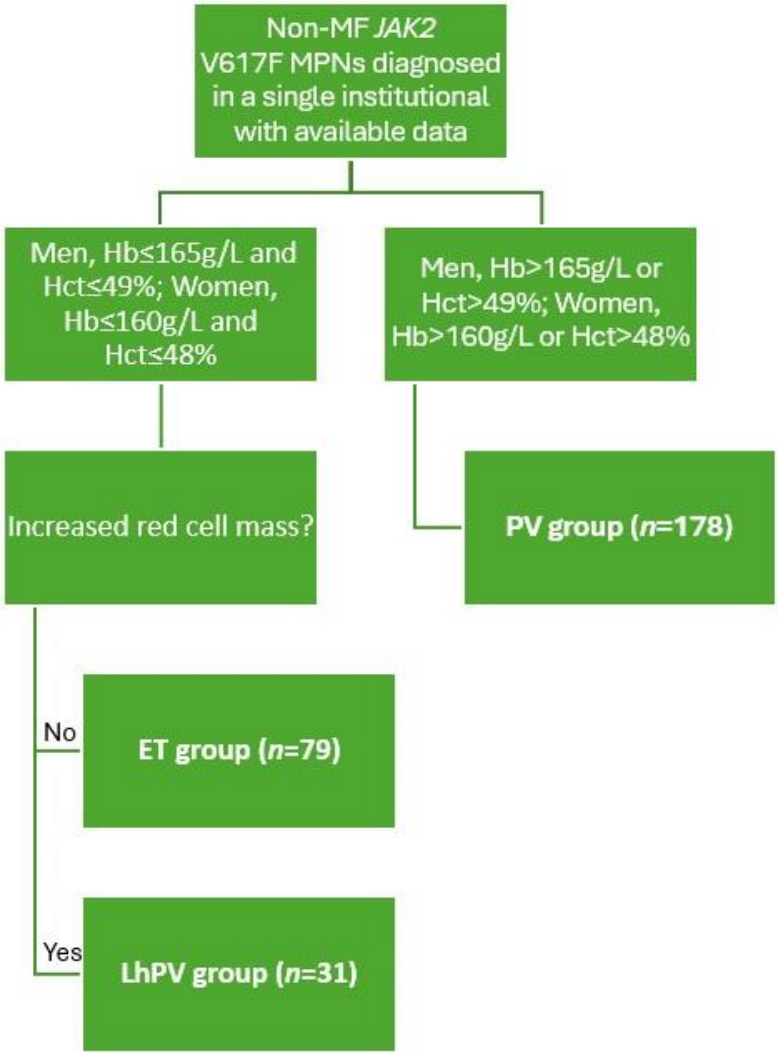
Supplemental Table S1. Bone marrow characteristics across study groups

	Global, *n*=153	ET, *n*=73	lhPV, *n*=15	PV, *n*=65	*P* value
Increased cellularity, *n* (%)	52(35)	16(22)	6(40)	30(48)	0.005
Increased megakaryocytes, *n* (%)	123(82)	61(84)	11(73)	51(81)	0.6
Increased granulocytic lineage, *n* (%)	13(9)	5(7)	0	8(13)	0.3
Increased erythroid lineage, *n* (%)	47(31)	16(22)	2(13)	29(45)	0.004
Megakaryocytes clusters, *n* (%) Loose Dense	 88(60) 23(16)	 48(66) 11(15)	 8(62) 3(23)	 32(52) 9(15)	 0.3 0.7
Megakaryocytes with cloud-like or hyperchromatic nuclei, *n* (%)	22(15)	12(16)	0	10(17)	0.4
Large megakaryocytes with staghorn-like nuclei, *n* (%)	82(57)	45(62)	9(69)	28(47)	0.2
Naked megakaryocyte nuclei, *n* (%)	1(1)	0	0	1(2)	0.5
Small megakaryocytes, *n* (%)	15(10)	3(4)	2(14)	10(16)	0.04
Paratrabecular megakaryocytes, *n* (%)	2(1)	1(1)	1(8)	0	0.2
Sinusoidal hyperplasia, *n* (%)	19(13)	6(8)	2(15)	11(18)	0.2
Intrasinusoidal hematopoiesis, *n* (%)	4(3)	0	1(8)	3(5)	0.06
Bone marrow fibrosis, *n* (%) Grade 0 Grade 1 Grade 2-3	 73(49) 69(47) 6(4)	 35(49) 33(46) 4(5)	 7(50) 6(43) 1(7)	 31(50) 30(48) 1(2)	0.5
Increased CD34+, *n* (%)	3(2)	1(1)	0	2(3)	0.7
Lymphoid aggregates, *n* (%)	16(11)	8(11)	1(8)	7(11)	1
Histologic diagnosis, *n* (%) PV ET Normal Unclassifiable	 37(24) 66(43) 28(18) 22(14)	 4(5) 44(60) 13(18) 12(16)	 5(33) 6(40) 2(13) 2(13)	 28(43) 16(25) 13(20) 8(12)	 <0.001 <0.001 0.9 0.9

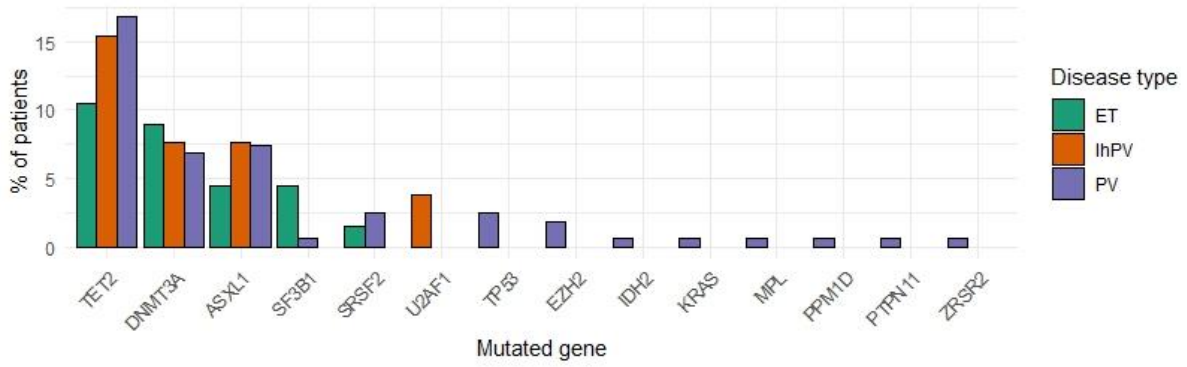
ET, essential thrombocythemia; lhPV, low-hemoglobin polycythemia vera; PV, polycythemia vera. Cellularity and erythroid lineage predominance differed significantly among the groups. A small proportion of patients (4%) had grade 2–3 bone marrow fibrosis at diagnosis; however, none fulfilled the histopathologic criteria for myelofibrosis. These cases were classified histologically as myeloproliferative neoplasm, unclassifiable.

Supplemental Figure S1. Study design



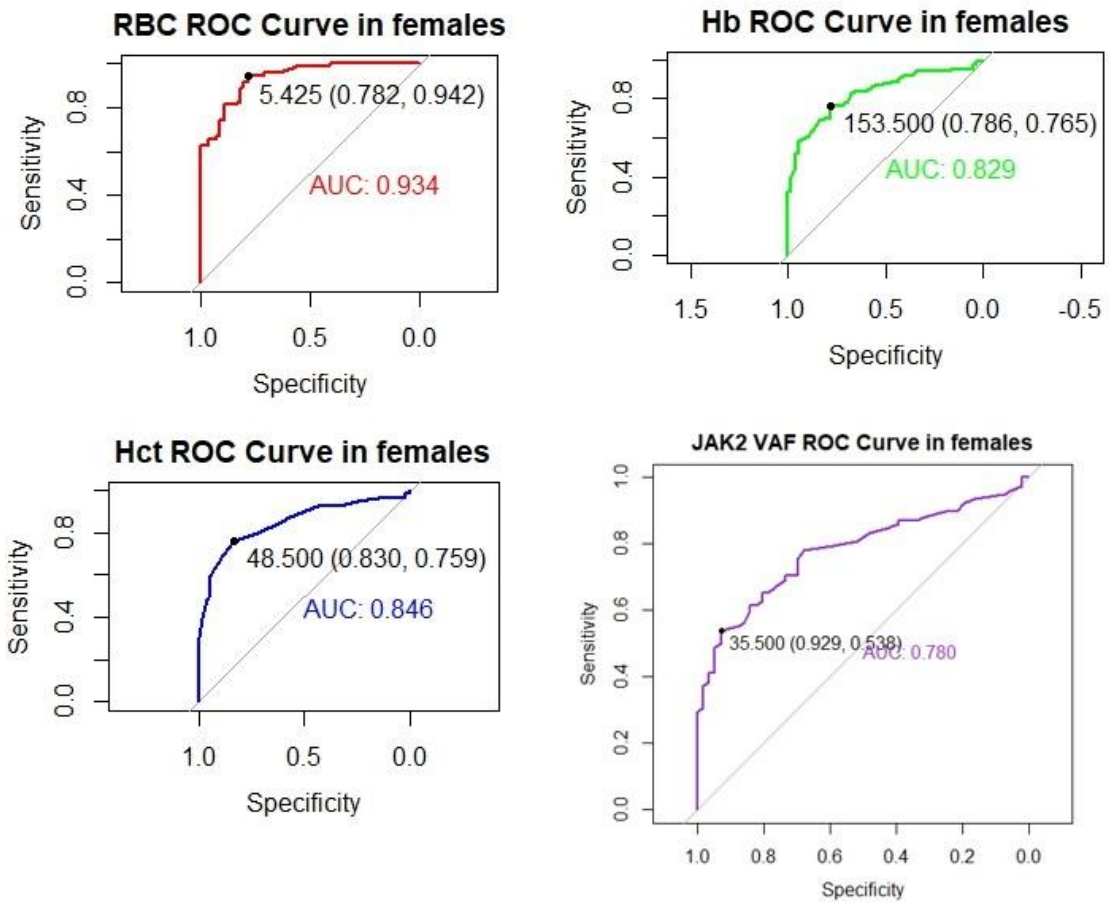
Abbreviations: MF, myelofibrosis; MPNs, myeloproliferative neoplasms; Hb, hemoglobin; Hct, hematocrit; ET, essential thrombocythemia; LhPV, low-hemoglobin polycythemia vera; PV, polycythemia vera

Supplemental Figure S2. Proportion of patients carrying additional mutations across study groups



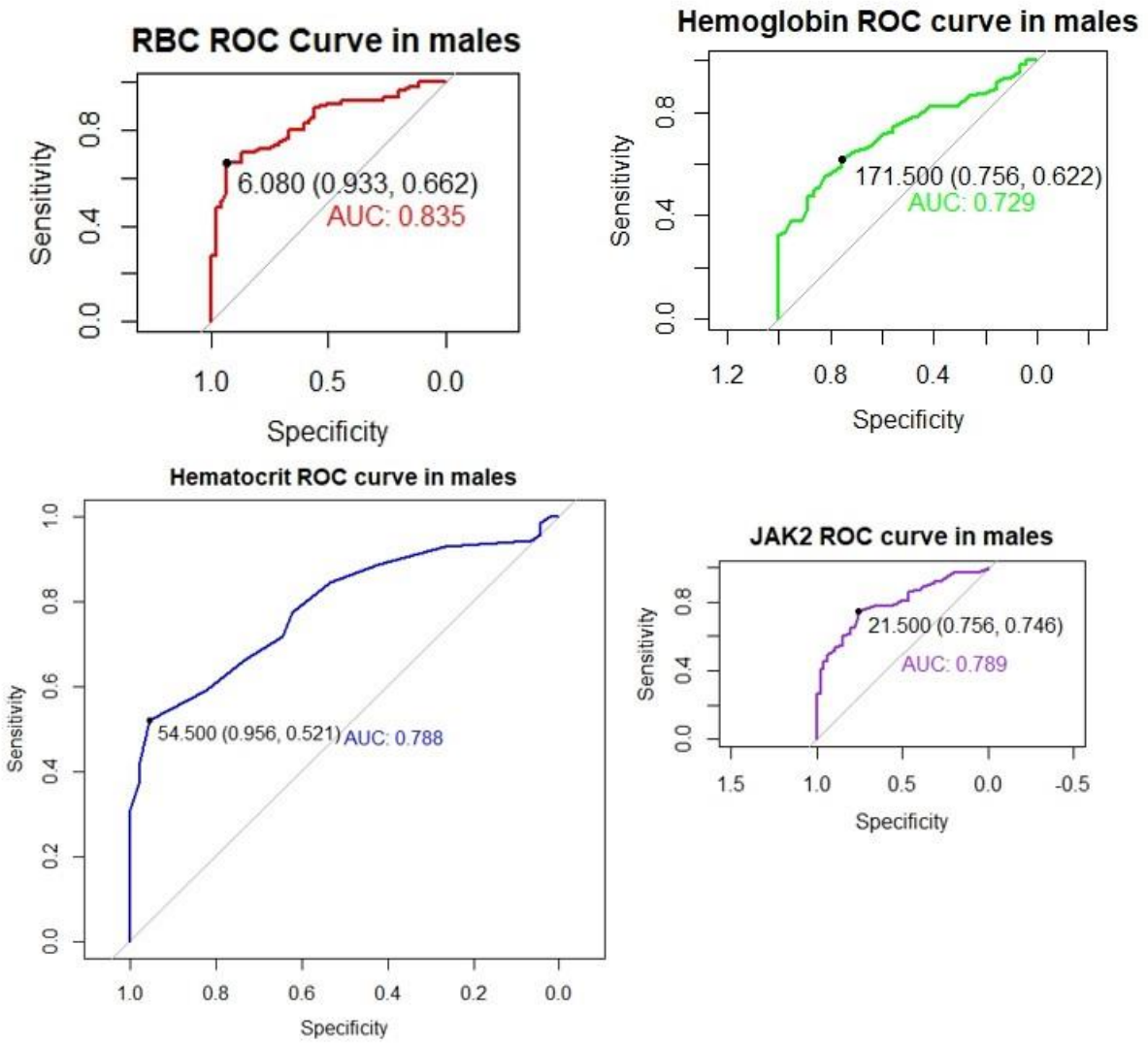
Additional mutations most commonly affect *TET2*, *DNMT3A* and *ASXL1*. Abbreviations: ET, essential thrombocythemia; IhPV, low-hemoglobin polycythemia vera; PV, polycythemia vera.

Supplemental Figure S3. ROC curves of the studied variables, with AUC and Youden Index value in women



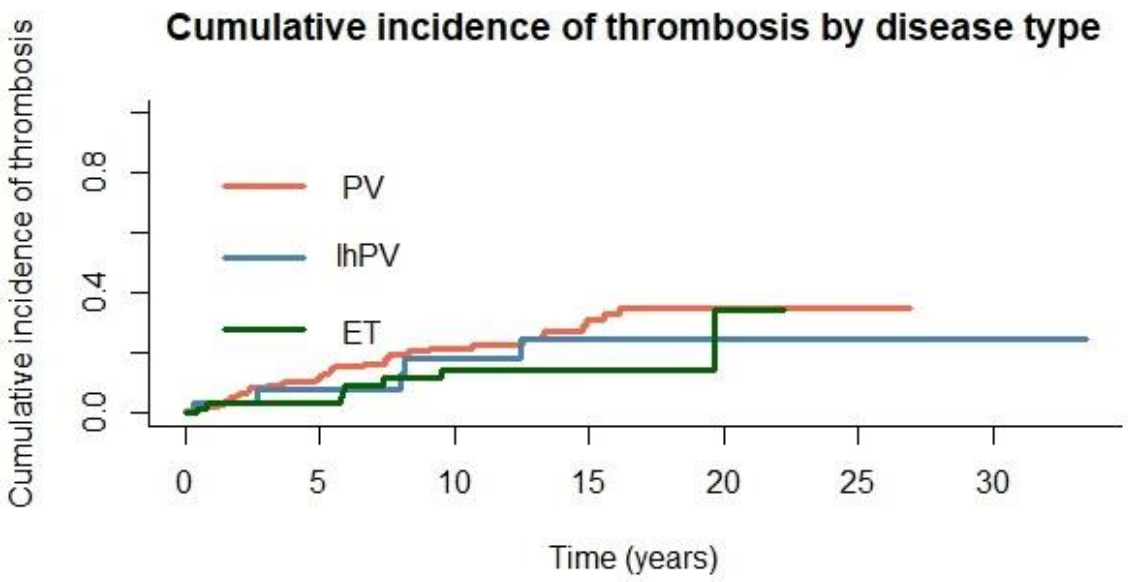
Abbreviations: ROC, receiver operation characteristic; AUC, area under the curve; RBC, red blood cells; Hb, hemoglobin, Hct, hematocrit; VAF, variant allele frequency.

Supplemental Figure S4. ROC curves of the studied variables, with AUC and Youden Index value in men



Abbreviations: ROC, receiver operation characteristic; AUC, area under the curve; RBC, red blood cells; Hb, hemoglobin, Hct, hematocrit; VAF, variant allele frequency.

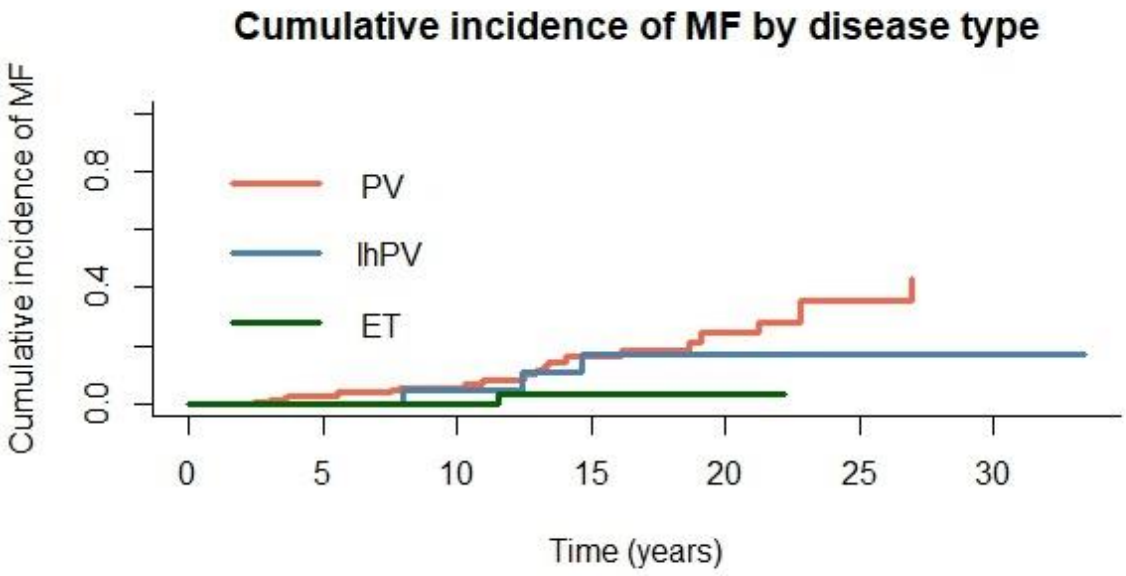
Supplemental Figure S5. Cumulative incidence of thrombosis by disease type



	5 years	10 years	15 years
General	9.8%	18.8%	25.1%
PV	12.6%	21.3%	29.8%
lhPV	7.5%	16.9%	22.3%
ET	4.2%	13.7%	13.7%

Cumulative incidence of thrombosis by disease type, with death as a competing event. Gray's test: $p = 0.17$. Abbreviations: PV, polycythemia vera; lhPV, low-hemoglobin polycythemia vera; ET, essential thrombocythemia

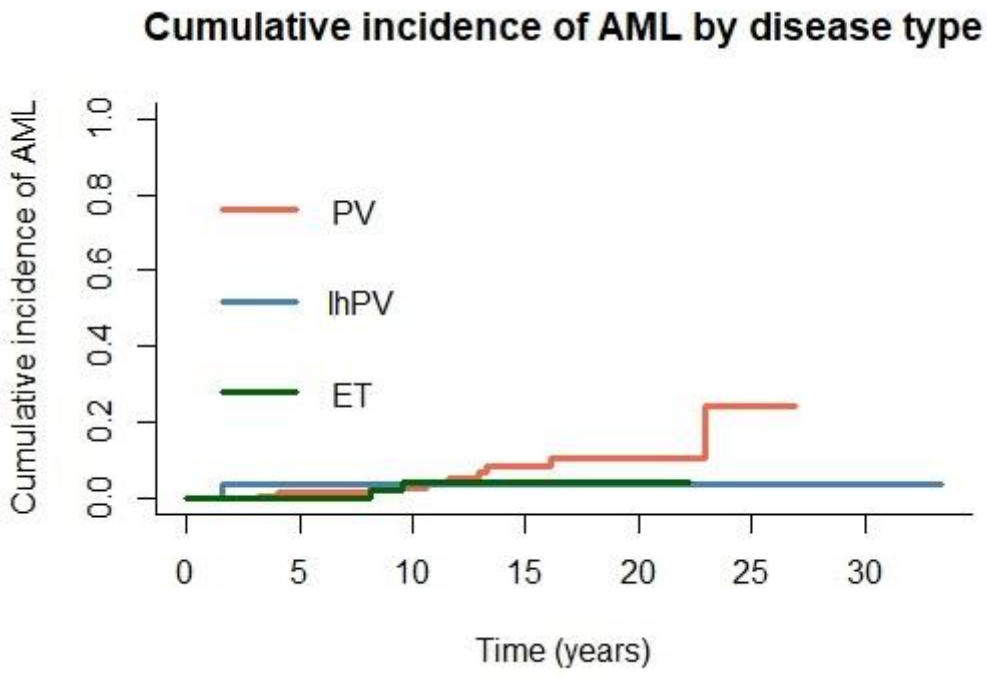
Supplemental Figure S6. Cumulative incidence of myelofibrosis transformation by disease type



	5 years	10 years	15 years
General	1.7%	4%	13.5%
PV	2.9%	5.6%	16.2%
lhPV	0%	4.9%	16.9%
ET	0%	0%	3.3%

Cumulative incidence of myelofibrosis transformation by disease type, with death as a competing event. Gray's test: $p = 0.06$. Abbreviations: MF, myelofibrosis; PV, polycythemia vera; lhPV, low-hemoglobin polycythemia vera; ET, essential thrombocythemia

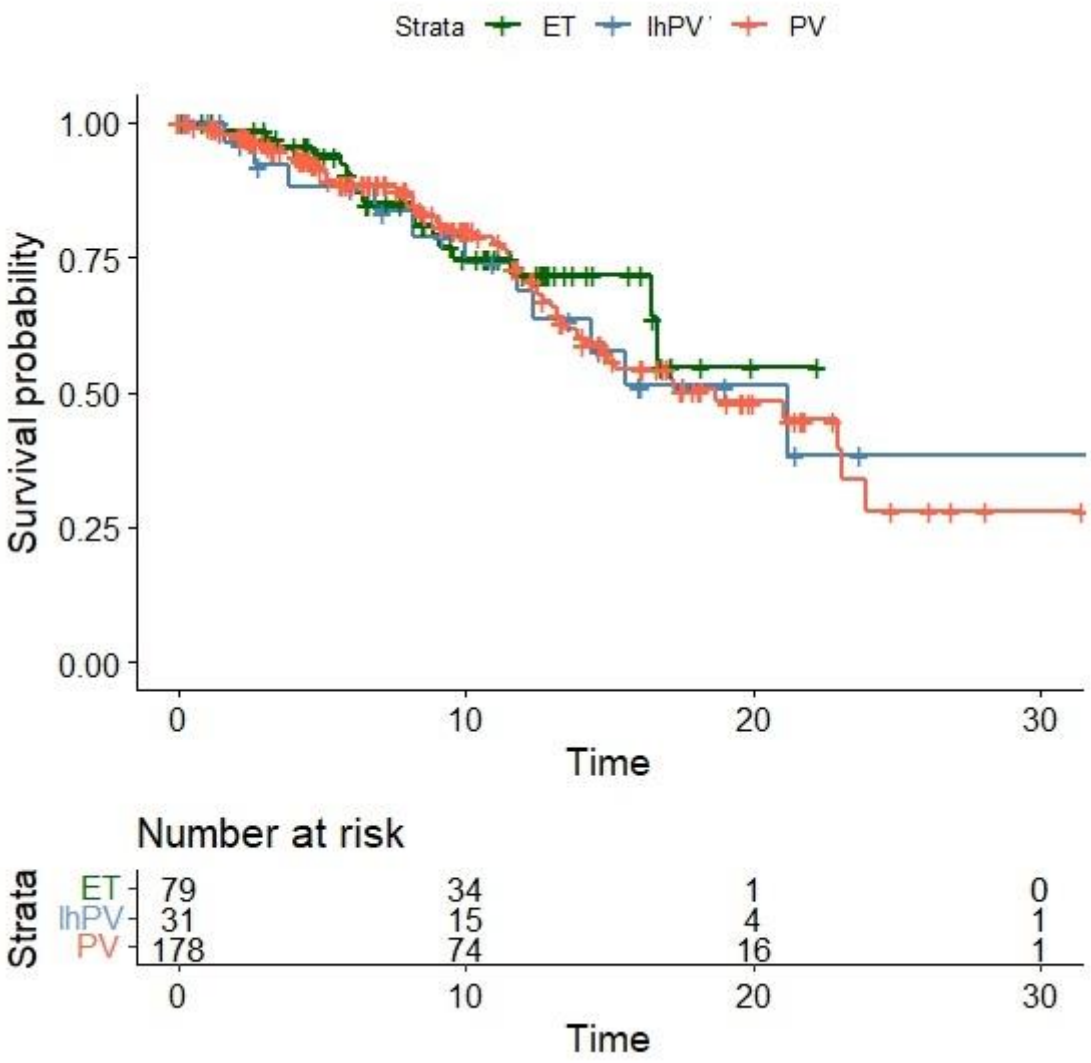
Supplemental Figure S7. Cumulative incidence of leukemic transformation by disease type



	5 years	10 years	15 years
General	1.2%	3.2%	6.9%
PV	1.4%	2.5%	8.5%
lhPV	3.9%	3.9%	3.9%
ET	0%	4.5%	4.5%

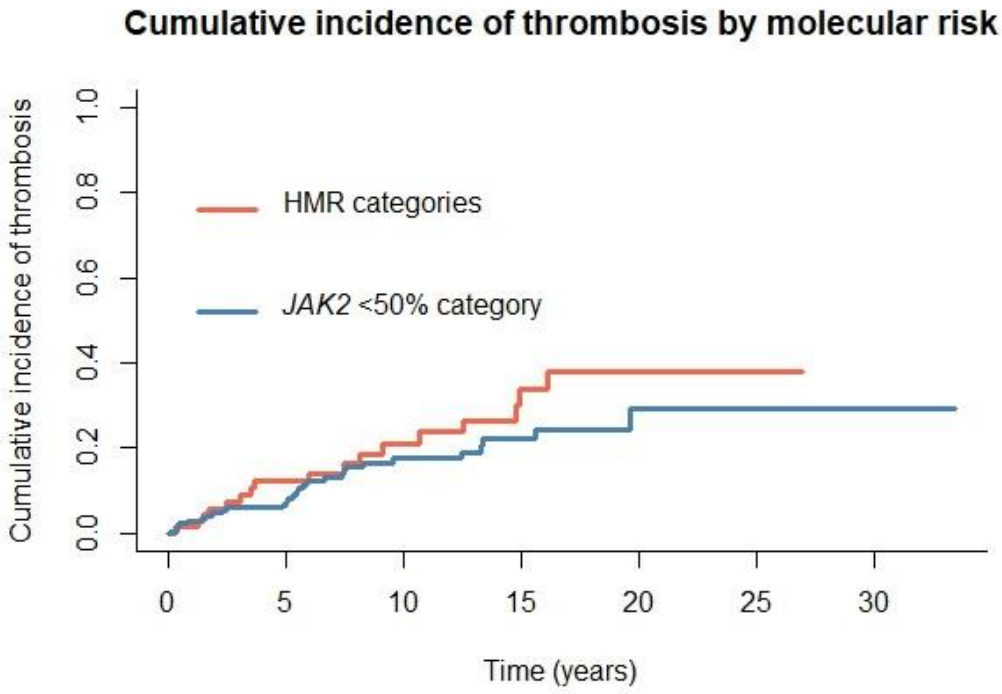
Cumulative incidence of leukemic transformation by disease type, with death as a competing event. Gray's test: $p = 0.62$. Abbreviations: AML, acute myeloid leukemia; PV, polycythemia vera; lhPV, low-hemoglobin polycythemia vera; ET, essential thrombocythemia

Supplemental Figure S8. Overall survival according to disease group



Kaplan–Meier estimates of overall survival in patients with ET, lhPV, and overt PV. Median overall survival was not reached in ET and was 16 and 17 years in the lhPV and PV groups, respectively (log-rank p = 0.085). Abbreviations: ET, essential thrombocythemia; lhPV, low-hemoglobin polycythemia vera; PV, polycythemia vera

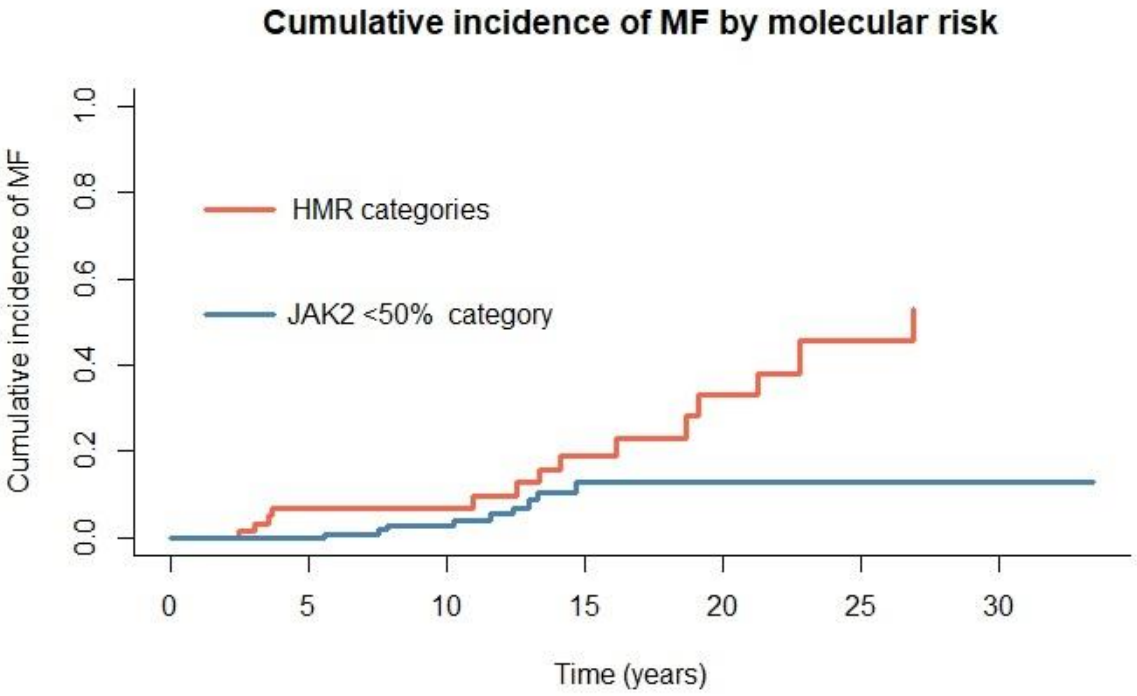
Supplemental Figure S9. Cumulative incidence of thrombosis by molecular risk category



	5 years	10 years	15 years
HMR	12.3%	21.1%	33.9%
JAK2 <50%	7.5%	17.6%	22.1%

Cumulative incidence of thrombosis by molecular risk, with death as a competing event. Gray's test: $p = 0.31$. Abbreviation: HMR, high molecular risk

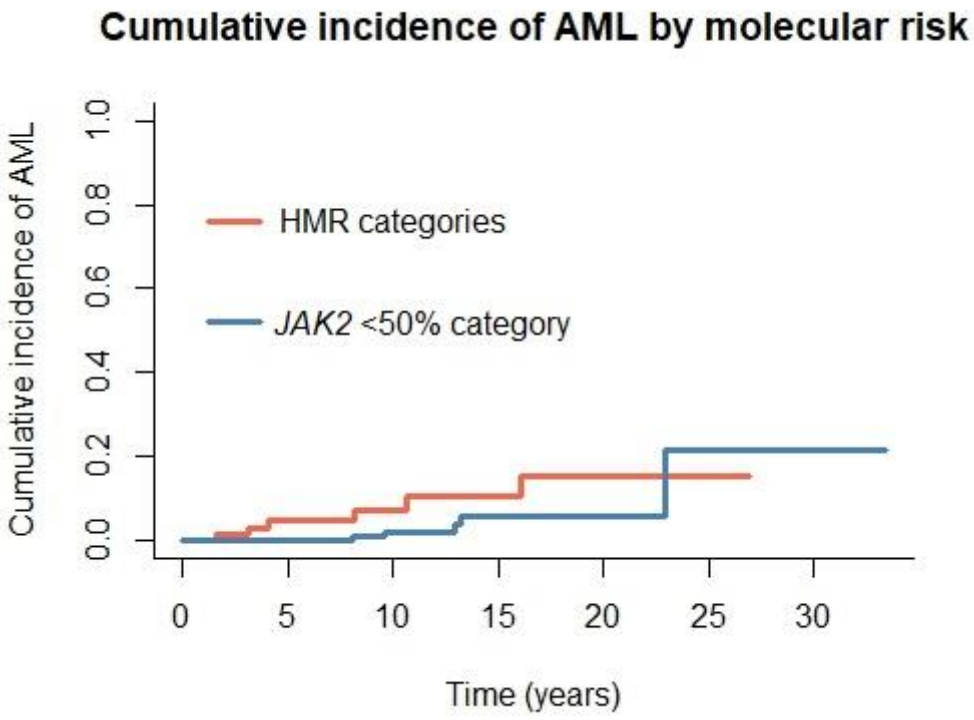
Supplemental Figure S10. Cumulative incidence of myelofibrosis transformation by molecular risk category



	5 years	10 years	15 years
HMR	6.7%	6.7%	18.8%
JAK2 <50%	0%	2.8%	12.7%

Cumulative incidence of myelofibrosis transformation by molecular risk, with death as a competing event. Gray's test: $p = 0.01$. Abbreviation: HMR, high molecular risk

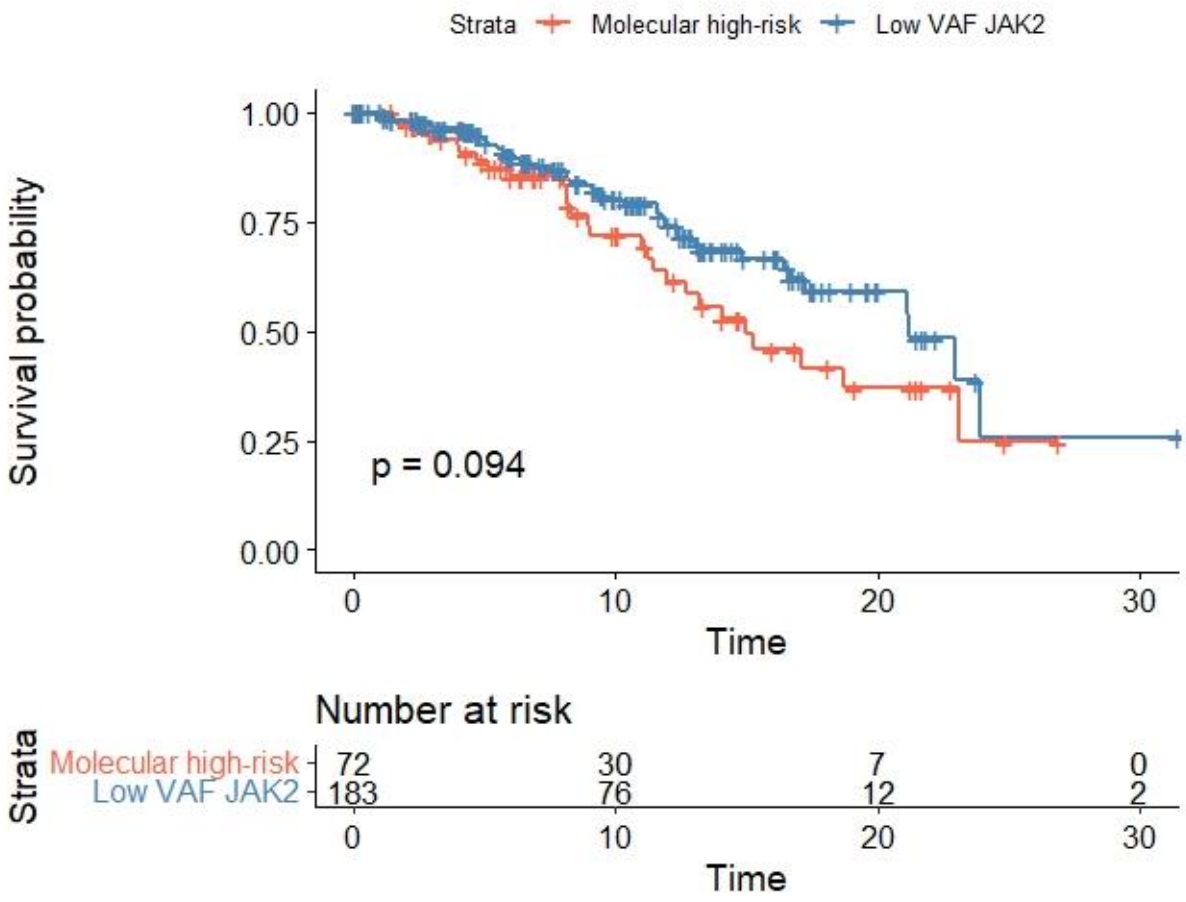
Supplemental Figure S11. Cumulative incidence of leukemic transformation by molecular risk category



	5 years	10 years	15 years
HMR	4.9%	7.3%	10.4%
JAK2 <50%	0%	2.2%	5.8%

Cumulative incidence of leukemic transformation by molecular risk, with death as a competing event. Gray's test: p = 0.09. Abbreviations: AML, acute myeloid leukemia; HMR, high molecular risk

Supplemental Figure S12. Overall survival according to disease group



Kaplan–Meier estimates of overall survival in high molecular risk categories vs *JAK2* <50% category (log-rank *p* = 0.094).

SUPPLEMENTARY INFORMATION

Sample collection and processing

DNA from peripheral blood (PB) samples stored in the Hematopathology Collection (Biobank Hospital Clínic-IDIBAPS; R121004-094) was used for this study. PB samples were obtained in 4ml or 10ml EDTA tubes and processed within 24 hours of extraction. Genomic DNA extraction was performed according to manufacturer instructions (QIAamp DNA kit, Qiagen, Hilden, Germany). DNA quality and concentration were assessed using NanoDrop or Qubit (ThermoFisher Scientific, Waltham, United States). Genomic DNA was stored at -20ºC.

Droplet digital PCR (ddPCR) assay for *JAK2* V617F mutation

Quantification of the *JAK2* V617F mutation was performed by ddPCR using a commercially available assay (Bio-Rad, Hercules, USA). Briefly, 50–100 ng of genomic DNA per sample were analyzed on the QX200 Droplet Digital PCR System (Bio-Rad). Droplets were generated using the QX200 Droplet Generator (Bio-Rad). Following PCR amplification, droplets were read on the QX200 Droplet Reader (Bio-Rad), which measures fluorescence in each droplet using a two-color detection system (FAM and HEX). Data acquisition and analysis were performed using QuantaSoft/QX Manager Software version 2.1 (Bio-Rad).

The variant allele frequency (VAF) was calculated based on a Poisson distribution model, as the ratio of mutant-positive droplets to the total number of target-positive droplets. The fractional abundance (%) of the mutation was calculated as follows:

Fractional abundance (%) = [mutant events / (mutant events + wild-type events)] × 100

Molecular analysis by Next-Generation Sequencing (NGS)

200-400ng of DNA per sample were used for NGS sequencing. Targeted sequencing by NGS was performed using the following commercial panels, which were used depending on availability at the time of testing:

- Customized Myeloid Solution panel MIC_v1, MiSeq System, Illumina Platform; Sophia Genetics, Rolle, Switzerland. This panel included the following 32 genes: *ABL1* (hotspot regions), *ASXL1* (hotspot regions), *BRAF* (hotspot regions), *CALR* (hotspot regions), *CBL* (hotspot regions), *CEBPA* (coding region), *CSF3R* (coding region), *CSNK1A1* (hotspot regions), *DNMT3A* (coding region), *ETV6* (coding region), *EZH2* (coding region), *FLT3* (hotspot regions), *HRAS* (hotspot regions), *IDH1* (hotspot regions), *IDH2* (hotspot regions), *JAK2* (coding region), *KIT* (hotspot regions), *KMT2A* (hotspot regions), *KRAS* (hotspot regions), *MPL* (coding region), *NPM1* (hotspot regions), *NRAS* (hotspot regions), *PTPN11* (hotspot regions), *RUNX1* (coding region), *SETBP1* (hotspot regions), *SF3B1* (hotspot regions), *SRSF2* (hotspot regions), *TET2* (coding region), *TP53* (coding region), *U2AF1* (hotspot regions), *WT1* (hotspot regions) and *ZRSR2* (coding region).
- Oncomine Myeloid Assay, Ion GeneStudio S5 System, Ion Torrent platform; ThermoFisher Sicentific, Waltham, United States. This panel included the following 40 genes: *ABL1* (hotspot regions), *ASXL1* (coding region), *BCOR* (coding region), *BRAF* (hotspot regions), *CALR* (coding region), *CBL* (hotspot regions), *CEBPA* (coding region), *CSF3R* (hotspot regions), *DNMT3A* (hotspot regions), *ETV6* (coding region), *EZH2* (coding region), *FLT3* (hotspot regions), *GATA2* (hotspot regions), *HRAS* (hotspot regions), *IDH1* (hotspot regions), *IDH2* (hotspot regions), *IKZF1* (coding region), *JAK2* (hotspot regions), *KIT* (hotspot regions), *KRAS* (hotspot regions), *MPL* (hotspot regions), *MYD88* (hotspot regions), *NF1* (coding region), *NPM1* (hotspot regions), *NRAS* (hotspot regions), *PHF6* (coding region), *PRPF8* (coding region), *PTPN11* (hotspot regions), *RB1* (coding region), *RUNX1* (coding region), *SETBP1*

(hotspot regions), *SF3B1* (hotspot regions), *SH2B3* (coding region), *SRSF2* (hotspot regions), *STAG2* (coding region), *TET2* (coding region), *TP53* (coding region), *U2AF1* (hotspot regions), *WT1* (hotspot regions) and *ZRSR2* (coding region).

- Oncomine™ Myeloid Assay GX v2, Ion Torrent Genexus System; ThermoFisher Scientific, Waltham, United States. This panel included the following 45 genes: *ABL1* (hotspot regions), *ANKRD26* (hotspot regions), *ASXL1* (coding region), *BCOR* (coding region), *BRAF* (hotspot regions), *CALR* (coding region), *CBL* (hotspot regions), *CEBPA* (coding region), *CSF3R* (hotspot regions), *DDX41* (hotspot regions), *DNMT3A* (hotspot regions), *ETV6* (coding region), EZH2 (coding region), *FLT3* (hotspot regions), *GATA2* (hotspot regions), *HRAS* (hotspot regions), *IDH1* (hotspot regions), *IDH2* (hotspot regions), *IKZF1* (coding region), *JAK2* (hotspot regions), *KIT* (hotspot regions), *KRAS* (hotspot regions), *MPL* (hotspot regions), *MYD88* (hotspot regions), *NF1* (coding region), *NPM1* (hotspot regions), *NRAS* (hotspot regions), *PHF6* (coding region), *PPM1D* (hotspot regions), *PRPF8* (coding region), *PTPN11* (hotspot regions), *RB1* (coding region), *RUNX1* (coding region), *SETBP1* (hotspot regions), *SF3B1* (hotspot regions), *SH2B3* (coding region), *SMC1A* (hotspot regions), *SMC3* (hotspot regions), *SRSF2* (hotspot regions), *STAG2* (coding region), *TET2* (coding region), *TP53* (coding region), *U2AF1* (hotspot regions), *WT1* (hotspot regions) and *ZRSR2* (coding region).

GRCh37 (hg19) was used as the reference genome. Variant calling and analysis were carried out using the SOPHia DDM™ platform, Ion Reporter Software, or Ion Torrent Genexus System. Only pathogenic and likely pathogenic variants with a variant allele frequency (VAF) of ≥2% were considered for further analysis.

Genomic classification proposed by Grinfeld et al. (1) was applied to all cases based on mutational data obtained by NGS and cytogenetic information on medical records. Patients were hierarchically assigned to four molecular subgroups: (1) *TP53* disruption or aneuploidy (*TP53* mutations, chr17p loss of heterozygosity [LOH], or chr5-/5q-); (2) ≥1 genetic aberration

in chromatin or spliceosome genes (including *EZH2*, *IDH1*, *IDH2*, *ASXL1*, *PHF6*, *CUX1*, *ZRSR2*, *SRSF2*, *U2AF1*, *KRAS*, *NRAS*, *GNAS*, *CBL*, chr7/7q LOH, chr4q LOH, *RUNX1*, *STAG2*, and *BCOR*); (3) homozygous *JAK2* mutation VAF ≥50% or chr9p LOH when available); and (4) heterozygous *JAK2* mutation (VAF <50%).

References:

1. Grinfeld J, Nangalia J, Baxter EJ, Wedge DC, Angelopoulos N, Cantrill R, et al. Classification and Personalized Prognosis in Myeloproliferative Neoplasms. New England Journal of Medicine. 2018;379(15):1416–30. doi:10.1056/nejmoa1716614 PubMed PMID: 30304655.