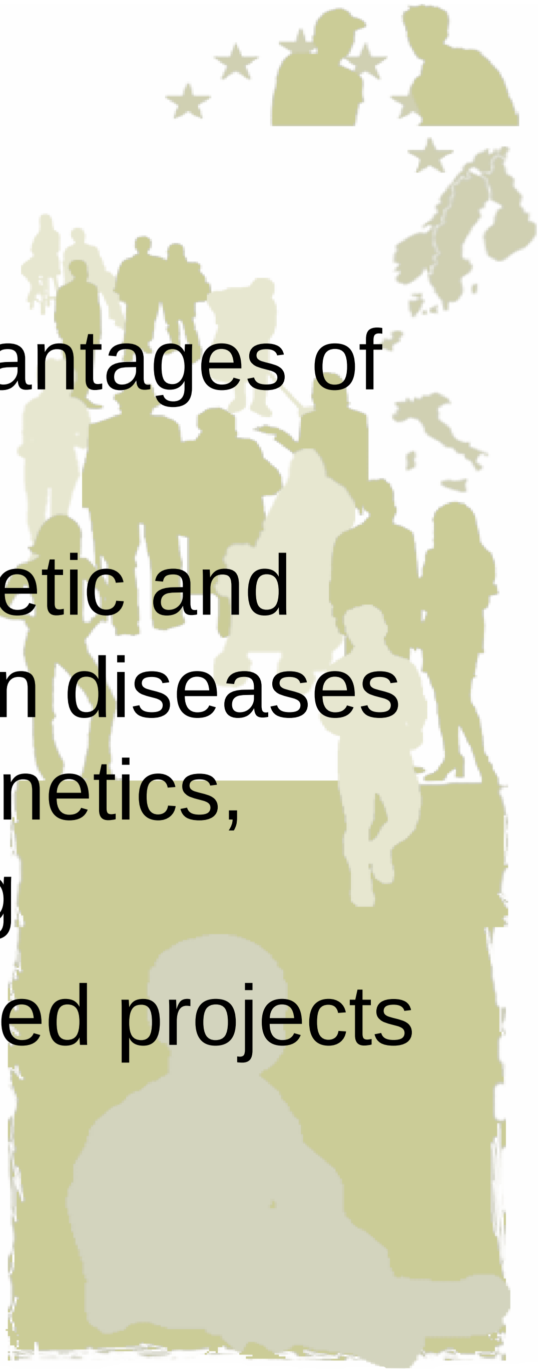
GenomEUtwin

(www.genomeutwin.org)

Genome-wide analyses of European twin and population cohorts to identify genes predisposing to common diseases

GENOMEUTWIN

GenomEUtwin



- Aimed to capitalize special advantages of Europe in population genetics
- Goal was to identify critical genetic and life-style risk factors for common diseases using European strengths in genetics, epidemiology and biocomputing
- An FP5 pilot project for Integrated projects
- Duration 2002-2007

EU: Genomics and Human Health

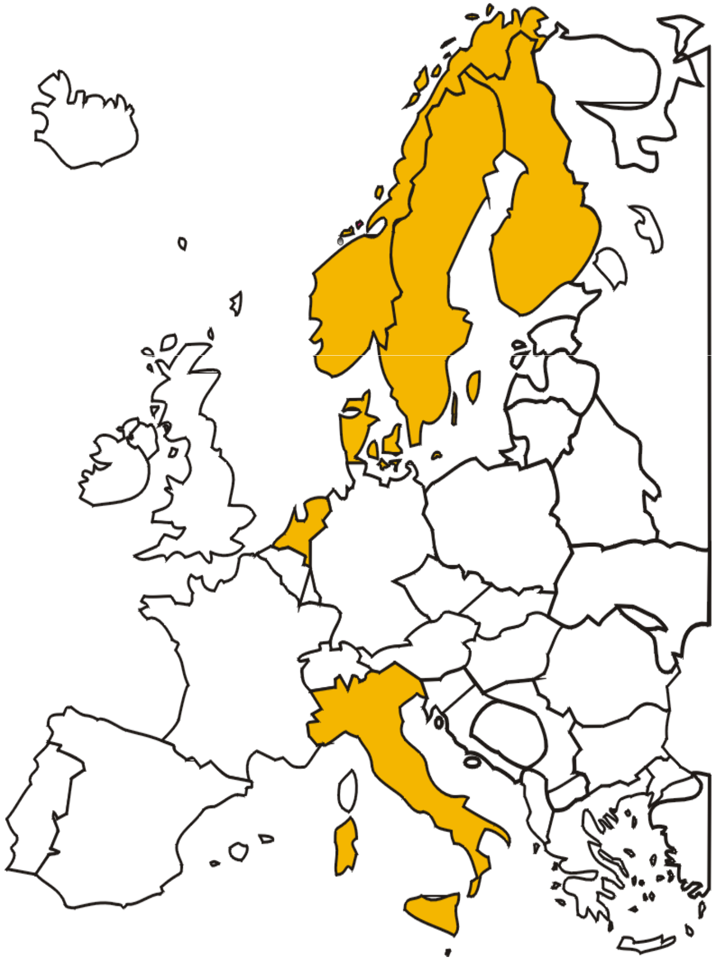
- In early 2000, EU called for proposals for project funding for projects utilizing genomics to promote human health
  - 5th framework: LIFE
- EU tries to be better in selecting the projects to be selected for funding by lessening the way politics affect the process

EU: Genomics and Human Health

- The process was designed to be transparent, based on the scientific merits and promoting European health
- A special emphasis was laid on the "added European value“
- Three projects were funded
  - Proteomics, Transgenics, Genetic Epidemiology

GENOMEUTWIN

GENOMEUTWIN: Twin Cohorts



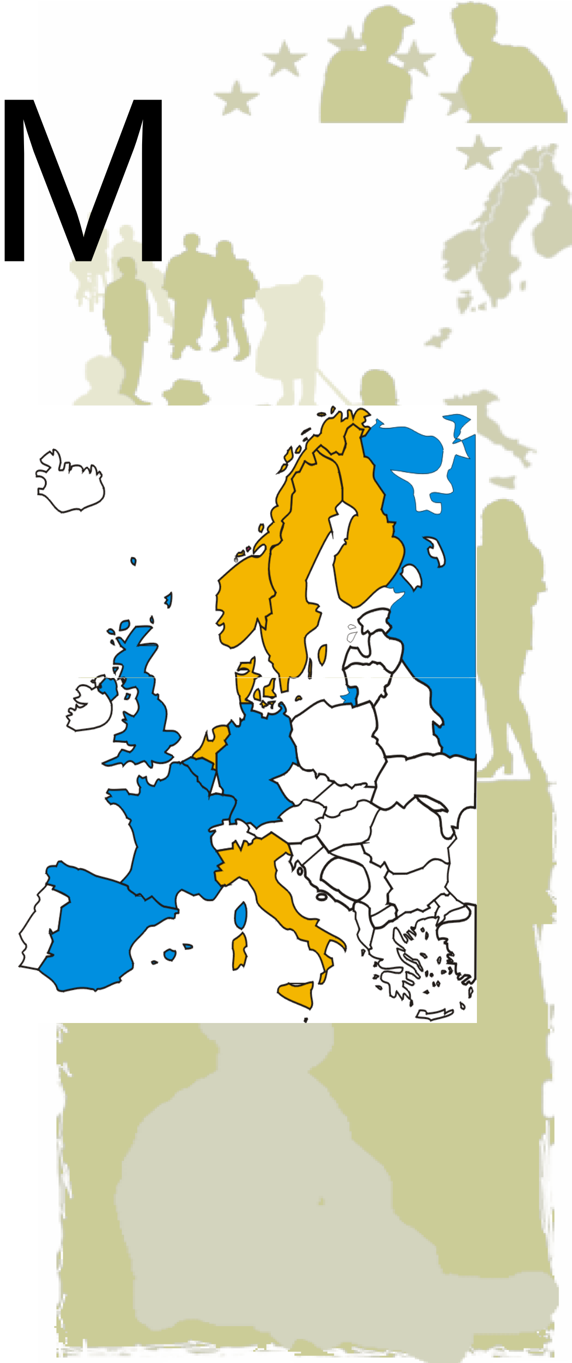
GENOMEUTWIN

Consortium

- In total, over 600 000 living twin pairs (2/3 dizygotic) are in the registers
- Construction of a federated database of the material with height, BMI and longevity as examples of quantitative data and CHD and migraine as qualitative data.
- Discordant pairs to be collected for all phenotypes and concordant affected for migraine.

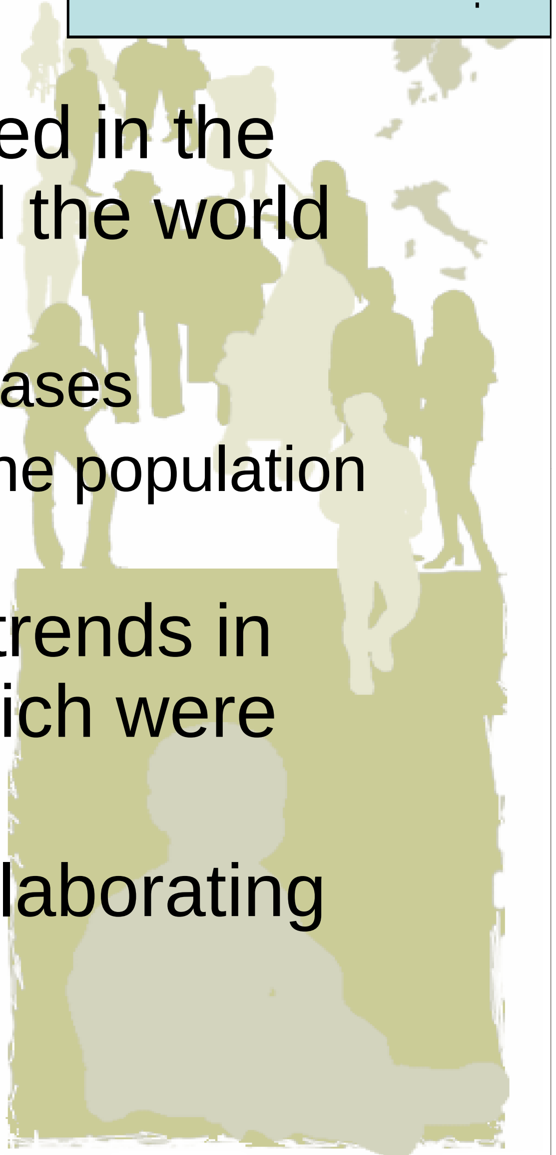
MORGAM

The Genetic Continuum
of the Largest Population Study
in the World: MONICA



GENOMEUTWIN

Background: MONICA



- The MONICA Project was established in the early 1980s in many centres around the world to:
  - **MONI**tor trends in **CA**rdiovascular diseases
  - relate these to risk factor changes in the population over a ten year period
- It was set up to explain the diverse trends in cardiovascular disease mortality which were observed from the 1970s onwards
- There were total of 32 MONICA Collaborating Centers in 21 countries

Background 2: MORGAM

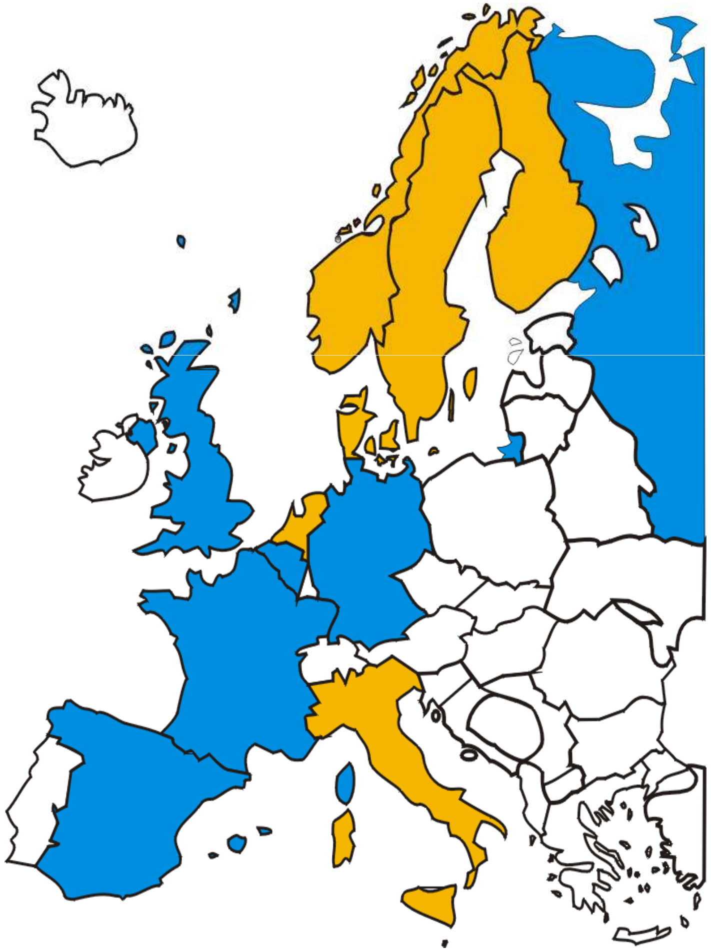
- The genetic part of MONICA: the cohort component
- A selected subset of MONICA centres and some additional centres meeting MONICA standards and having collected DNA
- 10 countries: Belgium, Denmark, Finland, France, Germany, Italy, Netherlands, Russia, Spain, Sweden, UK: 29 cohorts
- The populations are, however, being followed and screened for CVD end-points
- EU gave funds for so-called MORGAM project to
  - Construct a monograph of MONICA results
  - Collect DNA from participating centres for genetic studies

Design

- A case-*cohort* setting with several endpoints
  - Controls will be a random sub-sample of the cohort stratified by age and sex
  - Cohort: approximately 2x controls
  - The setting allows screening for several endpoints as well as direct determination of allele frequencies
- Many phenotypes in common with twin cohorts

GENOMEUTWIN

Twin cohorts + Morgam



GENOMEUTWIN

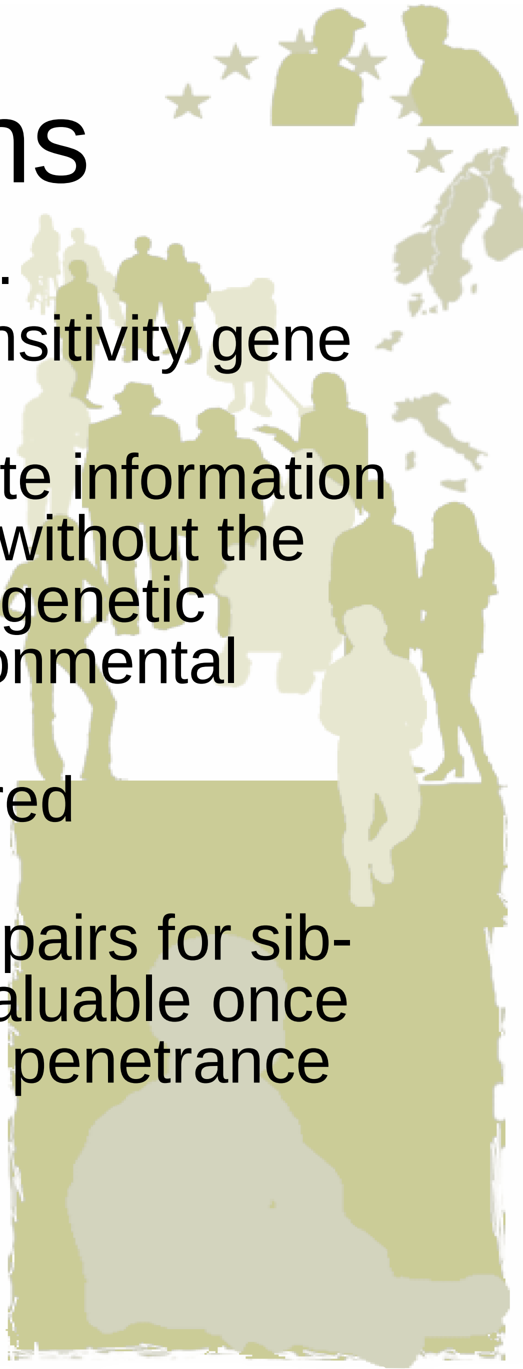
Advantages, 1

- The twin samples are unascertained for a disease phenotype
    - – Unselected representative population-based sample,
    - – Is it really? *Barker hypothesis *Twinning as a genetic trait
- Environmental aspects are controlled by the twin status.
- Easy access to core families

Advantages, 2

- Cohorts are well followed and recorded
  - Most followed with regular questionnaires from 70’s
  - All can be linked to national health care registries
  - The cohorts emerge from high quality European health system
- Questions:
  - Is diversity here a virtue or a hazard?
  - How to select phenotypes for study?
    - Heritability as an inclusion criteria

Monozygotic Twins



- Identical twins make the collection unique.
- A possibility to follow both level and of sensitivity gene effects versus environmental effects.
- Even though the MZ twins do not contribute information on the marker gene effect, they do allow (without the need to be typed) an estimate of the total genetic variance and resolve it from shared environmental variance.
- Family studies typically have ignored shared environmental effects
- Additionally, while the emphasis is on DZ pairs for sib-pair linkage, typing of MZ pairs will be invaluable once mutations have been found for evaluating penetrance and/or environmental sensitivity.

Ethical considerations

- Consent
- Data security
- Quality issues
- Informing participants
- Access to data and samples
- Public/Private

Initial Steering Group

Co-ordinator: Prof. Leena Peltonen, National Public Health Institute (PI)

Dr. Markus Perola, National Public Health Institute (NPHI), Helsinki (Secretary of the Steering Group)

Prof. Ann-Christine Syvänen, Uppsala University (Genotyping Core)

Prof. Kaare Christensen, University of Southern Denmark (Epidemiological Core)

Prof. Kirsten Ohm Kyvik, University of Southern Denmark (Danish Twin Registry)

Prof. Alun Evans, University of Belfast (Epidemiological Core and MORGAM)

Dr. Kari Kuulasmaa, National Public Health Institute, Finland (Database Core)

Prof. Nancy Pedersen, Karolinska Institutet, Sweden (Database Core and Swedish Twin Registry)

Dr. Jennifer Harris, The National Institute of Public Health, Norway (Ethical Issues Core and Norwegian Twin Registry)

Prof. Aarno Palotie, Finnish Genome Center (Genotyping Core)

Dr. Lodewijk Sandkuijl, Erasmus University, Netherlands (Statistical Core)

Prof. Dorret Boomsma, Vrije Universiteit, Netherlands (Dutch Twin Registry)

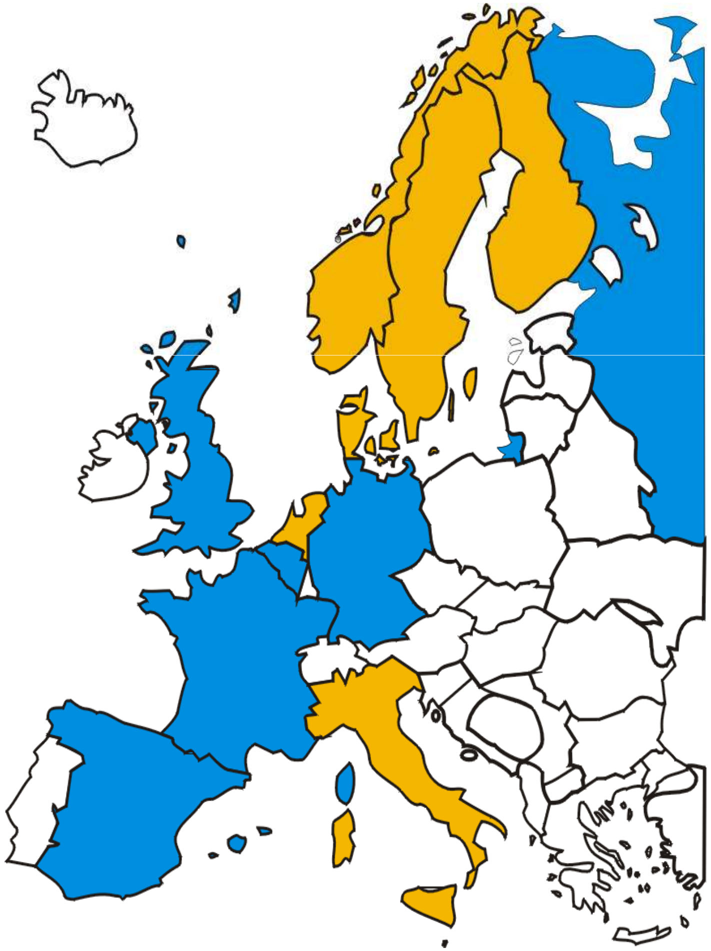
Prof. Jaakko Kaprio, University of Helsinki (Finnish Twin Cohort Study)

Dr. Antonia Stazi, Istituto Superiore di Sanità (Italian Twin Registry)

GENOMEUTWIN

European Populations studied in GenomEUtwin

Twin cohorts from
U.K. (T. Spector) and
Australia (N.Martin) also
included



GENOMEUTWIN

GenomEUtwin
research, networking, training

Population cohorts

- Australian twins
- Danish twins
- English twins
- Finnish twins
- Italian twins
- Dutch twins
- Norwegian twins
- Swedish twins
- MORGAM

Intellectual core facilities

- DNA isolation and genotyping (Helsinki,Uppsala)
- Epidemiological expertise (Odense)
- Database expertise (Stockholm)
- Biocomputing expertise (Leiden)
- Ethical and legal expertise(Oslo)

GENOMEUTWIN

Non-scientific accomplishments of GenomEUtwin

- Intensified collaboration between sites and groups
- Website central and valuable source of information
- Cataloguing effort for accessible study samples
- Harmonization of phenotypes and/or documentation of variants
- Federated database for pooling data and analyses
- Student exchanges and numerous workshops

Increased interaction and collaboration of European epidemiologists, geneticists & clinicians

- Many new EU projects largely leveraged by GenomEUtwin, using resources established by the project, including:
- Eurohead, GEHA, Euroclot, Diogenes
- Catalyzer of MORGAM expansion
- Significant contribution to P3G and to strategic planning of FP7

GENOMEUTWIN

Challenges of GenomEUtwin

- The database concept and the data transfer was labor intensive and the motivation of partners was not always the best possible
- Need for education of short vs.long term goals of the project, interactions with scientists of the groups
- Projects funded from other resources have greatly contributed. However, some previous commitments of partners make some data transfer issues problematic
- Reliance on mainly existing phenotypes, standardization problems and prior commitments to earlier ways of coding and treating data in each cohort
- The value of pooled data began to be recognized, leading to ENGAGE (2008 -)

Pooled studies of heritability and shared environment

- For many diseases, affected twin pairs are rare
- Need to pool over cohorts
- Genomeutwin facilitated such analyses

Twin similarity for life span at very old age

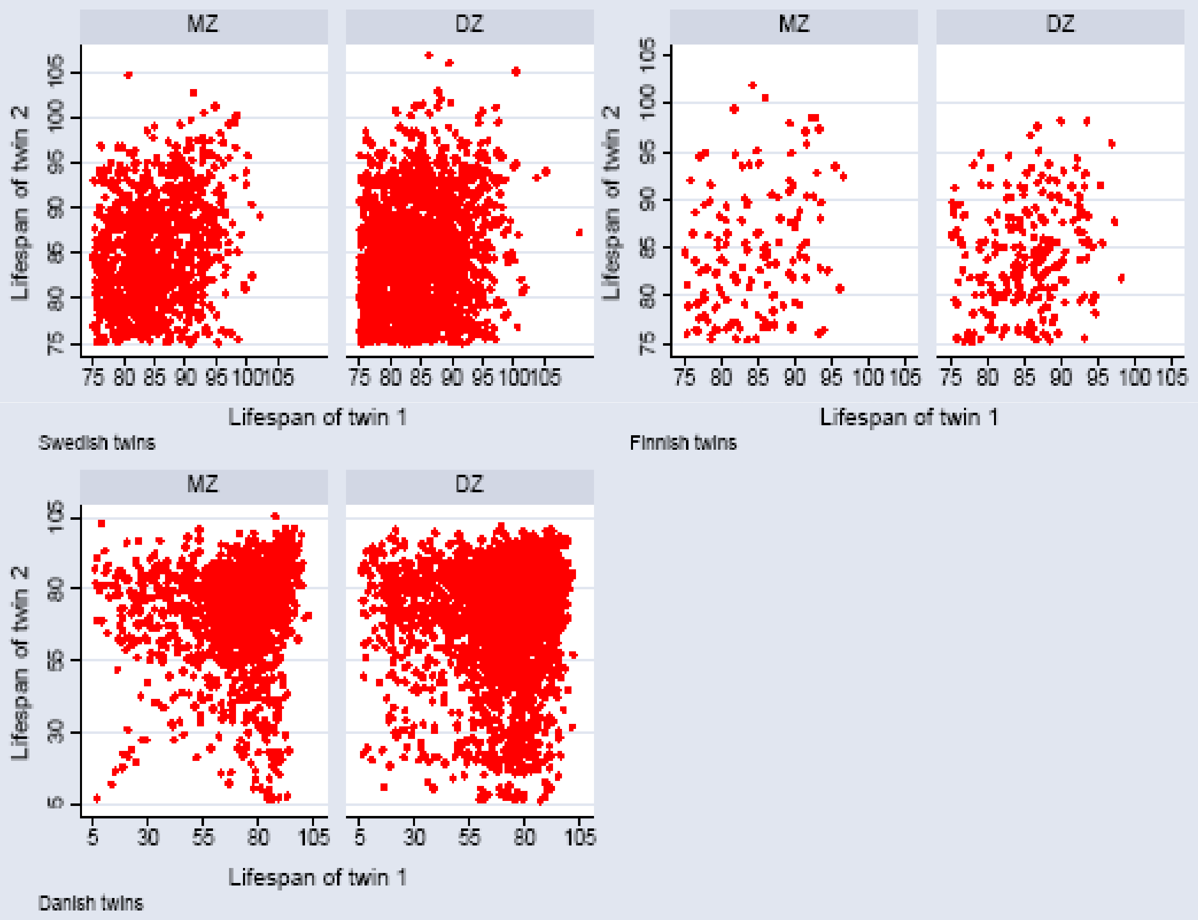


Figure 2: Twin-twin plot of the Swedish, Finnish and Danish cohort.

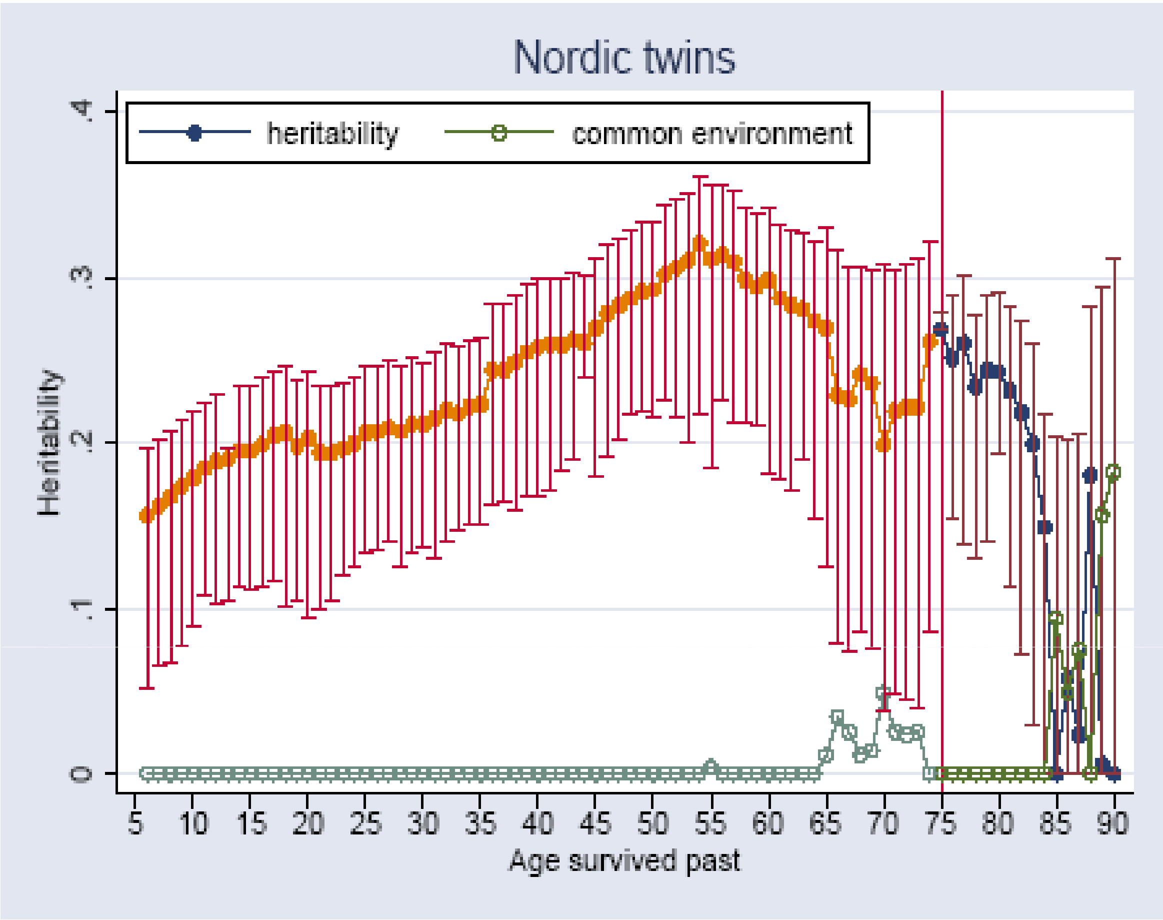
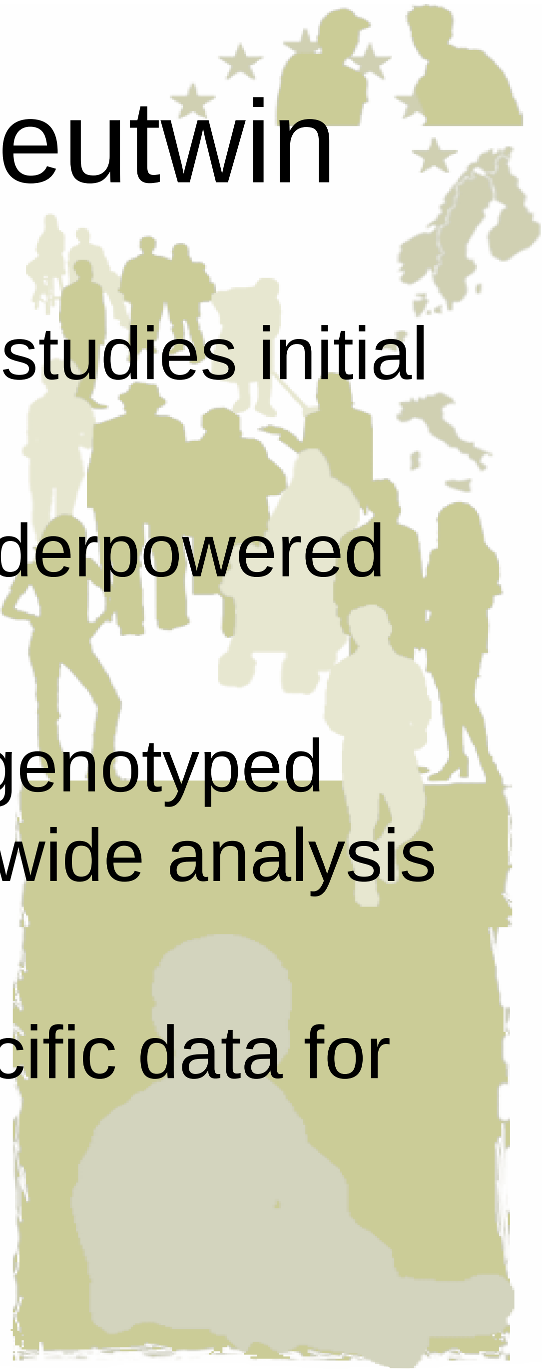


Figure 4: Heritability in lifespan of Danish, Finnish and Swedish twins by age survived past (obtained from ACE model). Estimates of amount of total variation due to common environmental effects are given as well. Swedish and Finnish twins enter when surviving past 75 years of age (full vertical line). 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

Genotyping in Genomeutwin

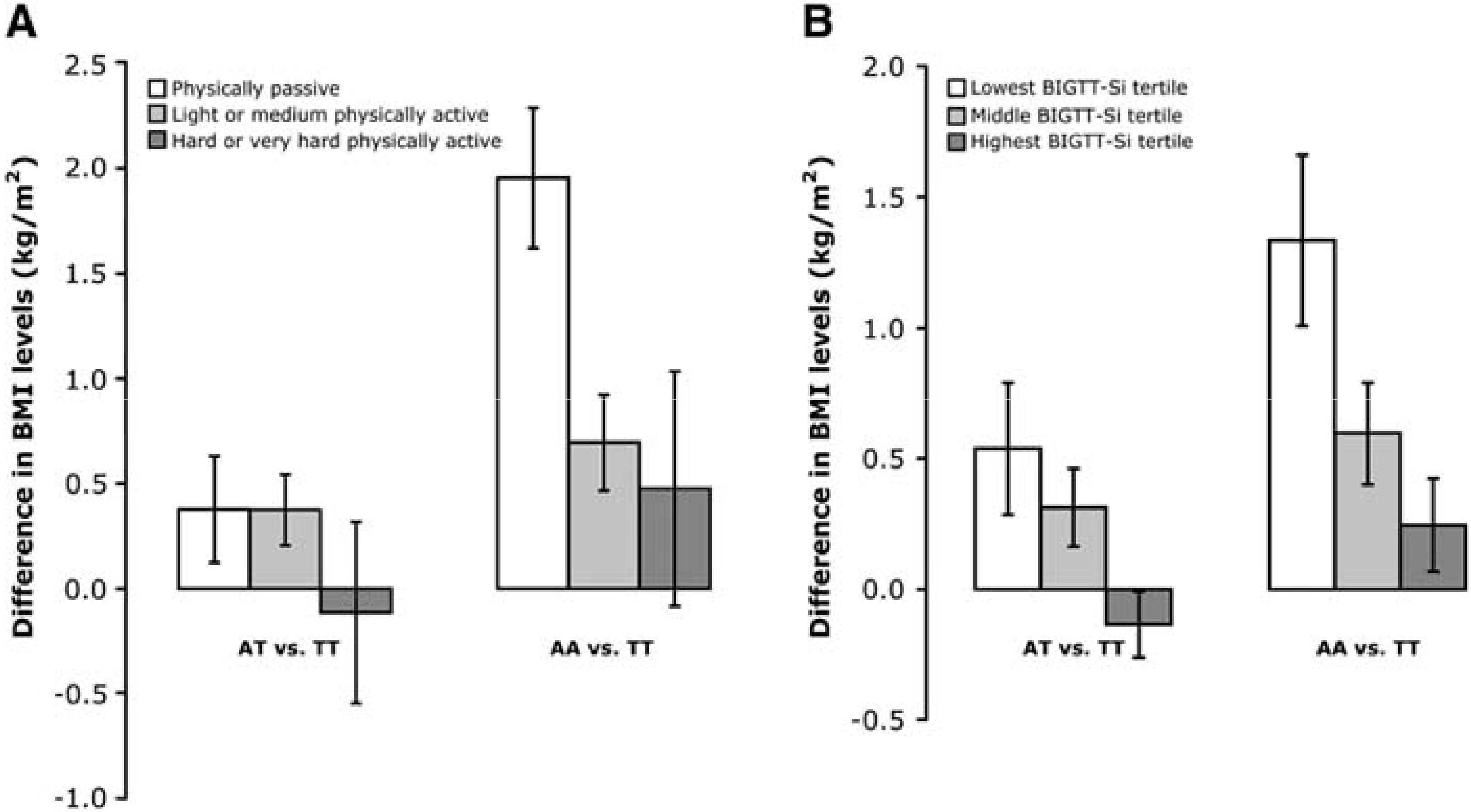


- Linkage scans and candidate gene studies initial focus
- Not too much came out of it; still underpowered
- GWAs came towards end of project
- C. 2000 monozygotic female pairs genotyped using Illumina 310 for first genome-wide analysis of variability genes
- Also provided much population specific data for population genetics

Use of twins to explore gene-environment interactions

- Modification of heritability estimates by measured variables
- High heritability of BMI, yet relatively few genes have been identified
- GxE ?
- Role of physical activity in modifying gene expression relevant in obesity

Specific genes, physical activity and BMI



Andreassen et al. Low Physical Activity Accentuates the Effect of the *FTO* rs9939609 Polymorphism on Body Fat Accumulation. *Diabetes* 2008

ENGAGE

European network for genetic and genomic epidemiology

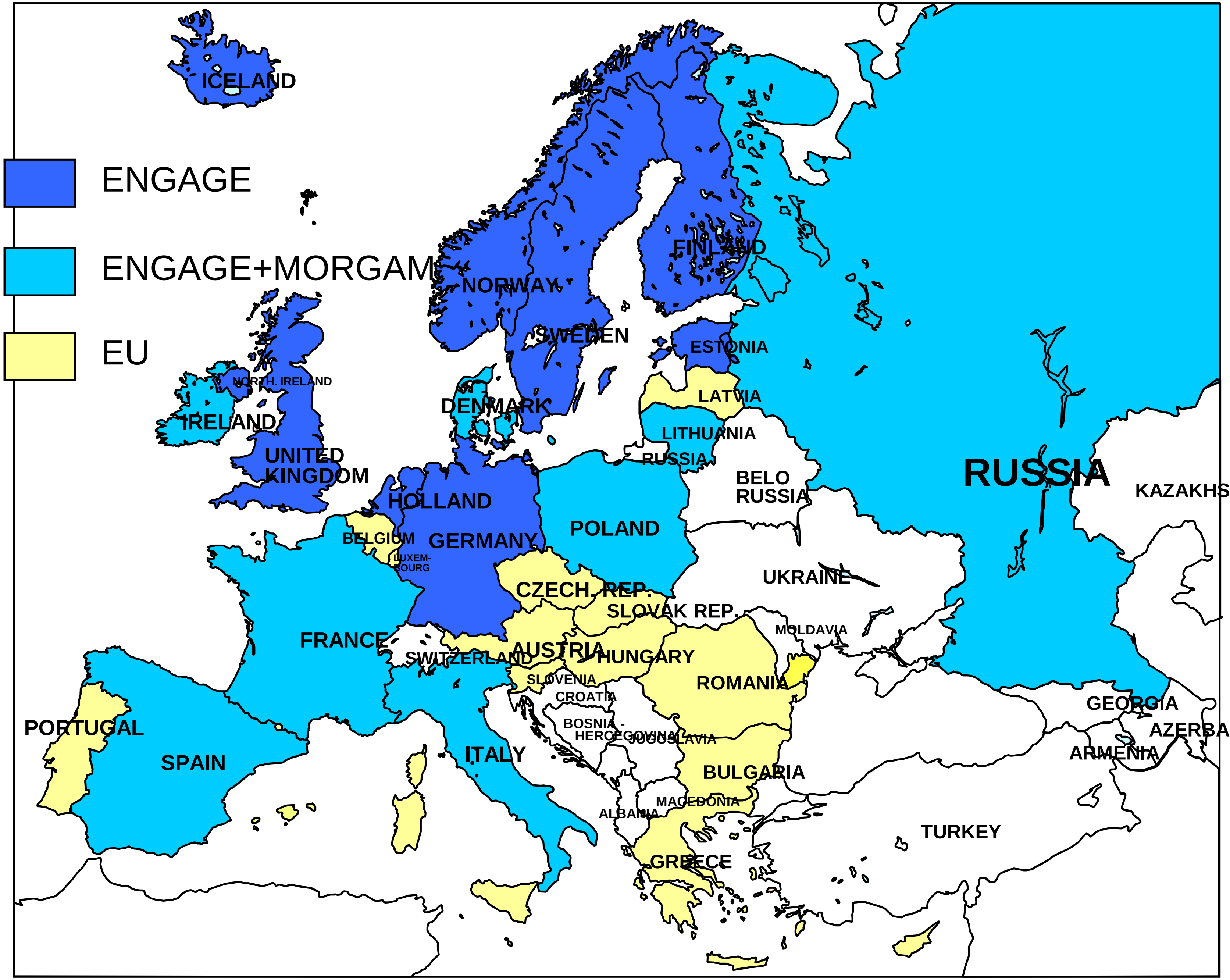
Website: www.euengage.org

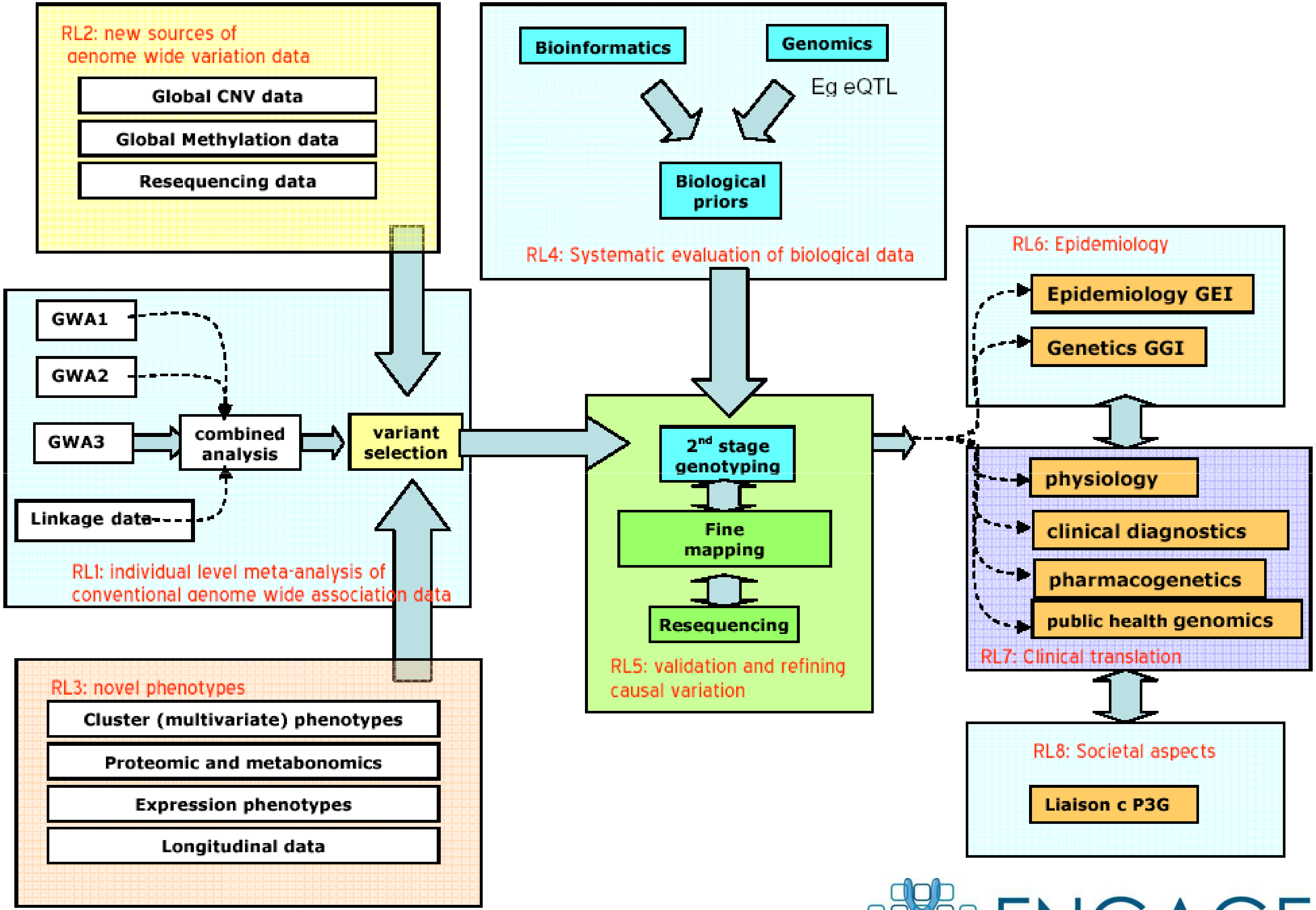
ENGAGE

Mission/Aim

To translate the wealth of data emerging from large-scale efforts in genetic and genomic epidemiology conducted in well-characterized European (and other) samples into information of relevance to future clinical advances.

ENGAGE


▇ ENGAGE
▇ ENGAGE+MORGAM
▇ EU
ICELAND
NORWAY
SWEDEN
FINLAND
ESTONIA
LATVIA
LITHUANIA
RUSSIA
NORTH. IRELAND
IRELAND
UNITED KINGDOM
DENMARK
HOLLAND
BELGIUM
LUXEM-BOURG
GERMANY
POLAND
BELO RUSSIA
RUSSIA
KAZAKHSTAN
UKRAINE
CZECH. REP.
SLOVAK REP.
MOLDAVIA
FRANCE
SWITZERLAND
AUSTRIA
HUNGARY
ROMANIA
SLOVENIA
CROATIA
BOSNIA - HERCEGOVINA
JUGOSLAVIA
GEORGIA
AZERBAIJAN
ARMENIA
PORTUGAL
SPAIN
ITALY
BULGARIA
MACEDONIA
ALBANIA
TURKEY
GREECE



RL2: new sources of genome wide variation data

Global CNV data

Global Methylation data

Resequencing data

Bioinformatics

Genomics

Eg eQTL

Biological priors

RL4: Systematic evaluation of biological data

GWA1

GWA2

GWA3

combined analysis

variant selection

Linkage data

RL1: individual level meta-analysis of conventional genome wide association data

2nd stage genotyping

Fine mapping

Resequencing

RL5: validation and refining causal variation

RL6: Epidemiology

Epidemiology GEI

Genetics GGI

physiology

clinical diagnostics

pharmacogenetics

public health genomics

RL7: Clinical translation

RL8: Societal aspects

Liaison c P3G

RL3: novel phenotypes

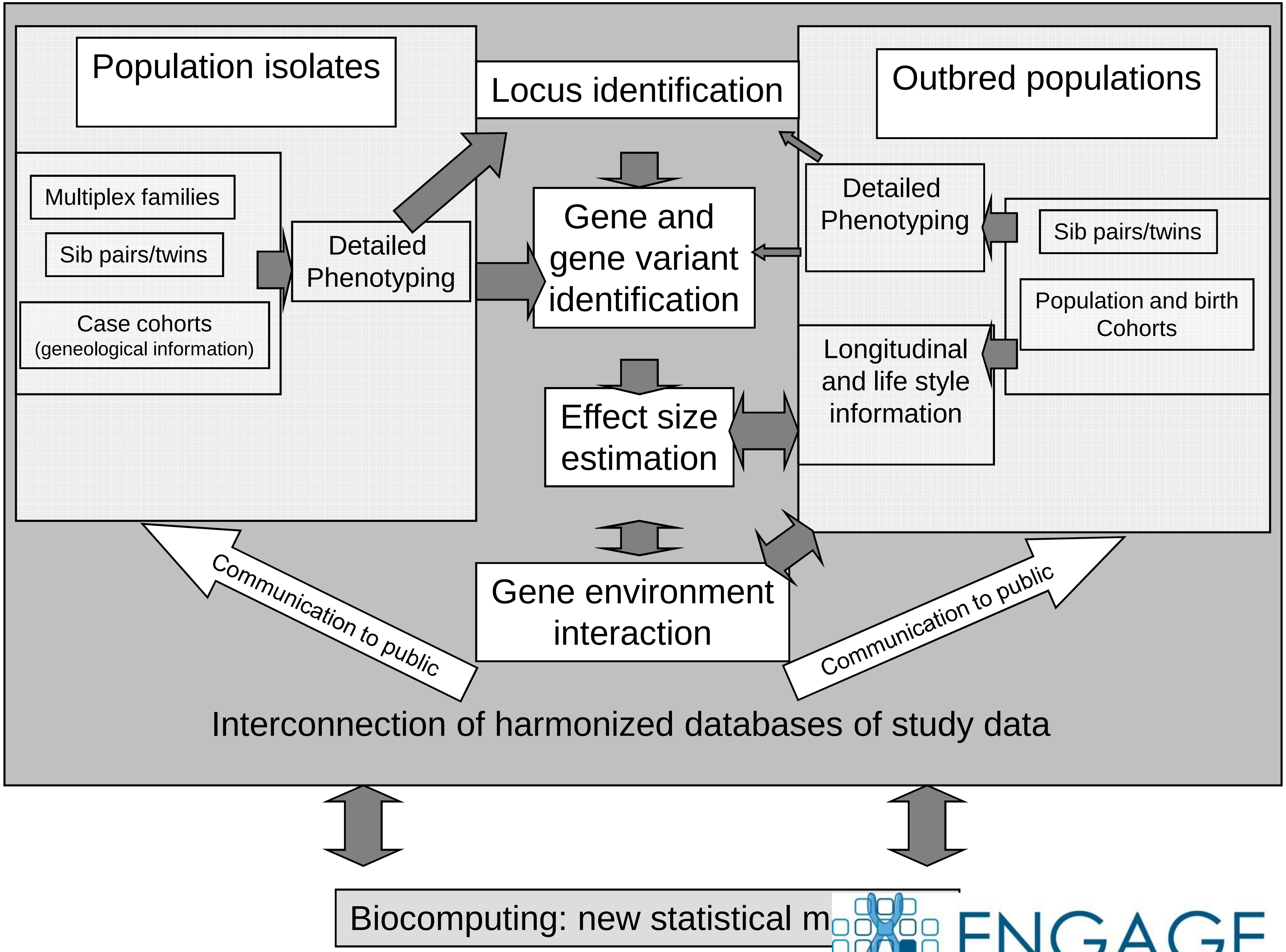
Cluster (multivariate) phenotypes

Proteomic and metabonomics

Expression phenotypes

Longitudinal data

ENGAGE



Population isolates

Multiplex families

Sib pairs/twins

Case cohorts
(geneological information)

Detailed Phenotyping

Locus identification

Gene and gene variant identification

Effect size estimation

Gene environment interaction

Outbred populations

Detailed Phenotyping

Sib pairs/twins

Population and birth Cohorts

Longitudinal and life style information

Communication to public

Communication to public

Interconnection of harmonized databases of study data

Biocomputing: new statistical m

ENGAGE

Twins are always popular !