Analysis of metabolomics data – examples from the study of the lipid metabolism

Olav M. Kvalheim

Dept. of Chemistry, University of Bergen

Faculty of Health Studies, University College of Sogn and Fjordane

Content of this lecture

- Analytical methods
- Early work on cancer diagnostics
- Lipoproteins, fatty acids & cardiovascular health
- Aging and gender effects

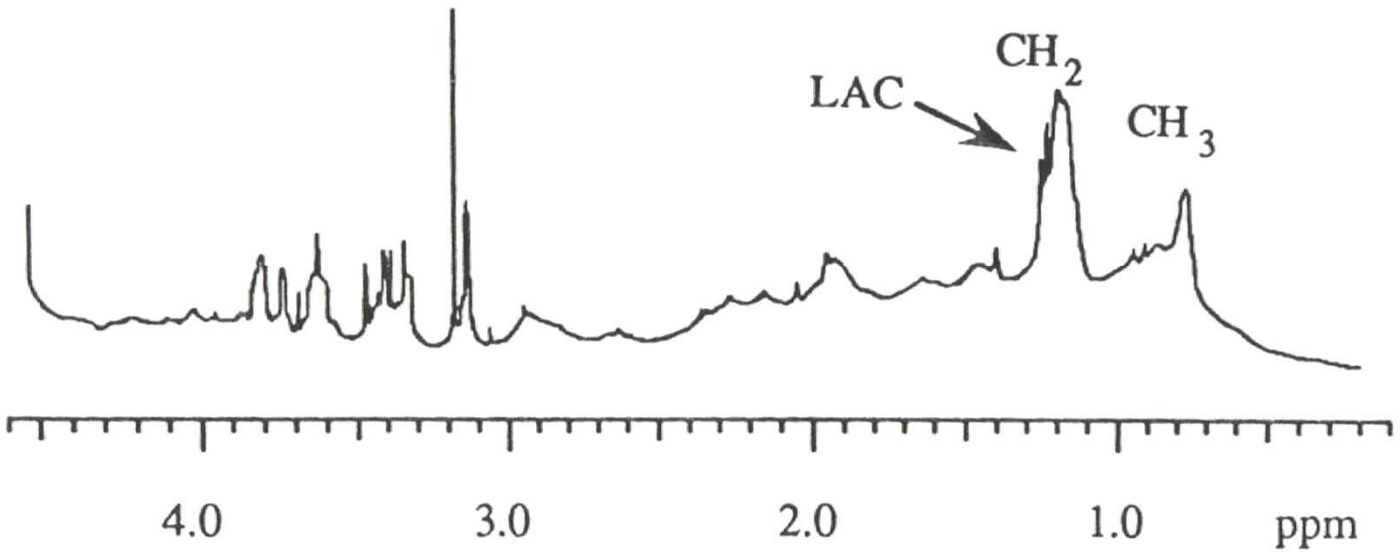
Analytical methods

- Proton NMR
- HPLC
- Multivariate Data Analysis

Metabolomics diagnosis of cancer - 1987-95

1. Proton NMR profiling of serum
2. Focused on lipoproteins, loss of weight => cancer disturbs lipid metabolism?
3. Whole digital peaks as input to PCA/SIMCA
4. Classification of clinical diagnosed cancer patients using model of healthy controls

Cancer project - serum proton spectrum (400 MHz)



Sletten E., Kvalheim O.M., Kruse S, Farstad M. and Søreide O. (1990), Detection of malignant tumours - Multivariate analysis of proton NMR spectra of serum. *European J. of Cancer* 26, 615-618.

Cancer diagnosis from serum proton NMR

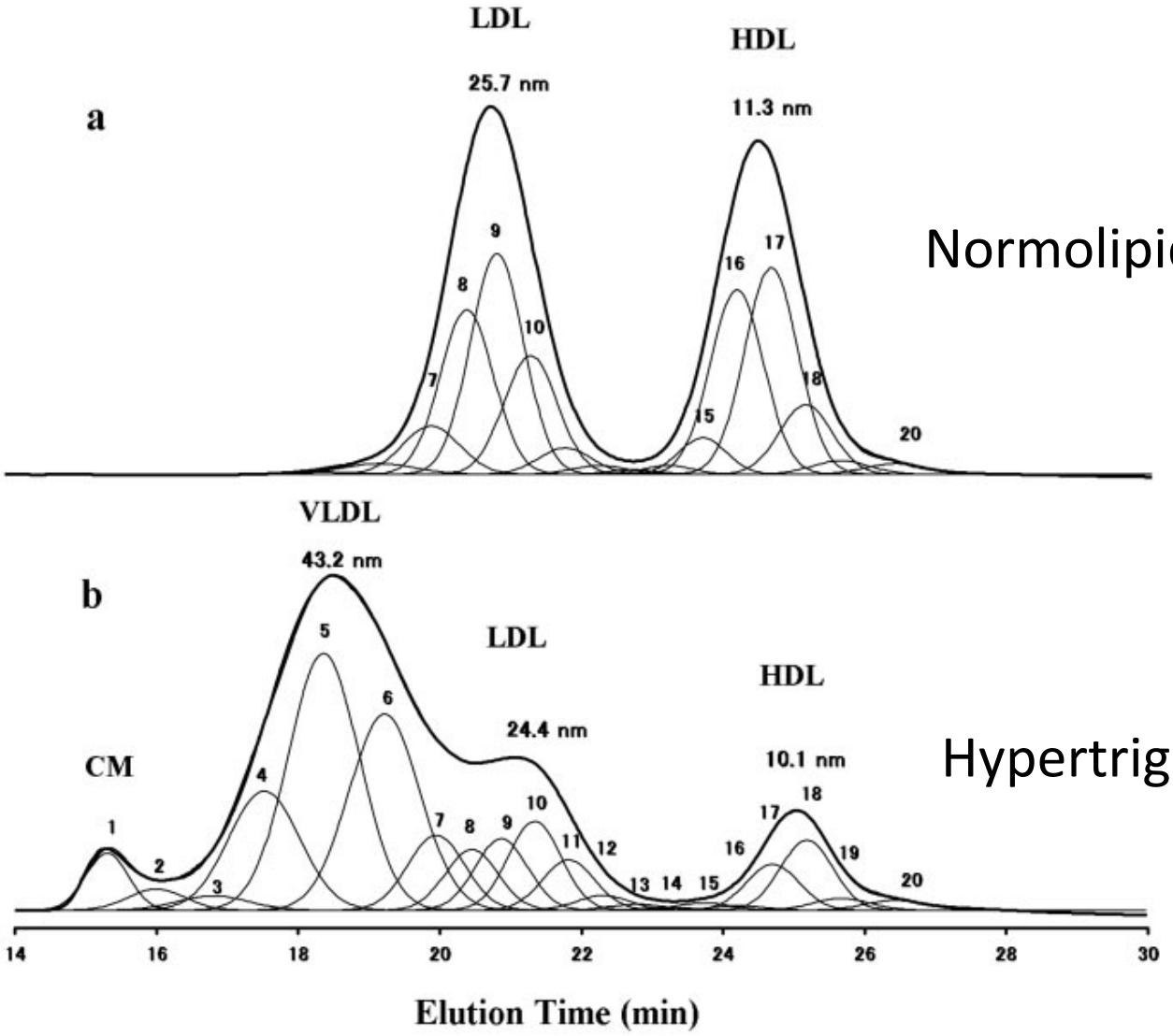
Fluge Ø., Gilje K.S., Sletten E., Kvalheim O.M., Skarland E., Halvorsen J.F., Farstad M. and Søreide O (1996) Proton nuclear magnetic resonance spectroscopy of serum from patients with colorectal neoplasia, *Eur. J. Surg. Onc.*, 22, 78-83.

=> Not enough information in the proton NMR spectral profiles to provide early diagnosis

But then:

Bro et al. (2015) *Metabolomics*: 80% correct diagnosis of breast cancer two years ahead of clinical diagnosis!

HPLC – 20 lipoprotein subclasses of TG & Cholesterol



Okazaki et al. (2005) *Arterioscler. Thromb. Vasc. Biol*., 25:578-584.

24 lipoprotein features

- Main classes: CM, VLDL, LDL, HDL, TC, TG
- Subclasses: VLDL-VL, VLDL-L, VLDL-M, VLDL-S (IDL), LDL-L, LDL-M, LDL-S, LDL-VS, HDL-VL, HDL-L, HDL-M, HDL-S & HDL-VS
- Average particle size: VLDL-Size, LDL-Size & HDL-Size
- Apolipoproteins: ApoA1 & ApoB

Lin et al. (2016) *Metabolomics*, 12:6
Rajalahti et al. (2016) *Metabolomics*, 12:51

Data analysis

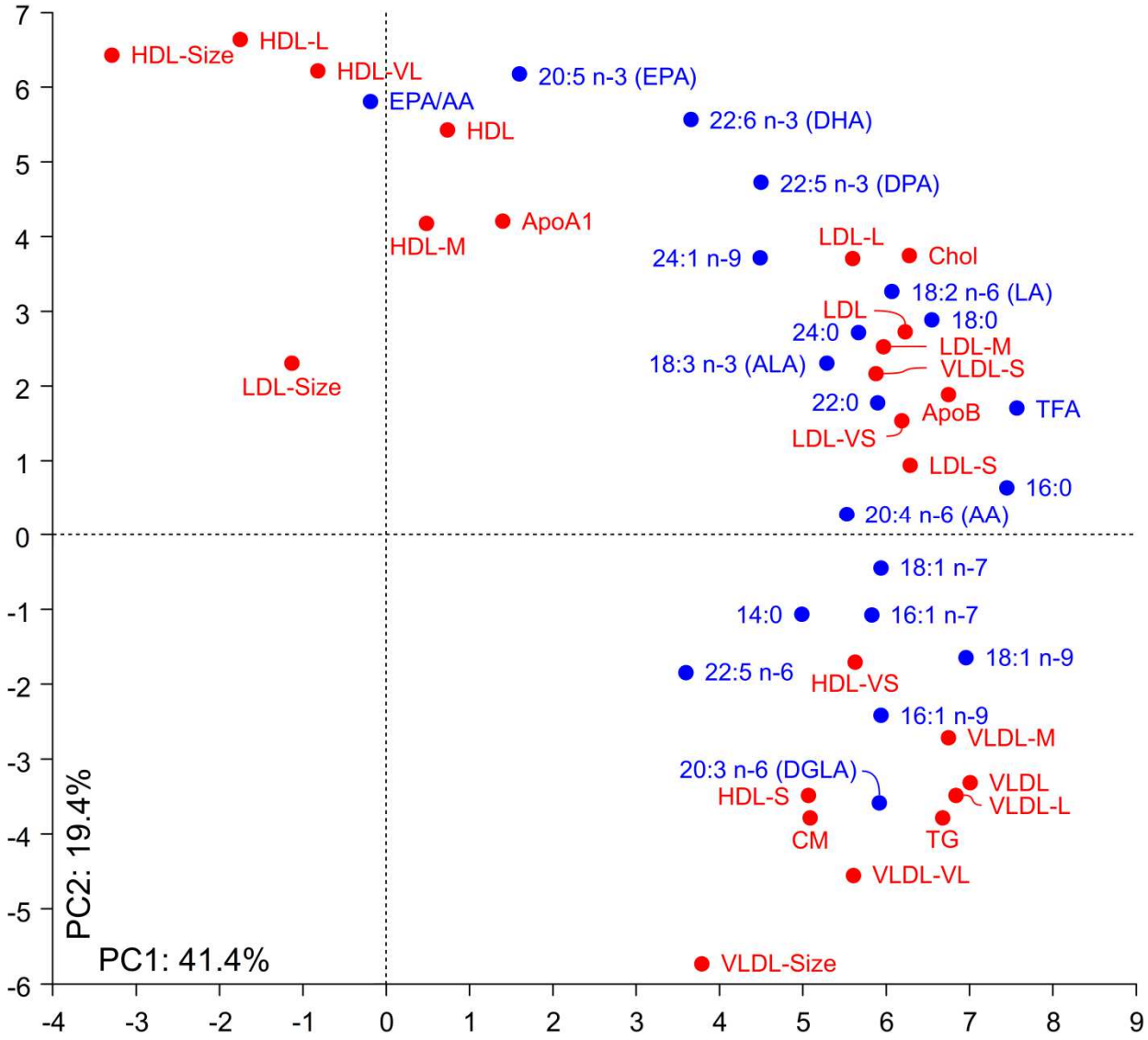
1. PLS/PLS-DA using RDCV
2. Target projection (TP) to single latent variable related to the *y*-information.
3. Selectivity ratio (SR) plots for interpretation.
4. Discrimination Variable (DIVA) plot/Rank sum classification rate to find biomarkers or patterns of biomarkers.

Lin et al. (2016) *Metabolomics*, 12:6
Rajalahti et al. (2016) *Metabolomics*, 12:51
Rajalahti et al. (2016) *Metabolomics*, 12:81.

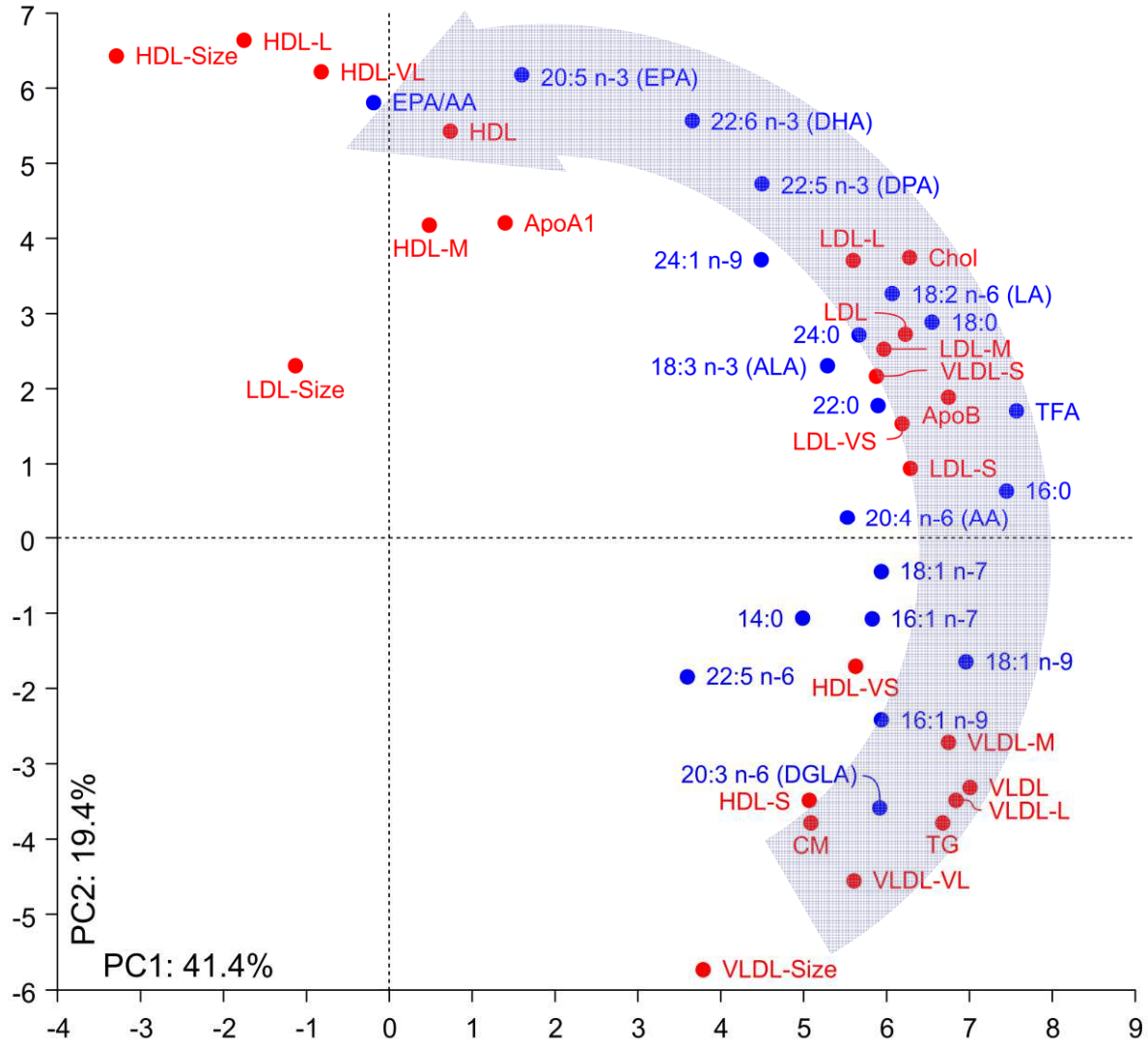
Lipoprotein and fatty acid patterns in healthy adults and implications for cardiovascular health

Lin et al. (2016) *Metabolomics*, 12:6.

Lipoprotein and fatty acid patterns - women



Lipoprotein and fatty acid patterns - women

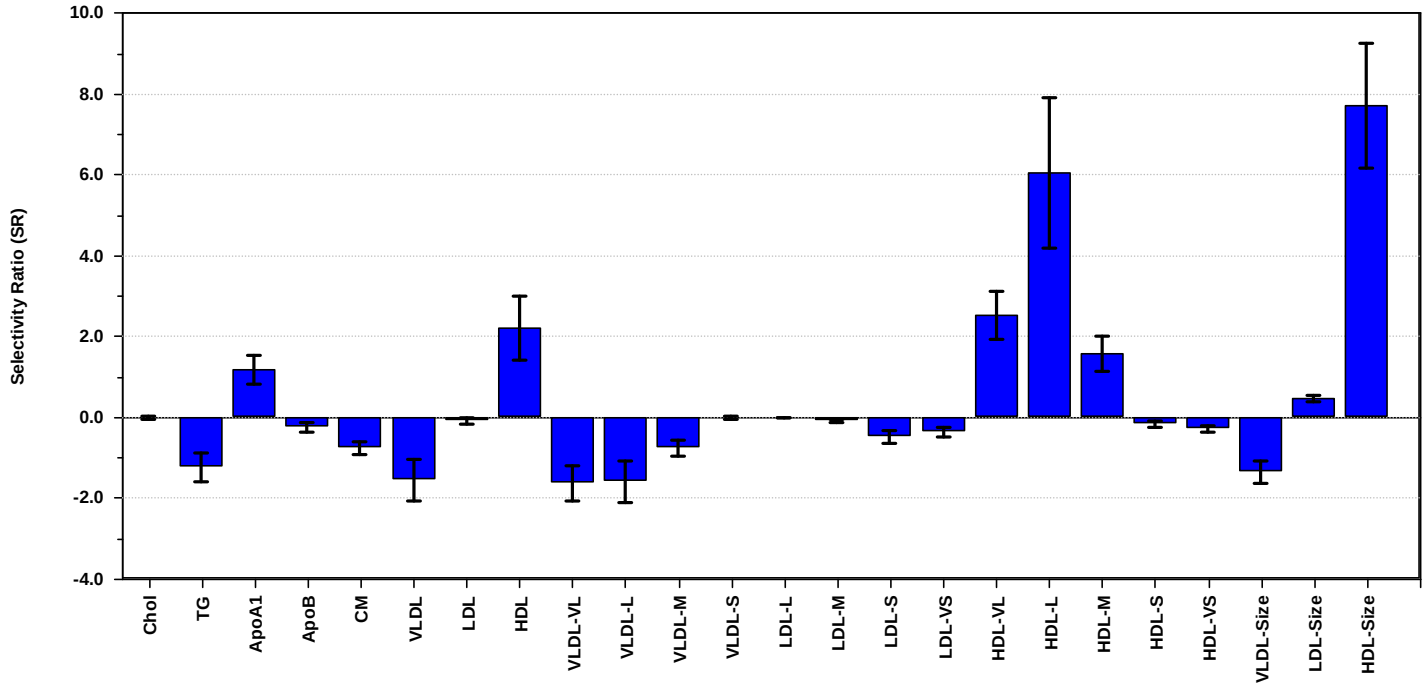


Modelling - PLS-DA/TP/SR

We have quantified the concentrations of all the biologically important fatty acids in blood and connected these to the lipoproteins (the transporters) by mathematical modelling.

This approach provides models that can be used to infer how elements of diet (e.g. EPA, DHA) impacts the lipoprotein pattern and, further on, how they impact on risk of CVDs.

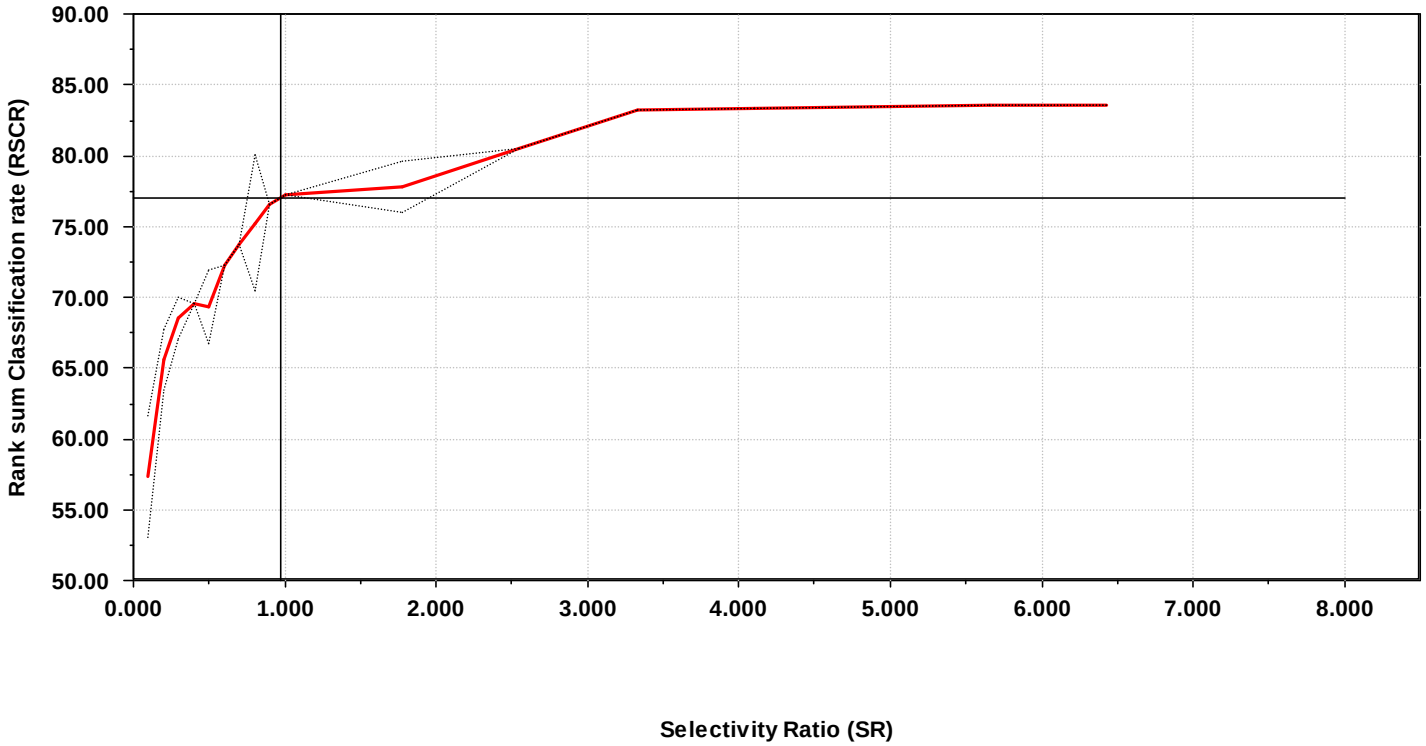
SR plot displaying gender differences in lipoproteins in healthy Norwegian adults



Positive bars => larger in women

Negative bars => larger in men

Discriminating variable (DIVA) plot showing Rank sum classification rate (RSCR) as a function of SR



Bonferroni corrected p-value is $7x10^{-11}$ for HDL-Size with SR=8.0 and RSCR=83%!

Lipoprotein and fatty acid patterns for children

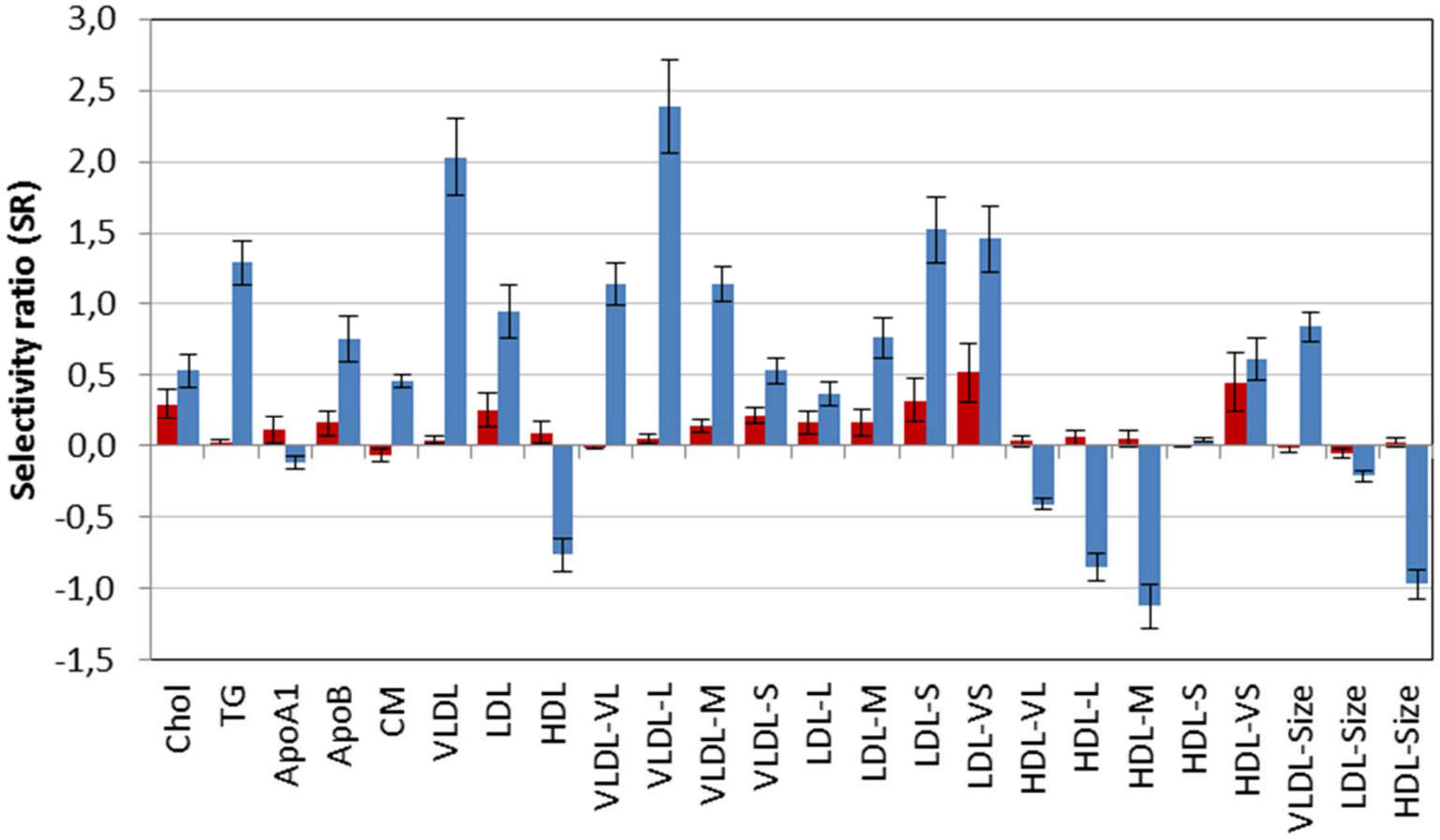
- Small gender differences
- Boys have larger HDL concentration than girls
- Girls have larger VLDL and TG concentrations than boys

Explanation: Boys are more physical active than girls

Changes in lipoprotein and fatty acid patterns from prepuberty to adulthood and during aging

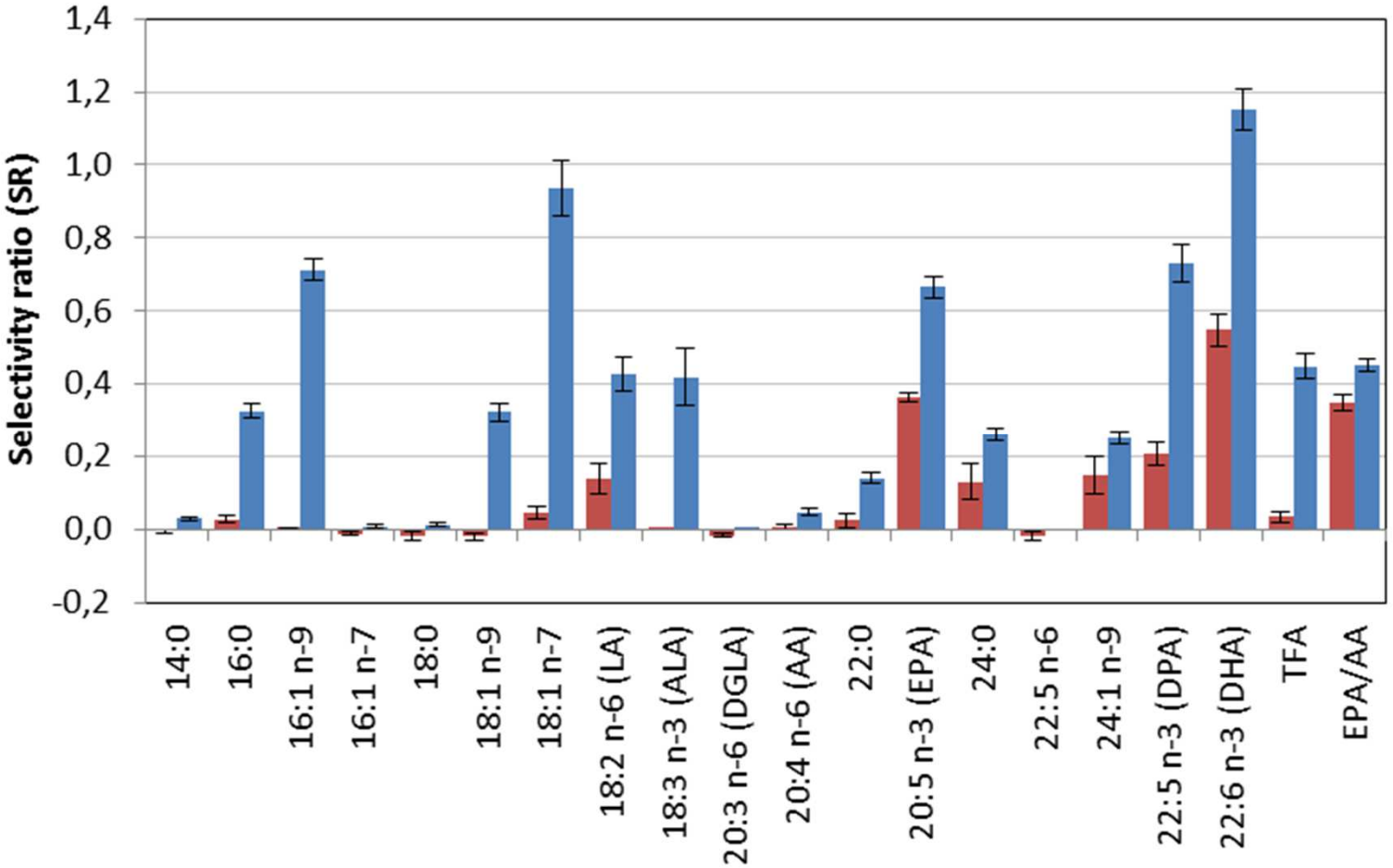
Rajalahti et al (2016), *Metabolomics*, 12:51.

Changes in lipoprotein pattern from prepuberty to adulthood



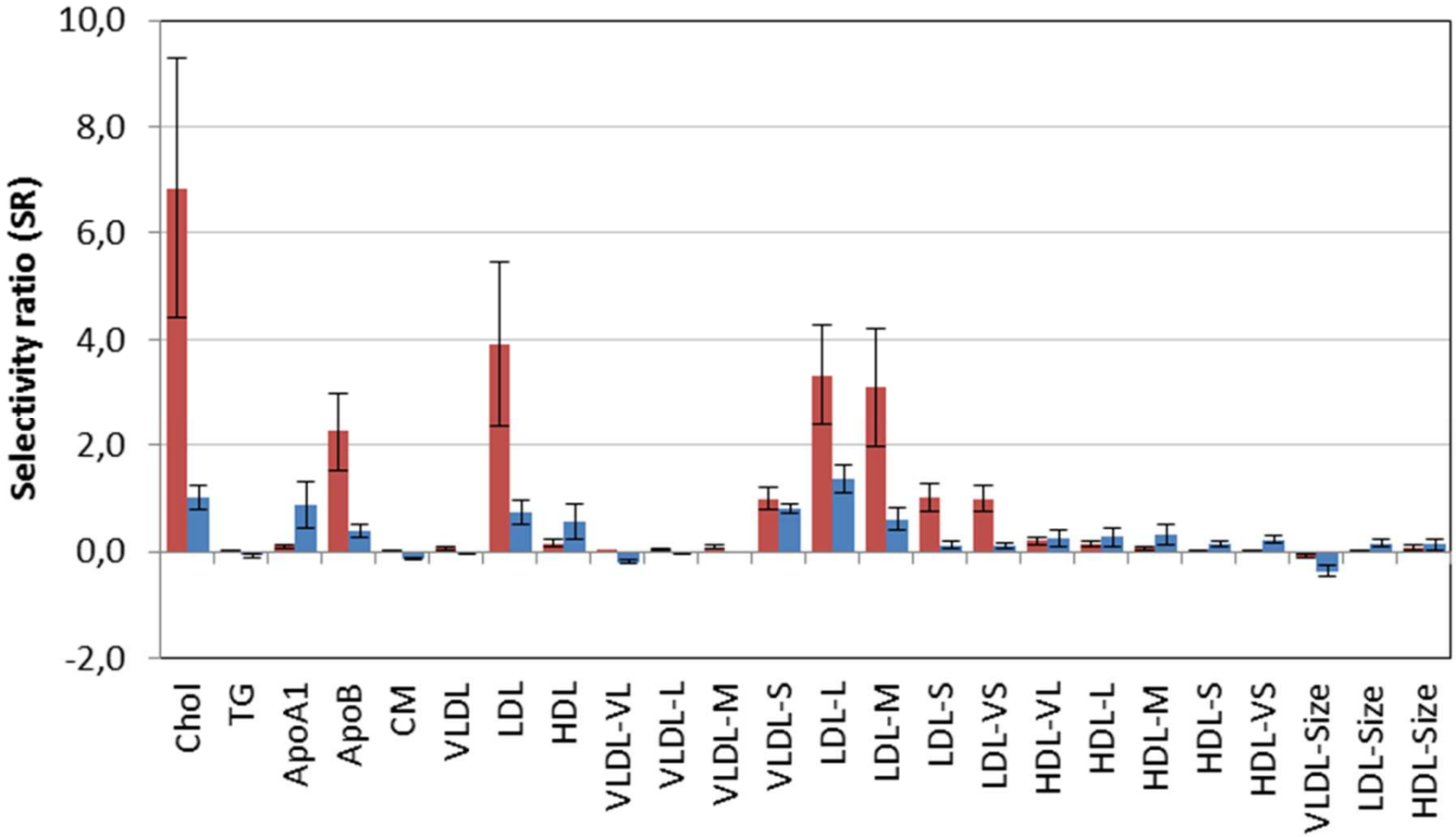
Blue => men, red => women

Changes in fatty acid pattern from prepuberty to adulthood



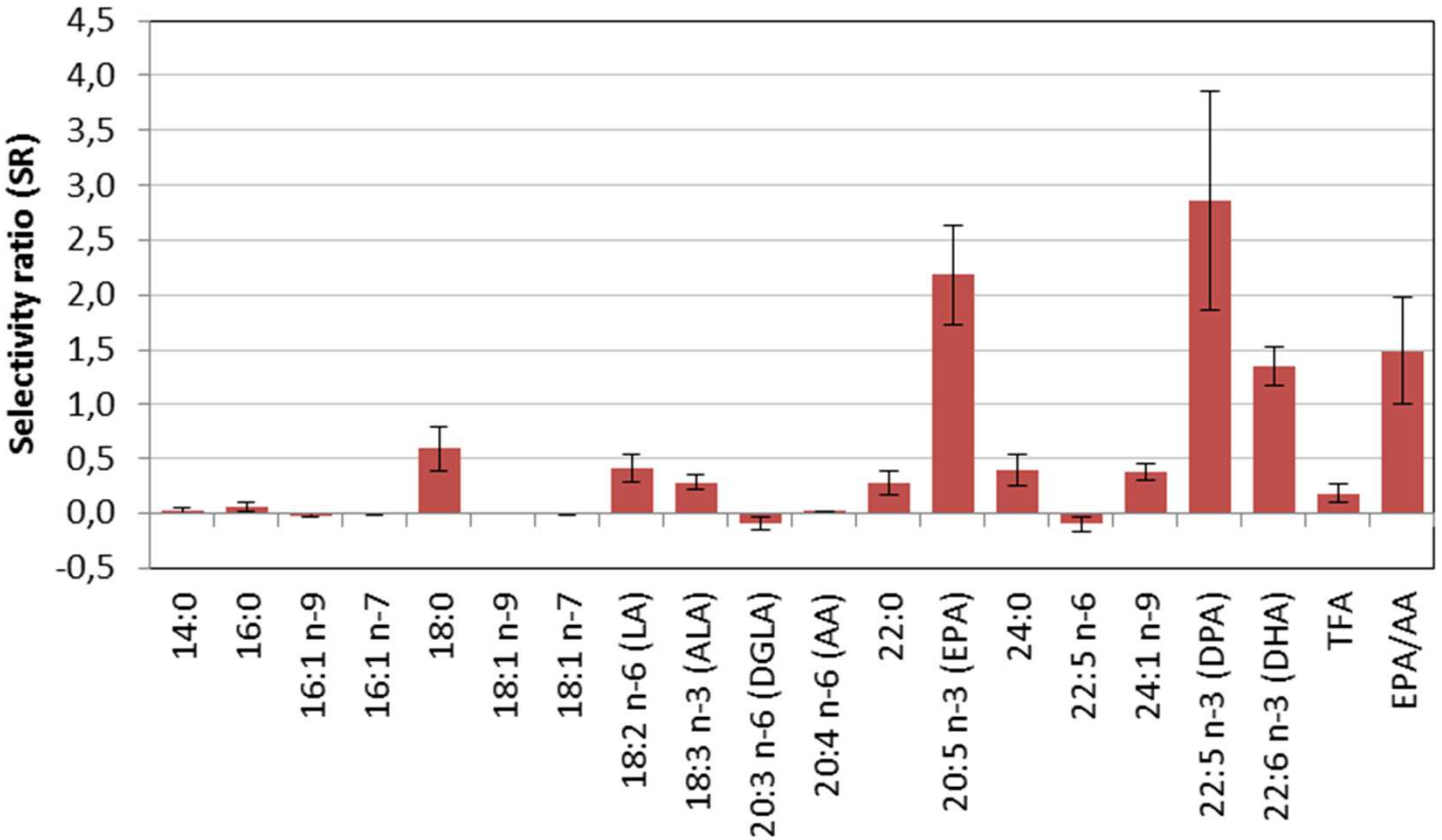
Blue => men, red => women

Changes in lipoprotein pattern with aging



Blue => men, red => women

Changes in fatty acid pattern in women with aging



Ongoing work, prepubertal children

- Changes in lipoprotein subclass pattern as a result of a physical activity intervention
- Effects of child obesity on lipoprotein pattern
- Cognition as a function of omega-3 fatty acid levels
- Genetic disorders: Hypercholesterolaemia

N > 1000 measured by proton NMR (lipoprotein and fatty acids) and GC (fatty acids) at two time points, before and after intervention

Ongoing work, morbidly obese subjects

- Effects of bariatric surgery – two kinds of surgery
- Intensive lifestyle intervention

N <25-30, but measurements at several time points

Co-workers

- Tarja Rajalahti (LP)
- Chenchen Lin (FA)
- Svein Are Mjøs (FA)
- Geir Kåre Resaland (PA)
- Eivind Aadland (PA)
- John Roger Andersen (lifestyle/obesity)
- Villy Våge (bariatric surgery)