Gene of the Month: The Calreticulin Gene and Myeloproliferative Neoplasms

Aoibhinn Clinton¹
Mary Frances McMullin¹,²

¹.Department of Haematology,
Belfast City Hospital,
Lisburn Road,
Belfast BT9, N. Ireland

². Centre for Medical Education,
Queen’s University Belfast,
Lisburn Road,
Belfast BT9, N. Ireland.

Correspondence to:
Mary Frances McMullin,
C Floor, Haematology,
Belfast City Hospital,
Queen’s University Belfast,
Lisburn Road,
Belfast BT9 7AB, N. Ireland.
Email m.mcmullin@qub.ac.uk

Keywords: Calreticulin gene, Calreticulin protein, Myeloproliferative Neoplasms
Word Count: 2873
Abstract

The Philadelphia negative myeloproliferative neoplasms include Polycythaemia Vera (PV), Essential Thrombocytopenia (ET) and Primary Myelofibrosis (PMF). Patients with these conditions were mainly thought to harbour JAK2V617F mutations or an MPL substitution. In 2013 two revolutionary studies identified recurrent mutations in a gene which encodes the protein calreticulin. (CALR). This mutation was detected in patients with PMF and ET with non mutated JAK2 or MPL but was absent in patients with PV. The CALR gene encodes the calreticulin protein, which is a multifactorial protein, mainly located in the endoplasmic reticulum in chromosome 19 and regulates calcium homeostasis, chaperones and has also been implicated in multiple cellular processes including cell signalling, regulation of gene expression, cell adhesion, autoimmunity and apoptosis. Somatic 52-bp deletions and recurrent 52-bp insertion mutations in CALR were detected and all resulted in frameshift and clusters in exon 9 of the gene. This review will summarise the current knowledge on the CALR gene and mutation of the gene in pathological conditions and patient phenotypes.

Localisation of the CALR Gene

Calreticulin is a ubiquitous protein which is present in all eukaryotic cells except erythrocytes as these lack endoplasmic reticulum. (1) This protein is encoded in humans by the Calreticulin (CALR) gene. Through detailed analysis of somatic cell hybrids, it was found that CALR gene is located on chromosome 19 (2). To date two CALR genes have been identified – CALR 1 and CALR 2. The exact function of CALR 2 gene has yet to be determined. The CALR 1 gene encodes the 46KDa protein (4). In humans, the CALR gene spans 4.6kb of the genomic DNA and is composed of nine exons. (5,6) and eight introns (3). The CALR protein has been found to have Ro/SSA ribonucleoprotein complex properties. Itoh et al found that within the CALR gene, separate genes encode the 60-kd form and the 52 –kd forms of Ro/SSA antoantibodies subsequently mapping to chromosome 1 (TROVE2;600063) and chromosome 11 (TRIM21;109092) respectively. (5,7)
Structure

A. The CALR gene

This gene is made up of nine exons and eight introns. It spans 3.6kb of human genomic DNA. The promoter sites of the CALR gene contain several putative regulatory sites which include AP-1 and AP-2 sites, GC rich areas which include an Sp1 site, a H4TF-1 site and also four CCAAT sequences. Both H4TF-1 and AP-2 recognition sequences have been found in genes which are active in cellular proliferation.

B. Calreticulin Protein

The calreticulin (CALR) protein is a multifactorial protein which is localised to the intracellular, cell surface and extracellular compartments. It regulates a wide variety of biological processes including antigen processing and presentation to the adaptive immune system, cellular proliferation, cell adhesion, uptake of CALR expressing cancer cells by dendritic cells and phagocytosis of apoptotic cells. It contains three main structural domains:

1. The first third of the protein is called the N Terminal. N terminal is responsible for interactions with other proteins and the C terminal. This domain contains eight anti-parallel β strands and interacts with the DNA binding site of steroid receptors and α integrins. Additionally the disulphide bonds formed by cysteine residues within this domain have been found to interact with the P domain to form the vital chaperone function of CALR.

2. The second third of the protein is the P Domain. This proline rich P domain contains two sets of three repetitive amino acid sequence regions. These regions form lectin like structures which are responsible for the protein folding function of CALR. This domain contains a region that binds calcium with high affinity. It also contains several KPEDWU repeats, several ‘PEST’ regions and a putative nuclear localisation signal.

3. Finally, the C terminus is rich in acidic amino acids and contains multiple calcium binding sites. It is a highly acidic region and 37 of the final 57 residues are aspartic or glutamic acid. This domain regulates the calcium levels within the endoplasmic reticulum (ER). It binds to calcium with high capacity and low affinity. This terminus contains a four amino acid (lysine, aspartate, glutamate and leucine) KDEL ER retention signal.
motif which prevents CALR from being secreted from the ER (12). These three domains are illustrated in Figure 2.

The CALR promoter region contains multiple binding sites for transcription factors. Many of these factors have been identified as essential modulators for CALR expression including NKx2.5 which encodes a homeobox-containing transcription factor and defects of which can result in Tetralogy of Fallot, atrial septal defects with atroventricular conduction defects and congenital hypothyroidism (13). MEF2C is a transcription activator and is present in the regulatory regions of multiple muscle specific genes. It binds specifically to the MEF2 element within these genes. It is involved in vascular development and cardiac morphogenesis and myogenesis (14), GATA6 (a member of the zinc transcription factor family which plays a vital role in organogenesis and cellular differentiation during vertebral development (15), and Evi-1 an oncoprotein and transcription factor which is involved in cell differentiation, apoptosis, proliferation, cellular development and haematopoiesis (11,16).

**Biological Functions of CALR protein**

(A) *Intracellularly*

CALR acts as a major calcium binding protein within the lumen of the endoplasmic reticulum, it binds >50% of calcium present in the ER (12). In doing so it ensures newly synthesised proteins are correctly folded, become proteosome resistant and therefore enables the proteins to carry out their required intracellular function. (12) CALR protein has been found to regulate p53 expression, localisation and function (12) and also possibly plays a role in transcription regulation within the nucleus of the cell. (17).

CALR protein has been found to act as a molecular chaperone within the ER where it binds to the newly synthesised glycoproteins thereby preventing their aggregation and facilitating correct protein folding and subsequent function. CALR works in harmony with calnexin and ERp57 to create a protein forming cycle which controls protein folding. (3)

Waser et al determined that the CALR gene, which regulates the CALR protein, is activated by A23187-, bradykinin dependent Ca2+ depletion of intracellular calcium or thapsigargin in both *in vivo* and *in vitro* experiments (17). It has also been determined that CALR mRNA and protein levels increase according to endoplasmic reticulum store depletion.
(B) Extracellularly

Outside the endoplasmic reticulum, CALR protein has been detected on the cell surface and within the cytosol. (18) Here it modulates the expression of N-cadherin and vinculin, two transmembrane proteins which are vital in cell adhesion. It also modulates the transcriptional activity of steroid receptors and other transcription factors and mediates the nuclear export of the glucocorticoid receptor. (3)

It has also been found that early disruption of the CALR protein can be embryonically lethal, showing decreased ventricular wall thickness and intertrabecular recesses within the ventricular walls. (19) This occurs due to the disruption of CALR which leads to abnormal ER calcium availability thereby impairing myofibrillogenesis. (20)

From this we can conclude that the CALR protein is a multifunctional protein which acts intracellularly where it plays a vital role in calcium homeostasis and also facilitates protein folding, quality and control. Extracellularly CALR plays a vital role in cell adhesion, modulation of gene expression, nuclear export, cardiogenesis and immunogenic cell death. (3)

Mutation of the CALR gene

In 2005, JAK2-V617F was discovered. Prior to this minimal information was available on the molecular pathogenesis of myeloproliferative neoplasms (MPN). In 2013, two studies carried out by Nangalia et al and Klampf et al (21,22) discovered somatic recurrent insertions/deletions which exclusively affected exon 9 in the CALR gene in 70-84% of wild type JAK2 and MPL Primary Myelofibrosis and essential thrombocytopenia. (21,22,23)

Klampfl et al (22) discovered the CALR gene mutation whilst carrying out exome sequencing on tumour samples on MPN patients DNA and matched CD3 + T lymphocyte in six patients lacking JAK2 and MPL mutations. (22) A recurring mutation was noted on the CALR gene. Further analysis by polymerase chain reaction found that 25-35% of patient who did not possess JAK2 or MPL mutations had mutations in the CALR gene. (22,24). Subsequent studies have also found that CALR mutations were detected in 60-80% of patients with essential thrombocythaemia (ET) and primary myelofibrosis.
(PMF) who lacked JAK2 and MPL mutations. (23) CALR mutations were not detected in patients with polycythaemia vera. (PV) (22,25)

Nangalia et al performed exome sequencing on 151 patients with MPN and identified CALR mutations in 70-84% of samples of MPN without JAK2 mutations. (21) They then further expanded their investigation to patients with other haematological malignancies and non-haematological malignancies and found that this mutation occurs at low rates in myeloid malignancies and not at all in non-haematological malignancies. (21)

Greater than 50 different CALR mutations have been detected to date, but the most common mutation types detected are type 1 variant/mutation (p,L367fs*47) which results from 52-bp deletion, and type 2 variant/mutation (p.K385fs*47) which results from a 5-bp TTGTC insertion within exon 9 of the gene. (5,25,26). Of patients, 45-53% harbour type 1 mutations and 32-41% of patients harbour type 2 mutations. (20). Both types of mutations cause a single base pair frameshift which results in the formation of a novel mutant C-Terminal peptide composed of a minimal 36 amino acid stretch which replaces the 27 amino acids which are lost from the normal sequence. (20) The last four amino acids of calreticulin (KDEL), which contain the ER retention signal become positively charged, the reticulum targeting KDEL sequence is abolished thereby disturbing it cellular localisation. (23)

A non mutated CALR C terminus is largely negatively charged, whereas the mutated terminus contains positively charged amino acids. Type 1 mutations eliminate all negatively charged amino acids whereby type 2 mutations maintain up to half of the positively charged amino acids. (26) By eliminating the negative charge on this C terminus the Ca $^{2+}$ binding function is impaired and the KDEL modif is lost therefore mutant CALR may have an altered subcellular localisation (22,24,26).

In MPN, several studies have found CALR mutations in patients which are located in haematopoietic stem and progenitor cells and are believed to activate the STAT5 signalling pathway. (18) By causing cytokine hypersensitivity, suggesting that mutated CALR has the ability to activate the haemopoietic cytokine signalling pathway (12).
Role of CALR in Pathological Conditions

The Philadelphia negative myeloproliferative disorders are a group of disorders which include PV, ET and PMF. PMF is characterised by abnormal proliferation of megakaryocytes, abnormal stem cell trafficking, deposition of fibrous connective tissues in the bone marrow and extramedullary haematopoiesis (27). ET is characterised by platelet overproduction resulting from hyperproliferation of megakaryocytes. PV is characterised by hyperproliferation of predominantly erythroid cells. (12).

Specific genetic mutations have been associated with myeloproliferative neoplasms. JAK2 mutations are almost invariable in PV. 50-60% of patients with PMF or ET harbour the JAK2 mutation and up to 10% harbour the Myeloproliferative leukaemia (MPL) virus oncogene. (22,28) Recently, alternative mutations have been detected in patients with sporadic ET or PMF without JAK2 or MPL alterations. (29).

How the mutated CALR, which is characterised by lower calcium binding activity and is independent of the endoplasmic reticulum motif KDEL (30), causes the overproduction of abnormal megakaryocytes and platelets is the major question. It has been hypothesised that the pathogenetic effects have been found to be somewhat attributable to the JAK/STAT signalling. This is supported by findings that showed a link between increased STAT5 phosphorylation and CALRdel52 mutation in interleukin 3 dependent murine Ba.F3 cells. This resulted in cells to become cytokine independent. (20) Recent studies in cell lines and mice of the mutant protein show the mutant activates the downstream pathways with cMPL(31,32) and show an mechanism where a mutated chaperone activates cytokine receptor signalling (33). A novel mechanism is conceived where the mutant chaperone CALR constitutively activates receptor signalling through the abnormal interaction with MPL (34).

A. CALR in primary myelofibrosis

In patients with PMF, CALR mutations are associated with better survival compared to those with JAK2 or MPL mutations. (35) However, Tefferi et al discovered that the prognostic benefit of CALR mutations is limited to patient with type 1 (52-bp) or type 1 variant of the CALR mutation rather than those with type 2 or type 2 variant mutations (36). The CALRdel52 mutation was
found to be more frequent in PMF than ET. \(^{(19)}\) The mutational frequencies of CALR in PMF have been found to be between 25-35\% and one study found that 27\% of patients harboured a CALR mutation, with 80\% harbouring type 1 and 11\% harbouring type 2. \(^{(22,35,36)}\) In PMF poor patient survival has been associated with the presence of ASXL1 mutations and also in patients with TYPE 2 CALR mutations \(^{(38)}\), with the absence of such mutations associated with a favourable survival rate.

**Patient phenotype:**

PMF patients with the CALR mutations are preferentially male, younger in age (less than 60 years of age), \(^{(24)}\) have a higher haemoglobin, leucocyte count, platelet count and an improved overall survival rate \(^{(37)}\) In comparison to patients with JAK2 and MPL mutations. \(^{(9)}\) A lower incidence of spliceosome mutations have also been noted in these patients. \(^{(29)}\) Patients with CALR mutations were noted to be less anaemic, thereby requiring fewer transfusions \(^{(24)}\). No evolutionary differences were noted between patients with the CALR mutations and JAK2 mutated genes \(^{(30)}\). See table 1 for further details on patient phenotypes.

**B. CALR in essential thrombocytopenia**

The frequency of CALR mutations in ET ranges between 15-24\% \(^{(30)}\). Two variants of the CALR mutation exist in ET, type 1, a 52-bp deletion (p.L367fs*46) and type 2a 5-bp TTGTC insertion (p.K385fs*47). Recently the frequencies of type 1 and type 2 mutations in ET have been found to be 46\% and 38\% respectively \(^{(36)}\). Type 2 has been associated with a high circulating blast percentage, increased platelet count \(^{(28)}\) leucocyte count, dynamic international prognostic scoring system and poorer survival rate. \(^{(29)}\).

**Patient phenotype**

Patient with CALR mutated ET were found to be male, younger in age (less than 60 years), higher platelet count (>1000 x 10\(^9\)/l), lower leucocyte count, lower haemoglobin and a lower incidence of cardiovascular complications or thrombosis risk than their JAK2 mutated counterparts \(^{(9,29,37)}\). No patient with the CALR mutation to date has been found to evolve from PET to PV or acute leukaemia \(^{(38)}\), whereas 29\% of patients with JAK2 mutation at
15 years were found to evolve. \(^{(30)}\) Patients harbouring CALR mutations have been found to have a higher incidence of myelofibrotic transformation as opposed to their JAK2 mutated counterparts \(^{(37)}\). In such patients pegylated interferon has been noted to have a positive clinical effect, the allele burden was noted to decrease form 43% - 19% with this treatment and therefore could be used as a biomarker to monitor therapeutic response. See table 1 for further details on patient phenotypes.

**Table 1: A comparison of patient phenotypes with both CALR and JAK2 mutations in ET and PMF\(^{(17,27)}\)**

<table>
<thead>
<tr>
<th></th>
<th>ET - CALR mutation</th>
<th>ET - JAK2 mutation</th>
<th>PMF – CALR mutation</th>
<th>PMF – JAK2 mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td>Younger &lt;50</td>
<td>Older &gt;50</td>
<td>Younger &lt;50</td>
<td>Older &gt;60</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td>male</td>
<td>female</td>
<td>male</td>
<td>male</td>
</tr>
<tr>
<td><strong>Platelets (x10^9/l)</strong></td>
<td>Higher &gt;800</td>
<td>Lower &lt;800</td>
<td>Lower (&lt;400)</td>
<td>Higher (&gt;400)</td>
</tr>
<tr>
<td><strong>Haemoglobin (g/l)</strong></td>
<td>Lower &lt;140</td>
<td>Higher &gt;140</td>
<td>Lower (&lt;120)</td>
<td>Higher (&gt;120)</td>
</tr>
<tr>
<td><strong>Leucocytes (x10^9/l)</strong></td>
<td>Lower (&lt;8.5)</td>
<td>Higher (&gt;8.5)</td>
<td>Higher (&gt;9)</td>
<td>Lower (&lt;9)</td>
</tr>
<tr>
<td><strong>Thrombosis Risk</strong></td>
<td>lower</td>
<td>higher</td>
<td>lower</td>
<td>higher</td>
</tr>
<tr>
<td><strong>Survival</strong></td>
<td>longer</td>
<td>Worse prognosis</td>
<td>longer</td>
<td>Worse prognosis</td>
</tr>
</tbody>
</table>

**C. CALR mutation in non-haematological conditions**

Calreticulin has been detected in the serum of patients suffering from Systemic Lupus Erythematous, Sjogren’s disease, coeliac disease, rheumatic disease and also various parasitic diseases. CALR has been found to associate with ribonucleoprotein complex Ro/SSA, an autoantigen which is found in most patients with Sjogren’s and SLE. It has also been found to interact with C1q, the
first component of complement, thereby activating the classical complement pathway. \(^{(39)}\)

**Diagnostic Criteria**

The revolutionary discovery of the \textit{CALR} mutation in MPN patients harbours prognostic relevance. The following tables illustrate the WHO 2008 diagnostic criteria for PV, ET and PMF. \(^{(40)}\)

\textit{Table 2: The 2008 WHO diagnostic criteria for Essential Thrombocytopenia} \(^{(40)}\)

<table>
<thead>
<tr>
<th>Major Criteria</th>
<th>Minor Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thrombocytosis &gt;450x10(^9)</td>
<td>Enlarged, mature megakaryocytes on bone marrow biopsy with no significant increase or left-shift of granulopoiesis or erythropoiesis</td>
</tr>
<tr>
<td>Patient does not meet the criteria for: PMF, PV, MDS, \textit{BCR-ABL1}-positive chronic myelogenous leukemia or another myeloid neoplasm</td>
<td>Presence of \textit{JAK2} V617F or other clonal marker or in its absence no evidence for a reactive thrombocytosis</td>
</tr>
</tbody>
</table>

All four must be present for a patient to be diagnosed with ET.

\textit{Table 3: The 2008 WHO diagnostic criteria for Polycythaemia Vera} \(^{(40)}\)

<table>
<thead>
<tr>
<th>Major Criteria</th>
<th>Minor Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemoglobin &gt;185 g/L in men, &gt;165 g/L in women or other evidence of increased red cell volume</td>
<td>Bone marrow biopsy showing hypercellularity for age with trilineage growth (panmyelosis) with prominent erythroid, granulocytic and megakaryocytic proliferation</td>
</tr>
<tr>
<td>Presence of \textit{JAK2} V617F or other functionally similar mutation (such as \textit{JAK2} exon 12 mutation)</td>
<td>Serum erythropoietin level below the reference range for normal</td>
</tr>
<tr>
<td>Endogenous erythroid colony formation \textit{in vitro}</td>
<td></td>
</tr>
</tbody>
</table>
For a patient to be diagnosed with PV they must possess both major criteria or the first major criteria and two minor criteria.

Table 4: The 2008 WHO diagnostic criteria for Primary Myelofibrosis\(^{(40)}\)

<table>
<thead>
<tr>
<th>Major criteria</th>
<th>Minor Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Megakaryocyte proliferation and atypia, usually with either reticulin and/or collagen fibrosis</td>
<td>Leukoerythroblastosis</td>
</tr>
<tr>
<td>Presence of (JAK2) V617F or other clonal marker (eg (MPL) W515K/L), OR in absence of a clonal marker, no evidence that the marrow fibrosis or other changes are reactive</td>
<td>Increase in serum LDH</td>
</tr>
<tr>
<td>Does NOT meet WHO criteria for any of the following: PV, (BCR-ABL1)+chronic myelogenous leukemia, MDS, or other myeloid neoplasm</td>
<td>Anaemia</td>
</tr>
<tr>
<td></td>
<td>Splenomegaly</td>
</tr>
</tbody>
</table>

Diagnosis of PML requires all three major criteria plus two minor criteria.

\(CALR\) mutation does not feature in these criteria due to its recent discovery. Its discovery not only provides new insights into the molecular basis of MPN but also new molecular approaches to their diagnosis and therefore inclusion into the above diagnostic criteria will be part of revised criteria of these diseases\(^{(41)}\).

**Conclusion**

William Dameshek first described myeloproliferative disorders in 1951 as a ‘related group of disease with a shared myelostimulatory factor accounting for overlapping clinical and laboratory features\(^{(42)}\). The \(JAK2\) V617F gene mutation was subsequently discovered 55 years later and noted to be present in 90-95% of PV, 40-60% of ET and PMF cases. This discovery has proved revolutionary in the diagnosis, understanding and clinical management of these patients. The recent discovery of the \(CALR\) gene mutation has further revolutionised our understanding of the myeloproliferative disorders. By identifying the role of \(CALR\) mutation in ET and PMF we have enabled the patients presenting with these conditions to be placed into one of two subgroups, A. \(CALR\) mutant ET/PMF which has an indolent clinical course and B. PMF/ET with non mutated \(JAK2\), \(CALR\) and \(MPL\) which is a very aggressive
myeloid neoplasm. (27) New studies have since determined that PCR amplification, followed by fragment length analysis can be used in patients with MPN for CALR mutations to determine their mutant allele burden (43) which could allow early diagnosis of a more aggressive phenotype and subsequent early or type specific management and therefore increase patient survival.

The discovery of $CALR$ mutations has enabled us to firstly distinguish that ET patients who present with the $CALR$ mutation are younger in age, male gender increased platelet counts and lower haemoglobin and leucocyte counts. These patients also have a significantly lower thrombosis risk and therefore may not require the introduction of antiplatelet therapy or hydroxyurea. Research carried out in 2015 by Alvaraz-Larren et al on $CALR$ positive ET patients who are at low risk of thrombosis determined that such patients did not benefit from low dose aspirin as the risk of bleeding offset the reduction in the rate of thrombosis (44). Hydroxyurea was not tested in this study. Additionally, these patients have been found to show a reduction in allele burden with the use of pegylated interferon thereby allowing this to be used as a biomarker for therapeutic response in these patients. (25) PMF patients presenting with $CALR$ mutations are now known to have a better prognosis and lower risk of evolution to acute leukaemia than patients with $JAK2$ mutations. By discovering this mutated $CALR$ gene we now can account for the genetic mutation in up to 90% of patients who present with MPN, decipher accurate management for these patients, thereby increasing patient safety and improving disease prognosis.

**Conflict of Interest:** No conflicts of interests to declare.
References


Legends to Figures

Figure 1
This diagram illustrates the human gene with 9 exons. The CALR gene is located on exon 9. (Adapted from Michalak et al\(^{(5)}\))

Figure 2:
This diagram illustrates the three functional domains of the CALR gene and their various functions. The N domain is where interaction with the DNA binding site of steroid receptors and \(\alpha\) integrins takes place. The P domain is rich in proline and forms lectin like structures which are responsible for CALR’s protein binding function and also binds Ca \(2^+\) with high affinity. The C domain is rich in acidic amino acids and regulates calcium levels within the endoplasmic reticulum. This domain terminates with the KDEL ER sequence.