Hematological alterations associated with the SNV rs10974944, part of the 46/1 haplotype, in patients from the Brazilian Amazon with BCR::ABL1-negative myeloproliferative neoplasms

Jhemerson F. Paes  
Programa de Pós-Graduação em Ciências Aplicadas à Hematologia, Universidade do Estado do Amazonas (UEA)

Dania G. Torres  
Programa de Pós-Graduação em Ciências Aplicadas à Hematologia, Universidade do Estado do Amazonas (UEA)

Deborah C. Aquino  
Programa de Pós-Graduação em Ciências Aplicadas à Hematologia, Universidade do Estado do Amazonas (UEA)

Emanuela V. B. Alves  
Programa de Pós-Graduação em Ciências Aplicadas à Hematologia, Universidade do Estado do Amazonas (UEA)

Erycka A. Mesquita  
Programa de Pós-Graduação em Ciências Aplicadas à Hematologia, Universidade do Estado do Amazonas (UEA)

Milane A. Sousa  
Programa de Pós-Graduação em Ciências Aplicadas à Hematologia, Universidade do Estado do Amazonas (UEA)

Nelson Abrahim Fraiji  
Programa de Pós-Graduação em Ciências Aplicadas à Hematologia, Universidade do Estado do Amazonas (UEA)

Leny N. M. Passos  
Programa de Pós-Graduação em Ciências Aplicadas à Hematologia, Universidade do Estado do Amazonas (UEA)

Rosângela S. Abreu  
Programa de Pós-Graduação em Ciências Aplicadas à Hematologia, Universidade do Estado do Amazonas (UEA)

George A. V. Silva  
Programa de Pós-Graduação em Ciências Aplicadas à Hematologia, Universidade do Estado do Amazonas (UEA)

Andréa M. Tarragô  
Programa de Pós-Graduação em Ciências Aplicadas à Hematologia, Universidade do Estado do Amazonas (UEA)

Lucivana P. de Souza Mourão (lpsouza@uea.edu.br)  
Programa de Pós-Graduação em Ciências Aplicadas à Hematologia, Universidade do Estado do Amazonas (UEA)

Article

Keywords:

Posted Date: February 2nd, 2024

DOI: https://doi.org/10.21203/rs.3.rs-3880113/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License

Additional Declarations: No competing interests reported.
Abstract

BCR::ABL1-negative myeloproliferative neoplasms are hematopoietic disorders characterized by panmyelosis. \textit{JAK2 V617F} is a frequent variant in these diseases and often occurs in the 46/1 haplotype. The G allele of \textit{rs10974944} has been shown to be associated with this variant, specifically its acquisition, correlations with familial cases, and laboratory alterations. This study evaluated the association between the 46/1 haplotype of \textit{JAK2} in patients with myeloproliferative neoplasms in a population from the Brazilian Amazon. Clinical, laboratory and molecular sequencing analyses were considered. Carriers of the G allele of \textit{rs10974944} with polycythemia vera showed an increase in mean corpuscular volume and mean corpuscular hemoglobin, while in those with essential thrombocythemia, there was an elevation in red blood cells, hematocrit, and hemoglobin. Associations were observed between \textit{rs10974944} and the \textit{JAK2 V617F}, in which the G allele (OR: 3.47; \textit{p} < 0.0001), CG genotype (OR: 8.4; \textit{p} = 0.002), and GG genotype (OR: 4.1; \textit{p} = 0.002) were associated with \textit{JAK2 V617F+} and an increase in variant allele frequency (GG: OR 13.1; \textit{p} = 0.004; G: OR: 6.0; \textit{p} = 0.0002). These results suggest an association between \textit{rs10974944} (G) and a status for \textit{JAK2 V617F+}_VAF ≥ 50%, and laboratory alterations in the erythroid lineage.

Introduction

Myeloproliferative neoplasms (MPNs) are clonal diseases characterized by hyperplasia of the myeloid lineage with effective maturation, which results in leukocytosis in peripheral blood, increased erythrocyte mass and possible progression to medullary fibrosis or leukemic transformation\cite{1}. They have an incidence rate of 6 cases per 100,000 individuals and mostly affect white males between 60 and 70 years of age\cite{2}. Polycythemia vera (PV), essential thrombocythemia (ET), and primary myelofibrosis (MF) are the most common \textit{BCR::ABL1}-negative MPNs, though differ in signs, symptoms, hematological and clinical alterations, and genetic findings\cite{3}.

\textit{JAK2 V617F} (dbSNP ID: \textit{rs77375493}) is the main genetic finding in MPNs and has a frequency of 95% in PV cases and between 50%-60% in ET and MF cases\cite{4}. This somatic variant triggers the substitution of valine by phenylalanine at codon 617, which alters the pseudo-kinase domain of the \textit{JAK2} protein and conditions a constitutive activation of the JAK/STAT signaling pathway\cite{5}.

Studies have shown a significant correlation between \textit{JAK2 V617F} and the 46/1 haplotype, a set of germline genetic variations distributed along chromosome 9p.24.1. This haplotype covers regions with a high number of genetic variants in \textit{JAK2} (exons 12 and 14) and is in linkage disequilibrium with the variant \textit{rs10974944} (C > G), located in intron 12 of the same gene\cite{6} (Fig. 1). Studies indicate that this genetic alteration is a factor that favors the acquisition of \textit{JAK2 V617F} by increasing the mutational rate of \textit{JAK2}, which can lead to DNA damage and replication errors\cite{7-9}. In addition to being identified in MPN patients of various populations, this haplotype has also been associated with more pronounced alterations in laboratory exams, presence of splenomegaly, inflammatory dysregulation, familial cases of MPNs (increasing the risk of developing any myeloproliferative neoplasm by 5 to 7 times) and abnormal methylation of the gene promoter\cite{10-13}. Therefore, the \textit{JAK2 46/1} haplotype confers predisposition to the development of myeloproliferative neoplasms associated with the \textit{JAK2 V617F} mutation (OR = 3.7; 95% CI = 3.1–4.3) and provides a conceptual framework in which a constitutional genetic component is associated with a substantial increase in the risk of acquiring a specific somatic mutation\cite{14}.

In this study, we performed genetic sequencing of intron 12 of the \textit{JAK2} gene to identify the \textit{rs10974944} variant (C > G), in strong linkage disequilibrium with the 46/1 haplotype, in 100 patients with \textit{BCR::ABL1}-negative myeloproliferative neoplasms (polycythemia vera: \textit{n} = 39; essential thrombocythemia: \textit{n} = 61) for whom clinical and laboratory information was available for clinical and laboratory characterization.

RESULTS

Characterization of the study population
The study included individuals clinically diagnosed with polycythemia vera (PV) \( (n = 39) \) or essential thrombocythemia (ET) \( (n = 61) \), whose clinical-laboratory characteristics are presented in the supplementary material. The female gender was more prevalent among individuals diagnosed with ET \( (n = 48, p = 0.002) \). The median age of the participants ranged between the fifth and sixth decades of life \( (p = 0.441) \).

Regarding hematological results, the medians of overall red blood cell count (RBC), hematocrit (Ht), hemoglobin (Hb), and total white blood cell count (WBC) were significantly higher in the PV group compared to the ET group \( (p < 0.05) \) (see Table SI). Other hematological markers, such as mean corpuscular volume \( (103.9 \text{ pg}, p < 0.0001) \), mean corpuscular hemoglobin \( (33.5 \text{ fl}, p < 0.0001) \), and overall platelet count \( (467,000 \times \text{cells/mm}^3, p < 0.0001) \), were also significantly elevated in the ET group compared to the PV group. Hemorrhagic events were more frequent in patients with ET compared to PV \( (p = 0.003) \), while the frequency of splenomegaly and thrombotic events did not differ significantly between PV and ET \( (p > 0.05) \) groups.

Regarding genetic findings, the presence of \( JAK2 \ V617F^+ \) was more frequent in patients with PV \( (58.9\%, p = 0.020) \) (Fig. 2a), and a variant allele frequency (VAF) of \( \geq 50\% \) was also more common in patients with this hematologic condition \( (41\%, p = 0.005) \) (Fig. 2b).

A greater frequency of patients with ET \( (95.1\%, p < 0.0001) \) received cytoreductive treatment in comparison to PV patients \( (66.6\%) \).

**Identified genetic variants**

Data on the allelic and genotypic frequency of \( rs10974944 \) (C > G) are presented in Figs. 2c and 2d. Of all the individuals included in the study, 63\% exhibited the \( rs10974944 \) variant (G): 26\% in homozygosity (GG) and 37\% in heterozygosity (CG). The GG genotype of \( rs10974944 \) was more prevalent in the PV group (36\%), whereas CG was more homogeneous between the groups (33.3\% in PV and 39.3\% in ET). Regarding allelic frequency, the G allele was more frequent in the PV (53.6\%) group, and the wild-type allele proved to be more prevalent in the ET (60.7\%) group.

Table 1 presents the hematological data of individuals with polycythemia vera and essential thrombocythemia stratified according to the absence or presence of the \( rs10974944 \) (CC and G carriers, respectively). In PV, G carriers showed significantly increased values for MCV and MCH \( (p = 0.030 \text{ and } p = 0.041, \text{respectively}) \), while in ET, patients with the variant exhibited elevated indices of RBC, Ht, and Hb with demonstrated statistical significance \( (p < 0.05) \).
Table 1: Laboratory characteristics of G carriers (rs10974944) and individuals without the variant who were diagnosed with polycythemia vera or essential thrombocythemia. Abbreviations: RBC: red blood cell count, Ht: hematocrit, Hb: hemoglobin, MCV: mean corpuscular volume, MCH: mean corpuscular hemoglobin, MCHC: mean corpuscular hemoglobin concentration, WBC: white blood cell count. Reference values: RBC: 3.9–5.3 million/mm$^3$, Ht: 36–48%, Hb: 12–16 g/dL, MCV: 80–100 fL, MCH: 27–33 pg, MCHC: 32–36 g/dL, WBC: 3,600 – 11,000 cells/mm$^3$, Platelets: 150,000-400,000 per mm$^3$.

Distribution of variants in patients stratified according to JAK2 V617F status and variant allele frequency

Considering the association of rs10974944 with JAK2 V617F, the genotypic frequency analysis of rs10974944 (C > G) was performed according to the positive (+) or negative (-) status of JAK2 V617F and its variant allele frequency (VAF), with data described in Table 2. Homozygous individuals for rs10974944 (GG) showed a significantly higher frequency of JAK2 V617F$^+$ status and a higher likelihood of being positive for this variant when compared to the CC genotype (42.2% vs 20%; OR: 8.4; 95% CI 2.6–24.8; p = 0.002). The same was true for the GG/CG genotypes (42.2%/37.8% vs 20%; OR 4.1: 95% CI 1.6–9.7; p = 0.002) and the G allele. We emphasize the association of the rs10974944 G allele with the V617F variant, which demonstrated a 3.4-fold higher probability of being present in JAK2 V617F$^+$ individuals compared to individuals carrying the C allele (61.1% vs 38.9%; OR 3.47; 95% CI 1.9–6.2; p < 0.0001).

Additionally, the analyses revealed that individuals with the GG genotype of rs10974944 had a 13.1-fold higher probability of having a VAF greater than 50% when compared to individuals with the CC genotype (75% vs 15%; OR 13.1; 95% CI 1.8–72.3; p = 0.004). Regarding the allele, carriers of the G allele showed a 6-fold higher risk of having a VAF of ≥ 50% compared to the wild-type allele (C) (82.5% vs 17.5%; OR: 6.0; 95% CI: 2.1–14.8; p = 0.0002). These results demonstrate an association between rs10974944 and the variation in VAF in JAK2 V617F.

Table 2: Distribution of single nucleotide variants (SNVs) in MPN patients stratified by JAK2 V617F status and variant allele frequency. Abbreviations: VAF: variant allele frequency. *In linkage disequilibrium with haplotype 46/1.

<table>
<thead>
<tr>
<th>Genotype/allele</th>
<th>JAK2 V617F Status (n = 93)</th>
<th>JAK2 V617F$^+$ VAF (n = 45)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive (n = 45)</td>
<td>Negative (95% CI)</td>
</tr>
<tr>
<td>rs10974944*; n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>9 (20)</td>
<td>24 (50)</td>
</tr>
<tr>
<td>CG</td>
<td>18 (37.8)</td>
<td>18 (37.5)</td>
</tr>
<tr>
<td>GG</td>
<td>19 (42.2)</td>
<td>6 (12.5)</td>
</tr>
<tr>
<td>C</td>
<td>35 (38.9)</td>
<td>66 (68.7)</td>
</tr>
<tr>
<td>G</td>
<td>55 (61.1)</td>
<td>30 (31.2)</td>
</tr>
</tbody>
</table>

Identified haplotypes

The linkage disequilibrium (LD) of rs10974944 and JAK2 V617F (rs77375493) is demonstrated in Fig. 3. Variants identified in the analyzed region were included in the haplotype analysis. When these genetic alterations are paired, they give rise to eight
haplotypes (Table 3). Haplotype analysis revealed that only haplotype 2
(rs10974944G/rs10815151C/rs10119004A/rs77375493T) was more prevalent in individuals with the PV phenotype (33.3%; OR: 3.3; 95% CI: 1.2–9.2; p = 0.0006), indicating that it is a possible marker associated with PV_JAK2 V617F+

Table 3
Haplotypes of JAK2 intron 12 present in individuals with polycythemia vera (PV) or essential thrombocythemia (ET).

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>rs10974944</th>
<th>rs10815151</th>
<th>rs10119004</th>
<th>rs77375493</th>
<th>PV</th>
<th>ET</th>
<th>Chi-Square</th>
<th>Odds ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>G</td>
<td>6 (15.3%)</td>
<td>17 (27.8%)</td>
<td>3.246</td>
<td>0.4 (0.1–1.2)</td>
<td>0.07</td>
</tr>
<tr>
<td>2</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>13 (33.3%)</td>
<td>8 (13.1%)</td>
<td>11.918</td>
<td>3.3 (1.2–9.2)</td>
<td>0.0006</td>
</tr>
<tr>
<td>3</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>G</td>
<td>6 (15.3%)</td>
<td>14 (22.9%)</td>
<td>1.894</td>
<td>0.6 (0.2–1.7)</td>
<td>0.168</td>
</tr>
<tr>
<td>4</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>6 (15.3%)</td>
<td>12 (21.3%)</td>
<td>0.495</td>
<td>0.7 (0.2–2.2)</td>
<td>0.481</td>
</tr>
<tr>
<td>5</td>
<td>C</td>
<td>C</td>
<td>A</td>
<td>G</td>
<td>1 (2.5%)</td>
<td>3 (4.9%)</td>
<td>1.006</td>
<td>0.5 (0.03–3.5)</td>
<td>0.315</td>
</tr>
<tr>
<td>6</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>2 (5.1%)</td>
<td>2 (3.2%)</td>
<td>1.14</td>
<td>1.5 (0.2–10.4)</td>
<td>0.285</td>
</tr>
<tr>
<td>7</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>2 (5.1%)</td>
<td>2 (3.2%)</td>
<td>0.28</td>
<td>1.5 (0.2–10.4)</td>
<td>0.597</td>
</tr>
<tr>
<td>8</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>G</td>
<td>1 (2.5%)</td>
<td>1 (1.6%)</td>
<td>0.035</td>
<td>1.6 (0.08–31.2)</td>
<td>0.851</td>
</tr>
<tr>
<td>9</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>G</td>
<td>1 (2.5%)</td>
<td>1 (1.6%)</td>
<td>0.004</td>
<td>1.6 (0.08–31.2)</td>
<td>0.949</td>
</tr>
</tbody>
</table>

Discussion

Myeloproliferative neoplasms have characteristic alterations in laboratory exams, as well as genetic findings that permit their identification and differentiation. Findings involving genetic alterations in introns are not yet fully understood, but this scenario is becoming of increasing interest for understanding the etiopathogenic aspects and the role of these DNA regions in these diseases.

Essential thrombocythemia proved to be the most frequent myeloproliferative neoplasm, which are findings that align with the premises established by Torres\textsuperscript{15}, who studied a population with BCR::ABL1-negative myeloproliferative neoplasms in the state of Amazonas (Brazil). Similar data were described by Macedo\textsuperscript{16}, who reported a similar scenario in patients from the states of Paraná and São Paulo who had the same hematologic malignancy, and these data converge with descriptions found in other countries\textsuperscript{17,18}

The age range of individuals was between the fifth and seventh decades of life, which is consistent with what is stated in other studies\textsuperscript{19,20}. The progressive accumulation of genetic variations in hematopoietic stem cells and the biological machinery of the
DNA repair system\textsuperscript{21,22}, an increase or decrease in telomeres\textsuperscript{23,24} and cumulative exposure to risk factors throughout life, such as smoking and obesity\textsuperscript{25,26}, may explain the prevalence of this age group in the context of myeloproliferative neoplasms.

Regarding clinical characteristics, polycythemia vera (PV) showed an equal proportion of men and women, while essential thrombocythemia (ET) revealed a majority of cases involving women, and these data are in line with the literature\textsuperscript{27,28}. Some studies have demonstrated that women have an increased risk of developing myeloproliferative neoplasms\textsuperscript{29} and a higher likelihood of developing cardiovascular complications and splenomegaly\textsuperscript{26}. The reason for this risk is uncertain, but changes in sex chromosomes, hormonal factors and gene expression may be possible contributors to this process\textsuperscript{28}. Laboratory data, and thrombotic and hemorrhagic events presented as expected for each neoplasm: PV demonstrated a higher prevalence of increased erythrogram values and ET showed changes in the megakaryocytic series, with a higher risk of hemorrhagic events, as described by the World Health Organization\textsuperscript{3}, and in other studies on the subject\textsuperscript{27,30}.

Regarding the genetic findings, PV demonstrates a higher prevalence of positive cases for the JAK2 V617F variant, since it is directly associated with the specific pathogenesis of this hematologic malignancy\textsuperscript{36} and plays a role in the constitutive activation of the JAK-STAT pathway\textsuperscript{5}. It is interesting to note that 58% of our PV population was positive for the variant, which may initially differ from findings commonly described in the literature that point to JAK2 V617F frequencies of over 70% in Brazilian, Korean, Chinese, Japanese, and European patients\textsuperscript{31–35}.

The limited number of PV JAK2 V617F\textsuperscript{+} patients identified in this study is related to cytoreductive therapy. We emphasize that 66.6% of patients with PV were on cytoreductive therapy, which acts to suppress and/or decrease the variant burden of JAK2 V617F through the inhibition of the myeloproliferative process of mutated hematopoietic cells, as noted in a recent study\textsuperscript{36}. This directly affects the sensitivity of the molecular detection methods used to identify the variant, underscoring the increased importance of incorporating molecular analysis for JAK2 V617F in the initial suspicions for MPNs.

In the literature, the germline haplotype 46/1, identified by the rs10974944 (C > G) variant, has a well-documented association with JAK2 V617F\textsuperscript{14,37–39} as also observed in our study. The high frequency of the G allele of rs10974944 in individuals positive for JAK2 V617F contributes to discussions about the non-random correlation between these two genetic alterations\textsuperscript{13,40} This relationship is in line with another finding from our study, haplotype 2 (rs10974944G/rs10815151C/rs1011004A/rs77375493T), which strengthens concepts based on the interaction between rs10974944 (C > G) and JAK2 V617F (rs77375493C > G). These propositions are in agreement with findings involving haplotype 46/1 in other Brazilian, Taiwanese, European, Chinese, and Japanese populations\textsuperscript{16,32–34,41}, indicating that the possible mechanisms preceding the acquisition of JAK2 V617F are not limited to a specific ethnic group; therefore, its evolutionary basis can be considered as a genetic predisposition factor for the disease\textsuperscript{8}.

Studies report a higher risk of individuals with the GG genotype of rs10974944 being positive for JAK2 V617F\textsuperscript{14,40,42}. Consistent with the results of the aforementioned studies, our population exhibited a four-fold increase in the risk of positive JAK2 V617F in individuals with the GG genotype of rs10974944. (OR: 4.1; 95% CI: 8-13.9). These findings support the hypothesis of hypermutability, which establishes haplotype 46/1 as a dysregulating agent of the JAK2 gene, which increases the risk of DNA replication errors and conditions a mutagenic scenario for the acquisition of variants with selective advantages, such as JAK2 V617F\textsuperscript{31–35}.

The association of rs10974944 (G) and the JAK2 V617F VAF suggests a possible involvement of haplotype 46/1 in clonal expansion. We identified a six-fold higher risk of individuals carrying the G allele of rs10974944 and JAK2 V617F VAF of ≥ 50%. Our data indicate that the marker of haplotype 46/1 may play a role not only in the acquisition of JAK2 V617F but is also attributed to clonal expansion, maintenance, and survival. Tefferi\textsuperscript{46} suggests that JAK2 V617F is not the initial clonogenic event in MPNs but rather one of several subclones derived from an ancestral clone. This is in accordance with the notes of Pardanani et al.\textsuperscript{47}, which support the hypothesis that this haplotype is located in a favorable cis regulatory environment, which facilitates the acquisition of JAK2 V617F, and which, in turn, is responsible for clonal expansion and the development of MPNs.
Furthermore, the possible role of acquired uniparental disomy, a genetic event that leads to mitotic recombination associated with neutral loss of heterozygosity of chromosome 9p in MPN patients, reducing both the haplotype and JAK2 V617F to a homozygous state\textsuperscript{14,48,49}, cannot be ruled out. In this context, cells with both variants theoretically have a selective advantage, which conditions greater myeloproliferative potential and favors the establishment of variant cells over healthy cells, thus explaining the increased VAF in individuals with the combination rs10974944 (G) + rs77375493 (T) (JAK2 V617F) in homozygosity.

Association between the elevation of hematological indices and the presence of 46/1 is observed in the literature\textsuperscript{16,33,50}; however, this is not a consensus among the scientific community\textsuperscript{8,53,61}. Our data show significant differences in MCV, MCH values in the PV group, and RBC, Hb, and Ht in TE carriers of the G allele of rs10974944, which has been observed in previous studies\textsuperscript{7,42,51}.

The present research is the first to analyze the 46/1 haplotype using the rs10974944 variant, present in intron 12 of JAK2, in a population from the Brazilian Amazon. The results of this study show that the rs10974944 (G) variant is associated with BCR::ABL1-negative myeloproliferative neoplasms, in patients positive for JAK2 V617F, especially those with PV, and a high allelic variant burden in these patients, and hematological alterations. Furthermore, the haplotype rs10974944G/rs10815151C/rs1011004A/rs77375493T was identified as a factor related to PV.

Materials and Methods

**Population:** One hundred individuals clinically diagnosed with BCR::ABL1-negative myeloproliferative neoplasms were included in the study. The study was conducted from February 2021 to January 2023. Laboratory analysis was performed at the Genomics Laboratory of the Foundation Hospital for Hematology and Hemotherapy of the State of Amazonas.

**Ethical Approval:** The study was submitted to and approved by the Ethics Committee of the Foundation Hospital for Hematology and Hemotherapy of the State of Amazonas under opinion No. 4,450,813 and certificate of ethical appreciation No. 39991420.6.0000.0009. Written informed consent was obtained from patients. This study complied with Resolution No. 466/2012 of the National Health Council for research involving human subjects and followed the parameters determined by the Declaration of Helsinki.

**Clinical and Laboratory Data:** Clinical data (gender, age, splenomegaly, thrombotic and/or hemorrhagic events) and laboratory data were obtained from medical records.

**Biological Sample and DNA Extraction:** Venous blood samples were collected in tubes containing EDTA, and DNA was extracted using Brazol (Lgcbio, Brazil), following the manufacturer's instructions, and stored at -80 °C.

Conventional PCR and PCR purification: For the amplification of the DNA region under analysis, a reaction with a final volume of 25 μL was used with 50-100 ng of genomic DNA, Buffer (1x), MgCl\textsubscript{2} (1.5 mM), forward primer CCAACTGAGTTTCCTTGCAG and reverse primer CTAGGTTAAGAGTATGTGGTCC (0.4 mM), dNTP mix (0.2 mM), and TAQ (1 U). The PCR products were separated on a 1.5% agarose gel. The PCR product, a 572 bp amplicon, was purified with polyethylene glycol (PEG 8000) (Promega).

**Nucleotide Sequencing and Sequence Analysis:** Approximately 5-30 ng of purified PCR product was applied to the sequencing reaction. Nucleotide sequencing was performed using BigDye\textsuperscript{®} Terminator v3.1 (Applied Biosystems), following the manufacturer's recommendations and the primers described above. The products were purified by the EDTA/Ethanol protocol and evaluated in the 3500 XL Genetic Analyzer\textsuperscript{®} automatic sequencer (Applied Biosystems, USA), with POP-7 polymer. The sequences were initially analyzed using the Sequencing Analysis software (Applied BioSystems [Thermo Fisher Scientific, São Paulo, Brazil]). Geneious 6.0.6 software (Biomatters, USA) was used to map the variants and obtain contigs for the comparison with the reference sequence Homo sapiens Janus kinase 2 (JAK2), (NCBI: NG_009904.1).

**Haplotype Analysis:** Haplotype frequencies were calculated using Haploviev software (v.4.2) as a measure of linkage disequilibrium (LD). Haplotypes with frequencies of <1% were not considered relevant for comparisons. Pairwise degree between
nucleotides was analyzed using the LD structure, considering \( r^2 \) values of >0.8 as strong LD, <0.8 as weak, and <0.1 as negative LD. Hardy-Weinberg equilibrium was calculated by comparing estimated and observed genotype frequencies using the \( \chi^2 \) test. SNVs with p-values of <0.001 were considered to be out of Hardy-Weinberg equilibrium.

**Statistical Analysis:** The obtained results were subjected to the Shapiro-Wilk normality test. Categorical variables were expressed as absolute value (n) and relative frequency (%) and were tested using the \( \chi^2 \) and Fisher's exact test with a 95% confidence interval. Numerical variables were expressed as median (Md) and interquartile range [IQR] with 75th percentile through GraphPad Prism v.9.0.2 software. For non-parametric variables, the Kruskal-Wallis test was performed. For both analyses, Dunn's post-test for multiple comparisons was also conducted using GraphPad Prism v.9.0.2 software. P-values of <0.05 were considered statistically significant.

**Declarations**

**Data Availability**

The datasets used and/or analyzed during the present study are available from the corresponding author on reasonable request. The GenBank accession number for the nucleotide sequence is PP208825.

**Acknowledgments**

The authors would like to thank the Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Our thanks also go to the Fundação Hospitalar de Hematologia e Hemoterapia do Amazonas (FHEMOAM) and the medical staff and laboratory professionals for their assistance in carrying out the work. Special thanks go to Dr Enedina Nogueira at the Genomics Laboratory CAM/UFAM for her support and the use of the infrastructure for the molecular analysis.

**Funding**

The present study was supported by the Fundação de Amparo à Pesquisa do Estado do Amazonas (Pro-Estado Program; grant nos. #002/2008, #007/2018 and #005/2019, and POSGRAD Program grant nos. #008/2023), Conselho Nacional de Desenvolvimento Científico e Tecnológico, and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.

**Author contributions statement**


**Competing interests**

The authors declare no competing interests.

**References**


**Figures**

**Figure 1**

The 46/1 haplotype, (a) located on chromosome 9p24.1, (b) encompasses the genes *JAK2, INSL6* and *INSL4*. (c) Variants in introns 10, 12, 14, and 15 are in strong linkage disequilibrium with the 46/1 haplotype and serve as markers for the detection of this haplotype.
Figure 2

Distribution of genetic data for (a) JAK2 V617F, (b) Variant allele frequency of JAK2 V617F+, and (c) Genotypic frequency and (d) Allelic frequency of rs10974944 in patients with polycythemia vera or essential thrombocythemia.
Figure 3

Linkage disequilibrium (LD) structure of JAK2 intron 12 in patients with polycythemia vera (PV) or essential thrombocythemia (ET). Numbers in the boxes indicate the value of the LD correlation coefficient ($r^2$) multiplied by 100. Lighter shades of boxes indicate a decreased $r^2$ value, strong LD is represented by the dark gray box. A discrete LD is observed between $rs10974944$ and $rs1081515$; $rs10974944$ and $rs10119004$; and $rs10974944$ and $rs77375493$ (JAK2 V617F).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial.pdf