Genomes for LIfe

Cohort study of Genomes of Catalonia

The *GCAT| Genomes for Life. Cohort Study of the Genomes of Catalonia* is an ambitious project that has been set up at the *Institute of Predictive and Personalized Medicine of Cancer* to study the role of genetic and environmental factors in the development of common diseases. Its vision is to contribute to making the management and personalized treatment of the patient faced with the disease, Predictive and Personalized Medicine, a reality.

GCAT T: (+34) 93 554 30 54 hevans@imppc.org

List of Contents

1. WHO IS BEHIND GCAT? 4
1.1 GCAT is a project of the Institute of Predictive and Personalized Medicine of Cancer 4
1.2 The IMPPC is in the process of a merger with the IGTP 4
1.3 The IMPPC and the IGTP are on the Campus Can Ruti 5
1.4 Banc de Sang i Teixits (BST) is the collaborating partner 5
1.5 The Ministry of Health of the Generalitat supports the project 5
1.6 Catalan Agency for Quality and Health Monitoring (AQuAS) ensures data protection 6
1.7 The population of Catalonia are the participants and the beneficiaries 6
1.8 Collaborating Scientists will provide the results 6
1.8 Funding 6
2. SCIENTIFIC LEADERS 7
2.1 Director of GCAT and of the IMPPC, Manuel Perucho PhD 7
2.2 Scientific Director GCAT, Rafael de Cid PhD 7
2.3 Scientific Director IGTP, Manel Puig, MD, PhD 8
2.4 Investigator GCAT-ICO, Victor Moreno, MD, PhD 8
2.5 Director of Healthcare, Blood and Tissue Bank, Lluís Puig, PhD 8
3. THE PROJECT 9
3.1 Background - Why do we need GCAT? 9
3.1.1 Increasing load of chronic diseases 9
3.1.2 A more precise Medicine, towards health care focused on Everyone 9
3.1.3 Multiple Diseases and Treatments 9
3.1.4 Relevance - Who will the GCAT Study? 9
3.2 There are already biobanks. What is the GCAT Added Value? 10
3.2.1 The power of different sources of information 10
3.2.2 GCAT is a Prospective Study, a Cohort. 10
3.2.3 One piece in a worldwide network 11
3.3 How will we do it? 12
3.3.1 Biological material: Blood and DNA 12
3.3.2 Health Survey 13
3.3.3 Electronic Medical Registries 13
4. PARTICIPANTS 14
4.1 Who can participate? 14
4.2 Who is not able to participate? 14
4.3 Why participate? 14
4.4 Participants will NOT: 14
4.4 In the case that they have given their consent Participants will: 14
4.5 How can I sign up? 14
5. CONFIDENTIALITY AND DATA PROTECTION LAW 15
5.1 Anonymous Participants 15
5.2 Informed Consent 15
6. PHASES OF THE PROGRAMME 16
7. FUTURE RESULTS -WHAT KIND OF INFORMATION WILL WE FIND? 17
7.1 Population research 17
7.2 Medical Practice 18
7.2.1 Simple inheritance conditions 18
7.2.2 Complex inheritance conditions 18
7.3 More efficient pharmacological treatments 18
7.4 Future Results - Tools for Medical doctors 19
7.4.1 Guides for Prediction and Recommendations for good health 19
7.4.2 Improved Diagnostics 19
7.4.3 Improved Predictions of Reactions to Treatment 19

We have all heard stories of somebody's grandparent who smoked all their life but never got ill, while other people never smoked and however died of lung cancer. We know about people who do lots of sport who have high levels of cholesterol. Similarly we have heard about cancer patients who have tried one treatment after another but none have worked when other more fortunate people recovered from the same illness after the first treatment given to them. How can doctors know which group a patient will be in? How can they know which treatment will work best?

It is to help us find answers to these questions that the **IMPPC has set up the GCAT Project.**

GCAT T: 93 544 3054 hevans@imppc.org

1. WHO IS BEHIND GCAT?

1.1 GCAT is a project of the Institute of Predictive and Personalized Medicine of Cancer (IMPPC)

The Institute of Predictive and Personalized Medicine of Cancer (IMPPC) aims to contribute to change from the current model of diagnostic and curative medicine to a new model of preemptive and tailored medicine, focusing our activities on the field of cancer research.

Historically, medicine has consisted of treating a patient after symptoms of an illness have appeared and patients with similar symptoms have been given the same treatment. The new model of predictive medicine is based on technological advances that now allow us to find global information from the human genome. It focuses on healthy individuals with a view to evaluating their susceptibility to certain diseases to ultimately generate preventive clinical practices. At the same time, medicine based on pharmacogenomic knowledge aims to predict the risk to disease development, predict the response to treatments, and to generate personalized treatments that vary for each patient in function of their genetic profile.

Within this new paradigm of medicine, the mission of the IMPPC is to identify the molecular bases that will allow prediction of how and when a cancer will develop, to predict the response to treatment, and to enable the design of personalized therapies for the individual cancer patient.

The activities of the IMPPC focus on:

- Discovery of molecular markers and tools for cancer diagnosis and prognosis
- Identification of genetic and epigenetic variants predictive of resistance/ susceptibility to cancer and response to treatment
- Contributing to improvements in cancer prediction, personalization, and eventually prevention, for the entire research community and society in general
  http://www.imppc.org/index.html

1.2 The IMPPC is in the process of a merger with the IGTP

The Institute Germans Trias i Pujol (IGTP) is the research institute of the Hospital Germans Trias i Pujol (known as Can Ruti). As a CERCA centre of the Generalitat (Autonomous Government of Catalonia) and recognized by the Institute Carlos III as a Centre of Excellence in Health Research the IGTP coordinates many research activities on the campus. Its mission is to create a multi-disciplinary and multi-institutional environment that opens the gates to highly efficient translational research with a view to improving people's health and quality of life.
http://www.germanstrias.org/cat/laFundacio.php

1.3 The IMPPC and the IGTP are on the Campus Can Ruti

The Campus Can Ruti is home to the teaching and reference hospital Germans Trias i Pujol, the teaching unit of the Autonomous University of Barcelona (UAB), several spin-off companies and a long list of research institutions including 4 CERCA centres. The synergies and multi-disciplinary nature of the campus make it an important node focused on health within the Barcelona research cluster.



1.4 Banc de Sang i Teixits (BST) is the collaborating partner

The BST is the collaborating partner in the GCAT project who ensures that the logistics of sample collection function to the highest standards. The Blood and Tissue Bank is a public company of the Ministry of Health of the Generalitat of Catalonia. Its mission is to guarantee the supply and correct use of blood and tissues in Catalonia; it is a reference centre for diagnostic immunology and the development of advanced therapies.

Banc de Sang i Teixits

http://www.bancsang.net/es_index/
http://www.bancsang.net/donants/genomes-for-life/

1.5 The Ministry of Health of the Generalitat supports the project

Generalitat de Catalunya
Departament de Salut

The Catalan Ministry of Health supports the GCAT Project fully. It is represented on the Governing Boards of the IMPPC, the IGTP and the BST and oversees that the research programs are in line with its research mission: To encourage and plan the research policy and generate innovation in Health in collaboration with other Ministries responsible for

research, universities, research centres and the Catalan Agency for Quality and Health Monitoring and other institutions.

1.6 Catalan Agency for Quality and Health Monitoring (AQuAS) ensures data protection

The Catalan Agency for Quality and Health Monitoring (AQuAS) will be the responsible for monitoring the connection of health records of the participants, ensuring total compliance with the Data Protection Law, protection of anonymity of participates and security of the data.

1.7 The population of Catalonia are the participants and the beneficiaries



Although individual participants will not receive the results of blood tests or reports on what is coded in their genome, their data, together will the data of others from the same geographical region, will provide a wealth of information, which will help doctors improve the health of people in the future. Participation is totally altruistic, for free and given with a full written informed consent. People who donate their data will be helping to build a powerful research tool that will be used to resolve the health problems suffered by most of us as we grow older such as neurodegenerative or cardiovascular diseases, cancer, diabetes etc.

1.8 Collaborating Scientists will provide the results



The huge canon of data contained in the GCAT is a valuable resource for biomedical research and a source for many benefits for society as a whole. Scientists from the IMPPC and other public research institutions will be able to request access to the data.

New data generated from collaboration will enrich GCAT Database with new layers of complexity, and further development of the project.

The scientific results will be published with the scientists of the GCAT team as contributing authors or acknowledgements of the resource, which will further increase the prestige of the project, the IMPPC, IGTP and the research in Catalonia in general. The publications will be available to the research and medical community providing information for use by researchers and doctors everywhere. This will lead to further practical results, for example: health guides, changes in treatment protocols, development of new diagnostic tools or screening programmes and so on.

1.8 Funding

The project is fully funded by "Sub-programme of dynamic actions for researchers and technology SNS.2010. File ADE10/00026 from the Health Research Fund (FIS), Ministry of Health, Social Services and Equality of Spain, of which M. Perucho is the principal investigator, dedicated to the implementation of the Program of Genomics and Epigenomics of Cancer Prediction.

2. SCIENTIFIC LEADERS

2.1 Director of GCAT and of the IMPPC, Manuel Perucho PhD



Our aim has always been to make a major contribution to a change in paradigm towards predictive medicine. The GCAT is an important step along this path

Born in 1948 en La Roda, Albacete, Spain, Dr Perucho completed his PhD in Biological Sciences at the Complutense University of Madrid, Spain in 1976 and went on undertake his postdoctoral research at the Max-Planck-Institut für Molekulare Genetik, Berlín. He joined the Cold Spring Harbor Laboratory, USA in 1979 and in 1982 started working as Assistant Professor in the State University of New York (SUNY) at Stony Brook, where he was made a tenured associate professor in 1987. Dr Perucho moved to California in 1988 when he joined the California Institute for Biological Research (CIBR) in La Jolla as Research Program Director until he moved to the Burnham Institute for Medical Research in La Jolla in 1995. There his research centered on Molecular pathogenesis of cancer of the microsatellite mutator phenotype and genomic instability in gastrointestinal cancer. http://www.imppc.org/research-activities/cancer-genetics-and-epigenetics/. In 2009 he moved to Barcelona as Director of the IMPPC.

Dr Perucho is the author of more than 150 publications, many in reputed high visibility journals, including Cell, Cancer Cell, Science, Nature, Nature Genetics, Nature Medicine, etc. His papers accumulate over 15,000 citations with an average of over 100 citations per paper.

2.2 Scientific Director GCAT, Rafael de Cid PhD



The GCAT will be the way to create valuable tools for physicians so that they can predict disease, diagnose it more accurately and better predict patients' responses to treatment

Rafael de Cid studied biology at the University of Barcelona. His PhD thesis was completed at the Genetics Department of the Hospital Clínic in Barcelona and at the Institute for Oncology Research (IRO). As a post-doctoral researcher he joined a multidisciplinary team at the Center for Genomic Regulation (CRG) in Barcelona to work on molecular epidemiology of complex diseases, where he led the analysis of gene X gene and gene X environment in complex diseases, principally in immunological disorders as asthma, atopy and psoriasis. He was involved in the development of a genotyping facility (CEGEN-Barcelona node) to give support to the scientific community in genetic association projects.

From 2007-2012 he was at the Centre National de Genotypage (CNG) in Paris France as a post-doctoral researcher. There he studied the role of genomic structural variants in genomic susceptibility of skin disorders. Then at GENETHON, as an associate researcher, he was involved in the development of high-throughput DNA-array approaches using CGH and Exome sequencing approaches for rare muscular disorders and their applications in diagnosis.

In 2012 he joined the IMPPC and started his own team to study genomics of chronic diseases with complex inheritance. http://www.imppc.org/research-activities/disease_genomics/

2.3 Scientific Director IGTP, Manel Puig, MD, PhD



Modern biomedical research can only work with large numbers of people from different disciplines working together to pool ideas, share resources and bring different skills to bear to improve the lives of patients.

Dr Puig graduated in medicine in 1988 from the Autonomous University of Barcelona and continued his training as an endocrinologist in several centres, including two years as a researcher at the University of Texas Health Science Center. Since 2010 he has been Chief Physician of the Endocrinology and Nutrition Service at the Germans and Trias University Hospital (Can Ruti) and since 2013 he has been the Scientific Director of the Germans Trias Institute. Dr Puig has published over 150 papers in international scientific journals and teaches extensively both in Spain and other countries. He is currently the president of the Spanish Society for Endrocrinology and Nutrition

2.4 Investigator GCAT-ICO, Victor Moreno, MD, PhD



GCAT is an opportunity to identify determinants of chronic diseases that will help us developing personalized preventive strategies

Dr. Moreno is Professor of Preventive Medicine and Director of the Cancer Prevention and Control Program at the Catalan Institute of Oncology-IDIBELL in Barcelona, Spain. He has long experience in conducting cancer prevention studies, focusing on genetic and molecular epidemiology of colorectal cancer. He has contributed to the identification of genetic susceptibility loci, and molecular mechanisms involved in colorectal cancer etiology and progression. He leads the Unit of Biomarkers and Susceptibility at Catalan Institute of Oncology, with strong expertise in biostatistics and bioinformatics.

2.5 Director of Healthcare, Blood and Tissue Bank, Lluís Puig, PhD

The BST is deeply committed to supporting research. This project combines the best of altruism in the participants (some of whom are blood donors and some of whom are not) and state-of the art technology



Dr Puig qualified from the University of Barcelona as a doctor specialised in haematology and haemotherapy. He has been the Director of the Red Cross Blood Bank in Barcelona, Director of the Blood Bank in Sant Pau Hospital and responsible for Quality Control in the Blood and Tissue Banc (BST). He is currently Healthcare Director at the BST.

3. THE PROJECT

3.1 Background - Why do we need GCAT?

3.1.1 Increasing load of chronic diseases

We the citizens of Europe are **living longer and longer and most of us will suffer from one of the major chronic disease**s in our lives and many of us from more than one of them. These illnesses, such as neurodegenerative or cardiovascular diseases, diabetes and cancer seriously reduce quality of life and wellbeing and the cost of treating them is an ever-increasing part of health spending in developed countries.

3.1.2 A more precise Medicine, towards health care focused on Everyone

Traditionally when a person develops a disease they go to the doctor and are treated; but **how can we know who is more at risk of this happening**? How can we know who is more susceptible (genetic risk factors) and who is more affected by their environment (diet, exercise, smoking, pollution etc.)? The GCAT projects aims to answer some of these questions. Once diagnosed we all receive the same treatment, but different drugs work differently for different people and for some they may not work at all, produce side-effects or interfere with other medicines. GCAT studies will also produce information about who will respond best to different types of treatments. This information is **essential for doctors to give each of us the best treatment**, it wills also avoid over-prescription of costly drugs that do not work for everybody.

3.1.3 Multiple Diseases and Treatments

An additional challenge for doctors treating the aging population is that studies are usually done on a group of people with one illness, but how do **the different risk factors for diseases (genetic and environmental) and medications affect each other**? The answers to these questions lie in analysing huge quantities of data with the latest tools (Big Data).

By studying information from a large number of individuals the GCAT project aims to separate out these different factors and provide the doctors of the future with tools to predict who is more at risk, what actions they can take to reduce risk and what treatments will be most effective for them. The GCAT Project will track and follow-up many variables for each individual sampled.

3.1.4 Relevance - Who will the GCAT Study?

The GCAT project takes its subjects from the Catalan region in the North-East Spain. The whole Catalan population in 2011 was 7.539.618 people. The individuals will come from the entire geographical area of the region.

The Catalan population is a panmictic meaning it is not separated into distinctive and different groups or isolated populations by geographic features such as mountains or ethnics. In fact it is very similar to the European population. Data from this group can be added to data in other projects to study the population of Spain, or European as a whole. Additionally it can be added to data from samples from groups from all over the world to study differences and similarities between people from different places.

3.2 There are already biobanks. What is the GCAT Added Value?

3.2.1 The power of different sources of information

GCAT will combine three sources of information: data from the complete genetic code (whole genome sequence), detailed information about lifestyle, and also information from medical files.



Fig 1. The combination of three types of information will allow investigators to study how each one affects the health of the individual.

3.2.2 GCAT is a Prospective Study, a Cohort.

The participants of the GCAT are not selected because they are patients, as is the case for many of the collections from hospitals at the moment; they are a cross-section of 40-65 year-olds living in one geographical area and their follow-up; a population- based cohort. After the initial "snapshot" of their genetic information, their lifestyle habits, family history, a connection with their personal medical records* will allow researchers to know who stays healthy and who gets ill and what treatments they receive. Only in nthis way can scientists start to discover what keeps some people fit into old age while others develop diseases.

* See section on data protection

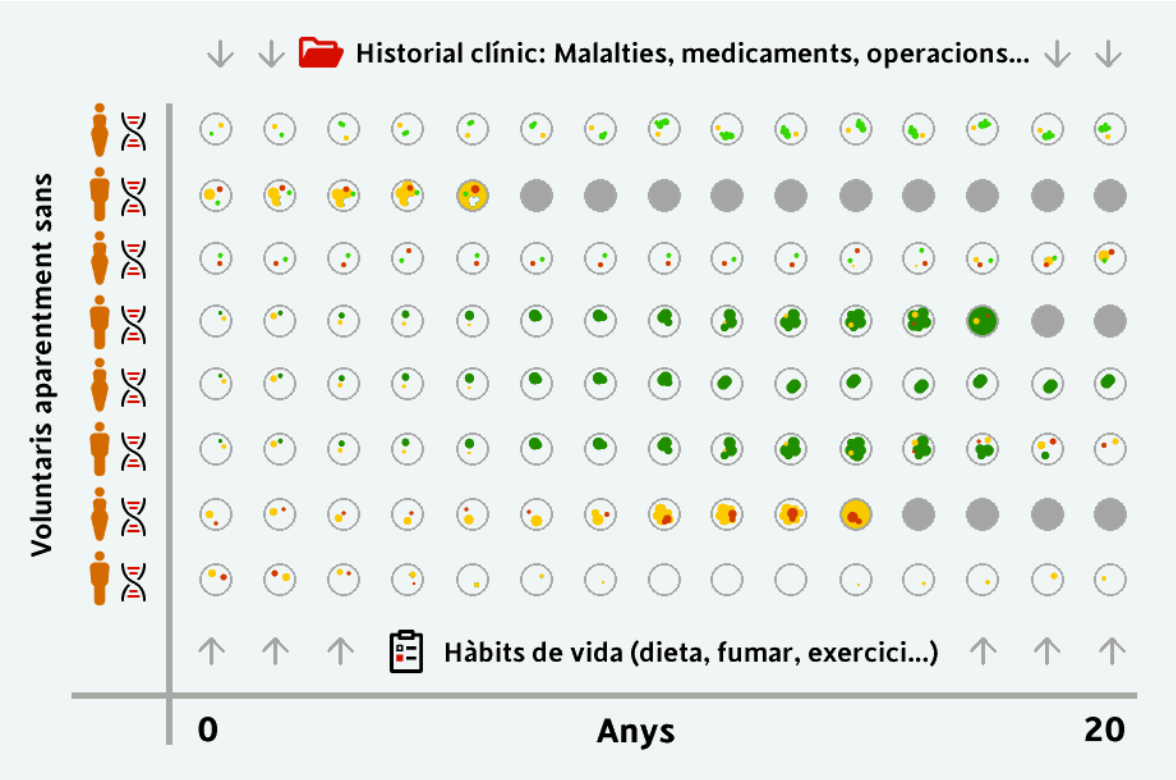


Fig. 2 An initial genetic snapshot will be followed to see what effects different habits have on health, what illnesses appear and what the responses to treatments are.

3.2.3 One piece in a worldwide network

The GCAT project is set up to comply with European standards, such as those of the Public Population Project in Genomics and Society (http://p3g.org/about-p3g) and the Biobanking and BioMolecular Resources Research Infrastructure (http://bbmri-eric.eu/about) so that data from this project can be combined with data from other similar studies. This means that researchers can potentially compare millions of people or look at different groups to study the effect of different factors in an efficient manner.



Fig 3. Map of other similar projects in Europe

3.3 How will we do it?

The GCAT characterizes whole genomes of participants to determine the complete genetic map of each individual. Genomic information will be used to determine, together with other pieces of personal health; life habits, family history and medical and clinical data, to study what is the correlation with health status of each individual.

After recruitment and the initial collection of epidemiological data from the volunteers in the cohort, the GCAT will follow up on all the individuals over the next 20 years. Along with the genomic information, follow up of health records will allow researchers to obtain health data in order to evaluate the associated risks for development, prognosis and treatment of illnesses.

Follow-up of the cohort means the register of data related to health:

A. Epidemiological data (periodic pro-active updates by volunteers)
B. Medical and clinical data (continual and passive via electronic records)
C. Mortality data (pro-active update on the part of GCAT)

3.3.1 Biological material: Blood and DNA

Blood sample is collected at each GCAT-BST donation centre

- A small blood sample, much smaller than a blood donation is collected and sent to the GCAT laboratories at the IMPPC.
- Blood is separated into, Serum, Plasma, Live cells, and DNA
- DNA will be analyzed completely to characterize the individual Genome; by Whole Genome Sequencing or Whole SNParray analysis.

3.3.2 Health Survey

- Body measurements

  - Blood pressure, heart rate
  - Weight and height
  - Hip and waist circumference

Body measures are taken by trained BST staff and added to the questionnaire

- Health survey

Information about personal social data

- Origen and Actual Residence city
- Family status, Familial Income
- Education and Working status

Information about life style habits

- Diet, exercise, smoking, alcohol, sun exposure,
- Social networking

Information about personal medical data

- Family history of disease
- Common Medical screening test
- Male and Female specific Health
- Drug use

Health survey data is collected via a self-administered questionnaire and participants are supported by trained BST staff.

Health survey questionnaires are adapted to men and women, and they are completed on-line from the BST donor centre. See the complete questionnaires here (http://www.gcatbiobank.org/participants/en_documents/

3.3.3 Electronic Medical Registries

Indirect follow-up of whole participants, by mining personal medical records will be done in accordance with the Catalan Health Public Agency from Ministry of health, to ensure absolute confidentiality, data protection.

* See section on Data Protection

3.3.4 Follow up - How often will participants need to give samples and complete questionnaires?

Initially there is only one blood sample and one questionnaire to be fulfilled.
Over the next 20 years participants will be follow up, and could be contacted to give another blood sample and/or answer some more questions about health status, always in a voluntary manner.

4. PARTICIPANTS

4.1 Who can participate?

- **Almost everyone** between 40 and 65 years of age, resident in Catalonia and not planning to leave in the next 5 years.
- You do NOT have to be blood donor to participate.

4.2 Who is not able to participate?

- Anybody who does not speak Spanish or Catalan (these people cannot answer the self-administered health questionnaire)
- People without a Social Security Health Card (these people may not have easy access to digital records here)
- People planning to live outside Catalonia in the next 5 years (these people's medical records may not be updated and they may not be able to give another blood sample in the future if necessary)

4.3 Why participate?

- Participating in the GCAT project is an altruistic act.
- Participants will contribute to a huge database of anonymized information, which will be used to improve the health of citizens in the future.
- They will receive general information about discoveries made from the data as a whole and tips for a healthy lifestyle.
- This is an opportunity to give and contribute to biomedical research for better health in the future.

4.4 Participants will NOT:

- Be sent blood/ DNA results or information about their own health
- Receive any information about diagnostics for illnesses they may or may not have

* See the section on data protection and incidental findings

4.4 In the case that they have given their consent Participants will:

- Receive information about clinical results from which they/ or their family can obtain any proved clinical advantage.
- Through Genetic Counselling Advisors GCAT will contact and give the advice of the convenience to consult a genetic specialist

4.5 How can I sign up?

If you are 40-65 years old, speak and understand Spanish or Catalan and have a Social Security health card you can join in.

- Call **93 557 35 00** and ask to participate in the GCAT
- Send an email to atencioaldonant@bst.cat
- Fill in this on-line form (in Catalan or Spanish) http://www.bancsang.net/donants/genomes-for-life/inscripcio_gcat/

You will be given an appointment at one of the collaborating BST Blood Donor Centres
http://www.gcatbiobank.org/participants/vull-colaborar/on-he-danar/ - centres BST

5. CONFIDENTIALITY AND DATA PROTECTION LAW

The data protection laws in Spain are among the strictest in Europe. The IMPPC and the GCAT Project comply with Data Protection Law and take protection of data and confidentiality very seriously.

5.1 Anonymous Participants

Participants identify themselves by their name and social security number when they give their blood sample and answer the questionnaire, these details are stored and they are given a unique ID number for the GCAT Project.

The key to the identity of participants will be accessible to the Scientific Director only through the BST, which holds and protects the database of personal data.

GCAT will only identify individuals by a non-identifiable code number. Neither the researchers working on the data or anybody else will know the identity of the participants.

5.2 Informed Consent

Participants must specifically authorize the use of their blood sample and data for the GCAT Project. They can also withdraw this consent totally, so that their samples and data are no longer used in studies on any health issue.

See the consent information here:
http://www.gcatbiobank.org/media/upload/arxius/Docs%20consent/Com_reclutament%20_GCAT_Annex_2a_Full_consent_MODEL_WEB_GCAT_Catalan_V2.pdf

See the consent form here:
http://www.gcatbiobank.org/media/upload/arxius/Docs%20consent/Consentiment_Copia%20web.pdf

5.3 Participants will not usually receive any results

GCAT Project is a research project, and DNA analyses are not diagnostic tests.

Participants can opt to receive information about studies in which their sample was included; in this case they might, for example, receive the overall results from a study of women over 50 who smoke if their sample was included.

5.4 Exceptional Results - Incidental findings

Whole genome analysis approach reveals all the genome information, and some unexpected research results, known as incidental findings, can be uncovered.

Participants can opt to receive any information about the personal results obtained, but only those results that might seriously affect their health or reproductive fitness or that of their children will be communicated to the participant.

In accordance to the American Academy of Medical Science, a list of genomic variants is selected on the basis of a clinical utility value.

GCAT has a protocol to evaluate the incidental finding and the opportunity to communicate to participants in accordance with the participant's wishes and the law.

We need to consider if:

- there is a considerable health risk
- actions can be taken to avoid it
- the participant has opted to be informed, in which case they will be invited to a session by a trained genetic counsellor

In this last case the Genetic Counsellor will explain the risks and consequences and actions that can be taken.

6. PHASES OF THE PROGRAMME

Project development phases

2012-2014	Design of the project and Set up of Lab facility
2014	BST collaboration agreement signed
2014-2017	Collection of samples and data from participants
2015	AQUAS collaboration agreement signed
2016-2035	Iniciation of follow-up of the cohorts

2015 Onwards

- Data mining of baseline health related data
  All collected samples are analyzed for initial health status

  Genomic variation analysis of the first 5.000 volunteers at the IMPPC to generate a map of genomic variation for the two million known genetic variants

  The currently known genome map, the resul to of other genome project,s will allow researchers to incorporate more than 7.000.000 known variants in the GCAT variation map,

- Whole Genome Sequencing of the first 1.000 participants
  A whole map of genome variation will be generated from 1000 individuals
  DNA samples will be analyzed for the whole genome sequence variation

  (Discovering the dark side of the genome, what was previously called junk DNA, but is now known to have a purpose, although this purpose is not yet known).

2016 Onwards

- Electronic Clinical and Medical Records Follow-up of the cohorts

  Integration of clinical data with genetic data for in-depth studies of different diseases and common risk traits will allow the analysis of disease and the

2015-2035 Long term Activities

- Genetic variation map within the population
- Genotype imputation reference GCAT panel
- Data mining of retrospective electronic health records phenotypes
- Data mining of prospective electronic health records phenotypes
- Contribution of genetic variants to disease and intermediate phenotypes
- Expansion of age range cohort: Family GCAT collection
- Development of a EHR-linked DNA collections resource
- Definition of biomarkers for predisposition, prognostics and prediction
- Development of tools for translating genomic to clinical useful information
- Educational programs for Participants

7. FUTURE RESULTS -WHAT KIND OF INFORMATION WILL WE FIND?

It has been established that genomic profiling, large enrolment of participants, genomic linking to medical records and informatics are the keystones of the development of stratified precision and further personalized medicine.

Molecular characterization and follow-up of a large cohort of individuals will generate a large number of combinations, that are needed for studying the interaction between the genetic inheritance we are born with (our genome), and the modifications to our genome (our epigenome) caused by the interaction with environmental factors to which we are exposed.

The objective of the GCAT project remains open to accommodate new hypotheses and future knowledge. Here below we detail some examples in which genome information is useful and could add an extra value to research and medical practice:

7.1 Population research

Some examples follow

The Genome variation map will allow investigation of the similarties and differences of the population compared with other continental super populations, such as the named European Ancestry populations as well with specific populations in similar environmental conditions.

For example, whether different populations have different propensities to high blood pressure.

Why are different ancestry populations affected differently by disease?. How can environment modulate theses differences? What are the molecular differences?

"American Native Indians have a prevalence of diabetes higher than other American populations. Is this related to genomic features fixed as consequence of population isolation? Can these differences identify new biological pathways involved in disease, and thus propose targets for new drugs?"

Even if a gene is identified in several populations, different genomic variants could be associated with the same disease, so the analysis in several populations reinforces the study, pointing to a the functional role of the gene involved

*" **ADAM33** gene has been linked to asthma in several populations but, in each populations different genetic variants have been observed associated with it"*

7.2 Medical Practice

The success of this initiative will be evaluated taking into account the degree of translation to clinical research and to clinical practice. Some simple conditions are largely incorporated in current medical practice, but the goal is incorporate all genome information making it standard practice for most common diseases.

The genomic era is now starting to take its first steps, and even when clearly promising, we need to be careful not to create immediate expectations and exaggerate hopes of medical applications.

Below are some examples of medical practices where genomic information is already used.

7.2.1 Simple inheritance conditions

These conditions are due to the affectation of one single gene, alterations in it have known consequences:

For many of these genetic conditions the general population is already screened examples include: endocrine diseases (congenital hypothyroidism), hemoglobin pathologies (thalassemias), congenital errors in metabolism of fatty acids, organic acids, urea or amino acids (such as Phenylketonuria),neuromusclar diseases (Duchenne) or cystic fibrosis.

7.2.2 Complex inheritance conditions

In these cases, the majority of common diseases, there are multiple factors at play and not just one gene. Here there are genetic risk factors and environmental risk factors, in many cases we don't know the function of the affected genes and don't know how variants we find might affect them.

The knowledge of risk factors could allow us to carry out preventive actions for particular groups at risk.

Example 1. Variants in the MCM6 gene affect the regulation of the gene for Lactase, which affects an individual's ability to digest milk. Not all populations are the same and knowing in which group a population is can allow for specific public health campaigns.

Example 2. There is a simple variant in a gene, which only has two forms. A DNA analysis allows doctors to identify people at risk of a fatal cardiovascular event.

7.3 More efficient pharmacological treatments

The knowledge of the genetic profile of an individual allows us to apply more efficient pharmacological treatments.

At the moment there are genetic tests accepted by medical authorities that allow patients to be stratified (divided into groups) for more efficient treatment. Most of these treatments are for cancer, but they also exist for some infectious diseases and cystic fibrosis.

Example 1. Carbamazepine
This drug is prescribed for depression. A variation in the HLA gene can mean that a patient suffers unpleasant side effects. In most European populations this variant is rare, but in Asian populations it is high, meaning that medical protocols can vary according to genetic knowledge.

Example 2 Cetuximab
Cetuximab is an anti-cancer drug that is very specific for EGFR receptors in cancer cells and prevents them multiplying. If a patient also has mutations in the KRAS gene the treatment with Cetuximab doesn't work, meaning that they undergo a very expensive treatment that will not work. Identifying who the drug will

benefit means patients can receive a drug that works and not waste time and that money is saved not medicating people who will not respond.

7.4 Future Results - Tools for Medical doctors

7.4.1 Guides for Prediction and Recommendations for good health

- Identification of genomic markers for people more at risk of a disease or more likely to live a healthy life
- Identification and quantification of risk activities that will increase certain people's risk of developing a disease considerably

7.4.2 Improved Diagnostics

- Identification of genomic markers for susceptibility to disease leading to new simple or non-invasive tests
- Identification of the interaction of certain behaviours with genomics to identify groups at high risk (for example who is at greater risk of skin cancers due to exposure)

7.4.3 Improved Predictions of Reactions to Treatment

- Guides for doctors about how different people will react to treatments to identify the correct treatment for each individual and avoid trying different treatments randomly until one works in order to avoid mental distress, side-effects and wasting time
- A more cost-effective use of treatments targeting individuals for the treatment that works for them

More information

GCAT Project
IMPPC Building
Campus Can Ruti
Badalona

Communications: Harvey Evans
(+34) 554 3054
hevans@imppc.org