

**“VICTOR BABEȘ” UNIVERSITY OF
MEDICINE AND PHARMACY TIMIȘOARA
FACULTY OF MEDICINE
DEPARTMENT VI – CARDIOLOGY**

LIGHEZAN DIANA LUISA



DOCTORAL THESIS

**CLINICAL, GENETIC AND BIOLOGICAL CORRELATIONS
IN THE PATHOGENESIS OF CRONIC
MYELOPROLIFERATIVE DISEASES**

A B S T R A C T

Scientific Coordinator
PROF. UNIV. DR. PETRESCU LUCIAN

**Timișoara
2021**

CONTENTS

List of published articles	VII
List of abbreviations	VIII
List of figures	IX
List of tables	X
Dedication	XII
Acknowledgments	XIII
INTRODUCTION	XV
 GENERAL PART	 1
1. Myeloid neoplasms	1
2. Myeloproliferative neoplasms	5
2.1. Overview	5
2.2. Chronic myeloid leukemia	5
2.2.1. Definitions and overview	5
2.2.2. Clinical characteristics	6
2.2.3. Laboratory findings	7
2.2.3.1. Chronic phase	7
2.2.3.2. Accelerated phase	8
2.2.3.3. Blastic phase	8
2.2.4. Treatment and prognostic	8
2.3. Polycythemia vera	9
2.3.1. Definitions and overview	9
2.3.2. Clinical characteristics	10
2.3.3. Laboratory findings	11
2.3.4. Treatment and prognostic	11
2.4. Primary myelofibrosis	12
2.4.1. Definitions and overview	12
2.4.2. Clinical characteristics	13
2.4.3. Laboratory findings	14
2.4.4. Treatment and prognostic	15
2.5. Essential thrombocythemia	16
2.5.1. Definitions and overview	16
2.5.2. Clinical characteristics	16
2.5.3. Laboratory findings	17

2.5.4. Treatment and prognostic.....	18
3. Genetics of myeloproliferative neoplasms	20
3.1. Somatic mutations.....	21
3.1.1. <i>BCR-ABL</i> , <i>JAK2</i> , <i>MPL</i> și <i>CALR</i> mutations.....	21
3.1.1.1. <i>BCR-ABL</i> mutations	21
3.1.1.2. <i>JAK2</i> mutations	22
3.1.1.3. <i>MPL</i> mutations.....	24
3.1.1.4. <i>CALR</i> mutations	25
3.1.2. Other somatic mutations.....	27
3.2. Germ-line variants.....	29
3.2.1. <i>JAK2</i> haplotype	30
3.2.2. Other germ-line variants	31
EXPERIMENTAL PART.....	32
1. Work hypothesis and general objectives	32
2. Study 1 – Assessment of <i>SH2B3</i> rs3184504 polymorphism as a risk factor in myeloproliferative neoplasms.....	35
2.1. Work hypothesis/objectives	35
2.2. Materials and methods	36
2.2.1. Patients and controls	36
2.2.1.1. Inclusion criteria.....	36
2.2.1.2. Exclusion criteria	36
2.2.2. Genotyping methods	39
2.2.2.1. DNA extraction	39
2.2.2.2. Somatic <i>JAK2</i> V617F mutation genotyping.....	39
2.2.2.2.1. Qualitative method: PCR tetra-primer	40
2.2.2.2.2. Quantitative method: quantitative Real time PCR (qPCR).....	42
2.2.2.3. Somatic <i>CALR</i> mutation genotyping	45
2.2.2.4. Somatic <i>BCR-ABL</i> mutation genotyping	46
2.2.2.4.1. Qualitative method: nested PCR for somatic <i>BCR-ABL</i> mutation detection.....	47
2.2.2.4.2. Quantitative method: real-time PCR for somatic <i>BCR-ABL</i> mutation detection and quantification	49
2.2.2.5. <i>SH2B3</i> rs3184504 polymorphism genotyping method	50
2.2.3. Statistical analysis	52
2.3. Results	53
2.3.1. Correlations between <i>SH2B3</i> rs3184504 genotypes and alleles, and MPN phenotypes	53
2.3.2. Correlations between <i>SH2B3</i> rs3184504 genotypes and alleles, and MPN molecular subtypes	55

2.3.3. Correlations between <i>SH2B3</i> rs3184504 polymorphism and various hematological and clinical features of patients with MPN	55
2.4. Discussions and conclusions	57
3. Study 2 - Assessment of <i>TET2</i> rs1548483 polymorphism as a risk factor in myeloproliferative neoplasms. The relationship between <i>TET2</i> rs1548483 polymorphism and somatic mutations and other germ-line variants	61
3.1. Work hypothesis/objectives	61
3.2. Materials and methods	63
3.2.1. Patients and controls	63
3.2.2. Genotyping methods	64
3.2.2.1. DNA extraction	64
3.2.2.2. Somatic <i>JAK2</i> V617F mutation genotyping	65
3.2.2.3. Somatic <i>CALR</i> mutation genotyping	65
3.2.2.4. Somatic <i>BCR-ABL</i> mutation genotyping	65
3.2.2.5. <i>TET2</i> rs1548483 polymorphism genotyping method	66
3.2.3. Statistical analysis	67
3.3. Results	68
3.3.1. Description of MPN subtypes	68
3.3.2. Association between <i>TET2</i> rs154843 SNP and MPN Subtypes–Allelic Model	69
3.3.3. Association between <i>TET2</i> rs154843 SNP and MPN Phenotypes–Genotypic Models	71
3.3.4. Association between <i>TET2</i> rs154843 SNP and MPN Molecular Subtypes–Genotypic Models	73
3.3.4.1. <i>JAK2</i> V617F mutations	73
3.3.4.2. <i>CALR</i> mutations	75
3.3.5. Epistatic Two-Way SNPs Interaction Stratified by MPN Subtypes	77
3.4. Discussions and conclusions	82
4. GENERAL DISCUSSIONS	87
CONCLUSIONS	91
LIST OF REFERENCES	95
APPENDIX	I

GENERAL PART

According to the latest WHO (World Health Organization) classification revised in 2016, the term *myeloid neoplasm* covers clonal changes of all myeloid cell lineages, thus comprising 8 distinct myeloid entities. The subcategory of chronic myeloid neoplasms includes the MPN category, which comprises the three major ("classic") subcategories of *BCR-ABL*-negative neoplasms - polycythemia vera (PV), essential thrombocythemia (ET), primary myelofibrosis (PMF), along with chronic myeloid leukemia (CML) and 3 other myeloid subcategories (chronic neutrophilic leukemia, chronic eosinophilic leukemia not otherwise specified, and myeloproliferative neoplasms, not otherwise specified).

In case of the *BCR-ABL*-negative neoplasms, the new WHO classification has stipulated the presence of one of the 3 known driver mutations, as a major diagnostic criteria for PV (*JAK2* V617F and exon 12), ET and PMF (*JAK2* V617F, *CALR*, *MPL*) .

In addition to these somatic mutations, modern sequencing techniques of the human genome allowed the detailed analysis of germ-line variants, specifically "single nucleotide polymorphisms" (SNPs), which increase the risk of developing an NMP.

EXPERIMENTAL PART

1. WORK HYPOTHESIS AND GENERAL OBJECTIVES

NMPs are pathologies caused by the clonal proliferation of myeloid progenitors and acquired somatic mutations (*JAK2*, *CALR* and *MPL* mutations) are the promoters of this myeloid cell overproduction. However, they cannot sufficiently explain the phenotypic diversity of these diseases.

The extensive description of driver mutations in the last years, has not been able to completely identify the factors that lead to such a heterogeneous phenotypic expression of MPNs. To date, no explanation is given for the question of why some patients with *JAK2* V617F somatic mutation develop a PV type MPN, while others develop ET. This observation highlights the possibility of a constitutional germ-line variant in these individuals, which could explain the diversity of phenotypic expression.

The experimental part of my doctoral thesis tried to assess the role of some inherited germ-line factors in NMP.

The objectives of my research were the following:

1. Characterization of somatic driver mutations (*JAK2*, *CALR* and *MPL*) in a representative cohort of patients with MPN.
2. Defining the contribution of the *SH2B3* gene polymorphism (study 1), respectively of the *TET2* gene polymorphism (study 2) to the expression of different MPN phenotypes.
3. Defining the contribution of the *SH2B3* gene polymorphism (study 1), respectively of the *TET2* gene polymorphism (study 2) to the occurrence of various somatic MPN-specific driver mutations.
4. Assessment of the interactions between *SH2B3* polymorphism and other germ-line variants such as *JAK2* rs10974944, *TERT* rs2736100, *MECOM* rs2201862, *HBS1L-MYB* rs9376092 and *THRB-RARB* rs4858647 in the occurrence of various phenotypes of MPN patients.
5. Evaluating the correlation between genetic constitutional status and the biological and clinical expression in a subgroup of MPN patients included in the study.

2. STUDY 1 – ASSESSMENT OF *SH2B3* RS3184504 POLYMORPHISM AS A RISK FACTOR IN MYELOPROLIFERATIVE NEOPLASMS

2.1. WORK HYPOTHESIS/OBJECTIVES

SH2B3 gene codes a protein named LNK, which promotes dysregulation of the JAK-STAT pathway, producing a proliferation of the myeloid elements which is the major feature of the MPN (84, 85).

The aim of this study was to determine whether the *SH2B3* rs3184504 polymorphism is associated with an enhanced potential of acquiring MPN-associated somatic driver mutations and developing the four major MPN - PV, ET, PMF, and CML.

2.2. MATERIALS AND METHODS

2.2.1. PATIENTS AND CONTROLS

This study included 1901 patients with various MPN and a control group of 359 individuals without any hematological malignancy. Out of the 1901 patients, 575 had PV, 798 had ET, 251 had PMF, and 277 had CML. Only patients with a molecularly proven driver mutation (*JAK2* V617F, *CALR* or *BCR-ABL1* mutation) were included. ET and PMF patients positive for *MPL* mutations were excluded, due to their low number.

2.2.2. GENOTYPING METHODS

All the genetic assays were performed on DNA obtained from peripheral whole blood, collected on EDTA, using various commercial kits (Wizard Genomic DNA Purification kit, Promega, Madison, WI, USA; Quick gDNA MiniPrep kit, Zymo Research, Irvine, CA, USA; PureLink Genomic DNA Mini Kit, Invitrogen, Thermo Fisher, Waltham, MA, USA).

The somatic driver mutations were analyzed in all patients, namely *JAK2* V617F using a tetra-primer PCR and a real-time PCR assay, *CALR* mutations were assessed using a simplex PCR. *BCR-ABL1* mutations were analyzed using a qualitative nested PCR and quantified using an automated, cartridge-based real-time PCR system.

SH2B3 rs3184504 polymorphism was genotyped in all individuals using a TaqMan SNP Genotyping Assay (assay number C__2981072_10), by a real-time PCR system (Applied Biosystems, Thermo Fisher, USA).

2.3. RESULTS

2.3.1. CORRELATIONS BETWEEN *SH2B3* RS3184504 GENOTYPES AND ALLELES, AND MPN PHENOTYPES

First, we analyzed the relationship between *SH2B3* rs3184504 genotypes, their alleles, and the four phenotypes included in the study, namely PV, ET, PMF, and CML. In case of rs3184504 genotypes, we analyzed both dominant and recessive genetic models, but only the recessive model yielded statistically significant results. The homozygous TT genotype was significantly associated with PV (OR = 1.54; 95% CI = 1.14-2.06; crude p-value = 0.004; adjusted p-value = 0.024), and with PMF (OR = 1.50; 95% CI = 1.04-2.12; crude p-value = 0.02; adjusted p-value = 0.04). The homozygous TT genotype also attained statistical significance when analyzing the whole cohort of *BCR-ABL1*-negative MPN (OR = 1.34; 95% CI = 1.03-1.74; crude p-value = 0.02; adjusted p-value = 0.04). There was no correlation with ET and CML phenotypes. We also analyzed the allelic model, but in this case the T allele showed a near significant association only in case of PV (OR = 1.19; 95% CI = 1-1.44; crude p-value = 0.06; adjusted p-value = 0.24), while all other comparisons

showed non-significant results. We also analyzed the dominant model (CT+TT versus CC genotype), but in this case, none of the comparisons performed, yielded significant results.

2.3.2. CORRELATIONS BETWEEN *SH2B3* RS3184504 GENOTYPES AND ALLELES, AND MPN MOLECULAR SUBTYPES

Then, we analyzed the relationship between *SH2B3* rs3184504 genotypes and their alleles, and the molecular subtypes to which the MPN patients included in the study belong. We analyzed both dominant and recessive genetic models, but again only the recessive model yielded statistically significant results. Because all patients with PV harbored *JAK2* V617F mutation, identical values for TT genotype were obtained as in the case of PV phenotype (OR = 1.54; 95% CI = 1.14-2.06; crude p-value = 0.004; adjusted p-value = 0.024). The TT genotype was also associated with *JAK2* V617F-positive PMF, albeit less significantly (OR = 1.57; 95% CI = 1.04- 2.33; crude p-value = 0.03; adjusted p-value = 0.08). Similar results for the TT genotype as in the case of *JAK2* V617F-positive PMF were also obtained when analyzing the whole cohort of MPN patients harboring *JAK2* V617F mutation (OR = 1.36; 95% CI = 1.04-1.77; crude p-value = 0.02; adjusted p-value = 0.08). However, the TT genotype was not associated with *CALR*-positive ET or PMF, *JAK2* V617F-positive ET, or CML. Again, the allelic model yielded non-significant results, except for PV (where all patients were *JAK2* V617-positive), in which the T allele showed the same near significant association (OR = 1.19; 95% CI = 1-1.44; crude p-value = 0.06; adjusted p-value = 0.24). We also analyzed the dominant model (CT+TT versus CC genotype), but in this case, none of the comparisons performed, yielded significant results.

2.3.3. CORRELATIONS BETWEEN *SH2B3* RS3184504 POLYMORPHISM AND VARIOUS HEMATOLOGICAL AND CLINICAL FEATURES OF PATIENTS WITH MPN

We also assessed whether *SH2B3* rs3184504 was correlated with hematological parameters displayed by MPN patients (hemoglobin, hematocrit, white blood cells count, platelets). The white blood cells count had higher values in patients with homozygous TT genotype. However, this was seen only in PV (median 12.95 x 10⁶/L in patients with TT genotype versus 11.45 x 10⁶/L in patients with CT+CC genotypes, Mann-Whitney test p-value = 0.03). All other comparisons yielded statistically non-significant results (p-value > 0.05).

The complete clinical information regarding the occurrence of major thrombosis was available in 375 patients with ET and 273 patients with PV. There were 97 patients with ET (25.9%) and 108 patients with PV (39.5%) who have experienced major thrombosis at diagnosis. The following events were considered as major thrombosis: stroke/transient ischemic attack, acute coronary disease, acute limb ischemia, splenic infarction, mesenteric infarction, deep venous thrombosis, splanchnic thrombosis, and cerebral sinus venous thrombosis. There was no correlation between *SH2B3* rs3184504 genotypes and alleles and the occurrence of major thrombosis.

2.4. DISCUSSIONS AND CONCLUSIONS

Our study showed a significant association between the TT homozygous genotype of *SH2B3* rs3184504 and *JAK2* V617F-positive MPN, but not *CALR*-positive MPN. Thus, we have confirmed on a large cohort of patients the contribution of *SH2B3* rs3184504 to the occurrence of *JAK2* V617F-positive MPN, placing it together with other polymorphisms defining the genetic predisposition to MPN, such as *JAK2* 46/1 haplotype, *TERT* rs2736100 or *MECOM* rs2201862.

3. STUDY 2 - ASSESSMENT OF *TET2* RS1548483 POLYMORPHISM AS A RISK FACTOR IN MYELOPROLIFERATIVE NEOPLASMS. THE RELATIONSHIP BETWEEN *TET2* RS1548483 POLYMORPHISM AND SOMATIC MUTATIONS AND OTHER GERM-LINE VARIANTS

3.1. WORK HYPOTHESIS/OBJECTIVES

Somatic *TET2* mutations are frequent events in MPN and CHIP (Clonal Hematopoiesis of Indeterminate Potential), whilst genetic germ-line variations at *TET2* locus seem to play a role in MPN predisposition. To address this issue, we have analyzed a large cohort of MPN patients, aiming to establish the additional contribution of the recently described *TET2* rs1548483 SNP to the occurrence of MPN phenotypes and their associated somatic mutations.

3.2. MATERIALS AND METHODS

3.2.1. PATIENTS AND CONTROLS

A total of 1601 patients diagnosed with MPN were included in this study: 431 with PV, 688 with ET, 233 with PMF and 249 with CML. The study included only patients with a molecularly proven driver mutation, *JAK2* V617F, *CALR* or *BCR-ABL1*. ET and PMF patients positive for *MPL* mutations were excluded, due to their low number. Among the 1601 included patients, 939 patients were explored in our previous work for other germ-line variants. Specifically 454 patients with ET, 337 patients with PV and 148 patients with PMF had been previously genotyped for *TERT* rs2736100, *MECOM* rs2201862, *HBS1L-MYB* rs9376092, and *THRB-RARB* rs4858647 polymorphisms and *JAK2* 46/1 haplotype. A total of 884 patients, namely 427 with ET, 316 with PV, and 139 PMF were successfully genotyped for all SNPs—the five above-mentioned ones and *TET2* rs1548483. The study also included 197 individuals with no hematological malignancy, representing the control group.

3.2.2. GENOTYPING METHODS

The same methods as mentioned in the former study, were used for DNA extraction and detection of somatic *JAK2* V617F, *CALR* și *BCR-ABL1* mutations. *TET2* rs1548483 SNP was genotyped in all patients and controls using a TaqMan assay (assay number C__7512138_20), as recommended by the manufacturer (Applied Biosystems, Thermo Fisher, Waltham, MA, USA), using Quant Studio 3 or 7500 Fast Dx real-time PCR systems (Applied Biosystems, Thermo Fisher, Waltham, MA, USA).

3.3. RESULTS

3.3.1. DESCRIPTION OF MPN SUBTYPES

The distribution of driver mutations was as follows: all PV patients were *JAK2* V617F positive ($n = 431$). Out of the ET patients, 525 (76,3%) were *JAK2* V617F positive, and 163 (23,7%) harbored *CALR* mutations. In case of PMF, 151 patients were *JAK2* V617F positive (64,8%) și 82 (35,2%) *CALR* positive.

3.3.2. ASSOCIATION BETWEEN *TET2* RS154843 SNP AND MPN SUBTYPES—ALLELIC MODEL

The genotype distribution of *TET2* rs154843 SNP satisfied the Hardy–Weinberg equilibrium in the analyzed reference population (p -value = 0.334 for the control group).

Out of the analyzed patient groups, the PV and PMF groups showed a significant increase of MAF compared to control group (OR = 1.70; 95% CI: 1.01–2.91; p -value =

0.046, and OR = 2.02; 95% CI: 1.14–3.57; p -value = 0.015, respectively). Taking into consideration molecular subtypes, *TET2* variant allele was associated with *JAK2* V617F-positive PV and PMF (OR = 1.70; 95% CI: 1.01–2.91; p -value = 0.046, and OR = 2.04; 95% CI: 1.10–3.77; p -value = 0.024, respectively). The association between *TET2* variant allele and *CALR* -positive PMF bordered the statistical significance (OR = 1.95; 95% CI: 0.95–4.02; p -value = 0.066), on the behalf of *CALR* type 2 mutations (OR = 2.98; 95% CI: 1.12–7.93; p -value = 0.035).

3.3.3. ASSOCIATION BETWEEN *TET2* RS154843 SNP AND MPN PHENOTYPES—GENOTYPIC MODELS

The variant genotypes of *TET2* SNP were significantly associated with an increased risk of PMF in the codominant and overdominant models tested (OR = 2.4; 95% CI: 1.3–4.43; p -value = 0.005; corrected p -value = 0.013 and OR = 2.41; 95% CI: 1.31–4.45; p -value = 0.005, corrected p -value = 0.0125 respectively). *TET2* variant genotypes remained an independent risk factor also after adjusting for age and gender (OR = 1.95, 95% CI: 1.02–3.73; p -value = 0.044 and OR = 1.96; 95% CI: 1.03–3.76; p -value = 0.041, respectively). In addition, there was a significant positive association between variant genotype of *TET2* SNP and risk of PMF in the dominant model (OR = 2.26; 95% CI: 1.24–4.12; p -value = 0.008; corrected p -value = 0.0133), this association bordering the statistical significance after the adjustment for age and gender (OR = 1.84; 95% CI: 0.97–3.47; p -value = 0.062).

3.3.4. ASSOCIATION BETWEEN *TET2* RS154843 SNP AND MPN MOLECULAR SUBTYPES—GENOTYPIC MODELS

3.3.4.1. *JAK2* V617F mutation

We observed a positive association between the variant genotypes of *TET2* SNP and *JAK2* V617F-positive PMF in the codominant, dominant and overdominant models tested (OR = 2.43; 95% CI: 1.26–4.69; p -value = 0.008; corrected p -value = 0.020; OR = 2.29; 95% CI: 1.2–4.37; p -value = 0.012; corrected p -value = 0.020 and OR = 2.44; 95% CI: 1.26–4.72; p -value = 0.008; corrected p -value = 0.020 respectively). This association became weaker after adjustment for age and gender, bordering the statistical significance in this case for all three models (OR = 1.99; 95% CI: 0.98–4.07; p -value = 0.056; OR = 1.88; 95% CI: 0.94–3.79; p -value = 0.076 and OR = 2.01; 95% CI: 0.99–4.01; p -value = 0.053, respectively). Also in the case of *JAK2* V617F-positive PV, a near significant association was observed in the dominant model (crude OR = 1.68; 95% CI: 0.95–2.96; p -value = 0.074). However, this association was no longer observed after the adjustment for age and gender (p -value > 0.05).

3.3.4.2. *CALR* mutation

The results highlighted a positive significant association between the variant genotype of *TET2* SNP and risk of *CALR* -positive PMF in the dominant and overdominant inheritance models tested (OR = 2.18; 95% CI: 1.02–4.66; p -value = 0.045 corrected p -value = 0.075 and OR = 2.33; 95% CI: 1.08–5.03; p -value = 0.031; corrected p -value = 0.075). However, this association was no longer observed after the adjustment for age and gender (p -value > 0.05). We then explored the possible different effect of *TET2* SNP on the two major types of *CALR* mutations, type 1 and type 2, respectively. No statistical associations were observed between *TET2* SNP and *CALR* type 1 mutations in any MPN subtype (p -value > 0.05 for all these comparisons). However, in the case of type 2 mutations, the association was significant in PMF, in codominant, dominant and overdominant models (OR = 3.75; 95% CI: 1.30–10.78; p -value = 0.014; corrected p -value = 0.030; OR = 3.53; 95% CI: 1.24–10.08; p -value = 0.018; corrected p -value = 0.030, and OR = 3.77; 95% CI: 1.31–10.84; p -value = 0.014; corrected p -value = 0.030, respectively). After the adjustment for age and gender, the associations were no longer significant (p -value > 0.05).

3.3.5. EPISTATIC TWO-WAY SNPS INTERACTION STRATIFIED BY MPN SUBTYPES

We also tested for epistatic interaction between *TET2* rs154843 and other SNPs (*TERT* rs2736100, *MECOM* rs2201862, *HBS1L-MYB* rs9376092, and *THRB-RARB* rs4858647), which we previously genotyped in PV, ET and PMF. Based on the log-likelihood ratio test (LRT), we identified epistatic interactions between *TET2* rs154843 and *HBS1L-MYB* rs9376092 in PV (pinteraction = 0.012 under overdominant model).

We also identified epistatic interactions between *TET2* rs154843 and *HBS1L-MYB* rs9376092 (pinteraction = 0.014 and pinteraction = 0.049 under codominant and overdominant model) and *JAK2* rs10974944, which tags the *JAK2* 46/ 1 haplotype (pinteraction = 0.037 under recessive model) in ET.

3.4. DISCUSSIONS AND CONCLUSIONS

Through this study we have reported a significant association between *TET2* rs154843 and PV and PMF phenotypes. Moreover, *TET2* rs154843 has been associated with *JAK2* V617F-positive PV and PMF, and type 2 *CALR* -positive PMF.

4. GENERAL DISCUSSIONS

In essence, chronic myeloid neoplasms are a group of heterogeneous, clonal entities that result from the proliferation of myeloid progenitors. These neoplasms tend to progress through a leukemic or a fibrotic transformation (119). Clonal proliferation encountered in *BCR-ABL1*-negative NMPs is promoted by somatic mutations, including *JAK2* V617F, *CALR* and *MPL*. Likewise, the *BCR-ABL1* mutation in CML. These mutations are promoters of clonal proliferation.

The introduction of somatic *JAK2* V617F, together with *CALR* and *MPL* mutation in the diagnostic algorithm, has led to an increased accuracy and a more efficient NMP diagnosis. If in the last 2 decades important discoveries have been made in the field of somatic mutations in MPN, the current focus relies on the analysis of an individual, inherited genetic status, which could bring more insight into MPN development.

In this thesis I have characterized two distinct germ-line variants that predispose to NMP: consisting of one *SH2B3* gene polymorphism in the first study and one *TET2* gene polymorphism in the second study. Then, I have analyzed the contribution of each polymorphism to the appearance of NMP phenotypes, namely PV, ET, PMF and CML. In the first study, I was able to demonstrate a significant association of the TT homozygous genotype (recessive model) of the *SH2B3* rs3184504 polymorphism with PV and PMF, but also with the entire *BCR-ABL1*-negative MPN cohort (100). These results are consistent with data published by Lesteven et al., who first described the association between the *SH2B3* rs3184504 polymorphism and MPN, reporting its association with MPN in the allelic model, with the T allele being the risk allele. In their study, the most significant association was also in the case of PMF. However, their study did not include patients with PV (87). Other following studies such as Olkhovskiy et al. included patients with PV, have reported a significant correlation only in patients with PV, while Chen et al. observed the association with all three types of *BCR-ABL1*-negative MPNs (PV, ET and PMF). Regarding the correlation between *SH2B3* rs3184504 and CML, we did not observe statistically significant results, similar to the data obtained by Olkhovskiy et al. but inconsistent with data from Chen et al., who reported a remarkable frequency of CC genotype in the group of CML patients (88, 89).

In the second study I have focused on the correlation between *TET2* gene polymorphism rs1548483 and the expression of different NMP phenotypes, demonstrating an association with PV (in the allelic model) and PMF (in the allelic model and in the dominant, codominant and overdominant genotypic models). Similar results were published by Hinds et al. who reported an

association of this SNP with MPN. In addition, their study revealed a weaker nominal association with CML and systemic mastocytosis.

A larger access to genomic molecular testing has allowed a more detailed analysis of the most common single nucleotide polymorphisms (SNPs). This genetic evaluation allowed the identification of germline SNPs, that provided more evidence in favor of an inherited cause of MPN, such as the *JAK2* 46/1 haplotype and the identification of many other SNPs in genes such as *TERT*, *MECOM*, *SH2B3*, *TET2*, *ATM*, *CHEK2*, *THRB-RARB*, *PINT* (54.75-83). Thus, on a subgroup of the analyzed cohort from the second study, I have studied the epistatic interactions between the *TET2* rs1548483 polymorphism and the germ-line variants such as *JAK2* rs10974944, *TERT* rs2736100, *MECOM* rs2201862, *HBS1L-MYB* rs93748 regarding phenotype expression in MPN. Following this analysis, we have discovered epistatic interactions between *TET2* rs1548483 and *HBS1L-MYB* rs9376092 in PV in the overdominant model, between *TET2* rs1548483 and *HBS1L-MYB* rs9376092 in ET in the codominant and overdominant model and between *TET2* rs1548483 and *JAK2* rs10974944 (*JAK2* 46/1 haplotype) in ET in the recessive model (110).

Regarding the analysis of other parameters correlated with the presence of a genetic predisposition, we have analyzed the correlation between the *SH2B3* rs3184504 polymorphism and various clinical and paraclinical parameters in the first study. These include hematological parameters such as hemoglobin, hematocrit, leukocyte and platelets count. Out of all the analyzed parameters, only leukocytosis was positively correlated with the *SH2B3* rs3184504 polymorphism (homozygous TT genotype, recessive genotypic model).

5. GENERAL CONCLUSIONS

In conclusion, in this doctoral thesis:

- I have characterized somatic driver mutations like *JAK2*, *CALR*, *MPL* and *BCR-ABL1* on a representative cohort of patients with MPN composed of 1901 individuals.
- I have characterized the contribution of the *SH2B3* gene polymorphism (study 1) to the appearance of different NMP phenotypes, namely PV, ET, PMF and CML. The TT homozygous genotype (recessive model) of the *SH2B3* rs3184504 polymorphism was significantly associated with PV and PMF and reached statistical significance when analyzing the entire *BCR-ABL1*-negative NMP cohort.
- I have characterized the contribution of the *TET2* gene rs1548483 polymorphism (study 2) to the appearance of different phenotypes of MPN. *TET2* rs1548483 polymorphism was correlated with PV and PMF in the allelic model. Also, variant genotypes containing at least one minor *TET2* rs1548483 allele were statistically significant associated with an increased risk of PMF in the dominant, codominant, and overdominant genotypic models.
- I have defined the contribution of the *SH2B3* rs3184504 polymorphism to the occurrence of various MPN specific somatic driver mutations. The TT homozygous genotype (recessive model) of the polymorphism was statistically significant associated with the entire *BCR-ABL1*-negative MPN cohort, *JAK2* V617F-positive PV and *JAK2* V617F-positive PMF.
- I have defined the contribution of *TET2* rs1548483 polymorphism in the appearance of various somatic driver mutations, specific to MPN. The minor allele of the *TET2* rs1548483 polymorphism was associated with *JAK2* V617F positive PV and PMF in the allelic model. Also, the *TET2* rs1548483 variant genotypes were positively associated with *JAK2* V617F PMF in the dominant, codominant and overdominant

genotypic models. In addition, *TET2* rs1548483 polymorphism was significantly associated with *CALR*-positive PMF, especially type 2 positive mutations, in the dominant, codominant, and overdominant genotypic models.

- I have analyzed the correlation between *SH2B3* rs3184504 polymorphism and various hematological laboratory findings (hemoglobin, hematocrit, leukocyte count, platelet count) and clinical findings (major thrombosis) characterising MPN patients. Out of all the analyzed parameters, only leukocytosis was positively correlated with the *SH2B3* rs3184504 polymorphism (homozygous TT genotype, recessive genotypic model)
- I evaluated the interactions between the *TET2* rs1548483 polymorphism and constitutional genetic variants like *JAK2* rs10974944, *TERT* rs2736100, *MECOM* rs2201862, *HBS1L-MYB* rs9376092 and *THRB-RARB* rs4858647 in the appearance of various phenotypic expressions in MPN patients. I observed epistatic interactions between *TET2* rs1548483 and *HBS1L-MYB* rs9376092 the PV overdominant genotype model, between *TET2* rs1548483 and *HBS1L-MYB* rs9376092 the ET codominant and overdominant model and between *TET2* rs1548483 and *JAK2* rs10974944 (*JAK2* 46/1 haplotype) in ET recessive model.

Hereby I confirm on a large cohort of patients the contribution of *SH2B3* rs3184504 and *TET2* rs1548483 to the occurrence of *JAK2* V617F-positive MPN, placing it together with other polymorphisms defining the genetic predisposition to MPN, such as *JAK2* 46/1 haplotype, *TERT* rs2736100 or *MECOM* rs2201862.

5.1 FUTURE RESEARCH DIRECTIONS

1. Performing a GWAS (Genome wide association study) to analyze the presence of other genetic variants in MPN patients. These data could be integrated into studies concerning epistatic interactions between several polymorphisms present in these patients.

2. Conducting methylation studies in genes like *SH2B3*, *TET2* and other, in order to analyze the functional effects of polymorphisms in these genes.

3. Analysis of additional genetic mutations, others than driver mutations such as *JAK2*, *CALR* and *MPL*, and the correlation with gene polymorphisms present in MPN patients.

4. Correlation of clinical data and hematological laboratory findings with the presence of additional mutations and gene polymorphisms, in order to optimise the description the phenotypic expression and evolution of MPN.

LIST OF PUBLISHED ARTICLES

1. **Lighezan DL**, Bojan AS, Iancu M, et al. *TET2* rs1548483 SNP Associating with Susceptibility to Molecularly Annotated Polycythemia Vera and Primary Myelofibrosis. *J Pers Med*. 2020;10(4):259.
doi:10.3390/jpm10040259
Impact factor 4.433 (2019)
First author
2. Trifa AP, **Lighezan DL**, Jucan C. et al. *SH2B3* (LNK) rs3184504 polymorphism is correlated with *JAK2* V617F-positive myeloproliferative neoplasms. *Rev Romana Med Lab*. 2020;28(3):267-77.
DOI:10.2478/rrlm-2020-0025
Impact factor 0.945 (2019)
Author with equal contribution to first author