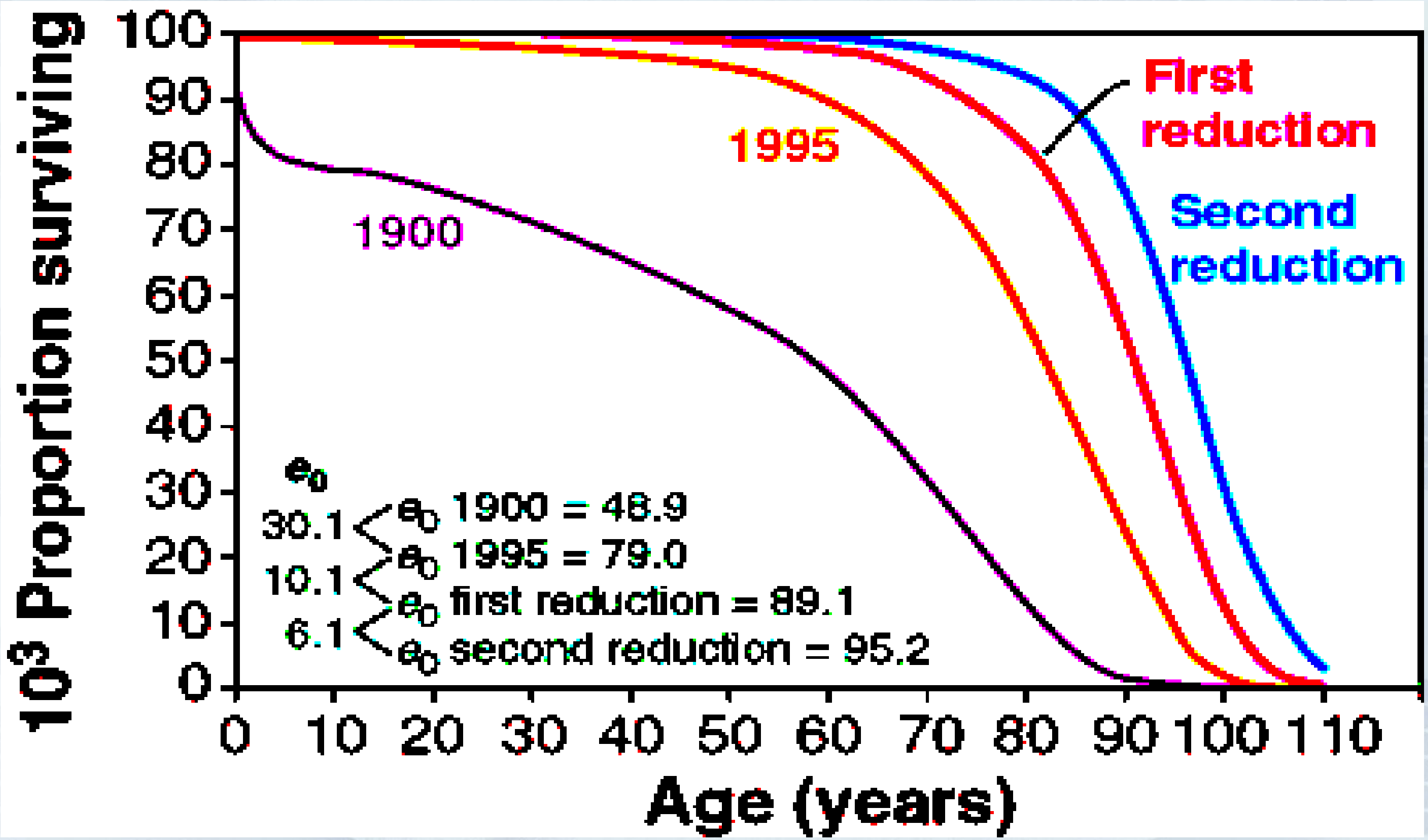
Gene & Environment Interactions: General principles and applications

Jaakko Kaprio

Professor of Genetic Epidemiology
University of Helsinki and National Public Health Institute

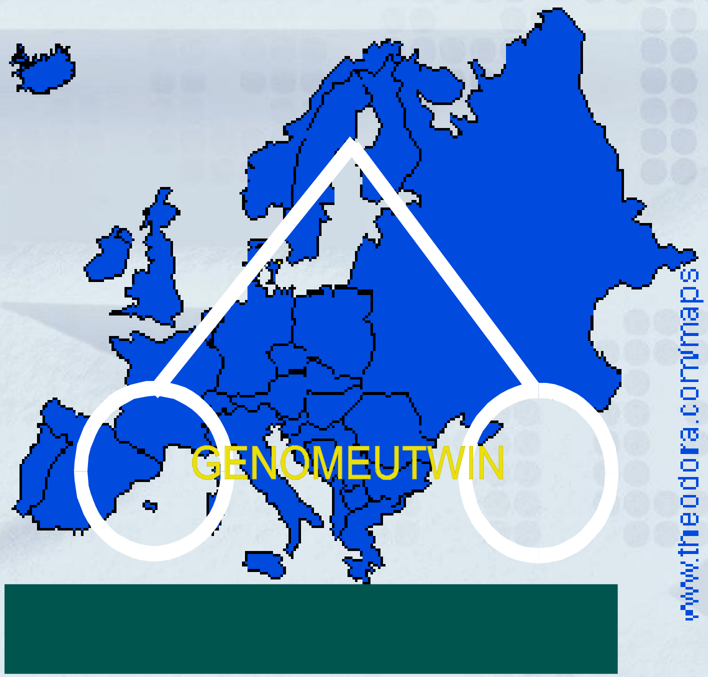
Development of life expectancy

(U.S females in 1990, 1995 and projected)



Olshansky et al., Science 2001; 291:1491

The European Consortium of Twin Registries for Analyses of Complex Traits -GENOMEUTWIN-



Twin similarity for life span at very old age

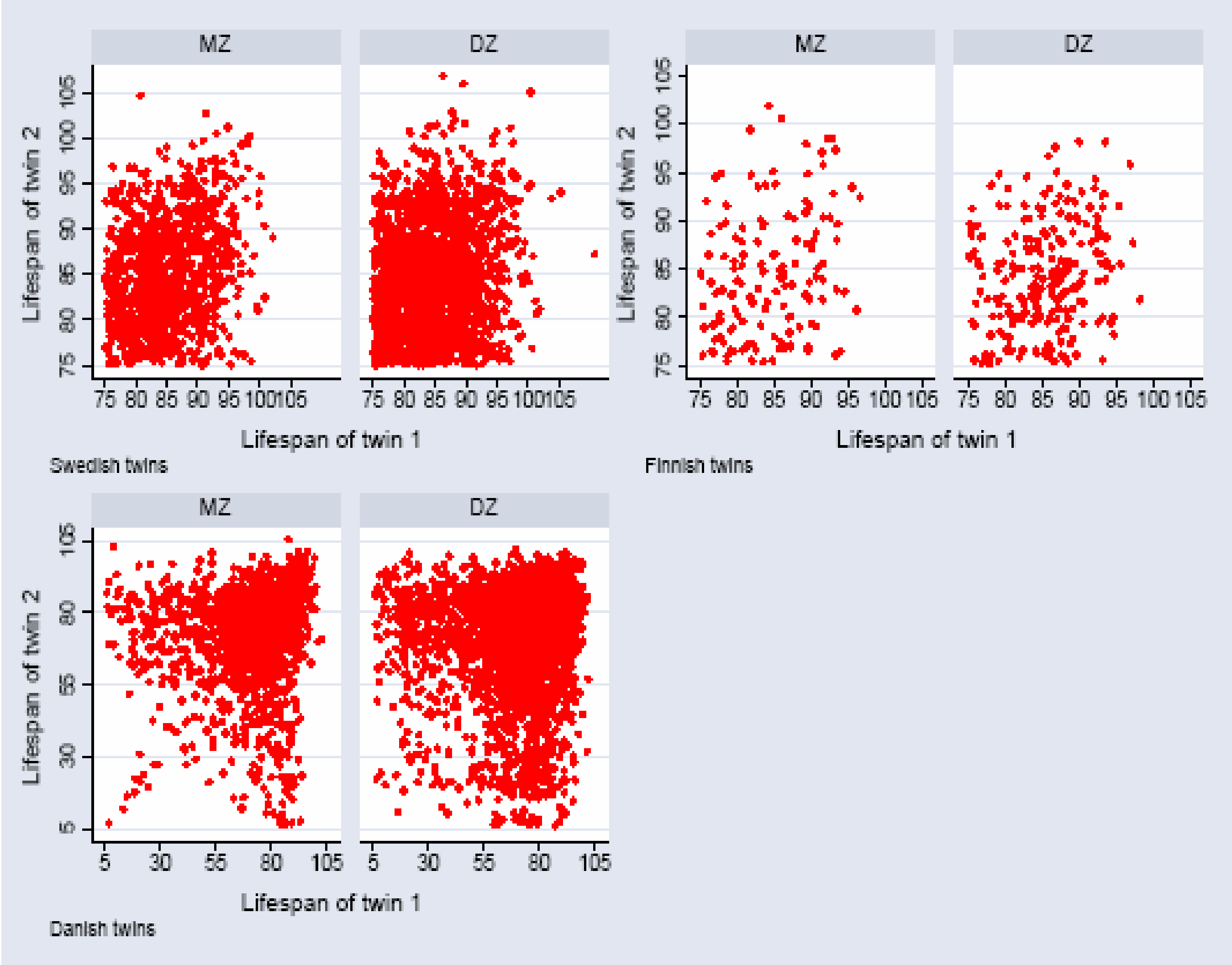


Figure 2: Twin-twin plot of the Swedish, Finnish and Danish cohort.

Model of covariances between relatives

Conceptual model of individual's phenotype:

$Y = \mu + G + C + E$

Hence, variance can be decomposed:

$\sigma^2 = \sigma^2 G + \sigma^2 C + \sigma^2 E$

Heritability is $\sigma^2 G/\sigma^2$ and genetic variance has several components:

$\sigma^2 G = \sigma^2 A + \sigma^2 D + \sigma^2 I$

For a relative pair:

$cov(Yij, Yik) = 2\varphi_{jk}\sigma^2 A + \Delta_{jk}\sigma^2 D + \varphi^2_{jk}\sigma^2 I + \gamma_{jk}\gamma\sigma^2 C + \delta_{jk}\sigma^2 E$

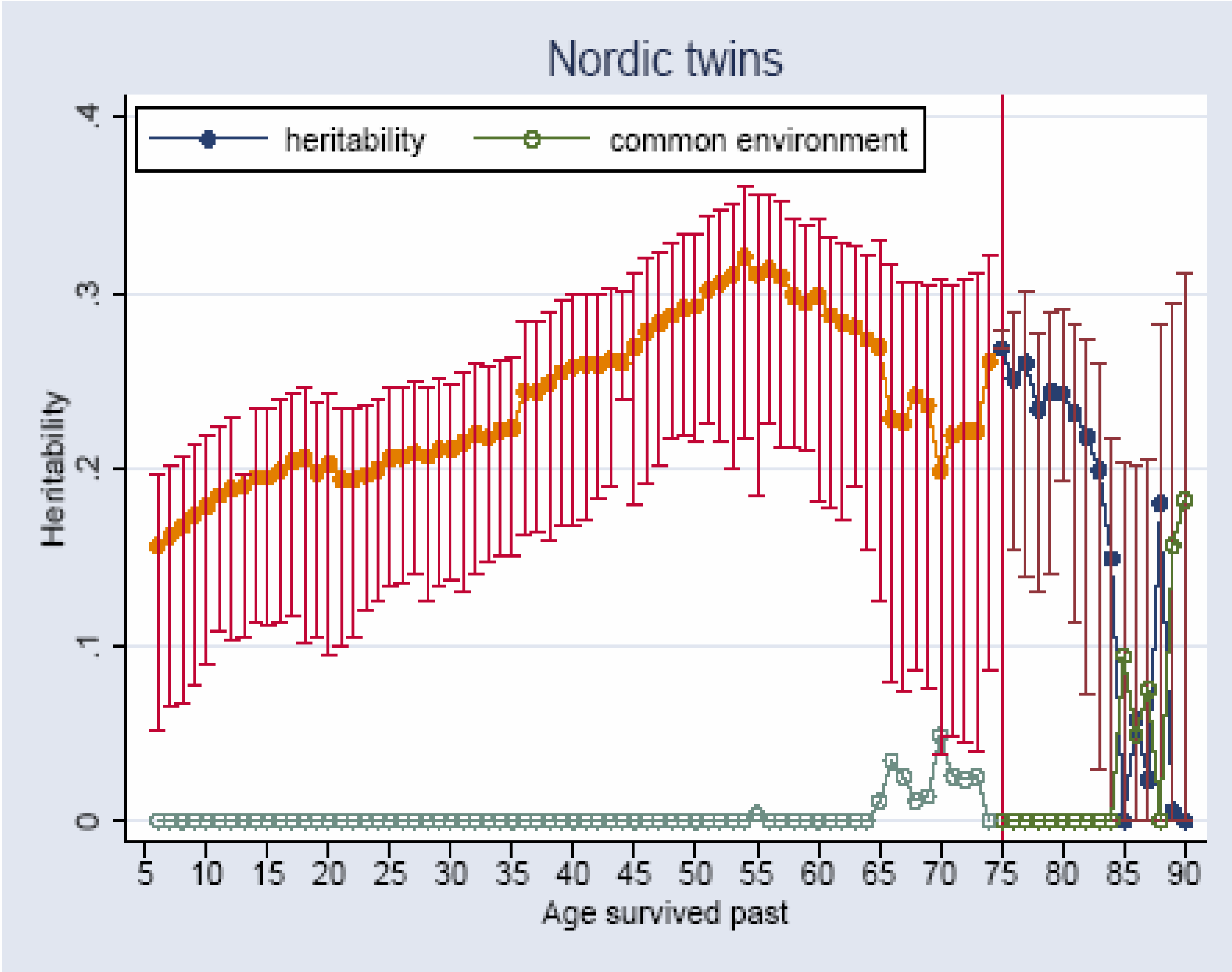


Figure 4: Heritability in lifespan of Danish, Finnish and Swedish twins by age survived past (obtained from ACE model). Estimates of amount of total variation due to common environmental effects are given as well. Swedish and Finnish twins enter when surviving past 75 years of age (full vertical line). Cohorts are combined assuming only equal standardized variance components across countries (see text). Vertical lines represent 95% confidence limits of heritability.

Hjelmborg et al, Human Genetics 2006

Better (?) conceptual model

$Y = \mu + G + Env$

Where $Env = C+E$

Hence, variance can be decomposed:

$\sigma^2 = \sigma^2 G + \sigma^2 C + \sigma^2 E$

Heritability is $\sigma^2 G/\sigma^2$ and genetic variance has several components:

$\sigma^2 G = \sigma^2 A + \sigma^2 D + \sigma^2 I$

Better model: $Y = \mu + G + Env + cov(G,Env)$

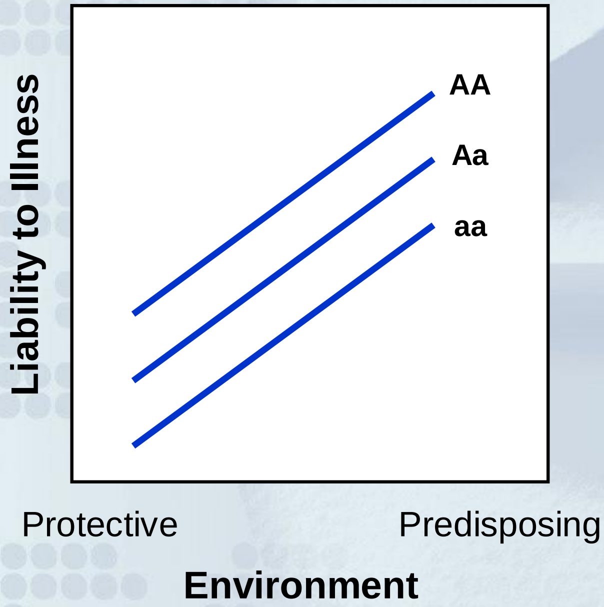
Traditional quantitative genetic analyses have usually assumed:

- no assortative mating
- **no gene-environment correlations (rGE)**
- no gene-environment interactions (GxE)

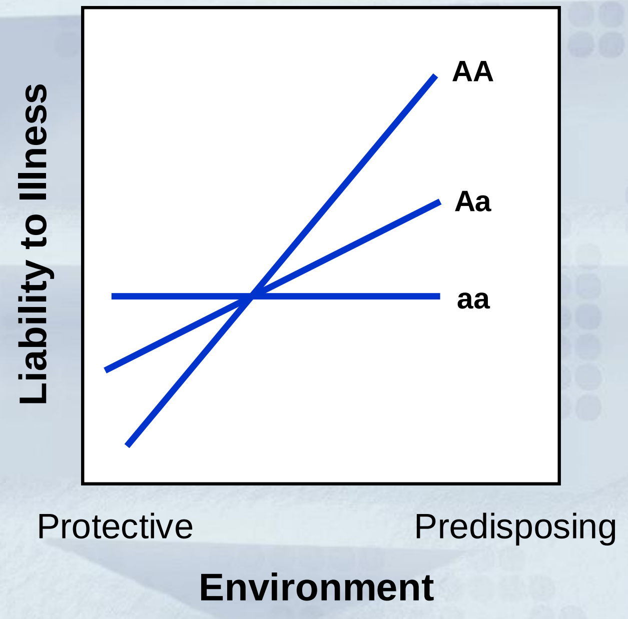
It is now clear that all three assumptions do not always hold, and we must consider a more complex interplay between nature and nurture.

Gene x Environment Interactions *Kendler & Eaves, 1986*

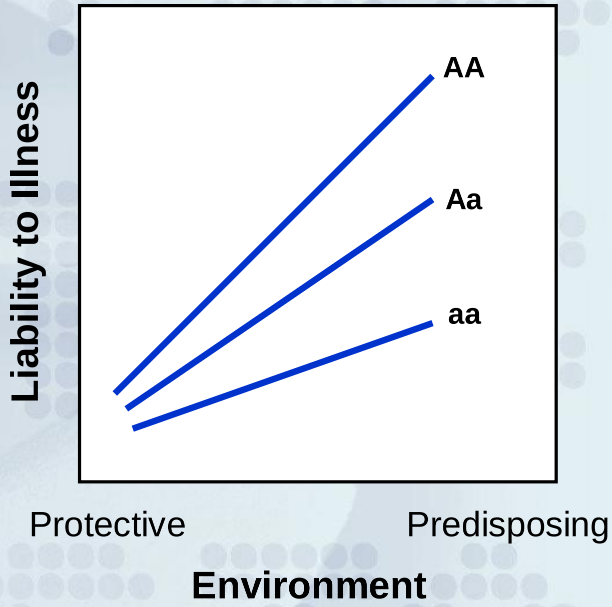
Genes and environment have additive, independent effects



Genes control degree of sensitivity to environmental influence



Genes control susceptibility to environmental pathogenesis



Gene-environment correlations refer to genetic effects on individual differences in **liability to exposure** to particular environmental circumstances.

(Background is the extensive evidence that environmental risk exposure is far from randomly distributed)

Gene-environment interactions concern genetically influenced individual differences in the **sensitivity** to specific environmental factors.

(Background is the extensive evidence of huge individual differences in vulnerability to all manner of environmental hazards)

GENE-ENVIRONMENT CORRELATIONS

Conceptualized under 3 main headings

1. Passive: meaning that genetically influenced characteristics of the **parents** help shape the rearing environments they provide for their children.
2. Active: meaning that genetically influenced characteristics of the individual serve to shape and select the environments they experience.
3. Evocative: meaning that genetically influenced characteristics of the **individual** serve to affect their interactions with other people and the responses that they elicit from others.

IMPLICATIONS OF ACTIVE / EVOCATIVE rGE

1. Some genetic effects are indirect, operating through effects on environmental risk
2. Insofar as the above apply, heritability estimates will be misleadingly high
3. Some effects of adverse environments are genetically mediated
4. Genes involved in individual differences in environmental risk exposure
5. Violations of Equal Environments Assumption (EEA) likely
6. Examples: genetic influences on Life events, Marital breakdown, Differential parenting, etc

From M. Rutter, 2002

Traditional quantitative genetic analyses have usually assumed:

- no assortative mating
- no gene-environment correlations (rGE)
- **no gene-environment interactions (GxE)**

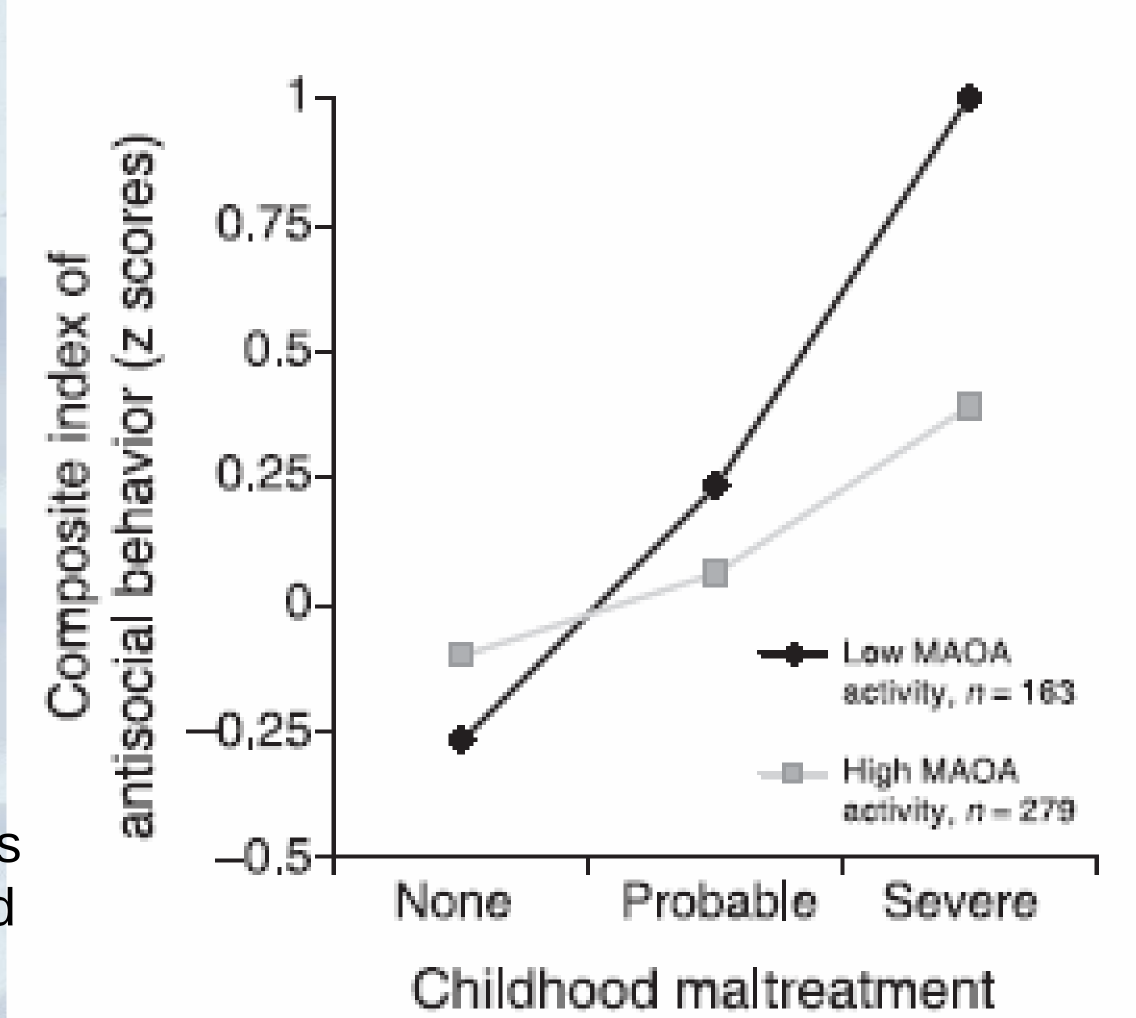
.

Examples of social x biological interactive effects

- Biology controls sensitivity to environment effects
  - E.g., family stress x serotonin metabolism => depression and anxiety risk (Caspi, Science 2003)
- Biology exaggerates social-context effects
  - E.g., amphetamines augment dominance behavior in dominant macaques and subordinate behavior in subordinates
- Social context generates undifferentiated risk; biology constrains pathologic specificity (
  - E.g., loss => depression in one child, but disordered conduct in another
  - E.g., childhood neglect => alcoholism in men, eating disorders in women

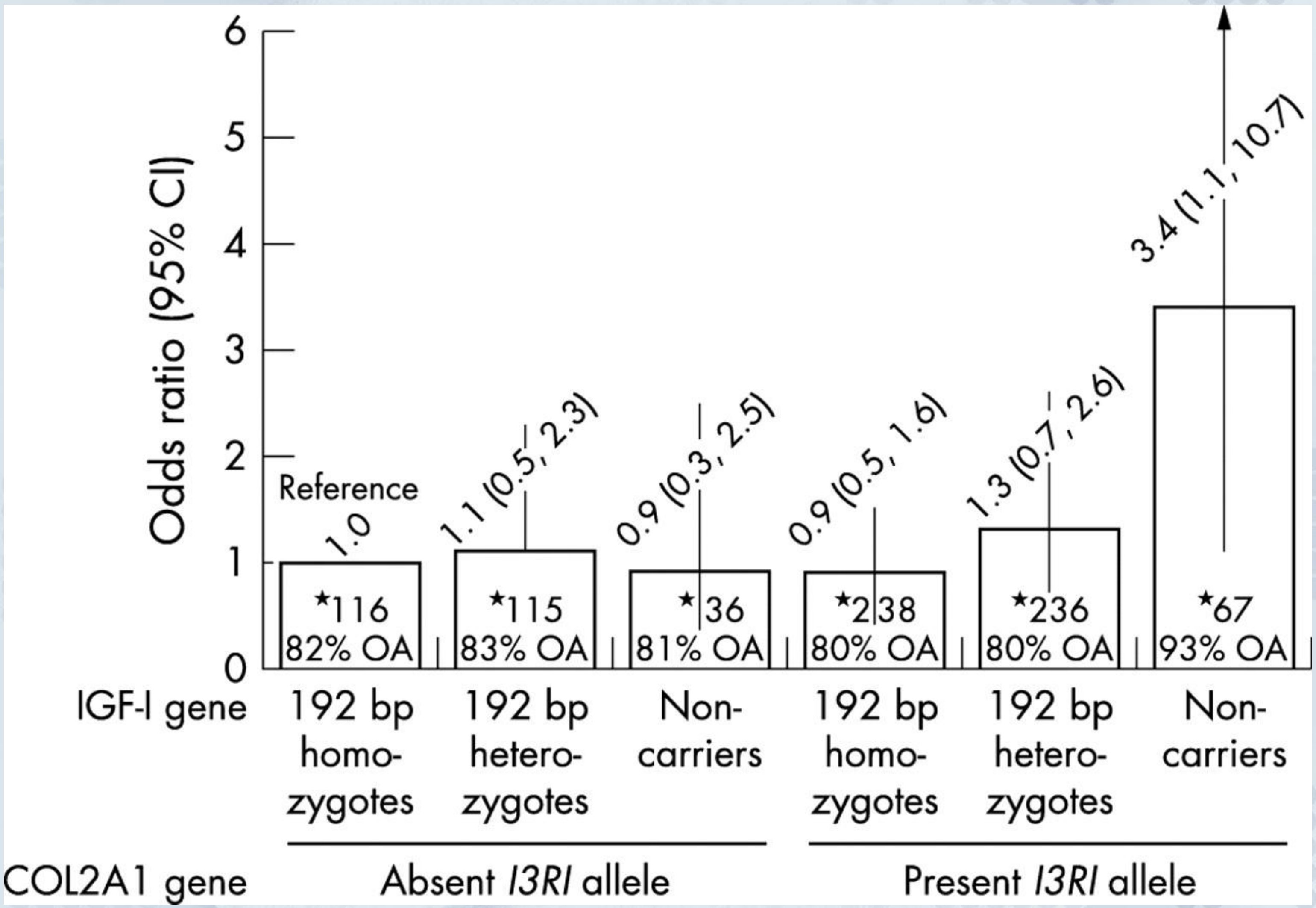
Examples

Means on the composite index of antisocial behavior as a function of *MAOA* activity and a childhood history of maltreatment. *MAOA* activity is the gene expression level associated with allelic variants of the functional promoter polymorphism, grouped into low and high activity; childhood maltreatment is grouped into 3 categories of increasing severity. The antisocial behavior composite is standardized (*z* score) to a M 0 and SD 1; group differences are interpretable in SD unit differences (d).



Caspi et al, Science 2002

Figure 1 Interaction between the IGF-I and COL2A1 genes in the occurrence of radiographic osteoarthritis (ROA). Risk estimates adjusted for age, sex, body mass index, and bone mineral density. *Total number by genotype and percentage of subjects with ROA. bp, base pair; CI, confidence interval; OA, osteoarthritis.



Zhai, G et al. Ann Rheum Dis 2004;63:544-548

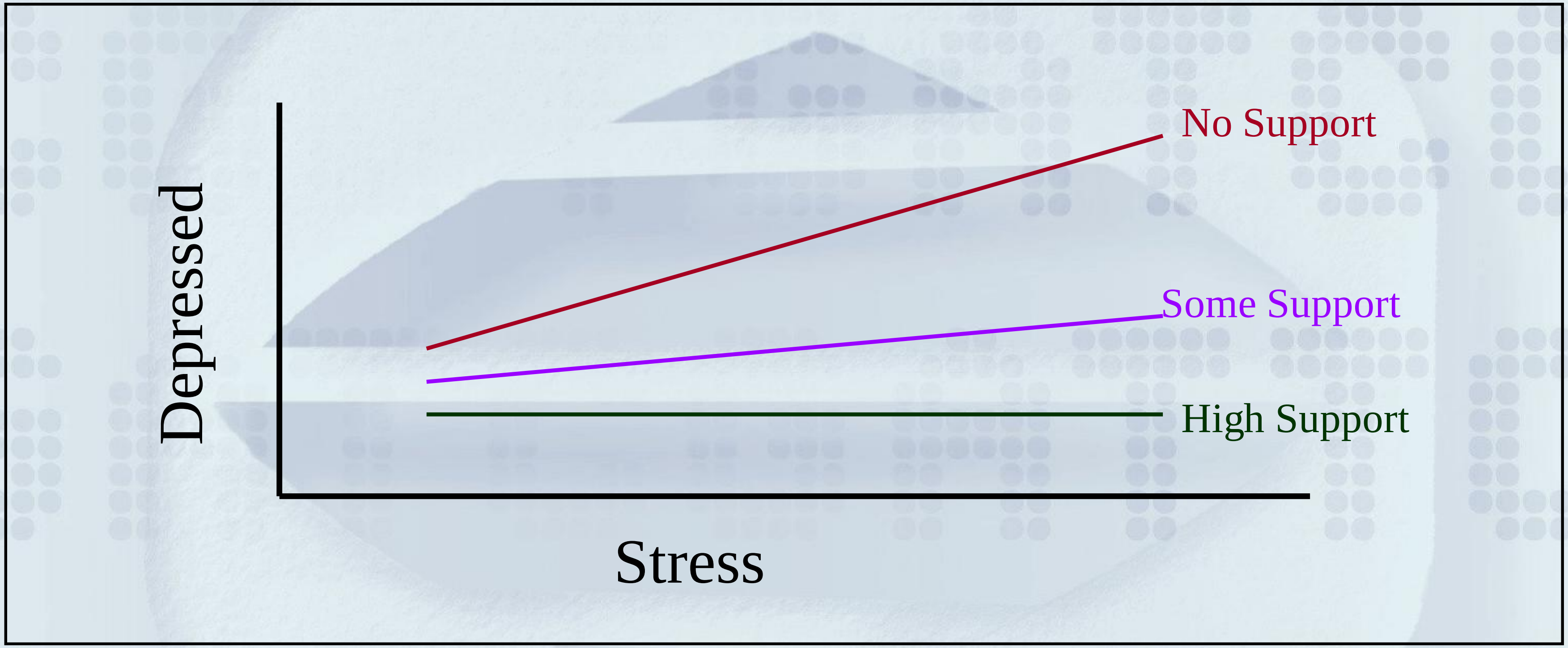
Statistical Interaction

- The outcome – exposure relationship is not constant in different conditions
  - Regression lines are not parallel

- In many studies of perceived social support, stress and depression, a three-way picture is found:
    - For persons who are not experiencing an important stress, the relation between social support and depression is small or zero.
    - For persons who have a modest amount of stress, social support seems to reduce risk of depression/level of depressiveness
    - For persons with high stress, social support has large inverse association with depression
  - This is called the stress-buffering effect of support.

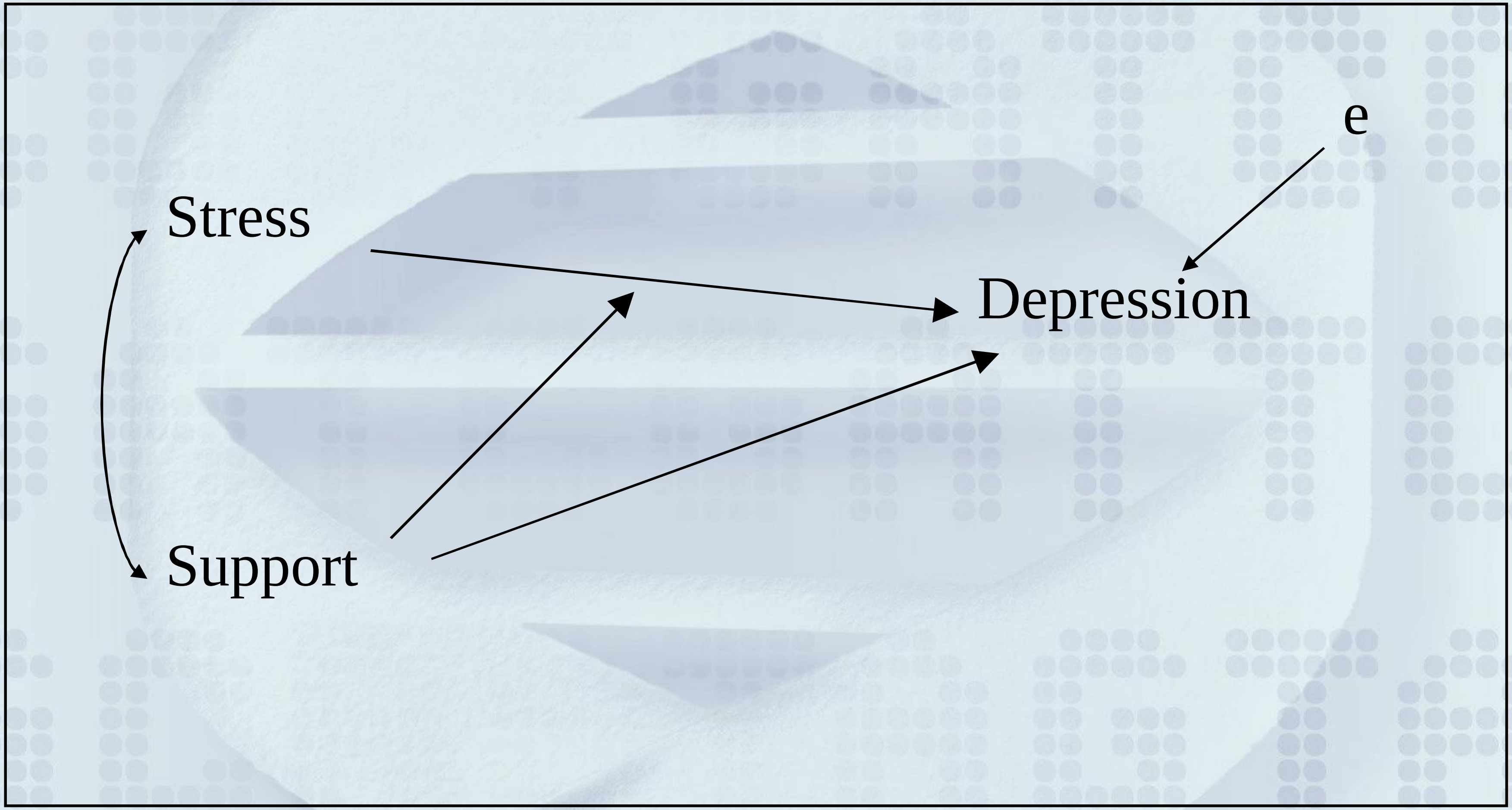
- Romanov et al, Acta Psychiatrica Scand 2003 for example from the Finnish Twin Cohort study

Two Pictures of Interaction



Regression coefficients of depressiveness on stress differ by level of social support

Two Pictures of Interaction



- The moderating effect of support on the stress – depression relationship

Representing Interaction in the Regression Equation

- One way to model how the effect of X_1 is systematically affected by X_2 is to include a multiplicative term in the regression equation
  - $Y=b_0+b_1X_1+b_2X_2+b_3(X_1*X_2)+e$

Interpreting the Multiplicative Model

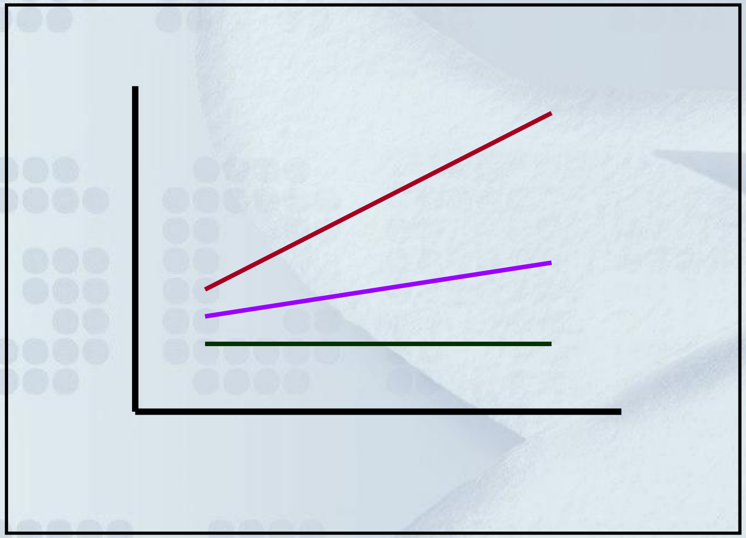
- $Y=b_0+b_1X_1+b_2X_2+b_3(X_1*X_2)+e$
- The effect (slope) of X_1 varies with different values of X_2
  - For $X_2=0$, the effect is b_1
  - For $X_2=1$, the effect is b_1+b_3
  - For $X_2=2$, the effect is b_1+2b_3

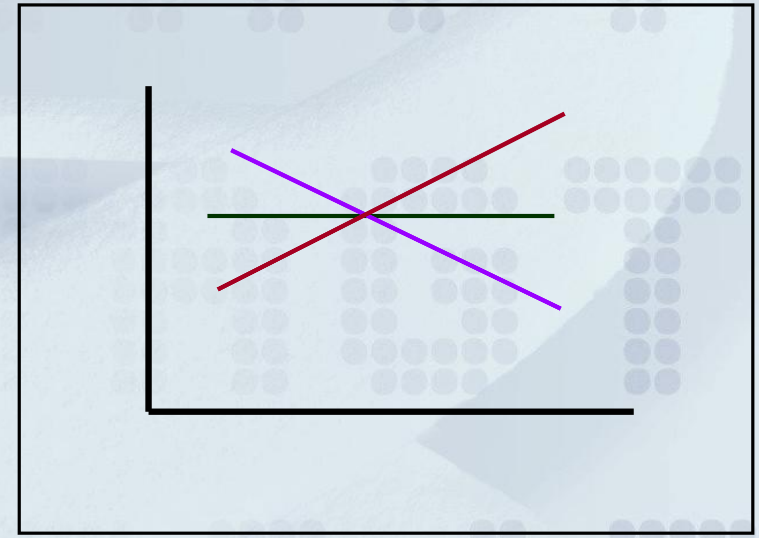
Scaling

- Scaling of the outcome variable can affect whether an interaction term is needed.
- If we have a simple multiplicative model in Y, it will be additive in Ln(Y)
    - $E(Y|X_1*X_2) = bX_1*X_2$
    - $E(\ln(Y)|X_1*X_2) = \ln(b)+\ln(X_1)+\ln(X_2)$
- When the variance seems to be related to the level of Y, the hypothesis of interactions being simple scaling functions needs to be considered

Scaling

- Scaling is especially important if the trajectories of interest do not cross in the region where data is available.





Interaction in Epidemiology

- Definition

  "Interaction is present when the incidence rate of disease in the presence of two or more risk factors differs from the incidence rate expected to result from their individual effects." -- MacMahon

Interaction
Additive Model

Disease Rates

Factor B \ Factor A	Absent	Present
Absent	3.0	9.0
Present	15.0	?

Factor B \ Factor A	Absent	Present
Absent	3.0	9.0
Present	15.0	21.0

Risk Difference

Factor B \ Factor A	Absent	Present
Absent	0	6.0
Present	12.0	18.0

Interaction
Multiplicative Model

Disease Rates

Factor B	Factor A: Absent	Factor A: Present
Absent	3.0	9.0
Present	15.0	?

Factor B	Factor A: Absent	Factor A: Present
Absent	3.0	9.0
Present	15.0	45.0

Relative Risks

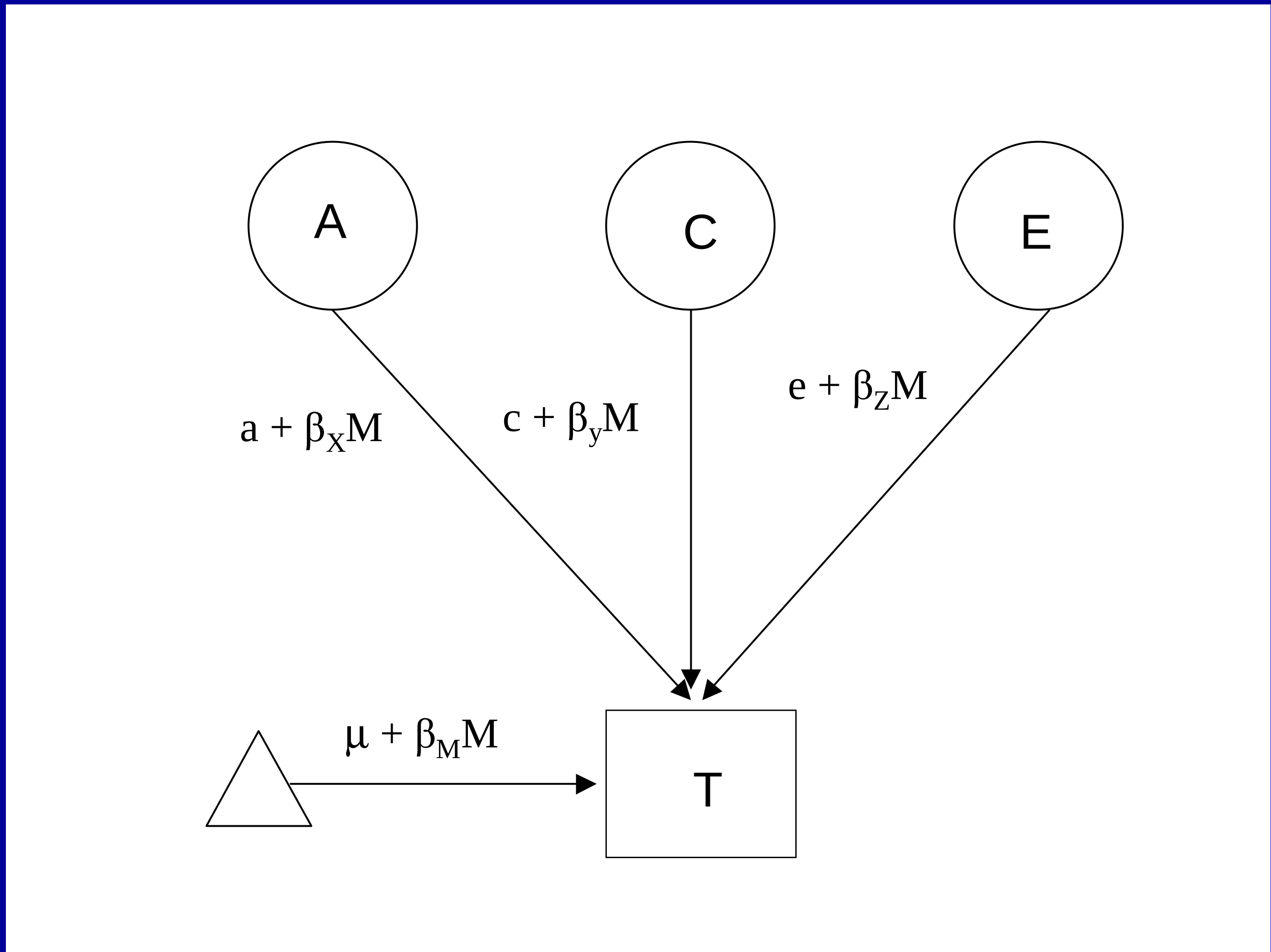
Factor B	Factor A: Absent	Factor A: Present
Absent	1.0	3.0
Present	5.0	15.0

More examples of biological and social/environmental interactions

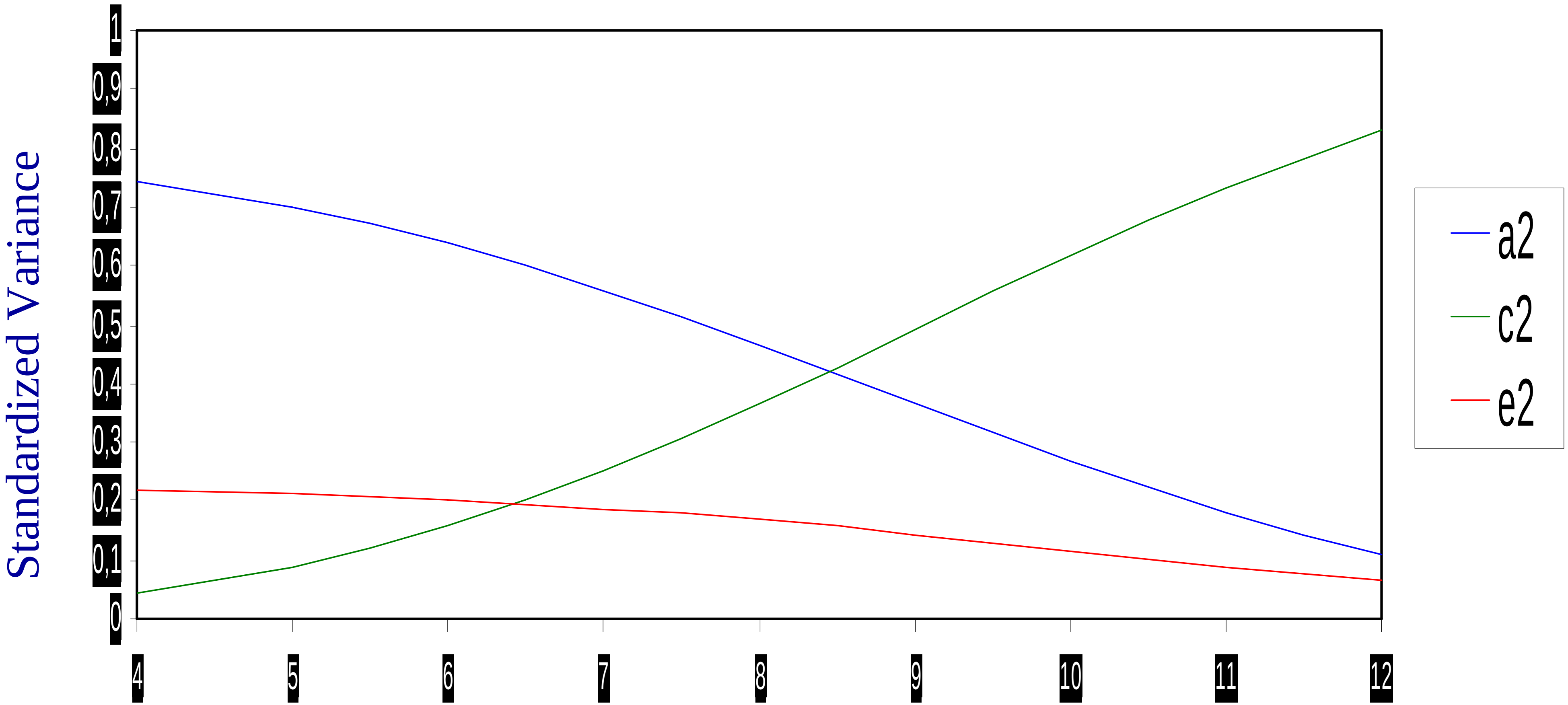
- Biological susceptibilities are amplified during rapid or intense contextual change
  - E.g., biological or gender-based vulnerabilities to depression during middle school transition as indexed by pubertal development

- Biology controls liability to experiencing predisposing environments
  - E.g., genes for skin color

More Complex Models
Main & Moderating Effects



Parental Monitoring and Smoking Quantity in 14 year olds



Parental Monitoring

Dick et al, J Abn Psychol, 2006

Traditional quantitative genetic analyses have usually assumed:

no assortative mating

no gene-environment correlations (rGE)

no gene-environment interactions (GxE)

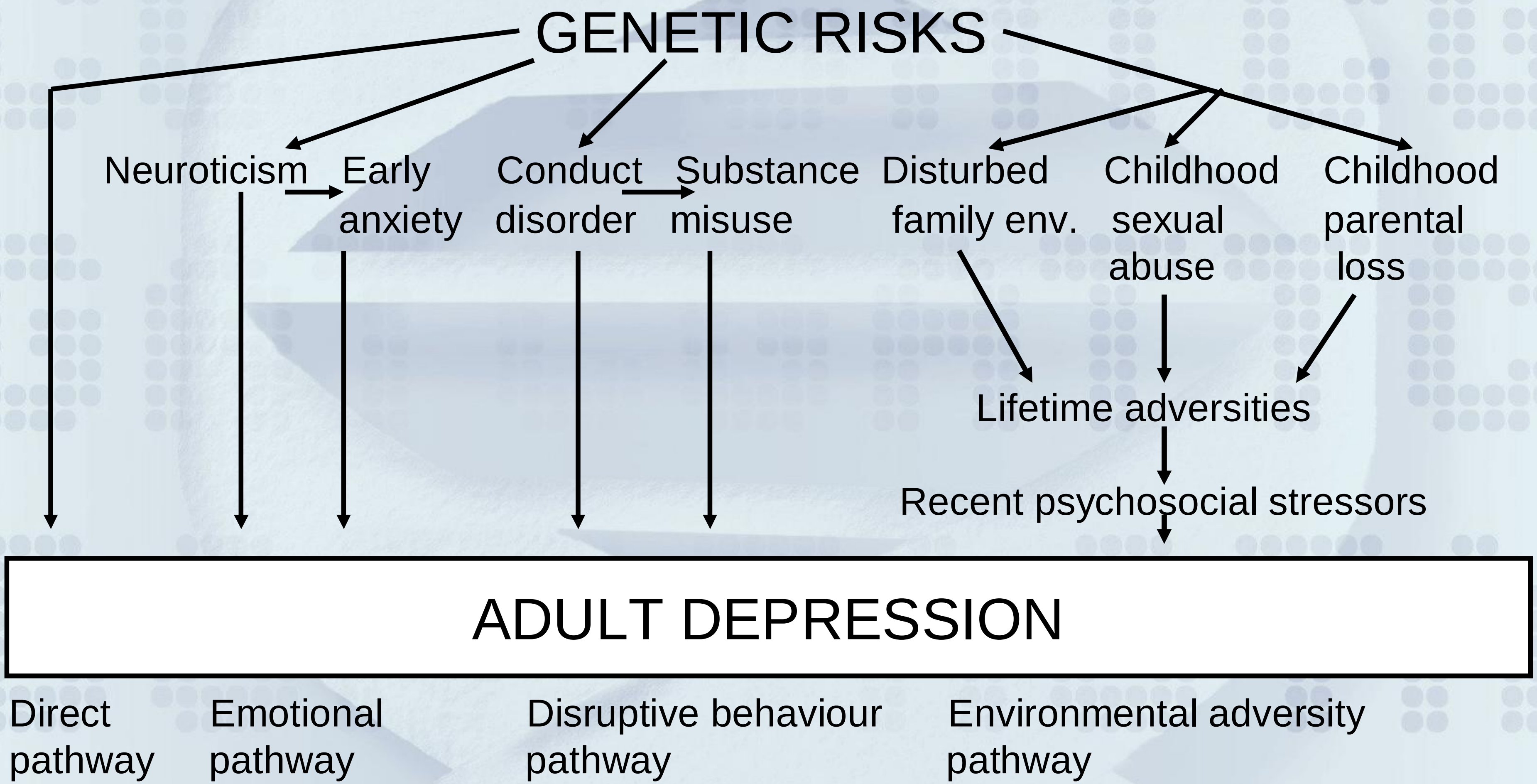
ASSORTATIVE MATING

First causes biasing effects on estimates of h^2 (downwards in twin studies and upwards in adoptee studies)

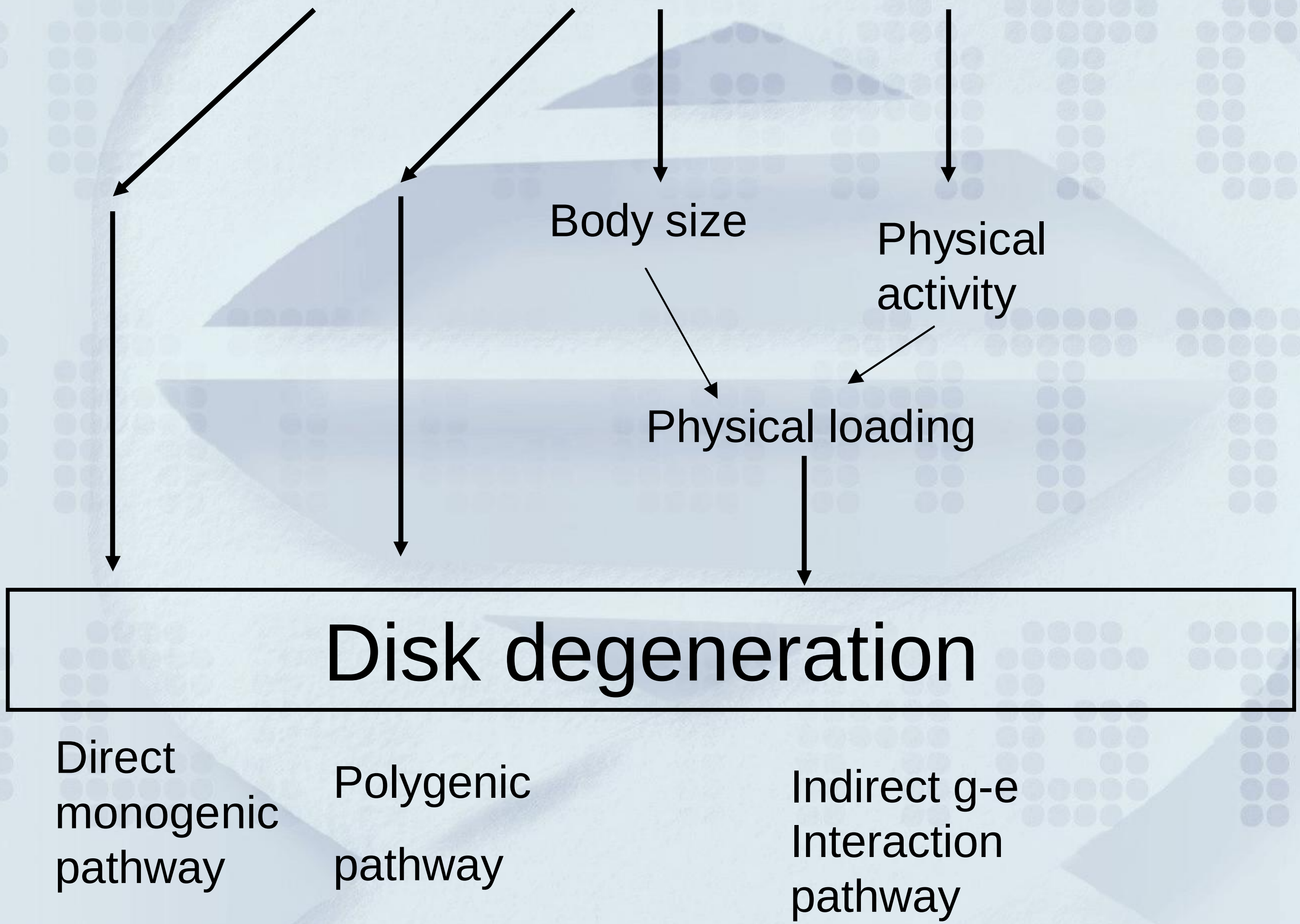
Also, it constitutes G effects (from the parent) on E as experienced by the child.

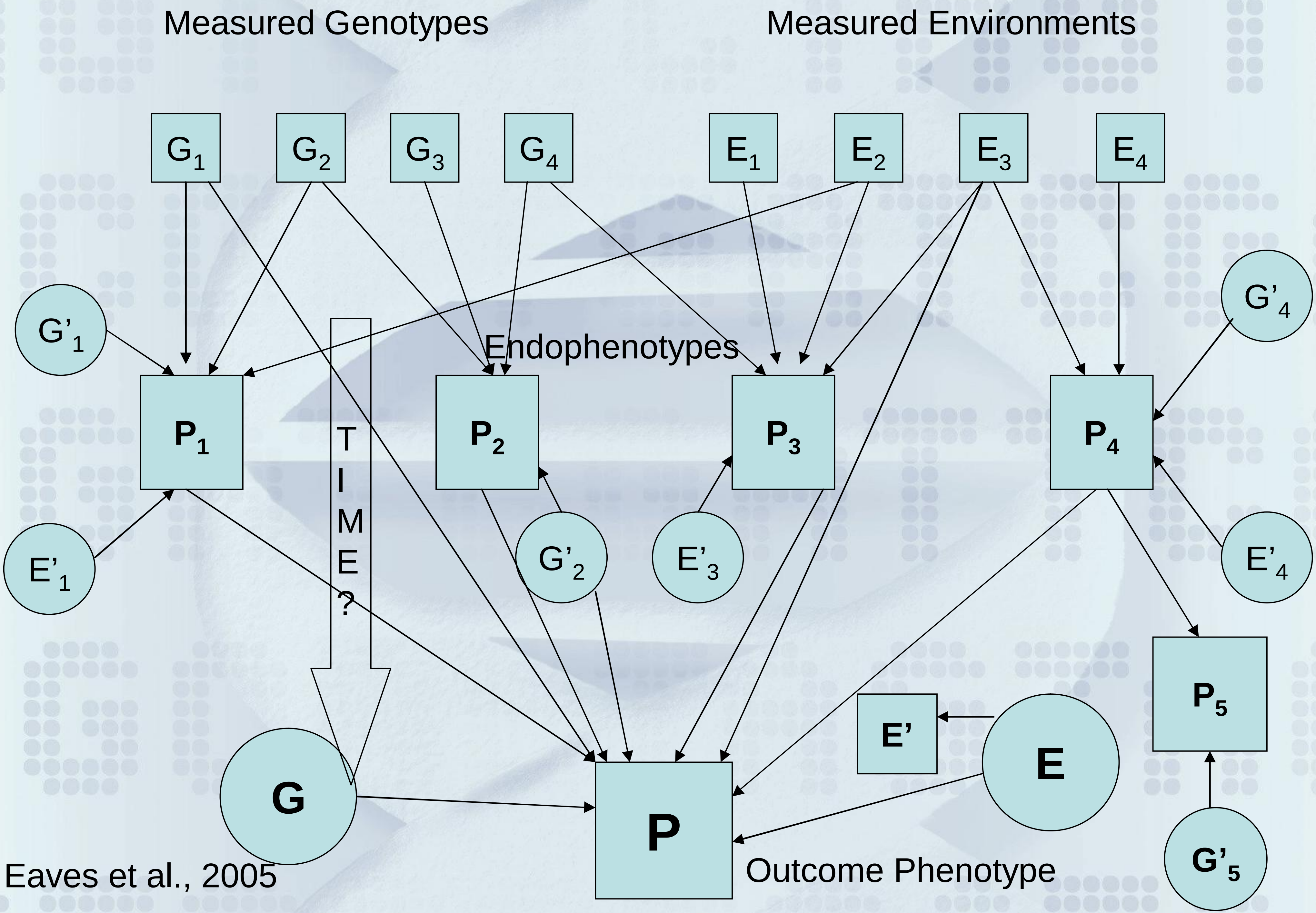
SIMPLIFIED CAUSAL PATHWAYS FROM GENES TO ADULT DEPRESSION

(from Kendler et al., 2002)



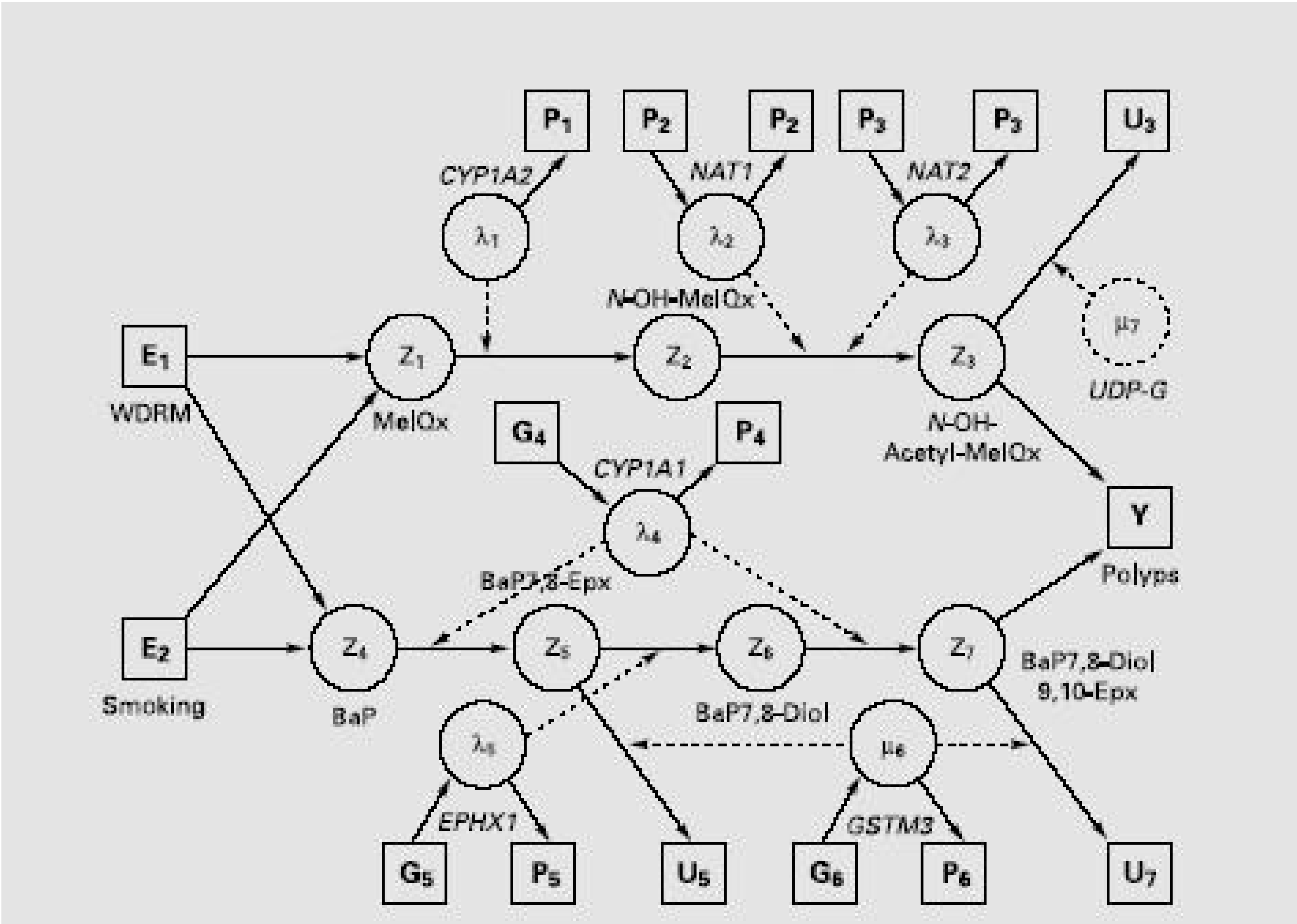
Genetic risks





Eaves et al., 2005

Pathway approach



Hung et al. Cancer Epid Biomarker Prev 2004 &
Conti et al. Human Heredity 2003

IMPLICATIONS OF rGE & GxE

<u>For research</u>

i) Analyses of genetic effects must take account of environmental risk exposure and evaluate rGE and GxE.

ii) Analyses of environmental effects must take account of genetic risks and evaluate rGE and GxE.

iii) Molecular epidemiology is crucial to study nature-nurture interplay.

From M. Rutter, 2002

IMPLICATIONS OF rGE & GxE

For clinical work

i) Conceptualization of disorders into those that are due to stress/adversity and those that are genetically determined is seriously misleading.

ii) People shape and select their environments as a result of how they behave, such behavior being genetically influenced.

iii) The huge individual differences in environmental risk exposure is due, in part, to rGE.

iv) The huge individual differences in response to stress/adversity is due, in part, to GxE.

From M. Rutter, 2002

An intensive four-day protocol

(Kirsi Pietiläinen, Leila Karhunen, Aila Rissanen, Anssi Sovijärvi and Hannele Yki-Järvinen)

MZ

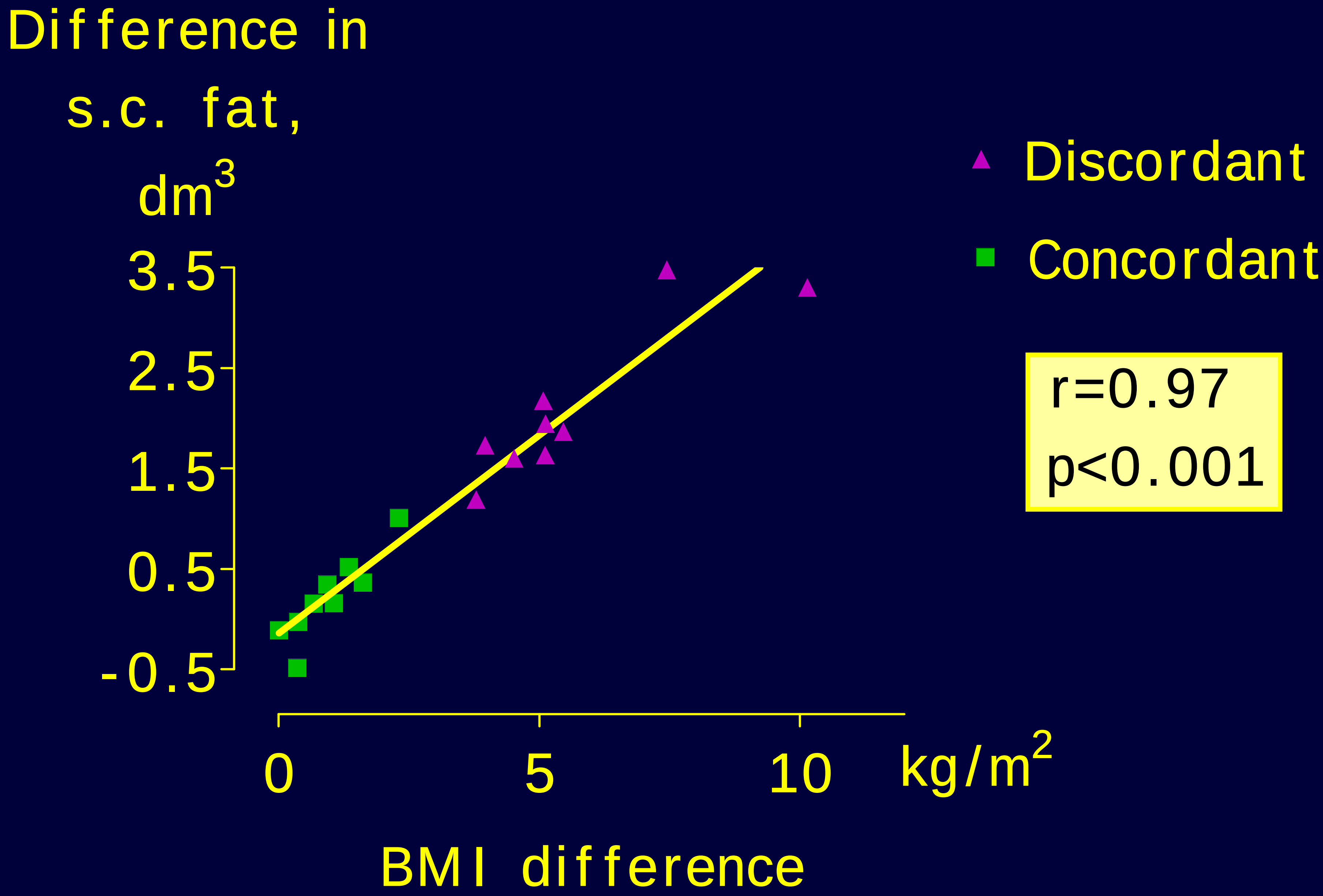
- Discordant twin pairs: 15 MZ twin pairs from FinnTwin16 cohort who differ for their BMI by 4-10 BMI units, i.e. 10-25 kg; one co-twin normal weight with a BMI~25 kg/m2 & one co-twin obese with a BMI~30 kg/m2.
- Concordant twin pairs: 6 MZ obese pairs with similar BMIs, and 5 MZ normal weight pairs with similar BMIs.

DZ

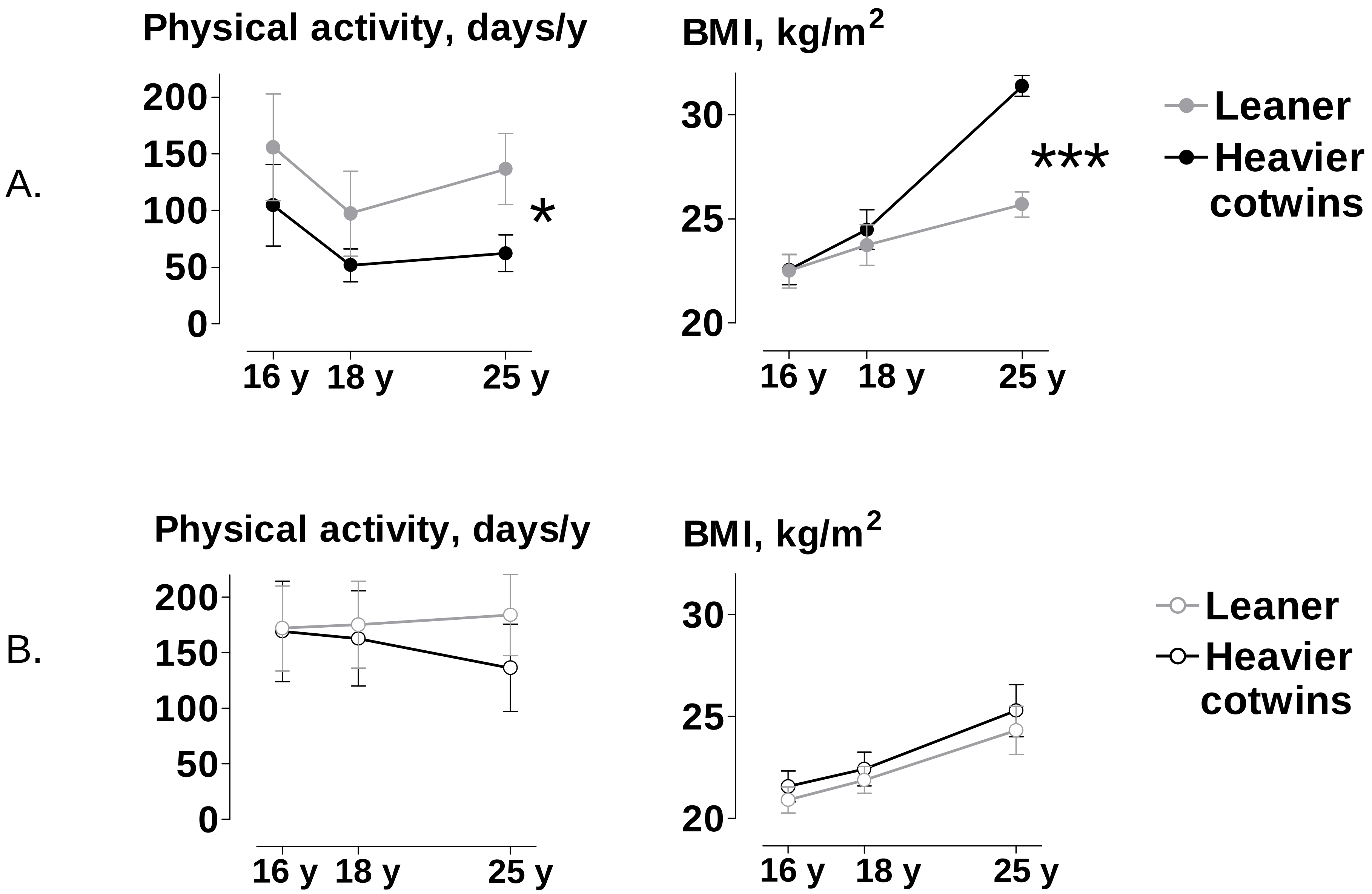
- Discordant twin pairs: 11 DZ twin pairs who differ for their BMI by 5-17 BMI units, i.e. 16-43 kg.

- Total to date: 74 subjects

Acquired obesity - increased s.c. fat

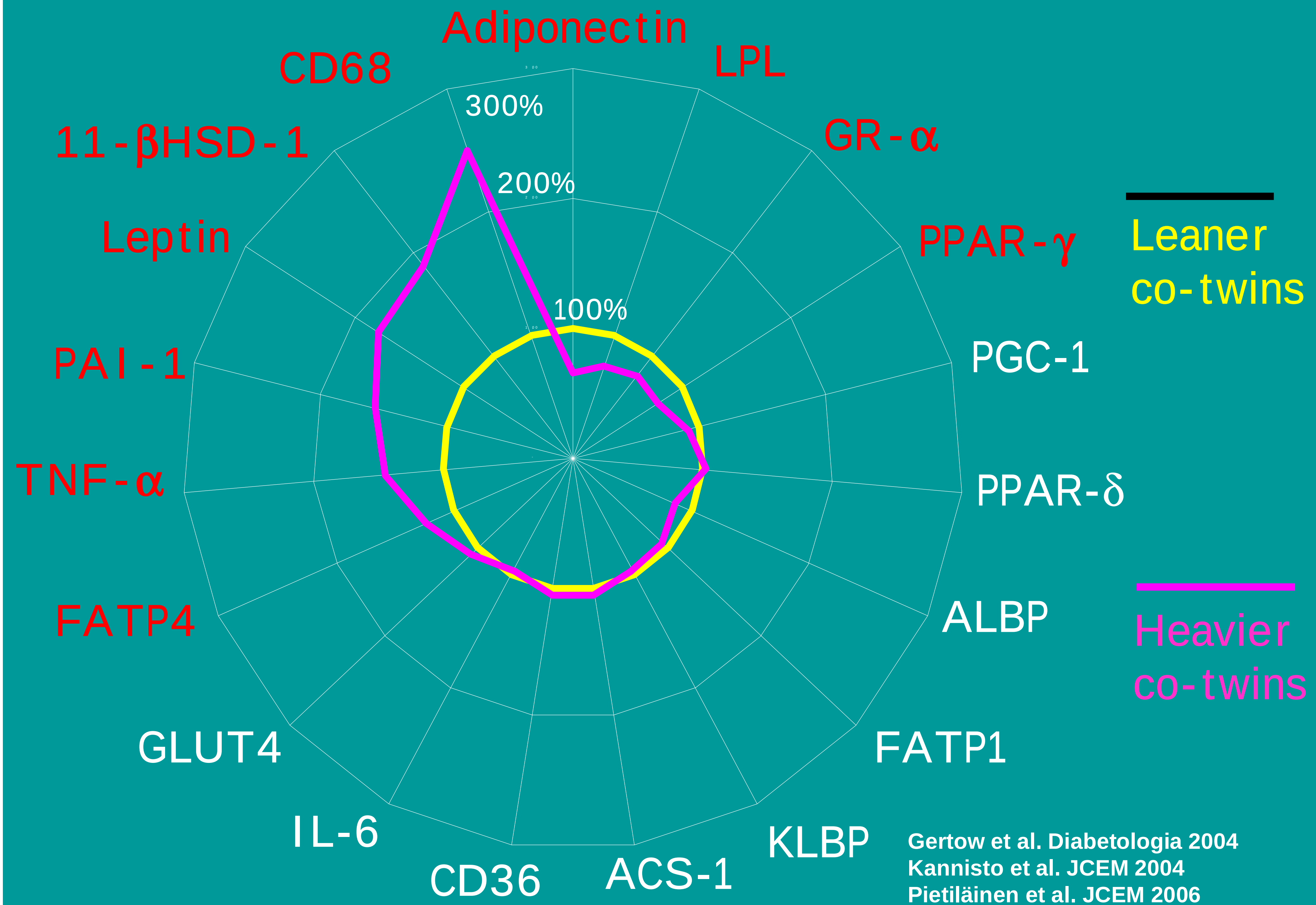


Physical activity and weight from adolescence to young adulthood in MZ twin pairs discordant (A) and concordant (B) for obesity as young adults



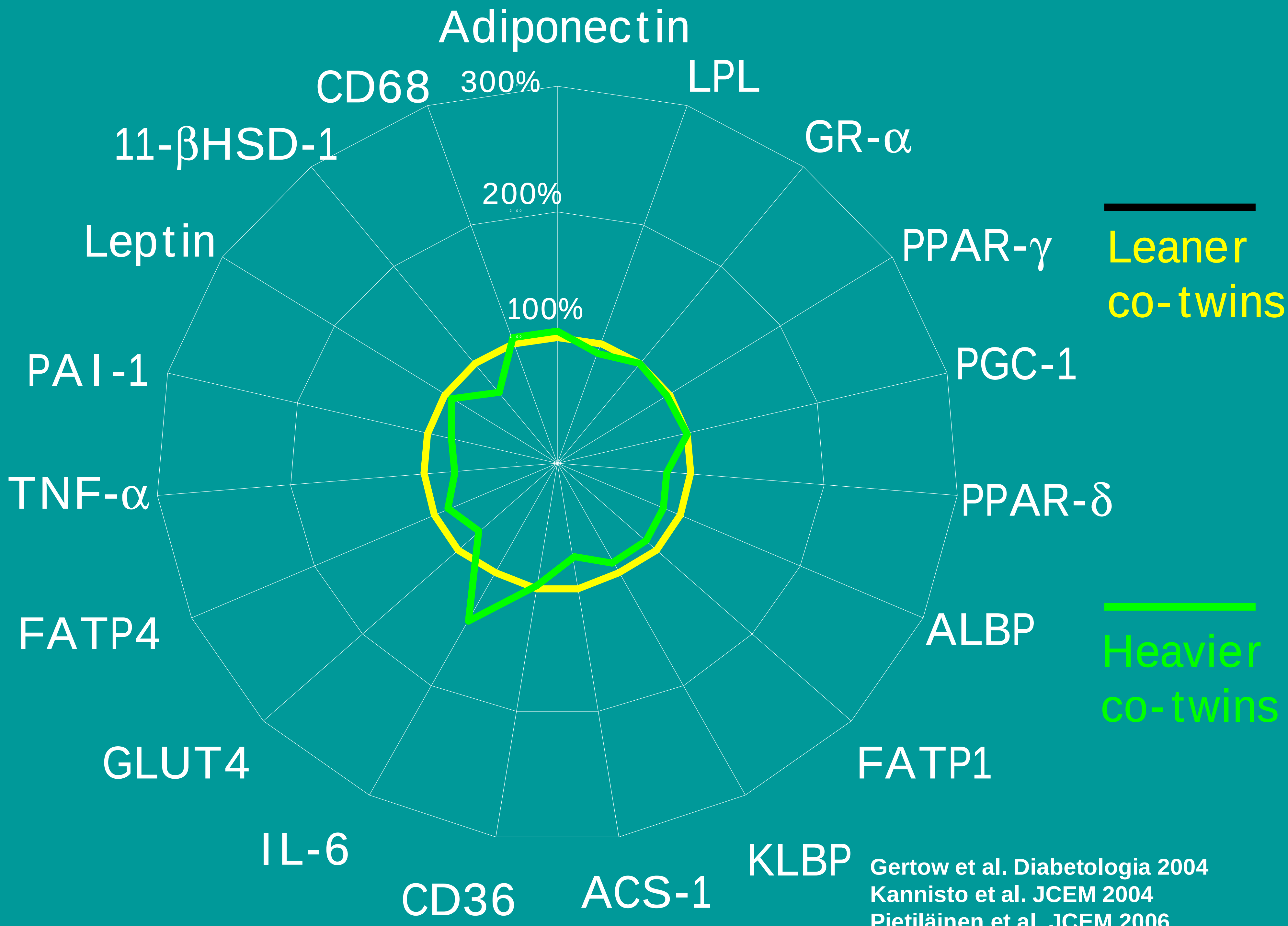
Pietiläinen et al, submitted

Discordant pairs: mRNA expression



Gertow et al. Diabetologia 2004
Kannisto et al. JCEM 2004
Pietiläinen et al. JCEM 2006

Concordant pairs: mRNA expression



Gertow et al. Diabetologia 2004
Kannisto et al. JCEM 2004
Pietiläinen et al. JCEM 2006

Discordance for smoking in MZ pairs

(picture in BMJ Oct.6,2001 issue)

In the smoker:
- rapid aging
- more premature deaths
-more cancers
-more atherosclerosis
-more intervertebral disc degeneration
- more coronary heart disease

Examples of developmental differences

- Pre- and postnatal effects
  - Maternal nutrition and well-being
  - Childhood living conditions
  - Individual lifestyle and occupational choices & circumstances
- Post-genomic changes
  - Somatic mutations
  - Epigenetic effects

MZ twin concordance in different diseases and age of onset

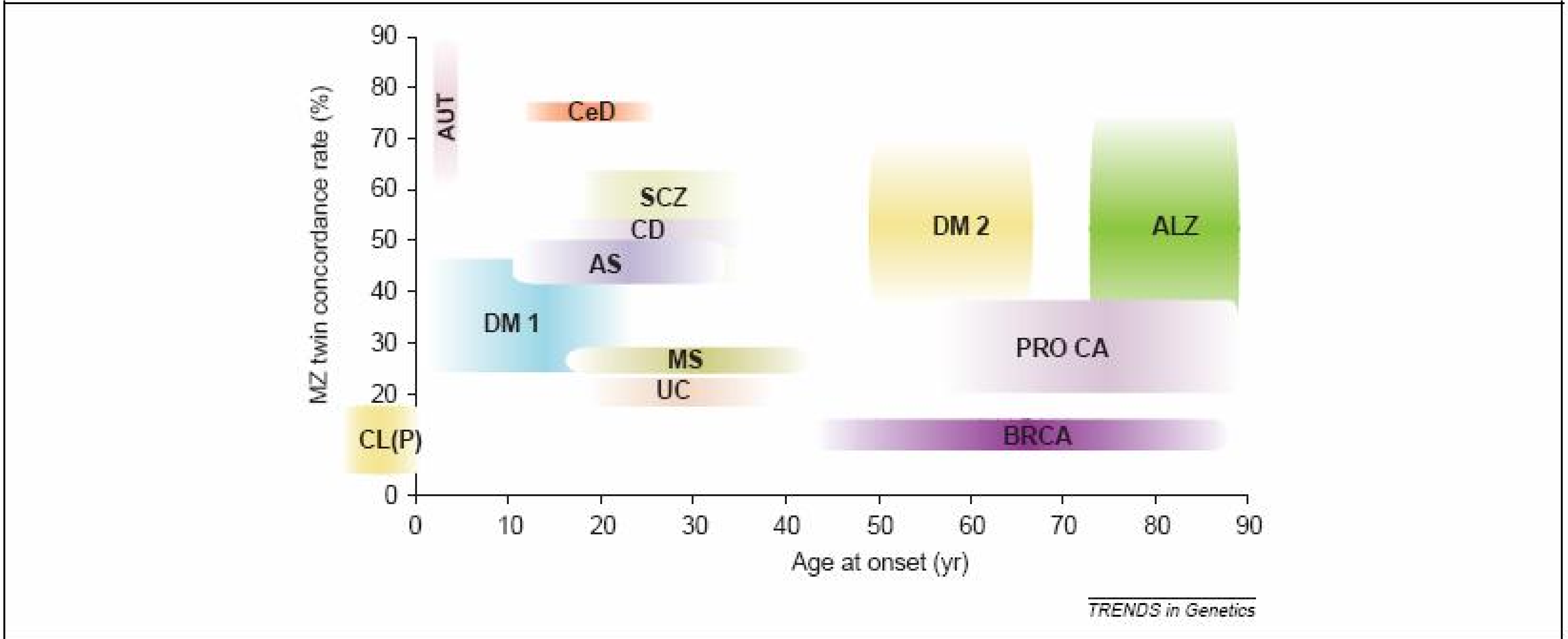


Figure 1. The age at disease onset and concordance rates of MZ twins. The left and right margins of each shape represent the range of age at onset, and the upper and lower margins represent the lowest and the greatest concordance rates of each disease. The data is taken from least two independent twin studies. Cleft lip and palate [CL(P)] occurs during embryogenesis and therefore placed to the left of the graph. Abbreviations: AUT, autism; CeD, celiac disease; SCZ, schizophrenia; CD, Crohn's disease; AS, asthma; DM1, type 1 diabetes mellitus; DM2, type 2 diabetes mellitus; MS, multiple sclerosis; UC, ulcerative colitis; BRCA, breast cancer; PRO CA, prostate cancer; ALZ, Alzheimer's disease. Primary references of twin concordance rates are available from the author.

Petronis A. Epigenetics and twins: three variations on the theme. Trends in Genetics (in press, 2006)

Epigenetic effects have been demonstrated to be associated with trait discordance in MZ pairs (e.g. Fraga et al, PNAS 2005), but quantification of effect is not established