From Registers to integrating -omics and GWAS

Jaakko Kaprio

University of Helsinki & National Public Health Institute, Finland

Structure

- Large cohorts and survival models

- Discordant pair design,
  - genetic mechanisms versus environmental causes of discordance
  - Teasing out consequences and causes

Linkage of the older (born before 1958) Twin Cohort to hospital discharge registry and National Insurance Institute medication registry to identify diabetes cases to end of 2004

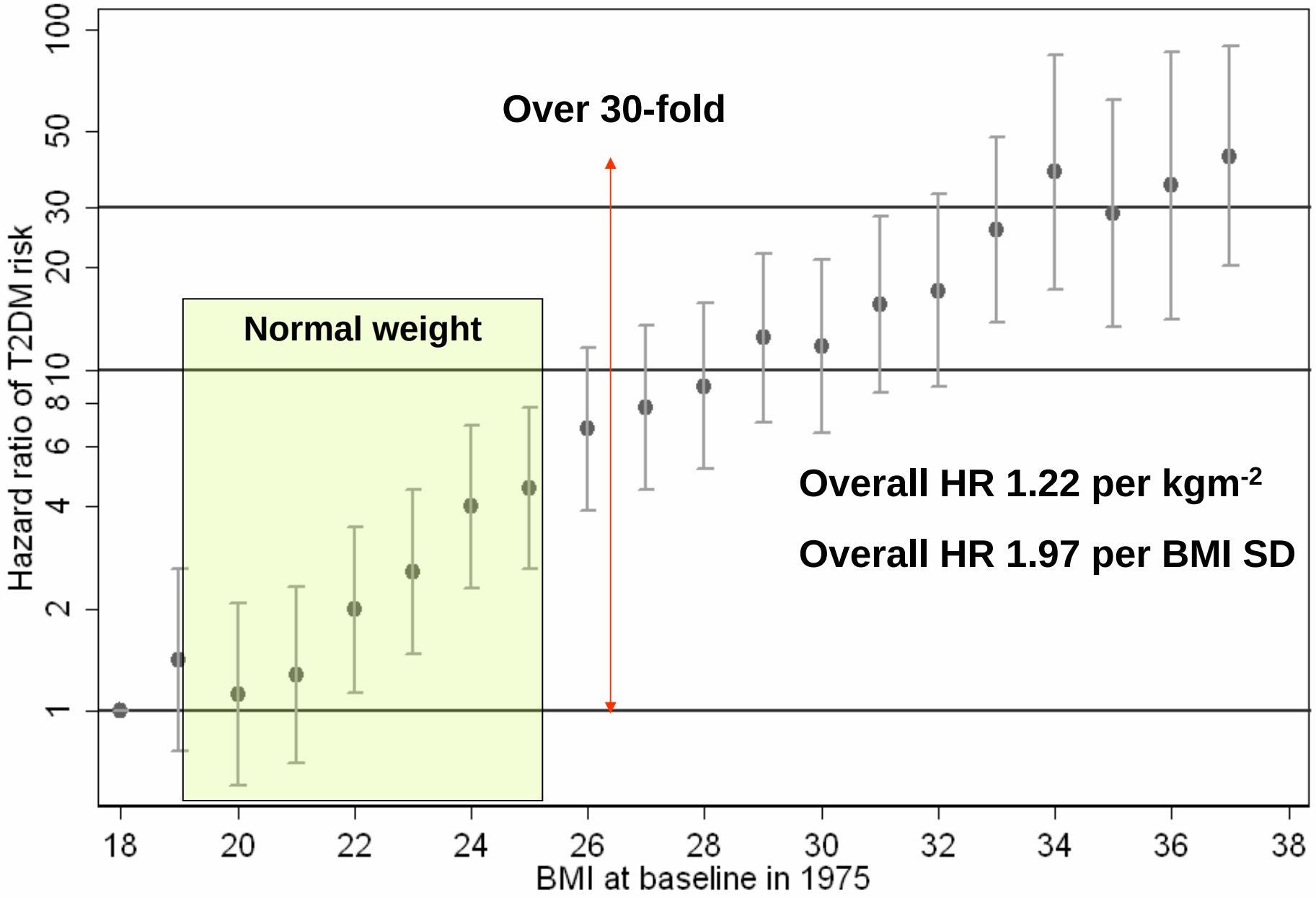
16430 twin pairs baseline cohort (MZ, DZ, XZ)

Total of 2336 diabetes cases, of which 2077 type 2 diabetes (rest are T1DM, gestational DM and secondary cases)

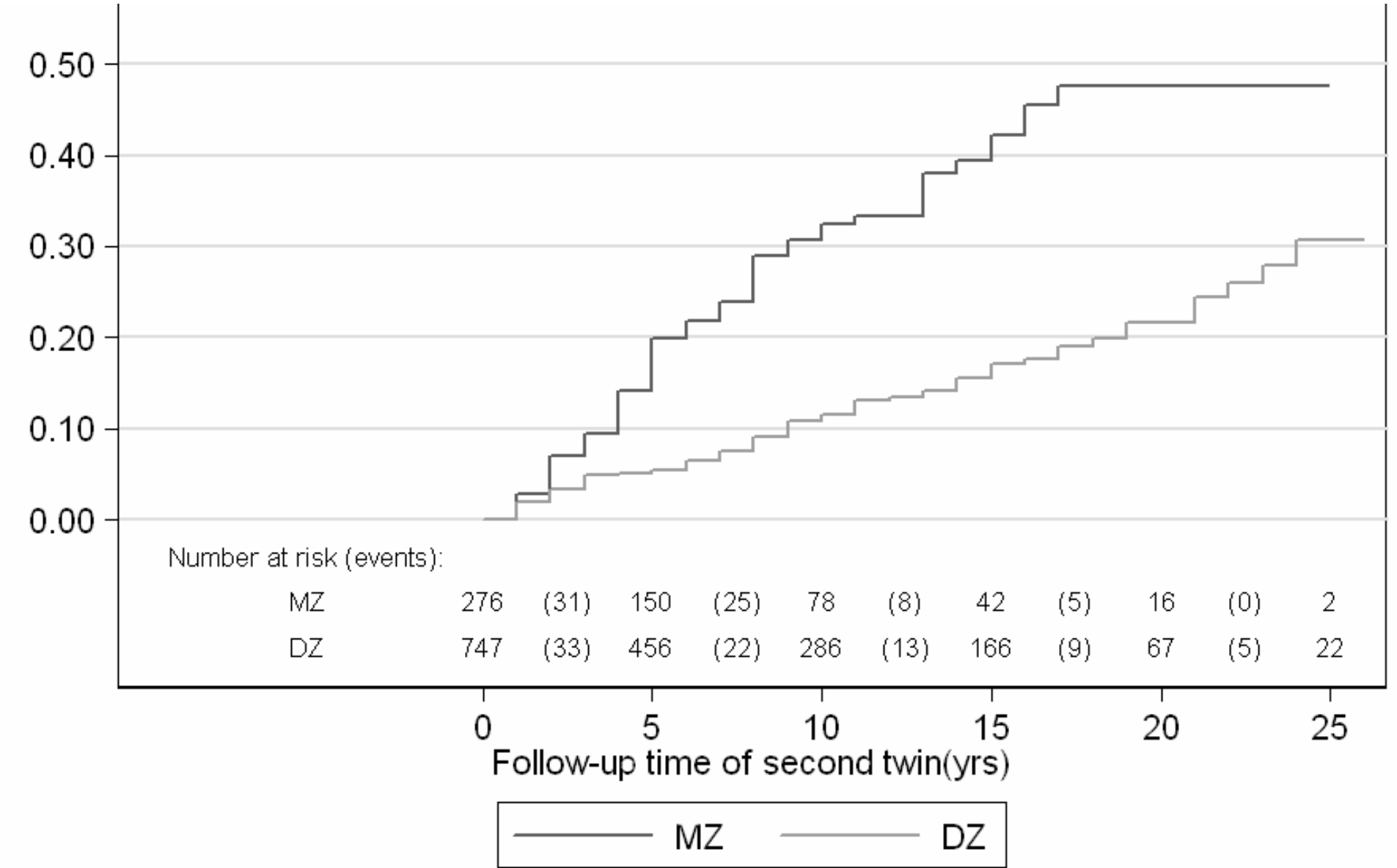
(Classified as described in Kaprio et al, Diabetologia 1992).
Lehtovirta et al, paper under review

Photo: Elina Ketola, Helsinki

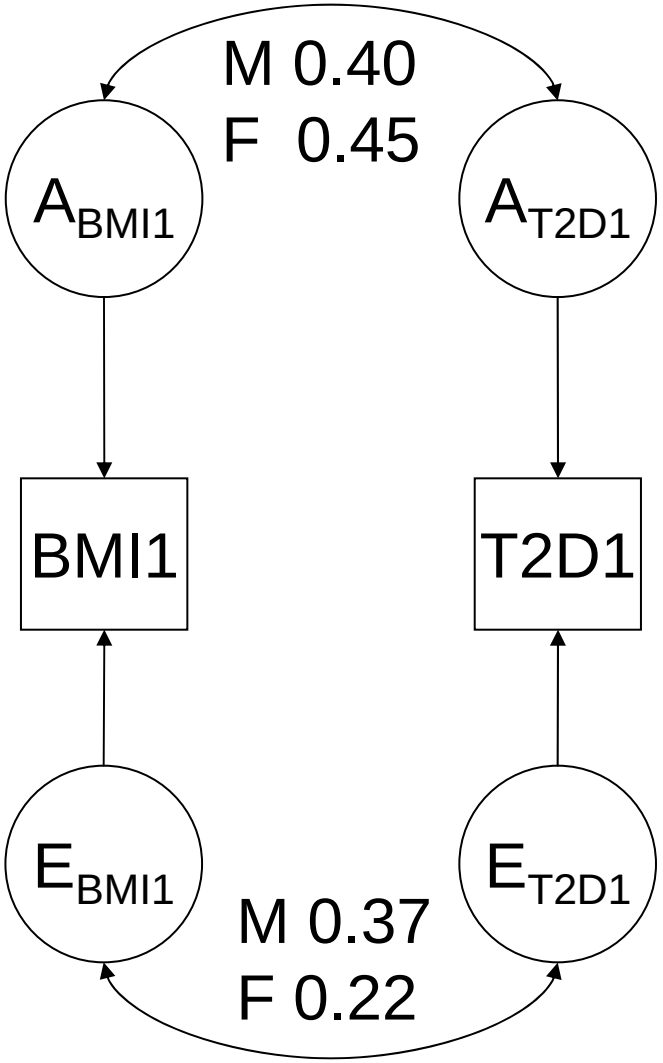
Risk of new T2 diabetes 1976-2004 by initial BMI level



Risk for T2D in the twins with affected cotwin

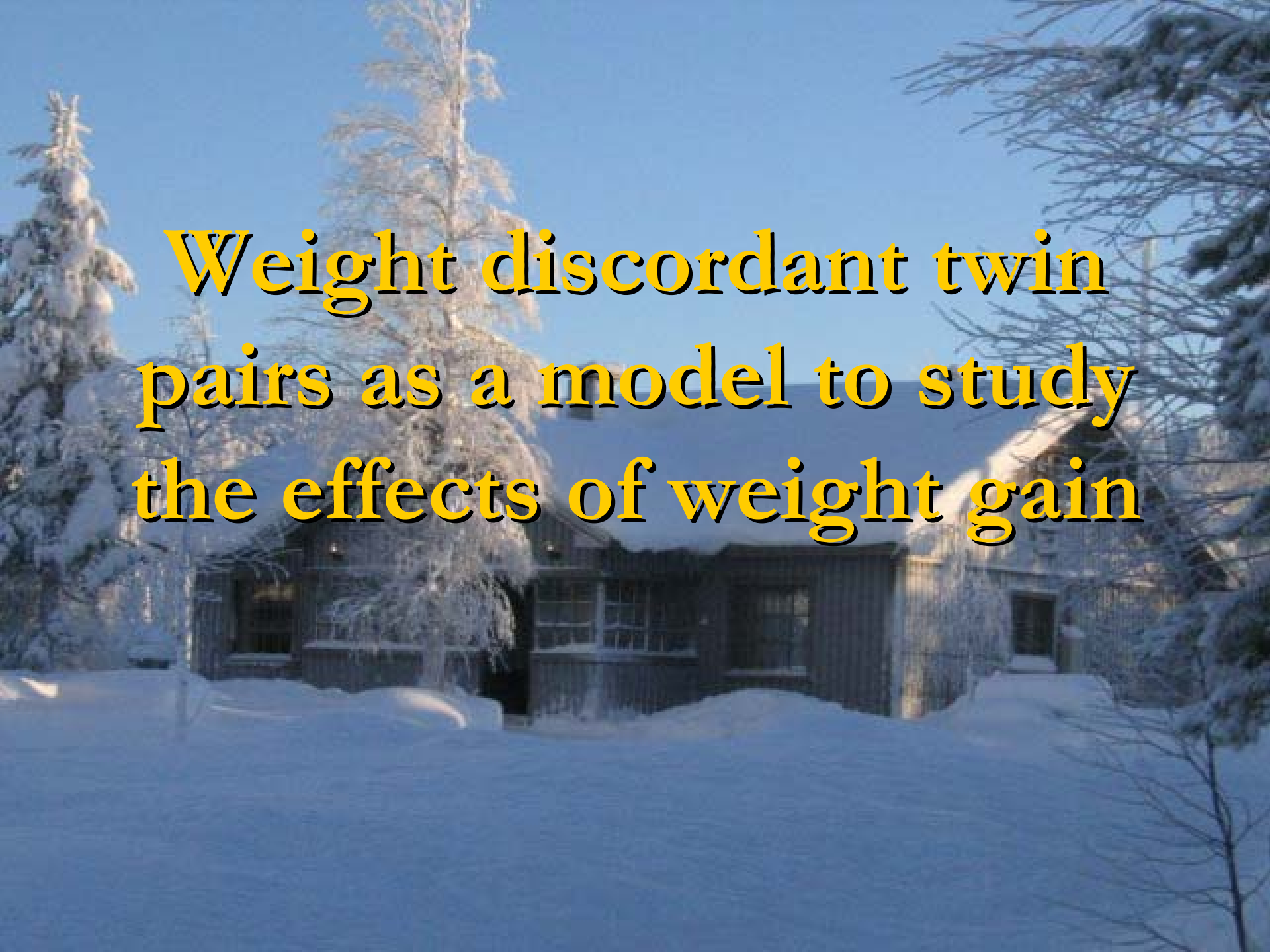


Bivariate variance component model

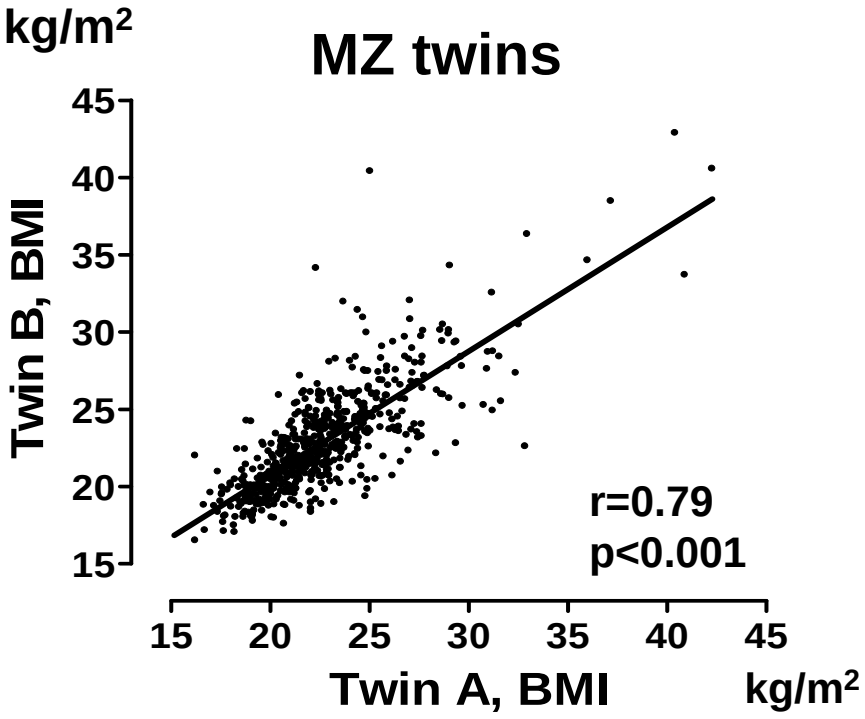


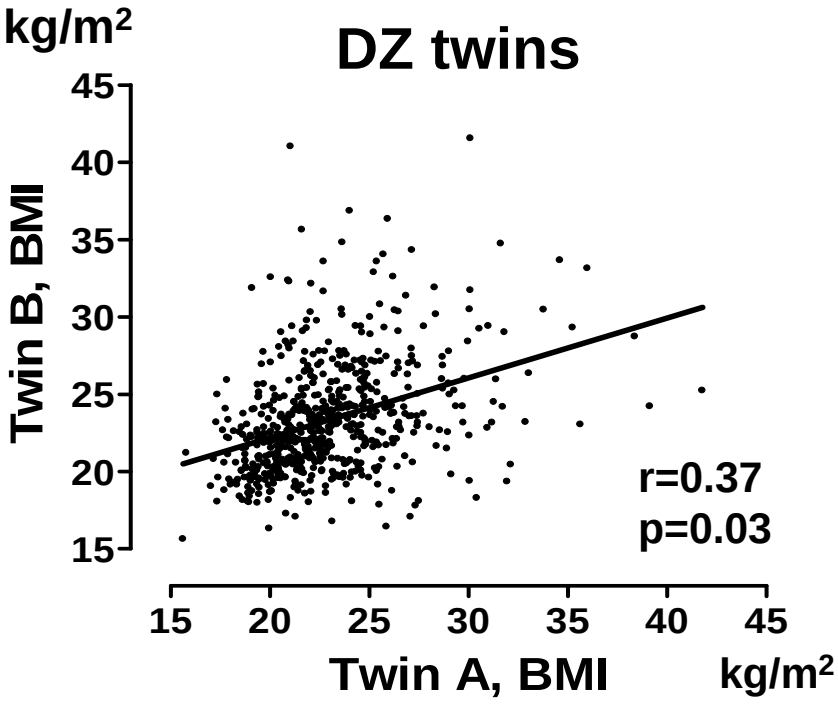
Twin 1

Lehtovirta et al, paper under review

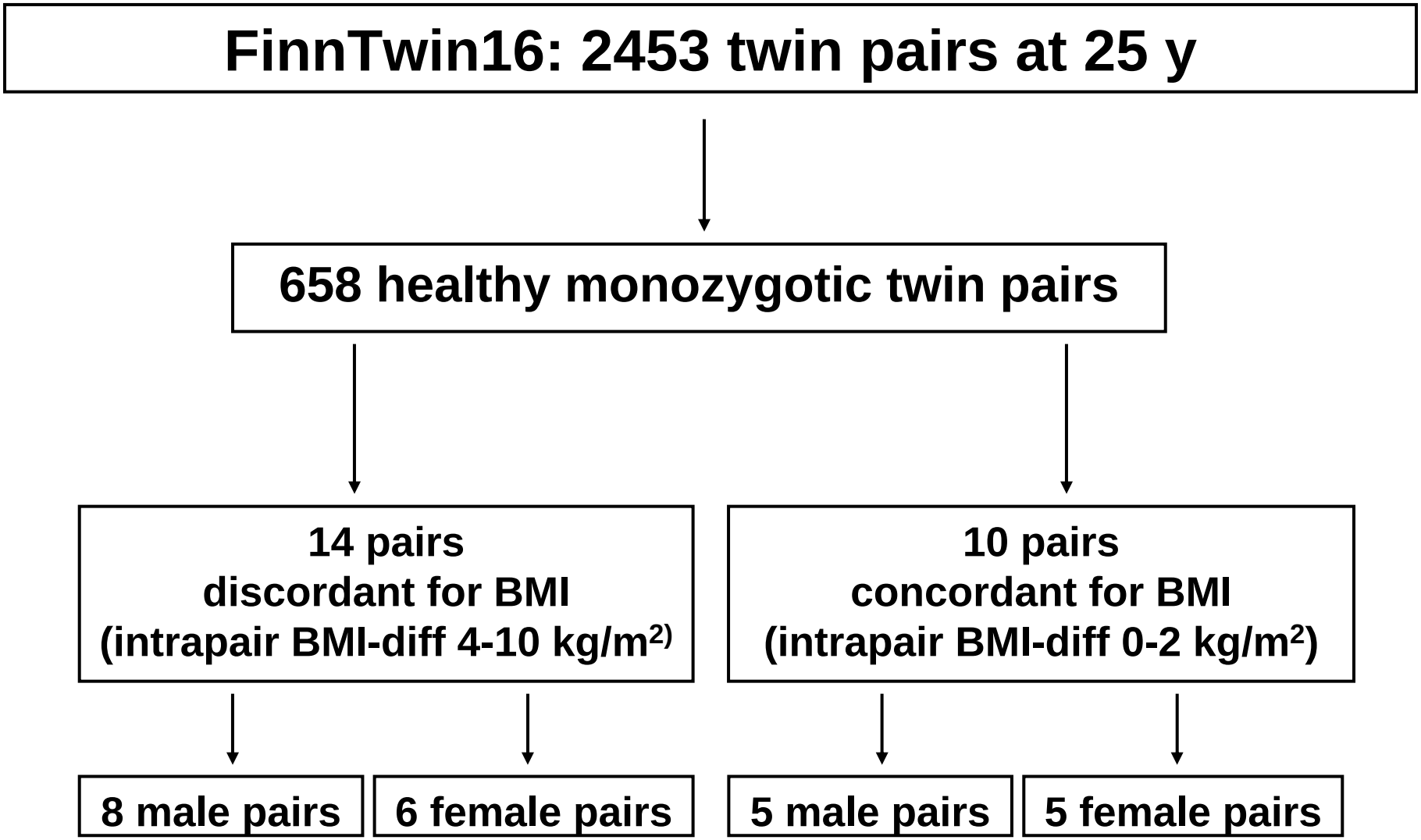

Weight discordant twin pairs as a model to study the effects of weight gain

FinnTwin16, 90% of all twins born in Finland 1975-79
n=2453 pairs at 25 y





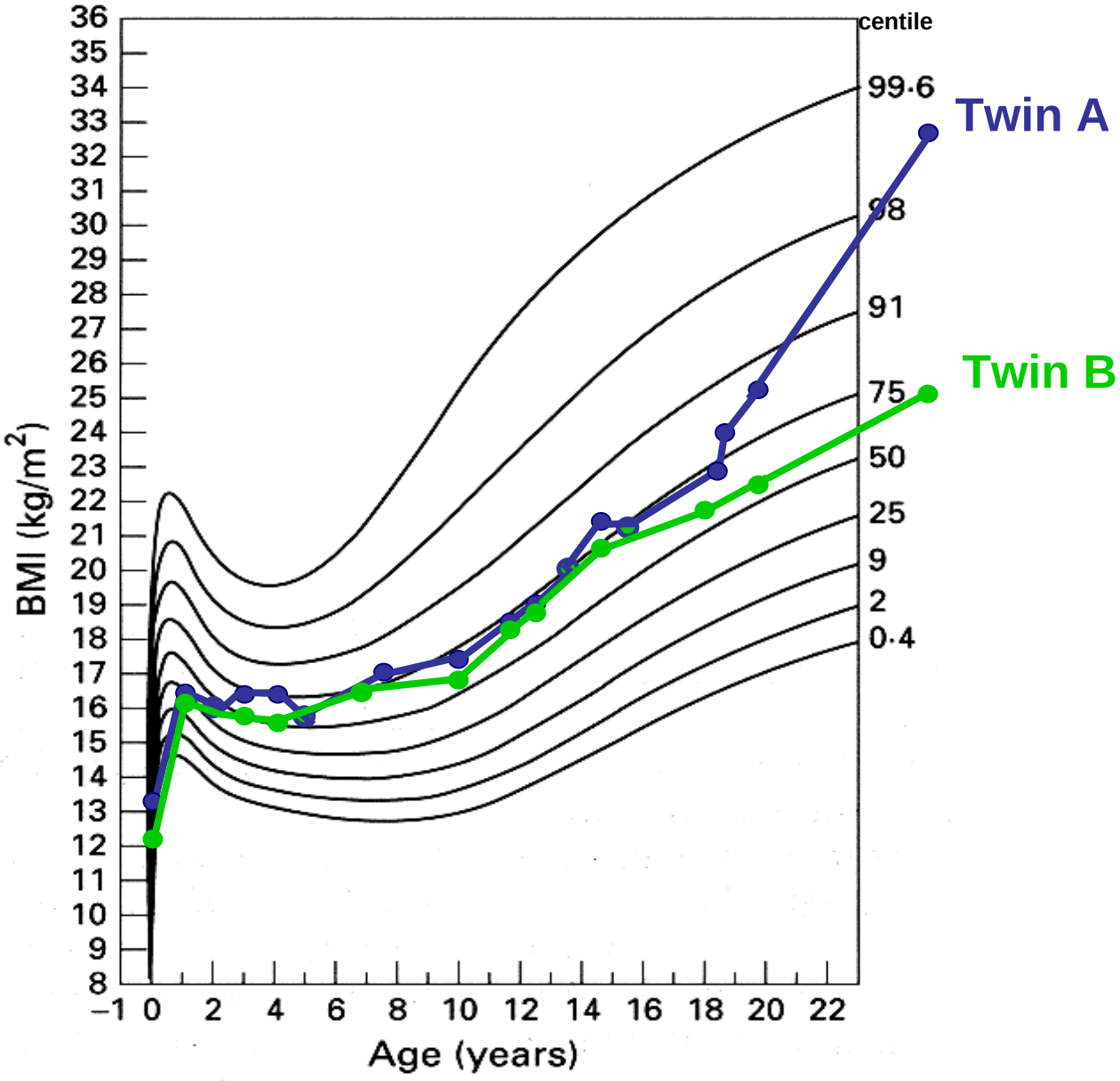
Kaprio J, FinnTwin16


FinnTwin16: 2453 twin pairs at 25 y
658 healthy monozygotic twin pairs
14 pairs
discordant for BMI
(intrapair BMI-diff 4-10 kg/m2)
10 pairs
concordant for BMI
(intrapair BMI-diff 0-2 kg/m2)
8 male pairs
6 female pairs
5 male pairs
5 female pairs

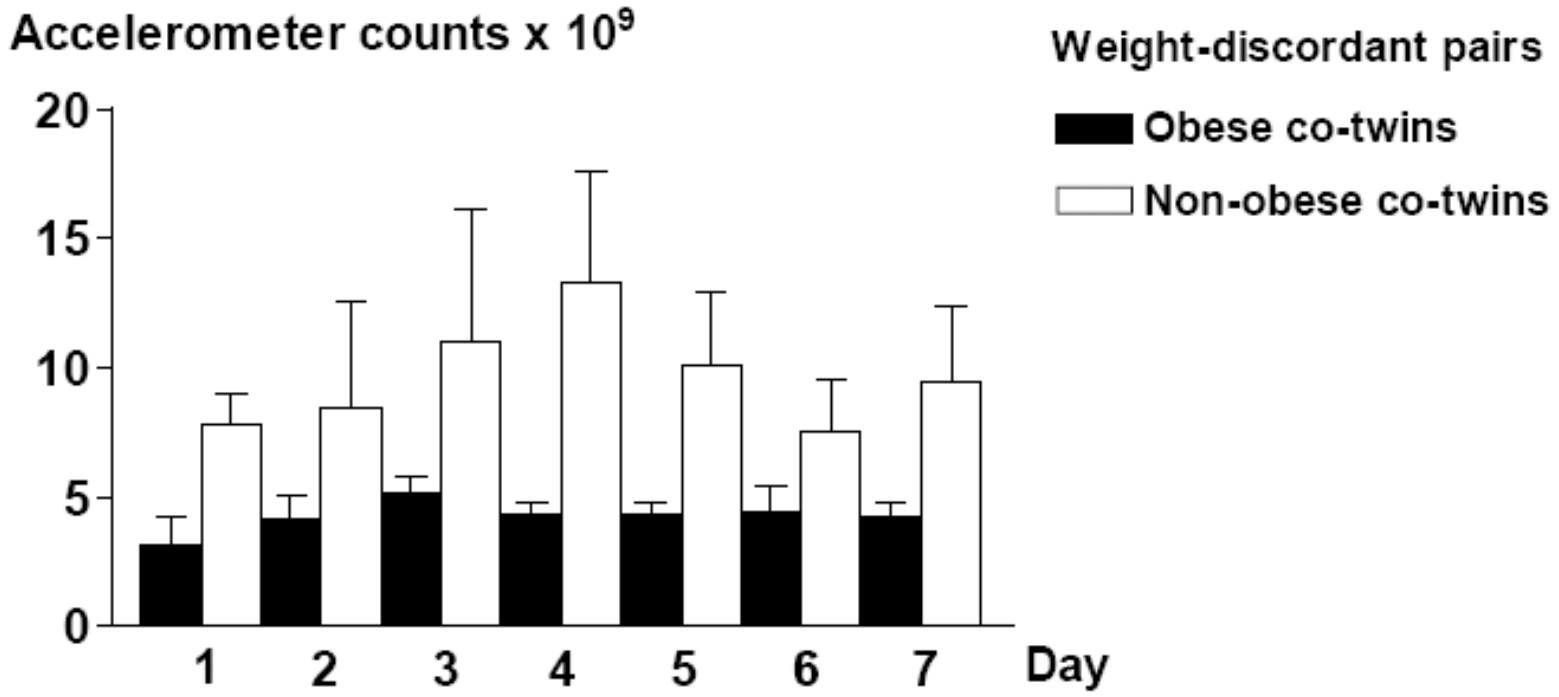
Measures for examination of metabolic features, behavioural characteristics, and physical fitness

- Fasting blood samples for DNA, routine hematology, chemistry and lipids, cytokines, neuropeptides, lipidomics etc
- body composition and anthropometrics by DXA, bioelectrical impedance, skinfolds, and circumferences
- body fat accumulation (subcutaneous and visceral fat content by MRI, intrahepatic and intramyocellular fat content by proton spectroscopy)
- adipocyte gene expression from subcutaneous fat biopsies, candidate genes and genome-wide microarray analyses
- in MZ discordant twins, mitochondrial DNA sequencing & telomere length assays

- intra-arterial endothelic function
- whole body insulin sensitivity under normoglycemic hyperinsulinemic conditions (the clamp technique)
- test meal with ghrelin and leptin assays
- PROP- and fat-tasting procedures
- resting energy expenditure (indirect calorimetry) and a 14 day-total energy expenditure in free living conditions (the doubly labeled water technique)
- Accelerometers & physical fitness by bicycle spiroergometer
- questionnaires and interviews on past and current food intake, food preferences, physical activity, use of alcohol and smoking, health-related attitudes, weight history, family history and quality of life
- structural and functional brain MRI, SPET
- autonomic nervous system assays
- structured psychiatric interview
- questionnaires and interviews of the parents

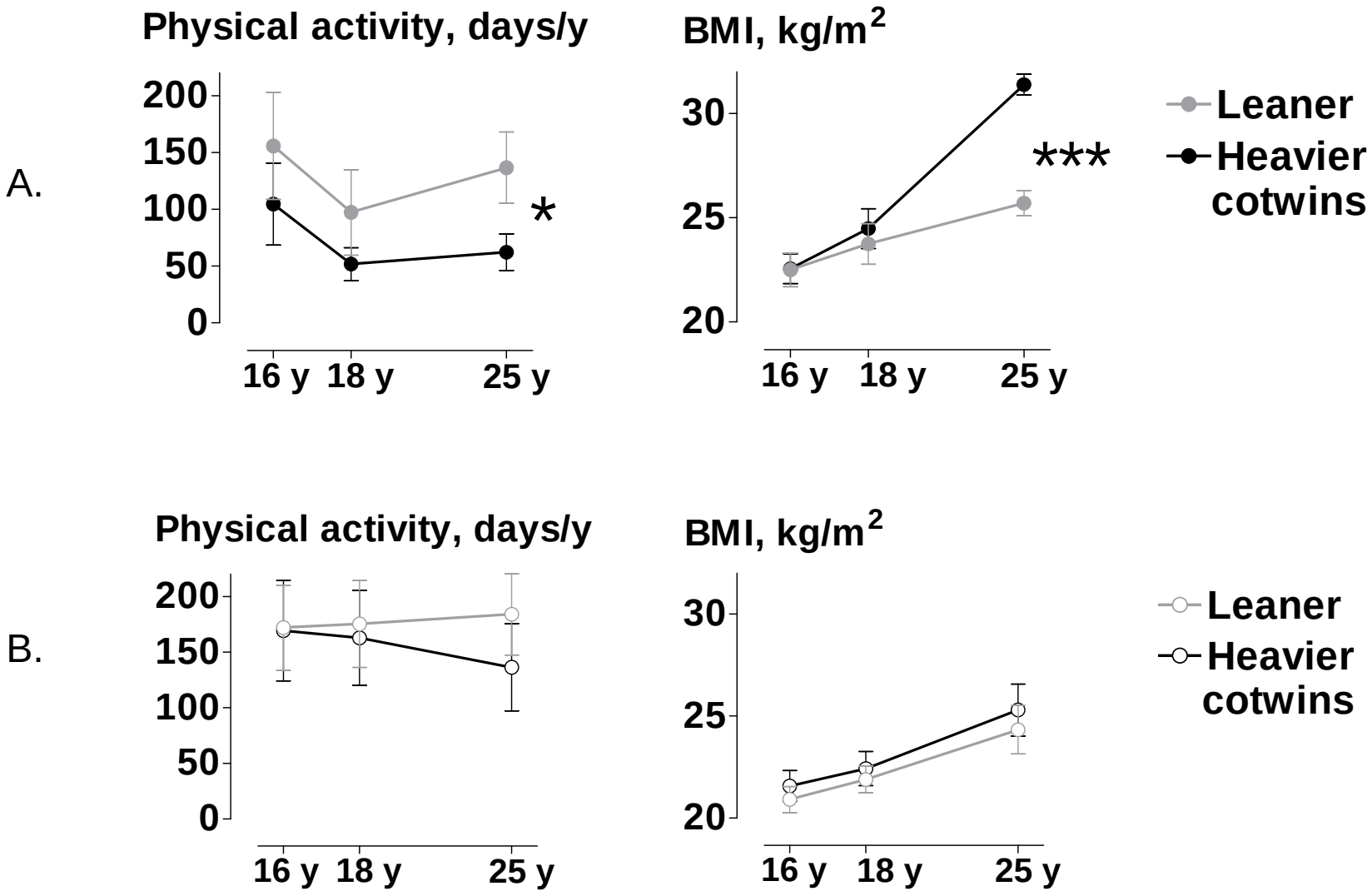


Pietiläinen et al. Twin Res 2004;7:421



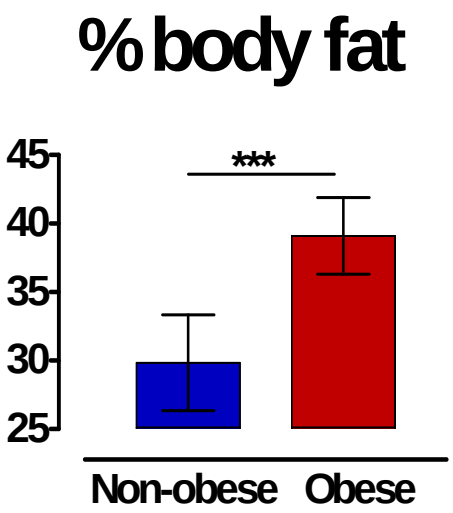
Pietiläinen et al, Obesity 2008

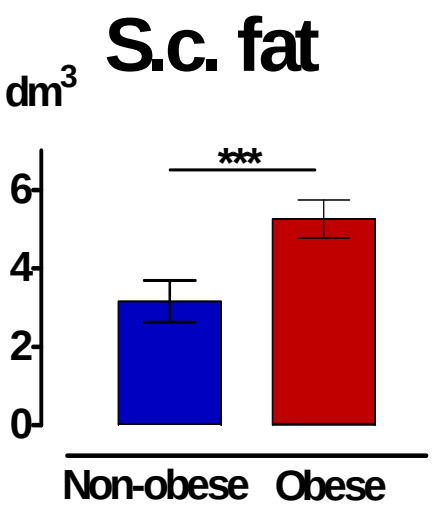
Physical activity and weight from adolescence to young adulthood in MZ twin pairs discordant (A) and concordant (B) for obesity as young adults

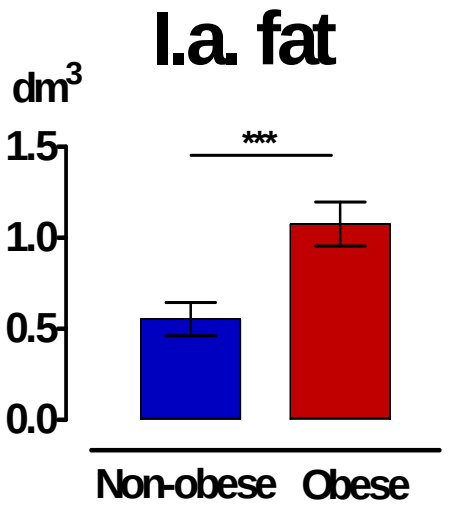


Pietiläinen et al, Obesity 2008

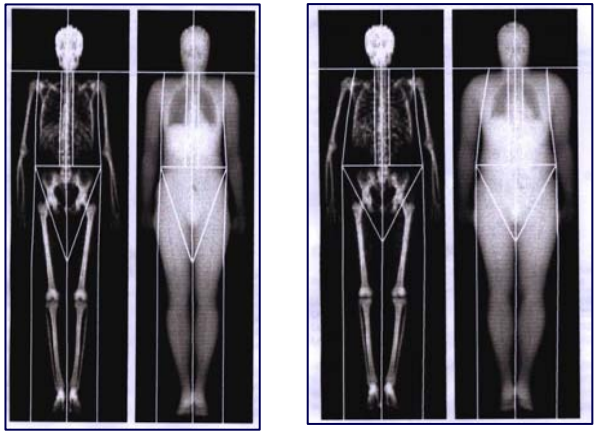
Obesity-discordant pairs, average weight-difference = 15 kg



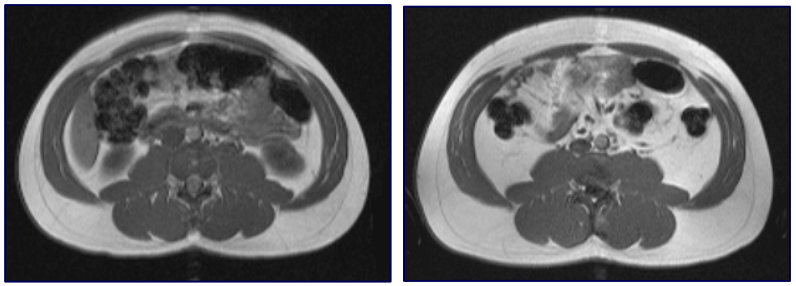




DEXA

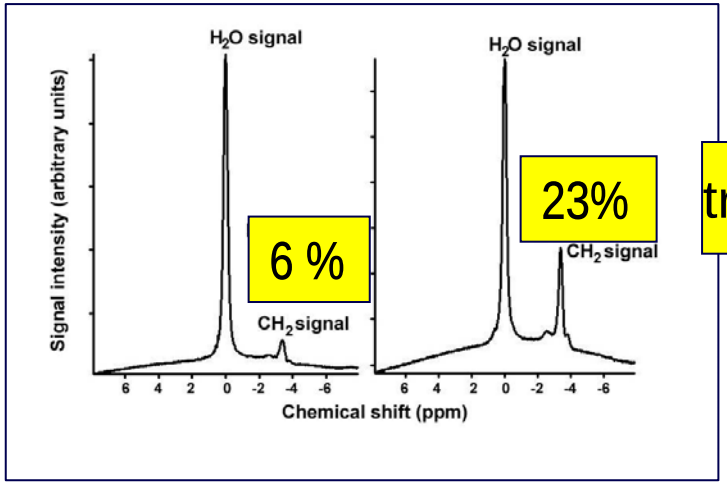


MRI

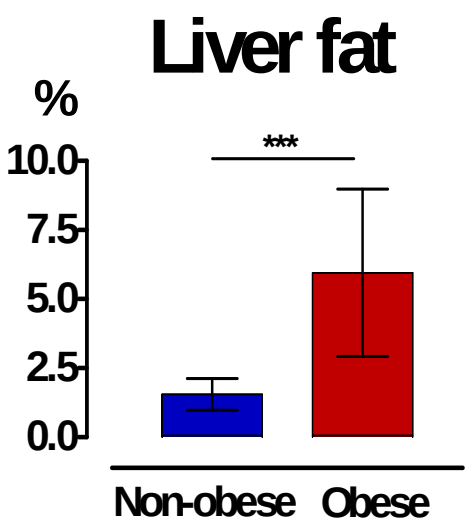


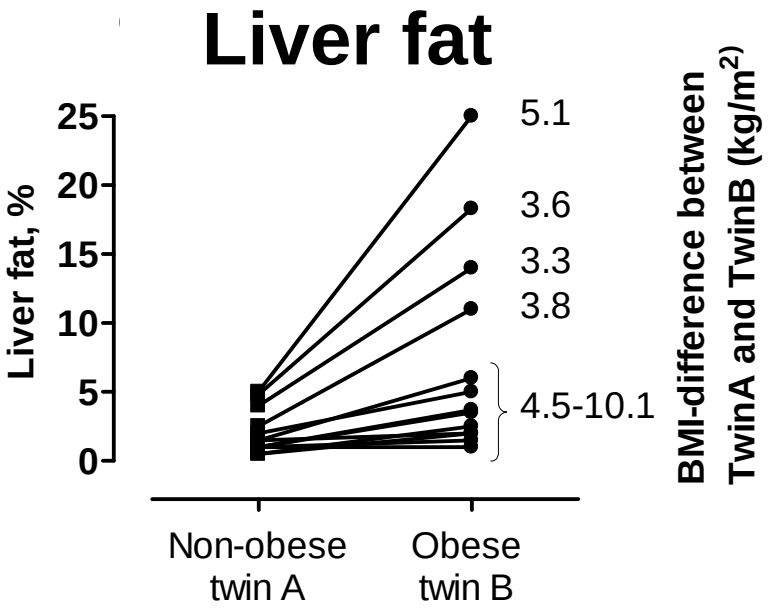
Pietiläinen et al. AJP-Endo 2005

MRI-spectroscopy



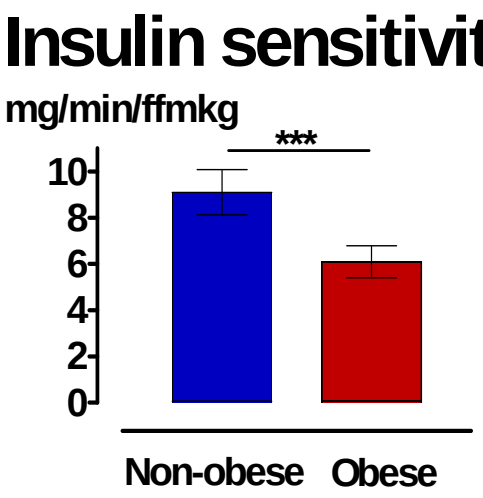
triglyceride peak

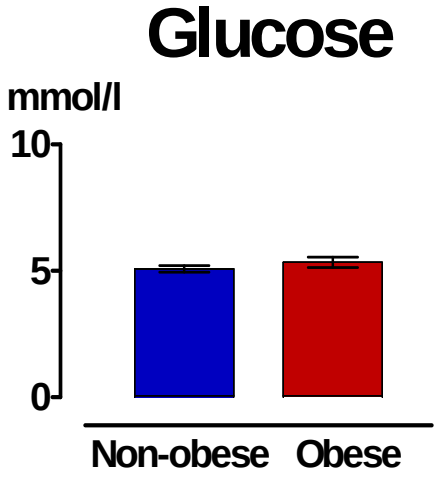




Pietiläinen et al. AJP-Endo 2005

Euglycemic, hyperinsulinemic clamp





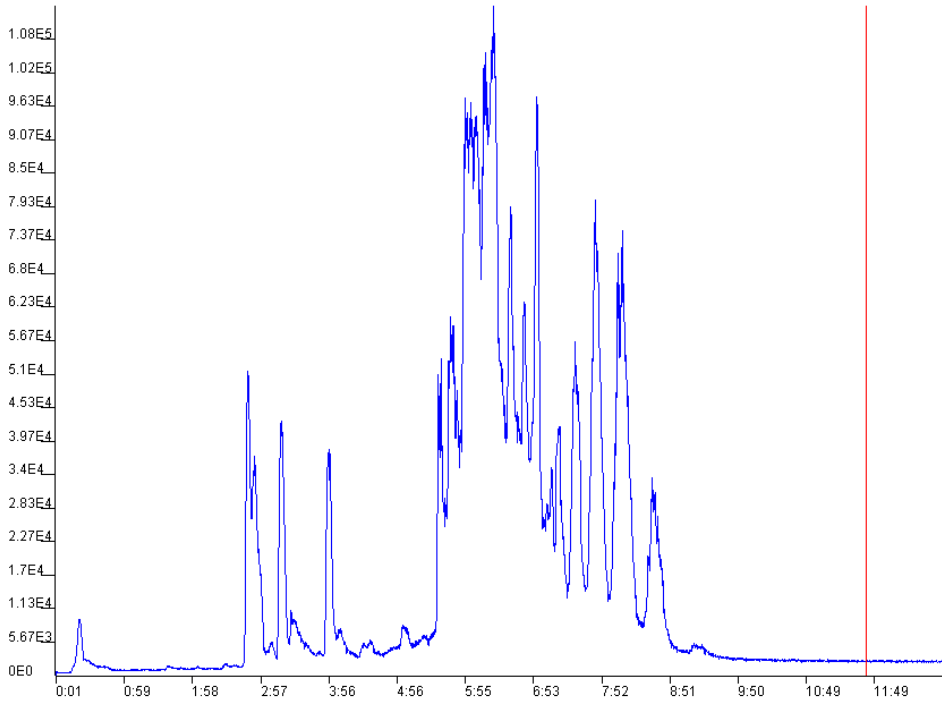
Pietiläinen et al. AJP-Endo 2005

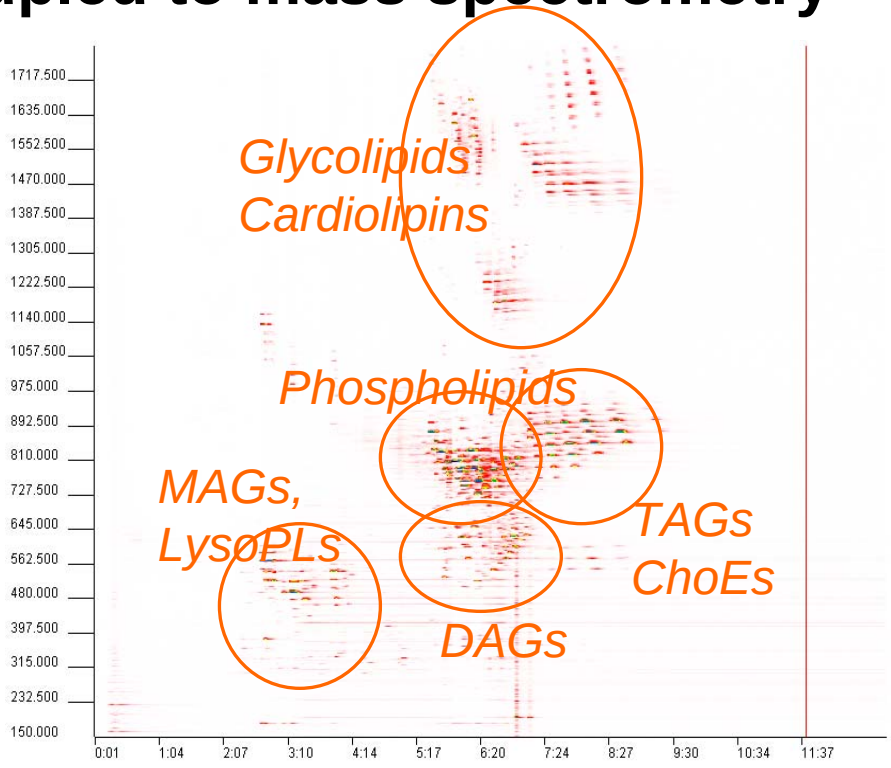
Classical serum lipids			
	Non-obese	**p**	**Obese**
LDL	**2.6 ± 0.2**	0.03	**2.8 ± 0.2**
HDL	**1.5 ± 0.1**	0.01	**1.4 ± 0.1**
TG	**1.0 ± 0.1**	0.03	**1.3 ± 0.1**

Lipidomics

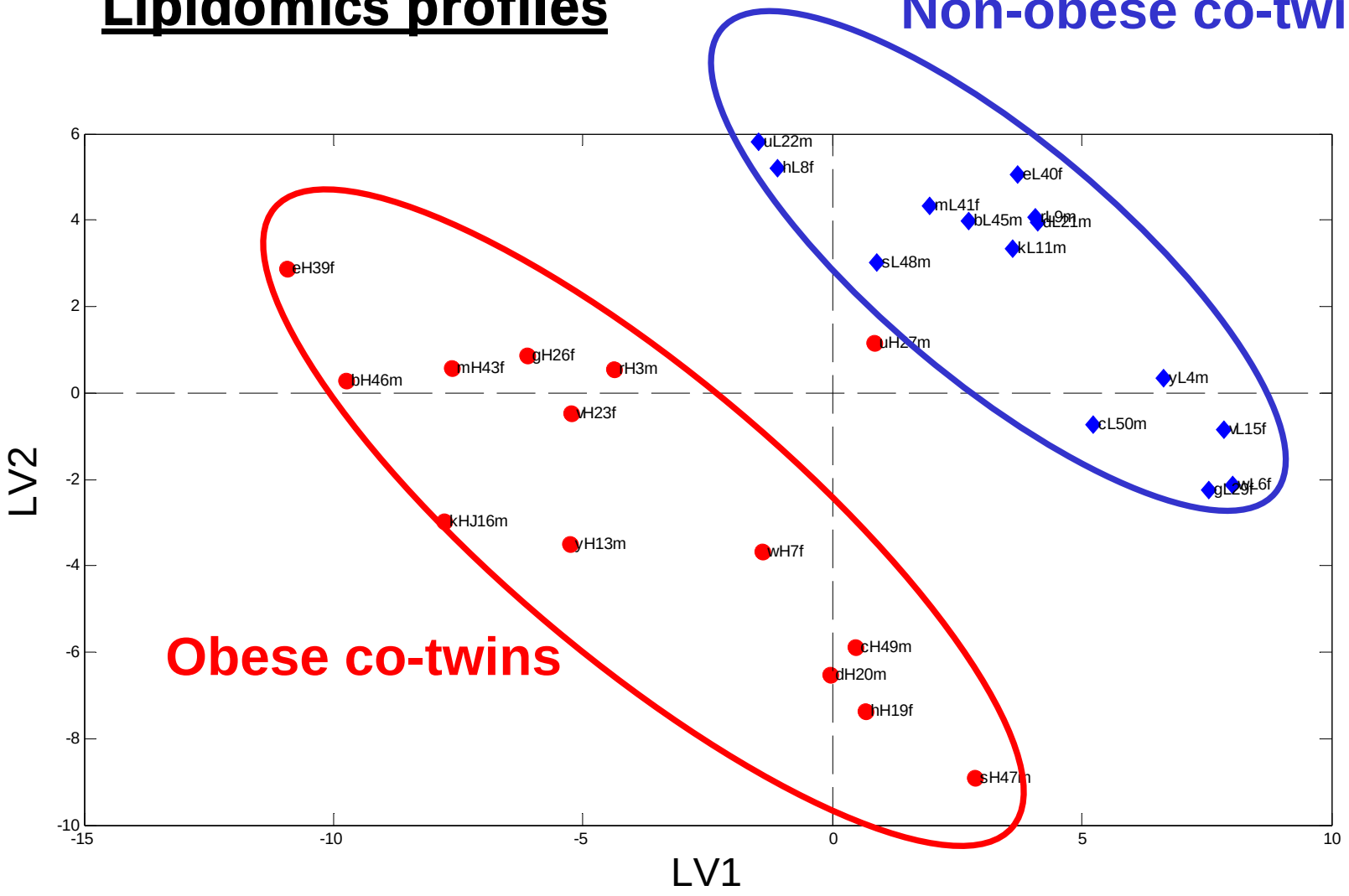
f-Plasma: 330 lipid molecular species

Liquid chromatography coupled to mass spectrometry



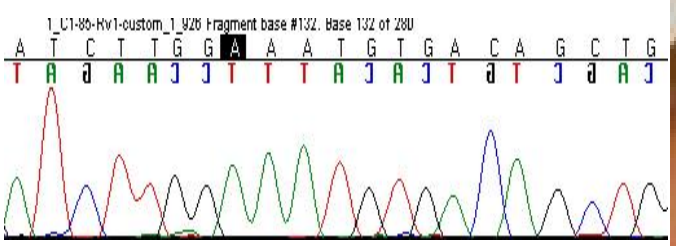


Lipidomics profiles

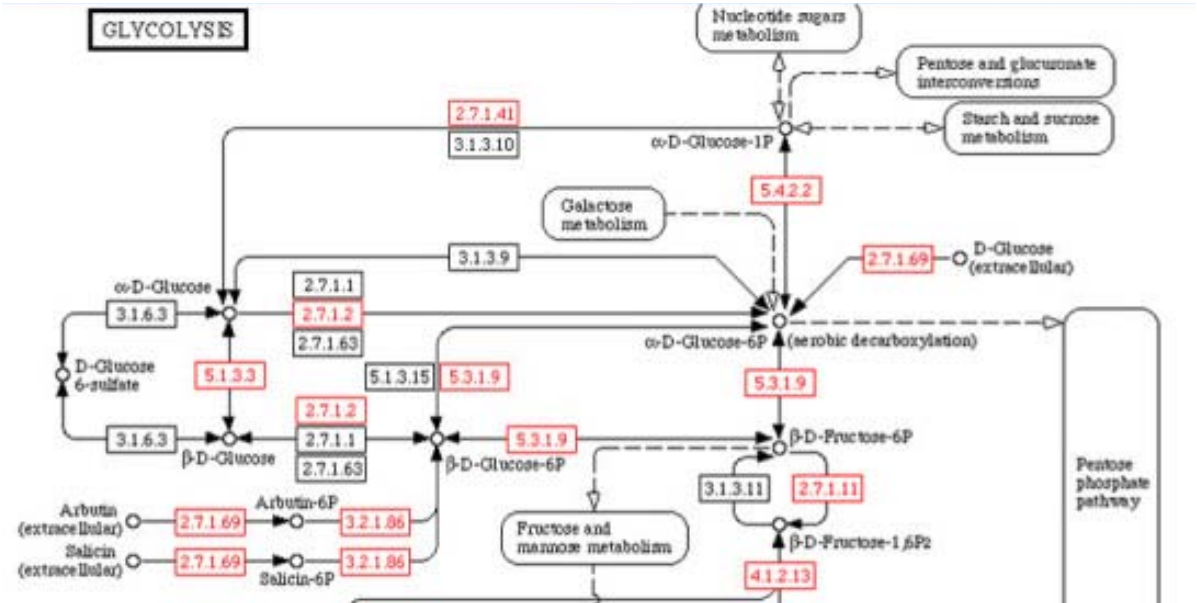


Pietiläinen et al. PLoS One 2007

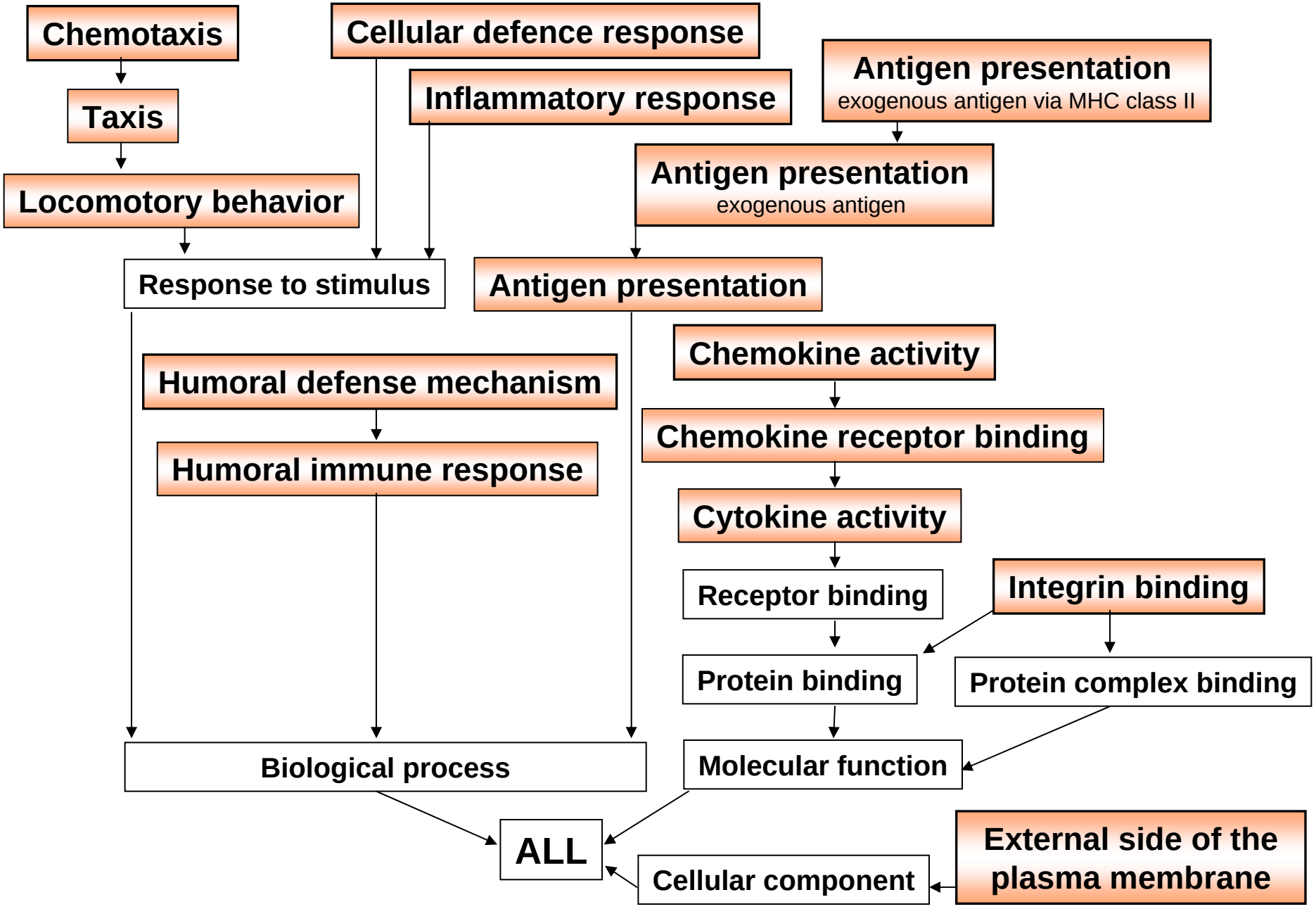
Adipose tissue gene expression



Pathway analyses



15 inflammation pathways up-regulated in obese fat



FinnTwin16
90% of twins born in Finland 1975-1979
2453 twin pairs at 25 y
658 MZ twin pairs

18 MZ twin pairs
Discordant for obesity
BMI difference >3 kg/m^2

Non-obese

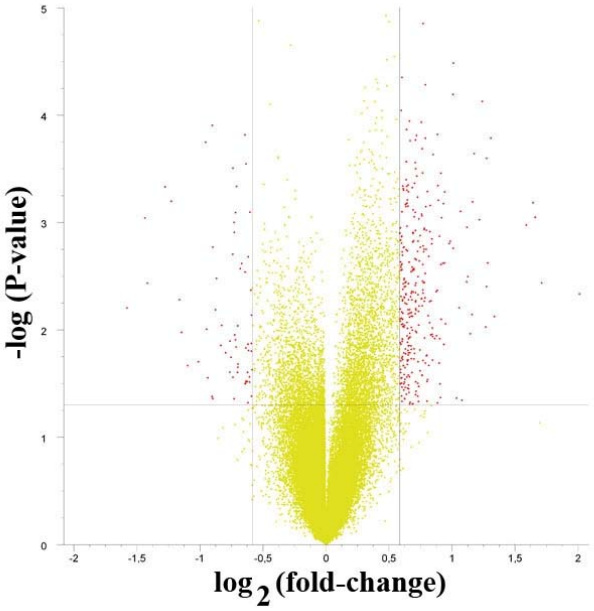
Obese

14 pairs participated

10-12 h fast

Fat biopsy	**Affymetrix, mtDNA sequence and copy nro**
Blood test	**Insulin, adiponectin, leptin, BCAAs, ketoacids**
Clamp	**Whole body insulin sensitivity**
DEXA	**Body composition**
MRI, spectroscopy	**Intra-abdominal and subcutaneous fat, liver fat**

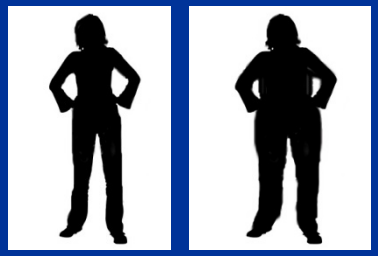
Differentially expressed genes in fat btw obese and non-obese



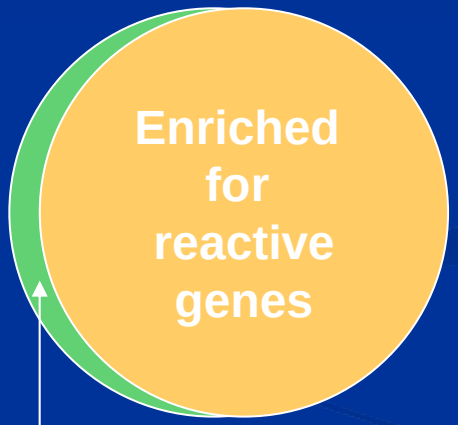
Enriching for ”Causative” Genes

Obesity *causative* genes

Obesity *reactive* genes



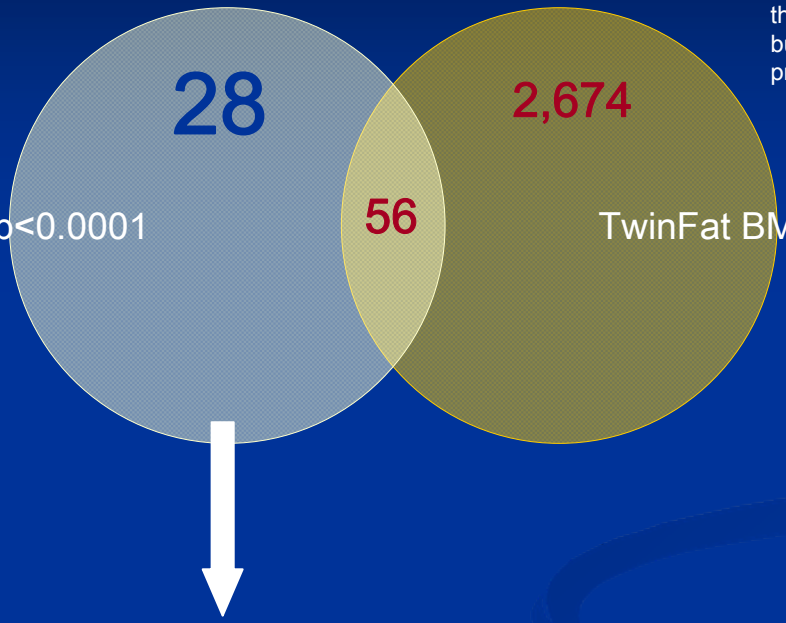
BMI discordant MZ twins greatly enriched for *reactive* genes



Causative genes?

Enriching for Causative genes

Note: Here I have used a much less stringent criteria for statistical significance in the TwinFat samples, in order to not miss any genes that are in fact reacting to obesity. This way we can increase the likelyhood that the genes left over really are not reacting to BMI, but rather (hopefully) "*causal*", or related to a causal process...



BMI *correlated* genes with p<0.0001

TwinFat BMI *reactive* genes with p<0.05

These 28 probes represent 27 unique genes/gene products

Retrieved genomic positions for these probes, added 1kb up- and down-stream sequence, then acquired the genotypes for the regional SNPs from two cohorts with GWA data available...

Testing for Association

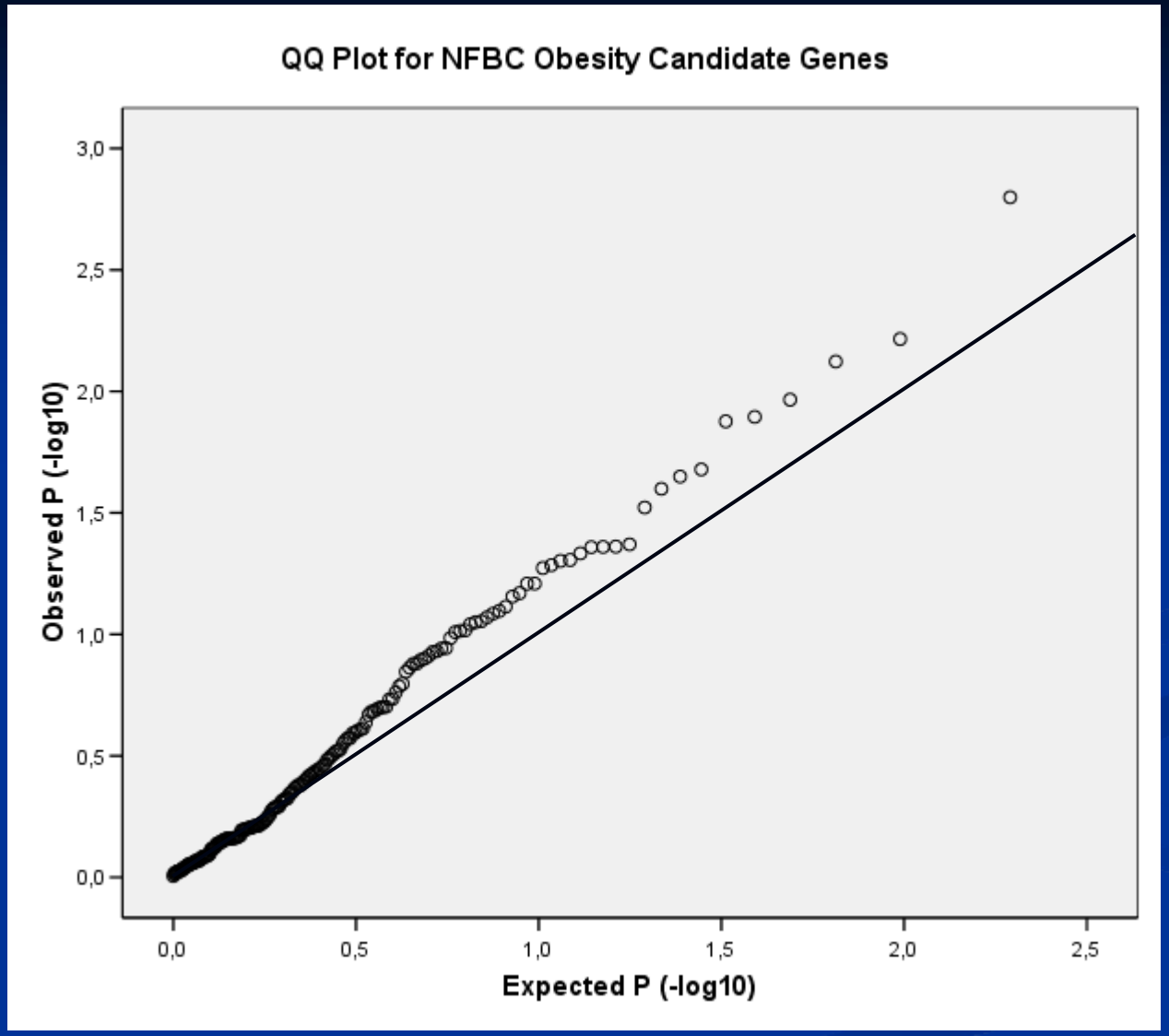
- Genotypes for 204 SNPs acquired from two GWA cohorts:

Northern Finland Birth Cohort (NFBC)
(N=4,762)

The Discovery cohort

MZ Twin GWA
(N=1,691)

The Replication cohort

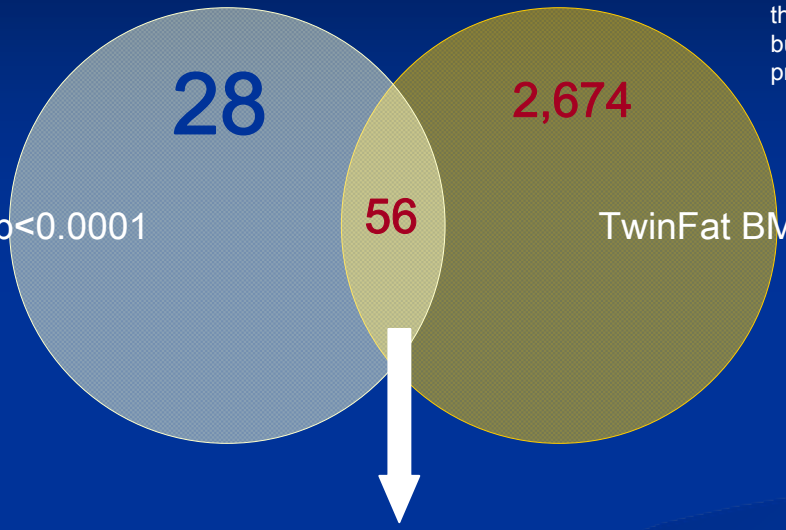


195 SNPs
27 genes
N=4,762

Many more points above the null hypothesis line (y=x) suggesting that our method of choosing genes for analysis is working...

What about the ”reactive” genes?

Note: Here I have used a much less stringent criteria for statistical significance in the TwinFat samples, in order to not miss any genes that are in fact reacting to obesity. This way we can increase the likelyhood that the genes left over really are not reacting to BMI, but rather (hopefully) ”*causal*”, or related to a causal process...



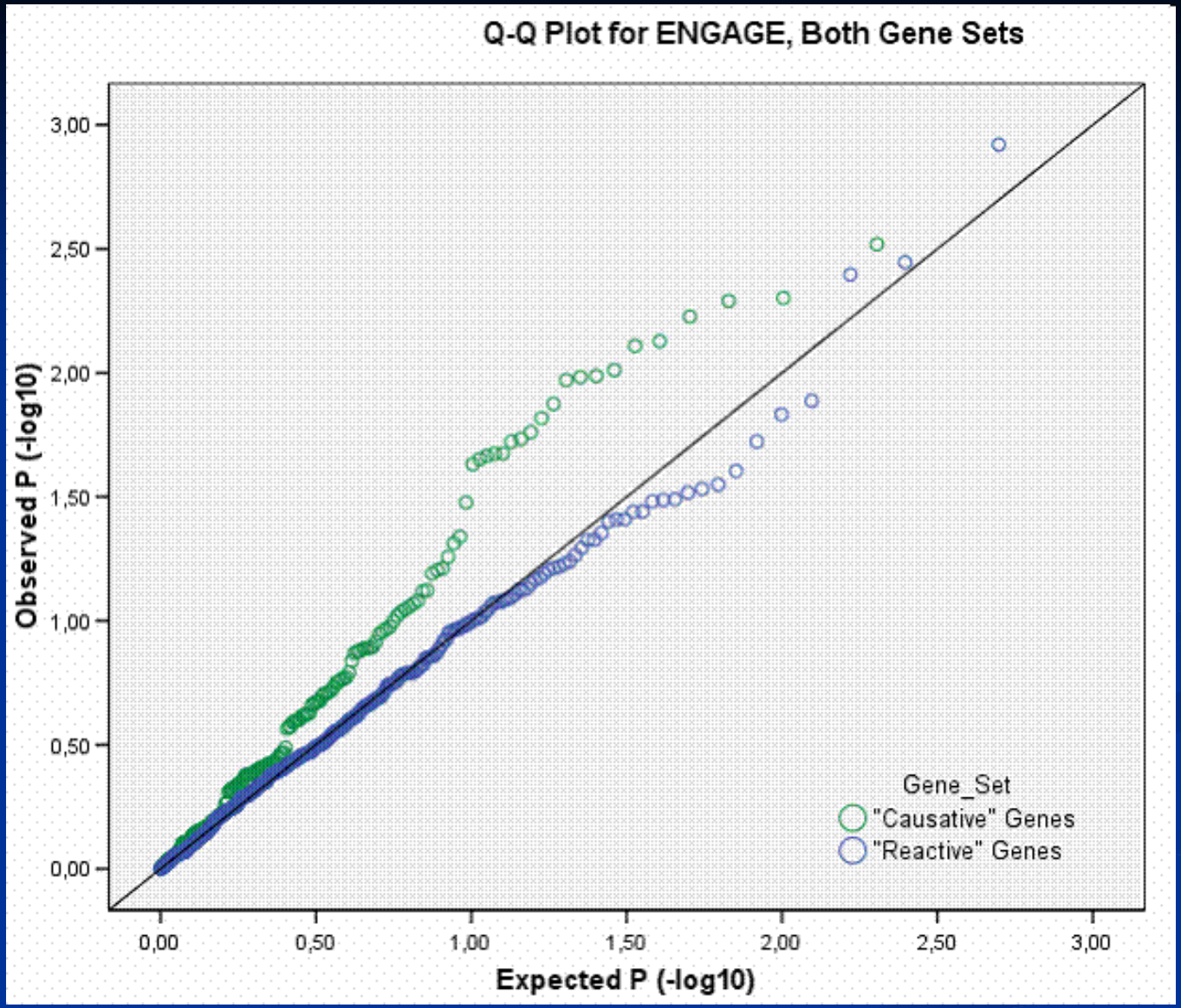
BMI *correlated* genes with p<0.0001

TwinFat BMI *reactive* genes with p<0.05

These 56 probes represent 43(?) unique genes/gene products

If the selection process has been successful, then the list of reactive genes should not yield as many significant associations as the “causative” genes list…

Taking the same gene sets and redoing the analysis in the ENGAGE population cohorts (N=21,000)…



Upon re-doing the analyses in the ENGAGE cohort, nearly 4 times as large, the trend strengthened –the set of genes that were associated with BMI has changed somewhat, but for example F13A1 is still the top hit among the “causative” genes

Conclusions

- Utilizing genome-wide expression data from these two different cohorts (MZ twins and unrelated samples) can succesfully identify BMI correlated genes whose expression is more likely to be under genetic, rather than environmental control in human samples.

- Using a list of candidate genes generated via this approach can in large scale GWA studies identify genetic polymorphisms that associate with the trait in question –here BMI.

- Here, of a total number of 27 candidate genes, 10 harbored SNPs that associated with BMI in the meta-analysis of 6,453 individuals.

- Many of these genes are known to play important roles in processes biologically relevant to the trait in question.

AKNOWLEDGEMENTS

Prof Aila Rissanen
Obesity Research

AssProf Kirsi Pietiläinen
Clinical Studies

Prof Leena Peltonen
Genetic Studies

MSc Jussi Naukkarinen
Genetic Studies

PhD Samuli Ripatti
Statistics

PhD Juha Saharinen
Bioinformatics

Prof Anu Wartiovaara
Mitochondria

NFBC, Genomeutwin
& Engage investigators

Dr Matej Oresic
Lipidomics

Prof Hannele Yki-Järvinen
Metabolic Studies

Jaakko.kaprio@helsinki.fi