TÍTULO DEL PROYECTO:

Caracterización del paisaje germinal y somático de las neoplasias mieloides asociadas a tratamiento (TRMN)

Somatic and germinal landscape of therapy related myeloid neoplasms

INVESTIGADOR PRINCIPAL: FRANCESC SOLÉ RISTOL

CENTRO: INSTITUT DE RECERCA CONTRA LA LEUCÈMIA JOSEP CARRERAS, BADALONA.

2024-2027

RESUMEN

Los síndromes mielodisplásicos o neoplasias mielodisplasicas (abreviado SMD) pueden ser de novo o secundarios a un cáncer previo. Actualmente se estima que 3 de cada 5 personas desarrollarán un cáncer a lo largo de su vida. Debido a los avances en medicina, cada vez hay más pacientes con cáncer que consiguen una curación. Sin embargo, los pacientes diagnosticados de cáncer y tratados con quimioterapia y/o radioterapia tienen un mayor riesgo de desarrollar neoplasias mieloides secundarias relacionadas con el tratamiento (*therapy related myeloid neoplasms*, TRMN), que son neoplasias hematológicas clonales muy agresivas y de mal pronóstico que afectan al linaje mieloide.

Asimismo, la predisposición genética podría estudiarse también para los SMD primarios, principalmente en aquellos casos en los que existe una edad temprana de aparición de la enfermedad o agregación familiar de neoplasias mieloides. Estos pacientes conforman la entidad de SMD con predisposición germinal (SMDg), que viene determinada por la presencia de mutaciones germinales en genes que predisponen al desarrollo de SMD.

Por ello, se propone el estudio a nivel somático y germinal de una serie de 50 pacientes con TRMN. Además, se evaluará también la prevalencia de la predisposición genética a SMD (SMDg) en 25 pacientes menores de 50 años de edad con SMD, con el objetivo de explorar los mecanismos patogénicos que desencadenan los diferentes escenarios de SMD associados a prediposición genética, así como proporcionar un diagnóstico preciso y diferencial. Los estudios del transcriptoma y el análisis (cito)genético mediante WES y Optical Genome Mapping en muestras de médula ósea de los pacientes con TRMN y SMDg, permitirán la determinación de las vías moleculares claves implicadas en la tumorogenesis.

En este proyecto se propone el estudio a nivel somático y germinal de una serie de 50 pacientes con TRMN para explorar los mecanismos patogénicos implicados en el desarrollo de esta enfermedad, así como evaluar la posible relación entre la progresión de los TRMN y la presencia previa CHIP o predisposición germinal. Esto permitirá desarrollar métodos adecuados para el diagnóstico precoz y diferencial de los diferentes escenarios de TRMN, y guiar las terapias en la práctica clínica de los tumores primarios en pacientes con riesgo de desarrollar un TRMN. Además, se evaluará la prevalencia de la predisposición genética a SMD (SMDg) en 25 pacientes menores de 50 años de edad con SMD, con el objetivo de explorar los mecanismos patogénicos que desencadenan los diferentes escenarios de SMD associados a prediposición genética, así como proporcionar un diagnóstico preciso y diferencial. Para ello, se llevarán a cabo estudios del transcriptoma y el análisis (cito)genético mediante secuenciación de exoma (WES) y Optical Genome Mapping en muestras de médula ósea de los pacientes con TRMN y SMDg con el fin de determinar las vías moleculares clave implicadas en la tumorogenesis.

OBJECTIVES:

In this proposal we aim to explore the different pathogenic mechanisms that drive TRMN, and gMDS scenarios and, in view of the background exposed above, we propose the following objectives:

Objective 1. Evaluation of the genetic landscape of TRMN and gMDS patients by WES

> **TRMN scenario:** (i) Evaluation of the somatic burden of mutations in the TRMN context in comparison with primary MDS to identify differences between both diseases, (ii) Description of the somatic landscape of TRMN to understand the association with primary tumor and treatment and (iii) Identification of germline predisposition prevalence in TRMN patients to complete the biomarker spectra for the early and differential diagnosis, treatment guidance and monitoring of TRMN-like patients.
>
> **MDS younger than 50 years old, gMDS scenario:** (i) Study of the genetic predisposition prevalence in candidate gMDS patients (ii) Evaluation of the mutational profile in those patients where mutations in none of the described genes linked to gMDS have been found for the identification of new potential driver genes, (iii) Clusterization of molecular alterations and signatures of gMDS for subtype stratification and comparison with primary MDS (WES from the host group) for their differential diagnosis and risk stratification.

Objective 2. Transcriptome studies by RNA-seq: (i) Differential expression analysis between MDS younger than 50 years old and TRMN samples with and without germline mutations, to determine which genes show a differential expression (upregulated or dowregulated), (ii) Functional enrichment analysis to determine if the differentially expressed genes are enriched/overrepresented in a particular metabolic pathway or cellular function, (iii) Integration of transcriptomic and genetic data to identify the key molecular driver pathways involved in each scenario and discover novel hub gens that promote tumorigenicity.

Objective 3. Detection of cytogenetic alterations by Optical Genome Mapping (OGM) technology: (i) Study of TRMN patients and MDS younger than 50 years old to precisely define structural alterations breakpoints and numeric alterations, (ii) Reevaluation of cytogenetic findings by OGM for a more precise prognostic stratification of patients and, (iii) Cost efectiveness study to determine OGM application in the rutinary cytogenetic evaluation of MDS or TRMN patients.

PATIENTS AND METHODS:

Project design and patient selection: This is a prospective and retrospective study. The preliminary recruitment of 50 TRMN patients and their corresponding genetic study has been carried out in the previous FIS project granted in 2020 “Caracterización genética de las neoplasias mieloides asociadas a tratamiento (TRMN)” (PI 20/00531). Among 50 cases of previous FIS project, we have avaliable stored samples from 20 patients (retrospective study). The prospective study will be carried out with 30 new TRMN cases and from 25 MDS younger than 50 years old. Recruitment of primary MDS and TRMN patients have also been conducted and are available for the current proposal.

Recruitment of TRMN and gMDS patients will be done mainly by Germans Trias and Pujol ICO-Hospital center (Badalona) and from collaborating centers belonging to the Spanish Group of Myelodysplastic Syndromes (GESMD). We expect a minimum of 25 MDS patients younger than 50 years old to fit in the enrolment criteria during the 3 years of the proposal:

Samples: Samples will be collected in a retrospective and in a prospective manner. BM will be used as source of tumoral DNA and RNA, while peripheral blood (PB) CD3+ lymphocytes will be used as a source of germline DNA.

For the prospective study, PB and BM samples will be collected from new TRMN and MDS younger than 50 years old patients. All the samples will be stored in “Colección de Muestras del IJC de Leucemias y otros cánceres hematológicos".

The retrospective samples of BM and PB (cell pellets in trizol or cryopreserved cells) of TRMN and MDS patients are stored in the Hematology Laboratory of the Germans Trias i Pujol Hospital (HGTiP) and in "Col·lecció de Mostres de l'IJC de Leucèmies i altres Malalties Hematològiques". Collections have been registered at the "Carlos III Health Institute" (Ref: C0000886, C0002922, C0006730, C0006786). For alive patients whose germline paired control sample is not available, PB will be obtained prospectively.

For all TRMN and MDS cases, total BM cells will be obtained by erythrocyte lysis from BM samples. In addition, CD3+ T-lymphocytes will be obtained by immunomagnetic enrichment from PB samples.

Ethical considerations: Ethics Committee (CEIC) of the Hospital Germans Trias i Pujol approved the genetic and molecular studies included in the previous FIS project for human sample usage of TRMN samples (ref. PI-19-075).

Patient data and related samples will be recorded in a single database. The confidentiality and traceability of the patient data will be guaranteed by the NorayBio laboratory management tool (Noray Bioinformatics S.L., IGTP-HUGTiPBiobanc). Only samples of patients who have signed an informed consent will be included in the project. We will follow the Helsinki Declaration and the national data confidentiality regulations (Ley Orgánica de Protección de Datos, LOPD) for medical collection. Also, the institutional Data Protection Officer (DPO) will ensure the compliance with the national and international regulations for Data Protection.

DNA extraction and quantification: DNA extraction will be performed using the Maxwell 16 Blood DNA Purification Kit (Promega). For OGM, the DNA will be extracted following manufacturer’s guidelines. DNA will be quantified using a fluorometric method (Qubit, Thermo Fisher Scientific). DNA purity will be assessed by spectrophotometry (Nanodrop® 2000, Thermo Fisher Scientific).

Targeted deep sequencing (TDS): TDS will be performed in BM samples from 25 MDS patients using a custom myeloid panel including 51 genes previously reported in myeloid neoplasms. TDS will be performed at a mean coverage of 1000x, to detect variants at a 3-5% sensitivity, which will be integrated to the results of the 500 MDS cases from the rest of the centers participating in the Umbrella Project. TDS libraries will be prepared from 50ng of gDNA using the KAPA HyperCap Workflow 3.0 chemistry (Roche), following manufacturer´s recommendations.

Whole exome sequencing (WES): WES will be performed in tumoral and germline paired samples of 30 newly recruited TRMN patients and 25 MDS younger than 50 years old patients to evaluate the genetic landscape in both series of patients. BM will be used as source of tumoral DNA, while PB will provide CD3+ lymphocytes obtained by immunomagnetic enrichment as a source of germline DNA. Mutational burden in BM and CD3+lymphocytes will be compared to identify variants of germline origin. Somatic profiles will be compared with the corresponding MDS subtypes sequenced to identify molecular biomarkers and key molecular drivers. WES will be performed at an external core facility (Novogene) using an Illumina instrument. A minimum depth of coverage of 100x will be requested for PB and BM samples.

WES analysis will be performed using in-house pipelines. Raw data will be aligned against the reference genome (hg19/GRCh37). PCR duplicates will be marked and removed using Picard tools at this point. Variant calling will be performed by Mutect and Strelka2 software. All variants will be annotated with ANNOVAR software using ENSEMBL, ExAC, dbSNP, Exome Variant Server, GenomAD, ClinVar and COSMIC. Variants will be filtered according to variant type, population frequency and damage predictors information. Validation of candidate variants will be performed by Sanger sequencing.

RNA-Seq: Transcriptome studies of 50 TRMN patients and 25 MDS younger than 50 years old patients will be performed by a collaboration of the group from Pamplona. BM samples from 5 healthy elderly donors will be used as control samples. In all cases, CD34+ sorted cells from BM will be used as a source of RNA, then libraries will be prepared following the Illumina Stranded Total RNA Prep Protocol following manufacturer´s recomendations. Briefly, ribosomal RNA is depleted and then RNA is fragmented and denatured. Reverse transcription then converts the remaining RNA into cDNA, while subsequent ligation and amplification steps add adapters for clustering and sequencing on an Illumina system. Libraries will be sequenced at a sequencing depth of 50 million reads per sample. Finally, RNA-Seq data will be analyzed with DESeq2 package for identification of differentially expressed genes.

Optical Genome Mapping (OGM): Optical genome mapping (OGM) will be performed in 75 patients (20 retrospective and 30 prospective TRMN and 25 MDS younger than 50 years old in the Microarrays Unit IJC. In the collaborative project (UMBRELLA), OGM will be applied in patients with normal cytogenetics, cases without mitosis and cases with abnormal karyotype (using PB and BM of the same patients) in 40 samples. The same number of cases will be performed in three other centers involved in the collaborative project. OGM is based on the analysis of ultra-high molecular weight (UHMW) DNA molecules, providing a high-resolution genome-wide analysis highlighting copy number and structural anomalies. DNA will be extracted from cryopreserved cells or BM frozen pellets stored with DNA stabilizer. Once the DNA is extracted, the molecules are enzymatically tagged in a specific sequence (CTTAAG) that is repeated at a periodicity of approximately every 5000 base pairs. These long labeled DNA molecules are loaded onto a chip that straightens the

molecules and ultimately pass as individual DNA molecules through a nanochannel array where images of each molecule can be obtained. These images, each with a unique banding pattern, are then converted into a digital representation that can be bioinformatically processed and assembled into a genome to find all sorts of structural rearrangements. The data acquired with Saphyr instrument will be processed by Bionano Access software, where the patterns of markers deviating from the reference become apparent and are converted into structural variants. Results obtained by OGM will be used to confirm the structural variants already detected by karyotype and to determine the presence of complex rearrengements not detected by the other conventional techniques.

CALENDAR

Semesters	At 3 years 1	2	3
Patient recruitment and sample collection			
Genetic characterization of TRMN and MDS (≤50y) patients by WES			
i. Evaluation of the somatic landscape in TRMN			
ii. Evaluation of germline predisposition prevalence in TRMN patients (define TRMN-like cohort)			
iii. Germline predisposition studies in gMDS and gene discovery			
TDS in 25 prospective MDS patients (UMBRELLA PROJECT)			
Molecular driver studies by RNA-seq			
i. Differential expression analysis			
ii. Enrichment analysis			
iii. Integration of transcriptomic data to identify the key molecular driver pathways			
Detection of cytogenetic alterations by OGM technology			
i. Study of TRMN and MDS (≤50y**)** normal karyotype			
ii. Study of TRMN and MDS (≤50y) patients with complex and non complex karyotype			
iii. Study of the UMBRELLA project (PB and BM of cases with abnormal karyotype, cases with normal karyotype and cases without metaphases)			
Analysis of the results and performance of association studies			

TRMN: therapy related myeloid neoplasm, WES: whole exome sequencing, OGM: optical genome mapping, TDS: targeted deep sequencing.