Gene expression profiling of CD34(+) cells from patients with myeloproliferative neoplasms

MUGDHA SHARMA1, CHANDRA BHAVANI2, SRINAG BANGALORE SURESH1, JOHN PAUL1, LOKENDRA YADAV3, CECIL ROSS1 and SWETA SRIVASTAVA3

1Department of Medicine, St. John's Medical College and Hospital; 2St. John's Research Institute, St. John's National Academy of Health Sciences; 3Department of Transfusion Medicine and Immunohematology, St. John's Medical College and Hospital, Bengaluru, Karnataka 560034, India

Received February 6, 2020; Accepted October 8, 2020

DOI: 10.3892/ol.2021.12465

Abstract. Myeloproliferative neoplasms (MPN) are clonal disorders characterized by the increased proliferation of hematopoietic stem cell precursors and mature blood cells. Mutations of Janus kinase 2 (JAK2), Calreticulin (CALR) and MPL (myeloproliferative leukemia virus) are key driver mutations in MPN. However, the molecular profile of triple negative MPN has been a subject of ambiguity over the past few years. Mutations of, methyllyotosine dioxygenase TET2, polycomb group protein ASXL1 and histone-lysine N-methyltransferase EZH2 genes have accounted for certain subsets of triple negative MPNs but the driving cause for majority of cases is still unexplored. The present study performed a microarray-based transcriptomic profile analysis of bone marrow-derived CD34(+) cells from seven MPN samples. A total of 21,448 gene signatures were obtained, which were further filtered into 472 upregulated and 202 downregulated genes. Gene ontology and protein-protein interaction (PPI) network analysis highlighted an upregulation of genes involved in cell cycle and chromatin modification in JAK2V617F negative vs. positive samples. A significant increase in total CD34 counts in JAK2V617F negative vs. positive samples (P<0.05) was also observed. Overall, the present data showed a distinct pattern of expression in JAK2V617F negative vs. positive samples with upregulated genes involved in epigenetic modification.

Introduction

In the year 1915, Dr William Damshek first coined the term ‘myeloproliferative disorders’ (MPD) to classify phenotypically varied neoplasms that originate from a myeloid progenitor cell, namely chronic myeloid leukemia (CML), polycythemia vera (PV), essential thrombocythemia (ET) and primary myelofibrosis (PMF) (1). Fialkow et al reported that MPD are clonal stem cell disorders affecting both lymphoid and myeloid lineage (2). However, the identification of the Philadelphia chromosome in CML delineated the classification of MPD (3). The incidence rates for PV (0.4-2.8/100,000/year) in European Union and US registries are higher compared with ET (0.38-1.7/100,000/year) and PMF (0.1-1/100,000/year) (4,5). The median survival of PV and ET is longer (8-10 years) compared with PMF (2-5 years) (4,5).

With advancement of molecular techniques, in 2005 four different research groups published their findings regarding a recurrent somatic mutation in the MPN other than CML (6-9). The JAK2 gene was affected at codon 617 of exon 14, resulting in a valine to phenylalanine substitution (9). This was followed by another JAK2 mutation identified in exon 12 that lead the World Health Organization to include it as a diagnostic criteria for breakpoint cluster region Abelson murine leukemia virus oncoprotein homolog 1 (BCRABL) negative MPNs (10,11). JAK2 is a tyrosine kinase belonging to the Janus family of proteins, which also includes JAK1, JAK3 and TYK2 (12,13). When

Correspondence to: Dr Sweta Srivastava, Department of Transfusion Medicine and Immunohematology, St. John’s Medical College and Hospital, Sarjapur Road, Bengaluru, Karnataka 560034, India
E-mail: sweta.s@stjohns.in

Dr Cecil Ross, Department of Medicine, St. John’s Medical College and Hospital, Sarjapur Road, Bengaluru, Karnataka 560034, India
E-mail: cecil.ross@stjohns.in

Abbreviations: BCR, breakpoint cluster region; ABL, Abelson murine leukemia viral oncogene homolog 1; AML, acute myeloid leukemia; DNMT, DNA methyltransferase; IL, interleukin

Key words: myeloproliferative neoplasm, CD34, JAK, microarray, epigenetics, chromatin, MYB, cell cycle
interacting with cytokines, these kinases are activated via phosphorylation and further trigger a signal transduction cascade, which involves the STAT transcription factors (STAT1-8) that translocate into the nucleus and stimulate transcription of downstream targets (12,14). Studies have demonstrated that the JAK-STAT signaling pathway is utilized by a variety of cytokines for proliferation, survival and differentiation of hematopoietic cells (15-17). The role of JAK-STAT signaling in hematopoiesis has been demonstrated in mice with a germline deletion of JAK2, leading to embryonic lethal effects due to disrupted erythropoiesis (18). Mutations in the pseudokinase domain of JAK2 results in an uninhibited phosphorylation of the tyrosine kinase and constitutive activation of the JAK-STAT pathway, which is central to the pathogenesis of MPNs that possess this gain-of-function mutation (19). JAK2 mutations are reported in ~95% of PVs, 50-60% of ETs and 50-60% of PMFs, which led to the investigation of differing phenotypes resulting from the same genetic aberration (7). MPNs without the aforementioned genetic abnormalities are relegated to the miscellaneous category and in the past few years, considerable work has been done to characterize JAK2V617F and BCR-ABL negative MPN (15,20,21). In this context various mutations in genes involved in cell signaling (including SH2B adaptor protein 3, E3 ubiquitin-protein ligase CBL, NRAS, NF1 and FLT3), DNA methylation [TET2, DNA methyltransferase (DNMT)3A], metabolism ([NADP] cytoplasmic isocitrate dehydrogenase 1 and 2; IDH 1 and 2), histone modification (EZH2 and polycomb group protein ASXL1) and RNA splicing (serine/arginine-rich splicing factor 2 (SRSF2), splicing factor 3B subunit 1 (SF3B1) and splicing factor (U2AF1) have been identified (15). These mutations provide an insight into the disease progression but have not been able to explain the pathogenesis of this MPN subtype. Therefore, the present study aimed to investigate the microarray-based expression profiles of bone marrow derived CD34(+) positive cells from JAK2V617F positive vs. JAK2V617F negative MPNs.

Materials and methods

Sample collection and sorting. Bone marrow samples were collected from seven patients with MPN as per guidelines and approval by Institutional Ethics Committee at St. John's Medical College and Hospital, Bengaluru (approval no. 103/2016). The median age of patients was 39 years (32-47 years). The samples were collected between June 2016 to July 2018. Patients >18 years old and classified according to the World Health Organization criteria for MPNs (20) were included in the study. CD34 cells were isolated from bone marrow samples using CD34 Microbead kit from Miltenyi Biotec, Inc., as per the manufacturer's instructions.

Microarray analysis. RNA was isolated from CD34 cells using TRIzol® (Invitrogen; Thermo Fisher Scientific, Inc.). Whole transcriptome analysis of the seven samples was performed using the Whole Transcript (WT) PLUS Reagent kit (Applied Biosystems; Thermo Fisher Scientific, Inc.). The assay workflow in the order included first-stand cDNA synthesis, second-strand cDNA synthesis, cRNA amplification, cRNA purification and quantification, 2nd cycle ss-cDNA synthesis, template RNA removal, ss-cDNA purification and quantification, fragmentation, terminal labelling and hybridization to WT. All the steps were performed as per manufacturer's instructions. In total, 100 ng RNA was used as the input.

The quality check of CEL file was performed using Transcriptome Analysis Console software version 4.0.1 (Thermo Fisher Scientific, Inc.) and the files were normalized using the Oligo-R-based Bioconductor package (22). The processing was done using RNA method for R (23).

The normalized RNA files were checked for any discrepancies using the top ten housekeeping gene and their expression values across the sample. Principal Component Analysis was performed to cluster the samples. The samples that were clustered together based on the housekeeping genes were used for further analysis.

The normalized text file was then subjected to the AltAnalyze (http://www.altanalyze.org/) automated tool for analysis. Differentially expressed genes in the JAK2V617F negative vs. JAK2V617F positive samples were identified using ±1.5 as the fold-change cut-off with a statistical significance of <0.05. Prune ontology was calculated using the z score (initial filtering 1.96 cut-off). The database used for analysis was Ensembl (https://asia.ensembl.org/). The pathway network was generated automatically using the AltAnalyze tool.

Bioinformatic analysis. Differentially expressed genes were calculated using ±1.5 fold-change and P<0.05. Database for Annotation, Visualization and Integrated Discovery (DAVID) (https://david.ncifcrf.gov/) tool version 6.8 was used for Gene Ontology (GO), and functional and pathway enrichment analyses (24,25). The protein-protein interaction (PPI) network was constructed using STRING database version 11 (https://string-db.org/) using minimum confidence 0.4 (26). The genes which were enriched for DNA modification and stemness properties based on GO analysis were fed into StemChecker tool (http://stemchecker.sysbiolab.eu/) for analyzing stemness markers and transcription factors for enriched genes (27). Hierarchical clustering and heatmaps were created using the ClustVis online tool (28).

Reverse transcription-quantitative (RT-q)PCR. RNA was isolated from cells using TRIZol® (Invitrogen; Thermo Fisher Scientific, Inc.). In total, 1 µg RNA was converted into cDNA using the MMLV enzyme (Invitrogen; Thermo Fisher Scientific, Inc.) and the files were normalised using the Oligo-R-based Bioconductor package (22). The processing was done using RNA method for R (23).

Flow cytometry. The bone marrow cells post red blood cell lysis (150 mM ammonium chloride, 10 mM potassium bicarbonate,
CD34(+) cells were labelled with 0.1 mM ethylenediaminetetraacetic acid (EDTA) and 1% Fetal Bovine Serum (FBS) were labelled with CD34 FITC (clone 581 cat. no. 555821; BD Biosciences) at room temperature for 30 min in dark and analyzed using a FACs Calibur instrument (BD Biosciences). The results were analyzed using FCS4 Express software version 6.06.0025 (De Novo Software).

### Statistical Analysis

The statistical significance for CD34(+) cells was calculated using an unpaired student’s t-test using SPSS v16 (IBM). P<0.05 was considered to indicate a statistically significant difference. All Q-PCR experiments were performed in triplicates (n=3). The differential expression of genes was calculated using one-way ANOVA and Benjamini and Hochberg false discovery rate (FDR).

### Results

**Gene expression patterns in CD34(+) cells isolated from JAK2V617F negative vs. positive neoplasms.** The role of JAK2V617F mutation is well established in driving a subset of the myeloproliferative neoplasms (6-9); however, the etiology of the JAK2V617F negative MPNs is unknown. To understand this, JAK2V617F mutant positive and negative MPNs were compared using microarray-based transcriptional analysis. The hematopoietic cancer stem cells are determinant of the cellular hierarchy of tumor progression in hematological malignancies; however, the underlying mechanism is poorly understood in most cases, except CML wherein the role of translocation and resultant chimeric protein is a driving force in hematopoietic stem cells (HSCs) (30). To understand the changes that may be driving the JAK2V617F negative MPNs, bone marrow derived CD34(+) cells were targeted in the present study. CD34(+) fractions were separated from patient samples and subjected to microarray analysis. Tables II and III present the baseline characteristics and mutation profile of the patients involved in the present study. Patients who were JAK2V617F negative were also negative for CALR type 1 and 2 mutations.

### Tables

| Table I. Primer sequences for selected genes. |
|---------------------|---------------------|---------------------|
| **Gene** | Forward primer, 5'-3' | Reverse primer, 5'-3' |
| GUSB | AGCCCCATTATTCAGAGCGGAG | CCAAAATGAGCTCTTCACACC |
| SUV39H1 | TATGACTGCCAAAATCGTG | TGACTCTCCCACGTACTCC |
| MYB | GCAGTGACGAGGATGATGAG | CTGTTCCATTCTGTTCCACC |

<p>| Table II. Baseline characteristics of the patients. |
|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|</p>
<table>
<thead>
<tr>
<th>ID</th>
<th>Age, years</th>
<th>Sex</th>
<th>Hb, g/dl</th>
<th>Total WBC counts, x10^9/l</th>
<th>Platelet count, x10^9/l</th>
<th>Janus kinase 2 V617F status</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>39</td>
<td>M</td>
<td>19.1</td>
<td>8.43</td>
<td>297</td>
<td>Absent</td>
</tr>
<tr>
<td>2</td>
<td>41</td>
<td>M</td>
<td>16.9</td>
<td>11.81</td>
<td>286</td>
<td>Present</td>
</tr>
<tr>
<td>3</td>
<td>39</td>
<td>M</td>
<td>17.5</td>
<td>7.71</td>
<td>337</td>
<td>Absent</td>
</tr>
<tr>
<td>4</td>
<td>35</td>
<td>M</td>
<td>17.9</td>
<td>9.77</td>
<td>209</td>
<td>Absent</td>
</tr>
<tr>
<td>5</td>
<td>28</td>
<td>M</td>
<td>23.2</td>
<td>4.59</td>
<td>108</td>
<td>Absent</td>
</tr>
<tr>
<td>6</td>
<td>47</td>
<td>M</td>
<td>17.5</td>
<td>6.63</td>
<td>35</td>
<td>Absent</td>
</tr>
<tr>
<td>7</td>
<td>32</td>
<td>M</td>
<td>20.4</td>
<td>9.89</td>
<td>13</td>
<td>Present</td>
</tr>
</tbody>
</table>

WBC, white blood cell, M, male.

<p>| Table III. Mutation profile of patients negative for JAK2V617F. |
|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|</p>
<table>
<thead>
<tr>
<th>ID</th>
<th>JAK2V617F status</th>
<th>MPL W515L status</th>
<th>MPL W515K status</th>
<th>MPL W515A status</th>
<th>MPL S505N status</th>
<th>CALR type 1, p.L367fs*46</th>
<th>CALR type 2, K385fs*47</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
</tr>
<tr>
<td>3</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
</tr>
<tr>
<td>4</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
</tr>
<tr>
<td>5</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
</tr>
<tr>
<td>6</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
</tr>
</tbody>
</table>

MPL, myeloproliferative leukemia virus; CALR, calreticulin; JAK2, Janus kinase 2.
and MPL mutations (W515L, W515K, W515A and S505N) (Table III).

The microarray analyses of CD34(+) cells isolated from JAK2V617F negative and positive samples resulted in 21,448 gene signatures. The analysis resulted in 472 upregulated and 202 downregulated genes. The maximally downregulated genes, with >3-fold difference were 3-oxo-5α-steroid 4-dehydrogenase 1 (SRD5A1), matrix metalloproteinase (MMP7), Wilms tumor protein (WT1) and Ankyrin repeat and death domain-containing protein 1B (ANKDD1B), while the maximally upregulated genes with >3-fold difference were X-ray repair cross-complementing protein (XRCC6), TUBB, histone-lysine N-methyltransferase SUV39H1 and pantetheinase VNN1 as shown in Table IV. The expression of SUV39H1 and MYB was validated in five samples used for microarray analysis as shown in Fig. S1 where in the expression of both the genes was higher in JAK2V617F negative samples compared with control samples.

Hierarchical clustering of significantly upregulated and downregulated genes was performed to understand the correlation between samples (Fig. 1). Hierarchical clustering separated seven samples into two different groups consistent with the experimental design. Red bars indicated a high level of expression, blue bars indicated a low level of expression and white bars indicated no significant difference (Fig. 1). There was a distinctive difference in the expression of genes in JAK2V617F positive and negative samples (Fig. 1). The expression pattern of genes was homogeneous within the mutation negative samples; however, there was heterogeneity within mutation positive samples.

Table IV. Fold change of selected genes in JAK2V617F negative vs. positive samples.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Fold-change</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRD5A1</td>
<td>-7.71</td>
</tr>
<tr>
<td>ANKDD1B</td>
<td>-5.13</td>
</tr>
<tr>
<td>MMP7</td>
<td>-3.31</td>
</tr>
<tr>
<td>WT1</td>
<td>-3.54</td>
</tr>
<tr>
<td>VNN1</td>
<td>22.39</td>
</tr>
<tr>
<td>SUV39H1</td>
<td>5.03</td>
</tr>
<tr>
<td>MYB</td>
<td>4.64</td>
</tr>
<tr>
<td>XRCC</td>
<td>3.38</td>
</tr>
<tr>
<td>TUBB</td>
<td>4.43</td>
</tr>
</tbody>
</table>

Gene Ontology and STRING analysis reveals distinct cell cycle and epigenetic signatures in JAK2V617F negative vs. positive samples. To understand the important biological processes that determine the phenotypes in these two subsets of MPNs, GO was performed. The enriched genes and common pathways were analyzed using DAVID (Fig. 2). The distribution of the transcripts in the GO terms revealed a distinct set of biological processes (BP) (Fig. 2A), including ‘mitotic nuclear division’, ‘DNA replication’ and genes involved in ‘cell division, migration and proliferation’. Notably, all the BPs that were identified were indicative of a DNA-related processes, such as histone exchange and G2/M transition. The cellular components category also indicated an enrichment of genes associated with the nuclear compartment, such as ‘nucleus’, ‘nucleoplasm’ and ‘nuclear matrix’ (Fig. 2B), suggesting alterations in events occurring...
in the nuclear compartment. The most significant GO terms describing molecular functions for the enriched genes were ‘protein binding’, ‘chromatin binding’ and ‘ATP binding’ (Fig. 2C).

Following the aforementioned observations, the interactions between the enriched proteins were analyzed using the STRING database to understand the important molecular processes that may be activated in the JAK2V617F (−) MPNs. The PPI network constructed for the upregulated genes is shown in Fig. 3A. Notably, the STRING analysis revealed the existence of a distinct network of genes represented by tight network clusters. A detailed analysis of the network indicated the involvement of genes in two main biological processes: Cell cycle (red) and nucleobase modification (blue) (Fig. 3A). The central tight network cluster included genes involved in the cell cycle and its regulation, such as CDK1, TUBB, G2 and S phase-expressed protein (GTSE1), borealin (CDCA8), SUV39H1 and centrin 2 (CETN2) (Fig. 3A). Nucleobase modifying proteins, such as nuclear pore complex protein NUP85, NUP62, MYB, SS18, nucleolar and coiled-body phosphoprotein NOLC1, SWI/SNF complex subunit SMARCC2 and SMARCE1, were also linked in the network (Fig. 3A). SUV39H1 encodes a histone methyltransferase that tri-methylates lysine 9 of histone H3, which leads to transcriptional gene silencing (31). SMARCC1 and SMARCE1 are part of the SWI/SNF chromatin remodeling complex involved in transcriptional activation of target genes in an ATP-dependent manner (32). The majority of the proteins in the network were involved in acetylation and phosphorylation (Fig. 3A). Proteins such as hypoxia up-regulated protein HYOU1, Isoleucine-tRNA ligase (IARS), tyrosine t-RNA synthetase YARS, dual specificity protein kinase TTK, acyl-coenzyme A synthetase ACSM3, CDK1, XRCC and carbamoyl-phosphate synthetase CAD marked for nucleotide binding and acetylation indicating towards the role of DNA modification in this subset of MPNs. However, unlike other hematological

Figure 2. A representation of Gene Ontology terms using Database for Annotation, Visualization and Integrated Discovery analysis for enriched genes under (A) biological processes (B) cellular components and (C) molecular function with corresponding P-value scores and P<0.05.
malignancies, such as acute myeloid leukemia (AML) (33), little information is available on epigenetic modulations and their role in regulating hematopoiesis and fibrosis in MPNs. A heatmap was constructed to visualize the differences in gene expression between samples for genes comprising the aforementioned tight cluster (Fig. 3B). There was a categorical difference between samples with the red bars indicating a high expression of genes in mutant negative samples and blue bars indicating decreased expression. A bar graph representing the fold-change between mutant positive and negative samples shows the significant relative difference in expression of selected few genes involved in cell cycle and nucleotide modification, such as MYB, SUV39H1, XRCC, SMARCC2, SMARCE1 and CDK1 (Fig. 3C).
A network analysis was also performed for downregulated genes, which highlighted the genes involved in developmental process, response to stimuli, multicellular organismal process and regulation of apoptosis (Fig. S2). Key genes involved in this network included muscarinic acetylcholine receptor (CHRM3), IL6, Kalirin (KALRN), MMP7 and ankyrin repeat domain-containing protein ANKRD1.

Enrichment of chromatin modifiers in JAK2V617F negative samples vs. positive samples. Analyses using the DAVID and STRING databases based on GO and network enrichment for the upregulated genes suggested an association between the significantly upregulated genes and stem-like properties. Enriched genes which clustered into DNA modification and stemness properties (based on DAVID GO analysis) were further analyzed using StemChecker for evaluating the stemness signatures and transcription factors associated with these genes. Seven genes (12%) were linked to the hematopoietic signature (Fig. 4A) while forty-seven (76%) genes were related to the embryonic signature (Fig. 4A). The abilities of self-renewal, maintenance and differentiation of stem cells serve as a core reservoir of cancer initiation and development and tumor growth (34). Elevated numbers of embryonic or lineage-specific stem cells can lead to deregulation and altered differentiation of the progeny (34). The expansion of one specific lineage in mutant negative MPNs might be attributed to an altered hematopoietic stem cell. Notably, 76% of genes matched with embryonic stem cells (ESC) (Fig. 4A). ESC belongs to the primitive class of stem cells and are capable of differentiating into various cell types (35). Owing to their diverse nature, even a minute perturbation in these cells can lead to notable changes in cellular physiology (35). The transcription factors (TFs) associated with stem cells were also analyzed (Fig. 4B). Overexpression of stem-cell specific TFs may contribute to the pathological self-renewal characteristics of cancer stem cells (36). The majority of the genes mapped to E2F4 and NANOG TF studies by Boyer et al (37) (Figs. 4B and S3) which are involved in cell cycle and pluripotency respectively (38,39). To further evaluate the correlation between upregulated genes and stemness and chromatin modification, a network was constructed using STRING (Fig. 5A). The central tight cluster mainly comprised of genes implicated in the cell cycle, nucleobase modification and chromatin modification (Fig. 5A). These genes included MYB, tubulin γ, probable ATP-dependent RNA helicase DDX41, TUBB, IARS, SS18, CDCA8 and TTK (Fig. 5A). The heat map representing the relative expression of genes involved in the cluster indicated a higher expression of these genes in mutant negative samples (Fig. 5B). There was heterogeneity in mutant positive samples, but mutant negative samples had a homogenous profile. The bar graph representing the fold-changes of a select few key chromatin modifiers from the PPI network shows the significant relative difference in expression between mutant positive and negative samples such as chromatin modification protein (YEATS2), MYST-associated factor (MEAF6), methyl CpG binding protein (MBD1), nuclear receptor corepressor (NCOR2), PHD finger protein (PHF10) and SS18 (Fig. 5C). SUV39H is a well-known epigenetic marker and has an established role in MPN (40); however, the role of MYB is unknown in this context. Therefore, the expression of MYB was evaluated in JAK2V617F negative samples. There was a ~10-fold increase in MYB expression in the patient samples compared with controls (Fig. 6).

On the basis of the aforementioned observations, the stem cells in patient samples were further evaluated. CD34 was selected as the marker for interest as it is the most characterized and reported HSC marker shown to be involved in various hematological malignancies (41). Using flow cytometry, it was demonstrated that there was a significant increase in CD34 counts in JAK2V617F negative samples compared with positive samples (P<0.05; Fig. 7C), which concurred with the aforementioned bioinformatic analysis.

Discussion

The identification of the JAK2V617F mutation distinguished BCR-ABL negative MPN into two broad categories of JAK2 mutant positive and negative neoplasms (15,21). The activation and role of the JAK-STAT pathway has been well characterized in MPNs (19). It is the mutation negative subset of MPN that has drawn considerable interest in last few years. Seminal studies have identified mutations in CALR, MPL, ASXL1 and
Figure 5. Integrative analysis of genes involved in DNA modification and the cell cycle. (A) Network of 109 upregulated filtered genes using STRING at minimum confidence 0.4. Each node represents a protein. A tight cluster within the network was enriched with nucleobase modifying (red) and cell cycle (blue) proteins. (B) Heatmap of the representative genes encompassed the tight cluster were confirmed to be upregulated in JAK2V617F N1-5 vs. JAK2V617F P1-2 samples. ClustVis webtool based on R software was used for the analysis. Color key shows the differential expression of the genes. (C) Graphical representation of significant fold-change difference of few selected genes from the corresponding tight cluster and heatmap in JAK2V617F negative samples from microarray data. N, negative; P, positive; JAK2, Janus kinase 2.
TET2 genes as drivers in certain a subset of JAK2V617F mutation negative neoplasms (15,21,42). It is important to identify the factors that drive this cohort of mutant negative neoplasms. The MPN initiating cell is likely a hematopoietic stem or progenitor cell that, under perturbed conditions, deviates from the normal hematopoietic differentiation (43). To understand the molecular events that mediate tumor progression in mutant negative MPNs, the present study compared the transcriptional profiles of CD34(+) cells from JAK2 mutant positive and JAK2 mutant negative samples. Microarray based transcriptional analysis revealed the enrichment of a set of genes unique to the JAK2 mutant negative samples.

The present gene expression data revealed distinct signatures in mutant negative and positive samples. The GO analysis and STRING network of the upregulated genes identified two key networks of proteins involved in the cell cycle and chromatin modification. The present data revealed that there was an association between genes that are involved in transcriptional activation, DNA damage response and checkpoint regulation, microtubule stabilization and spindle assembly. These results conformed to previous findings reporting that dysregulation of events in the mitotic cell cycle are involved in multiple cancer types, including those of hematopoietic origin (44). CDK1 is involved in the G1/S to G2/M transition by binding to M phase cyclins. It is also involved in phosphorylation of key genes, such as MYB, HIST1H3F, EZH2, histone H3-like centromeric protein and nucleophosmin, that are further involved in downstream DNA modification (45-50). GTSE1 accumulates in the nucleus and binds the tumor suppressor protein p53, shutting it out of the nucleus and repressing its ability to induce apoptosis. GTSE1 is only expressed in the S and G2 phase of the cell cycle (51).

Notably, the other key network identified consisted of genes involved in the modification of chromatin. Mutations in genes involved in epigenetic regulation, such as TET2, ASXL1, DNMT3a, EZH2 and IDH1, have been previously reported in MPNs (52). Epigenetic changes control the expression of DNA in numerous ways, including DNA methylation, acetylation and modification of histones. These are important epigenetic regulatory processes that control various cellular events, including stemness maintenance and differentiation (52). The covalent modifications of histones at N-terminal lysine residues can lead to transcriptional activation and repression. Histone methylation markers, such as H3K4, are associated with transcription activation, while H3K9 and others are associated with transcriptional repression (53). The role of individual histone modifications within MPNs is currently unknown. Similarly, polycomb-related proteins, such as PRC2 and ASXL1, have been shown to be involved in chromatin remodeling in MPNs (52,54). The present study reported the enrichment of chromatin modifying enzymes, such as YEATS2, MEAF6, SMARCE1, SMARCC2, NCoR2, SUV39H1 in JAK2 negative samples. Reports suggest that these genes regulate the epigenetic landscape of cells (31,32,55-59). Similarly, SUV39H1, NCoR2 and CDK1 are involved in acute promyelocytic leukemia pathogenesis and treatment response (60). SUV39H1, CDK1 and XRCC are involved in chromosome architecture and chromatin assembly (31,45,61-63). MBD1 bind to methylated DNA and repress transcription of target genes (64). PHF10, SMARCC2, SMARCE1 and SS18 form part of chromatin remodeling complex altering the DNA-nucleosome biology (32,65,66). MYB, a helix-turn-helix DNA-binding transcriptional regulator, drives leukemia progression in AML and mixed lineage leukemia-fusion leukemia. MYB targets genes involved in myeloid differentiation, cellular proliferation, cell cycle, apoptosis and cell signaling (67). MYB activation by c-MYC via DNMT-1 upregulation promotes the progression of leukemia and lymphoma stem cells (68). The present study validated the expression of SUV39H1 and MYB in the samples sent for array. Since the role of SUV39H1 has been previously explored in MPN (40), the expression of MYB in mutation negative samples was evaluated. A >10-fold increase in MYB expression in JAK2V617F negative samples was observed; however, the exact function of MYB and the underlying molecular mechanisms in the pathogenesis of mutation negative MPN need to be further evaluated. Since there was an enrichment of genes involved in epigenetic modification, the association between these genes and stemness was investigated. Using StemChecker, 7 and 47 genes linked to the hematopoietic and embryonic signatures were reported, respectively. In total, 24 and 17% of these genes were under the control of E2F4 and NANO2 transcription factors, respectively. JAK2 is involved in the self-renewal of mouse ES cells by regulating the H3Y41 phosphorylation on the NANO2 promoter (69). TET2 also regulates the differentiation of ESC by inducing methylation at NANO2 promoter (70). The increased level of the CD34 hematopoietic stem cell marker in mutation negative MPN was correlated with the increase in stem-like properties, which may be due to altered epigenetics in these samples. The exact mechanism and chain of events that result in the gain of stemness and epigenetic properties in these subsets of MPN are yet to be investigated.

Due to an increasing number of studies investigating the role of epigenetics in MPN, clinical trials have already begun to supplement existing treatment regimens for MPN. Histone deacetylase inhibitors LBH589 and ITF2357 are in phase I and II clinical studies (71,72). DNMT inhibitors azacytidine and decitabine are also in phase II trials as a treatment approach in MPN (73,74).

The underlying causes regulating the initiation and pathogenesis of JAK2 negative MPN are multidimensional. Over the years several studies have identified driver mutations (15,21),

Figure 6. Graphical representation of fold-change difference of MYB expression normalized to control in JAK2V617F negative samples (n=8; three technical replicates). S, sample; JAK2, Janus kinase 2.
which has improved our understanding of MPN and aided in diagnosis and treatment of the disease. The present study focused on the epigenetics and the signaling between the associated genes rather than mutations. In the absence of driver mutations, the study highlighted the role of key dysregulated events that might promote the pathogenesis of MPN. To the best of our knowledge, the study is first of its kind to observe a distinct epigenetic signature in JAK2V617F negative patients in an Indian cohort. The altered expression of genes involved in epigenetic regulation and a further crosstalk between the epigenome and the transcription machinery may result in aberrant hematopoiesis that further contributes to disease progression. The limitation of the present study that should be addressed in future research is the evaluation and comparison of absolute CD34 counts in the bone marrow and circulating blood of patients who are positive and negative for the JAK2V617F mutation. Additional work needs to be conducted to resolve the exact function of these genes in disease progression. A further investigation into the MPN-specific epigenome will lead to an improved characterization of JAK2V617F negative MPN and help improve prognosis and treatment.

Acknowledgements

The authors would like to thank Dr Sitalakshmi Subramanian, (Department of Transfusion Medicine and Immunohematology, St. John's Medical College and Hospital), for support with use of the flow cytometer facility at St. John's Medical College and Hospital.

Funding

The present study was funded by The Rajiv Gandhi University of Health Sciences (grant no. M-73:2015-2016). Mugdha Sharma is supported by fellowship from Council for Scientific and Industrial Research, Government of India.
Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request. The additional datasets generated and/or analyzed during the current study are available in the StemChecker repository, [http://stemchecker.sysbiolab.eu/].

Authors' contributions

MS conceptualized the project, designed and executed experiments, contributed to data analysis, manuscript and figure preparation. CB contributed to the data analysis and manuscript preparation. SBS performed the flow cytometry and analyzed the data. JP contributed to the sample collection and clinical data analysis. LY contributed to molecular data analysis. CR conceptualized the project and contributed to the clinical data analysis. SS conceptualized the project, contributed to experimental design, data analysis and manuscript preparation. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Bone marrow samples were collected from patients as per guidelines and approval by Institutional Ethics Committee at St. John's Medical College and Hospital, Bengaluru in accordance with Declaration of Helsinki (approval no. 103/2016). All patients provided written informed consent for participation in the study.

Patient consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

References