

Quality of Life and Management of Living Resources

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

GENOMEUTWIN

Action line: Generic Activities Area 8.5

Proposal number: QLRT-2001-01254

Table of contents

1. OBJECTIVES AND EXPECTED ACHIEVEMENTS	3
2. PROJECT WORKPLAN	4
2.1 Introduction	4
2.2 Description of work packages	5
2.3 Project structure, planning and timetable.	
List of participants	12
Table 1. Work package list	13
Table 2. List of milestones	14
Table 3. List of Deliverables	15
Figure 1. Management in relation to workpackages and milestones	16
3. ROLE OF PARTICIPANTS	17
4. PROJECT MANAGEMENT AND COORDINATION, SCIENTIFIC ADVISORY BOARD	22
5. EXPLOITATION AND DISSEMINATION ACTIVITIES	24
6. ETHICAL ASPECTS AND SAFETY PROVISIONS	24
7. ONGOING EC FUNDED PROJECTS	26

1. OBJECTIVES AND EXPECTED ACHIEVEMENTS

The European population cohorts are of special significance in the current era of genomic research aimed at understanding normal and disease traits in man. New avenues of investigation have opened thanks to several recent advances including: the completion of the human genome sequence, increased understanding of sequence variants, low cost genome-wide tools to monitor these variants, and powerful multivariate statistical tools. The scientific yield from these tools is, however, limited when used to analyse human study samples with inherent problems of poor epidemiological study design, ascertainment bias of the study samples, lack of longitudinal follow up of the phenotypes studied and a tremendous amount of noise created by environmental variation throughout development, but particularly during foetal life and childhood.

Scientific objectives and expected achievements of this proposal are:

- 1) To develop novel strategies to utilize maximally the unique features of twin cohorts, including the availability of longitudinal data and ample information about lifestyle and environmental factors, in the characterization of complex traits.
- 2) To utilize the synergy between the twin cohorts and the representative population cohorts from the same countries, in studies of genetic and environmental predictors of traits.
- 3) To develop (in collaboration with the biotechnological industry) new molecular methods for high-throughput genotyping by analysing sub-samples of twin cohorts selected for specific traits (stature, BMI, migraine, coronary heart disease and stroke), requiring genotyping of hundreds of multiallelic and SNP markers in thousands of subjects.
- 4) To develop novel mathematical strategies to combine information on genetic profiles underlying common traits (multiple genes and their various alleles with different impacts on disease outcome) and to estimate the role of genetic and environmental factors in the disease process in selected traits using dizygotic (DZ) and monozygotic (MZ) twins and population cohorts.
- 5) To develop statistical methods for longitudinal data analyses of these life-span data sets to identify genes and/or environmental risk factors expressing themselves only at certain ages.
- 6) To identify major genetic factors predisposing to migraine, coronary heart disease and stroke in European populations.

Cooperation objectives and expected achievements of this proposal are to:

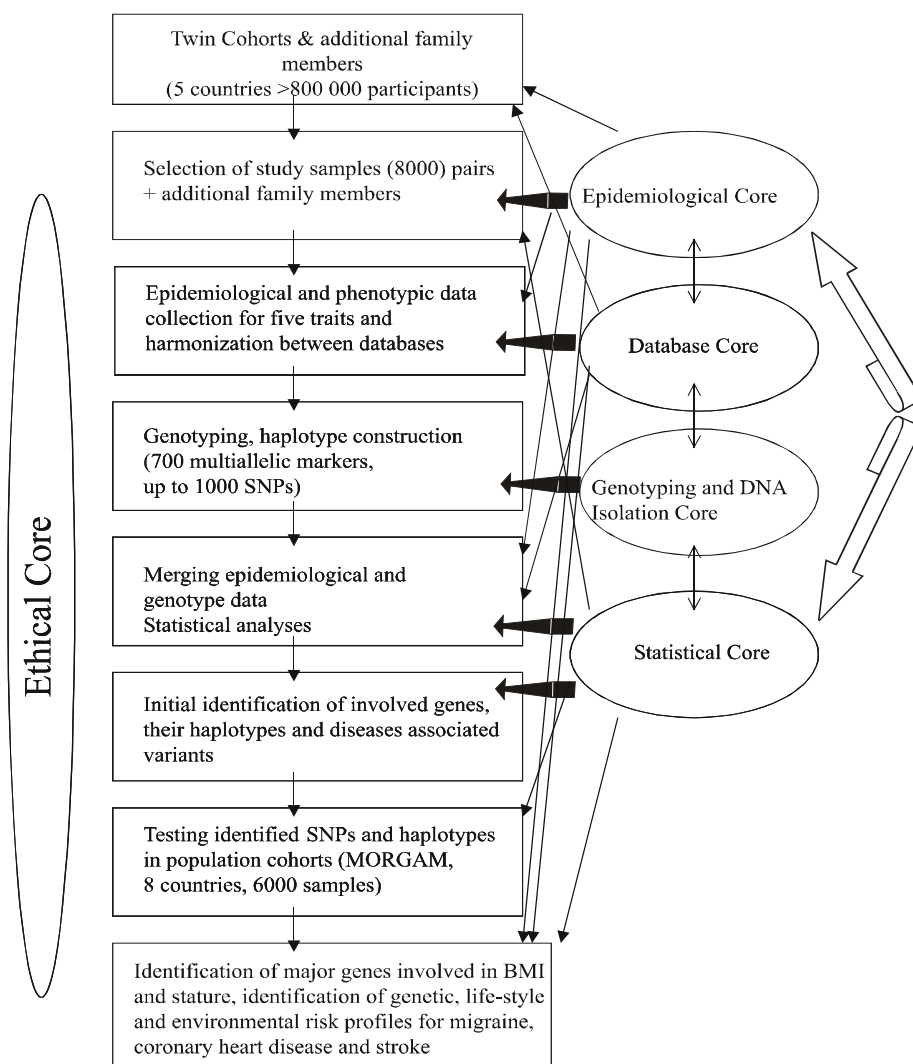
- 1) Create a solid infrastructure and research environment for European-wide studies using representative population cohorts and population-based twin registries that are particularly beneficial for genetic research when compared to other types of cohorts. Phenotyped twins offer an ideal study sample for selecting concordant and discordant pairs for linkage studies. Combined with the MORGAM population cohorts, a unique study sample is produced for genetic investigations of human traits and diseases. This setting provides an ideal training environment in genetic epidemiology.
- 2) Harmonise and standardize the epidemiological, phenotypic and genotypic databases in participating centres to facilitate pooled analyses of genetic, phenotypic and epidemiological data.
- 3) Establish quality control and data security systems for these databases in order to provide the highest possible level of data confidentiality.
- 4) Establish a centralized DNA extraction and genotyping facility for collected DNA samples and create a common policy of informed consent and other ethical and legal issues involved.
- 5) Build the intellectual and computational infrastructure and expertise to produce new information of “genetic profiles”, predisposing to common diseases. The involved projects create training opportunities for tens of students and fellows in genetic epidemiology in Europe.

Time Table: Due to the complexity of the preparatory phase, involving extensive harmonization of phenotypic, epidemiological and genotyping databases across the population cohorts as well as intellectual core facilities, we consider highly justified to extend the program to four years.

2. PROJECT WORKPLAN

2.1 Introduction

Large, unbiased study samples are needed to verify the role of any genetic variant behind common traits. The participating six twin cohorts form an amazing collection of over 0.6 million pairs of twins. Over 30 000 DNA samples with informed consents for genetic studies of common diseases have already been stored from these population-based twin cohorts. Just as the twin studies, the WHO MONICA (www.ktl.fi/monica) collaboration has a long and successful history of conducting multicentre research, which makes it ideal for embarking on genetic studies of complex, multifactorial diseases. Studies targeted to cardiovascular traits are now being undertaken in MORGAM, a prospective case-cohort study. MORGAM cohorts include approximately 6000 individuals, drawn from population-based cohorts consisting of more than 80 000 participants who have donated DNA samples. These unique study samples will be analysed in four intellectual core facilities of this integrated project using accumulated expertise by partners in genetics, epidemiology and biostatistics and the summarized flow sheet of the project workplan is provided below.



2.2 Description of work packages

Work package number and title: 1; Genetic tools: Genotyping and SNP-selection

Start date: 0 months

Completion date: 48 months

Partner responsible: P 1 (419m), P 3B (13+248m), P 10 (108+54m)

Other partners involved: P 9, P 2 – P 8

Total person months of WP 1: 842

Objectives

Develop strategies to select optimal markers, including SNP variants to be monitored in the cohort samples. Select the markers from multiple databases and produce some new ones with targeted sequencing of biologically relevant candidate genes. Produce genome-wide genotypes and SNP genotypes initially for 7000 DZ and 1000 MZ twin pairs and later for the MORGAM population cohort. Establish high quality database and quality control system for the data produced. Develop molecular methods for genotyping and allelic haplotype building.

Description of work

For the pilot study to test the high-throughput genotyping, database systems and quality control methods, we will genotype 2000 twin pairs with 400 multiallelic markers to produce a genome-wide skeletal map for twin-cohort based analyses of stature and BMI. This will help us to establish best genotyping strategies, test database integration with the epidemiological and phenotypic data and develop the quality control methodology across the databases. The sib pair structure of twin cohorts helps us to minimize the genotypic errors. Importantly it provides us with detailed information of the allele and haplotype frequencies in different European populations and helps us to design the best set of multiallelic markers for the next stage. In the next stage of the study aiming at identification of major genes for migraine, CHD and stroke, we will genotype 6000 twin pairs with 400 multiallelic markers. Parallel with this effort we start the selection of SNPs based on the existing information of the informativeness and genome-wide SNP haplotypes and predicted biological significance. Throughout the project we will develop SNP detection and allele calling technologies and develop databases and statistical strategies to combine the SNP genotypes with multiallelic genotypes, which will facilitate monitoring for LD and allelic haplotypes in our cohorts.

Deliverables

Solid deliverables for the project include 3.2 million error-checked multiallelic genotypes and some 3 million SNP genotypes. This will result in the increased information of allelic and haplotype frequencies in different European populations. We will develop novel biocomputing methods for SNP selection as well as technologies for SNP-detection. The work package will produce a unique genotype database that can form the basis for genetic epidemiological studies in Europe for years.

Expected results and relevant corresponding milestones

Novel biocomputing methods for SNP identification and haplotype construction
Established allelic frequencies of and haplotype frequencies of analyzed genes in different European populations. Some understanding of genome-wide LD intervals in European populations.
Genotype database for large population cohorts, which, when properly data-protected can be made accessible to investigators throughout Europe

Work package number and name: 2; Combining epidemiological and genotype information to explore the etiology of five selected traits

Start date or starting event: 0 months

Completion date: 48 months

Partner responsible: P 2 (100+78m), P 3 (156+212m), P 4 (150+120m), P 5 (130+7.8m), P 6 (150m), P 7 (70+108m), P 8 (78+989.6m)

Other partners involved: P 1, P 9

Total person months of WP 2: 2349.4

Objectives

The postgenomic era of biomedical research is armed with molecular tools to define the genetic components of common diseases. We aim to use these tools to analyze unique population cohorts in Europe to define and characterize the genetic, environmental and life-style components in the etiology of stature, BMI, migraine, CHD and stroke. We aim to assess and evaluate the general epidemiological value of twin cohorts and perform careful quantitative genetic analysis of the traits. Using environmental and lifetime event data and genetic information we aim to use novel statistical analyses to produce genetic and life style risk profiles for these traits. The population cohorts with large amounts of the lifetime data will be used to evaluate context-dependency and gene-environment interactions in these complex traits.

Description of work

We will use detailed analysis of epidemiological data in twin registers to address the inheritance of selected traits and the general epidemiological value of twin cohorts. For the latter we have an excellent opportunity, especially for cardiovascular traits due to availability of the MORGAM population cohort. The DNA samples already available facilitate immediate production of genotypic data for 2000 twin pairs to test and further develop our experimental strategies. In this study sample we will address stature and BMI. In the study sample of 8000 twin pairs selected for three disease traits we will address the selected disease traits of migraine, CHD and stroke. In years 3 and 4, when genome-wide genotyping data is available, we will analyze the phenotypic and epidemiological data combined with genotypic information to identify gene-environment and gene-gene-interactions and the role of level genes and variability genes.

Deliverables

We will publish comparative analyses of the epidemiological relevance of twin register and population cohort data especially in traits associated with CVD. We will also produce detailed information of the epidemiology and phenotypic features of five selected traits in different populations of Europe. Finally, when combined with the genotypic information produced we will create genetic profiles for these common traits and identify at least the major genes and their variants behind these traits. These profiles provide the basis for the evaluation of gene-environment interaction behind these traits. Such profiles are of outmost importance for health care related industries as well as national health care and prevention programmes in all European countries.

Expected results and relevant corresponding milestones

Basic information of genetic profiles important for stature, BMI, migraine, CHD and stroke.
Information of allelic diversity of disease genes in different European populations.

Some understanding of gene-gene and gene-environment interactions in these traits.

Work package number and name: 3; Biostatistical analyses

Start date or starting event: 0 months

Completion date: 48 months

Partner responsible: P 9 (288+16m)

Other partners involved: P 1, P 2, P 3, P 4, P 5, P 6, P 7, P8

Total person months of WP 3: 304

Objectives

There is a lack of statistical tools to face the analytical and quantitative challenges of modern, large genetic epidemiological studies. We will develop new analytical methods and statistical strategies to maximize the use of quantitative and longitudinal phenotypic and environmental data of our twin and population cohort databases. We will combine the power of linkage and association analyses in the studies of common traits. We will develop QTL and multivariant analyses to address the gene-gene and gene-environment relationships and produce new statistical programmes to incorporate multiple variants in the analyses of the genetic background of common diseases. This will benefit the whole scientific community and proliferate the outcomes of this Proposal.

Description of work

We will encounter problems like multiple testing, genetic heterogeneity, pleiotropism, epistasis, genotype by environment and life event interaction as well as correlated environmental parameters as integral problems of genetic studies of the disease traits we are analyzing. Further, immense size of the data in our cohort databases and merging from our genotype efforts means we have to further develop currently existing statistical strategies as well as develop new ones. We will develop improved pedigree analysis software with novel and/or advanced algorithms and parallel processing. We will develop more powerful NPL statistics that can be combined across case-control and core family data. We will develop QTL methods and software able to handle pleiotropy and longitudinal data as well as multivariant traits. These methods will be further developed to accommodate gene-gene and gene-environment data. Our immense amount of genotypic data will necessitate development of new error-correction and haplotype building strategies over wide genetic intervals as well as novel approaches to statistical interpretation of multiple tests.

Deliverables

We will improve the performance of existing computer programs to reflect the needs of our study cohorts (first year). We will produce novel statistical programs based on new algorithms and computational strategies developed by us (second year). Using these statistical tools we will analyze huge amounts of genetic, phenotypic and environmental data in our databases (third and fourth year). These new programs developed to a user-friendly format will be made accessible in the web for the scientific community (fourth year). Some computer programs evaluating the genetic and environmental risk profiles will be developed to fulfil the needs of pharmaceutical and other health-care related industries.

Expected results and relevant corresponding milestones

Novel statistical methods meeting the demands of large-scale studies of modern genetic epidemiology. New user-friendly computer programs to handle both qualitative and quantitative genetic and environment and lifestyle data

Work package number and name: 4; Linking European clinicians and clinical and epidemiological data collection
Start date or starting event: 0 months
Completion date: 36 months
Partner responsible: P 1 (160)
Other partners involved: P 2 (43.1m), P 3A (30m), P 4 (34.6m), P 5 (32m), P 6 (30m), P 7 (32m)
Total person months of WP 4: 361.7

Objectives

Although multiple research collaborations exist between European biomedical research centres and between European groups and leading non-European groups in genetics of common diseases and QTL traits, rather rarely have they resulted in a conceptual or technical infrastructure facilitating research “larger than the sum of its parts” or common database structure accessible to all participants and shared with the research community at large. We aim to maximally utilize and intensify the numerous collaborative links of participating centres with clinical and epidemiological research groups around the world by linking them to the facilitating infrastructure created by this integrated project. We aim to network individual research groups with only partially overlapping interest with the project by creating a system that facilitates distribution of statistical analyses of their data and even pooled data analyses over the Internet.

Description of work

Because of a long-established cooperation between twin cohorts and MORGAM population cohorts we can start our network operations immediately. Epidemiological and phenotypic data collection will be harmonized, evidence-based methods implied in the standardized quality control of the data and internationally recognized measures launched among participating clinical centres for all specific trait determinants essential for the research component of the integrated project. Special emphasis is placed on quality control of the data to be entered in the databases and the level and reliability of the details collected of clinical phenotypes and lifetime events.

Deliverables

We will start the harmonization of cohort data collection for essential environmental and life-style events affecting BMI and stature between the twin cohorts and MORGAM population cohorts (first year). Building on this experience we will continue the harmonization and quality control testing for multiple parameters relevant to studied disease traits (migraine, CHD, stroke) (second and third year). During the fourth year we aim to launch a supportive computer-assisted system accessible to all research groups for reliable data collection for genetic epidemiological studies including quality control standards. The infrastructure can be utilized in the future to analyze the genetics of various traits in addition to those listed here. Such traits include cardiovascular, autoimmune and neuropsychiatric traits.

Expected results and relevant corresponding milestones

Integrated network of twin cohorts and population cohorts of MORGAM
Harmonized and to some extent standardized data collection and quality control criteria for analyzed traits between the cohorts. Web-assisted help and guidance for epidemiological and phenotypic data collection including necessary quality control support.

Work package number and name: 5; Database networking and harmonization**Start date or starting event: 0 months****Completion date: 42 months****Partner responsible: P 1 (100m), P 7 (30m)****Other partners involved: P 2, P 3, P 4, P 5, P 6, P 8****Total person months of WP 5: 130****Objectives**

The ambitious goals of our research component can only be met if we a) create an infrastructure for database networking between the participating centres, b) build an integrative database system facilitating effortless combination of genetic and epidemiological and phenotypic data, c) carry out pooled data analyses beyond the possibilities and facilities of individual centres. We aim to develop a computer infrastructure facilitating data mining of different existing databases. An important component of the system will be effortless merge of genetic and environmental and life-event data. We will to maximal extent harmonize the databases in participating centers and create a good structure for database management.

Description of work

All the participating population cohorts and data analysis centers have their own databases and expertise in data quality control and statistical analyses of the data. We will have frequent meetings between the local database managers and major database experts of the database core to evaluate the current database structures and identify the best, feasible and user-friendly system to harmonize database structures. This will require intensive interactions between data producer sites, statistical experts and analytical personnel. This is established by the operation of our Database Core that has representatives from participating population cohorts, genetic analysis centers and the Statistical Core, responsible for the data analyses of the integrated project. The process of database networking and harmonization also involves multiple social, legal and ethical issues requiring expertise in various firewall systems to guarantee data confidentiality but without compromising or complicating the research efforts.

Deliverables

We will produce a plan for a harmonized database structure between twin cohorts and MORGAM population cohorts (first year). We will create a system facilitating genetic data mining and effortless interrogation of population data (second year). We will implement this plan and test its operational maturity and data confidentiality protection system by testing data on stature and BMI (second and third year). We will develop and expand the system to facilitate the complex statistical analyses of three disease traits in the complete data set (fourth year).

Expected results and relevant corresponding milestones

Harmonized database structure in twin cohorts and MORGAM population cohorts
Database structure facilitating integration of genetic, epidemiological and phenotypic data and advanced statistical analyses of these data. Guidelines for data confidentiality and firewall structure of large epidemiological and genetic databases.

Work package number and name: 6; Training and mobility

Start date or starting event: 0 months

Partner responsible: P 3 (10m), P 6 (67.7m)

Completion date: 48 months

Other partners involved:

Total person months of WP 6: 77.7

Objectives

It is imperative that European research groups obtain “state of art” training in modern genetic epidemiology. The future of Europe is dependent on the experts in this rapidly developing field, which has implications for so many areas of life. We aim to build an educational system and productive training programme for both for pre-and postdoctoral researchers and for senior clinicians and scientists to increase their expertise in this field. This training and mobility component operates in close interaction with the research and networking components.

Description of work

Our intellectual core system creates an ideal infrastructure for training and exchange of junior and senior scientists. They can spend various amounts of time during their career to obtain training in epidemiology, high throughput genotyping, database management and operations, statistical analyses or ethical issues of genetic epidemiological research. A more formal training programme includes courses, inter-institute exchanges, workshops and hands-on training at participating centres. We will recruit a Director of the training programme in order to guarantee good coordination and integration of educational systems in the participating countries and centres and to guarantee optimal training for pre-and postdoctoral students within the programme. The workshops and international symposia arranged annually on the topics crucial for the integrated project will maximize the education of scientists of participating centres and other centers in the project’s integrated research. It is imperative that the training is closely connected with national educational systems since scientists of the participating centres are leaders of the graduate schools of epidemiology, statistics and genetics in their own countries.

Deliverables

We will establish our training and mobility plan during the first year. We will create a website for information of all the courses, meetings and workshops in the participating centres (first year). We will hold workshops, symposia, and training sessions forming the skeleton for our predoctoral and postdoctoral training programme (second-fourth year). We will design special targeted courses for integrated research in genetic epidemiology for post-doctoral scientists, clinicians and advanced scientists (second-fourth year). We will arrange a special workshop series in ethical, legal and social issues of genetic epidemiological research and biobanks (second-fourth year).

Expected results and relevant corresponding milestones

A website will be established that is accessible to everybody concerning the training possibilities in genetic epidemiology in participating centres and institutes. We will develop a comprehensive training programme, tightly integrated to national educational system, in modern genetic epidemiology.

Work package number and name: 7; Scientific Advisory Board

Start date or starting event: 0 months

Partner responsible: P 1 (5m)

Completion date: 48 months

Other partners involved:

Total person months of WP 7: 5

Objectives

To guarantee the excellence of the scientific work as well as the top quality of the networking and training components. To monitor the legal and ethical issues of the project. To evaluate the progress of the work and the deliverables of the individual workpackages and the complete project.

Description of work

A Scientific Advisory Board (SAB) will meet once a year with the steering committee and all the partners and provide scientific guidance for the work and a critique of the results produced. Throughout the project the SAB will provide expert advice to the Co-ordinator, the Steering Committee and the leadership of intellectual core facilities. SAB will provide special advice on the management of the whole project and in the quality control issues of data collection and analyses. SAB will evaluate annually the progress of the project as whole and also the integration of the work packages into the goals of the project.

Deliverables

Yearly evaluation reports of the scientific progress and the quality of the networking and education component. Expert advice in special issues of data collection and analyses, including the ethical and legal issues. Guidance in the organization of the networking and training components.

Expected results and relevant corresponding milestones

Every 10th month of each project year a meeting with the project Partners will be held and the SAB will produce a progress report of the project to the Commission. Selected SAB members, based on their field of expertise, will participate in the annual, targeted training and networking workshops.

2.3 Project structure, planning and timetable

List of participants

Participant 1: National Public Health Institute, Helsinki, Finland

Co-ordinator: Prof. Leena Peltonen, (PI)

Mailing address: National Public Health Institute, Department of Molecular Medicine, Biomedicum, P.O. Box 104, FIN-00251 Helsinki Finland

e-mail: leena.peltonen@ktl.fi *tel:* +358-9-4744 8393 *fax :* +358-9-4744 8480

Participant 2: University of Southern Denmark, Denmark

Prof. Kirsten Kyik, Danish Twin Registry –

Mailing address: Syddansk Universitet Campusvej 55 DK-5230 Odense M Denmark

e-mail: KChristensen@health.sdu.dk *tel :* + 45 65 50 30 49 *fax :* + 45 65 95 77 66

Participant 3 (A and B): University of Helsinki, Finland

Prof. Jaakko Kaprio, (Participant 3A; Finnish Twin Cohort Study)

Mailing address: Dept. of Public Health, P.O. Box 41, FIN-00014 University of Helsinki, Finland

e-mail: jaakko.kaprio@helsinki.fi *tel :* +358 9 191 27 595 *fax :* +358 9 191 27 600

Prof. Aarno Palotie, (Participant 3B; Finnish Genome Center)

Mailing address: Finnish Genome Center, University of Helsinki, Biomedicum, Haartmaninkatu 8,

e-mail: aarno.palotie@helsinki.fi *tel :* +358-9-1911 *fax :* +358 9 1912 5478

Participant 4: Istituto Superiore di Sanità, Italy, Dr. Antonia Stazi, Italian Twin Registry

Mailing address: Istituto Superiore di Sanità, Viale Regina Elena 299, 00161 Rome- Italy

e-mail: Stazi@iss.it *tel :* +39 06 49903181 *fax :* +39 06 49387173

Participant 5: Free University, Netherlands

Prof. Dorret Boomsma, Dutch Twin Registry

Mailing address: Vakgroep Psychonomie, Vrije Universiteit, De Boelelaan 1111, 1081 HV AMSTERDAM, Netherlands

e-mail: di.boomsma@psy.vu.nl *Tel:* 020-444.8789 *Fax:* 020-444 8832

Participant 6: The Norwegian Institute of Public Health, Norway

Dr. Jennifer Harris, Norwegian Twin Registry

Mailing address: Norwegian Institute of Public Health PO.Box 4404 Nydalen 0403 OSLO, Norway

e-mail: Jennifer.harris@folkhelsa.no *tel* +47 23 40 82 14: *fax :* +47 23 40 82 52

Participant 7: Karolinska Institutet, Sweden

Prof. Nancy Pedersen, Swedish Twin Registry

Mailing address: The Swedish Twin Registry, Department of Medical Epidemiology, Box 281, Karolinska Institutet, SE-171 77 Stockholm, Sweden.

e-mail: Nancy.Pedersen@mep.ki.se *tel:* +46-8-728 74 18, *fax:* +46-8-31 49 75

Participant 8: University of Belfast

Prof. Alun Evans

Mailing address: Dept. of Epidemiology and Public Health, Mulhouse Building, Grosvenor Road, Belfast BT12 6BJ, United Kingdom

e-mail: a.evans@qub.ac.uk *tel :* +44 28 90894614 *fax :* +44 28 9023 1907

Participant 9: University of Leiden, Netherlands

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

Mailing address: Dept. of Medical Statistics, P.O.Box 9604, 2300 RC Leiden, The Netherlands

e-mail: sandkuijl@lumc.nl *tel :* +071 5276803, *fax :* +071 5276799

Participant 10: Uppsala University, Sweden

Prof. Ann-Christine Syvänen

Mailing address: Department of Medical Sciences, Uppsala University, Ing 70, 3 tr, Foavd 2, UAS

75185 Uppsala, Sweden

e-mail: Ann-Christine.Syvanen@medsci.uu.se Phone: +46-18-6112959 Fax: +46-18-6112519

Prof.Ulf Pettersson

Mailing address: Dag Hammarskjölds väg 20, Rudbecklaboratoriet, 751 85 UPPSALA

e-mail: Ulf.Pettersson@genpat.uu.se tel: +46-18-471 45 04 fax: +46-18-471 48 08

Table 1. Work package list

Work package No	Work package title	Responsible participant No	Person months	Start month	End month	Deliverable Numbers
WP 1	Genetic tools: Genotyping and SNP-selection	P1, P3b & P10	842	0	48	1 & 2
WP 2	Combining epidemiological and genotypic information to explore the etiology of five selected traits	P2-P8	2349.4	0	48	3 & 4
WP 3	Biostatistical analyses	P9	304	0	48	5,6
WP 4	Linking European clinicians and clinical and epidemiological data collection	P1	361.7	0	36	7,8,9
WP 5	Database networking and harmonization	P1 & P7	130	0	42	10,11
WP 6	Training and mobility	P3 and P6	77.7	0	48	12
WP 7	Scientific Advisory Board	P1	5	0	48	1-12
	Total		2262.4			

Table 2A. List of milestones, total project

Milestone No	Title	Delivery date	Partici pants	Description
MS1	6.2 million genotypes	48	P1 & P10	More than 6 million genotypes from European populations is produced for statistical analysis
MS2	Genetic profiles	48	P2-P8	Basic analyses and information of genetic profiles important for stature, BMI, migraine, CHD and stroke finished.
MS 3	Computer programs	48	P9	Beta versions of novel computer programs delivered for scientific use.
MS 4	Database completion	42	P1& P7	A common, secure database for all relevant scientific information in this project is established
MS 5	Informatics	12	P3	A multi-access level, secure website is established for the project, which includes the draft manual for the project
MS 6	Training	48	P3	A comprehensive training program, in modern genetic epidemiology is integrated to national educational systems.

Table 2B First year results

Result	Title	Delivery date	Partici pants	Description
MS 1	1 million genotypes	12	P1 & P10	Allele reading and genotyping system established, quality controls at place
MS 2	Genetic profiles	-	P2 - P8	Cluster analyses performed on phenotypic variants to form the basis for profile analyses
MS 3	Computer programs	12	P9	A set of statistical tools designed for initial analyses, simulations performed
MS 4	Database	12	P1&P7	A database structure established
MS 5	Informatics	12	P3	A multi-access level, secure website is established for the project, which includes the draft manual for the project
MS 6	Training		P3&P6	A teaching program designed, first courses organized

Table 3A. List of Deliverables, all 4 years

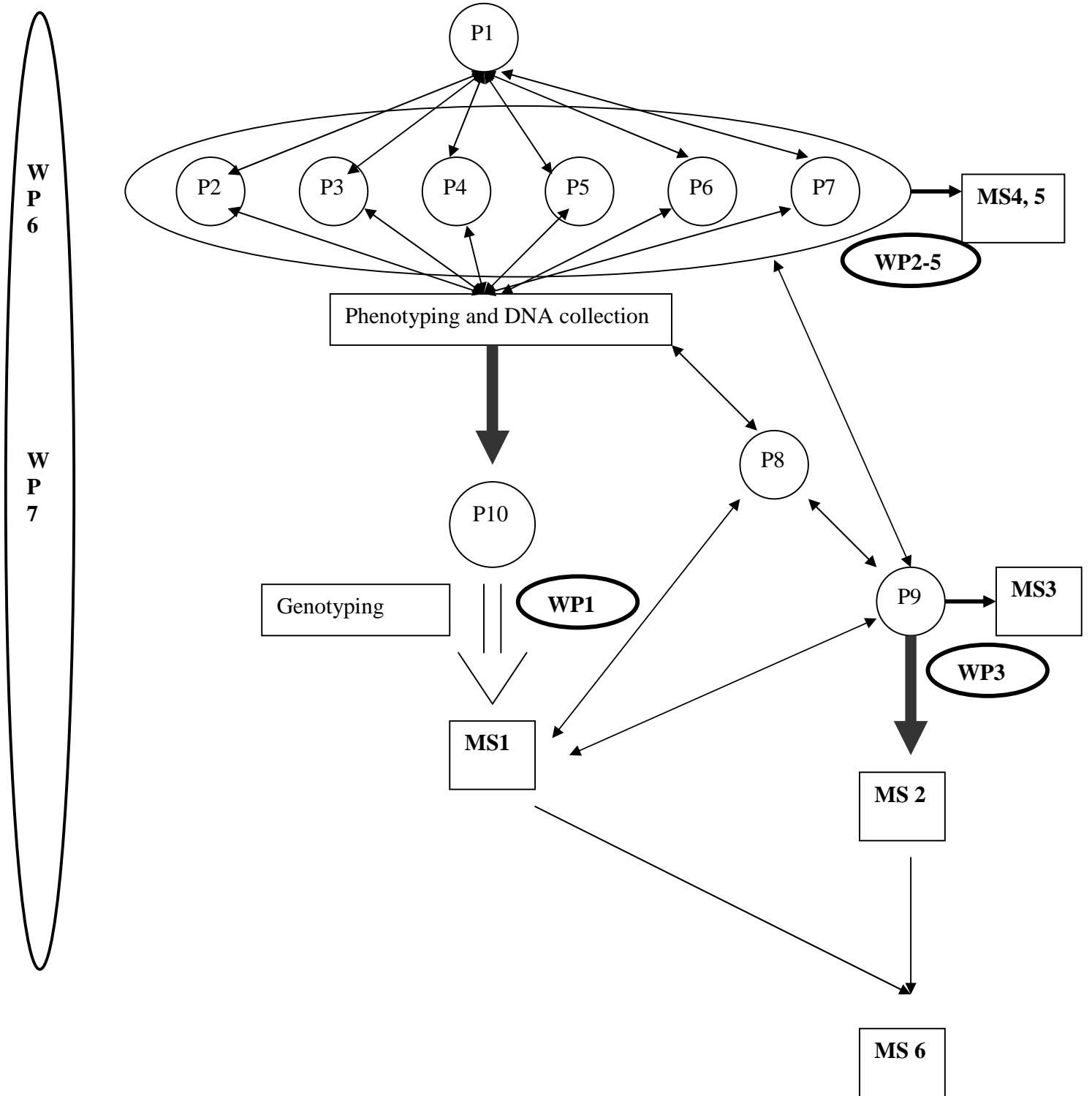
Deliverable No	Deliverable title	Delivery date	Partners involved	Nature	Dissemination level	Dissemination target
D 1	3.2 million error-checked multiallelic genotypes and some 3 million SNP genotypes	40	P1,3B,10	R	RE	Scientific community
D 2	genotype database for genetic epidemiological studies in Europe	48	P1,P7	R	RE	Scientific community, bioindustry
D 3	Epidemiology and phenotypic features of five selected traits in Europe	24	P2-8	R	PU	Scientific community, bioindustry, health management
D 4	Genetic profiles for five common traits	48	P1,7,9	R	RE & PU	Scientific community, bioindustry, health management
D 5	Novel statistical methods for complex traits	32	P9	R	PU	Scientific community
D 6	New Computer programs for applying D5 methods	48	P1,7,9	R,O	PU	Scientific community, bioindustry
D 7	Harmonization protocol for inter-European population samples with genetically informative data	24	P2-8	R	RE	Scientific community, bioindustry, health management
D 8	Networking documentation	48	P1-8	R	PU	Health management
D 9	Ethical & legal guidelines for produced data	34	P6	R	PU	Scientific community, health management
D 10	Plan for a harmonized database structure, collection of all QTL phenotype information accessible in cohorts	12	P1,7,	R	CO	Scientific community, bioindustry, health management
D 11	Fully operational harmonized database management system	30	P 1,7	R,O	RE	Scientific community, bioindustry, health management

D 12	Pre- and postdoctoral training programme	24	P3,6	O	PU	Scientific community
-------------	--	----	------	---	----	----------------------

Table 3A. List of First year deliverables

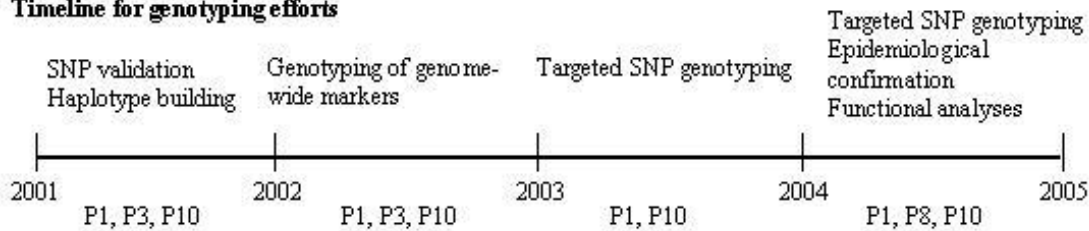
Deliverable No	Deliverable title	Partners involved	Percentage completed
D 1	3.2 million error-checked multiallelic genotypes and some 3 million SNP genotypes	See above	10%
D 2	genotype database for genetic epidemiological studies in Europe		30%
D 3	Epidemiology and phenotypic features of five selected traits in Europe		30%
D 4	Genetic profiles for five common traits		0%
D 5	Novel statistical methods for complex traits		10%
D 6	New Computer programs for applying D5 methods		10%
D 7	Harmonization protocol for inter-European population samples with genetically informative data		40%
D 8	Networking documentation		10%
D 9	Ethical & legal guidelines for produced data		25%
D 10	Plan for a harmonized database structure, collection of all QTL phenotype information accessible in cohorts		60%
D 11	Fully operational harmonized database management system		50%
D 12	Pre- and postdoctoral training programme		25%

Figure 1: Management structure in relation to workpackages and milestones

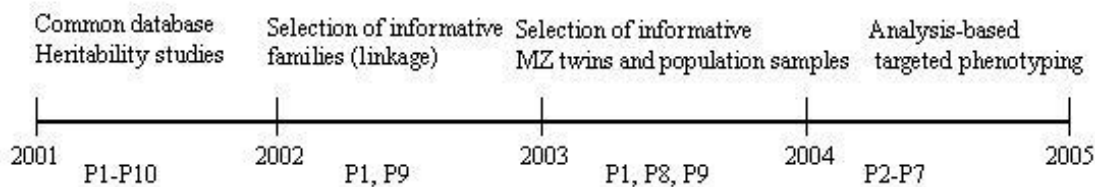


Time tables

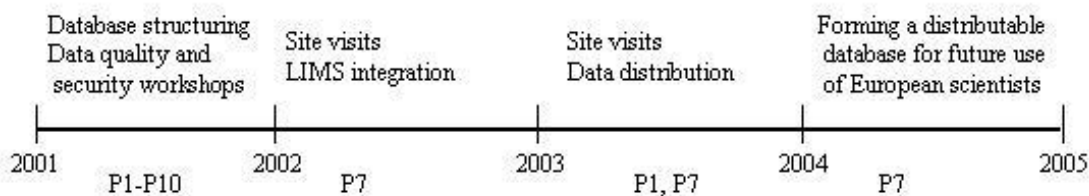
Timeline for genotyping efforts



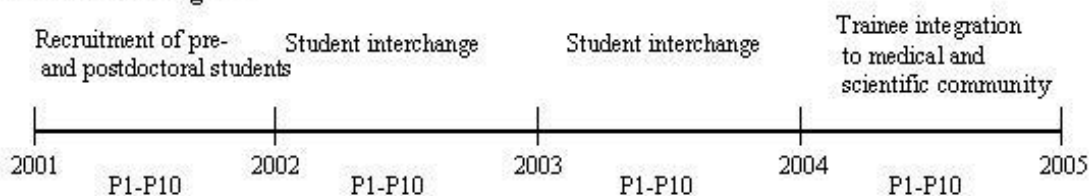
Timeline for phenotyping efforts



Timeline for database development



Network Training Plan



3. ROLE OF PARTICIPANTS

3. ROLE OF PARTICIPANTS

The whole project effort (non-EC funded months within parentheses)

	WP 1	WP 2	WP 3	WP 4	WP 5	WP 6	WP 7	TOTAL
P 1 (684m+48m for coordination)	419	-	-	160	100	-	5	732.0
P 2 (143.1m+78m)		100 (+78)		43.1				221.1
P 3A (212m+196m)		156 (+212)		30		10		408.0
P 3B (13m+248m)	13 (+248)							261.0
P 4 (184.6m+120m)		150 (+120)		34.6				304.6
P 5 (162m+7.8)		130 (+7.8)		32				169.8
P 6 (257.7m)		150		30		67.7		257.7
P 7 (132m+108m)		70 (+108)		32	30			240.0
P 8 (78m+989.6m)		78 (+989.6)						1067.6
P 9 (288m+16m)			288 (+16)					304.0
P 10 (108m+54m)	108 (+54)							162.0
TOTAL	842	2349.4	304	361.7	130	77.7	5	4127.8 (EC 2310.4+ No EC 1817.4)

First year effort

WP 1	1	2	3	4	5	6	7
P 1 (192m)	139			45	30		2
P 2 (35.9m)		45.6		9.9			
P 3 (52m) A		96		6		2	
B	72						
P 4 (40.6m)		65		5.6			
P 5 (48m)		37.8		13			
P 6 (72.2m)		50		11.2		11	
P 7 (24m)		37		12.0			
P 8 (12m)		262					
P 9 (72m)			76				
P 10 (42m)	42						
	253	593.4	76	102.7	30	13	2

Participant 1: National Public Health Institute (NPHI), Helsinki, Finland

Co-ordinator: Prof. Leena Peltonen, (PI)

Mailing address: National Public Health Institute, Department of Molecular Medicine, Biomedicum, P.O.Box 104, FIN-00251 Helsinki, Finland

e-mail: leena.peltonen@ktl.fi *tel:*+358-9-47448393 *fax* : +358-9-47448480

Scientific Team

Kari Kuulasmaa, PhD, Veikko Salomaa, MD, PhD, Sangita Kulthinal, PhD, Matti Niemelä, MD, Zygimantas Cepaitis, Engineer, Markus Perola, MD, PhD, Janna Saarela, MD, PhD, Juha Muilu, PhD, Juha Saharinen, PhD, Mervi Alanne, Bioengineer, Pekka Ellonen, Bioengineer, Juri Ahokas, System Administrator

Contractual links to other participants: none

Objectives: NPHI will be responsible for coordinating the project and harmonizing the overall goals of the project. NPHI will also be responsible for extraction, logistics, storage and aliquoting of the DNA collected as well as 50% of the high-throughput SNP genotypings. The genotyping database for the whole project is maintained in NPHI under supervision of a database expert. NPHI will also maintain the MORGAM database, securing its quality and confidentiality, assess the quality of the phenotypic and genotypic data for the MORGAM cohorts and co-ordinate the MORGAM data analyses. This expertise will also be used in the general database management of the whole project. The co-ordinator will also organize the connections with the scientific advisory board and orchestrate the operations of the ethical and legal core and oversee with the steering group the operations of the other intellectual core facilities.

Workplan: The overall administration and management in relation to workpackages and milestones is provided in Figure 1. The co-ordinator will meet monthly with the steering group and core directors. She will seek advice from the scientific advisory board in special issues concerning research strategies and consult them also in networking and training issues. Together with the steering group she will make decisions of the expansion of the project (inclusion of additional twin centers). NPHI will house major databases for phenotyping (MORGAM), genotyping and epidemiological (MORGAM) data. The development and quality control of these databases is the responsibility of P 1 in the close collaboration with P 7 that has the responsibility of the harmonization of all epidemiological databases. NPHI will establish a high-throughput SNP-genotyping platform in the Biochip unit of the Biomedicum. This unit has extensive experience in SNP-genotyping, SNP haplotyping and data quality control programs. The genotyping will be performed in the harmonized manner with P 10 (for methodological details see P 10). The genotype database will collect the genotypes of STR markers from P3B and P10 and use Mendel-based, haplotype guided error correction programs for quality control of the stored data. NPHI has over 10 years of experience in the database maintenance of the MORGAM project (directed by Dr. Kuulasmaa) and they will provide special expertise on the database quality control and data security issues to P7. Genotype database of NPHI will be managed by Dr. Muilu with 7 years of experience of database construction and quality control.

Participant 2: The Danish Twin Registry, University of Southern Denmark, Denmark

PI: Prof. Kirsten Ohm Kyvik

Mailing address: Syddansk Universitet Campusvej 55 DK-5230 Odense M Denmark

e-mail: KChristensen@health.sdu.dk *tel* : + 45 65 50 30 49 *fax* : + 45 65 95 77 66

Scientific Team: Kirsten Ohm Kyvik, Associate Professor MD, PhD, Managing Director, Kaare Christensen, Professor MD, PhD and Dr.Med.Sci., Niels V. Holm, MD, PhD, Axel Skytthe, Research Assistant Professor M.Sc., PhD

Contractual links to other participants: none

Objectives: The role and contribution of The Danish Twin Registry will be to implement the objectives of the proposal locally with respect to creation of infrastructure, harmonized and standardized epidemiological phenotypic and geno-type databases as well as quality control and data security systems. The participant will contribute environmental and life time event data as well as genetic information that will be used to produce genetic and lifestyle risk profiles and to evaluate context-dependency and gene-environment interactions in the study traits. Further, to collaborate with respect to the other objectives of the proposal including training and mobility & networking

Workplan (in common to P2, P3A, P4-P7):

The twin centres (P2, P3A, P4-P7) will contribute DNA from an average of 350 pairs per centre for the pilot genotyping for a total of 2000 pairs. Next a total of 6000 pairs for which data on stature, BMI, CHD, stroke and migraine have been collected will be identified to be used for the identification of liability genes; DNA samples will be collected if not already available. The relevant epidemiological, longitudinal risk information for these traits, CHD and stroke that has already been collected over the past years and decades will be identified, extracted and stored in a standard format. It will be made available to permit combining genotypic and environmental information in a comparative fashion over all centres and in collaboration with partner 8. This procedure will form the technical basis for later pooling of data on other traits, such as cardiovascular, autoimmune and neuropsychiatric traits. Each centre will act to collect additional data to further enrich the phenotypic and genotypic database, and will participate in infrastructure construction. Furthermore the twin centres may have additional centre-specific tasks that are presented in detail under the partner's role.

Participant 3 (A and B): University of Helsinki, Finland

A: PI: Prof. Jaakko Kaprio, Participant 3A: Finnish Twin Cohort Study

Mailing address. Department of Public Health, PO Box 41, Mannerheimintie 172, FIN-00014

University of Helsinki, Helsinki, Finland

e-mail: jaakko.kaprio@helsinki.fi *tel* +358 9 191 27 595 *fax* : +358 9 191 27 600

Scientific Team (A): Markku Koskenvuo, M.D., Ph.D., Kauko Heikkilä, Phil. lic., Anja Häppölä, M.Health Admin., Karri Silventoinen, Ph.D., Kirsi Pietiläinen, M.D., M.Sc.

Contractual links to other participants (A): none

Objectives (A): The role and contribution of the Finnish Twin Cohort study group will be to implement the objectives of the proposal locally with respect to creation of infrastructure, harmonized and standardized epidemiological phenotypic and genotypic databases as well as quality control and data security systems. Further, to coordinate the training and mobility component, with close liaison to national graduate and specialist training programmes. Finally, to act as networking contact for twin research groups which are not part of the Proposal, but wish to become involved in the proposed research and training activities.

Workplan (A): See P2

B: PI, Director: Prof. Aarno Palotie, Participant 3B; Finnish Genome Center

Mailing address: Haartmaninkatu 8, 00014 Helsingin yliopisto, Finland

e-mail: apalotie@mednet.ucla.edu *tel:* +358 9 1911 *fax* : +358 9 1912 5478

Scientific Team (B): Päivi Lahermo, PhD, Mika Salonoja, Msc, Juha Muilu, PhD, Maija Wessman, PhD, Elisabeth Widen, MD, PhD, Anna Rautanen, MSc, Timo Miettinen, BSc, Kari Tuomainen, BSc.

Contractual links to other participants (B): no

Objectives (B): The Finnish Genome Center will be responsible for genotyping most of the highly polymorphic multiallelic markers used in this study and in cooperating with the SNP analyses with Uppsala and NPHI.

Workplan (B): Genotyping of 60% of multiallelic (STR) markers is performed in the Finnish Genome Centre via a collaborative arrangement of the National Public Health Institute. The Finnish Genome

Center is directed by Dr. Aarno Palotie and has extensive experience of the genome-wide scans in large study samples (typically 3-8 96-well format DNA sample plates per study) using 500 markers and has optimized the denser set of 900 markers. The Finnish Genome Center is equipped with two automated, 96 capillary Megabase 1000 sequencers providing a theoretical genotyping capacity of 2 milj genotypes/year, enabling that the required 1 million genotypes can be produced in 6-12 months. Liquid handling is performed using Robbins and Tecan Microdispensers. Marker amplifications are performed using thermal cyclers operating in 384-well format. The lab workflow is controlled by a SQL*GT laboratory information management system (LIMS), which also stores all the genotypes, sample information, phenotypes and pedigrees. The genotypes are read using the Genetic Profiler program. The Genome Centre has eight full time employees in their bioinformatics unit responsible for data management, software development and statistical analyses. All statistical analyses are strictly supervised by statisticians of this consortium. In addition to a cluster of networked PC, Unix and Linux workstations, the centre is equipped with two midsize Sun Microsystems Unix servers with four processors and 2GB of central memory in each and 0.5 TB of fully mirrored, fault-tolerant disk space. The data is backed up every night to two physically separate tape robots with a total capacity of 10 TB.

Participant 4: Italian Twin Registry, Istituto Superiore di Sanità

PI: Dr. Antonia Stazi, *Mailing address:* Istituto Superiore di Sanità, Viale Regina Elena 299, 00161 Rome- Italy

e-mail: Stazi@iss.it *tel:* +39 06 49903181 *fax :* +39 06 49387173

Scientific Team: Dr. S.Giampaoli, Dr. L. Nisticò, Dr. M. Salvetti, Dr. R Tosi, Dr. F. Giubilei, Dr. C. Franceschi.

Contractual links to other participants:

Objectives: The Italian Twin Registry will implement the objectives of the Proposal locally, creating infrastructure, harmonised and standardised epidemiological phenotypic and genotypic databases as well as quality control and data security systems. It will participate in full scientific collaboration with respect to the other objectives of the proposal, including training and mobility.

Workplan: See P2

Participant 5: Netherlands Twin Registry, Vrije Universiteit, Netherlands

PI: Prof. Dorret Boomsma, Dutch Twin Registry

Mailing address: Vakgroep Psychonomie, Vrije Universiteit, De Boelelaan 1111, 1081 HV AMSTERDAM, Netherlands

e-mail: di.boomsma@psy.vu.nl *tel:* 020-444 8789 *fax:* 020-444 8832

Scientific Team: Dr. Eco de Geus, Dr. Toos van Beijsterveldt, Dr. Elles Mulder.

Contractual links to other participants:

Objectives: The Netherlands Twin Registry will implement the objectives of the Proposal with respect to creation of infrastructure, harmonized and standardized epidemiological phenotypic and genotypic databases as well as quality control and data security systems. We will participate in full scientific collaboration with respect to all objectives of the Proposal including training and mobility.

Workplan: See P2

Participant 6: The Norwegian Twin Registry, National Institute of Public Health, Norway

Dr. Jennifer Harris

Mailing address: Norwegian Institute of Public Health PO.Box 4404 Nydalen 0403 OSLO, Norway

e-mail: Jennifer.harris@folkhelsa.no *tel* +47 23 40 82 14: *fax:* +47 23 40 82 52

Scientific Team: Dr. K. Skjold Rønningen, M.D., Dr. C. Stoltenberg, M.D. with expertise in genetic and perinatal epidemiology, Dr. E. Røysamb, Dr. T. Reichborn-Kjennerud, M.D, and Dr. W. Nystad.

Contractual links to other participants:

Objectives: The role of The Norwegian Twin Panel will be to implement the objectives of the Proposal locally with respect to creation of infrastructure, harmonized and standardized epidemiological phenotypic and genotypic databases as well as quality control and data security systems. We will participate in full scientific collaboration with respect to the other objectives of the Proposal including training and mobility; and will also act as the coordinating centre for the Ethnical Issues and Consent Core.

Workplan: See P2

Participant 7: Karolinska Institutet, Sweden

Prof. Nancy Pedersen, Swedish Twin Registry

Mailing address: The Swedish Twin Registry, Department of Medical Epidemiology, Box 281, Karolinska Institutet, SE-171 77 Stockholm, Sweden.

e-mail: Nancy.Pedersen@mep.ki.se *tel:* +46-8-728 74 18 *fax:* +46-8-31 49 75

Scientific Team: Dr. Paul Lichtenstein, Professor Juni Palmgren and Assoc. Prof. Jan-Eric Litton.

Contractual links to other participants: none

Objectives: The Swedish Twin Registry (STR) will create a database and organise its infrastructure in order to harmonize and standardize our phenotypic and genotypic databases and also secure quality control and data security systems. This participant will also serve as a co-ordinator of the database networking and harmonization of the whole project. Further, to coordinate the training and mobility component, we will work with national graduate and specialist training programmes.

Workplan: See P2

Participant 8: Queen's University Belfast, UK

Prof. Alun Evans

Mailing address: Dept. of Epidemiology and Public Health, Mulhouse Building, RVH Grosvenor Road, Belfast BT12 6BJ, United Kingdom

e-mail: a.evans@qub.ac.uk *tel :* +44 28 90237153 *fax :* +44 28 9023 1907

Scientific Team:

Prof. P Maxwell, Dr. C Patterson, Dr. J Yarnell, Dr. Kari Kuulasmaa, Dr. François Cambien, Prof. Marco Ferrario, Dr. Denis Shields, Dr. Veikko Salomaa, Prof. Kjell Asplund, Prof. Hugh Tunstall-Pedoe and Prof. Elja Arjas. Also the individual MORGAM Participating Centres represent a powerful resource of epidemiological experience.

Contractual links to other participants: none

Objectives: Queen's University Belfast is the MORGAM Coordinating Centre, which is directed by Prof. Alun Evans, MD FRCP FFPHMI. The Investigators in MORGAM will employ their epidemiological expertise to clarify the prognostic efficiency of novel genetic and other risk factors, in concert with the twin cohort studies. The emphasis will be on cardiovascular diseases.

Workplan: MORGAM will be integrated with the twin cohorts and fully complementary to them, but in a sense, in assembling prospective cardiovascular and other events in a case-cohort design, will be free-standing and can investigate many different hypotheses relating to candidate genes and other factors. MORGAM is responsible for collating data from individual MORGAM cohorts, carrying out the statistical analyses and some targeted genotyping. It will develop statistical techniques to cope with the massive datasets generated. Queens University Belfast coordinates the multicenter MORGAM in more than ten centers in different European countries. The collection of epidemiological, especially follow-up data is very work-intensive, which is reflected by the high number of work months

accredited for this participant (P8). However, this is handled and funded by the respective individual MORGAM centers and thus we do not request funding for this work from EU.

Participant 9: University of Leiden, Netherlands

Dr. Lodewijk Sandkuijl,

Mailing address: Dept. of Medical Statistics, P.O.Box 9604, 2300 RC Leiden, The Netherlands

e-mail: sandkuijl@lumc.nl *tel :* +071 5276803 *fax :* +071 5276799

Scientific Team: Dr. L.A. Sandkuijl, prof. H.C. van Houwelingen, Dr. H. Putter

Contractual links to other participants: none

Objectives: The core is directed by Dr. Lodewijk Sandkuijl, a world-renowned statistical geneticist, who has had a major impact in the development of novel statistical tools for genetic analyses of diseases. He will coordinate the data analyses in individual centres and be responsible for pooled data analyses and harmonization and quality control of analytical strategies.

Workplan: Dr. Sandkuijl or his statisticians will visit individual centres bimonthly to guarantee the integration and harmonization. Also, the statisticians will be stationed periodically in different centres to ensure the training and integration works.

Participant 10: Uppsala University, Sweden

Prof. Ann-Christine Syvänen and Professor Ulf Pettersson

Mailing address: Molecular Medicine, Entz 70, 3rd floor, Uppsala University Hospital, 75185 Uppsala, Sweden

e-mail: Ann-Christine.Syvanen@medsci.uu.se *tel :* +46 18 61 12 959 *fax :* +46 18 61 12519

Scientific Team: Dr. Martha Alarcón-Riquelme, Dr. Tomas Axelsson, Prof. Ulf Gyllensten, Prof. Ulf Pettersson

Contractual links to other participants: none

Objectives: Uppsala University will be responsible for about half of the high-throughput SNP genotyping and for developing automated microarray technology for multiplexed, cost-effective SNP analyses. In addition, Uppsala University will contribute with one third of the multiallelic (STR) genotype analyses.

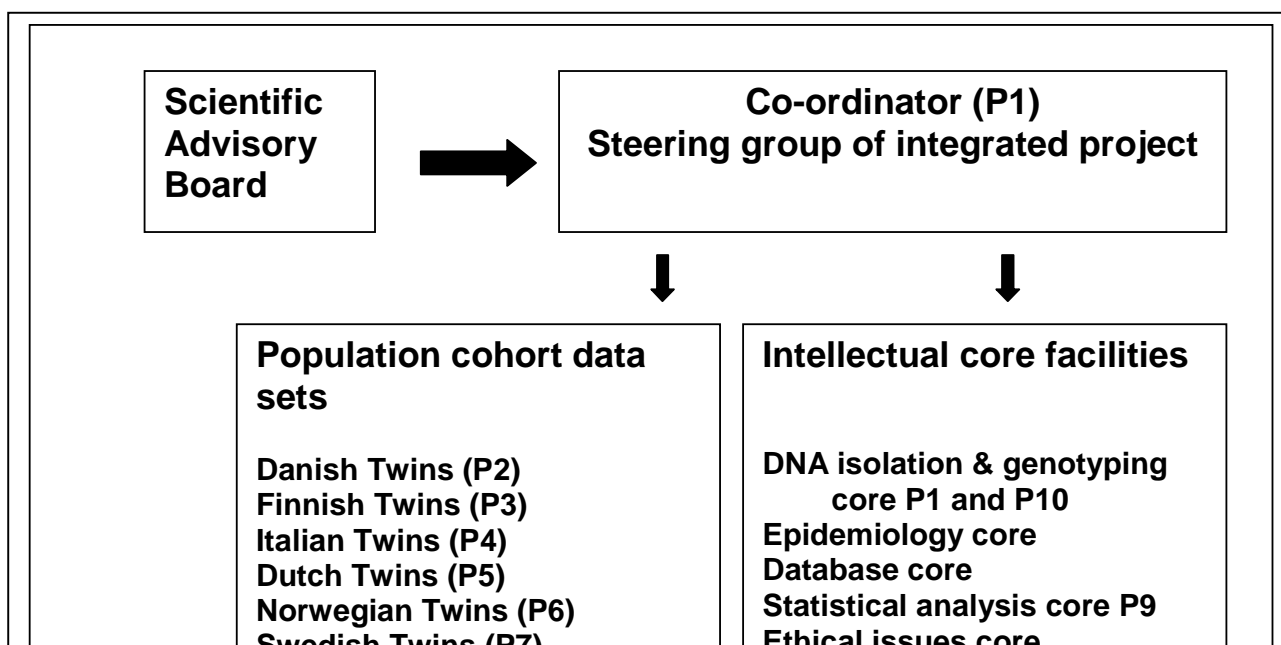
Workplan: The SNP genotyping analyses and technology development will be performed by the unit for molecular medicine, headed by professor Ann-Christine Syvänen, a pioneer in one of the central technologies for genotyping SNPs. A Swedish national facility providing high-throughput SNP-genotyping services, is well equipped with state of the art instrumentation for microarray technology (Cartesian array spotter, ScanArray 5000 four colour array scanner) and high-throughput SNP genotyping (PCR unit with two Packard Multiprobe robotic workstations and twelve 384 well blocks; FP Analyst AD) and a Nautilus laboratory management system. The research group is at the international forefront of developing microarray-based SNP genotyping methods. In addition to genotyping SNPs, special emphasis will be on developing approaches for multiplex PCR amplification and on automating the microarray-based SNP-genotyping system to establish a streamlined technology platform for high-throughput SNP-assays.

The STR genotyping analyses will be performed by the national center for genotyping, established in 1998, and headed by prof Ulf Gyllensten and prof Ulf Pettersson. The center is well equipped with robotics (10 ABI Catalyst 877), 3 4-block PCR instruments, capillary sequencers (one ABI 7700, one Megabase, ABI 377) and a Taqman ABI 7700 instrument, and has the capacity of performing 2 million genetic analyses per year. The center has effective biocomputing equipment and expertise, including 10 Unix workstations, and a number of softwares for data handling and statistical analyses. Software development is pursued to achieve higher throughput of data and a link to the center for parallel

computing at the Royal Institute of Technology has been established to secure storage of data. A panel of 600 microsatellite markers has been established at the center as a resource for gene mapping. The genotyping efforts in Uppsala will be performed closely co-ordinated with the corresponding activities at the Finnish Genome Center and the National Public Health Institute in Helsinki. Shared training and database facilities will harmonize the functions of these two Core components.

4. PROJECT MANAGEMENT AND COORDINATION, SCIENTIFIC ADVISORY BOARD

Overall management of the project is described in the graph below. Co-ordinator, together with the steering group (consisting of PI:s of each participating centre) is responsible for the overall strategy of the integrated project, for the progress of the scientific, training and networking components and for the monitoring of the budget. Co-ordinator will be responsible for the contacts with EC administration. To guarantee the excellence of the project's scientific work, networking and training, a Scientific Advisory Board (SAB) is formed. It will meet during month 10 in the annual meeting of project partners and have a special session with the steering committee more often via teleconferencing methods. We propose the following distinguished scientists with top expertise (the special area of expertise for each SAB member is given in the parenthesis) in the essential project areas of the project to form the SAB: Dr. Gert-Jan van Ommen, Leiden University (molecular genetics, genome-wide technologies); Dr. John Bell, Oxford University (molecular technologies, genetics of common diseases), Albert Hofmann, Erasmus University, Rotterdam (epidemiology, molecular epidemiology); Dr. John Blangero, University of Texas, San Antonio (statistical genetics, QTL statistics); Dr. David Evered, IARC, Lyon (ethical and social aspects of genetic research). The steering group will meet bimonthly to evaluate the project components and the budgetary issues. Co-ordinator can seek advice also from the SAB members especially in the quality control issues of different components. Co-ordinator will meet administrators of participating institutes once a year (month 9), before the budget plan for the next year is produced. Annual meeting of participating centers (see below: Project reporting , monitoring and review) will provide a forum for the exchange of information and for the meeting of the SAB with the project participants and provide the necessary information of the progress for SAB reports. We will organize small workshops targeted to special methodological and analytical issues to be organized 1-2 times per year. These are part of our networking and training program and also serve for development of technologies and methodologies of intellectual core facilities. The directors of the intellectual cores will be advising in topics of the databases. All the meetings and workshops will be scheduled 3 months in advance and the EC administration will be informed 2 months in advance of any event. The annual meeting will also involve a discussion between the EC representative and co-ordinators of the funded Genomic Centers that will help to harmonize numerous administrative issues and integrate research, training and networking components of these Centers.



Project reporting, monitoring and review

The following timetable will apply for the yearly reporting, project monitoring and review.

Project monitoring:

Year 1

The first progress report covering the period between month 1 to 10 should be sent by the coordinator to the Commission Scientific Officer in charge of the project at the latest by the end of month 10. At the same time, the project coordinator should send the scientific plan (modified Technical Annex) and budget breakdown for year 2. These documents will be forwarded by the Commission services to the Scientific Advisory Board (SAB) for assessment. The discussion on the first progress report and the evaluation of the new scientific plans for year 2 between the members of the consortium, the SAB and the Commission representative(s) will take place during the 1st annual meeting of the consortium (Month 12). An objective of this meeting is to reach a consensus on the scientific programme for year 2. The cost statements for the first year of the project should be sent to the Commission Scientific Officer within 2 months after the completion of year 1.

Year 2 and subsequent years

For years 2 and subsequent years, the same procedure will apply.

Final report

The final report and cost statement for the last year of the project should be sent to the Commission Scientific Officer within 2 months of completion of the project. The final report will be analysed by the SAB for their opinion on the success of the project.

Mid-term review

A mid-term review of the progress of the project will be instigated by the Commission services and organised by the coordinator and representatives of all participants. The Commission will be assisted by one or more external experts who shall be subject to confidentiality agreements. The mid-term review will take place at the second annual meeting (month 24) between the Commission services, the project consortium and the SAB. In advance of the review meeting, the Commission services will provide the coordinator with (i) a mid-term review manual detailing the review procedures and (ii) the template and guidelines to be used by the external experts in preparing their report.

Contract amendments

Request for contract amendments will be processed by the Commission services only once a year (months 13-15, 25-27, 37-39), excepting for special circumstances (bankruptcy, etc...).

5. EXPLOITATION AND DISSEMINATION ACTIVITIES

The health care industry, including drug development and disease prevention actions, is one of the largest growth areas in any developed country. The prevention and effective treatment of common, late onset diseases is one of the key goals of society. The expected outcome of this proposal will have a tremendous strategic impact on many areas of the European health care-associated industry. We will produce innovations that will have wide-ranging applications and have tremendous potential to improve the competitiveness of Europe. The results will be equally essential for the future of pharmaceutical companies and for many other fields of industry, including the food industry. The educational component of the project will produce the seedbed for future generations of public health policy makers and clinicians fluent in genetics, biocomputational analyses and epidemiology. Such experts are badly needed if the competitiveness of Europe, in many areas of life, is to be increased and also to make evidence-based decisions and judgments affecting the health of communities and individual patients. The steering group shall provide, by contract signature, a publishable summary of the project, which can be easily disseminated and distributed to the public. In addition, not later than the first report, the co-ordinator shall provide to the Commission a publishable poster targeted to a non-specialist audience and summarising the main features of the entire project.

6. ETHICAL ASPECTS AND SAFETY PROVISIONS

The coordinator shall implement the research project in full respect of the legal and ethical national requirements and code of practice. Wherever authorisations have to be obtained from national bodies, these authorisations shall be considered as documents relevant to the project under Article 27 of Annex II of the contract. Copies of all relevant authorisations shall be submitted to the Commission prior to commencement of the relevant part of the research project. The co-ordinator shall take all measures to assure that appropriate environmental safety provisions are fulfilled in the course of the project by all contractors, particularly those related to the deliberate release into the environment of genetically modified organisms. In addition the co-ordinator shall take all measures to assure for all contractors that, when dealing with biological material, strict safety procedures are in place in compliance with national and EU regulations on biosafety. All work must be carried out in compliance with national and EU regulations of safety.

This is an observational study and does not involve experimentation on human beings or animals. However, for a study of this magnitude, which includes genotyping of samples from large epidemiological cohorts, the ethical questions are of paramount relevance, and will be carefully reviewed by both external and internal boards. The principles of the Helsinki declaration and the Council of Europe's (COE) Convention on Human Rights and Biomedicine will be followed. We foresee that the coming years will witness a large amount of activity world-wide in the definition of ethical issues and formulation of policy to ensure proper handling of these issues. Therefore, during the tenure of our project we will monitor changes proposed by the COE to ensure that the conduct of our project reflects the highest level of ethical standards in biomedicine. Furthermore, each partner's national or local judicial scientific-ethical requirements and laws on data protection will be strictly adhered to. Any individual in the proposed research may decline to complete a questionnaire or may decline subsequent participation without penalty of any kind. Requests not to participate are honoured. Data to be collected are for research purposes only. All questionnaire and registry data are obtained in coded format, based on the assignment of research codes to each twin, for each individual member within each family and every participant in the population cohorts. This assignment is made at the

individual level, so that coding will be unique to the research files, and no one other than immediate local project staff can ever identify any individual. No outsiders can identify individuals or families in the subject samples. No names appear in any research data files. Only the researchers designated as data trustees at the individual study centres will have access to coded information linking the name and identification number of participants. Each person is assigned her/his own study number before DNA extraction, which is then used to identify all phenotypic and genotypic data at all times; name or other identifiable data is not used at the Genotyping Center at any time. Informed consent will be (is) collected at the point of participation. As a general rule, the genotypic data will not be linked to personal identification data. Because the study results are meant for scientific purposes only, the genotype or other genetic data will not be released to participants. In the population cohort component, in special circumstances, based on a strong recommendation of local Ethics Committees, the local Principal Investigators could release genetic results to the participants. Terms regarding the release of health or other DNA information to the study participants will be fully covered in the consent procedures where applicable and for new data collection procedures will comprise a separate consent item.

To emphasize the importance and essence of proper ethical conduct, the integrated project has an ethical core that will consist of leading experts and population cohort representatives. This Ethical Core will include representatives from each of the participating centers and will also include ethical and legal experts (who are currently being recruited), other than the scientists participating in this study. Including participants from each center safeguards the networking and ethical training influence of the Ethical Core such that it impacts all centers equally –the ethical training of the scientists participating in the analyses is of paramount importance for the structure of the whole project. To secure independence such that the ethical aspects of the project are unbiasedly addressed it is important that the ethical core includes outside experts as described above.

We would like to emphasize that the role of the Ethical Core is advisory; final decisions will be based on the laws established in each respective country by their national ethical committees. Note that the Ethical Core is not designed to act as an outside monitor –it will be wholly integrated in the project and is responsible that highest ethical standards are harmonically practised in all centers. This will also entail application of this study from each center to their national ethical boards for review and study permission. These reviews should reflect expert and outsider opinions within each country. Because this is an ongoing project it will be necessary for the Ethical Core to establish an internal monitoring process. This will assure that the principles set forth by the core are followed throughout the project period, will determine if there are new problems or ethical issues that the Core must address, and will guarantee that new actions and principles developed world-wide in biomedicine are considered for their relevance to this project.

Finally, in agreement with the recommendations provided to us in the Evaluation Summary Report, we recognize the importance of changing our procedure such that enough blood will be collected at least to some extent provide study materials for future potential RNA and protein analyses (frozen leucocytes and serum). We plan to collect 40 ml of blood in future collections to guarantee this option.

7. ONGOING EC FUNDED PROJECTS

University of Helsinki (P3A):

Factors in Healthy Eating. FP5 QLK-1999-916 .An EU-funded multicentre project, PI Janet Treasure, Institute of Psychiatry, London and collaborators include Prof. Dr. Johannes Hebebrand Philips University Marburg, Germany, Prof. Aila Rissanen Helsinki University Central Hospital, Prof. Dr. Sandro Sorbi, University of Florence, Italy, Dr. med Andreas Karwautz, University Clinic of Neuropsychiatry of Childhood and Adolescence, Vienna, Austria. Period of funding 1.4.00-31.3.03

University of Belfast (P 8):

Proposal acronym: MONICA: The Cohort Component. Period of funding: 01.05.98-30.04.02. Project ID: BM H4 CT98-3183. Type of Study: Shared Cost Action. This project is coordinated by Professor Alun Evans: other partners are Dr Kari Kuulasmaa, Dr François Cambien, INSERM SC7, Paris, France; Dr Denis Shields, The Royal College of Surgeons, Dublin, Ireland; Professor Hugh Tunstall-Pedoe, Cardiovascular Epidemiology Unit, Dundee, Scotland; and Professor Kjell Asplund, Department of Medicine, University of Umea, Sweden.

Karolinska Insitutet (P 7):

Cross-national determinants of quality of life and health services for the elderly (CLESA): EU funded. Coordinator: Dr. Stefania Maggi National Research Council – (CNR) Centre on Ageing, Italy, with participating centers in Finland, the Netherlands, Spain, Israel and Sweden.

Uppsala University (P 10):

Microarray technology for SNP-analyses,

- 1) SNP mapping resources for the functional genomics of Drosophila. QLRI-CT-2001-00004, FLYSNP, Nov 1, 2001 – October 30, 2004, Coordinator: Dr Barry J Dickson, Institute for Molecular Pathology, Vienna, Austria. Partner 2: Uppsala University, Molecular Medicine, Dept of Medical Sciences, PI A-C. Syvänen. Other Partners in Germany, Hungary, Spain, UK,
- 2) Research infrastructure to promote primate molecular biology. QLRT-2001-01325, INPRIMAT, Proposal submission Oct. 2001 Contract negotiations in progress, April 2002. Coordinator Prof. Xavier Domingo-Roura, Universitat Pompeu Fabra, Barcelona, Spain. Partner 4: Uppsala University, Molecular Medicine, Dept. of Medical Sciences, PI A-C. Syvänen, Other partners in Italy, Germany, UK, the Netherlands, France.