Inserm

La science pour la santé
From science to health

Submission deadline: April 2, 2018

CROSS-CUTTING PROGRAM

GENOMIC VARIABILITY 2018

Research Projects in the Field of Genomic Variability

Online submission https://eva3-accueil.inserm.fr

Contact variabilite-genomique@inserm.fr

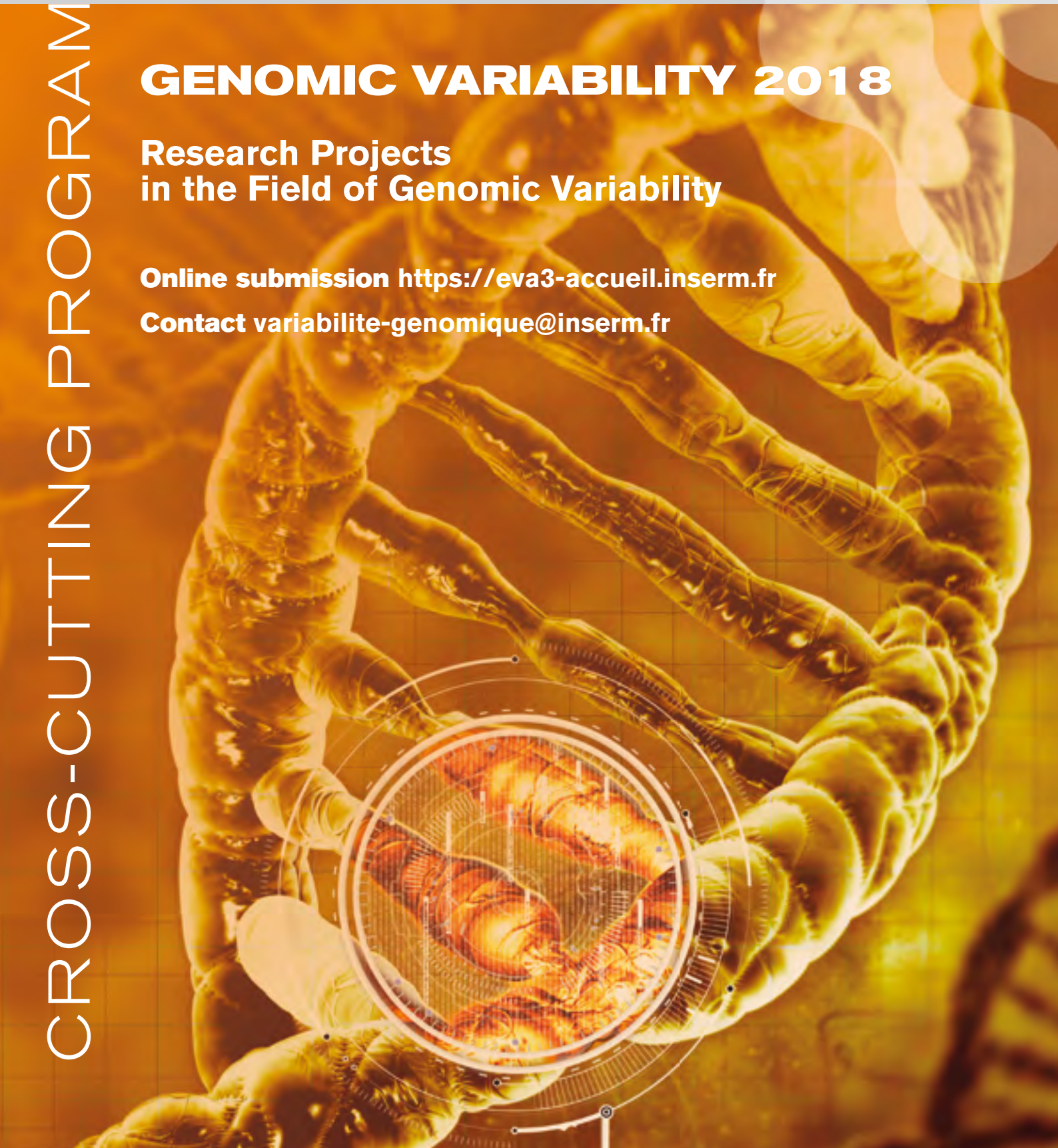

TABLE OF CONTENTS

Challenges and issues in the field of health and biology are constantly changing and opening up new directions for innovation. In this context and in line with its missions, which consist of accelerating progress in knowledge, supporting integrated and multidisciplinary research, and ensuring a continuum between fundamental and clinical research, Inserm is putting in place cross-cutting scientific programs with the following objectives:

- to build scientific communities in specific high-priority fields and bring forth national interdisciplinary consortia that will build on the skills and expertise of Inserm teams;
- to make French biomedical research a leading player in these fields by accelerating knowledge acquisition, transfer, and value-creation, including by potentially working with industry from the initial creation of cross-cutting programs.

These unifying programs aim to create a new dynamic in innovative fields by developing complementary skills for exploring research niches that have as yet been little studied. Funding will only be provided to collaborative projects that spread across a group of activities: several workpackages to be implemented by a consortium of teams. These programs are open to both academic and industrial partnerships with a flexible approach: such partnerships may be agreed across the whole program, or across one or more of the program workpackages only.

Scientific questioning at the frontier of biological knowledge, new technological opportunities, pooling the strengths of Inserm teams working in the field of the cross-cutting program, and potential added benefits in terms of societal value-creation constitute the determining elements for implementing these programs.

1 GENERAL PRESENTATION

The role of genes has been demonstrated in numerous human diseases but, until recently, the means for exploring variations in the human genome remained fairly limited and only a small part of genetic variability was understood. The recent development of new sequencing techniques has enabled much more detailed characterization, capable of identifying all the variants present in the genome of an individual or a collection of individuals. These genetic variants impact gene function to varying degrees, from a complete absence of the protein coded by the gene to variations in the level of protein expression in different tissues. Improved understanding of the impact of different variants and of genotypes on phenotypes remains key to using genetics for therapeutic or predictive ends. Having health measures and "omics" data for a collection of well-characterized individuals would be a major step to achieving this.

France has major resources in terms of cohort studies and in particular cohorts derived from the general population, such as the Constances study, formed of a random sample of 200,000 adults aged 18 to 69. This cohort, classed as a National Infrastructure in Biology and Health by the Investments for the Future program, aims to be an infrastructure provider for the research community. Following completion of a self-administered questionnaire at home, participants included in the cohort benefit from a health examination collecting health data (clinical examination, blood test, measurement of blood pressure, weight, height, and height/hip ratio, electrocardiogram and spirometry, eye and hearing examination, etc. ; for an exhaustive list see the documents available on the study website **www.constances.fr**). Monitoring is both active, through an annual self-administered questionnaire, and passive, through the extraction of data from national sources such as health insurance and hospital databases. The acquisition of innovative biological measures for this cohort as well as whole genome sequencing data aims to be a driver for improving France's international position in the field of modern epidemiology in the "omics" era.

2 OBJECTIVES

The objective of the cross-cutting program Genomic Variability 2018 is to understand the role of genes and their variants in the development of disease, based both on longitudinal monitoring of cohorts of individuals and their phenotyping, and by enabling the development of new methods for analyzing longitudinal data. This knowledge base will contribute to improving interpretation of the association between genetic variants detected in an individual and a disease or endophenotype, and to better comprehending the role of genes in phenotypic variability.

In order for this cross-cutting program to truly add value to Inserm groups, it must spur the **formation of a multidisciplinary consortium** combining groups of clinicians, geneticists, epidemiologists, bioinformaticians, biostatisticians, and biologists invested in the central question of the interpretation and impact of genomic variations on the phenotype, as well as mathematicians and statisticians who will develop innovative methods of data analysis.

The cross-cutting program will be part of the French Plan for Genomic Medicine 2025, which envisions genomic medicine in which sequencing will be used to guide diagnosis, contribute to prognosis, and choose the treatment best-suited to the patient. It will be carried out in close partnership with the General Population project of the plan directed by Emmanuelle Génin,

which primarily aims to respond to the need for "filtering out" the exomes/genomes of individuals affected by diseases with a genetic component, in order to eliminate genetic variants shared by the general population, by providing reference data on a panel of individuals representative of the populations from which patients are derived. Its secondary objectives are to answer questions from population genetics research about genetic diversity in France, as well as genetic epidemiology and statistics questions, by developing new methods for analyzing sequencing data in order to enable comparisons of the "burden" of rare gene variants between people with an illness and the general population, particularly for the study of complex diseases. This project is partly based on data previously collected from various genetic inheritance projects supported by France Génomique and LabEx GENMED, which have enabled the sampling and sequencing of almost 2,000 individuals representative of different French regions as selected on the basis of the birthplaces of their parents and grandparents in defined geographic areas. The funding that will be allocated to the General Population project by the French Plan for Genomic Medicine 2025 will enable it to continue recruiting individuals from the general population selected from the Constances cohort and to sequence the entire genome of these individuals to create a database of genetic variations in our populations.

Recent genetic research has undermined the idea that a single gene might control the nature or be the cause of a disease. Even for monogenic diseases in which so-called "causal" mutations are known, this hypothesis is questioned by the observation of strong phenotypic heterogeneity among different carriers of the same mutation, who may be in perfect health or conversely develop very severe forms of the disease. Having access to richer phenotypic information than the simple "disease/no disease" dichotomy is therefore a key asset for improving understanding of the role of genes in the phenotype. An important and too-often neglected aspect of this, due to the difficulty of acquiring long-term data, is the longitudinal dimension and the monitoring of phenotypes over time among the same individuals, followed by the integration and interpretation of this data. The cross-cutting program on genomic variability offers three workpackages. The first must lead to the formation, based on the resources of the Constances cohort and the General Population project of the French Plan for Genomic Medicine 2025, of a cohort of individuals contributing to the identification of genetic variants shared by the general population, and for which different phenotypes and biomarkers will be measured and monitored over time. The second workpackage will aim to develop innovative methods for integrating heterogeneous data types. The third and final workpackage will seek to interpret these results in order to better understand the role played by genes in phenotypes.

WORKPACKAGE 1 Phenotyping and sequencing in the Constances cohort

The objective of this workpackage is to **form a cohort of individuals representative of the various French geographic regions** whose genome will be sequenced as part of, and funded by, the General Population pilot project of the French Plan for Genomic Medicine 2025. The individuals in this cohort will be characterized by a **group of biomarkers and endophenotypes** consistent with the study of different diseases. The idea is to create a **genetic resource** that will serve as a reference panel to which data from other cohorts can be compared, in particular cohorts of patients affected by different diseases. The proposed biological measures must not require testing that is too invasive to be requested from the volunteers included in the Constances cohort as part of a monitoring study that has already been scheduled.

This workpackage thus aims to overcome a primary obstacle: the acquisition, from the same individuals, of different biological data such as the profiles of gene expression over time and the profiles of the expression of proteins or different metabolites that would enable **a link to be made between constitutional genetic information and protein expression and function.**

→ Deliverables

- Endophenotypes of interest that will be useful in the study of different diseases and can be compared with data that is already available or currently being acquired on cohorts of patients affected by these diseases to test well-supported hypotheses.
- Biological measures and biomarkers enabling the study of the impact of genetic variations on protein expression.
- The creation of databases of genetic variations and phenotypes that can be paired with the Constances databases.

→ Target audience

Geneticists, biologists, or epidemiologists involved in research projects investigating genetic factors associated with diseases that can be characterized by endophenotypes or by specific biological measures.

WORKPACKAGE 2 Methodological developments for the integrated and longitudinal analysis of genomic data with the Constances data

This workpackage aims to **develop new methods** that will enable analysis of the genomic and phenotypic data collected from the Constances cohort.

The first obstacle that this workpackage will look to overcome is that of the methodological developments for **integrating longitudinal information into genetic analyses.** The majority of studies seeking to demonstrate the role of a gene in a disease neglect the progressive dimension of the phenotype over time. Even when it is available, longitudinal information is very rarely taken into account in genetic association studies that define the phenotype (the dependent variable) with a binary (affected/not affected) or quantitative (measured at a given moment) measure and very often consider explanatory variables to be fixed and time-independent. Many variables are however modified over time, including for example levels of gene expression or in many cases environmental exposure. To analyze this data, it is often necessary to define these variables by a unique measure that overlooks their temporal dimension, leading to a loss of information. A second obstacle is the **integration and pooling of data.** This would enable communities of geneticists specializing in human disease to be brought together more easily and quickly with communities working on other types of presentation, in particular phenotypic.

→ Deliverables

- The development of new methods (from machine learning and data mining, mathematical models, statistical methods, etc.) for analyzing genetic, genomic, and presentation data in a way that takes into account their temporal dimension.
- The development of computing programs and procedures and a conceptual framework for the integrated analysis of heterogeneous “omics” data.
- The development of recommendations concerning good practice in the integrated analysis of “omics” data.
- Dynamic modeling of molecular networks and the development of mathematical and computing approaches to predict in silico the effect of genetic variants on various molecular, cellular, tissue, and other levels.

→ **Target audience**

Mathematicians, biostatisticians, geneticists, epidemiologists, computing experts, bioinformaticians, biologists, and clinicians from Inserm teams able to provide necessary knowledge, skills, or expertise.

WORKPACKAGE 3 Interpretation of the role of genetic variants on phenotypes

This workpackage aims to discover which genetic variants may be associated with the phenotypes measured in the Constances cohort and to better understand the role that genes may play in these phenotypes. It will use data generated from workpackage 1 and the methods of data analysis and integration developed in workpackage 2 in order to **overcome the obstacle of interpreting the role of genetic variants.** Association studies enable correlations to be shown between the presence of a given genetic variant in a gene and a particular phenotype, but these correlations are not proof of a direct effect from the genetic variant on the phenotype. To demonstrate this causal link, it is necessary to access biologist expertise and develop new, simple, and low-cost functional analysis methods to validate the hypotheses generated from mining big data. The objective of this workpackage is therefore to link phenotypic variations to genes and to identify within these genes **the key genetic variants and underlying molecular mechanisms.**

→ **Deliverables**

- The development of new causal inference methods applicable to genetic data and longitudinal phenotypic measures.
- The implementation of new methods for functional screening of the effect of genetic variants, enabling the direct effect of the alteration of the gene by the variant to be shown.
- The study of the regulation of gene expression: research into the transcription and splicing regulatory sequences.

→ **Target audience**

Bioanalysts, geneticists, biologists, statisticians, and epidemiologists.

3 OPERATIONS OF THE CROSS-CUTTING PROGRAM

Governance and organization

The cross-cutting program: based on the formation of a scientific consortium, organized around scientific workpackages, each composed of a number of scientific teams that may vary depending on the objectives. This consortium will be led by a scientific coordinator and guided by the heads of each workpackage.

The scientific expert committee: responsible for selecting the letters of intent, producing recommendations for the directions of the cross-cutting program, advising on alignments across the teams for the formation of workpackages, and approving the final scientific project. It is composed of international outside experts and directors of the relevant Inserm thematic institutes.

The scientific monitoring committee: responsible for monitoring the progress of the scientific project. It is composed of the consortium scientific coordinator and the scientific leads from each workpackage.

The program management committee: responsible for managing the running of the program, including the budget, and for approving proposals from the scientific monitoring committee for activities relating to the implementation of the overall program strategy. It is composed of the legal representative of the coordinating institution and the directors of the thematic institutes relevant to the theme of the program, in this case the Public Health Institute and the Genetics, Genomics and Bioinformatics Institute.

Establishing the cross-cutting program

Preparing the cross-cutting program

A working group composed of field experts has compiled a list of the relevant scientific issues in order to bring together complementary skills. Their discussions have culminated in the proposed program definition.

Putting in place the consortium

The consortium is organized around workpackages. Participation in the different workpackages will take place in two stages: an initial selection of candidates by the scientific expert committee on the basis of letters of intent (**see Evaluation criteria p. 10**), followed by a stage of co-construction of workpackages (see below).

Submission of the letter of intent

A single researcher or a research team may submit a letter of intent. This will specify the program workpackage being applied for and will describe the way in which the skills and expertise of the researcher or the team may contribute to overcoming one or more of the conceptual and/or technological obstacles identified as high-priority scientific components of the cross-cutting program.

Co-construction of the workpackages

Following selection of letters of intent, the coordinating institution (Inserm), based on the proposals and recommendations of the scientific expert committee, will invite the selected candidates to group themselves by workpackage and contribute to drafting a scientific project. This project will be presented to the steering committee and the scientific expert committee, as part of a discussion seminar.

Following this seminar, and taking onboard the committee recommendations, the consortium scientific coordinator will file a final scientific project with the coordinating institution (Inserm) that details the contribution of each team, the objectives, and the expected added benefits. Following formation of the program, a 3-year funding plan will be detailed and external funding sources identified.

Monitoring the consortium

The steering committee will organize an annual general meeting that brings together, in addition to its own committee, the scientific expert committee and the scientists involved in the consortium. During this meeting, participants will present and discuss the progress of the cross-cutting program, the next stages to tackle and, if necessary, propose new directions for research.

4 ELIGIBILITY CRITERIA AND EVALUATION OF LETTER OF INTENT

Eligibility criteria

To be considered eligible, the letter of intent must satisfy the following conditions:

- the letter of intent to participate in the consortium must respond to the objectives of this call for projects and fit into at least one of the workpackages described above;
- the researcher must be a researcher or tenured teacher-researcher working within an official Inserm team. They may, for the purposes of the project, propose a collaboration **with a researcher or a team** from other institutions, with the agreement of such parties;
- the researcher must specify:
    - their time commitment to the project;
    - the resources, particularly in terms of staff or equipment, that they intend to use as part of the cross-cutting program, in agreement with the responsible party from the partner institutions.

Evaluation criteria

After verifying eligibility, letters of intent will be submitted for evaluation by the scientific expert committee. Letters of intent not meeting the eligibility criteria will not be evaluated.

The evaluation criteria are as follows:

- **Quality and originality of the proposed research**
    - Clarity of objectives and research hypotheses
    - Innovative nature and advancement on current state of the art
- **Skills/expertises**
    - Relevance of skills to the program objectives
    - Possibility of combining skills within a broad network
- **Excellence of the research team(s)**
    - International recognition
    - Skills of the team leaders within their discipline
- **Quality of the research environment**
    - Human resources to be used by the program
    - Infrastructure available to carry out the program
- **Innovation/competition**
    - Innovative nature of the project in relation to international scientific issues or in relation to international competition
- **Expected added benefits**
    - Impact of added benefits in terms of knowledge and overcoming technological obstacles
    - Role of the project in the consortium construction in response to international calls

5 ELIGIBILITY CRITERIA OF THE FINAL PROJECT

To be considered eligible, the final project must satisfy the following conditions:

- the project must respond to the objectives of the cross-cutting program;
- each workpackage must include at least two teams with complementary skills, at least one of which must come from an Inserm research unit;
- the consortium coordinator must be significantly involved in the project.

6 PROGRAM CALENDAR

Opening of the project submission website		March 1, 2018
Submission deadline for the letter of intent	Online submission of the letter of intent	April 2, 2018
Meeting of scientific expert committee to select the letters of intent		April 2018
Co-construction of the workpackages		May 2018
Seminar and presentation of the workpackages		June 2018
Submission deadline for the final project	Project submission to the coordinating institution	July 2018

7 OPERATING PROCEDURES OF THE CONSORTIUM

Coordination of the consortium

The coordinating institution of the consortium is Inserm. Inserm is responsible for implementing the chosen project within the cross-cutting program and, if necessary, formalizing collaboration between the partner institutions (public or private), in particular by drawing up an agreement relating to the consortium, the production of project deliverables – including the production

of scientific reports –, the organization of progress meetings, and the communication of results. The partner institutions designate the public or private entities in which the partner units are involved as part of the cross-cutting program. The partner units designate in particular the research units, services, and teams involved in fulfillment of the project and placed under the responsibility of one or more partner institutions.

If non-Inserm partner units are involved in the consortium, they must have prior consent for their administrative supervision.

Duration of the project

Duration of the project is 3 years.

Scientific reports

The scientific coordinator of the consortium will provide scientific reports to the coordinating institution according to the Charter of good practices and the procedures defined below.

They will be sent as per the following schedule:

- a progress report 6 months after the project has started;
- a report halfway through the project;
- a final report no later than 2 months after the end of the project.

The scientific evaluation of intermediate and final scientific reports by the steering committee may lead Inserm to request additional information, to suspend the project, or to end financial support if the project is not being run properly or funding is being used for another project.

Responsibilities of the scientific coordinator

The consortium scientific coordinator must inform Inserm and its partners, if necessary, via the steering committee, of any substantial modification of the research project or any difficulties hindering project completion.

The consortium scientific coordinator must also participate actively in the project monitoring procedures organized by Inserm (presentation seminars, colloquia, etc.).

Publications – communication

All publications resulting from the research project must include the following funding statement:

« **Inserm cross-cutting program Genomic variability 2018** »

or

« **Programme transversal Inserm Variabilité génomique 2018** »

These publications are sent to Inserm for reference as soon as possible and at the latest five (5) days following publication.

Intellectual property

The results of the project belong to Inserm and to the project partner institutions.

The rules of ownership and the use of results from the project are defined as follows:

- between various partners in the context of a joint research facility: the applicable rules are those generally in force between the said various partners (in particular those of a joint research agreement);
- between various partners associated with several research structures, these rules will be defined in a separate consortium agreement.

8 RULES FOR SUBMISSION

Submitting the letter of intent
The submission of your application involves a mandatory stage: registration on the Eva Inserm website and online submission of the letter of intent.

This submission procedure, through the Eva Inserm website, includes:

- providing candidate information (surname, first name and email) enabling you to receive a user code and password providing access to a secure Eva personnel space;
- uploading the letter of intent to the Eva website.

Submission deadline: April 2, 2018

You are strongly advised not to wait for the closing deadline before submitting your letter of intent.

Submission of the final project
This will be submitted to the coordinating institution, Inserm.

9 PUBLICATION OF RESULTS

The list of candidates selected from the letters of intent will be published on the Eva Inserm website **https://eva3-accueil.inserm.fr**

10 CONTACTS

For further information you can contact:

- in relation to scientific and technical matters: the Public Health Institute and the Genetics, Genomics and Bioinformatics Institute **variabilite-genomique@inserm.fr**
- for questions relating to online submission **eva@inserm.fr**

Inserm is an organisation dedicated to biological and medical research as well as human health.

It is involved in the entire range of activities from the research laboratory to the patient's bed.

Inserm is a founding member of Aviesan, the French National Alliance for Life Sciences and Health.

Graphic design, layout: Myriem BELKACEM, Disc Inserm – Image of frontcover from vitstudio - Fotolia. Editorial secretary: Marie-Charlotte FERRAN, Annie METAIS, Disc Inserm. February 2018