The role of the *JAK2* GGCC haplotype and the *TET2* gene in familial myeloproliferative neoplasms

Damla Olcaydu,[1]* Elisa Rumi,[2]* Ashot Harutyunyan[1], Francesco Passamonti,[2] Daniela Pietra,[2] Cristiana Pascutto,[2] Tiina Berg,[1] Roland Jäger,[1] Emma Hammond,[3] Mario Cazzola,[2] and Robert Kralovics[1,4]

[1]Center for Molecular Medicine (CeMM) of the Austrian Academy of Sciences, Vienna, Austria; [2]Department of Hematology, Fondazione IRCCS Policlinico San Matteo of Pavia, Pavia, Italy; [3]Centre for Clinical Immunology and Biomedical Statistics, Murdoch University and Royal Perth Hospital, Perth, Australia; and [4]Division of Hematology and Blood Coagulation, Department of Internal Medicine I, Medical University of Vienna, Vienna, Austria

Citation: Olcaydu D, Rumi E, Harutyunyan A, Passamonti F, Pietra D, Pascutto C, Berg T, Jäger R, Hammond E, Cazzola M, and Kralovics R. The role of the JAK2 *GGCC haplotype and the* TET2 *gene in familial myeloproliferative neoplasms. Haematologica 2011;96(3):367-374. doi:10.3324/haematol.2010.034488*

Supplementary Methods

Statistical analysis

Statistical analysis of single nucleotide polymorphism associations in familial and sporadic myeloproliferative neoplasms (MPN) was performed using the R statistical software (version 2.9.2)[1] in conjunction with the SNPassoc package for R (version 1.6-0).[2] Individuals with reactive conditions (n=203) were used as a demographically matched control population. Allelic association analysis for *rs10974944* comparing the allele frequencies in patients with familial and sporadic MPN was carried out using the statistical program PLINK (v1.06).[3] The approximate penetrance in familial MPN was calculated assuming that all affected members have a "causative hypothetical mutation" in a heterozygous state (dominant type of inheritance). Based on this assumption, the probability that each healthy member has of carrying the mutation was calculated (e.g. the healthy child of an affected member has a 0.5 probability of carrying the MPN-causing mutation, a sibling also has a 0.5 probability) and the number was summed up to an artificial "number of healthy carriers". The number of affected members was then divided by the sum of the number of affected members and the number of healthy carriers, resulting in the value of penetrance in familial MPN.

The penetrance of the *JAK2* GGCC haplotype was estimated by dividing the prevalence of MPN in the general population (10^{-4}) by the frequency of *JAK2* GGCC haplotype carriers (approximately 50%, homozygous and heterozygous carriers taken together).

Malignancies occurring either before or after the diagnosis of MPN were recorded and statistical analysis was carried out using the STATA/SE (version 9.2) and STATISTICA (version 8.0) programs. The relative risk of developing cancer among patients with familial MPN compared to sporadic cases was estimated by means of the odds ratio. Statistical adjustments by age and *JAK2* haplotype were performed using the Mantel-Haenszel method. Patients were grouped into four age categories with approximately equal numbers of subjects in each group to determine the differences in occurrence of malignancies according to age at diagnosis of MPN. The same analysis was conducted after patients were categorized into three groups with similar numbers of study subjects according to age at last follow-up. The incidence of malignancies after a diagnosis of MPN in patients with familial and sporadic MPN was compared using incidence rate ratio calculations. Age- and *JAK2* haplotype-adjusted Mantel-Haenszel incidence rate ratio estimates were calculated in order to correct for the influence of age at the diagnosis of MPN and *JAK2* haplotype on the statistical evaluation. Kaplan-Meier estimates and log-rank tests were conducted in order to define the difference in cancer-free survival between patients with familial or sporadic MPN.

References

1. Team RDC. R: A Language and Environment for Statistical Computing. 2009; ISBN 3-900051-07-0.
2. Gonzalez JR, Armengol L, Sole X, Guino E, Mercader JM, Estivill X, et al. SNPassoc: an R package to perform whole genome association studies. Bioinformatics. 2007;23(5):644-5.
3. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559-75.

Online Supplementary Table S1. Primer sequences for sequence analysis of *TET2*, *CBL* and *MPL* genes in familial MPN.

Gene	Exon	Forward (5'-3'), 10μM	Reverse (5'-3'), 10μM	PCR conditions
TET2	exon 3_1	CAGTTTGCTATGTCTAGGTATTCCG	AGGCCCACTGCAGTTATGTG	
	exon 3_2	TGAACCTTCTCTCTCTGGGC	GTCTGTGCGGAATTGATCTG	
	exon 3_3	CACACATGGTGAACTCCTGG	AAGCAATTGTGATGGTGGTG	
	exon 3_4	TCTGTTCAGGTTCCAGCAG	TGCTGGCAGTTGTCCTGTAG	
	exon 3_5	GCCTCAGAATAATTGTGTGAACAG	TTTTGGAACTGGAGATGTTGG	95°C-5 min.
	exon 3_6	AAATTCCAACATGCCTGGG	TTCACCATGAAAACATTCTTCC	*10x touch down (-1°C/cycle):*
	exon 3_7	TCCCAGAGTTCACATCTCCC	AGTTGCGCAGCTTGTTGAC	94°C-30 sec.
	exon 3_8	TTTTGCAGGAAACAAGACCC	AAACTGCTTCAGATGCTGCTC	65°C-55°C-30 sec.
	exon 3_9	TTAAGGTGGAACCTGGATGC	AGCCTTTACAAATTGCTGCC	72°C-30 sec.
	exon 4	GCCCTTAATGTGTAGTTGGGG	ACAATTGTTTACTGCTTTGTGTG	26x:
	exon 5	AACCGTTCATTTCTCAGGATG	GGCATGAGTCTTTGATCTGG	94°C-30 sec.
	exon 6	CATGCTGATAAATGTTGCCC	ATAAATGTAAAAGTGCACGCTG	55°C-30 sec.
	exon 7	TCAATTGAAATAGTTCTGTGTGTGG	AAAATAGTGTGTATCTACAGTTTGGG	72°C-30 sec.
	exon 8	GGGATTCAAAATGTAAGGGG	TGCAGTGGTTTCAACAATTAAG	72°C-10 min.
	exon 9	TGTCATTCCATTTTGTTTCTGG	CCATGTGTTTTGGGAAGGAC	8°C-store
	exon 10	CTAGGCCACCAACACAAATC	CAGAACTTACAAGTTGATGGGG	
	exon 11_1	CGTATATCACTAGTGGAGTTTCTTACC	ACAGATCCATCGGCTGAGAC	
	exon 11_2	TCTAATCCCATGAACCCTTACC	CCTTGTTTTGGAGATGCAGG	
	exon 11_3	GATGGCCACTTCATGGGAG	ATGCTGGTAAAAGACGAGGG	
	exon 11_4	GCTTTCTGGATCCTGACATTG	ACTGTGACCTTTCCCCACTG	
CBL	exon 8	AGGACCCAGACTAGATGCTTTC	GGCCACCCCTTGTATCAGTA	as above, *touch down annealing*
	exon 9	CTGGCTTTTGGGGTTAGGTT	TCGTTAAGTGTTTTACGGCTTT	*temperature modified to 67°C-57°C*
MPL	exon 10	AGAGTAGGGGCTGGCTGG	AGGTGACGTGCAGGAAGTG	as for *TET2*

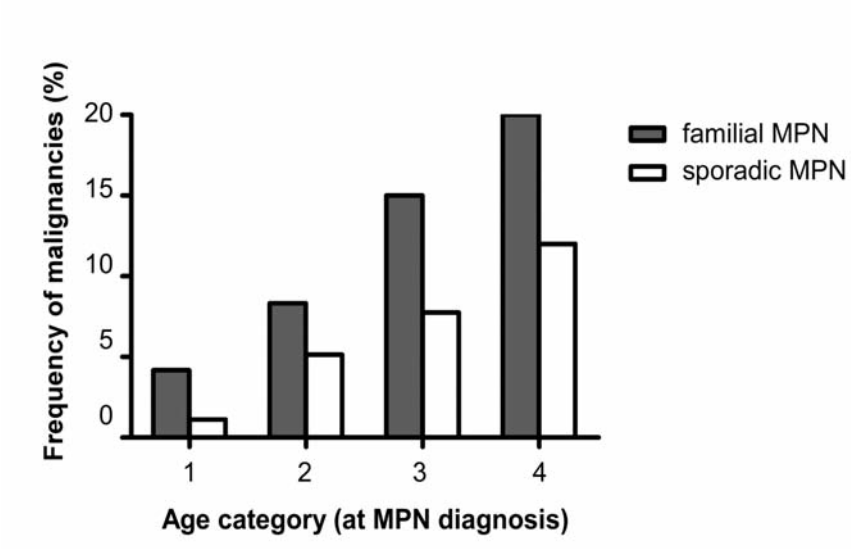
Online Supplementary Table S2. Clinical characteristics at diagnosis of 772 MPN patients investigated for familial clustering.

N. (%)	Familial MPN 88 (11)	Sporadic MPN 684 (89)	Total 772
Diagnosis, N. (%)			
Polycythemia vera	36 (41)	222 (32)	258
Essential thrombocythemia	36 (41)	364 (53)	400
Primary myelofibrosis	14 (16)	64 (9)	78
Post-polycythemia vera myelofibrosis	1 (1)	20 (3)	21
Post-essential thrombocythemia myelofibrosis	0 (0)	12 (2)	12
Unclassifiable MPN	0 (0)	2 (1)	2
Chronic myeloid leukemia	1 (1)	0	1
Sex, N. (%)			
Male	42 (48)	295 (43)	337 (44)
Female	46 (52)	389 (57)	435 (56)
Age, years			
Median	52.9	51	51.4
Range	16.2-77.8	13-85.7	13-85.7
Thrombosis,* N. (%)			
Before diagnosis	7 (7.9)	68 (9.9)	75 (9.7)
At diagnosis	7 (7.9)	41 (5.9)	48 (6.2)
Hemorrhage,* N. (%)			
Before diagnosis	1 (1.1)	25 (3.6)	26 (3.3)
At diagnosis	1 (1.1)	8 (1.1)	9 (1.1)

**thrombosis includes events before and at diagnosis; hemorrhage includes events before and at diagnosis.*

Online Supplementary Table S3. Occurrence of malignancies in patients with familial and sporadic MPN.

Site of cancer	Familial MPN (88, 11%)	Sporadic MPN (684, 89%)	Total (772)
Before diagnosis, N. (%)	3 (3.4)	18 (2.6)	21 (2.7)
Genito-urinary	2	1	8
Gastro-intestinal	0	1	1
Breast	0	5	5
Lung	0	1	1
Skin	0	2	2
Non-Hodgkin's lymphoma	1	0	1
Thyroid	0	1	1
Parotid	0	1	1
Vocal cords	0	1	1
Post-diagnosis, N. (%)	7 (7.9)	24 (3.5)	31 (4)
Genito-urinary	1	10	11
Gastro-intestinal	0	6	6
Breast	1	3	4
Lung	2	2	4
Brain	1	0	1
Skin	0	2	2
Chronic lymphocytic leukemia	1	0	1
Non-Hodgkin's lymphoma	1	1	2



Online Supplementary Figure S1. Frequency of malignancies in familial and sporadic MPN according to age at diagnosis of MPN. The cohort of 772 patients was grouped into four categories (1-4) according to the age at MPN diagnosis (<40 years, 40-55 years, 55-65 years, >65 years). Gray bars represent familial cases and the white bars sporadic cases. In patients younger than 40 years at MPN diagnosis the frequency of malignancies was 4.17% among familial cases and 1.12% among sporadic cases (P=0.24); in the second age group (40-50 years) the frequency of malignancies was 8.33% among familial cases and 5.14% among sporadic cases (P=0.51); in patients 55-65 years old the frequency of malignancies was 15% among familial cases and 7.75% among sporadic cases (P=0.28); in patients older than 65 years the frequency of malignancies was 20% among familial cases and 12% among sporadic cases (P=0.31). For each age category the frequency of malignancies in patients with familial MPN was higher than that observed in those with sporadic MPN, however, the difference was not statistically significant.