Genetic Epidemiology at the intersection between function and disease

Florian Kronenberg

Institute of Genetic Epidemiology, Medical University of Innsbruck

1

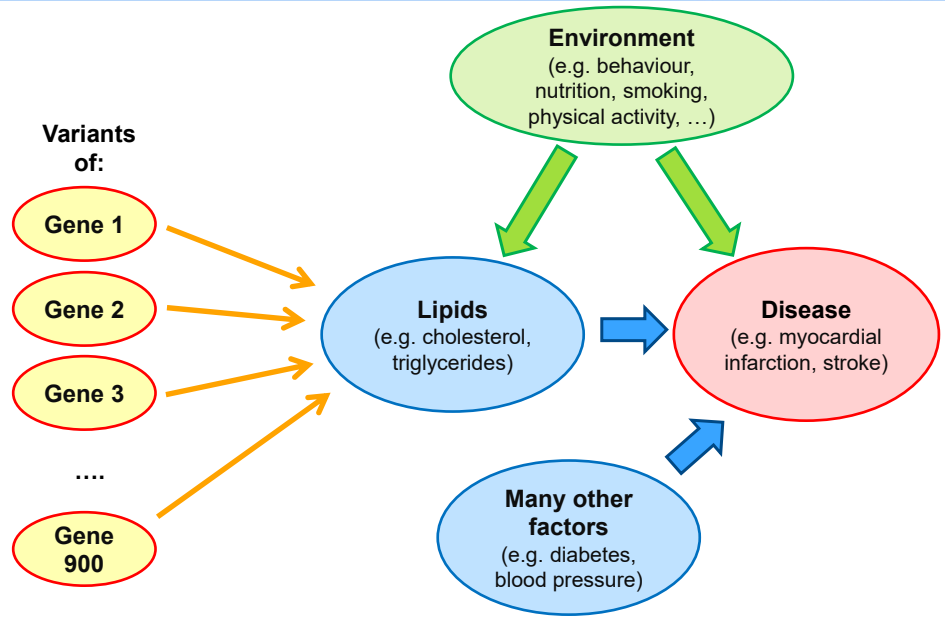
Overview

1. **Background**
2. Association studies
3. Genomewide association studies (GWAS)

GENEPI INNSBRUCK

2

How is health and disease determined?



GENEPI INNSBRUCK

3

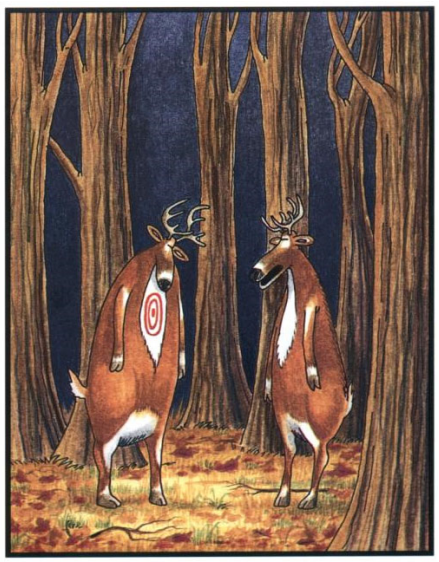
Why are we interested in "new" genes?

By Victor A. McKusick, M.D., Baltimore, Maryland

Ann. Int. Med. 49:556-567, 1958

Study of genetic factors is important:

(1) because potentially it will permit **recognition of genetic susceptibles**, for more effective application of preventive measures,



"Bummer of a birthmark, Hal."

GENEPI INNSBRUCK

4

Why are we interested in "new" genes?

By Victor A. McKusick, M.D., Baltimore, Maryland

Ann. Int. Med. 49:556-567, 1958

Study of genetic factors is important:

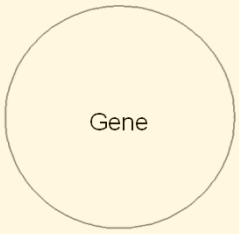
(2) because from our **understanding of the mechanism** whereby the gene or genes operate in these disorders can come preventive or therapeutic measures for breaking the chain leading to disease.

"This could be the discovery of the century. Depending, of course, on how far down it goes."

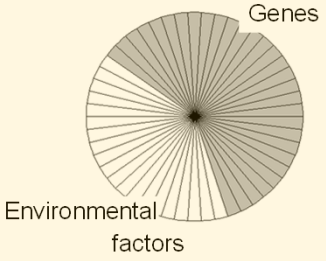
GENEPI INNSBRUCK

5

Monogenic and complex diseases



Monogenic diseases:
e.g. Morbus Huntington



Complex diseases:

e.g. Diabetes, myocardial infaction, overweight, cancer, …

Environmental factors are e.g. smoking, physical activity, nutrition, education, sun exposition, ….

GENEPI INNSBRUCK

6

Knockout versus small changes by polymorphisms

Knock-out

- Pronounced effects
- Animals: great models but not necessarily to extrapolate to humans
- Humans: often very rare cases

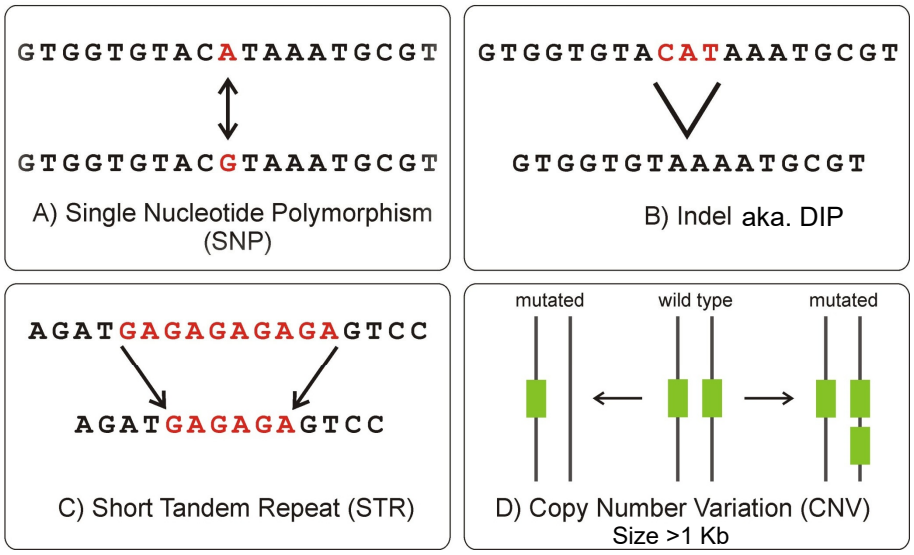
Polymorphism

- Small effects
- Usually investigated in humans
- Real in vivo conditions
- Thousands of people can be studied easily
- Sample sizes of thousands are required

GENEPI INNSBRUCK

7

Genetic variability

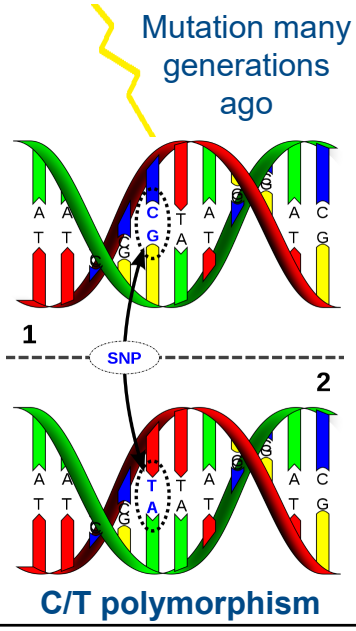


GENEPI INNSBRUCK

8

Single Nucleotide Polymorphisms (SNPs)

- Variations of single base pairs (bp) in the DNA sequence
- Heritable and stable.
- Account for 90% of the genetic variability
- Every 300 – 1000 bp
- At least 3 – 4 million SNPs per individual
- 10,000 – 11,000 non-synonymous SNPs per individual
- 700 million SNPs are described in databases



C/T polymorphism

GENEPI INNSBRUCK

9

Single Nucleotide Polymorphism (SNP)

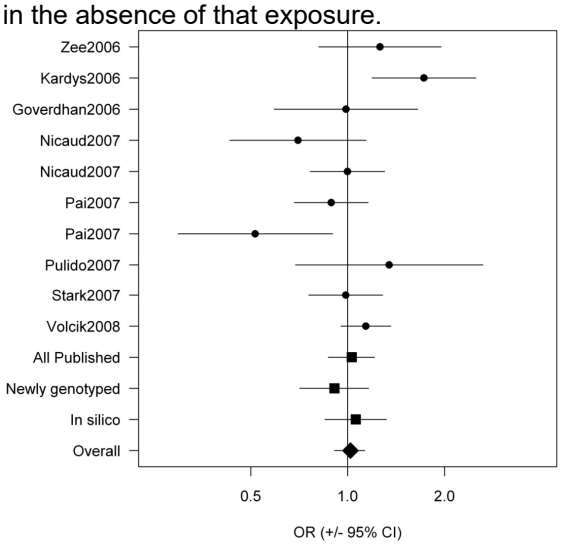
- **Coding SNPs within a gene**
  - synonymous exchanges: without influence on protein
  - non-synonymous exchanges: resulting in an AA exchange
- **SNPs within the regulatory regions:**
  - when and why a gene will be switched on or off
  - effect on quantity of protein production
- **SNPs within the untranslated regions**
  - with influence on mRNA stability
- **SNPs in intergenic regions**
  - functional consequences have to be evaluated

GENEPI INNSBRUCK

10

Some basics from epidemiology

- **Odds ratio**
  - Represents the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure.
  - Values between 0 and infinite (∞)
  - 1.00 = same odds
  - 1.50 = 50% higher odds
  - 2.00 = 100% higher odds
  - 0.50 = 50% lower odds
- **95% confidence interval (CI)**
- **Hazard ratio**
  - In case of prospective studies
- **Meta-analysis**
  - Combining data from more than one study



GENEPI INNSBRUCK

11

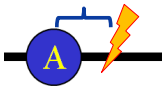
Overview

1. Background
2. **Association studies**
3. Genomewide association studies (GWAS)

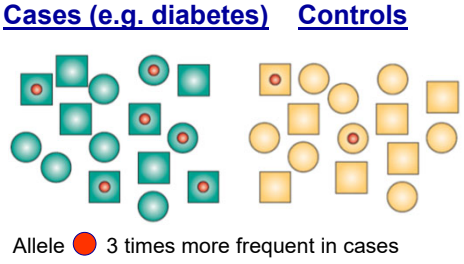
GENEPI INNSBRUCK

12

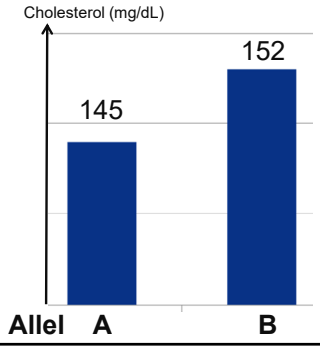
Principle of association studies



- **Qualitative analysis:**
  Preferential association of an allele with a disease status



- **Quantitative analysis:**
  Carriers of a various alleles differ in the mean values of the investigated parameter (e.g. cholesterol level)

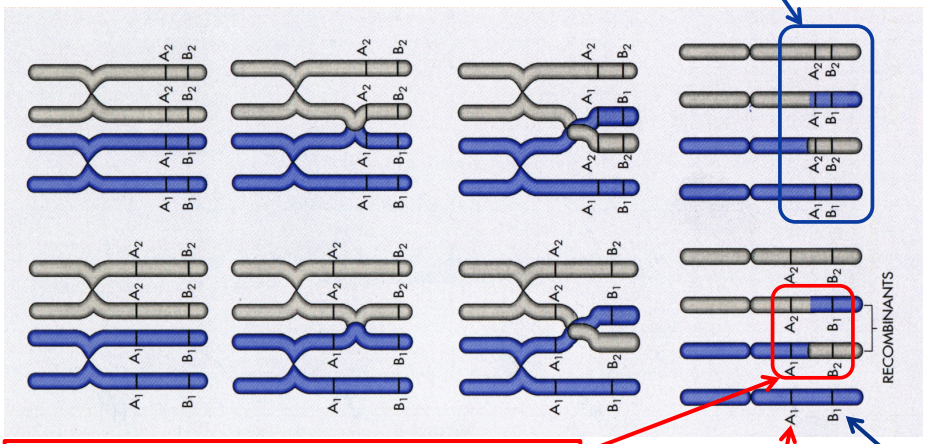


GENEPI INNSBRUCK

13

Remember: crossover and recombination during meiosis I

No recombination of alleles of the gene loci A and B since crossover outside the region between A and B



Recombination of alleles in the two gametes due to crossover

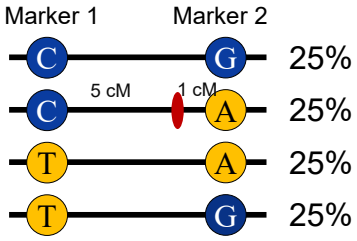
"Wanted" disease locus

Genotyped marker locus

Jorde, Carey, Bamshad, White: Medical Genetics

14

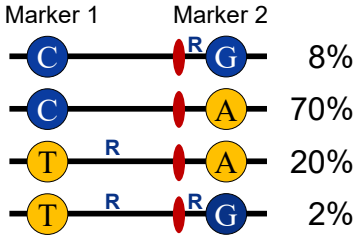
Basis of association: Linkage disequilibrium



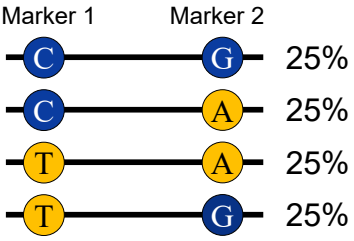
	Patients	Controls
Allele 2A	**90%**	**50%**
Allele 1C	78%	50%

After many generations

Patients



Controls

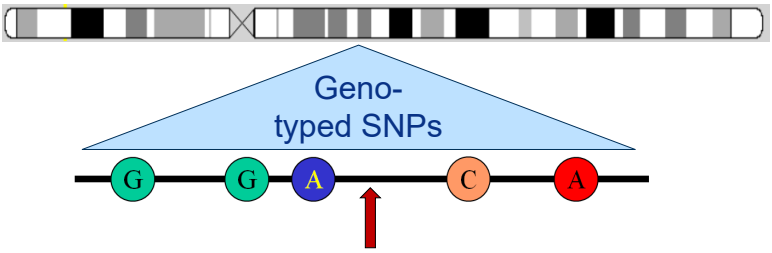


GENEPI INNSBRUCK

R … Recombination

15

Indirect association



Disease-causing variant we are searching

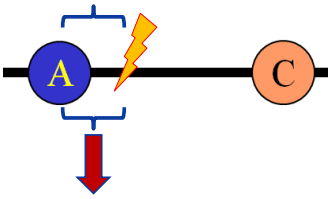
ACTAGAGCTACTACGAGGGACTAC...TACGAGCATCGACTA...GAGG
TAGAGCTATA...TTCT AGGCTA...CTACGATCGATC...ACGTAG...

High correlation

The genotyped SNP is a marker of the „non-genotyped“ disease-causing variant we are searching

The investigated genetic variant is in linkage disequillibrium with the causal variant

Due to the small distance there are rarely crossovers and recombinations during meiosis



Allele A can be observed more frequently in the diseased group

GENEPI INNSBRUCK

16

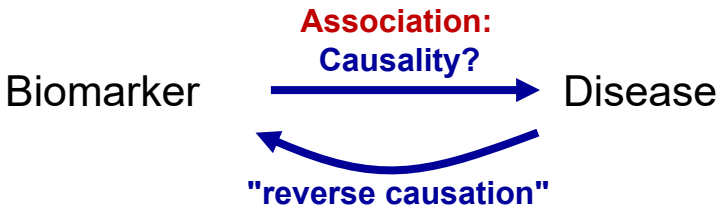
Mendelian Randomization study:

a great tool to support causality

GENEPI INNSBRUCK

17

Biomarker for diseases: causality or consequence?

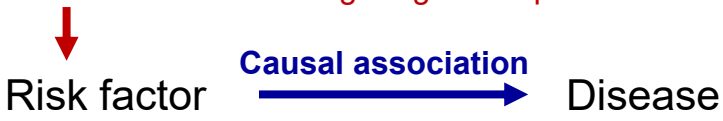


GENEPI INNSBRUCK

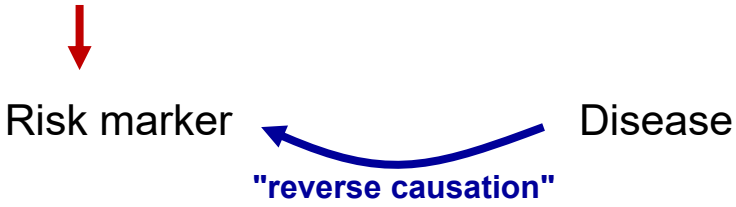
18

Biomarker for diseases: causality or consequence?

- Intervention including drug development



- No direct reason for intervention or drug development
- Useful for diagnostic purposes?



GENEPI INNSBRUCK

19

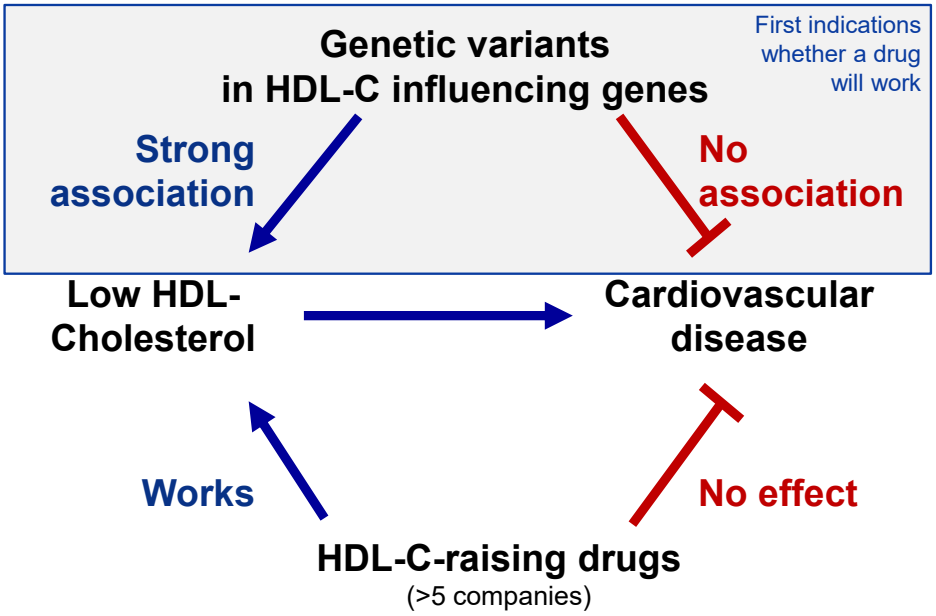
The big question for biomarkers

- You found an association with a disease
- **Risk factor or risk marker?**
- Classical epidemiological studies with prospective observation will last a long time and will not prove causality
- You have to decide now whether to go for drug development or not
- Worst case scenario: after 10-15 years of development the drug flops
- One reason might be that it is only a risk marker and not a risk factor.

GENEPI INNSBRUCK

20

Example of a flop: CETP inhibitors

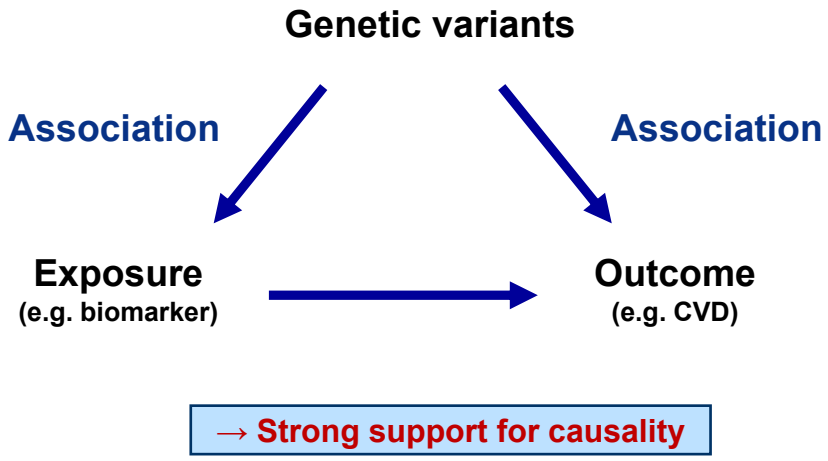


GENEPI INNSBRUCK

Kronenberg: Kidney Int. 89:747-9, 2016

21

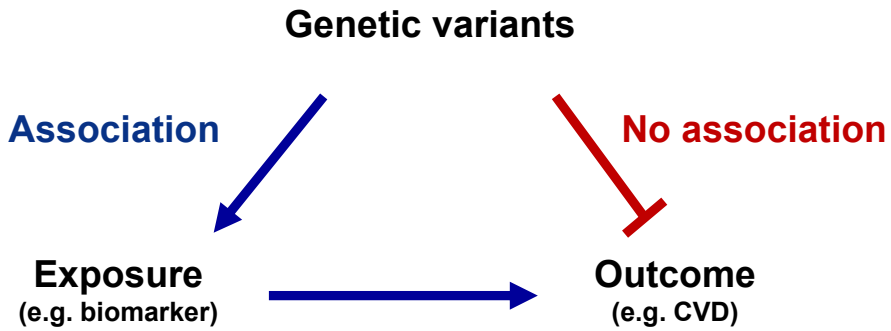
Mendelian randomization approach



GENEPI INNSBRUCK

22

Mendelian randomization approach

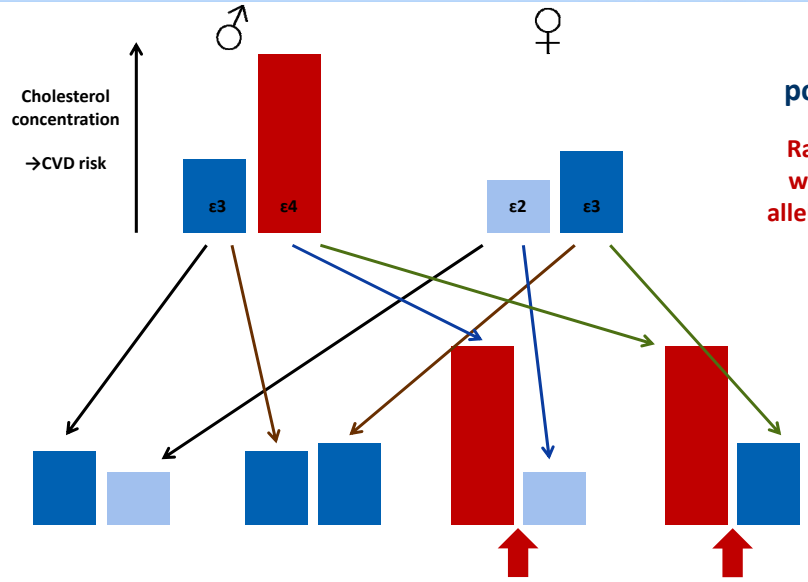


→ Causality is unlikely if the study is sufficiently powered

GENEPI INNSBRUCK

23

Mendelian randomization at the time of conception



ApoE polymorphism

Random decision which of the two alleles is transmitted to the child

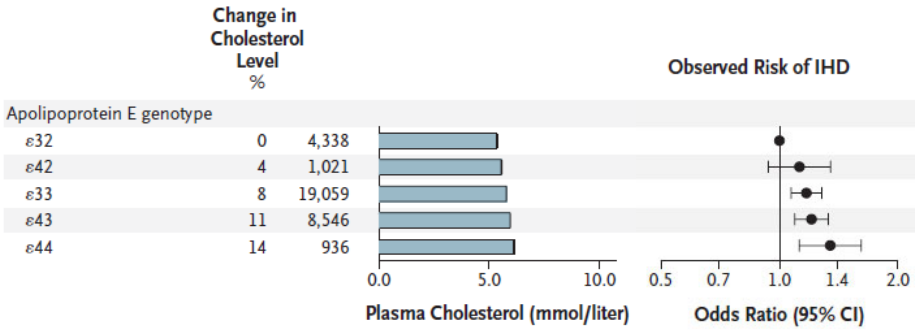
Increased risk for CVD due to lifelong exposure to higher cholesterol concentrations

GENEPI INNSBRUCK

Adapted from Cardiovasc.Drugs Ther. 30:87-100, 2016

24

ApoE, cholesterol and risk for CVD



GENEPI INNSBRUCK

Zacho et al.: NEJM 359:1897-1908, 2008

25

Example: Lipoprotein(a) - Lp(a)

The first example in history
a Mendelian randomization study has been performed
(in the early 1990-ies)

GENEPI INNSBRUCK

26

Experience of a young widow

- Husband 39 years of age
- Loses consciousness, cardiac arrest, revival not successful

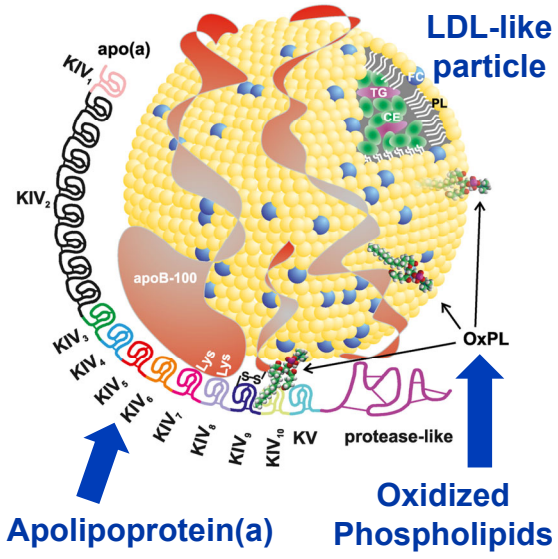
- No classical risk factors
- Healthy lifestyle, physically active
- Health checkup on a yearly basis

- Autopsy: most severe heart disease
- **Very high Lp(a) concentrations**

GENEPI INNSBRUCK

27

Lp(a) - the next target to fight cardiovascular disease



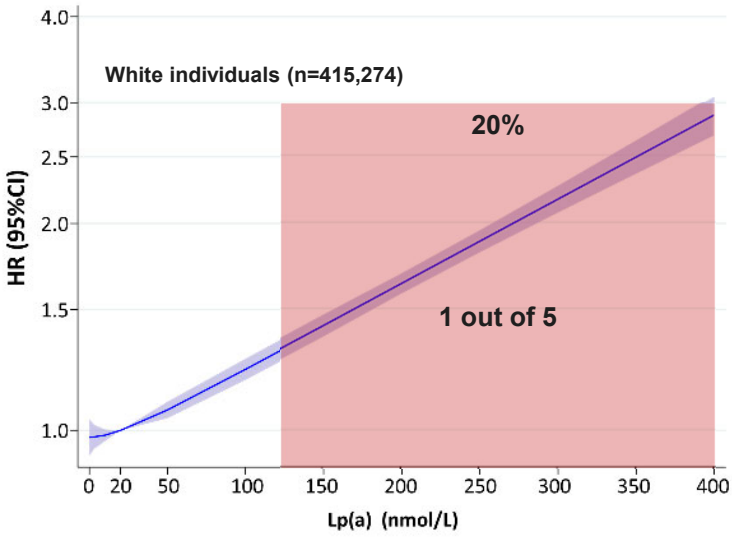
- **Described by K. Berg in 1963**
- **Concentrations are mainly genetically determined**
- **Shows only minor correlations with other lipoproteins (if any)**
- **An independent and one of the most important genetic cardiovascular risk factors**
- **One Lp(a) particle is about 6 times more atherogenic than one LDL particle**

GENEPI INNSBRUCK

28

The association between Lp(a) and major CVD* outcomes is <u>continuous</u>

independent of ethnicity → Lp(a) measurement is relevant globally



The higher the concentrations the higher the risk

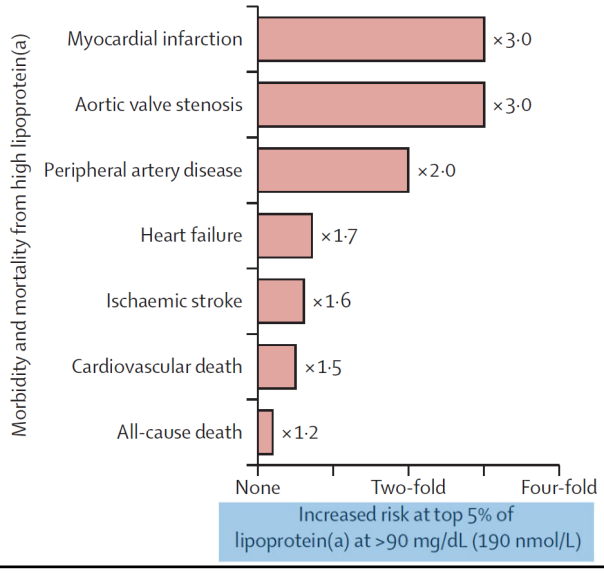
EAS

Kronenberg, Mora, Stroes et al.: Eur. Heart J. 43:3925-46, 2022

29

Lp(a) - association with various CVD endpoints

Data from Copenhagen studies



Enormous public health relevance:

- These endpoints are frequent
- High Lp(a) concentrations are frequent
- Similar population-attributable risk as smoking, diabetes and hypercholesterolemia

Please do not get confused by the two units! Never name an Lp(a) concentration without the unit!

Nordestgaard & Langsted: Lancet 404: 1255-64, 2024

30

Strong support for causality by Mendelian randomization studies

Genetic variants associated with increased Lp(a) concentrations
(e.g. small apo(a) isoforms, SNPs such as rs10455872, rs3798220)

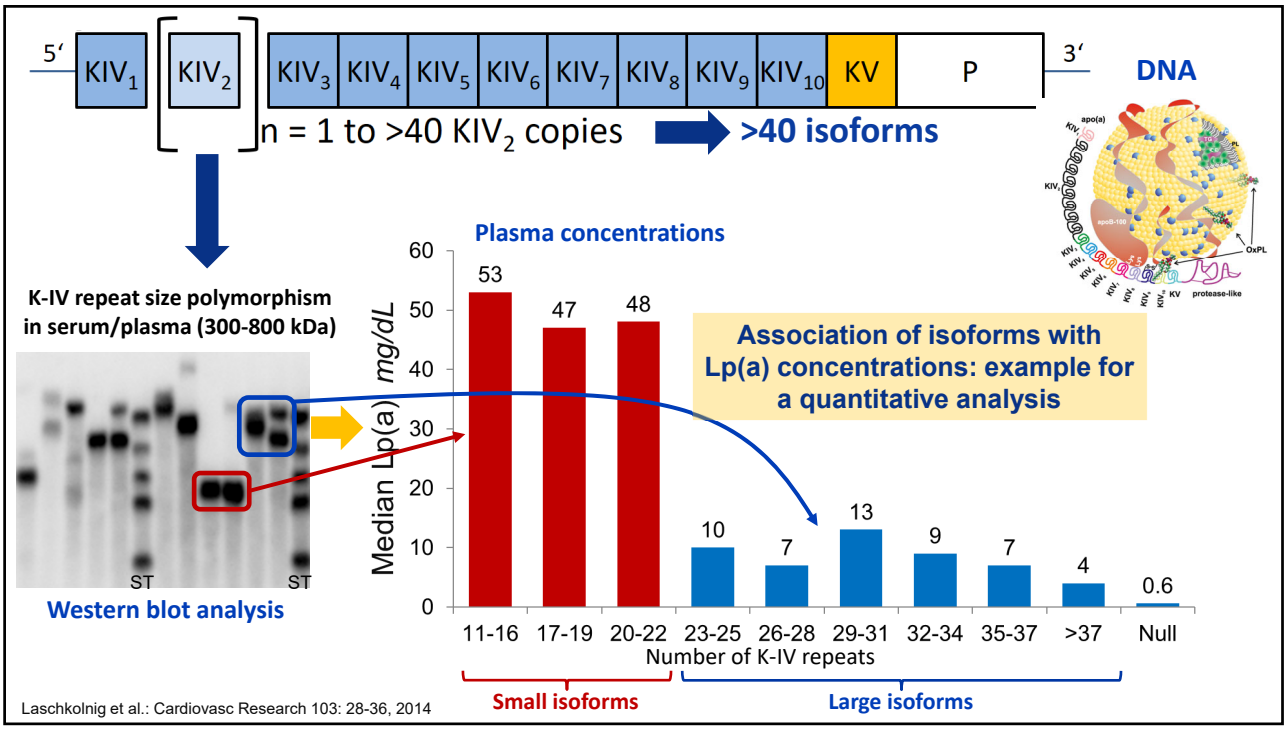
Lifelong elevated Lp(a) concentrations

Increased cardiovascular risk

Strong support for causal association

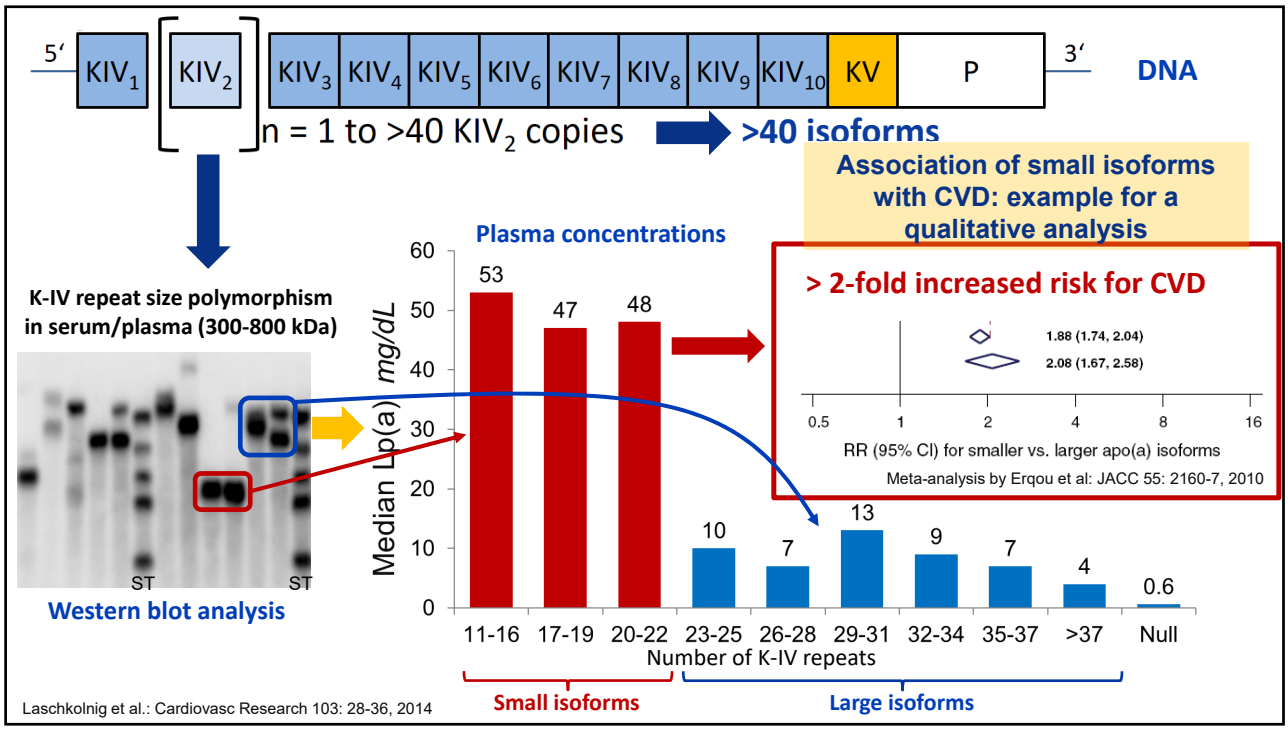
Kronenberg & Utermann: J. Int. Med. 273: 6-30, 2013

31



Laschkolnig et al.: Cardiovasc Research 103: 28-36, 2014

32



33

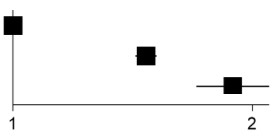
ESC European Society of Cardiology
European Heart Journal (2022) 43, 3925–3946 https://doi.org/10.1093/eurheartj/ehac361
SPECIAL ARTICLE
Miscellaneous

Lipoprotein(a) in atherosclerotic cardiovascular disease and aortic stenosis: a European Atherosclerosis Society consensus statement

Genetic variants associated with increased Lp(a) concentrations …

… are associated with increased cardiovascular risk

LPA genetic score Number of Lp(a)-increasing variants *	No. participants	No. events	HR (95%CI)
0 (median Lp(a) 14 nmol/L)	358,464	20,610	1.00 [reference]
1 (median Lp(a) 146 nmol/L)	77,655	6,167	1.47 [1.43, 1.52]
2 (median Lp(a) 262 nmol/L)	4,249	410	1.89 [1.70, 2.10]



* variants of rs10455872 and rs3798220 described by R. Clarke

EAS

Kronenberg, Mora, Stroes et al.: Eur. Heart J. 43:3925-46, 2022

34

EAS Consensus panel recommendations for Lp(a) testing

- **Lp(a) should be measured <u>at least once in all adults</u> to identify those with high cardiovascular risk**
- **Check the family in case of high Lp(a) of the index patient since Lp(a) concentrations is genetically determined**
- **Genetic testing is not required**

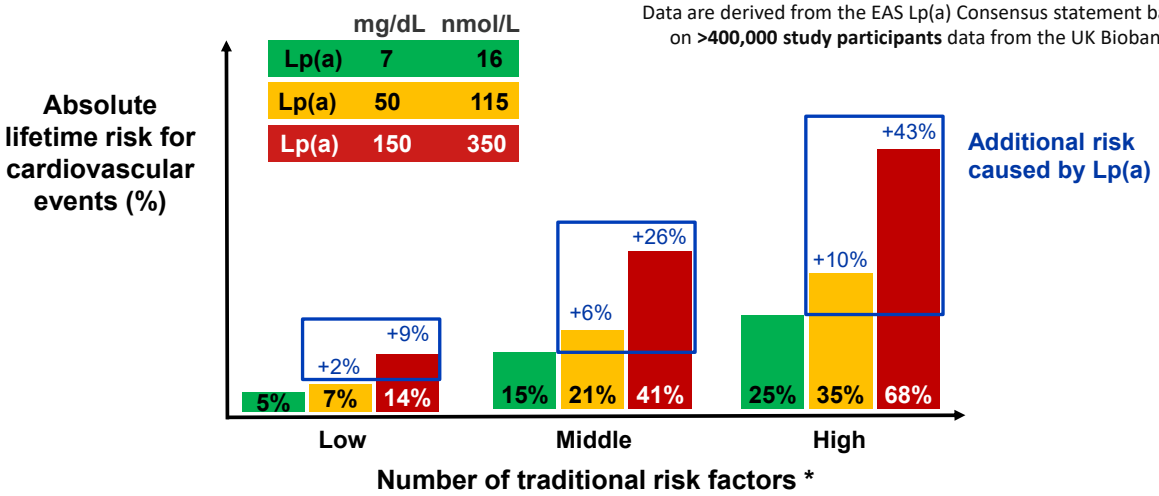
EAS

Kronenberg, Mora, Stroes et al.: Eur. Heart J. 43:3925-46, 2022

35

Lp(a) concentrations and cardiovascular risk

Data are derived from the EAS Lp(a) Consensus statement based on **>400,000 study participants** data from the UK Biobank



* **Traditional risk factors** are age, sex, blood cholesterol, blood pressure, smoking, diabetes, a family history of heart attacks in early life and body mass index

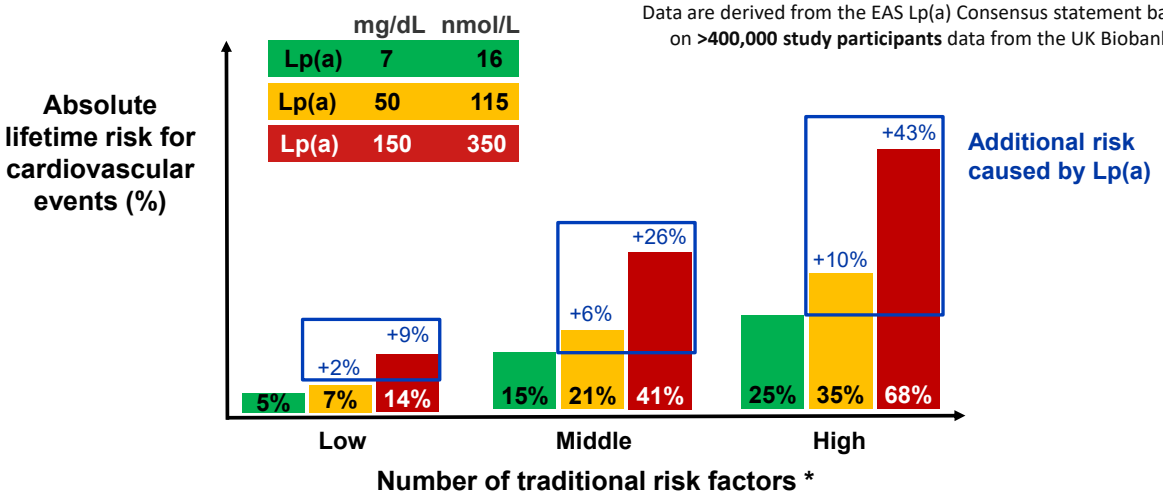
GENEPI INNSBRUCK

Kronenberg F: Curr. Atheroscler. Rep. 26:75-82, 2024

36

Lp(a) concentrations and cardiovascular risk

Data are derived from the EAS Lp(a) Consensus statement based on **>400,000 study participants** data from the UK Biobank



If Lp(a) level is not considered in risk estimation, ASCVD risk might be underestimated substantially

GENEPI INNSBRUCK

37

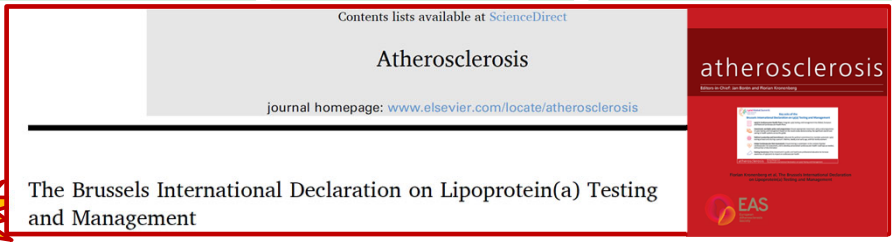
Endorse the declaration

Patronage of the Polish EU presidency

Key opinion leaders in the Lp(a) field	People with lived experience of high Lp(a)	Members of the EU-Parliament	Members of the WHO, EAS, IAS, WHF

Cost-effectiveness analysis of Lp(a) testing

NEW

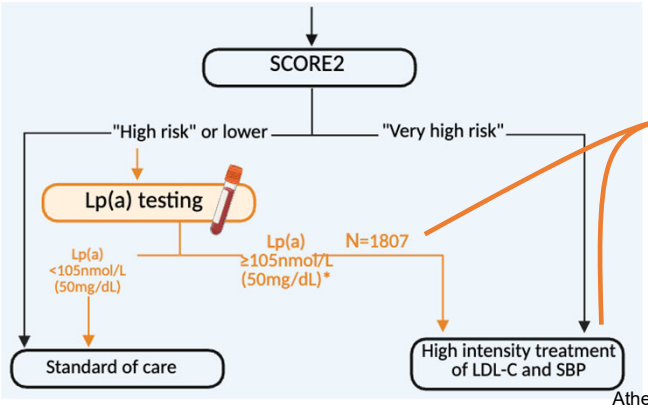


38

Primary prevention population aged 40-69 years

Standard of care versus introduction of Lp(a) testing

N=10,000



Applied to Austria for this generation:

- approx. 22,000 heart attacks
- approx. 5,000 strokes
- Cost savings of approx. € 1.4 billion

Testing followed by high-intensity treatment would prevent:

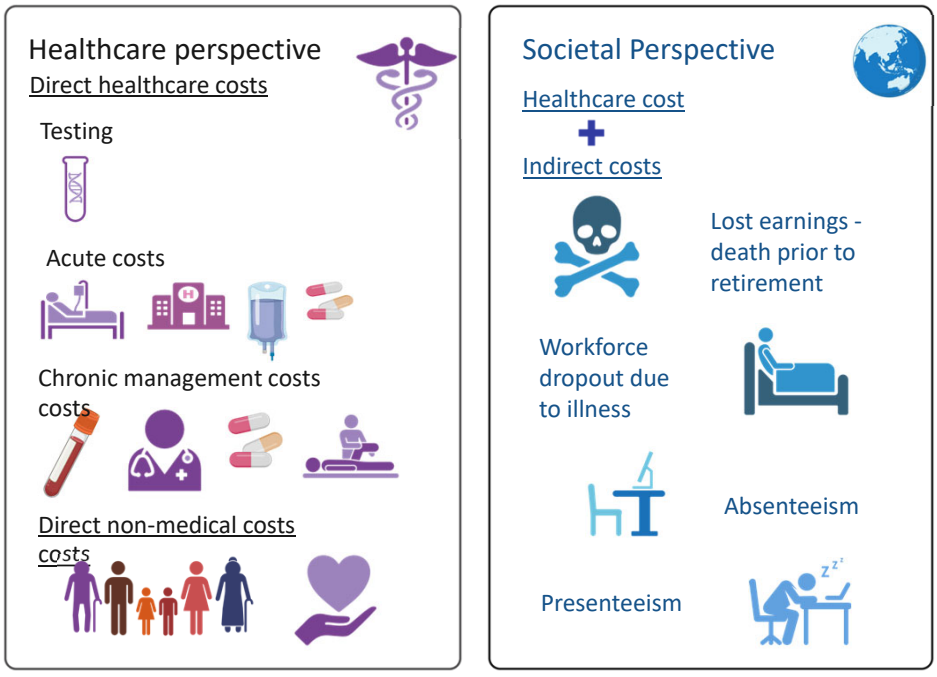
- 60 incident MIs,
- 13 incident strokes,
- 26 premature deaths
- e.g. 217 and 255 quality-adjusted life years gained in Australia and the UK

Atherosclerosis: https://doi.org/10.1016/j.atherosclerosis.2025.120447

atherosclerosis

Morton et al.: Lp(a) testing for the primary prevention of cardiovascular disease in high-income countries: a cost-effectiveness analysis

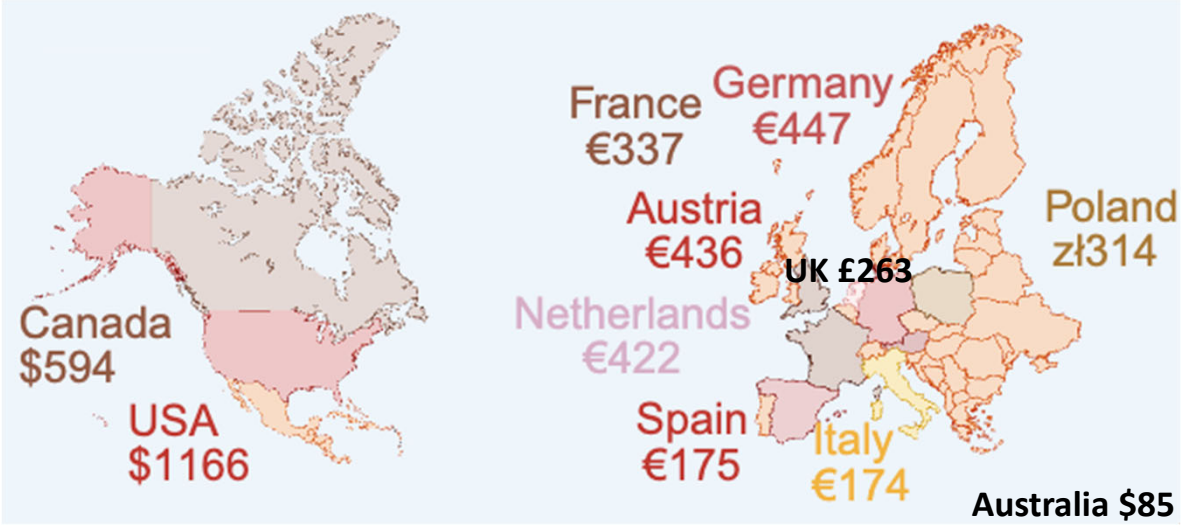
39



Morton et al.: Atherosclerosis: https://doi.org/10.1016/j.atherosclerosis.2025.120447

40

Costs saved per person



Morton et al.: Atherosclerosis: https://doi.org/10.1016/j.atherosclerosis.2025.120447 provided by Prof. Zanfina Ademi

atherosclerosis Morton et al.: Lp(a) testing for the primary prevention of cardiovascular disease in high-income countries: a cost-effectiveness analysis

41

Key asks of the Brussels International Declaration on Lp(a) Testing and Management

- **Lp(a) should be part of Cardiovascular Health Plans**
- **Policy and programmes for Lp(a) testing and management to save costs**
- **Political Leadership and Commitment for systematic Lp(a) testing with full reimbursement**
- **Global Cardiovascular Risk Assessment including Lp(a)**
- **Raising Awareness about Lp(a)**

Details can be found here:
https://doi.org/10.1016/j.atherosclerosis.2025.119218

Lp(a) Global Summit
International Task Force

You can endorse the Brussels International Declaration at:
https://fhef.org/brussels-international-declaration/

Each vote is a signal

atherosclerosis Kronenberg et al.: The Brussels International Declaration on Lp(a) Testing and Management

42

Overview

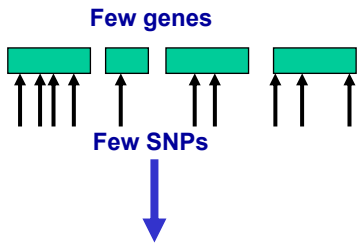
1. Background
2. Association studies
3. **Genomewide association studies (GWAS)**

GENEPI INNSBRUCK

43

Candidate gene approach vs. GWAS

Candidate gene approach



Association with phenotype

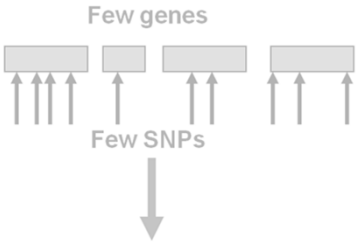
- Hypothesis-driven
- Biochemical or physiological *a priori* knowledge
- Few genes identified

GENEPI INNSBRUCK

44

Candidate gene approach vs. GWAS

Candidate gene approach



Association with phenotype

- Hypothesis-driven
- Biochemical or physiological *a priori* knowledge
- Few genes identified

Genomwide association study GWAS

23 chromosomes

0.5 – 10 million SNPs

Association with phenotype

- **"Hypothesis-free"** (unbiased)
- No *a priori* knowledge
- New pathways
- Small effects detectable
- **Very large sample sizes required**

GENEPI INNSBRUCK

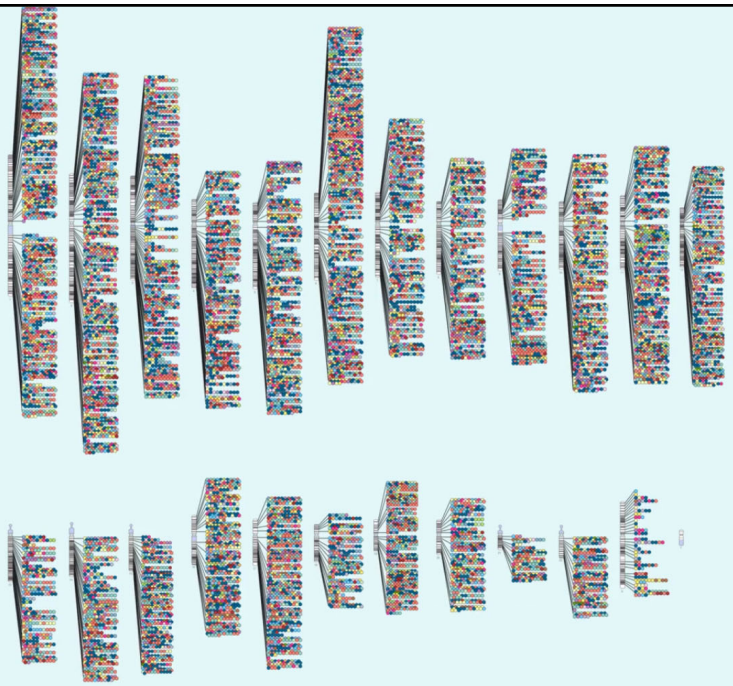
45

2018 Apr

Associations: 69,885

Studies: 5,152

Papers: 3,378



www.ebi.ac.uk/gwas

46

Gain in detected genes by GWAS

Examples for metabolic traits

Disease	before 2007	2007 onward
Type 2 DM	3	50
Body mass index	1	30
Glucose or insulin	1	15
Fat distribution	0	20
Lipids	16	95
Total	**21**	**202**

7 examples of autoimmune diseases

Disease	before 2007	2007 onward
Ankylosis spondylitis	1	13
Rheumatoid arthritis	3	30
Systemic lupus eryth.	3	31
Type 1 DM	4	40
Multiple sclerosis	1	51
Crohn's disease	4	67
Ulcerative colitis	3	44
Total	**19**	**277**

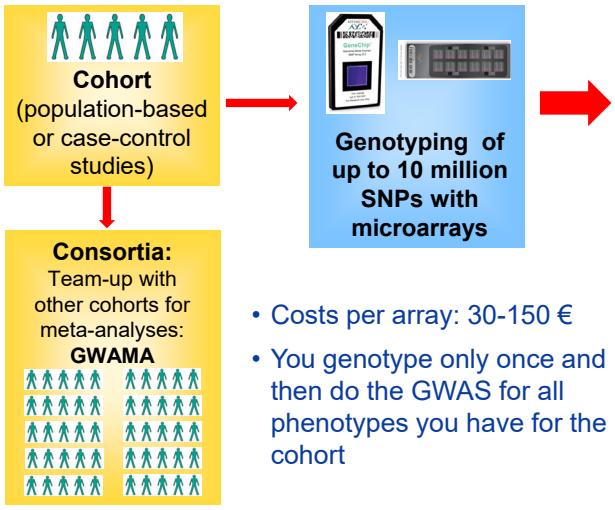
Since 2012 the number of known genes has further increased by 5- to 10-fold

GENEPI INNSBRUCK

Visscher et al.: Am.J.Hum.Genet. 90:7-24, 2012 (updated)

47

Design and cost-performance ratio



- Costs per array: 30-150 €
- You genotype only once and then do the GWAS for all phenotypes you have for the cohort

Association with phenotypes (e.g.)

- BMI
- Waist
- Blood pressure
- QT interval
- Smoking

Lab values

- Lipids
- Kidney function
- CRP
- Hemoglobin
- …

Diseases

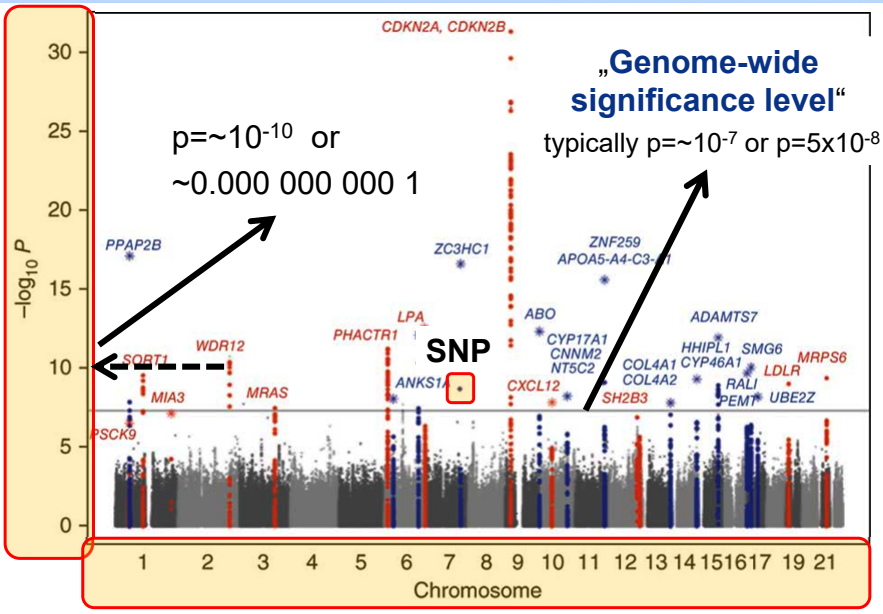
- CAD
- Stroke
- Ankle-brachial-index
- Cancer types

Whatever has a genetic component and is measured

GENEPI INNSBRUCK

48

The Manhattan Plot



GENEPI INNSBRUCK

Schunkert et al.: Nature Genetics 43:333-338, 2011

49

Genome-wide association studies (GWAS)

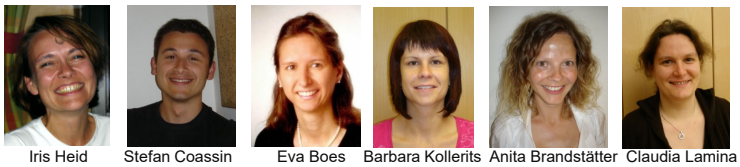
- **Examples:**
  - Lipids
  - Type 2 diabetes mellitus
  - Blood pressure
  - BMI

GENEPI INNSBRUCK

50

GWAS: Lipids

- **Consortium:**
  - Established during fall 2005
  - Together with Helmholtz-Zentrum München
  - 11 members studying various phenotypes
- **Innsbruck Group:**
  - Lipid metabolism
  - HDL-C as a starting point
  - Quantitative trait considered more powerful
- **Population and Genotyping:**
  - 1644 population-based subjects from KORA
  - Affymetrix 500K SNP chip

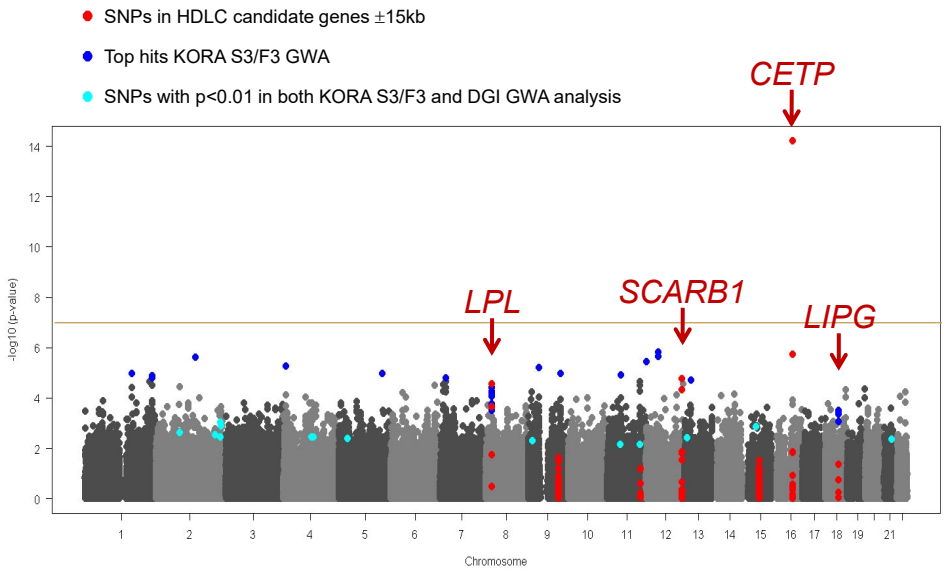


Iris Heid Stefan Coassin Eva Boes Barbara Kollerits Anita Brandstätter Claudia Lamina

GENEPI INNSBRUCK

51

GWAS: HDL cholesterol



GENEPI INNSBRUCK

Heid et al.: Circ. Cardiovasc. Genet. 1:10-20, 2008

52

GWAS on Lipids (TC, HDLC, LDLC, TG): next steps

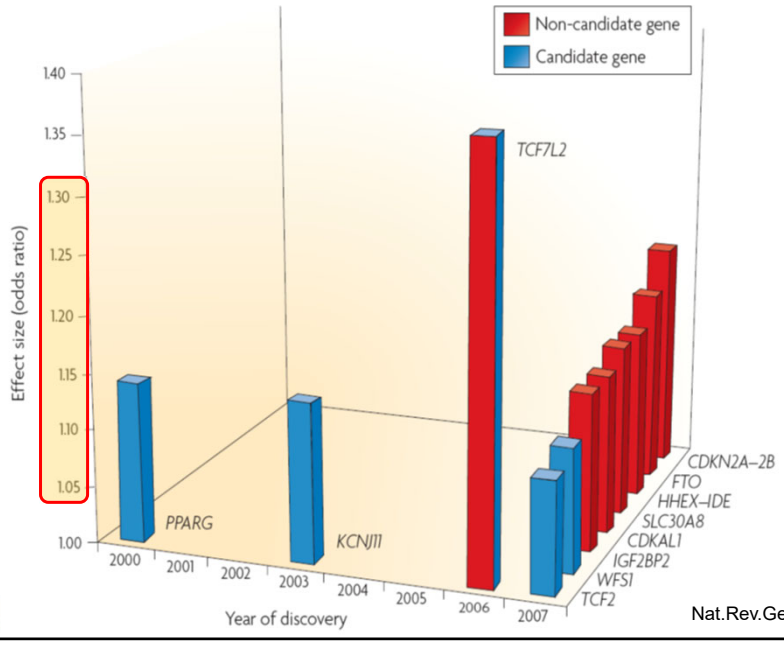
- **First own GWAS** *(Circ. Cardiovasc. Genet. 2008)*
  1644 probands: found nothing new
- **Engage Consortium:** *(Nature Genetics 2009)*
  - 22,000 probands: **22 genes found associated**
- **Global Lipids Genetics Consortium:** *(Nature 2010)*
  - >100,000 probands: **95 genes found associated**
- **Global Lipids Genetics Consortium:** *(Nature Genetics 2013)*
  - >188,000 probands: **roughly 155 genes found associated**
- **Global Lipids Genetics Consortium:** *(Nature 2021)*
  - 1,65 million probands: **>900 Gene**

- **The more individuals you investigate, the more genes you will find**
- **Functional characterisation for most of the genes has to be done**

GENEPI INNSBRUCK

53

GWAS: Type 2 diabetes mellitus



2017:

≈ 86 genes

2019:

≈ 250 genes

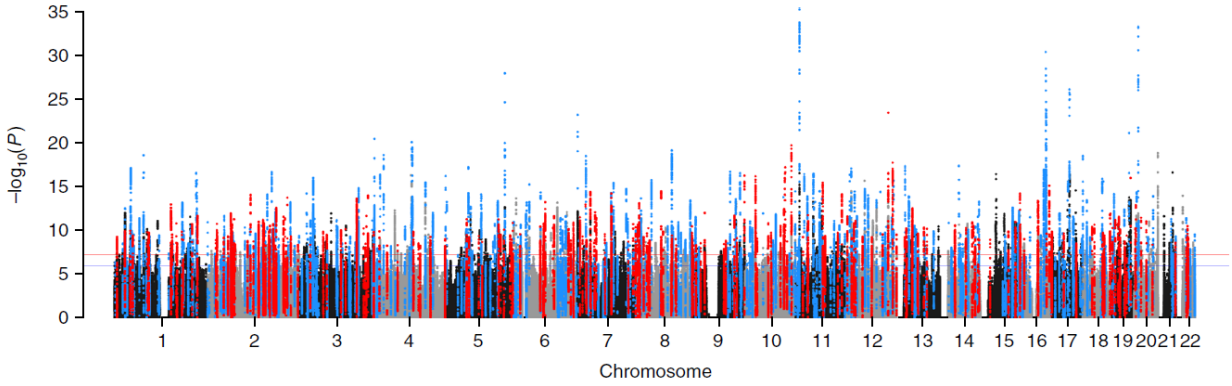
T. Frayling:
Nat.Rev.Genet. 8:657-62, 2007

GENEPI INNSBRUCK

54

GWAS and blood pressure traits

- Phenotypes: systolic and diastolic BP, pulse pressure
- > 1 million study participants
- 901 genetic loci in total (535 novel)



GENEPI INNSBRUCK

Evangelou et al.: Nature Genet. 50:1412-25, 2018

55

Genes for body mass index and overweight

Own behavior (lifestyle)

Genetics

- 536 genetic loci detected
- Many of them play a role in the brain by
  - ⇨ **Regulation of appetite**
  - ⇨ Neuronal component of overweight
- Explain roughly 5% of BMI

GENEPI INNSBRUCK

56

Contribution of single genes to overweight

- **Very few with strong effects:** risk increase by 10 to 30% per allele
- **More with moderate effects:** risk increase by 3 to 10% per allele
- **Many more with tiny effects:** risk increase by 0.1 to 3% per allele

GENEPI INNSBRUCK

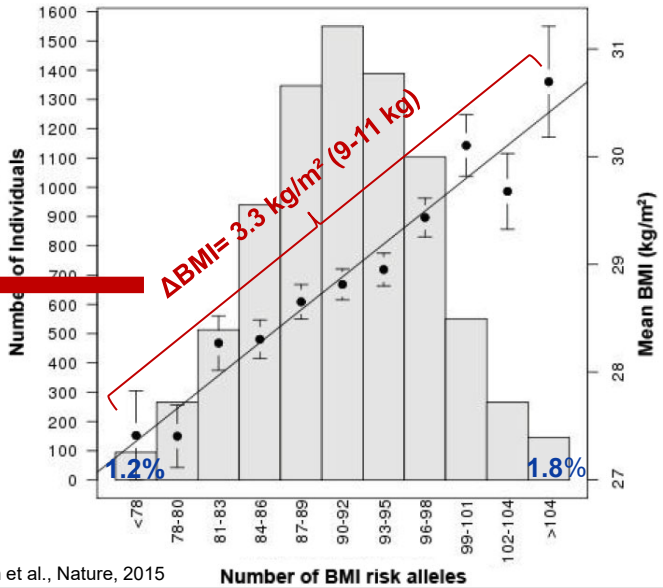
57

Development of SNP-Risk-Scores

GENEPI INNSBRUCK

58

Risk score of 97 SNPs for BMI: 0-194 risk variants



GENEPI INNSBRUCK

Shungin et al., Nature, 2015

59

Risk score for 65 SNPs for diabetes

Many risk variants (top quintile)

2.7-fold increased risk to develop a diabetes in the future

Few risk variants (bottom quintile)

GENEPI INNSBRUCK

Talmud et al., Diabetes 64:1830-40, 2015

60

Where is the reward?

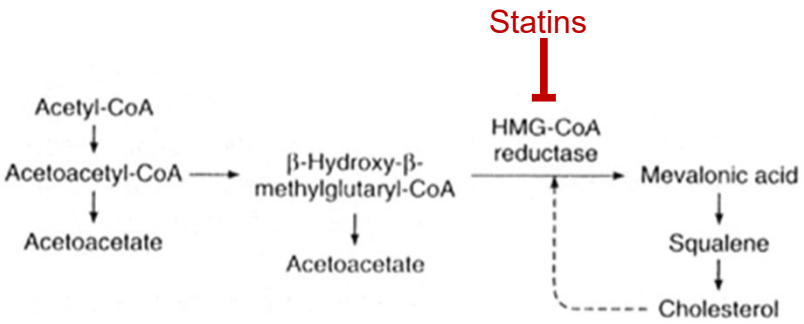
Can a single gene explaining less than 1% of the traits' variance still be useful for anything?

GENEPI INNSBRUCK

61

Statins: HMG-CoA-Reductase-Inhibitors

- **Mechanism of action**
  - Inhibition of HMG-CoA-Reductase: this enzyme catalyzes the conversion of HMG-CoA to mevalonic acid: an early and rate-limiting step in cholesterol biosynthesis.
  - Results in higher expression of LDL receptor which decreases LDL cholesterol



GENEPI INNSBRUCK

62

Polymorphisms in HMG-CoA-R gene region

- **GWAS results for HMG-CoA-reductase**
  - Very small effects
  - Were not detected in the first GWAS
  - This gene was only detected after investigation of at least 10.000 subjects
  - Single polymorphisms explain far less than 1% of the cholesterol concentrations within a population
  - Nevertheless, the most sucessful drug target for lipid metabolism
- **Other drug targets within the 900 lipid genes?**
  - CETP, ABCA1, PCSK9
  - Others?

GENEPI INNSBRUCK

63

Conclusions on GWAS

- An **hypothesis-free approach**
- Never before such a gain in gene-phenotypic information
- New genes for CAD, diabetes, cancer, kidney function…
- Odds ratios between 1.02 and 1.40
- To have the equipment is only the smallest step
- Very large studies of well phenotyped cohorts are necessary
- Works only within a very well constructed network between genetics, epidemiology, statistics, informatics, genomics
- Data sharing (a lot is already on the web)
- Non-coding SNPs and "gene deserts" can no longer be neglected
- A lot to learn about regulatory regions
- Functional characterization of "new" genes will need decades

GENEPI INNSBRUCK

64

Why are we searching these many genes?

Improvement of risk prediction (genetic risk scores)

Identification of new drug targets

- PCSK9 increases LDL cholesterol: discovered by genetic studies
- PCSK9 inhibitors lower LDL cholesterol by 60%

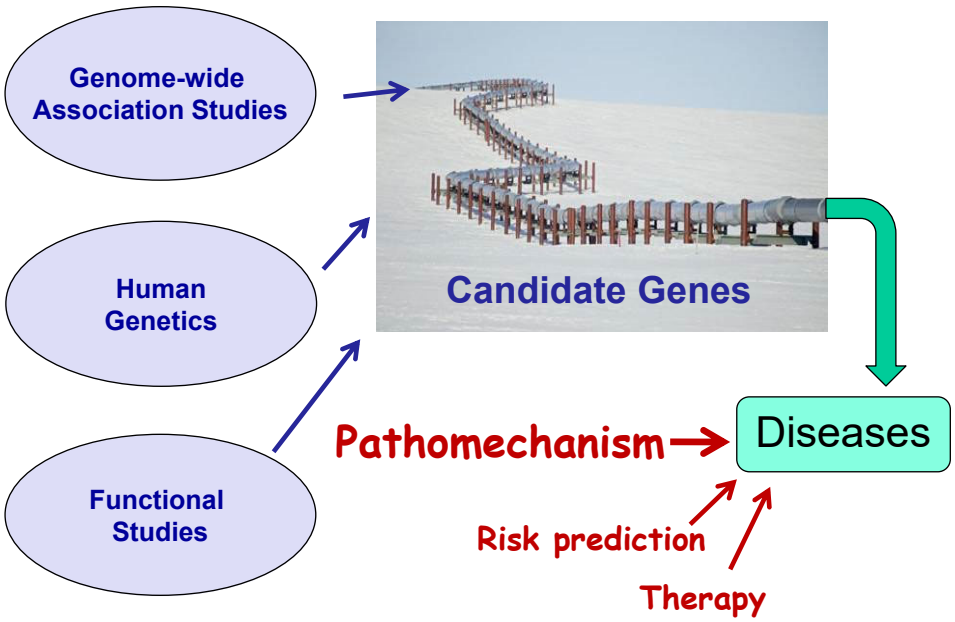
Exclusion of drug targets

- CETP increases the "good" cholesterol
- Development of CETP inhibitors to increase HDL cholesterol
- Billions of investment without lowering of heart attacks
- Genetic studies would have predicted the failure of these drugs
- Newer developments of CETP-inhibitors lower LDL-C and Lp(a)

GENEPI INNSBRUCK

65

Gene hunting: an interdisciplinary approach



GENEPI INNSBRUCK

66