PROTOCOL SYNOPSIS

1. Full title	Genetic & Environmental Determinants Of Immune Phenotype Variance: A Longitudinal Assessment
2. Acronyme or short title	MI-V3
3. Study design	The study is a single center interventional study without investigational product
4. Scientific leaders	Darragh DUFFY and Lluis QUINTANA-MURCI
5. Research rational	Susceptibility to infections, disease severity, and response to medical therapies and vaccines are highly variable from one individual to another. Individual heterogeneity in the immune response can have an enormous impact on the likelihood to respond to therapy or the development of side effects secondary to vaccine administration. Because of the complexity of immune responses in the individual and within the population, it has not been possible thus far to define the parameters (genetic or environmental) that constitute a healthy immune system and its natural occurring variability. The overall aim of the *Milieu Intérieur* study is to define the parameters that characterise a healthy immune response and its natural variation across individuals, and in doing so, inform clinical strategies for managing disease. To achieve this, in 2012, a total of 1000 healthy volunteers, descendants of mainland French persons for at least three generations, split equally by sex (1:1 sex ratio) and stratified across five-decades of life were recruited. With the first data collection (clinical study), we were able to gain key fundamental insights. The project has also established a rich sample and data repository, supporting ongoing integrative research in systems immunology and personalized medicine. To provide a longitudinal assessment of the immune variation we observed in the *Milieu Intérieur* cohort, we aim to perform a follow-up study of the cohort, 10 years after the initial study.
6. Principal Objective	To determine the effects of aging, genetics and environmental exposures on immune variation in a previously well-defined healthy population
7. Secondary Objectives	To identify *how* aging, genetics and environmental exposures affect immune variability To assess health impacts occurring in a previously well-defined healthy population

	To extend an existing bio-collection (ID RCB : 2018-A00861-54) with longitudinal follow-up samples
8. Primary evaluation criterion	Identify factors (genetic, epigenetic and environmental) that contribute to the observed heterogeneity in immune responses at the individual and population level over time: • To assess changes in circulating immune cell populations and proteins. • To assess changes in induced proteomic and transcriptomic immune responses after stimulation by pattern-recognition receptors agonists (PRR agonists) or immune stimulators. • To assess changes in the commensal microbiota (nasal swab and stools samples). • To assess changes in blood DNA methylation since the initial study. • To assess molecular and cellular markers of aging.
9. Secondary evaluation criteria	• To determine functional mechanisms behind genotype-to-phenotype associations. • To assess health impacts occurring in a previously well-defined healthy population • To provide well-characterized controls to disease cohorts.
10. Study population	Subjects recruited in the *Milieu Intérieur* study in 2012
11. Number of subjects or biological samples Justification of the number	We recruited a cohort of 1000 healthy subjects for the *Milieu Intérieur* in 2012 through the CRO Biotrial. According to Biotrial, 874 subjects can be contacted from this cohort. All contactacble subjects (874) that took part in the *Milieu Intérieur* study will be asked to participate. For sufficient statistical power for genetic analysis we will recruit a minimum of 200 subjects. We aim to recruit as many donors as possible to maximize statistical power.
12. Inclusion criteria	• Any individual who provided samples for the initial *Milieu Intérieur* study • Ability to give their informed consent in writing • Subjects who, according to the investigator, can and will comply with the requirements of the protocol. • Affiliated to the French social security or assimilated regimens
13. Non-inclusion criteria	1. Individuals benefiting from a legal protection measure 2. Individuals unable to provide expressed informed consent for participation. 3. Pregnant and breastfeeding women
14. Biological samples to be collected, Nature, amount and routing circuit	• 100 mL of blood: immune cell phenotyping, DNA extraction, protein assessment, functional immune assays, transcriptomic and epigenetic analysis, single cell RNA-Seq • Nasal swab: 16S sequencing, cytokine assays, antibody assays • Fecal samples: shotgun sequencing, mass spectometry

15. Main personal data to be collected,	Demographic data, medical history, vaccination history, socio-economic status, health-related habits, genetic profiling.
16. Data collection, transfer, record and management	Data Collection – Data will be collected using REDCap Data Transfer – Data will be stored in a secure Institut Pasteur server, under controlled access. Recording - Each donor has a unique numeric ID, that is pseudo-anonymised. Each sample has a specific unique barcoded number. Information on the samples is contained in a LIMS system, a specialized software for managing biological sample collections. For each data generation protocol, a specific sample tracking protocol is implemented to prevent potential data labelling errors. Each protocol is verified by at least 2 members of the MI team before implementation. Management - Upon data generation, all data files are stored in the *Milieu Intérieur* project space at Institut Pasteur, under password protection access. Following primary data analysis and cleaning, and verification by at least 2 MI team members "clean" analyzed data files are uploaded to the *Milieu Intérieur* LabKey data warehouse by MI data managers, who are the only members with such access. From LabKey, MI consortium members have access to all data sets under agreed consortium terms of usage.
17. Patient participation, visits duration	1 visit, 3 hours
18. Subject compensation	250 € for all subjects. A travel allowance will be provided to subjects based on the distance travelled (0,50 €/km travelled).
19. Statistical analysis	Data analysis steps are specific to each type of the generated data, due to the heterogenous nature of the data. However, general steps include initial data quality control checks, removal of low-quality samples, normalization procedures, imputation steps to replace missing values, and data visualisation using dimensional reduction approaches. These steps will be performed either by the scientist responsible for generation of the data or the MI team. For data integration approaches (eg genetic association studies), association between diverse sets of factors will be tested using specific methods in causal inference and machine learning.
20. Study Calendar	▪ Provisional starting date of inclusion: January 2022 ▪ Patient participation duration: 1 visit, 3 hours. 1 week for participants to send fecal samples. ▪ Study provisional duration: 10 months for inclusions, 20 years for analyses. ▪ Biobanks duration: 20 years after the last visit and 2 years after the last publication

21. Number of centres	1 centre study, Biotrial, France