Department of Services Area
SC Transfusion Medicine -Director: Daniele Prati
Tel. 02.5503.4687
AND-email:centrotrasfusionale@policlinico.mi.it

DEFINING THE GENOMIC LANDSCAPE OF MASLD

(Title of the study in English "DEFINING THE GENOMIC LANDSCAPE OF METABOLIC STEATOTIC LIVER DISEASE")

Acronym: DETECTIVE

Protocol version number: v.2.0

Date: 08/18/2025

Promoter:IRCCS Ca' Granda Foundation Maggiore Hospital Polyclinic,
Via Sforza 28, 20122 Milan, Italy

Coordinating Center:SC Transfusion Medicine
IRCCS Ca' Granda Foundation Maggiore Hospital Polyclinic,
Via Sforza 35, 20122 Milan, Italy

Principal Investigator: Luca Vittorio Carlo Valenti

FLOW CHART

	Screening	Intervention
Period	(-t1)	(t1)
Enlistment		
informed consent	x	
inclusion/exclusion criteria	x	
Administration of the intervention		
biological sample collection *		x
phenotypic data collection **		x

(*) the biological sample has already been previously collected in the context of the SERENA, REASON or LIVER-BIBLE studies

(**) These are data previously collected in the context of the SERENA, REASON or LIVER-BIBLE studies

LIST OF ABBREVIATIONS

THERE IS: Ethics Committee

THERE: Informed Consent

CRF: Case Report Form, data collection form

GCP: Good Clinical Practice, good clinical practice

WGS: Whole Genome Sequencing

MASLD: Metabolic dysfunction-associated steatotic liver disease

HCCHepatocellular Carcinoma

LoFLoss of function

CHIP: Clonal hematopoiesis of indeterminate potential

PRSPolygenic risk score

AST: Aspartate aminotransferase

ALT: Alanine aminotransferase

GGTGamma-glutamyl transferase

RESPONSIBILITY (role of the promoter and collaborators)

Sponsor: IRCCS Ca' Granda Foundation, Ospedale Maggiore Policlinico, Via Sforza 28, 20122 Milan, Italy

Coordinating Center: Transfusion Medicine Department, IRCCS Ca' Granda Foundation, Ospedale Maggiore Policlinico.

The Principal Investigator (PI) for this study is Professor Luca Vittorio Carlo Valenti, Head of the Biological Resources Center and the Precision Medicine Unit, Department of Transfusion Medicine, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan. As PI, he will be responsible for patient enrollment and study coordination.

Internal collaborations

Structure	*Participant name*	*Role and functions in the study*

SC Transfusion Medicine, IRCCS Ca' Granda Foundation Maggiore Polyclinic Hospital	*Dr. Daniele Prati* *Dr. Alessandro Cherubini* *Dr. Giulia Periti* *Dr. Serena Pelusi* *Dr. Sara Margarita* *Dr. Elia Casirati* *Dr. Francesco Malvestiti* *Dr. Stefania Mira* *Dr. Francesca Iemma*	*Statistical analysis of data from Whole Genome Sequencing*

External collaborations (biological sample analysis, data analysis, diagnostic procedures, etc.)

Institution	*Operational unit*	*Role and functions in the study*
Human Technopole	*National Facility for Genomics*	*Whole Genome Sequencing (WGS)*

INDEX

1. INTRODUCTION

1.1 Background and rationale

Metabolic Dysfunction-Associated Steatotic Liver Disease (MASLD) is a leading cause of chronic liver disease globally, with a prevalence exceeding 30% in the population. MASLD is closely associated with insulin resistance and cardiometabolic conditions and, in 20-30% of cases, can progress to steatohepatitis (MASH), which is characterized by progressive liver damage and inflammation [1]. In patients at highest risk, the disease can lead to the onset of advanced fibrosis, cirrhosis and hepatocellular carcinoma (HCC). One of the main problems in the clinical management of MASLD is the absence of specific risk biomarkers and the lack of effective treatments, especially for patients with advanced stage disease [2]. MASLD has a large and well-documented genetic component, with studies identifying several common variants associated with this pathology, such as those in the PNPLA3, TM6SF2 and MBOAT7 genes [3]. However, these variants identified so far explain only a small part of the heritability of MASLD, also suggesting the contribution of rare loss-of-function (LoF) variants [4-5]. Furthermore, scientific evidence indicates that the accumulation of somatic variants, both in hepatocytes and myeloid cells, could also play a key role in the progression of MASLD [6]. In particular, clonal hematopoiesis of indeterminate potential (CHIP), which is a condition characterized by the presence of hematopoietic clones with somatic mutations often associated with leukemia and cardiovascular diseases [7], could favor the onset of hepatocellular carcinoma [8]. However, the evidence available to date is still limited and requires further investigation and studies on larger cohorts.

The study in question therefore aims to further investigate this aspect by analyzing the genetic profile using a Whole-Genome Sequencing (WGS) approach; peripheral blood DNA samples from patients with advanced MASLD and peripheral blood DNA samples from controls presenting various associated metabolic risk factors will be sequenced.

Additionally, 80 liver tissue samples from patients with advanced MASLD will be sequenced to identify specific somatic mutations. Expected outcomes of this study include the identification of novel genetic variants associated with MASLD progression, improved risk stratification through the development of polygenic risk scores, and the identification of potential therapeutic targets. This

study represents a fundamental step in understanding the biology of MASLD and could have important clinical implications for disease management.

2. OBJECTIVE/S/HYPOTHESIS OF THE EXPERIMENTATION

A careful review of the scientific literature shows that MASLD has a strong genetic component, but the evidence available to date only partially explains the predisposition to the advanced form of the disease. Furthermore, it is known that somatic variants, particularly clonal hematopoiesis of indeterminate potential (CHIP), may contribute to the progression of MASLD to cirrhosis and hepatocellular carcinoma (HCC).

2.1 Primary objective

The primary objective of this study is toto identify genetic variants (germline and/or somatic) associated with the risk of developing advanced MASLD (fibrosis ≥2 and/or hepatocellular carcinoma) in the Italian population.

2.2 Secondary objectives

The secondary objectives of this study are:

- To characterize the contribution of somatic variants (in hepatocytes and hematopoietic cells) that are involved in the progression of advanced MASLD;
- Evaluate and develop polygenic risk scores (PRS) for risk stratification and identification of specific subphenotypes;
- To study the associations between genetic variants and metabolic phenotypes related to MASLD;
- Identify relevant therapeutic targets for the development of precision medicine strategies in advanced MASLD.

3. STUDY DESIGN

3.1 Study design

The proposed study is a single-center, national, non-pharmacological interventional study with genetic analysis and is non-profit.

The total number of samples for which Whole Genome Sequencing (WGS) is planned is 2880, distributed as follows:

- 800 peripheral DNA samples from patients with advanced MASLD attending the Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico and collected as part of the SERENA study (study promoted by the Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, approved by the CE Milano Area2 on 07.26.2017);
- 2000 peripheral DNA samples from controls presenting metabolic risk factors collected as part of the Liver-BIBLE study (a study promoted by the Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, approved by the CE Milano Area2 on 01.07.2020);
- 80 pathological liver tissue samples isolated from patients with advanced MASLD in the context of the REASON clinical study (study promoted by the Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, approved by the CE Milano Area2 on 07.07.2021);

3.2 Inclusion criteria

Specific inclusion criteria for patients with advanced MASLD:

- Patients with advanced MASLD defined as liver fibrosis ≥2 and/or development of HCC;
- patients enrolled in the SERENA study and where applicable also in the REASON study;
- Liver biopsy for suspected non-alcoholic steatohepatitis (NASH) at the time of diagnosis;
- Cholecstectomies
- age [40-70 years]
- patients who have signed the informed consent

Specific inclusion criteria for control group:

Blood donors from the LIVER-BIBLE study aged between 40 and 70 years in the presence of overweight or obesity (Body mass index-BMI>25Kg/m2) and at least two of the following risk factors:

- impaired fasting blood glucose or diabetes mellitus (blood glucose ≥100 mg/dl)

- dyslipidemia (triglycerides≥150mg/dl, HDL<45/55 in M/F)
- arterial hypertension

3.3 Exclusion criteria

Specific exclusion criteria for patients with advanced MASLD:

- Positivity for chronic viral hepatitis (HCV-RNA and/or HBsAg);
- Positivity for other liver diseases such as autoimmune and viral hepatitis (hepatitis B and C), hereditary hemochromatosis, alpha-1-antitrypsin deficiency, Wilson's disease.

Specific exclusion criteria for the control group:

Subjects suffering from chronic degenerative diseases will be excluded, with the exception of well-controlled hypertension and type 2 diabetes mellitus not requiring pharmacological therapy (as is already common practice for eligibility to donate blood) and donors aged > 65 and < 40.

4. PROCEDURES RELATING TO THE STUDY

4.1 Patient recruitment

Patient enrollment will take place at the outpatient clinics of the Transfusion Medicine Unit of the Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico during routinely scheduled visits.

If obtaining consent for the clinical study is not compatible with the conduct of the study, consent will be obtained remotely via telephone, as described in chapter 9.2.3.

In all other cases, the informed consent process will be managed in person.

Patients who enroll in the DETECTIVE study will not undergo any additional procedures, but will consent to the use of data and biological samples previously collected in the SERENA, REASON, and Liver-Bible studies.

4.2 Sample and data collection

Our research is based on two fundamental pillars, represented by rigorously selected study cohorts.

The SERENA project involves 800 patients with advanced MASLD, whose peripheral blood DNA will be sequenced (WGS, 20x coverage). To further deepen our understanding of disease progression, a subset of 35 of these patients will undergo even more detailed analysis: we will sequence at 50x 80 pathological liver tissue samples from the REASON study, including samples with advanced fibrosis and, where present, hepatocellular carcinoma (HCC).

On the control side, the Liver-BIBLE Milano project offers us a valuable cohort of 2,000 individuals matched by sex and metabolic risk factors. These participants are selected from blood donors aged 40 to 70 years who meet at least two criteria for metabolic syndrome. Their phenotypic characterization is extremely thorough, ranging from anthropometric assessment to metabolic and inflammatory profiles, from noninvasive liver injury diagnostics to OMICs analysis. This integrated approach, combining genetic and clinical data from well-defined cohorts, will allow us to conduct robust comparisons between cases and controls, significantly advancing our understanding of advanced MASLD.

Specifically, biological samples will be collected as follows:

- Peripheral Blood:Peripheral blood DNA samples will be collected from all 800 MASLD patients and all 2000 metabolic controls.
- Pathological Liver Tissue:A total of 80 pathological liver tissue samples will be collected from a subset of 35 patients with advanced MASLD. These samples will be derived from the REASON study and will include tissue with advanced fibrosis (n=30) and, when present, hepatocellular carcinoma (HCC) samples (n=12).

Clinical and phenotypic data associated with the biological samples collected:

- For Patients with Advanced MASLD, aimed at understanding the progression of the disease, its complications and related factors**:**

- Detailed Clinical Phenotype: Includes a detailed medical history focused on the history of MASLD, including diagnosis, progression, presence and severity of specific symptoms, history of pharmacological or surgical treatments, and relevant comorbidities, particularly metabolic and cardiovascular ones.
- Biomarkers of Liver Injury and Function: Liver enzymes such as alanine aminotransferase (ALT), aspartate aminotransferase (AST), gamma-glutamyl transpeptidase (GGT), and alkaline phosphatase (ALP) levels, indicators of hepatocellular injury or cholestasis.
- Liver synthetic function: Bilirubin, albumin, prothrombin time/INR, which reflect the liver's synthetic capacity.
- Non-invasive biomarkers of liver damage: data from tools such as Fibroscan or predictive score calculations (e.g., FIB-4, NAFLD Fibrosis Score) to assess the presence and degree of steatosis, steatohepatitis (NASH), fibrosis, and cirrhosis.
- Metabolic Risk Factors: such as data on glucose, lipids (total cholesterol, HDL, LDL, triglycerides), insulin, HbA1c and indices of insulin resistance (e.g. HOMA-IR), which are essential for outlining the metabolic profile underlying MASLD.
- Inflammatory Biomarkers: Such as C-reactive protein (CRP) and ferritin, useful for assessing systemic and/or hepatic inflammatory status.
- Hepatocellular Carcinoma (HCC) Data: For patients who have developed HCC, specific information on diagnosis, staging, oncological treatments and follow-up is collected, essential for studying clonal evolution.

- For Metabolic Controls, aimed at providing an extremely in-depth baseline characterization of their general health and metabolic risk profile, in the absence of advanced MASLD, to serve as a robust comparison:

  - Detailed Anthropometric Assessment: Height, weight, waist circumference, BMI, for a precise quantification of adiposity.
  - Detailed Medical History: Lifestyle, eating habits (diet), complete medical history, with particular attention to the presence of metabolic risk factors (hyperglycemia, hypertension, dyslipidemia, increased adiposity) which are inclusion criteria.

- Complete Blood Count: Evaluation of the cellular components of the blood.
- Coagulation Parameters: Includes factors such as FVIII, VWF, protein C (PC) and D-dimer.
- Standard Clinical Biochemistry: Includes assessments of general liver and renal function, as well as inflammatory biomarkers.
- Complete Metabolic Profiles: Glucose, lipids (total cholesterol, HDL, LDL, triglycerides), insulin and HbA1c, to quantify metabolic risk factors.
- Liver Biomarkers: Although not targeted to progression of advanced MASLD, baseline biomarkers of liver function are included to ensure the absence of significant non-MASLD liver disease.
- Non-Invasive Liver Injury Assessment (Fibroscan): Essential to confirm the absence of significant liver fibrosis in these controls, although they may have steatosis or early NASH.
- OMIC characterization (in a subset): This is a distinctive feature of this cohort, which includes circulating lipidomics, metabolomics, and metagenomics. It provides a very high level of molecular detail on their metabolic profiles and gut microbiome, crucial for understanding gene-environment-metabolism interactions.
- Prospective Follow-up Assessments: The longitudinal nature of the Liver-BIBLE Milano study means that follow-up data will be updated over time, providing information on the evolution of their metabolic and health status.

4.3 Sample analysis

The genetic analyses for this study will be conducted in close collaboration and under the supervision of Human Technopole.

As anticipated, the methodology envisaged for conducting this study involves whole-genome sequencing on DNA derived from peripheral blood (20x coverage) and pathological liver tissue (50x coverage) to characterize the genetic and somatic variants that influence disease progression.

The DNA will be isolated at the Foundation and shipped to the National Facility (NF) in batches of 3-4 plates at a time.

To identify associations between individual genetic variants and the burden of likely loss-of-function (LoF) variants at specific genetic loci, comparing participants with advanced MASLD to controls, analyses will be performed using a genome-wide regression model implemented in REGENIE, with additive genetic models accounting for age, sex, body mass index (BMI), and ancestry (BWI 1-10). To ensure the robustness of the results, the genome-wide significance threshold will be set at $p<5E-8$. If the study is not powered to identify new significant associations, suggestive hits ($p<5E-6$) will be replicated in additional cohorts.

In addition to defining the risk of advanced MASLD compared to controls as the primary outcome, we will thoroughly examine the genetic contribution to various endophenotypes. These will include metabolic risk factors (such as insulin resistance indices), liver enzymes (AST, ALT, GGT), and indicators of liver damage (noninvasive predictors of steatosis, steatohepatitis, fibrosis, cirrhosis, and HCC). The results of these analyses will be crucial for improving risk stratification and identifying specific disease subphenotypes through the development of polygenic risk scores, which have previously proven effective in identifying subsets of patients at specific risk for liver or cardiovascular events. This approach will also aim to identify therapeutic targets with a high probability of clinical success.

As regards CHIP (Clonal Hematopoiesis of Indeterminate Potential) and somatic variants, their identification will take place through a semi-automated pipeline. CHIP assignment will be assessed in a panel of specific genes (including NADK, GNB1, CBL, KMT2D, RHEBL1, PPM1D, JMJD6, MFSD11, DNMT3A, SF3B1, ASXL1, NOL4L, GNAS, RUNX1, AF015262.1, RPL34P3, U2AF1, TET2, CUX1, SH2B2, BCOR, BCORL1, MCAM, TP53, TLMET23, SRSF2, NPM1P46, LINC01426, EZH2P1, F4, AT1PB2, ZNF316, ANAPC1, JAK2), while for somatic variants in hepatocytes, relevant scientific literature will be consulted. Variant identification will be performed using three callers: Mutect, Vardict, and Freebayes. The impact of inherited genetic variation on both CHIP risk and related phenotypes will also be assessed.

All unused material at the National Facility for Genomics will be returned and stored at the Fondazione IRCCS Ca Granda. Subject to the patient's specific consent for biobanking, these residues may be retained beyond the study's conclusion for future research purposes at the Foundation's POLI-MI Biobank.

5. ENDPOINT

5.1 Primary Endpoint

Identifying the association between germline variants and the presence of advanced MASLD in cases versus controls.

5.2 Secondary Endpoints

Other main endpoints of the study in question are:

1) Identification of somatic mutations in hepatocytes and hematopoietic cells in order to study their association with the development of HCC.

2) Use of PRS (existing in the literature and new) in predicting patterns of advanced MASLD among cases and controls

3) Analysis of the association between genetic variants and clinical and biological phenotypes of MASLD, considering metabolic markers, liver enzymes (AST, ALT, GGT) and non-invasive indicators of steatosis and fibrosis (Fibroscan)

4) Identification of functional variants (e.g. LoF) in genes involved in lipid metabolism, liver inflammation, oxidative stress associated with advanced MASLD.

5. DURATION / TIMELINE OF THE STUDY

Study start date: 09/2025

Enlistment End Date: December 2025

Data analysis end date: 07/2026

Final Study Report: 08/2026

6. STATISTICAL ANALYSIS

6.1 Sample size

The total number of samples for which Whole Genome Sequencing (WGS) is planned is 2,880, including 800 subjects with advanced MASLD, 2,000 controls with at least two characteristic criteria of metabolic syndrome but free of advanced liver disease, and 80 liver tissue samples for somatic variant analysis. The sample size estimate is based on the study's primary objective, which is the identification of genetic variants (germline and/or somatic) associated with MASLD.

The study is descriptive with a power proportional to the number of subjects included.
This number was defined on the basis of the maximum availability of subjects that could be enrolled in the clinical cohorts with the aim of maximizing the statistical power and scientific value of the results.Preliminary data on common variants are available that confirm the statistical power of the cohort to identify phenotype-associated variants at the genome-wide level..
To confirm the adequate statistical power of the study, an analysis was performed using the R statistical analysis software package called “genpwr” by setting the error at a significance level of 5x10−8 conventionally used to define significance at the genomic level (https://cran.r-project.org/web/packages/genpwr/index.html) and the number of cases and controls in the detective cohort. This method is based on an extension of that reported by Gauderman (2002) (doi:10.1093/aje/155.5.478) and Gauderman (2002) (doi:10.1002/sim.973) and described in: Moore CM, Jacobson S, Fingerlin TE. Power and Sample Size Calculations for Genetic Association Studies in the Presence of Genetic Model Misspecification. American Society of Human Genetics. October 2018, San Diego. For common variants, combinations with variant allelic frequency (MAF) ≥ 0.1 and OR ≥ 2.0 were considered, while for low-frequency and rare variants with a greater impact on protein function, for which the focus is on identifying variants with a greater impact on disease risk, MAF values between 0.01 and 0.05 and ORs up to 15 were explored. The results confirm that based on the expected number of 800 cases and 2000 new controls (without considering the 1000 controls already available) the study will have a power greater than 95% to detect ORs > 2 for common variants with a frequency of 0.1 (power >80% for OR 1.3 for variants even with a frequency of 0.1, and much greater power for even more frequent variants, Graph 1). Furthermore, for low-frequency and rare variants, it will have a

power greater than 80% to detect ORs already > 5 for variants with MAF 0.01 that are difficult to type with array-based methods (see Chart 2), although variants with ORs > 8 are typically expected in this case, for which it will be possible to identify rare variants typeable only with WGS approaches, on which the DETECTIVE study is focused. Furthermore, a preliminary GWAS analysis that considered only common variants typed with Illumina GSA3.0 arrays using fibrosing steatohepatitis defined by Fibrotic NASH Index (FNI) as the main outcome (Chart 3), conducted in a part of the cohort that will be the subject of the study (N=2300) not only confirmed the impact of the two major loci for steatohepatitis already known in the literature (PNPLA3 and TM6SF2), but also highlighted the presence of two new loci not yet described, further strengthening the robustness of the proposed study.

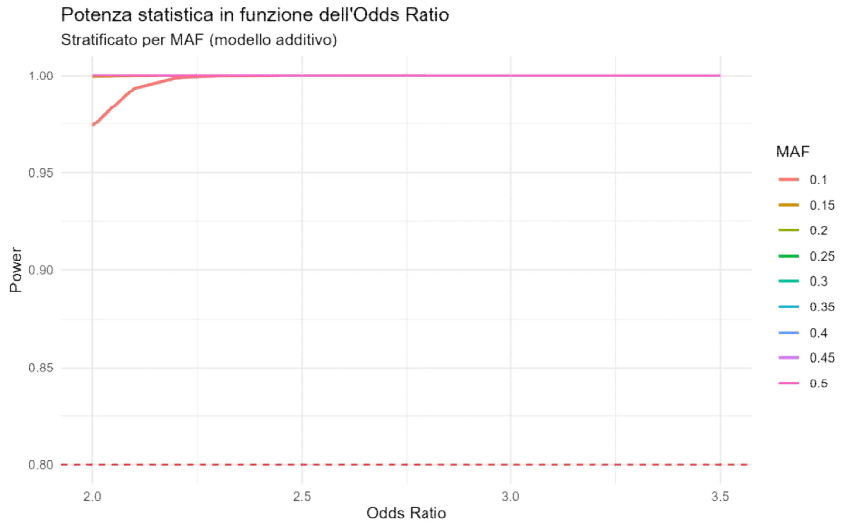


Figure 1. Statistical power to identify significant genomewide associations as a function of the OR for different minor allelic frequencies (MAF) of common variants in the Detective cohort

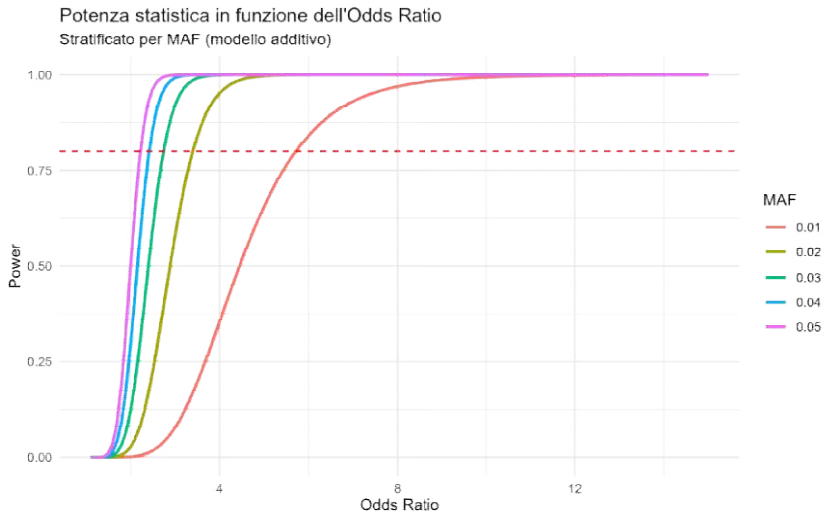
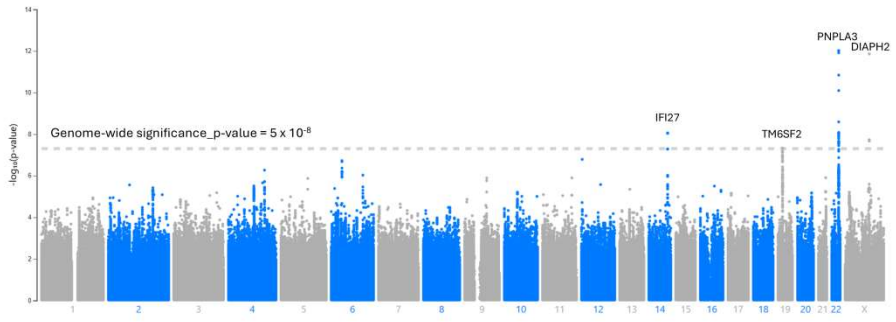


Figure 2. Statistical power to identify significant genomewide associations as a function of the OR for different MAFs of low-frequency/rare variants in the Detective cohort.



Graph 3. Manhattan plot showing the preliminary results of a genomewide association study (GWAS) for common variants in a first part of the Detective cohort (FOGS Study) using as an outcome the risk of fibrosing steatohepatitis assessed by FNI score.The results were obtained using genomic regression models adjusted for age, sex, body mass index (BMI), presence of diabetes and ethnic background (assessed by principal components PC 1-10).

6.2 Data analysis

The statistical analysis of this study will be conducted in collaboration with Human Technopole. The primary analysis will follow a rigorous approach based on genome-wide regression models between germline variants and the presence of advanced MASLD. REGENIE software will be used, which allows for the evaluation of large-scale whole-genome regression in large genome-wide association studies. The secondary analysis involves the development and risk stratification by constructing polygenic risk scores specific to advanced MASLD and its subphenotypes.
The statistical analyses of the study will be conducted using R and Python software.

7. ADVERSE EVENTS

The project does not involve the administration of drugs or other substances, nor does it involve invasive clinical procedures. Therefore, no adverse events are expected.

8. RISK/BENEFIT ASSESSMENT

This study does not anticipate an immediate benefit to participants, but the results will be fundamental for the identification of specific therapeutic targets and for the development of new personalized medicine strategies.

9. STUDIO MANAGEMENT

9.1 Data collection and management

Each participant will be assigned a unique code upon enrollment. Data de-identification will be performed in such a way that people accessing the database will not be able to trace the subjects' identities in any way. Only local investigators will be able to trace the identities of enrolled subjects. The data required for the study will be recorded in a dedicated eCRF in a Data Management System validated according to national regulations, provided by the Foundation's Scientific Directorate. The platform used will be RedCap (Research Electronic Data Capture).
The REDCap Consortium is composed of >1,000 institutional partners worldwide (research institutions, universities, ministries, etc.). The consortium supports a secure web application

(REDCap) designed exclusively to support data acquisition for research studies. The REDCap application allows users to quickly and securely create and manage online databases and is currently in use for more than 110,000 projects with approximately 150,000 users covering numerous research areas across the consortium.

Through REDCap, this study will implement: a) user-level identification, with specific restrictions based on role in the study b) real-time data validation and integrity checking c) patient de-identification before data export d) centralized data storage with daily backup, on a secure server within the Foundation's IT infrastructure.

9.2 Regulatory aspects and ethical considerations

9.2.1 Approval by the Competent Authority

In accordance with applicable regulations, the principal investigator must obtain approval from the appropriate Competent Authority before initiating the clinical study.

This study will be conducted in accordance with ICH/GCP (International Conference on Harmonization/Good Clinical Practice) regulations and all applicable laws, including the Helsinki Declaration of June 1964, as amended by the last World Medical Association General Assembly in Seoul, 2008.

9.2.2 Ethics Committee Approval

The investigator must ensure that the protocol has been reviewed and approved by the local independent Ethics Committee (EC) before starting the study.

The CE must also review and approve the informed consent (IC) form and all written information received from the patient prior to enrollment in the study.

Should it be necessary to modify the protocol and/or the IC during the study, the investigator will be the guarantor and therefore the person responsible for ensuring the review and approval of such modified document as requested by the CE.

The content of these changes will be implemented only after the EC has approved them. Until then, it will be necessary to refer to the previous version of the already approved document.

9.2.3 Informed consent (IC)

The investigator or other personnel designated by him are responsible for informing the subjects about all aspects and procedures of the study.

Informed consent will be collected by the doctor during the visit.

The informed consent process must comply with applicable regulatory procedures. The investigator (or designated staff member) and the subject must date and sign the informed consent form before the patient begins any study-related procedures. The subject will receive a copy of the informed consent, dated and signed by both parties; the original copy will be retained in designated study archives. Neither the investigator nor designated staff member may in any way coerce or influence a subject to participate or continue participating in the study. A subject's decision to participate in the study must be completely voluntary. The investigator and designated staff member must emphasize to the subject that they may withdraw their consent at any time without penalty or loss of any benefits to which they may be entitled.

Written or oral information relating to the study, including the written consent form, must not contain any language that forces the subject to waive (even apparently) his or her legal rights, or that would exonerate the investigator, institution, or sponsor from liability for negligence.

If participant access to the Foundation is not foreseeable within a timeframe compatible with the conduct of the study, given that this study requires the use of biological samples previously biobanked at the Foundation (patients have consented to the preservation of biological material for future research activities), consent to participate in the study, and the related data processing, may be obtained remotely via telephone.

During the telephone call, the Investigator, or a delegate, will provide the participant with information about the study and data processing methods, answer any questions, and verify the email address to which study documentation should be sent. The telephone call will be tracked using the document "INSTRUCTIONS AND FORM FOR ACQUIRING CONSENT BY

TELEPHONE CALL - PARTICIPATION IN CLINICAL TRIALS," which will be archived with the study documentation.

9.3 Duties of the investigator

In accordance with applicable local regulations, the investigator must submit periodic reports regarding the progress of the study at his/her site to the CE and notify the CE of study closure. Periodic reports and closure notification are the investigator's responsibilities.

9.4 Study monitoring

In accordance with applicable regulations and Good Clinical Practice (GCP), the monitor must periodically visit or contact the center. The duration, nature, and frequency of these visits/contacts depend on the recruitment rate, the quality of the center's records, and their adherence to the protocol.

Through these contacts, the monitor must:
• monitor and evaluate the progress of the study
• examine the collected data
• conduct the verification of the source document
• identify each problem and related solutions

The purposes of the monitoring activity are to verify that:
• the rights and well-being of the subject are respected
• the study data are accurate, complete and verifiable from the original documents
• the study is conducted in accordance with the protocol and any approved amendments, GCP and applicable regulations

The experimenter must:
• give the monitor direct access to all relevant documentation
• dedicate part of his time and his staff to the monitor to discuss the monitoring results and any other possible aspects.

The monitor must also contact the center prior to the start of the study to discuss the protocol and data collection procedures with the staff.

9.5 Study quality assurance

As the sponsor, Fondazione IRCCS Ca' Granda, Ospedale Maggiore Policlinico may conduct a quality control review of the study at its discretion. In this case, the investigator must allow the monitor direct access to all relevant documentation and dedicate part of their time and staff to the reviewer to discuss the monitoring results and any other aspects of the study.

Additionally, Regulatory Authorities may conduct inspections. In this case, the investigator must grant the inspector direct access to all relevant documentation and dedicate part of their time and personnel to the inspector to discuss the monitoring results and any other aspects of the study.

9.6 Closure of the study

Upon study closure, the monitor and the investigator must activate a series of procedures:

• review all study documentation

• reconcile study data

• reconcile all clarifying reports.

9.7 Document storage

In accordance with current national regulations, the investigator must retain a copy of all documentation and store it in a dry, safe place after the study has been closed.

9.8 Disclosure of information regarding scientific discovery

9.8.1 Confidentiality

The investigator and other personnel involved in the study must maintain all information related to the study (including the protocol, data obtained, and all documentation produced during the study) and must not use such information, data, or reports for purposes other than those

described in the protocol. These restrictions do not apply to:

1) information that becomes publicly available, not due to negligence on the part of the investigator or his staff;

2) information requiring confidential disclosure to CE for the sole purpose of evaluating the study;

3) information that must be disclosed in order to obtain appropriate medical care for a study subject.

9.8.2 Publications and Intellectual Property Rights on the Results of the Study

a) SINGLE-CENTER STUDY which includes collaborations:

Publications:

As the Promoter, the Foundation will ensure the dissemination and publication of the study results, even in the event of negative results, without any restrictions and guaranteeing the collaborating center visibility proportional to its actual participation.

Each scientific journal or publication containing the study results and data must indicate the role and participation of the collaborating center and the Foundation, in proportion to their actual contribution to the study and their role. The data may be published in aggregate form or otherwise anonymized, so as not to allow the identification of the data subject.

Intellectual Property Rights:

The Parties acknowledge that, in conducting the collaboration within the scope of the study, data, information, know-how, and inventions (whether patentable or not) owned by each Party may be used and shared. Each Party retains exclusive ownership of the same, even if it grants the other a non-exclusive and free right of access and use, solely for the purposes of carrying out the activities

covered by the study and limited to the duration of the study. It is understood that this right of use does not include the right to sublicense to third parties.

In accordance with current legislation, the data and results generated within the Study will be the property of the Sponsor, unless specific agreements are made between the Sponsor and the collaborating center(s).

10. COMPENSATION AND DAMAGES IN CASE OF DAMAGES

In the event of any undesirable events or damages arising from participation in research, our Institute's insurance policy also extends to cover the subjects participating in research projects.

11. FINANCIAL AGREEMENTS

Study procedures exceeding normal clinical practice described in this study will be covered by funding from the HUMAN TECHNOPOLE NATIONAL FACILITY FOR GENOMICS CALL FOR ACCESS 24-G-PILOT.

12. DISCLOSURE ON CONFLICTS OF INTEREST

The experimenters declare no conflicts of interest.13. REFERENCES

1. Younossi ZM, Golabi P, Paik JM, Henry A, Van Dongen C, Henry L. The global epidemiology of nonalcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH): a systematic review. Hepatology 2023;77:1335–47.https://doi.org/10.1097/HEP.0000000000000004

2. European Association for the Study of the Liver (EASL). Electronic address: easloffice@easloffice.eu , European Association for the Study of Diabetes (EASD), European Association for the Study of Obesity (EASO), European Association for the Study of the Liver (EASL). EASL-EASD-EASO Clinical Practice Guidelines on the management of metabolic dysfunction-associated steatotic liver disease (MASLD). J Hepatol 2024:S0168-8278(24)00329-5.https://doi.org/10.1016/j.jhep.2024.04.031.

3. Dongiovanni P, Stender S, Pietrelli A, Mancina RM, Cespiati A, Petta S, et al. Causal relationship of hepatic fat with liver damage and insulin resistance in nonalcoholic fatty liver. J Intern Med 2018;283:356–70.https://doi.org/10.1111/joim.12719

4. Pelusi S, Baselli G, Pietrelli A, Dongiovanni P, Donati B, McCain MV, et al. Rare Pathogenic Variants Predispose to Hepatocellular Carcinoma in Nonalcoholic Fatty Liver Disease. Sci Rep 2019;9:3682.https://doi.org/10.1038/s41598-019-39998-2.

5. Verweij N, Haas ME, Nielsen JB, Sosina OA, Kim M, Akbari P, et al. Germline Mutations in CIDEB and Protection against Liver Disease. N Engl J Med 2022;387:332–44.https://doi.org/10.1056/NEJMoa2117872.

6. Ng SWK, Rouhani FJ, Brunner SF, Brzozowska N, Aitken SJ, Yang M, et al. Convergent somatic mutations in metabolic genes in chronic liver disease. Nature 2021;598:473–8.https://doi.org/10.1038/s41586-021-03974-6

7. Marnell CS, Bick A, Natarajan P. Clonal hematopoiesis of indeterminate potential (CHIP): Linking somatic mutations, hematopoiesis, chronic inflammation and cardiovascular disease. J Mol Cell Cardiol. 2021 Dec;161:98-105. doi: 10.1016/j.yjmcc.2021.07.004. Epub 2021 Jul 21. PMID: 34298011; PMCID: PMC8629838.

8. Marchetti A, Pelusi S, Marella A, Malvestiti F, Ricchiuti A, Ronzoni L, et al. Impact of clonal hematopoiesis of indeterminate potential on hepatocellular carcinoma in individuals with steatotic liver disease. Hepatology 2024. https://doi.org/10.1097/HEP.0000000000000839