Somatic CALR Mutations in Myeloproliferative Neoplasms with Nonmutated JAK2


Abstract

Background
Somatic mutations in the Janus kinase 2 gene (JAK2) occur in many myeloproliferative neoplasms, but the molecular pathogenesis of myeloproliferative neoplasms with nonmutated JAK2 is obscure, and the diagnosis of these neoplasms remains a challenge.

Methods
We performed exome sequencing of samples obtained from 151 patients with myeloproliferative neoplasms. The mutation status of the gene encoding calreticulin (CALR) was assessed in an additional 1345 hematologic cancers, 1517 other cancers, and 550 controls. We established phylogenetic trees using hematopoietic colonies. We assessed calreticulin subcellular localization using immunofluorescence and flow cytometry.

Results
Exome sequencing identified 1498 mutations in 151 patients, with medians of 6.5, 6.5, and 13.0 mutations per patient in samples of polycythemia vera, essential thrombocythemia, and myelofibrosis, respectively. Somatic CALR mutations were found in 70 to 84% of samples of myeloproliferative neoplasms with nonmutated JAK2, in 8% of myelodysplasia samples, in occasional samples of other myeloid cancers, and in none of the other cancers. A total of 148 CALR mutations were identified with 19 distinct variants. Mutations were located in exon 9 and generated a +1 base-pair frameshift, which would result in a mutant protein with a novel C-terminal. Mutant calreticulin was observed in the endoplasmic reticulum without increased cell-surface or Golgi accumulation. Patients with myeloproliferative neoplasms carrying CALR mutations presented with higher platelet counts and lower hemoglobin levels than patients with mutated JAK2. Mutation of CALR was detected in hematopoietic stem and progenitor cells. Clonal analyses showed CALR mutations in the earliest phylogenetic node, a finding consistent with its role as an initiating mutation in some patients.

Conclusions
Somatic mutations in the endoplasmic reticulum chaperone CALR were found in a majority of patients with myeloproliferative neoplasms with nonmutated JAK2. (Funded by the Kay Kendall Leukaemia Fund and others.)
THE MYELOPROLIFERATIVE NEOPLASMS

are chronic myeloid cancers that are characterized by the overproduction of mature blood cells, and that may evolve into acute myeloid leukemia.1,2 In addition to chronic myeloid leukemia with the BCR-ABL fusion gene, the three most common myeloproliferative neoplasms are essential thrombocythemia, polycythemia vera, and myelofibrosis.

Many patients with a BCR-ABL–negative myeloproliferative neoplasm carry a Janus kinase 2 (JAK2) V617F mutation.3–6 The JAK2 V617F mutation or JAK2 exon 12 mutations are found in most patients with polycythemia vera,7,8 whereas the JAK2 V617F mutation is found in only 50 to 60% of patients with essential thrombocythemia or myelofibrosis.9,10 Tests for JAK2 mutations have greatly simplified the diagnosis of myeloproliferative neoplasms and are now firmly embedded as front-line tests in national and international guidelines.11–14 However, distinguishing essential thrombocythemia with nonmutated JAK2 from the much more common reactive thrombocytosis remains a major diagnostic challenge.

Additional mutations have been identified in patients who have myeloproliferative neoplasms with or without JAK2 mutations. These mutations fall into two main classes: signaling mutations that activate the thrombopoietin receptor (MPL) or inactivate negative regulators such as LNK15,16 and mutations in epigenetic regulators of DNA methylation (TET2, DNMT3A, and IDH1/2)17–20 or chromatin structure (EZH2 and ASXL1).21,22 However, these additional mutations affect only a small minority of patients, genomewide data are lacking, and the pathogenesis of myeloproliferative neoplasms without JAK2 or MPL mutations remains obscure.

We have used massively parallel sequencing to characterize the mutational landscape of 151 patients with myeloproliferative neoplasms and have identified the gene encoding calreticulin (CALR) as a new cancer gene that is mutated in the majority of patients with myeloproliferative neoplasms and nonmutated JAK2.

METHODOLOGY

STUDY SAMPLES

We obtained samples from patients with myeloproliferative neoplasms and other cancers and from healthy controls after obtaining written informed consent. The study was approved by the ethics committee at each participating study center. For exome sequencing, we obtained tumor samples from blood granulocytes and constitutional samples from cultured T cells (in 34 patients), isolated T cells (in 42), or buccal swabs (in 75). (Additional details are provided in the Supplementary Appendix, available with the full text of this article at NEJM.org.) We classified all myeloproliferative neoplasms in accordance with the criteria of the British Committee for Standards in Haematology.11,12 The diagnoses of other hematologic neoplasms were made in accordance with 2008 World Health Organization categories.13

DNA SEQUENCING, CLONAL ANALYSIS, AND PROTEIN EXPRESSION

We performed exome sequencing of matched tumor and constitutional samples of DNA using standard protocols and conducted bioinformatic analyses to identify somatically acquired mutations, as described previously.23,24 We confirmed the presence of somatic sequence variants using custom target-capture methods and parallel sequencing combined with manual data curation. We used Sanger sequencing of CALR exon 9 combined with analyses of exome or genome sequencing data to determine the mutation status of CALR in an additional 1345 hematologic cancers, in 1517 other cancer samples, and in 550 control samples.

We cultured hematopoietic colonies, as described previously,25 and genotyped them using Sanger sequencing for all mutations identified on exome sequencing for that patient. Expression constructs, cell lines, and other methods are described in the Supplementary Appendix.

STATISTICAL ANALYSIS

We calculated the significance of recurrently mutated genes, taking into account mutation frequency, gene size, and mutation and sequence context. P values were adjusted for a false discovery rate and are expressed as a q value, with a value of less than 0.05 indicating statistical significance. Clonal heterogeneity was inferred from exome-sequencing data on the basis of a Bayesian Dirichlet process.26 We evaluated the statistical significance of differences between subtypes of myeloproliferative neoplasms using standard methods, which are detailed in the Supplementary Appendix.
**RESULTS**

**MUTATIONAL LANDSCAPE IN THE PATIENTS**

We analyzed the results of exome sequencing of DNA from granulocytes and constitutional DNA obtained from purified T cells or buccal cells in 168 patients with myeloproliferative neoplasms. The identification of appropriate constitutional DNA samples is a challenge among patients with myeloproliferative neoplasms, since circulating T cells and buccal cells may be contaminated by neoplastic cells (Fig. S1 in the Supplementary Appendix). For patients with the JAK2 V617F mutation, we measured the JAK2 V617F allele burden, using a quantitative polymerase-chain-reaction (PCR) assay to identify contaminated constitutional DNA samples, with in vitro T-cell cultures being necessary in some instances. In cases in which constitutional DNA was limited, the use of a combination of unamplified and whole-genome amplified DNA improved exome coverage (Fig. S2 in the Supplementary Appendix). We developed a bioinformatics workflow to enable the identification of high-confidence somatic variants (Fig. S3 in the Supplementary Appendix), and complemented this with orthogonal verification of mutations by targeted resequencing and manual curation.

Of the 168 samples of myeloproliferative neoplasms that were sequenced, 151 samples (48 samples of polycythemia vera, 62 of essential thrombocythemia, 39 of myelofibrosis, and 2 of unclassifiable myeloproliferative neoplasms) fulfilled the bioinformatic criteria and were taken forward for analysis (Fig. S3 and Table S1 in the Supplementary Appendix). The average exome sequencing coverage was 141×.

We identified and validated 1498 somatic mutations in 150 samples, with a range of 1 to 32 mutations per sample (Fig. 1A, and Table S2 in the Supplementary Appendix); no mutations were identified in the sample obtained from Patient PD8641, who had essential thrombocythemia. A total of 99 of the mutations we identified were in the V617F mutation and was found in all 48 patients with polycythemia vera, 35 of 62 patients with essential thrombocythemia (56%), and 27 of 39 patients with myelofibrosis (69%). Mutations in epigenetic regulators were as follows: TET2, 25 somatic variants in 22 patients; DNMT3A, 13 somatic variants in 12 patients; ASXL1, 13 somatic variants in 12 patients; EZH2 somatic variants in 4 patients; and IDH1/2 somatic variants in 3 patients (Fig. 1B).

Mutations in genes encoding components of the splicing machinery were found in 9 patients (U2AF1 in 4 patients, SF3B1 in 3 patients, and SRSF2 in 2 patients), and mutations in the gene encoding the thrombopoietin receptor (MPL) were identified in 7 patients, all of whom had essential thrombocythemia or myelofibrosis with unmutated JAK2. Mutations in genes that are reported to be mutated in myeloproliferative neoplasms at low frequencies were found in 4 patients (CBL, in 1 patient; NFE2L2, in 2 patients; and SH2B3/LNK, in 1 patient) (Fig. 1B). Missense somatic mutations in CHEK2, which have not been reported previously in myeloproliferative neoplasms, were found in 1 patient each with polycythemia vera, essential thrombocythemia, and myelofibrosis (q = 0.008) (Table S3 in the Supplementary Appendix). We also assessed pairwise associations between genes mutated in myeloproliferative neoplasms (Fig. S5 in the Supplementary Appendix). Mutations in ASXL1 were found to be comutated with genes involved in RNA splicing (U2AF1 and SRSF2), and mutations in CALR and EZH2 were also comutated.

The significantly higher number of mutations in patients with myelofibrosis was consistent with the concept that this disorder is a more advanced stage of disease than either essential thrombocythemia or polycythemia vera (P < 0.001 for the comparison of myelofibrosis with essential thrombocythemia and P = 0.008 for the comparison of myelofibrosis with polycythemia vera; pairwise t-test with Bonferroni correction for both comparisons). The mutational spectrum showed a predominance of C→T transitions (Fig. S4 in the Supplementary Appendix) and was similar to that observed in myelodysplasia and several epithelial cancers.23,27

Mutations in genes encoding components of the splicing machinery were found in 9 patients (U2AF1 in 4 patients, SF3B1 in 3 patients, and SRSF2 in 2 patients), and mutations in the gene encoding the thrombopoietin receptor (MPL) were identified in 7 patients, all of whom had essential thrombocythemia or myelofibrosis with unmutated JAK2.
A Number and Type of Mutations

- **PV**
- **ET**
- **MF**
- **MPN-U**

No./of/Somatic/Variants

<table>
<thead>
<tr>
<th>JAK2</th>
<th>CALR</th>
<th>TET2</th>
<th>DNMT3A</th>
<th>ASXL1</th>
<th>MPL</th>
<th>EZH2</th>
<th>SF3B1</th>
<th>IDH1/2</th>
<th>TP53</th>
<th>SF3B1</th>
<th>IDH2</th>
<th>SH2B3</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>120</td>
<td>60</td>
<td>30</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>120</td>
<td>110</td>
<td>90</td>
<td>120</td>
<td>120</td>
<td>120</td>
</tr>
</tbody>
</table>

B Type of Mutation

- **Silent**
- **Other**
- **Essential splice**
- **Missense**
- **Nonsense**
- **Indel**

C Subtype of Myeloproliferative Neoplasm with Mutation

- **MPN-U**
- **MF**
- **ET**
- **PV**

<table>
<thead>
<tr>
<th>JAK2</th>
<th>CALR</th>
<th>TET2</th>
<th>DNMT3A</th>
<th>ASXL1</th>
<th>MPL</th>
<th>EZH2</th>
<th>SF3B1</th>
<th>IDH1/2</th>
<th>TP53</th>
<th>SF3B1</th>
<th>IDH2</th>
<th>SH2B3</th>
</tr>
</thead>
<tbody>
<tr>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
</tr>
</tbody>
</table>
in SRSF2 were comutated with genes that encode epigenetic modifiers (TET2, IDH1, and ASXL1) — a profile strikingly similar to the associations seen in myelodysplasia.\(^28\)

**RECURRENT SOMATIC CALR MUTATIONS**

A striking pattern of somatic mutations was observed in CALR, a gene that had not previously been recognized as an oncogene mutated in any form of cancer. CALR mutations were identified in 26 patients (Fig. 1B and 1C). All the mutations were indels in exon 9 and had a remarkable association with disease, since they were found in 26 of 31 patients with essential thrombocytemia or myelofibrosis and nonmutated JAK2 or MPL (84%; 95% confidence interval [CI], 66 to 94) but in none of the 120 patients with JAK2 or MPL mutations (Fig. 2A, and Fig. S5 in the Supplementary Appendix). There were two common variants: L367fs*46, which resulted from a 52-bp deletion flanked by 7 base pairs of identical sequence; and K385fs*47, which resulted from a 5-bp insertion and represented an inverse duplication of the five nucleotides preceding the insertion (Fig. 2B).

Exome sequencing revealed low coverage at CALR (median depth, 10 reads) (Fig. S6 in the Supplementary Appendix). We therefore also performed Sanger sequencing of exon 9, which detected CALR mutations in 10 to 15% of cells (Fig. S6 in the Supplementary Appendix). This confirmed the absence of CALR exon 9 mutations in all 120 patients who had JAK2 or MPL mutations. Of 6 patients initially lacking a mutation in CALR, JAK2, and MPL on exome sequencing, 1 patient was found to have low-level mutated CALR on Sanger sequencing.

Among patients with essential thrombocytemia, those with CALR mutations, as compared with those with JAK2 mutations, presented with significantly higher platelet counts (P<0.001 by the Wilcoxon rank-sum test) and lower hemoglobin levels (P=0.02 by Student’s t-test) (Table S4 in the Supplementary Appendix). Patients with CALR mutations had a significantly higher incidence of transformation from essential thrombocytemia to myelofibrosis than did those with JAK2 mutations (P=0.03 by Fisher’s exact test) (Table S4 in the Supplementary Appendix). There were no significant between-group differences in rates of survival, although the number of deaths was small.

Exome sequencing showed that 146 of 151 patients with myeloproliferative neoplasms (97%) had mutations in JAK2, MPL, or CALR in a mutually exclusive manner (q<0.01) (Fig. S5 in the Supplementary Appendix). The results of bone marrow histologic analyses were reviewed in 4 of the 5 patients lacking mutations in all three genes and were consistent with a myeloproliferative neoplasm. In addition, 4 of these patients (PD9420, PD9845, PD8635, and PD7441) had clonal somatic mutations, an indication that these patients had a clonal myeloid disorder (Table S2 in the Supplementary Appendix).

To further characterize CALR mutations in myeloid and other cancers, we performed Sanger sequencing of CALR exon 9 in 1397 samples (including 52 control samples) and evaluated whole-exome or whole-genome sequencing data for 502 solid tumors, 498 control samples, and 1015 cancer cell lines (Table 1). CALR mutations were present in 110 of 158 patients with myeloproliferative neoplasms lacking JAK2 or MPL mutations (70%; 95% CI, 62 to 77), including 80 of 112 patients with essential thrombocytemia (71%), 18 of 32 patients with primary myelofibrosis (56%), and 12 of 14 patients with progression of essential thrombocytemia to myelofibrosis (86%). No CALR mutations were found in 511 myeloproliferative neoplasms with JAK2 or MPL mutations. Secondary acute myeloid leukemia developed in 2 patients with CALR mutations (1 each with essential thrombocytemia and myelofibrosis, both with the K385fs*47 variant), with persistence of the CALR mutations after
**Figure 2. CALR Mutations in Myeloproliferative Neoplasms.**
Panel A shows the number of CALR mutations found in myeloproliferative neoplasms (polycythemia vera, PV; essential thrombocytopenia, ET; or myelofibrosis, MF) that have mutated JAK2, mutated MPL, or nonmutated JAK2 or MPL. Panel B shows the results of validation on polymerase-chain-reaction (PCR) assay and Sanger sequencing of the two most common CALR variants — deletion (L367fs*46) and insertion (K385fs*47) — in patients with myeloproliferative neoplasms. The patients are indicated by their patient-identification numbers above the results of gel electrophoresis of PCR products of patients’ granulocytes (Gran) and T cells. The sequencing traces show the heterozygous mutation of CALR. The shaded region on the left highlights a homologous DNA sequence flanking the common CALR deletion, and the green arrow on the right highlights the inverse tandem duplication of five bases in the common CALR insertion. Panel C shows the genomic location of CALR deletions. The numbers indicate the number of patients with each disease. Panel D shows the genomic location of CALR insertions (shown in red). aCML denotes atypical chronic myeloid leukemia, CMML chronic myelomonocytic leukemia, and MDS myelodysplastic syndromes.
CALR Mutations in Myeloproliferative Neoplasms

transformation. CALR mutations were identified in 10 of 120 patients with myelodysplastic syndromes (8%; 95% CI, 4 to 15) (Table 1). One patient each with chronic myelomonocytic leukemia and atypical chronic myeloid leukemia had CALR mutations. No mutations were found in control samples, lymphoid cancers, solid tumors, or cell lines (Table 1, and Table S5 in the Supplementary Appendix).

Overall, CALR exon 9 mutations were identified in 148 patients (Table S6 in the Supplementary Appendix). All mutations were indels with 19 distinct variants: 14 deletions, 2 insertions, and 3 complex indels (Fig. 2C and 2D). In these 148 patients, variants: 14 deletions, 2 insertions, and 3 complex indels (Fig. 2C and 2D). In these 148 patients, variants L367fs*46 and K385fs*47 were the most common CALR mutations (in 67 patients [45%] and 61 patients [41%], respectively). The remaining 20 patients were found to have 17 unique variants. L367fs*46 was found more frequently in myelofibrosis than was K385fs*47 (P = 0.009 by chi-square test).

**Mutant Protein with a Novel C-Terminal**

Calreticulin is a highly conserved protein with multiple reported functions. Within the endoplasmic reticulum, the protein ensures appropriate folding of newly synthesized glycoproteins and modulates calcium homeostasis.29,30 Outside the endoplasmic reticulum, calreticulin is also found in intracellular, cell-surface, and extracellular compartments, where it has been implicated in diverse biologic processes, including proliferation, apoptosis, and immunogenic cell death.31-34 Calreticulin has three main structural and functional domains: an N-terminal lectin-binding domain that is interrupted by a proline-rich P domain and an unstructured C-terminal acidic domain that contains multiple calcium-binding sites.30,35

All the CALR indel mutations that were detected are predicted to generate mutant proteins with a novel C-terminal (Fig. 3A). The extent of the C-terminal alterations vary, but all 19 distinct variants share a loss of a sequence of 27 amino acids with a concomitant gain of a novel peptide consisting of 36 amino acids. These alterations result in the loss of most of the C-terminal acidic domain and the KDEL signal. (The KDEL amino acid sequence [Lys-Asp-Glu-Leu] is present on some resident endoplasmic reticulum proteins and enables retrieval of these proteins from the Golgi apparatus back to the endoplasmic reticulum.) This loss of function raises the possibility of compromised retention or retrieval in the endoplasmic reticulum. All CALR mutations shift the reading frame by one base pair, a pattern of mutation that is very different from that observed in known tumor-suppressor genes such as TET2 and ASXL1 (Fig. 3B).36,37 Such genes are affected by nonsense mutations and indels, with the latter generating +1 and +2 base-pair frame-shifts that frequently cause premature protein termination. The remarkably stereotypical pattern of CALR mutations implies a strong selective pressure to generate the mutant C-terminal.

To confirm that CALR mutations result in a mutant protein product, CALR variant constructs (L367fs*46, deletion, and K385fs*47, insertion) were transiently expressed in human embryonic kidney (HEK) 293T cells. Immunoblotting confirmed the expression of mutant CALR proteins (Fig. 4A). Immunofluorescence studies of COS-7 cells showed that mutant CALR did not accumulate in the Golgi apparatus (Fig. 4B) and colocalized with exogenously expressed nonmutant CALR (Fig. 4C). Similar studies comparing mutated and nonmutated CALR myeloid cells from two patients with CALR-mutated essential thrombocythemia also showed no differences in the pattern of endogenous CALR (Fig. 4D).

The up-regulation of cell-surface CALR is reported to mediate the phagocytosis of blasts in acute myeloid leukemia and might be a consequence of altered cellular localization of CALR due to loss of KDEL. We measured cell-surface CALR in hematopoietic 32D cells expressing CALR in hematopoietic 32D cells expressing mutant CALR variants (L367fs*46 and K385fs*47) and found no differences from nonmutant CALR expression (Fig. S7 in the Supplementary Appendix). Consistent with these results, there was no significant increase in cell-surface CALR in peripheral-blood leukocytes from five patients with mutated CALR (with estimated tumor burdens of 40 to 100% in granulocytes from exome sequencing), as compared with normal controls (Fig. 4E). Our data do not exclude the possibility of a partial loss of mutant CALR from the endoplasmic reticulum through the secretory pathway.

**CALR Mutations in the Hematopoietic-Stem-Cell Compartment**

CALR mutations were identified in patient-derived granulocyte–macrophage colonies and in erythroid colonies (Fig. S8 in the Supplementary Appendix), indicating that the mutations occur in a...
<table>
<thead>
<tr>
<th>Tumor Type</th>
<th>No. of Samples</th>
<th>CALR Mutations</th>
<th>Occurrence</th>
<th>Frequency % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Exome screening</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myeloproliferative neoplasms</td>
<td>151</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polycythemia vera</td>
<td>0/48</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Essential thrombocytocemia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With JAK2 or MPL mutation</td>
<td>0/40</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without JAK2 or MPL mutation</td>
<td>18/22</td>
<td>82 (59–94)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myelofibrosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary myelofibrosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With JAK2 or MPL mutation</td>
<td>0/22</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without JAK2 or MPL mutation</td>
<td>4/5</td>
<td>80 (30–99)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>After essential thrombocytocemia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With JAK or MPL mutation</td>
<td>0/2</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without JAK or MPL mutation</td>
<td>4/4</td>
<td>100 (40–100)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>After polycythemia vera</td>
<td>0/6</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unclassifiable myeloproliferative neoplasm</td>
<td>0/2</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Follow-up screening</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myeloproliferative neoplasm</td>
<td>669</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polycythemia vera</td>
<td>0/217</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Essential thrombocytocemia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With JAK2 or MPL mutation</td>
<td>0/138</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without JAK2 or MPL mutation</td>
<td>80/112</td>
<td>71 (62–79)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myelofibrosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary myelofibrosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With JAK2 or MPL mutation</td>
<td>0/97</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without JAK2 or MPL mutation</td>
<td>18/32</td>
<td>56 (38–73)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>After essential thrombocytocemia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With JAK2 or MPL mutation</td>
<td>0/18</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Without JAK2 or MPL mutation</td>
<td>12/14</td>
<td>86 (56–98)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>After polycythemia vera</td>
<td>0/41</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other myeloid disorder</td>
<td>389</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myelodysplastic syndromes</td>
<td>10/120</td>
<td>8 (4–15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refractory anemia</td>
<td>5/53</td>
<td>9 (4–21)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RARS</td>
<td>3/27</td>
<td>11 (3–30)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refractory anemia with excess blasts</td>
<td>2/17</td>
<td>12 (2–38)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refractory anemia (5q–)</td>
<td>0/4</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RARS and thrombocytosis</td>
<td>0/6</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCMD</td>
<td>0/10</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCMD and ringed sideroblasts</td>
<td>0/3</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
multipotent progenitor capable of generating both myeloid and erythroid progeny. Consistent with these results, genotyping studies of three patients showed that CALR mutations were present in flow-sorted, highly enriched hematopoietic stem cells (HSC; lin− CD34+CD38− CD45RA− CD90+), common myeloid progenitors (lin− CD34+CD38+CD90−CD10−FLK2+CD45RA+), granulocyte–macrophage progenitors (lin− CD34+CD38+CD90−CD10−FLK2−CD45RA+), and megakaryocyte–erythroid progenitors (lin−CD34+CD38−CD90−CD10−FLK2−CD45RA+) (Fig. S8 in the Supplementary Appendix). These data are consistent with CALR mutations arising in the HSC compartment.

To ascertain whether mutation in CALR is an early event, we used exome-sequencing data to infer the fraction of cells bearing mutations and predict clonal relationships using a Bayesian Dirichlet process (Fig. S9 in the Supplementary Appendix). The results suggest that CALR muta-

Table 1. (Continued.)

<table>
<thead>
<tr>
<th>Tumor Type</th>
<th>No. of Samples</th>
<th>CALR Mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Occurrence</td>
<td>Frequency</td>
</tr>
<tr>
<td></td>
<td>no./total no.</td>
<td>% (95% CI)</td>
</tr>
<tr>
<td>Chronic myelomonocytic leukemia</td>
<td>1/33</td>
<td>3 (0–18)</td>
</tr>
<tr>
<td>Atypical chronic myeloid leukemia</td>
<td>1/29</td>
<td>3 (0–20)</td>
</tr>
<tr>
<td>Chronic myeloid leukemia</td>
<td>0/28</td>
<td>0</td>
</tr>
<tr>
<td>Acute myeloid leukemia</td>
<td>0/48</td>
<td>0</td>
</tr>
<tr>
<td>Eosinophilic disorder</td>
<td>0/2</td>
<td>0</td>
</tr>
<tr>
<td>Systemic mastocytosis</td>
<td>0/114</td>
<td>0</td>
</tr>
<tr>
<td>Idiopathic erythrocytosis</td>
<td>0/5</td>
<td>0</td>
</tr>
<tr>
<td>Transient abnormal myelopoiesis</td>
<td>0/10</td>
<td>0</td>
</tr>
<tr>
<td><strong>Lymphoid cancer</strong></td>
<td>287</td>
<td></td>
</tr>
<tr>
<td>Acute lymphoblastic leukemia</td>
<td>0/62</td>
<td>0</td>
</tr>
<tr>
<td>Multiple myeloma</td>
<td>0/73</td>
<td>0</td>
</tr>
<tr>
<td>Chronic lymphocytic leukemia</td>
<td>0/77</td>
<td>0</td>
</tr>
<tr>
<td>Diffuse large B-cell lymphoma</td>
<td>0/25</td>
<td>0</td>
</tr>
<tr>
<td>MALT lymphoma</td>
<td>0/26</td>
<td>0</td>
</tr>
<tr>
<td>Splenic marginal-zone lymphoma</td>
<td>0/24</td>
<td>0</td>
</tr>
<tr>
<td><strong>Solid tumors</strong></td>
<td>502</td>
<td></td>
</tr>
<tr>
<td>Breast</td>
<td>0/190</td>
<td>0</td>
</tr>
<tr>
<td>Lung</td>
<td>0/28</td>
<td>0</td>
</tr>
<tr>
<td>Prostate</td>
<td>0/30</td>
<td>0</td>
</tr>
<tr>
<td>Colorectal</td>
<td>0/110</td>
<td>0</td>
</tr>
<tr>
<td>Skin</td>
<td>0/25</td>
<td>0</td>
</tr>
<tr>
<td>Bone</td>
<td>0/80</td>
<td>0</td>
</tr>
<tr>
<td>Renal</td>
<td>0/26</td>
<td>0</td>
</tr>
<tr>
<td>Meningioma</td>
<td>0/13</td>
<td>0</td>
</tr>
<tr>
<td><strong>Cell lines</strong></td>
<td>1015</td>
<td></td>
</tr>
<tr>
<td>Bone marrow derived stem cells</td>
<td>0/1015</td>
<td>0</td>
</tr>
<tr>
<td><strong>Control samples</strong></td>
<td>550</td>
<td></td>
</tr>
<tr>
<td>Bone marrow derived stem cells</td>
<td>0/550</td>
<td>0</td>
</tr>
</tbody>
</table>

* Mutations in the gene encoding calreticulin (CALR) were identified from whole-exome sequencing and validated by targeted resequencing in the exome screening cohort. For the follow-up screening, CALR mutations were identified on Sanger sequencing of CALR exon 9 for myeloproliferative neoplasms, other myeloid disorders, lymphoid cancers, and a subset of 52 control samples. Solid tumors, cell lines, and the remaining 498 control samples were screened for CALR mutations from existing whole-exome or whole-genome sequencing data. MALT denotes mucosa-associated lymphoid tissue, RARS refractory anemia with ringed sideroblasts, and RCMD refractory cytopenias with multilineage dysplasia.
### A  Functional Domains of CALR Protein

<table>
<thead>
<tr>
<th>Amino acids</th>
<th>Human</th>
<th>Mouse</th>
<th>Zebradash</th>
<th>Drosophila</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N Domain</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P Domain</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C Domain</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KDEL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Deletions

- **L367fs**
  - QDEE-----------------QRTRMRTRMRMRMRMRMRMRMRMPAPRTSCREACLQGWTEA | 411
- **E370fs**
  - QDEEQRLKE-----------------VMRTRMRTRMRMRMRMRMRMRMPAPRTSCREACLQGWTEA | 411
- **L367fs**
  - QDEE-----------------RRRQTRMRTRMRMRMRMRMRMRMPAPRTSCREACLQGWTEA | 416
- **L367fs**
  - QDEE-----------------RRRQTRMRTRMRMRMRMRMRMRMPAPRTSCREACLQGWTEA | 413
- **K368fs**
  - QDEEQRL-----------RRTRMRTRMRMRMRMRMRMRMPAPRTSCREACLQGWTEA | 408
- **L367fs**
  - QDEEQR-----------RRRRQTRMRTRMRMRMRMRMRMPAPRTSCREACLQGWTEA | 417
- **R366fs**
  - QDEEQR-----------RRRRQTRMRTRMRMRMRMRMRMPAPRTSCREACLQGWTEA | 417
- **E371fs**
  - QDEEQRLKE-------------QRTRMRTRMRMRMRMRMRMRMPAPRTSCREACLQGWTEA | 418
- **K368fs**
  - QDEEQRL-------------------RRTRMRTRMRMRMRMRMRMRMPAPRTSCREACLQGWTEA | 409
- **E370fs**
  - QDEEQRLK-----------------MRTRMRTRMRMRMRMRMRMRMPAPRTSCREACLQGWTEA | 405
- **D373fs**
  - QDEEQRLK----------RRRRQTRMRTRMRMRMRMRMRMPAPRTSCREACLQGWTEA | 418
- **K374fs**
  - QDEEQRLKED----------RRRRQTRMRTRMRMRMRMRMRMPAPRTSCREACLQGWTEA | 425
- **E371fs**
  - QDEEQRLKE-------------QRTRMRTRMRMRMRMRMRMRMPAPRTSCREACLQGWTEA | 416

#### Insertions

- **K385fs**
  - QDEEQRLKEEEEDKKRKEEEEAEDRRMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA | 430
- **K385fs**
  - QDEEQRLKEEEEDKKRKEEEEAEDRRMRTKMRMRRMRRTRRKMRRKMSPARPRTSCREACLQGWTEA | 430

#### Complex

- **R376fs**
  - QDEEQRLKEEEEDKKRKEEEEAED---RKRRQTRMRTRMRMRMRMRMRMPAPRTSCREACLQGWTEA | 429
- **K385fs**
  - QDEEQRLKEEEEDKKRKEEEEAED--RRTRMRTRMRMRMRMRMRMPAPRTSCREACLQGWTEA | 430
- **E381fs**
  - QDEEQRLKEEEEDKKRKEE---QRTRMRTRMRMRMRMRMRMPAPRTSCREACLQGWTEA | 428

### B  Comparison with Other Common Mutations in MPN

<table>
<thead>
<tr>
<th>Percent Total Variants</th>
<th>ASXL1</th>
<th>TET2</th>
<th>CALR</th>
<th>ASXL1</th>
<th>TET2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frameshift by 2 bp</td>
<td>3</td>
<td>7</td>
<td>148</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Frameshift by 1 bp</td>
<td>4</td>
<td>3</td>
<td>8</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>Nonsense</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Inframe</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Missense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CALR Mutations in Myeloproliferative Neoplasms

Understanding the molecular basis for essential thrombocytopenia and myelofibrosis in patients without JAK2 mutations has been a major goal in the field of myeloproliferative neoplasms. In this study, we found that CALR, a previously unrecognized oncogene, is mutated in 70 to 84% of patients with essential thrombocytopenia or myelofibrosis without JAK2 or MPL mutations. We also describe a comprehensive characterization of the mutational landscape of BCR-ABL-negative myeloproliferative neoplasms. We found that the median number of somatic mutations per patient was similar to that seen in other hematologic cancers (e.g., acute myeloid leukemia, acute lymphoblastic leukemia, and chronic lymphocytic leukemia, in which the median number of somatic mutations ranges from 8 to 13) but was substantially lower than the median number seen in most common epithelial cancers (range, 30 to 70). There was a long tail of cancer genes that have been implicated previously in myeloid cancers, with each gene mutated in a small fraction of patients with myeloproliferative neoplasms, as well as somatic mutations in CHEK2, which have not been described previously in myeloproliferative neoplasms.

CALR is present in the endoplasmic reticulum, where it forms a key component of the quality-control machinery that ensures proper glycoprotein folding and also contributes to calcium homeostasis. In addition, CALR functions outside the endoplasmic reticulum, in the cytoplasm, at the cell surface, and in the extracellular matrix, to influence a variety of processes, including proliferation, apoptosis, phagocytosis, and immune responses. Although calreticulin has been implicated in promoting the survival of malignant cells and in their ability to generate immune responses, to our knowledge, somatic mutations affecting endoplasmic reticulum chaperones have not been reported in any cancer. We describe 19 different CALR mutations (all deletions, insertions, or both) that generate a +1 base-pair frameshift predicted to replace the canonical C-terminal by a novel protein sequence. The normal CALR C-terminal domain is highly acidic and contains multiple calcium-binding sites together with a retrieval sequence (KDEL) that is typical of endoplasmic reticulum resident proteins. In contrast, the new predicted C-terminal domain of mutant CALR is basic, lacks multiple calcium-binding sites, and has no KDEL sequence. Mutant CALR did not have an altered Golgi or cell-surface accumulation of CALR, which was consistent with previously reported KDEL-independent retention mechanisms in the endoplasmic reticulum. However, our data indicate that mutant CALR probably has acquired novel functionality within or outside the endoplasmic reticulum.

Current diagnostic guidelines for patients with myeloproliferative neoplasms without JAK2 or MPL mutations rely on the ruling out of reactive...
Nonmutant CALR/GalT-YFP

Deletion

Insertion

10 µm

CALR-Positive/Cells/(%)

0

G

L

Control

FSC

CALR

Mutated

CALR/GFP-KDEL

CALR-FLAG NM Merge+DAPI

C

CALR-GFP-KDEL/CALR-Variant-FLAG

CALR-GFP-KDEL

CALR-FLAG Del Merge+DAPI

PDI CALR Merge+DAPI

PDI CALR Merge+DAPI

Control

CALR

Mutated

CALR

G

1.47%

3.67%

L

2.75%

4.56%

Mutated

Additional Genetic Events over Time

ET (PD6563)

PET-MF (PD4063)

ET (PD8632)

ET (PD7442)

ET (PD6561)

Nonmutant

Nonmutant

Nonmutant

Nonmutant

Nonmutant

Nonmutant

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

COL4A2 splice

OR5212 P320A

COL1A1 A1194T

LRTM1 P22P

CPA4 Y87Y

OR4C6 V122M

GKS Y403C

CELF5 S447C

ARHGAP32 A544T

B4GALNT2 V472M

GTPBP2 F411L

KNTC1 A886A

INFAR1 N44D

DIP2C A815A

AHGAP32 A544T

B4GALNT2 V472M

GTPBP2 F411L

KNTC1 A886A

INFAR1 N44D

DIP2C A815A

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47

Cal R K385fs*47
Figure 4 (facing page). Localization of Calreticulin and Clonal Heterogeneity in Patients with CALR Mutations.

Panel A shows immunoblotting of transiently transfected human embryonic kidney (HEK) 293T cells analyzed for FLAG (a polypeptide protein tag), CALR (detecting endogenous as well as transfected calreticulin), and beta-actin. The CALR deletion is the L367fs*46 variant, and the CALR insertion is the K385fs*47 variant. The CALR expression construct is shown above the immunoblots. CMV-Pr denotes CMV promoter. Panel B shows confocal photomicrographs of COS-7 cells transiently expressing FLAG-tagged CALR variants and a Golgi reporter (galactosyltransferase fused to yellow fluorescent protein [GalT-YFP]). Red indicates FLAG; green, Golgi; and blue, nucleus ([4',6-diamidino-2-phenylidole dihydrochloride [DAPI]]). The images show that nonmutant and mutant CALR have an endoplasmic reticulum localization pattern with no increased accumulation in the Golgi. Panel C shows confocal photomicrographs of COS-7 cells transiently coexpressing nonmutant (NM) CALR tagged with green fluorescent protein (GFP) and FLAG-tagged CALR variants. Merge images show that FLAG-tagged CALR variants colocalize with nonmutant CALR to the endoplasmic reticulum. Panel D shows confocal photomicrographs of myeloid cells from CALR mutant and nonmutated granulocyte–macrophage colony-forming unit colonies derived from a patient with essential thrombocytemia with the CALR K385fs*47 mutation; the cells have been stained for protein disulfide isomerase (PDI, a resident protein of the endoplasmic reticulum) and endogenous CALR. Panel E shows flow cytometric analysis indicating the degree of CALR cell-surface expression in granulocytes (G) and lymphocytes (L) from a healthy control and from a patient with mutated CALR. The graph shows the percentage of viable cells expressing cell-surface CALR in peripheral blood from healthy controls (triangles) and patients with mutated CALR (circles). FSC denotes forward scatter. Panel F shows clonal structures in five patients (indicated by their patient-identification numbers) with myeloproliferative neoplasms with mutated CALR, as determined on genotyping of hematopoietic (erythroid) colonies. Each circle represents a clone, with nonmutant clones shown in white and mutant clones in brown. The earliest detectable clone is represented at the top of each diagram, with subsequent subclones shown below. Somatic mutations that were acquired in each subclone are indicated beside the respective nodes and represent those that were acquired in addition to mutations present in earlier subclones. Numbers of colonies that were identified for each node are shown inside the circles. ET denotes essential thrombocytemia, and PET-MF post-ET myelofibrosis.

Causes of thrombocytosis combined with examination of bone marrow–biopsy specimens.\textsuperscript{11,13,14}

In our exome subgroup, mutations in CALR, JAK2, or MPL were present in 146 of 151 patients with myeloproliferative neoplasms (97%). CALR mutations were present in 26 of 31 patients without JAK2 or MPL mutations in the exome subgroup (84%) and in 110 of 158 patients in the follow-up subgroup (70%), with the lower prevalence in the latter group possibly reflecting limitations of Sanger sequencing of whole-blood samples. Detection of CALR mutations in peripheral blood could potentially be used as a diagnostic tool in the same way that tests for JAK2 mutations have simplified and improved the accuracy of diagnosis of patients with myeloproliferative neoplasms worldwide.\textsuperscript{11-14} Further research is needed to explore these potential uses.

Supported by the Kay Kendall Leukaemia Fund (including a clinical research fellowship to Drs. Nangalia and Godfrey), the Wellcome Trust (including fellowships to Drs. Campbell and Ron), Leukemia and Lymphoma Research, Cancer Research UK, the National Institute for Health Research Cambridge Biomedical Research Centre, the Cambridge Experimental Cancer Medicine Centre, and the Leukemia and Lymphoma Society of America; a postdoctoral research grant from the Research Foundation–Flanders (to Dr. Van Loo); a grant from Associazione Italiana per la Ricerca sul Cancro (AIRC)–Gruppo Italiano Malattie Mieloproliferative (to Drs. Vannucchi and Guglielmelli); and a Canadian Institutes of Health Research Fellowship and the Lady Tata Memorial Trust grant (to Dr. Kent).

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank members of the Cambridge Blood and Stem Cell Bank (Cambridge), Tayside Cancer Tissue Bank (Dundee), and Haematology Oncology Diagnostics Laboratory, Addenbrooke’s Hospital (Cambridge) for technical assistance; Sarah Moody, Alex Clipson, Ming Wang, Giovanni Cazzaniga, Andrea Biondi, Jan Zuna, Helena Kemptski, and Mike Groves for sample ascertainment; Stefan Marciniak, Steffen Preisler, and Bertie Gootgens for helpful advice and discussion; Janice Prick for technical assistance; and Hesham Eidaly and James Watkins for review of histologic analyses.

APPENDIX

REFERENCES


Copyright © 2013 Massachusetts Medical Society.