Molecular & Genetic Epidemiology

Kostas Tsilidis, PhD

ktsilidi@uoi.gr / k.tsilidis@imperial.ac.uk

@KTsilidis

Gene-Environment Interaction

Learning Objectives

- Understand how studying gene-environment interaction may contribute to our understanding of disease aetiology and illustrate this concept using at least one established example.

- Recognise the major challenges in studying gene-environment associations.

Outline

- Gene-Environment Interaction
  - Conceptual Overview
  - Rationale
  - Challenges
  - Study designs
  - Established Examples

Definitions of gene-environment interaction

- *"Variation in the measure of effect of an environmental risk factor on an outcome according to genotype"*
- *"Joint effect of one or more genes with one or more environmental factors that cannot be readily explained by their separate marginal effects"*
- Examples: Individuals with different genotypes could differ in terms of:
  - Susceptibility to the health effects of exposures such as diet, smoking, drinking, sedentary lifestyle, etc.
  - Responses to life events such as trauma
  - Responses to medications (pharmacogenomics)

Thomas D. Nat Rev Gen 2010;259-272

Types of gene-environment interaction (I)

Model	Interpretation
No interaction	The **same effect** of the exposure on the outcome in individuals with different genotypes
Statistical interaction	A departure from a pure main effects model observed in one or a few studies
Positive interaction or synergism	**Greater effect** of the exposure on the outcome in individuals with a genotype of interest than in individuals with other genotypes
Negative interaction or antagonism	**Smaller effect** of the exposure on the outcome in individuals with a genotype of interest than in individuals with other genotypes
Multiplicative interaction	Interaction observed in multiplicative/relative measures of effect (e.g., OR, RR, HR, etc.)
Additive interaction	Interaction observed in additive/absolute measures of effect (e.g., RD, etc.)

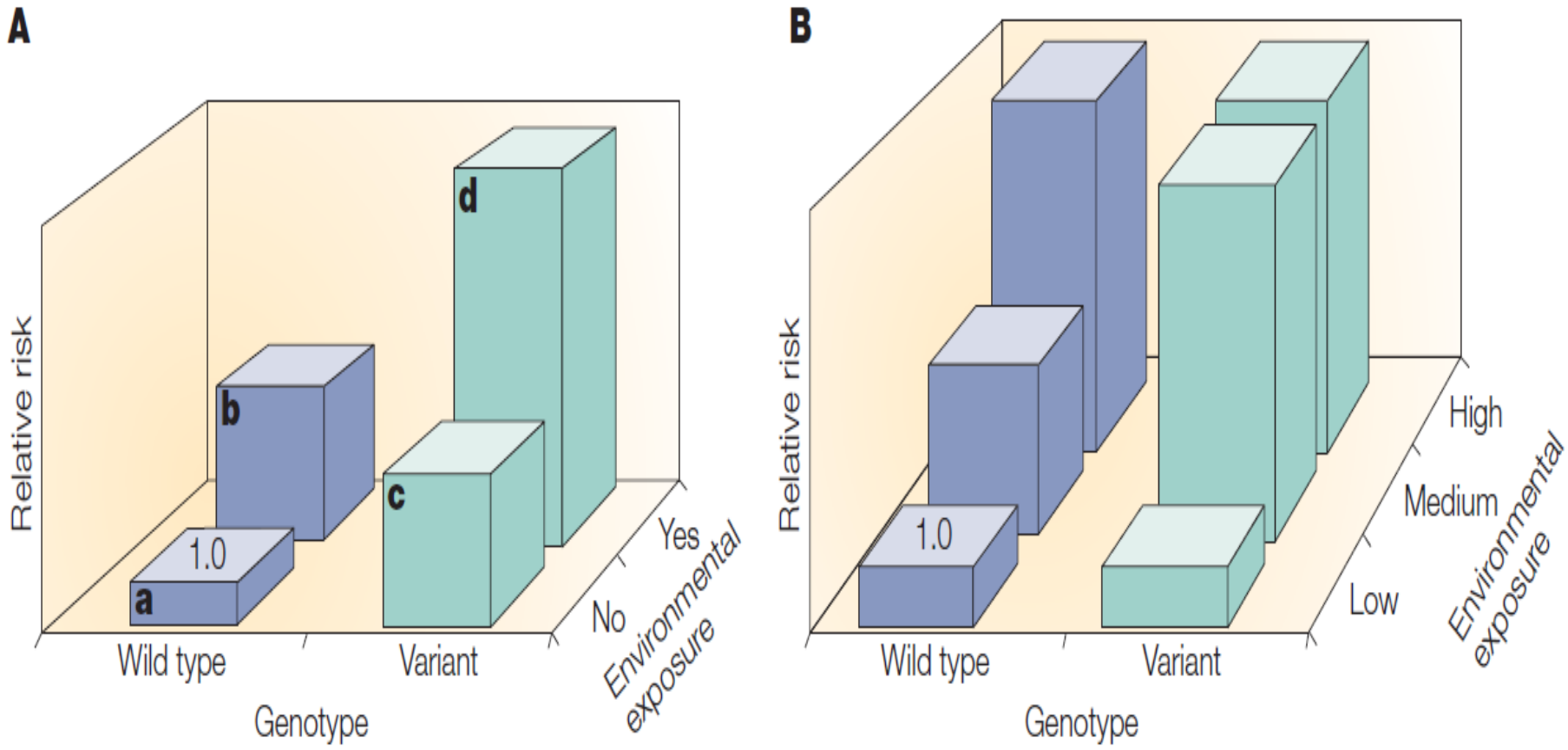
Types of gene-environment interaction (II)

Model	Interpretation
Quantitative interaction	Interaction in which the effects of the exposure on the outcome go in the same direction for different genotypes, but differ in magnitude
Qualitative interaction	Interaction in which the effects of the exposure on the outcome go in opposite directions (e.g., deleterious in carriers and protective in non-carriers) for different genotypes
Biological or causal interaction	An interaction that is present in nature (and is supported by the totality of the evidence)

Uses of gene-environment interaction

- Understanding biological mechanisms and pathways
  - Tobacco smoking – *NAT2* – bladder cancer
- Understanding heterogeneity in results across studies
- Identifying novel genes acting only through interactions
  - Could explain missing heritability (e.g., genetic susceptibility to air pollution in childhood asthma)
- Predicting individual risk of disease or prognosis
  - Optimal mammographic screening interval for *BRCA1* or *BRCA2* mutation carriers
  - Folate supplementation for colorectal cancer risk could depend on *MTHFR*
- Choosing the best treatment for an individual based on genetic predisposition
  - Statins – *SLCO1B1* - cardiomyopathy

Models of gene-environment interaction



Hunter DJ. Nat Rev Gen 2005;287-298

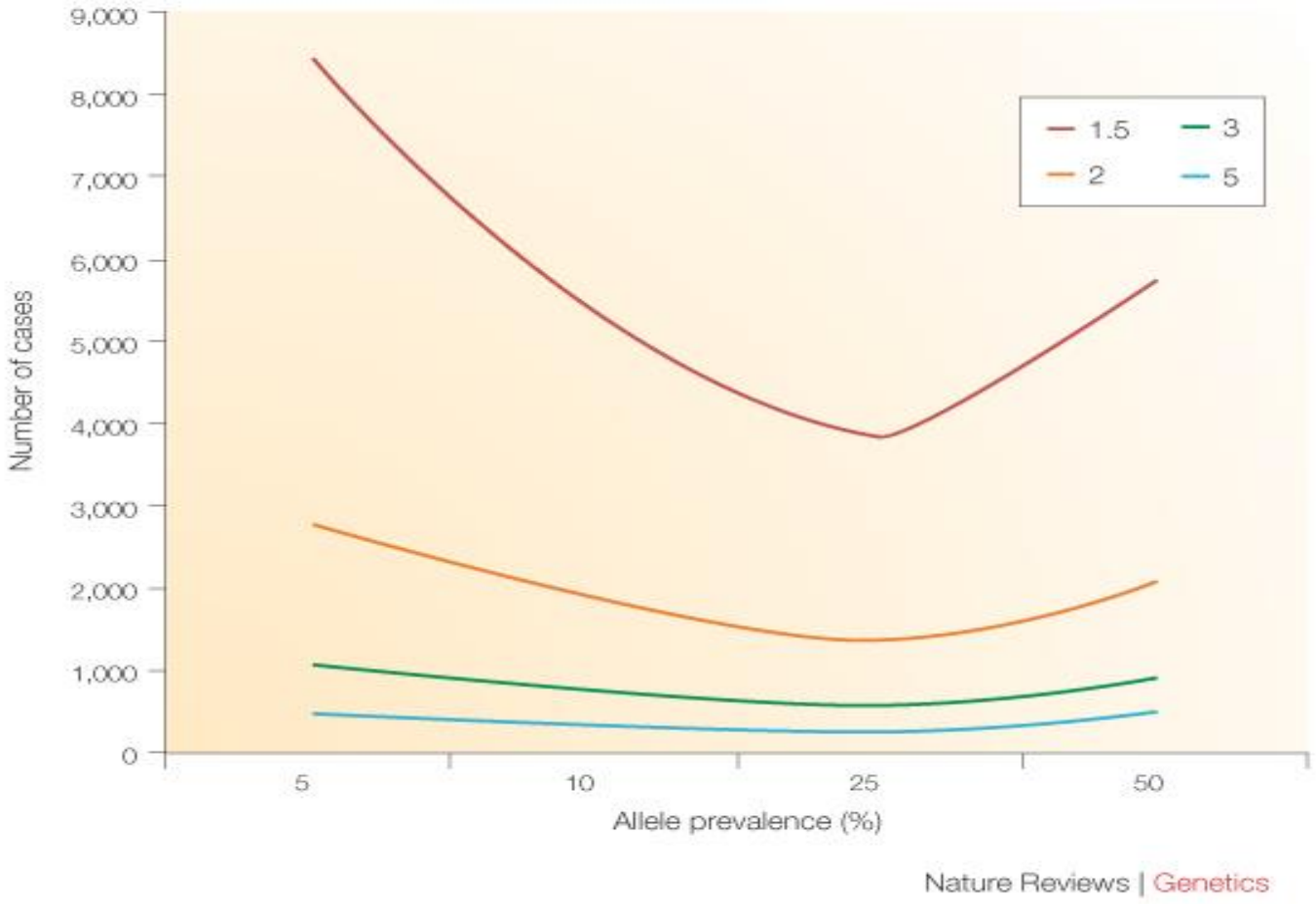
Challenges of gene-environment interaction

- Exposure assessment
  - Multidimensional, time-varying exposures
  - Interactions will be biased only if measurement errors are differentially related to both exposure and genotype

- Sample size and power
  - Sample size requirements can be enormous
  - Some of the poor replication ability of GxE interactions are due to underpowered studies

- Heterogeneity and replication

Example of Sample Size Issue for detecting ONE interaction for a dichotomous trait and a 10% exposure prevalence



Hunter DJ. Nat Rev Gen 2005;287-298

Study designs for gene-environment interaction (I)

Design	Approach	Advantages	Disadvantages	Settings	Examples
Basic epidemiologic designs					
Cohort	Comparison of incidence of new cases across groups defined by E and G	Freedom from most biases; clear temporal sequence of cause and effect	Large cohorts and/or long follow-up needed to obtain sufficient numbers of cases; possible biased losses to follow-up; changes in exposure may require recurring observation	Common Ds or multiple end points; commonly used in biobanks	*ITGB3* × fibrinogen in platelet aggregation in Framingham cohort[154]
Case–control	Comparison of prevalence of E and G between cases and controls	Modest sample sizes needed for rare Ds; can individually match on confounders	Recall bias for E; selection bias, particularly for control group	Rare Ds with common E and G risk factors	*CYP1A2*, *NAT2*, smoking and red meat in colorectal cancer[57]
Case-only	Test of G–E association among cases, assuming G–E independence in the source population	Greater power than case–control or cohort	Bias if G–E assumption is incorrect	G×E studies in which G–E independence can be assumed	Radiotherapy × DNA repair genes in second breast cancers[32]
Randomized trial	Cohort study with random assignment of E across individuals	Experimental control of confounders	Prevention trials for D incidence can require very large sample sizes	Experimental confirmation for chronic effects	Albuteral and *B2AR* in asthmatics[126]

Thomas D. Nat Rev Gen 2010;259-272

Study designs for gene-environment interaction (I)

- **GxE: Standard interaction test 1DF:**

$$\text{Logit}(\Pr(D=1|G))=\beta_0+\beta_G G+\beta_E E+\beta_{GxE} GxE+\beta_C C$$

- **E|G (case only): E|G association in cases e.g. case-only analysis**

$$\text{Logit}(\Pr(G=g|E,D=1))=\beta_0+\beta_E E+\beta_C C$$

Design	Approach	Advantages	Disadvantages	Settings	Examples
Basic epidemiologic designs					
Cohort	Comparison of incidence of new cases across groups defined by E and G	Freedom from most biases; clear temporal sequence of cause and effect	Large cohorts and/or long follow-up needed to obtain sufficient numbers of cases; possible biased losses to follow-up; changes in exposure may require recurring observation	Common Ds or multiple end points; commonly used in biobanks	*ITGB3* × fibrinogen in platelet aggregation in Framingham cohort[154]
Case–control	Comparison of prevalence of E and G between cases and controls	Modest sample sizes needed for rare Ds; can individually match on confounders	Recall bias for E; selection bias, particularly for control group	Rare Ds with common E and G risk factors	*CYP1A2*, *NAT2*, smoking and red meat in colorectal cancer[57]
Case-only	Test of G–E association among cases, assuming G–E independence in the source population	Greater power than case–control or cohort	Bias if G–E assumption is incorrect	G×E studies in which G–E independence can be assumed	Radiotherapy × DNA repair genes in second breast cancers[32]
Randomized trial	Cohort study with random assignment of E across individuals	Experimental control of confounders	Prevention trials for D incidence can require very large sample sizes	Experimental confirmation for chronic effects	Albuterol and *B2AR* in asthmatics[126]

Thomas D. Nat Rev Gen 2010;259-272

Study designs for gene-environment interaction (II)

Design	Approach	Advantages	Disadvantages	Settings	Examples
Hybrid designs					
Nested case–control	Selection of matched controls for each case from cohort members who are still D-free	The freedom from bias of a cohort design combined with the efficiency of a case–control design; simple analysis	Each case group requires a separate control series	Studies within cohorts requiring additional data collection	Antioxidants × MPO in breast cancer[155]
Case–cohort	Unmatched comparison of cases from a cohort with a random sample of the cohort	Same advantages as nested case–control; the same control group can be used for multiple case series	Complex analysis	Studies within cohorts with stored baseline biospecimens	*APOE* and smoking for CHD in Framingham offspring cohort[156]
Two-phase case–control	Stratified sampling on D, E and G for additional measurements (for example, biomarkers)	High statistical efficiency for subsample measurements	Complex analysis	Substudies for which outcome and predictor data are already available	*GST* genes and tobacco smoking in CHD[47]
Counter-matching	Matched selection of controls who are discordant for a surrogate for E	Permits individual matching; highly efficient for E main effect and G×E interactions	Complex control selection	Substudies in which a matched design is needed	Radiotherapy × DNA repair genes in second breast cancers[49]
Joint case-only and case–control	Bayesian compromise between case-only and case–control comparisons	Power advantage of case-only combined with robustness of case–control	Some bias when G–E association is moderate	G×E studies for which G–E independence is uncertain	*GSM1*, *NAT2*, smoking and diet in colorectal cancer[34]
Family-based designs					
Case–sibling (or –cousin)	Case–control comparison of E and G using unaffected relatives as controls	More powerful than case–control for G×E; immune to population stratification bias	Discordant sibships difficult to enroll; overmatching for G main effects	Populations with potential substructure	*GSTM1* × air pollution in childhood asthma[17]
Case–parent triad	Comparison of Gs for cases with Gs that could have been inherited from parents, stratified by case's E	More powerful than case–control for G×E; immune to population stratification bias for G main effects	Difficult to enroll complete triads; possible bias in G×E if G and E are associated within parental mating types	Substructured populations, particularly for Ds of childhood	*TGFA* × maternal smoking, alcohol and vitamins in cleft palate[157]
Twin studies	Comparison of D concordance between MZ and DZ pairs in different environments	No genetic data required; can be extended to include half-siblings, twins reared together or apart, or to compare discordant pairs on measured G and E	Used mainly to identify interactions with unmeasured genes; assumption of similar E between MZ and DZ pairs	Exploratory studies of potential for G×E before specific genes have been identified	Concordance of insulin levels in relation to non-genetic variation in obesity[158]

Study designs for gene-environment interaction (III)

Design	Approach	Advantages	Disadvantages	Settings	Examples
GWA designs					
Two-stage genotyping	Use of high-density panel on part of a case–control sample to select a subset of SNPs with suggestive Gs or G×E interaction for testing; the SNPs are tested using a custom panel in an independent sample, with joint analysis of both samples	Highly cost efficient	Only part of sample has GWA genotypes	GWA studies for which complete SNP data on all subjects is not needed	None identified
Two-step interaction analysis	Preliminary filtering of a GWA scan for G–E association in combined case–control sample, followed by G×E testing of a selected subset	Much more powerful for G×E or G×G interactions than a single-step analysis	Can miss some interactions	GWA studies with complete SNP data and focus on G×E	G × *in utero* tobacco in childhood asthma

Thomas D. Nat Rev Gen 2010;259-272

GxE Interaction: Testing for Additive/Multiplicative Effects

Stratum	Cases	Controls
Gene (G+), Environment (E+)	a	b
Gene (G+), No Environment (E-)	c	d
No Gene (G-), Environment (E+)	e	f
No Gene (G-), No Environment (E-)	g	h

Identifying GxE Interaction

Strata	Cases	Controls	Odds Ratio (OR)
G+E+	a	b	ah / bg
G+E-	c	d	ch / dg
G-E+	e	f	eh / fg
G-E-	g	h	1 (Ref)

GxE Interaction:
4 groups defined by genotype and exposure

COHORT STUDY:	G+ E+	G+ E-	G- E+	G- E-
Affected	a	b	e	f
Unaffected	c	d	g	h
Risk	a/(a+c)	b/(b+d)	e/(e+g)	f/(f+h)
Relative risk	$RR_{G+} = \frac{a/(a+c)}{b/(b+d)}$		$RR_{G-} = \frac{e/(e+g)}{f/(f+h)}$	
Risk difference	$RD_{G+} = a/(a+c) - b/(b+d)$		$RD_{G-} = e/(e+g) - f/(f+h)$	

Test for interaction: Is the effect of the exposure on the outcome the same in people with and without the high-risk genotype?

Multiplicative scale: No interaction implies $RR_{G+} = RR_{G-}$
Additive scale: No interaction implies $RD_{G+} = RD_{G-}$

Example: Factor V Leiden Mutations, Oral Contraceptive Use, and Venous Thrombosis

Strata	Cases	Controls	OR
G+E+	25	2	34.7
G+E-	10	4	6.9
G-E+	84	63	3.7
G-E-	36	100	Reference
Total	155	169	

Vanderbroucke et al., The Lancet 1994

Factor V Leiden Mutations, Oral Contraceptive Use, and Venous Thrombosis

Evidence for Interaction?

Strata OR

G+E+ 34.7

G+E- 6.9

G-E+ 3.7

G-E- Ref

$OR_{Interaction}$ =

34.7 / 6.9 x 3.7 = 1.4

Risk of thrombosis in women using OCs is much greater among those with Factor V Leiden Mutations than those without

Examples of gene-environment interactions

Table 2 | **Selected examples of gene–environment interactions observed in at least two studies**

Gene symbol	Variant(s)	Environmental exposure	Outcome and nature of interaction
Genes for skin pigmentation (for example, *MC1R*)	Variants for fair skin colour	Sunlight or ultraviolet light B	Risk of skin cancer is higher in people with fair skin colour that are exposed to higher amounts of sunlight
CCR5	Δ-32 deletion	HIV	Carriers of the receptor deletion have lower rates of HIV infection and disease progression
MTHFR	Ala222Val polymorphism	Folic acid intake	Homozygotes for the low activity Ala222Val variant are at different risk of colorectal cancer and adenomas if nutritional folate status is low
NAT2	Rapid versus slow acetylator SNPs	Heterocyclic amines in cooked meat	Red meat intake is more strongly associated with colorectal cancer among rapid acetylators
F5	Leiden prothrombotic variant	Hormone replacement	Venous thromboembolism risk is increased in factor V Leiden carriers who take exogenous steroid hormones
UGT1A6	Slow-metabolism SNPs	Aspirin	Increased benefit of prophylactic aspirin use in carriers of the slow metabolism variants
APOE	*E4* allele	Cholesterol intake	Exaggerated changes in serum cholesterol in response to dietary cholesterol changes in *APOE4* carriers
ADH1C	γ-2 alleles	Alcohol intake	Inverse association between ethanol intake and myocardial infarction; risk is stronger in carriers of slow-oxidizing γ-2 alleles
PPARG2	Pro12Ala	Dietary fat intake	Stronger relation between dietary fat intake and obesity in carriers of the Pro12Ala allele
HLA-DPB1	Glu69	Occupational beryllium	Exposed workers who are carriers of the Glu69 allele are more likely to develop chronic beryllium lung disease
TPMT	Ala154Thr and Tyr240Cys	Thiopurine drugs	Homozygotes for the low-activity alleles of *TPMT* are likely to experience severe toxicity when exposed to thiopurine drugs
ADRB2	Arg16Gly	Asthma drugs	Arg16Gly homozygotes have a greater response in the airway to albuterol

Hunter DJ. Nat Rev Gen 2005;287-298

Examples of 'established' GxE Interactions

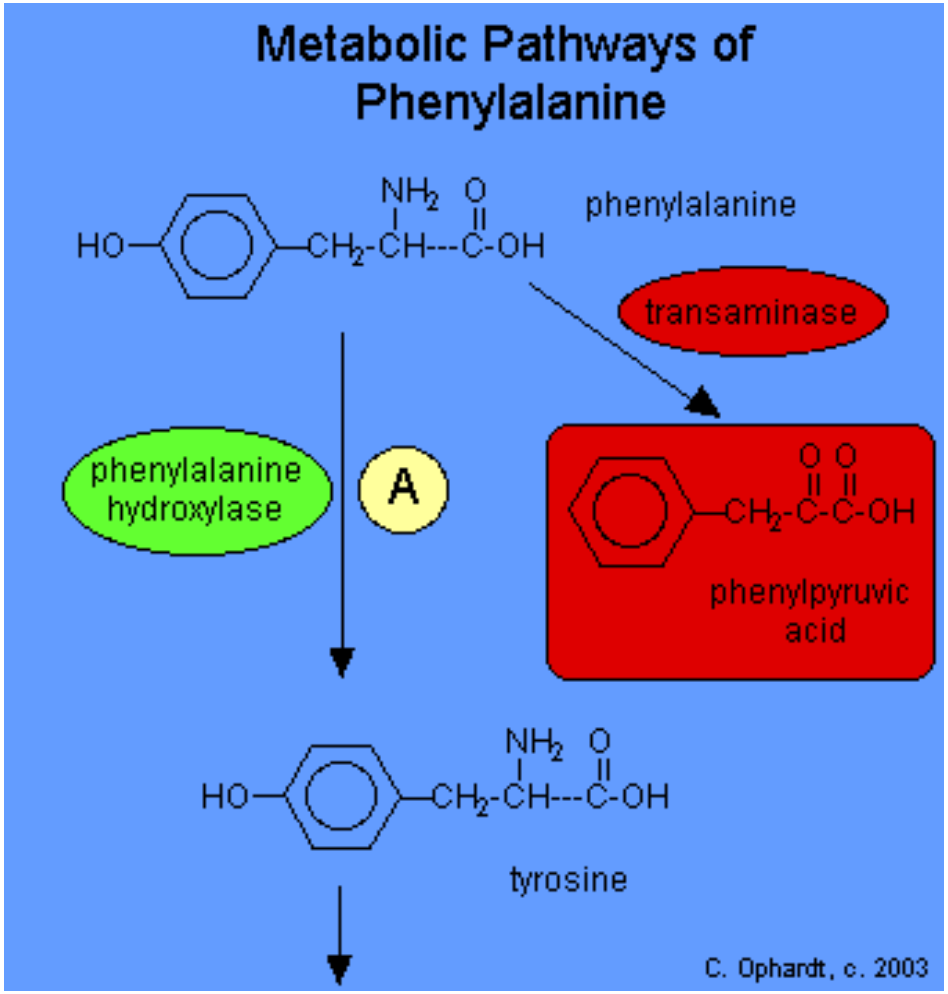
- Have any GXE Interactions been identified with certainty?
  - Few Established Examples to Date
    - Phenylketonuria
    - Lactose Intolerance
    - Smoking, NAT2 and Bladder Cancer
    - Coffee, GRIN2A, and Parkinson's Disease?

Phenylketonuria

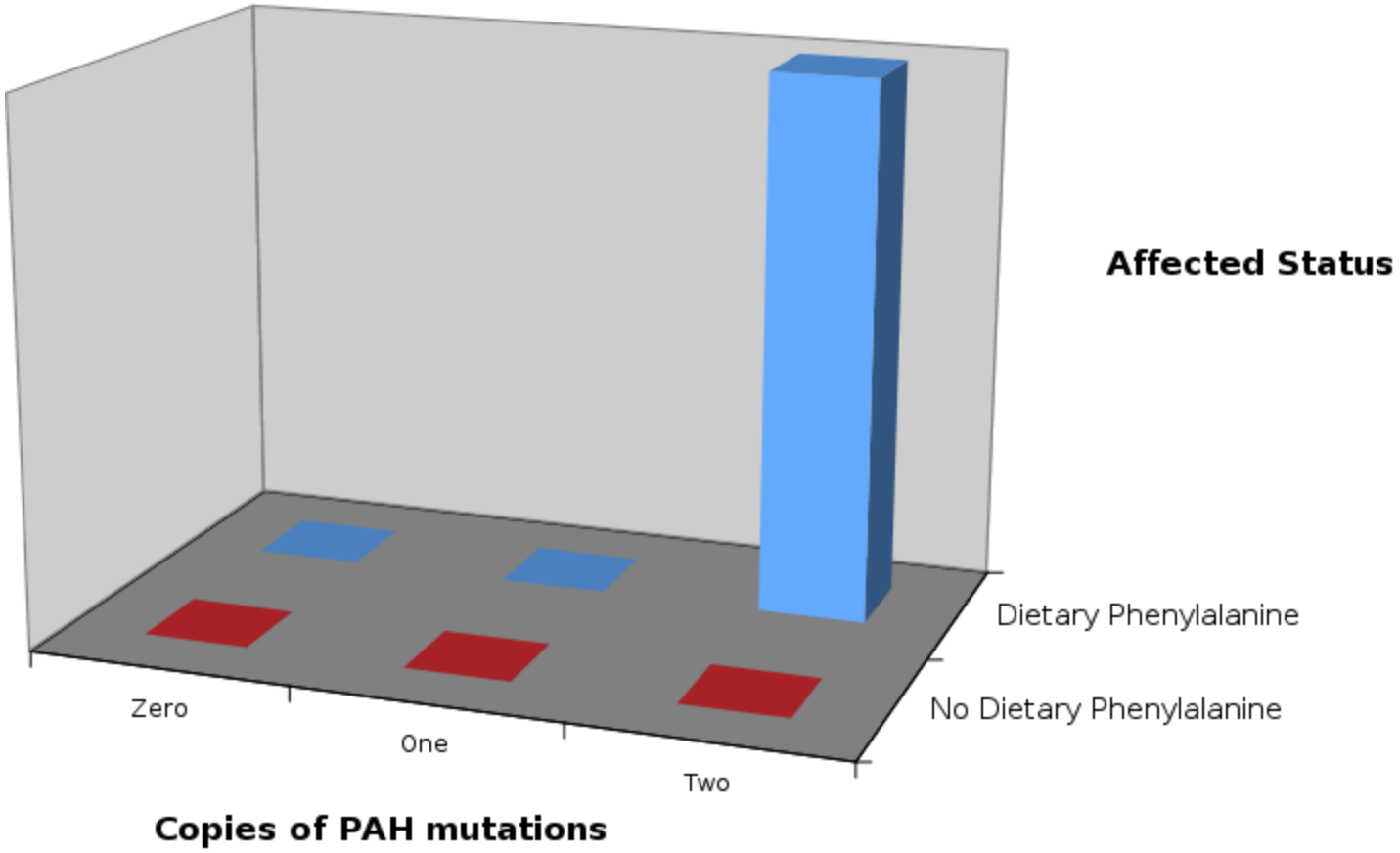
- Mental retardation and seizures
- 1/15,000 live births
  - 1/100,000 in Finland
  - 1/2,600 in Turkey

- Mutations in Phenylalanine Hydroxylase (PAH) (G)
- Dietary Phenylalanine (E)

- Both are necessary
- Neither is sufficient for disease

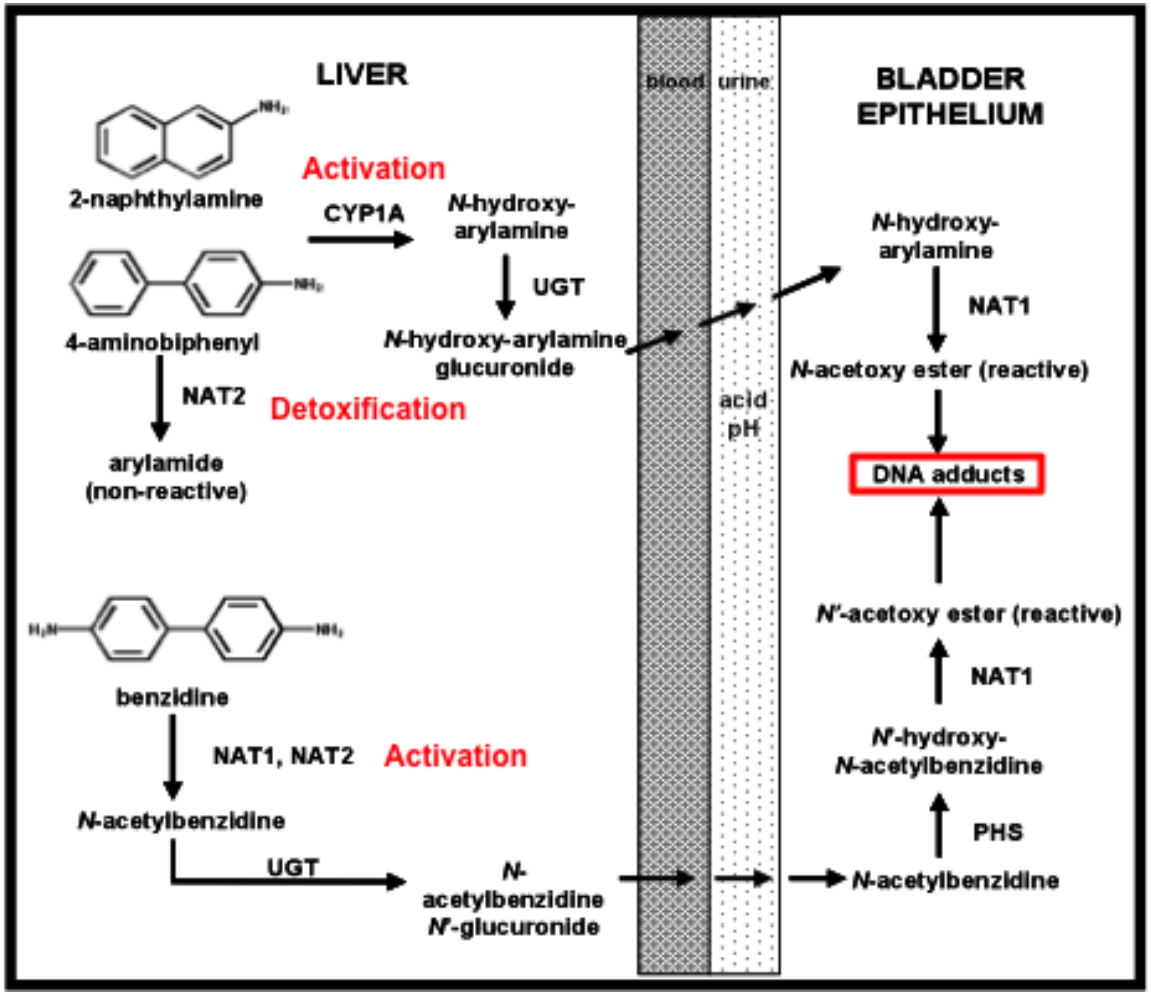


Phenylketonuria: Example of Gene-Nutrition Interaction



Common variation in metabolizing genes could modify the effects of arylamine exposure

Metabolism of aromatic amines and bladder carcinogenesis



Strong in vitro and in vivo evidence that interaction exists

Lancet 2005; 366: 649–59

NAT2 slow acetylation, *GSTM1* null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses

Montserrat García-Closas, Núria Malats, Debra Silverman, Mustafa Dosemeci, Manolis Kogevinas, David W Hein, Adonina Tardón, Consol Serra, Alfredo Carrato, Reina García-Closas, Josep Lloreta, Gemma Castaño-Vinyals, Meredith Yeager, Robert Welch, Stephen Chanock, Nilanjan Chatterjee, Sholom Wacholder, Claudine Samanic, Montserrat Torà, Francisco Fernández, Francisco X Real, Nathaniel Rothman

NAT2 slow acetylation increases bladder cancer risk by 40%

OR=1.4 95%CI (1.2-1.6)

Study	Year	Country	Cases
Taylor (blacks)	1998	USA	15
Karakaya	1986	Turkey	23
Miller	1983	USA	26
Su	1998	Taiwan	27
Woodhouse	1982	UK	30
Lower	1979	USA	34
Cartwright	1984	Portugal	47
Horai	1989	Japan	51
Jaskula-Sztul	2001	Poland	56
Lower	1979	Denmark	71
Ishizu	1995	Japan	71
Hsieh	1999	Taiwan	73
Dewan	1995	India	77
Giannakopoulos	2002	Greece	89
Kaisary	1987	UK	98
Evans	1983	UK	100
Mittal	2004	India	101
Roots	1989	Germany	102
Hanssen	1985	Germany	105
Kim	2000	Korea	112
Peluso	1998	Italy	114
Lower	1979	Sweden	115
Ladero	1985	Spain	130
Risch	1995	UK	189
Hung	2004	Italy	201
Taylor(whites)	1998	USA	215
Mommsen	1985	UK	228
Okkels	1997	Denmark	254
Tsukino	2004	Japan	325
Brockmöller	1996	Germany	374
Gu	2005	USA	504
García-Closas	Current	Spain	1134

	Cases	Odds ratio (95% CI)	p
All studies (n=31)	5091	1·4 (1·2–1·6)	<0·0001
Studies of white populations (n=22)	4216	1·4 (1·3–1·5)	<0·0001
Europe (n=18)	3437	1·4 (1·3–1·6)	<0·0001
USA (n=4)	779	1·1 (0·8–1·6)	0·58
Studies of Asian populations (n=6)	659	1·5 (0·8–2·6)	0·20

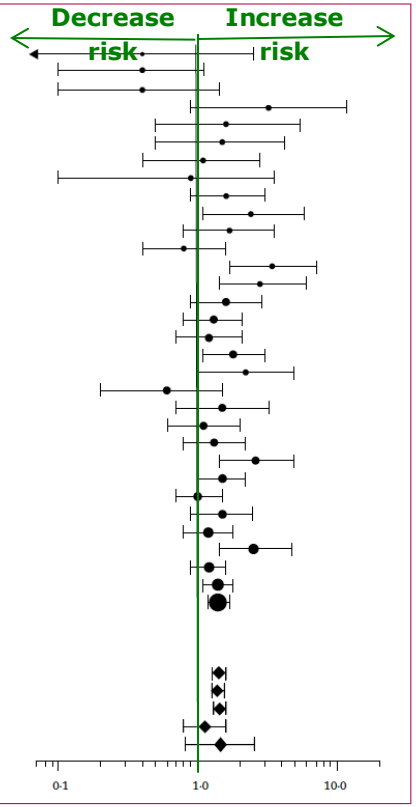


Figure 2: **Meta-analysis of studies of *NAT2* slow-acetylation genotype and bladder-cancer risk**
Numbers of cases are individuals with *NAT2* information.

GSTM1 deletion increases bladder cancer risk by 50%

OR=1.5 95%CI (1.3-1.6)

Study	Year	Country	Cases
Romkes	Unpublished	USA	13
Bell (blacks)	1993	USA	16
Heckbert	1992	USA	29
Daly	1993	UK	53
Mungan	2000	Netherlands	61
Šalagovic	1999	Slovakia	76
Georgiou	2000	Greece	89
Lin (whites)	1994	USA	89
Chern	1994	UK	95
Zhong	1993	UK	97
Aktas	2001	Turkey	103
Srivastava	2004	India	106
Moore	2004	Argentina	114
Törüner	2001	Turkey	121
Jeong	2003	Korea	126
Peluso	2000	Italy	130
Steinhoff	2000	Germany	135
Schnakenberge	2000	Germany	157
Kang	1999	Korea	174
Hung	2004	Italy	201
Bell (whites)	1993	USA	213
Kim	2002	Korea	216
Lee	2002	Korea	232
Okkels	1996	Denmark	234
Tsukino	2004	Japan	325
Karagas	2005	USA	354
Brockmöller	1996	Germany	374
García-Closas	Current	Spain	1139

	Cases	Odds ratio (95% CI)	p
All studies (n=28)	5072	1·5 (1·3–1·6)	<0·0001
Studies of predominantly white populations (n=18)	3539	1·4 (1·2–1·6)	<0·0001
Europe (n=13)	2841	1·5 (1·2–1·7)	<0·0001
USA (n=5)	698	1·3 (1·1–1·5)	0·02
Studies of Asian populations (n=6)	1073	1·4 (1·2–1·7)	<0·0001

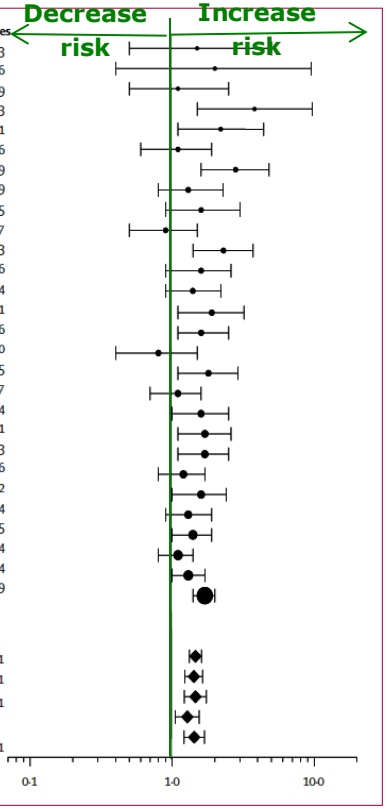
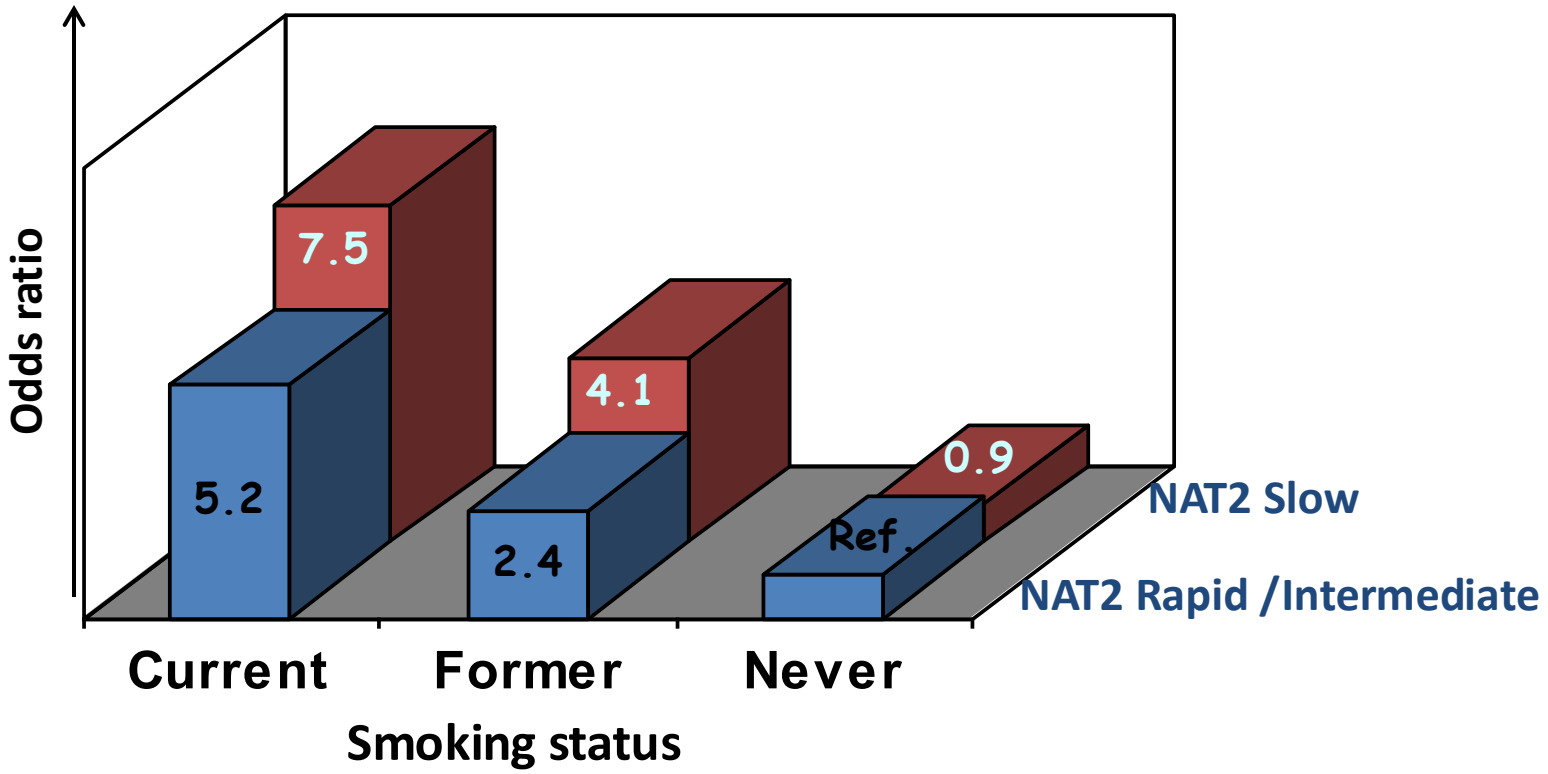


Figure 4: **Meta-analysis of studies of *GSTM1* null genotype and bladder-cancer risk**
Number of cases for studies in Engel et al[8] are based on table 1 of that paper.

NAT2 slow acetylators are at higher risk of developing bladder cancer from smoking

Joint effect of smoking and NAT2 acetylation on bladder cancer risk: Spanish Bladder Cancer Study



Garcia-Closas M et al. Lancet, 2005; 366: 649-658.

Coffee, *GRIN2A* and Parkinson's Disease?

- Coffee shown to be inversely associated with PD in observational studies (though not all benefit equally)

- Conducted GWAS (>800,000 SNPs; agnostic)

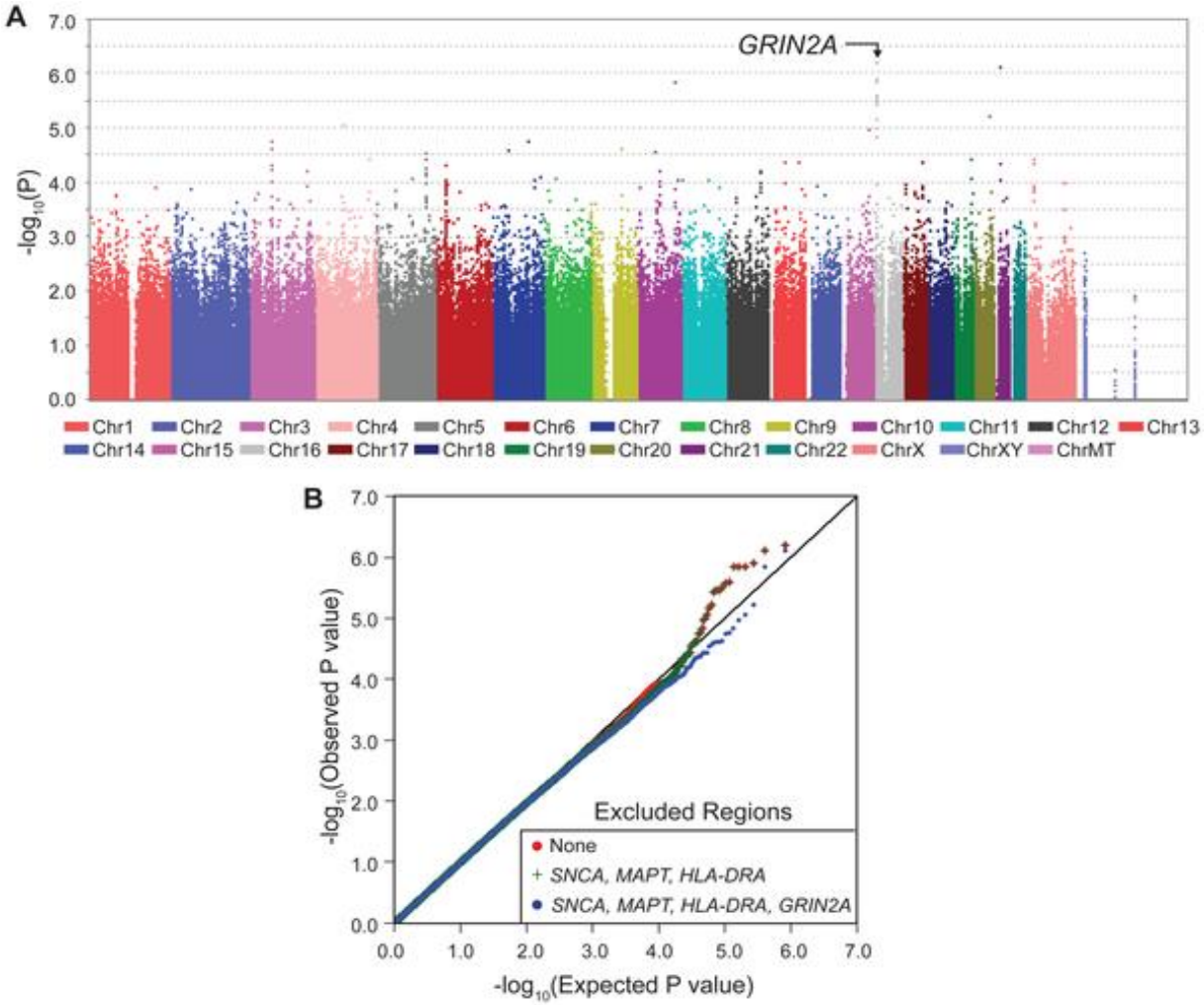
- 1,458 persons with PD and 931 without PD from the NeuroGenetics Research Consortium (NGRC),

- *GRIN2A* as a novel PD modifier gene. *GRIN2A* encodes a subunit of the NMDA-glutamate-receptor which is well known for regulating excitatory neurotransmission in the brain and for controlling movement and behaviour.

- Proof of concept that inclusion of environmental factors can help identify genes that are missed in GWAS.

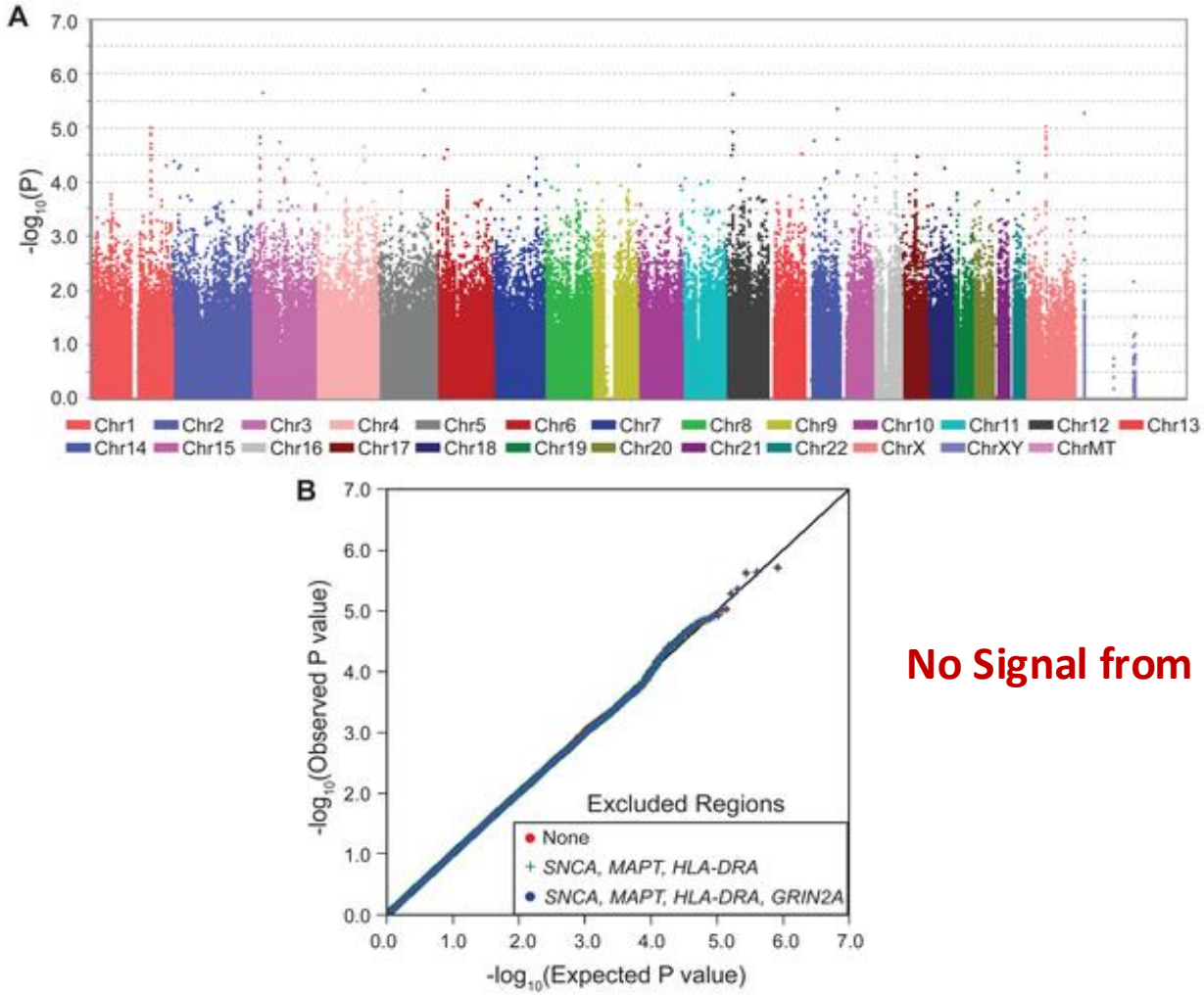
Hamza TH, et al. (2011) Genome-Wide Gene-Environment Study Identifies Glutamate Receptor Gene GRIN2A as a Parkinson's Disease Modifier Gene via Interaction with Coffee. PLoS Genet 7(8): e1002237.

GWAS in heavy coffee-drinkers.



Hamza TH, Chen H, Hill-Burns EM, Rhodes SL, et al. (2011) Genome-Wide Gene-Environment Study Identifies Glutamate Receptor Gene GRIN2A as a Parkinson's Disease Modifier Gene via Interaction with Coffee. PLoS Genet 7(8): e1002237. doi:10.1371/journal.pgen.1002237
http://www.plosgenetics.org/article/info:doi/10.1371/journal.pgen.1002237

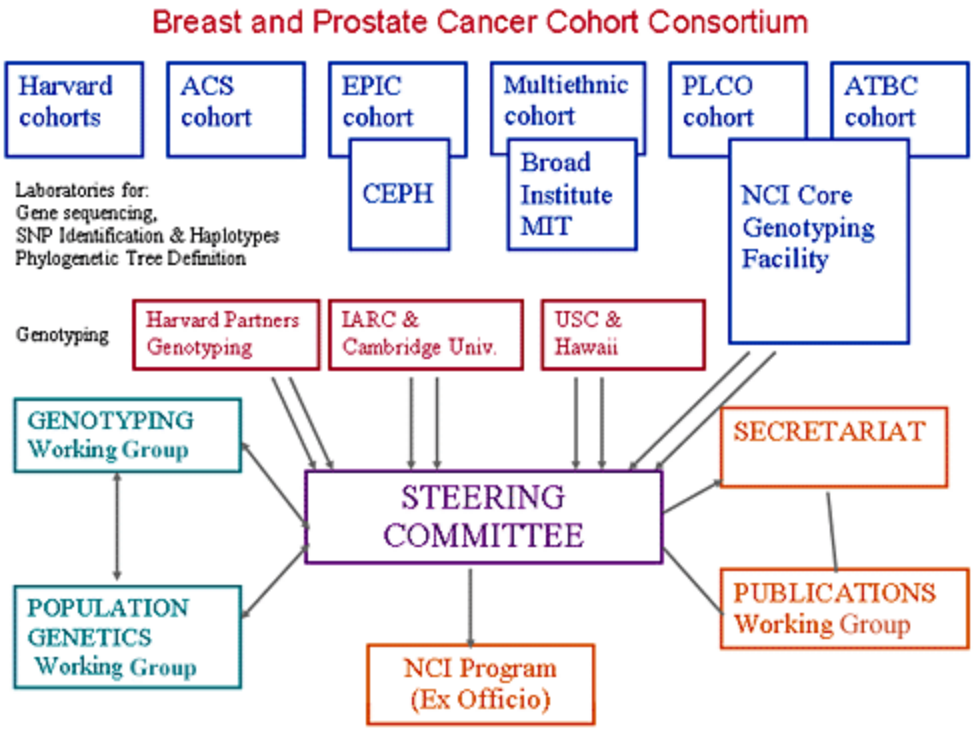
GWAS in light coffee-drinkers



No Signal from GRIN2A

Hamza TH, Chen H, Hill-Burns EM, Rhodes SL, et al. (2011) Genome-Wide Gene-Environment Study Identifies Glutamate Receptor Gene GRIN2A as a Parkinson's Disease Modifier Gene via Interaction with Coffee. PLoS Genet 7(8): e1002237. doi:10.1371/journal.pgen.1002237
http://www.plosgenetics.org/article/info:doi/10.1371/journal.pgen.1002237

Breast and Prostate Cancer Cohort Consortium (BPC3)



http://epi.grants.cancer.gov/BPC3/structure.html

BPC3 cohorts

Table 1 | **Cohort studies in the Breast and Prostate Cancer Cohort Consortium**

Cohort study	Year of blood collection	Initial study design	Number of men with blood sample	Number of women with blood sample	Breast cancer*	Prostate cancer*
American Cancer Society						
Cancer Prevention Study-II	1998	Prospective follow-up study	17,411	21,965	503; 503	1,207; 1,209
Harvard University						
Physicians' Health Study	1982	Randomized trial of aspirin and β-carotene	14,916‡	NA	NA	1,118; 1,447
Nurses' Health Study	1989	Prospective follow-up study	NA	32,826‡	1,100; 1,953	NA
Health Professionals Follow-up Study	1993	Prospective follow-up study	18,410‡	NA	NA	707; 701
Womens' Health Study	1993	Randomized trial of aspirin and vitamin E	NA	28,263‡	705; 705	NA
International Agency for Research in Cancer						
European Prospective Investigation into Cancer and Nutrition	1992	Prospective follow-up study	139,207‡	249,327‡	1,719; 2,844	953; 1,320
Universities of Southern California and Hawaii						
Multiethnic Cohort	1996	Prospective follow-up study	Blood collection ongoing	Blood collection ongoing	1,617; 1,962	2,320; 2,399
National Cancer Institute						
Prostate, Lung, Colon, Ovary	1993	Randomized trial of screening	32,338‡	32,339‡	NA	1,306; 1,668
α-Tocopherol, β-carotene	1991	Randomized trial of β-carotene and vitamin E	NA	26,593‡	NA	1,058; 1,058
Total for all studies	NA	NA	222,282	391,313	5,644; 7,967	8,669; 9,802

*Indicates that data are shown as cases; controls. ‡Indicates that blood specimens were collected before diagnosis. NA, not applicable.

Hunter DJ, et al. Nat Rev Cancer 2005;5:977-85

BPC3: GxE interaction studies for prostate cancer

American Journal of Epidemiology

DOI: 10.1093/aje/kwr423

Original Contribution

Interactions Between Genome-wide Significant Genetic Variants and Circulating Concentrations of Insulin-like Growth Factor 1, Sex Hormones, and Binding Proteins in Relation to Prostate Cancer Risk in the National Cancer Institute Breast and Prostate Cancer Cohort Consortium

Konstantinos K. Tsilidis*, Ruth C. Travis, Paul N. Appleby, Naomi E. Allen, Sara Lindstrom, Fredrick R. Schumacher, David Cox, Ann W. Hsing, Jing Ma, Gianluca Severi, Demetrius Albanes, Jarmo Virtamo, Heiner Boeing, H. Bas Bueno-de-Mesquita, Mattias Johansson, J. Ramón Quirós, Elio Riboli, Afshan Siddiq, Anne Tjønneland, Dimitrios Trichopoulos, Rosario Tumino, J. Michael Gaziano, Edward Giovannucci, David J. Hunter, Peter Kraft, Meir J. Stampfer, Graham G. Giles, Gerald L. Andriole, Sonja I. Berndt, Stephen J. Chanock, Richard B. Hayes, and Timothy J. Key

* Correspondence to Dr. Konstantinos K. Tsilidis, Department of Hygiene and Epidemiology, School of Medicine, University of Ioannina, Ioannina 45110, Greece (e-mail: kostas.tsilidis@ceu.ox.ac.uk).

Initially submitted July 13, 2011; accepted for publication October 27, 2011.

Genome-wide association studies (GWAS) have identified many single nucleotide polymorphisms (SNPs) associated with prostate cancer risk. There is limited information on the mechanistic basis of these associations, particularly about whether they interact with circulating concentrations of growth factors and sex hormones, which may be important in prostate cancer etiology. Using conditional logistic regression, the authors compared per-allele odds ratios for prostate cancer for 39 GWAS-identified SNPs across thirds (tertile groups) of circulating concentrations of insulin-like growth factor 1 (IGF-1), insulin-like growth factor binding protein 3 (IGFBP-3), testosterone, androstenedione, androstanediol glucuronide, estradiol, and sex hormone-binding globulin (SHBG) for 3,043 cases and 3,478 controls in the Breast and Prostate Cancer Cohort Consortium. After allowing for multiple testing, none of the SNPs examined were significantly associated with growth factor or hormone concentrations, and the SNP-prostate cancer associations did not differ by these concentrations, although 4 interactions were marginally significant (*MSMB*-rs10993994 with androstenedione (uncorrected $P = 0.008$); *CTBP2*-rs4962416 with IGFBP-3 (uncorrected $P = 0.003$); 11q13.2-rs12418451 with IGF-1 (uncorrected $P = 0.006$); and 11q13.2-rs10896449 with SHBG (uncorrected $P = 0.005$)). The authors found no strong evidence that associations between GWAS-identified SNPs and prostate cancer are modified by circulating concentrations of IGF-1, sex hormones, or their major binding proteins.

BPC3: GxE interaction studies for prostate cancer

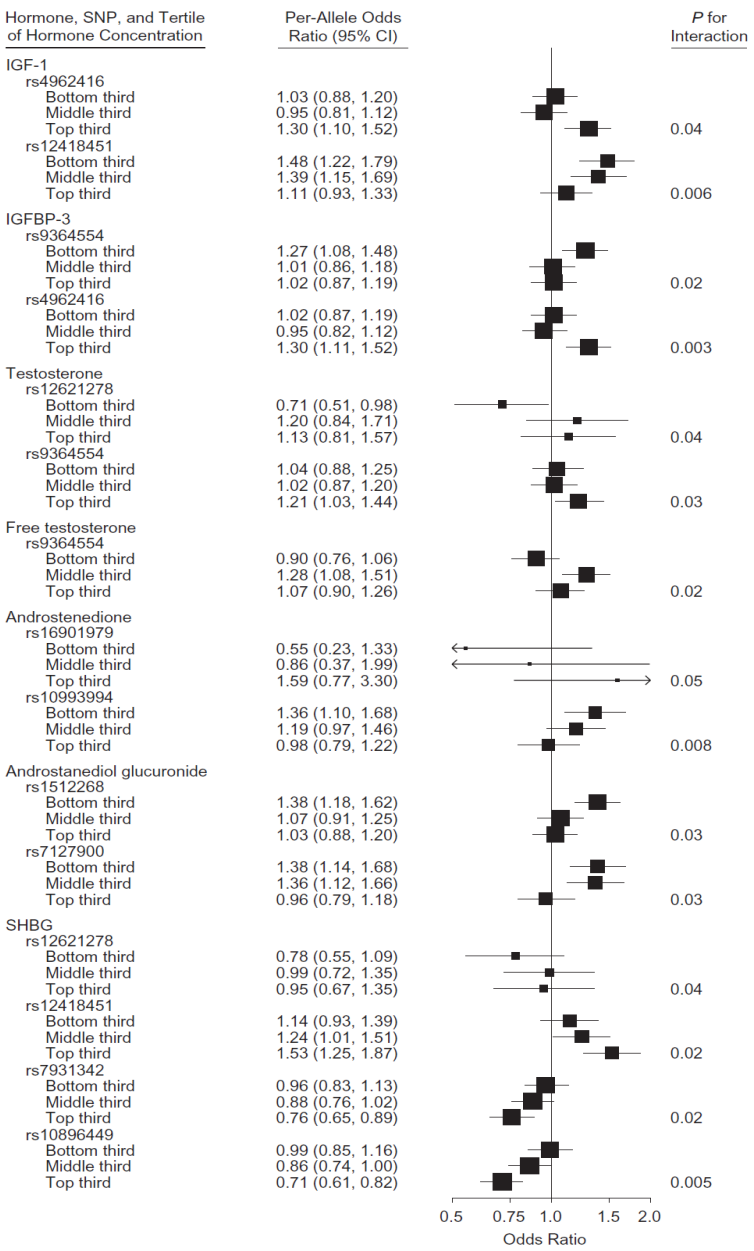


Figure 1. Per-allele associations between single nucleotide polymorphisms (SNPs) identified in genome-wide association studies and risk of prostate cancer, according to circulating concentrations of insulin-like growth factor and steroid sex hormones, for the 15 nominally significant interactions in the Breast and Prostate Cancer Cohort Consortium. Results were obtained from a conditional logistic regression model using cohort-specific thirds of the hormone concentrations (see Web Table 1), matched for age at blood draw, cohort, and country (within the European Prospective Investigation into Cancer and Nutrition), and adjusted for age at blood draw (years; continuous) and body mass index. The *P* values for interaction were calculated using 1-df likelihood ratio tests based on per-allele odds ratios and a continuous hormone variable. Conventional *P* values are shown; all *P* values were nonsignificant after allowance for multiple testing. Bars, 95% confidence interval (CI). (IGF-1, insulin-like growth factor 1; IGFBP-3, insulin-like growth factor binding protein 3; SHBG, sex hormone-binding globulin).

GxE Interaction: Epigenetics

- The study of reversible heritable changes in gene function that occur without a change in the sequence of nuclear DNA (focus of today's talk has been non-reversible heritable changes….)

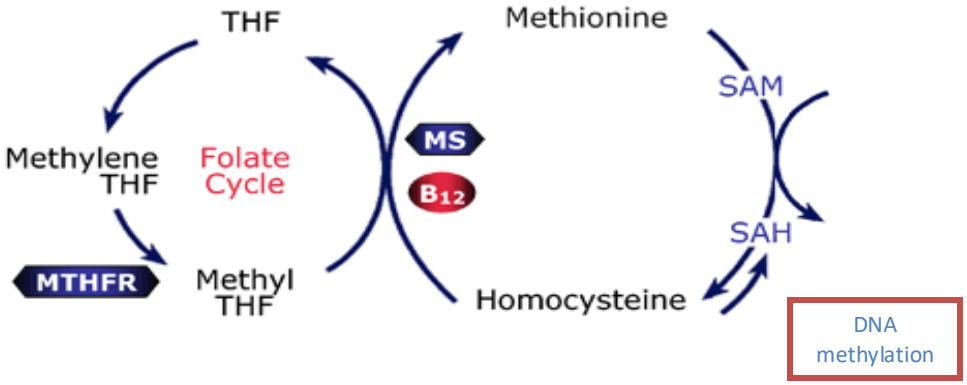
- Gene-regulatory information that is not expressed in DNA sequences but that is transmitted from one generation (of cells or organisms) to the next (e.g. as methylation changes to DNA structure)

- Strongly influenced by environmental exposures such as diet (in utero nutrition etc)

- Likely to influence GxE interactions….future studies may incorporate epigenetic data into GxE estimations….and beyond.

One-carbon metabolism and dietary methyl donors

Various dietary micronutrients can impact DNA methylation (e.g. Folate, choline, Vitamin B)



SAM ***S-adenosylmethionine***

OPEN ACCESS Freely available online PLoS one

Periconceptional Maternal Folic Acid Use of 400 μg per Day Is Related to Increased Methylation of the *IGF2* Gene in the Very Young Child

Régine P. Steegers-Theunissen[1,2,3,4*], Sylvia A. Obermann-Borst[1], Dennis Kremer[6], Jan Lindemans[5], Cissy Siebel[7], Eric A. Steegers[1], P. Eline Slagboom[6], Bastiaan T. Heijmans[6]

Table 3. *IGF2* DMR methylation in the child and independent factors of the mother and the child.

Factors	Mother	*P*-value	Child	*P*-value
Folic acid use	+4.5% (1.8)	0.014	-	-
Female sex	-	-	+2.0% (1.6)	0.232
Age	−0.4% (0.8)	0.585	−0.7% (1.0)	0.478
Birth weight	-	-	−1.7% (0.8)	0.034
Gestational age	-	-	−0.9% (0.8)	0.276
Biochemistry				
SAM, μmol/L	+1.7% (0.8)	0.037	+1.2% (0.8)	0.129
SAH, μmol/L	+0.8% (0.8)	0.331	+0.1% (0.8)	0.882
SAM/SAH	+0.0% (0.8)	0.985	+0.3% (0.8)	0.717

Linear Mixed Model analysis. Data are presented in percentage (standard error) of mean change in relative methylation. For independent quantitative parameters the change in relative methylation is given per SD-change in that parameter. The p-value of the significant association of periconceptional folic acid use and *IGF2* DMR methylation was additionally adjusted for maternal education. The p-value for the significant association between maternal SAM and IGF2 DMR methylation was also adjusted for maternal education and the SAM concentration of the child. The p-value for the significant association between IGF2 DMR methylation and birth weight was additionally adjusted for periconceptional folic acid use and gestational age at delivery.

doi:10.1371/journal.pone.0007845.t003

The Dutch Famine: 1944-45

•September – November 1944: German administration placed embargo on all food transports to the western Netherlands.

•Unusually harsh and early winter froze canals, Much agricultural land became battlefields and was destroyed

•People of the Western Netherlands subsisted on <1000 calories/day, Typical daily rations consisted of potatoes, bread, and sugar beets.

•More than 4.5 million people affected. Deaths of 18,000 Dutch people were attributed to malnutrition as the primary cause and in many more as a contributing factor.

•May 1945: famine ended with the liberation of the country by the Allies. But the famine's legacy lived on.
•Long-lasting effects on infant mortality, infant birth size, mental health, and the development of chronic diseases such as diabetes, obesity, and coronary heart disease and neurological conditions including schizophrenia and depression

•Associations dependent on sex of the exposed individual and timing of the exposure during gestation

•Grandchildren of women who were pregnant during the famine were also smaller than average