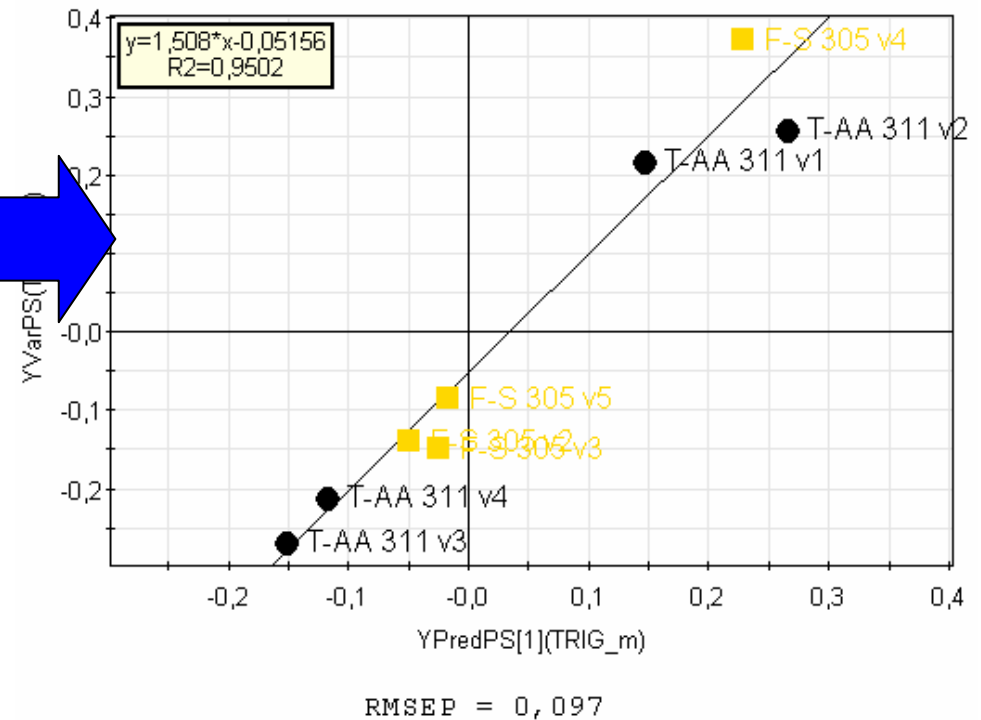
Prediction of triglycerides

Raw NMR data vs Metabolic effect modelling

Prediction of triglycerides
Raw NMR data (major bias)



Prediction of triglycerides
Metabolic effect modelling

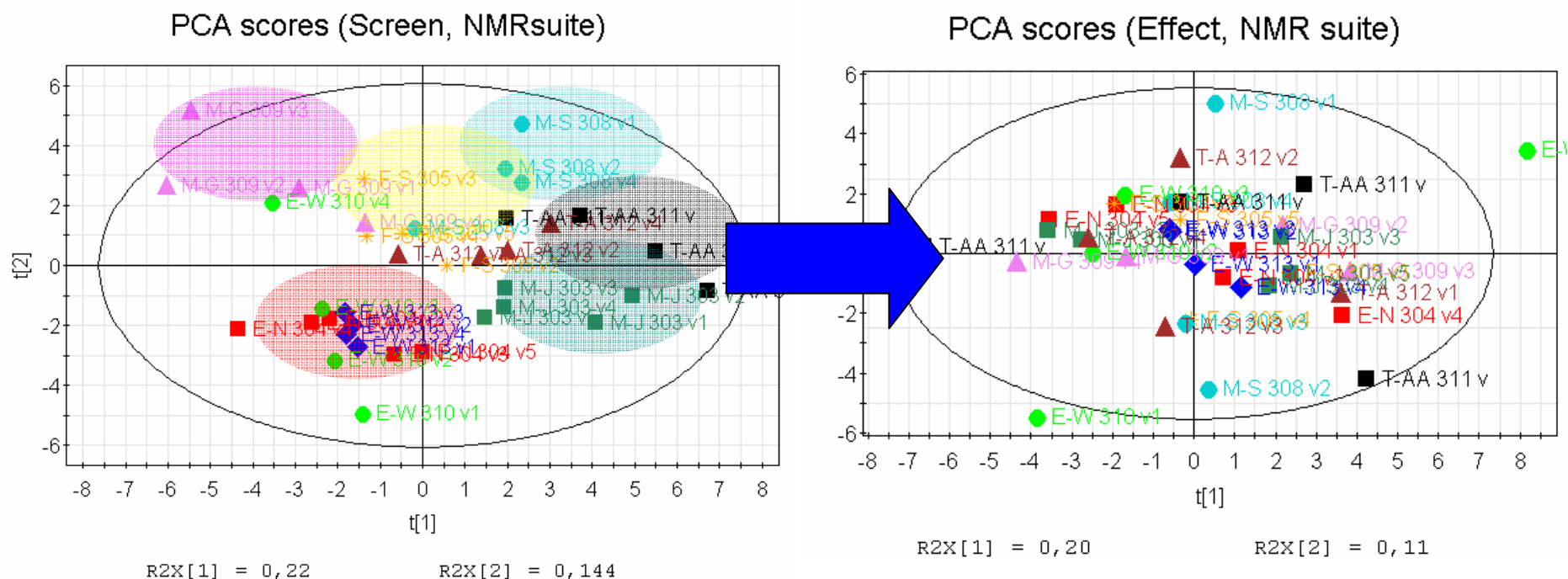


2008-06-11

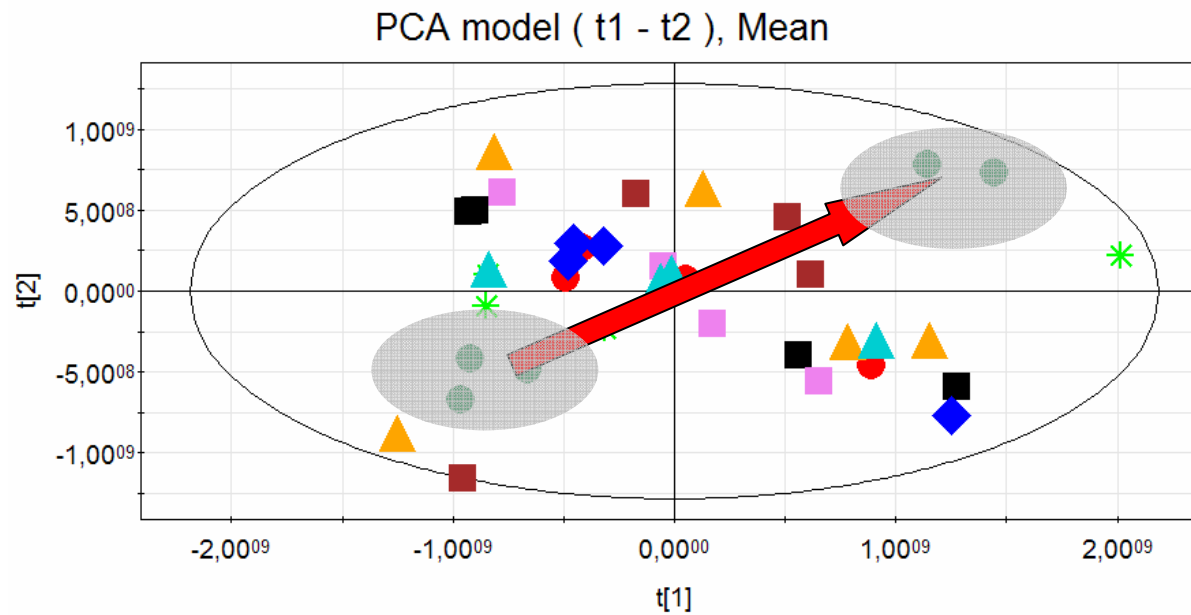
61

Case study 2: Modelling dynamic metabolomic time series data

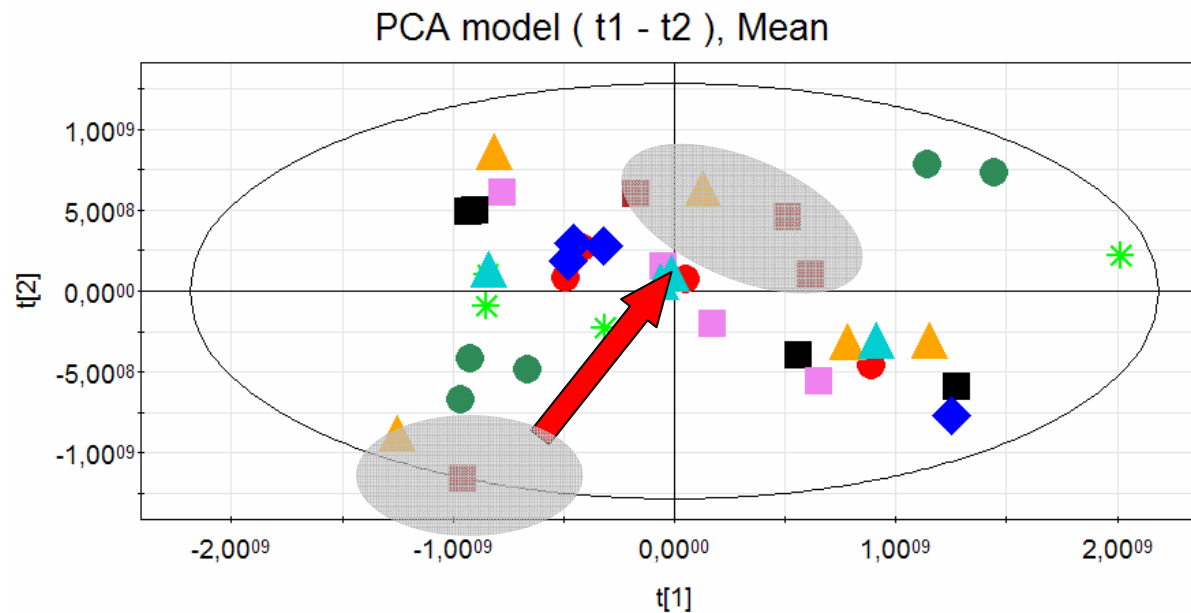
- Assumption 1: The metabolic baseline can be different over all individuals
- Assumption 2: The metabolic effect of the treatment is similar over all individuals



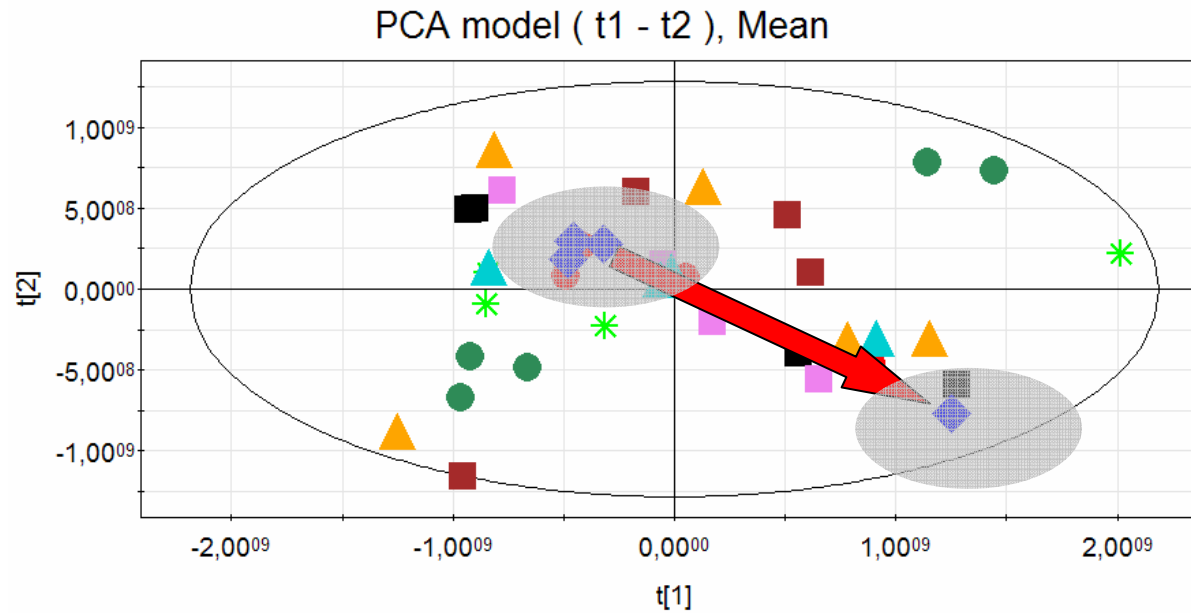
Case study 2: Overview of all individuals/samples ... after pre-processing/filtering



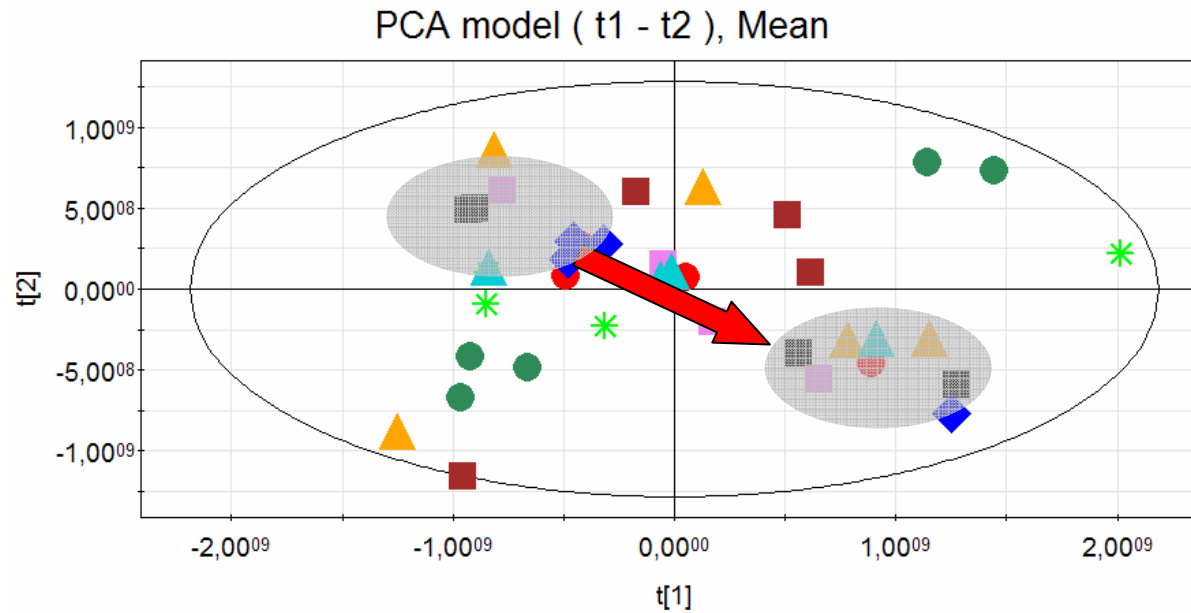
Case study 2: Overview of all individuals/samples ... after pre-processing/filtering



Case study 2: Overview of all individuals/samples ... after pre-processing/filtering

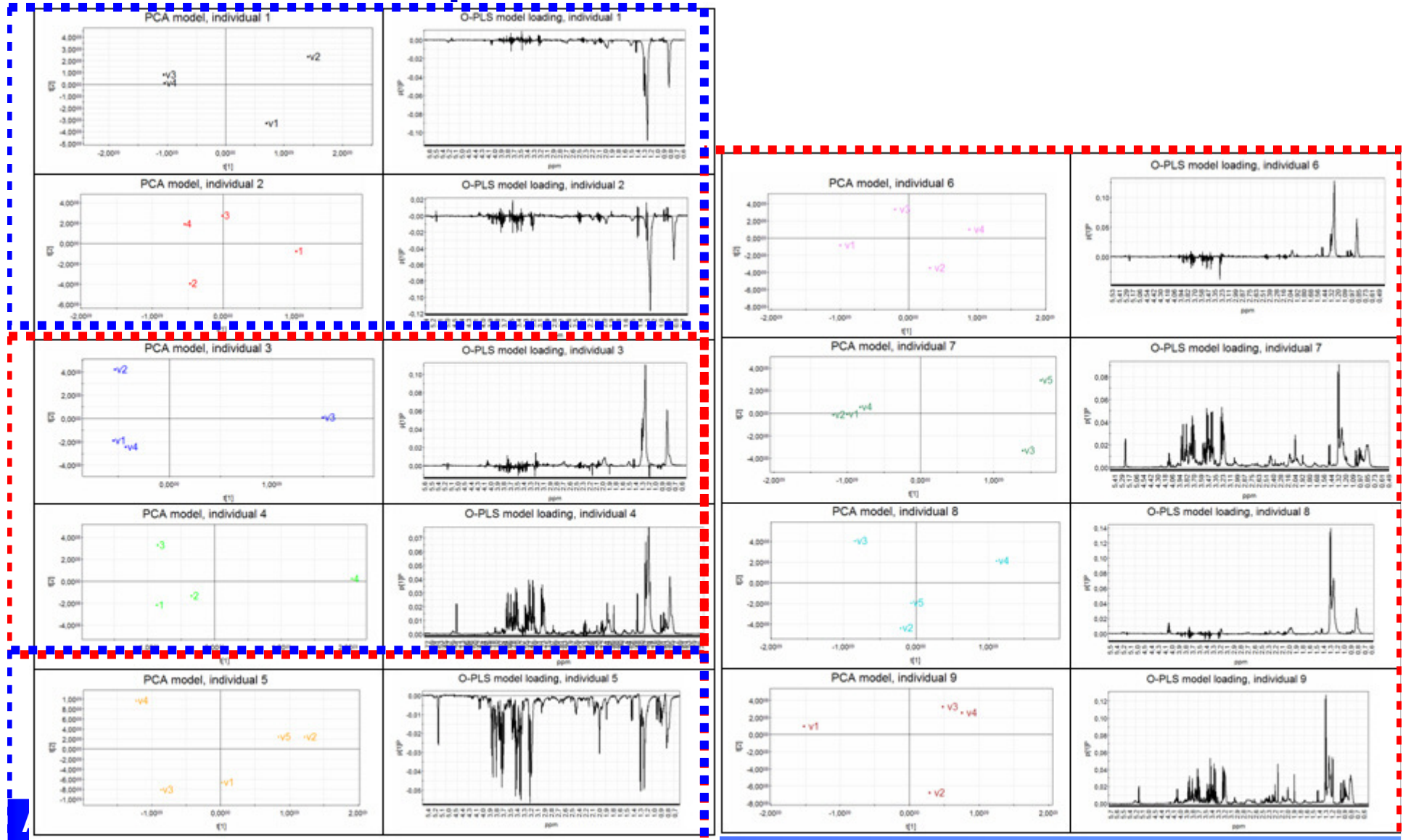


Case study 2: Overview of all individuals/samples ... after pre-processing/filtering



Case study 2: Metabolic profiles for each individual

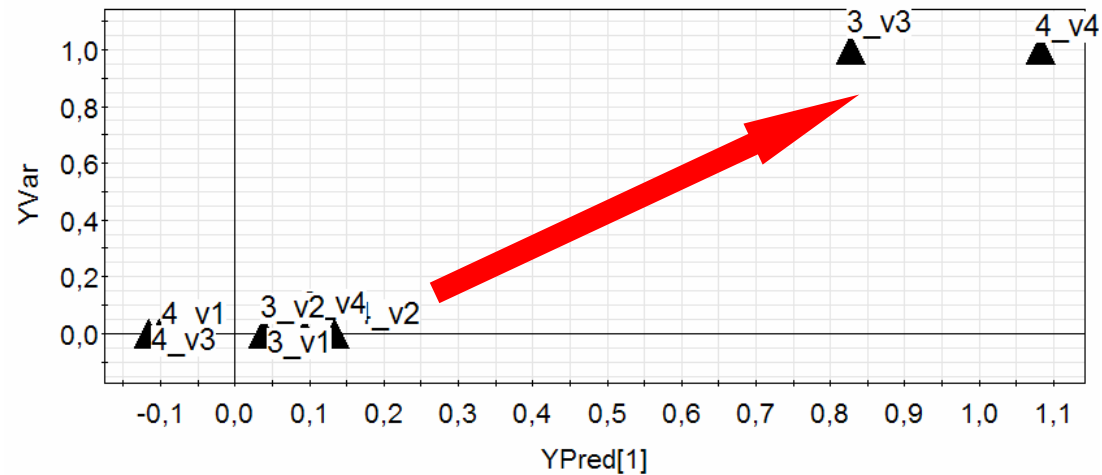
Assumption: the effect of foodstuff treatment



Case study 2: Prediction of effect in NMR profile

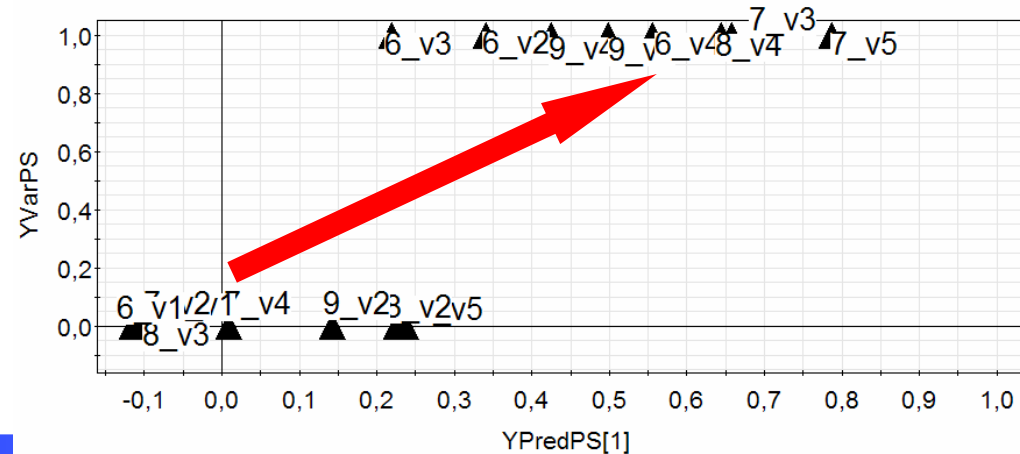
Model

Observed vs Predicted (model)
Individuals 3 and 4



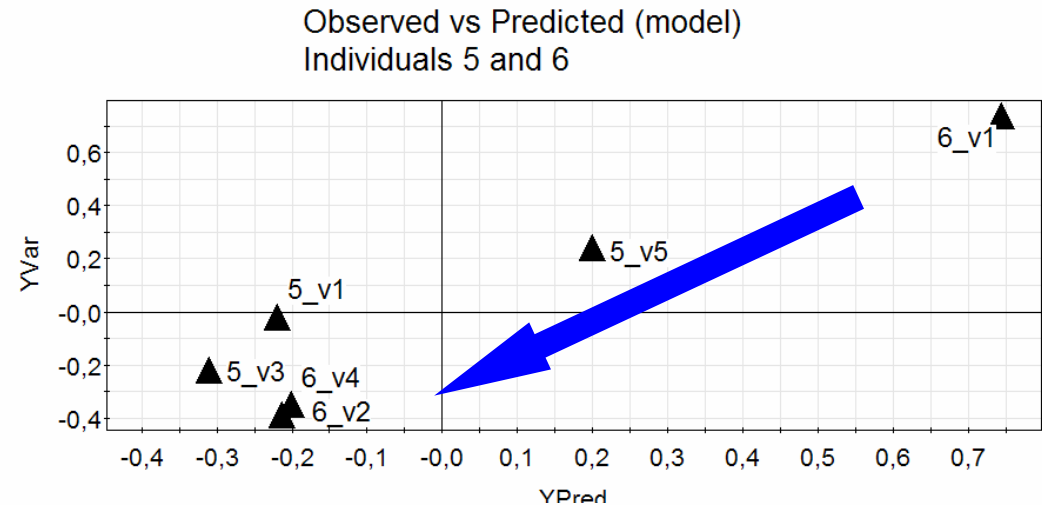
Prediction

Observed vs Predicted (Prediction set)
Individuals 6,7,8,9

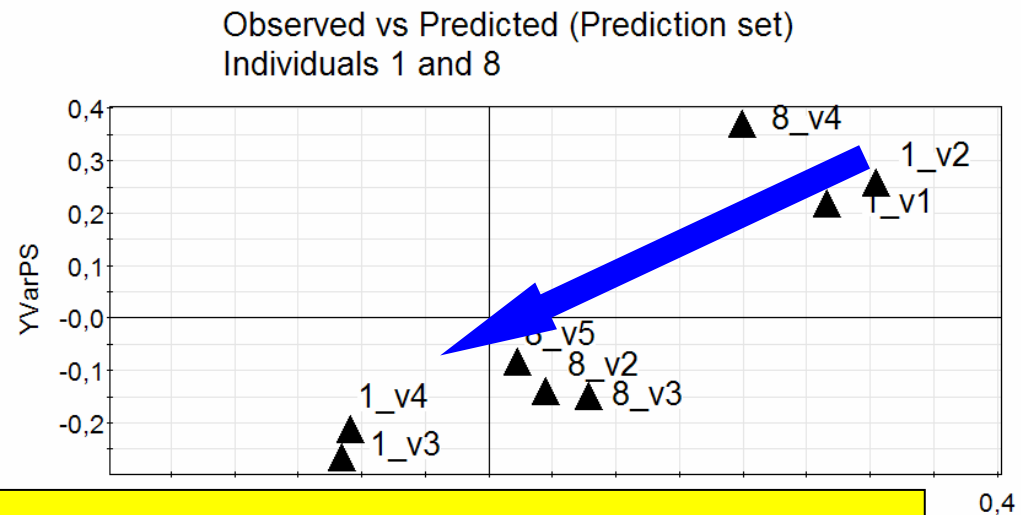


Case study 2: Prediction of health effect in endogenous metabolites ("biomarkers")

Model



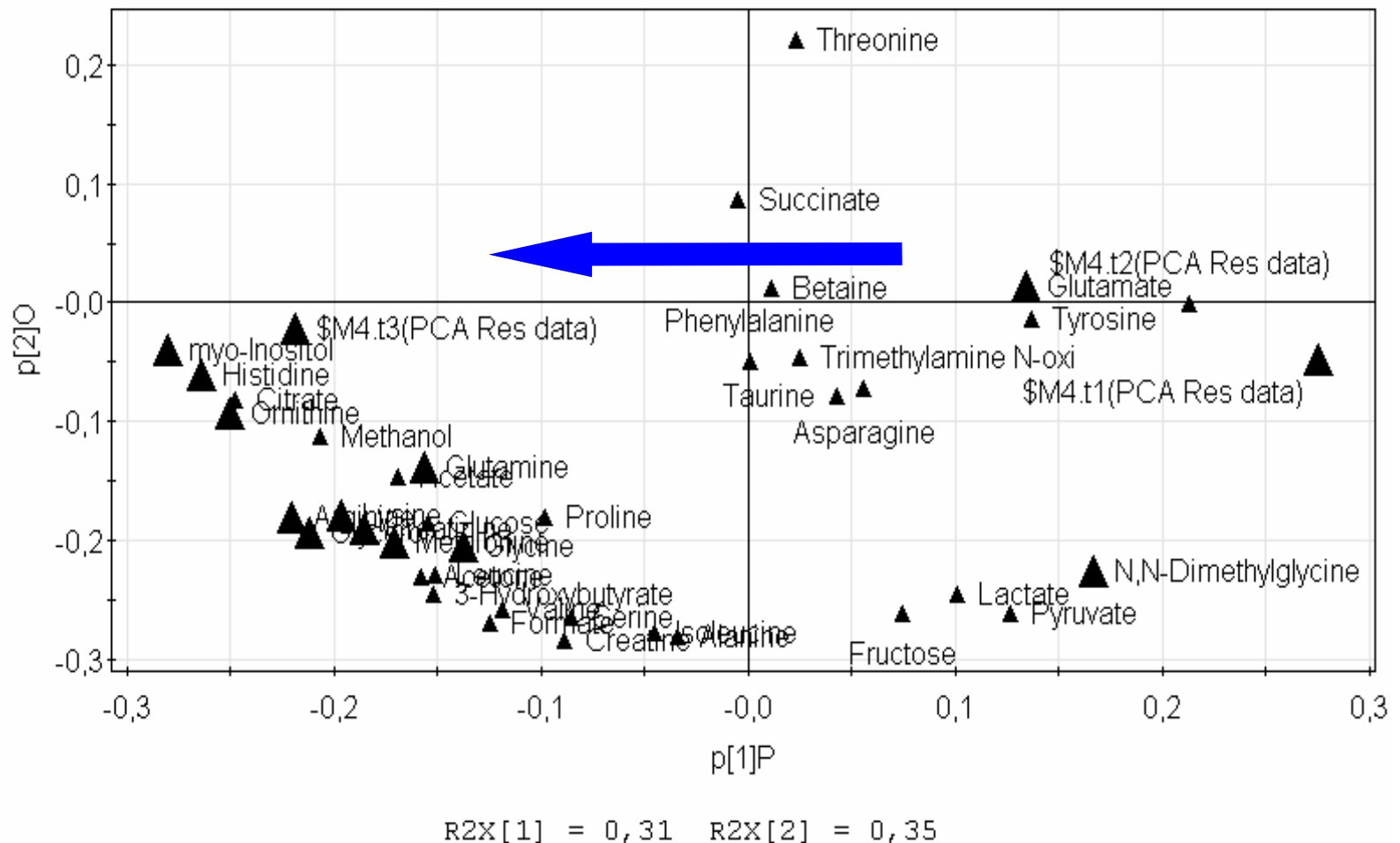
Prediction



Health effect shown, as metabolite concentration *decreases* after consumption of foodstuff

Case study 2: Analysis of health effect in endogenous metabolites "biomarkers")

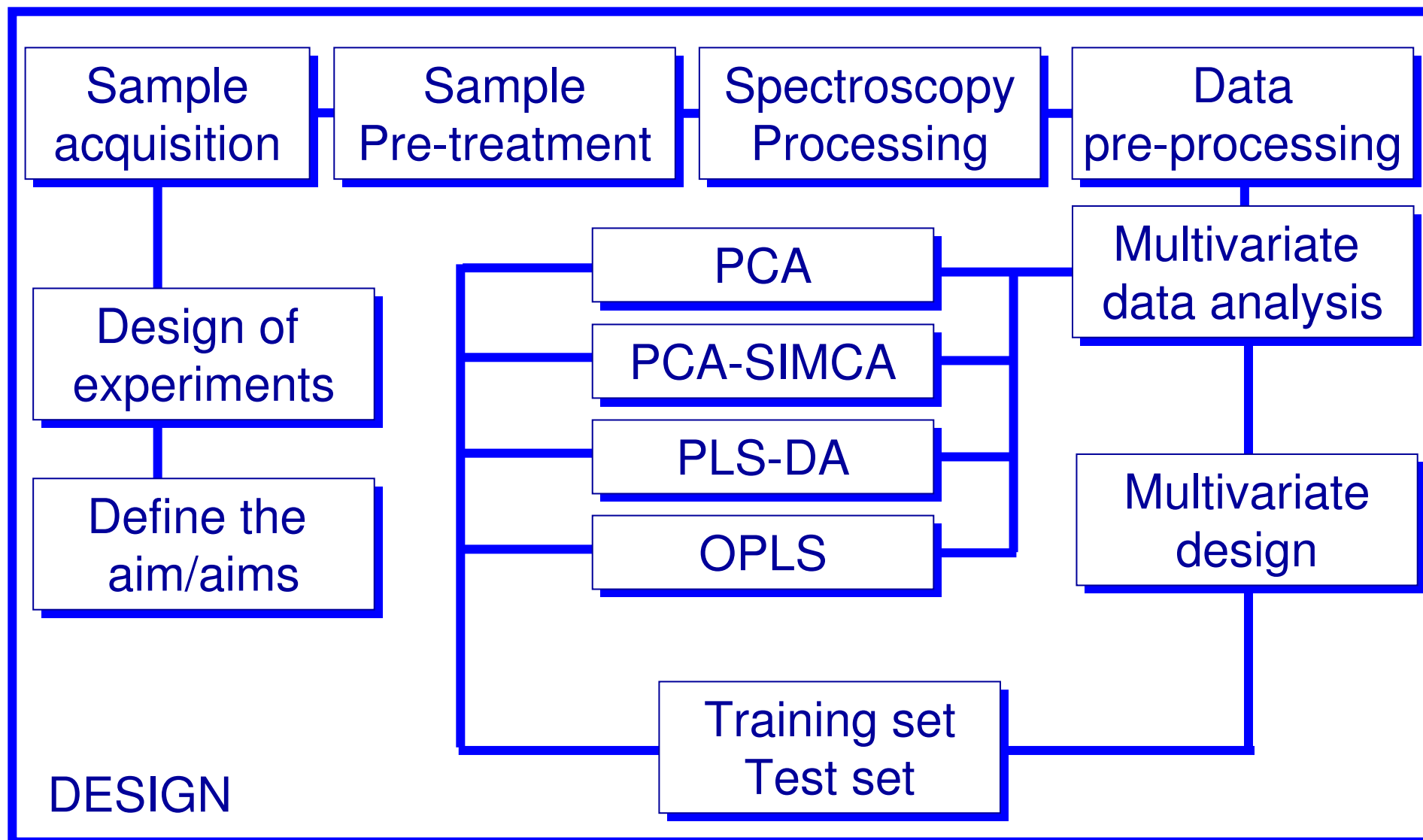
OPLS loading plot



Case study 2: Conclusions

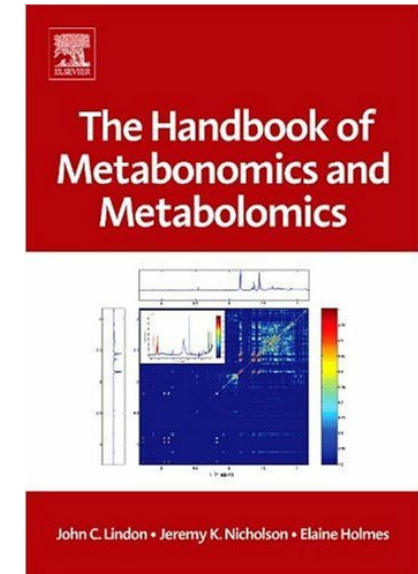
- Effect of food supplement established
 - ❑ Increase in Myo-Inositol
 - ❑ Decrease in triglycerides
- Opens new possibilities for development of functional foods
 - ❑ Proof of physiological health effect
 - ❑ Target identification

General analysis strategy



Strategy

1. Formalize the aim
 - What do we want?
2. Selection of objects
 - Design of Experiments (DOE)
 - Multivariate design (MVD)
3. Sample preparation and profiling of human and animal samples
 - *In vivo*, *in vitro* samples
 - Blood, Urine, Cerebral Spinal Fluid (CSF)
 - Synovial fluid (joint), Bowel fluids, Feces, Tissues
4. Integration and evaluation of collected data
 - Exploratory analysis, Interpretation & Visualization
 - Prediction models
 - Patterns
 - Target identification
5. Identify and define IPR opportunities and strategies



Acure omics

Projects in disease diagnostics

Osteoarthritis (patent application filed)

Rheumatoid arthritis (patent application filed)

Fibromyalgia

Chronic Fatigue Syndrome

Type 1 Diabetes (patent application filed)

Services CRO

Biofluid profiling

Biomarker identification

Toxicity monitoring

Summary

- Historical data
 - ❑ Always a good starting point for analysis (PCA)
 - ❑ Determine data structure, preferred format/output
 - ❑ Get acquainted with the process and the data
 - ❑ Find “hidden” information
 - ❑ Good starting point for discussions
- Determine
 - ❑ The aim – is there a defined stop criteria (yield, purity, accepted batches, ID important/”sensitive” variables etc.)
 - ❑ Important to define prior to investigation
 - ➔ know when to stop
 - ❑ The experimental domain (variables, settings, responses)

Summary

- Design of Experiments (DoE)
 - ❑ Simplify analysis
 - ❑ Ensure a systematic variation in the investigated experimental domain
 - ❑ Small design within the defined limits
 - ❑ (Design in historical data)
- Analysis
 - ❑ PCA (SIMCA etc.)
 - ❑ PLS, PLS-DA
- Next step
 - ❑ Aim(s) reached
 - ❑ Important variables
 - ❑ New optimal experiments

Acknowledgements

- Per Lek, AcurePharma
- Thomas Moritz, UPSC
- Johan Trygg, Umeå University
 - Rasmus Madsen, Umeå University
- Elisabeth Seifert, AcurePharma
- Jon Gabrielsson, AcureOmics
- Johan Olsson, Uppsala University
- Mattias Hedenström, Umeå University
- Katrin Lundstedt-Enkel, Uppsala University