haematologica
Journal of the Ferrata Storti Foundation

Germline bi-allelic *SH2B3/LNK* alteration predisposes to a neonatal juvenile myelomonocytic leukemia-like disorder

by Chloé Arfeuille, Yoann Vial, Margaux Cadenet, Aurélie Caye-Eude, Odile Fenneteau, Quentin Neven, Adeline A. Bonnard, Simone Pizzi, Giovanna Carpentieri, Yline Capri, Katia Girardi, Lucia Pedace, Marina Macchiaiolo, Kamel Boudhar, Monia ben Khaled, Wadih Abou Chahla, Anne Lutun, Mony Fahd, Séverine Drunat, Elisabetta Flex, Jean-Hugues Dalle, Marion Strullu, Franco Locatelli, Marco Tartaglia, and Hélène Cavé

Received: July 13, 2023.
Accepted: November 7, 2023.

Citation: Chloé Arfeuille, Yoann Vial , Margaux Cadenet, Aurélie Caye-Eude, Odile Fenneteau, Quentin Neven, Adeline A. Bonnard, Simone Pizzi, Giovanna Carpentieri, Yline Capri, Katia Girardi, Lucia Pedace, Marina Macchiaiolo, Kamel Boudhar, Monia ben Khaled, Wadih Abou Chahla, Anne Lutun, Mony Fahd, Séverine Drunat, Elisabetta Flex, Jean-Hugues Dalle, Marion Strullu, Franco Locatelli, Marco Tartaglia, and Hélène Cavé. Germline bi-allelic SH2B3/LNK alteration predisposes to a neonatal juvenile myelomonocytic leukemia-like disorder.
Haematologica. 2023 Nov 16. doi: 10.3324/haematol.2023.283917 [Epub ahead of print]

Germline bi-allelic *SH2B3/LNK* alteration predisposes to a neonatal juvenile myelomonocytic leukemia-like disorder

Chloé Arfeuille[1,2], Yoann Vial[1,2], Margaux Cadenet[1,2], Aurélie Caye-Eude[1,2], Odile Fenneteau[3], Quentin Neven[4], Adeline A Bonnard[1,2], Simone Pizzi[5], Giovanna Carpentieri[5], Yline Capri[6], Katia Girardi[7], Lucia Pedace[7], Marina Macchiaiolo[8], Kamel Boudhar[9], Monia ben Khaled[10], Wadih Abou Chahla[11], Anne Lutun[12], Mony Fahd[4], Séverine Drunat[1], Elisabetta Flex[13], Jean-Hugues Dalle[4], Marion Strullu[2,4], Franco Locatelli[7,14], Marco Tartaglia[5], Hélène Cavé[1,2]

[1]Département de Génétique, Unité de Génétique Moléculaire, Hôpital Robert Debré, Assistance Publique des Hôpitaux de Paris (AP-HP), Paris, France

[2]INSERM UMR_S1131, Institut de Recherche Saint-Louis, Université Paris-Cité, Paris, France

[3]Service d'Hématologie Biologique, Hôpital Robert Debré, Assistance Publique des Hôpitaux de Paris (AP-HP), Paris, France

[4]Service d'Onco-Hématologie pédiatrique, Hôpital Robert Debré, Assistance Publique des Hôpitaux de Paris (AP-HP), Paris, France

[5]Molecular Genetics and Functional Genomics, Bambino Gesù Children's Hospital IRCCS, 00146 Rome, Italy.

[6]Département de Génétique, Unité de Génétique clinique, Hôpital Robert Debré, Assistance Publique des Hôpitaux de Paris (AP-HP), Paris, France

[7]Department of Hematology/Oncology and Cell and Gene Therapy, Bambino Gesù Children's Hospital IRCCS, 00146 Rome, Italy.

[8]Rare Diseases and Medical Genetics, Bambino Gesù Children's Hospital IRCCS, 00146 Rome, Italy.

[9]Service de réanimation néonatale, Hôpital Central de l'Armée, Alger, Algérie

[10]University of Tunis El Manar, Faculty of Medicine of Tunis, 1007, Tunisia. Pediatric Immuno-Hematology Unit, Bone Marrow Transplantation Center Tunis, Tunis, Tunisia

[11]Service d'Hématologie Pédiatrique, Centre Hospitalier Universitaire de Lille, Lille, France

[12]Service d'Hématologie Pédiatrique, Centre Hospitalier Universitaire d'Amiens, Amiens, France

[13]Department of Oncology and Molecular Medicine, Istituto Superiore di Sanità, 00161 Rome, Italy

[14]Department of Pediatrics, Catholic University of the Sacred Hearth, 00168 Rome, Italy.

Running heads: SH2B3/LNK in JMML

Corresponding author:

Hélène Cavé, UF de Génétique Moléculaire, Hôpital Robert Debré, 48, Boulevard Sérurier, 75019 Paris, France. Tel: 33 (0)1 40 03 57 11; Fax: 33 (0)1 40 03 22 77; e-mail: helene.cave@aphp.fr

Contributions

CA designed experiments, produced and analysed data, wrote the manuscript. YV, MC, ACE, AAB, SP, GC produced and analysed data. OF reviewed cytological data. YC, PD, KG, LP, MM, LB, KB, MbK, WAC, AL, MF, JHD provided clinical assessments. QN collected and reviewed clinical data. SD and EF supervised functional analyses, MS and FL coordinated and reviewed clinical data. MT supervised genomic analyses, coordinated the project, raised funds, wrote the manuscript. HC designed and coordinated the project, supervised analyses and data collection, raised funds, wrote the manuscript. All authors reviewed the manuscript.

Data sharing statement

Whole genome sequencing, RNA-sequencing and DNA methylation data are deposited on Sequence Read Archive (SRA) (https://www.ncbi.nlm.nih.gov/sra). Access number SUB13849483

Acknowledgement

We deeply thank patients and families for their participation to the research. We thank Stacey Foulane and Aldjia Assous for help in analyses and clinical data collection respectively. We thank Dr Bruno Cassinat (Laboratoire de Biologie Cellulaire, Hôpital Saint-Louis) for *in vitro* culture of myeloid progenitors. We thank the CRB-K of the Robert Debré Hospital (CRB-K, BB-0033-00076).

Funding

This work was supported, in part, by grants from AIRC (IG-21614 and IG-28768), EJP-RD (NSEuroNet), and Italian Ministry of Health (RF-2021-12374963).

Conflict of interest

The authors declare no conflict of interest

Abstract

Juvenile myelomonocytic leukemia (JMML) is a rare, generally aggressive myeloproliferative neoplasm affecting young children. It is characterized by granulomonocytic expansion, with monocytosis infiltrating peripheral tissues. JMML is initiated by mutations upregulating RAS signaling. Approximately 10% of cases remain without an identified driver event. Exome sequencing of 2 unrelated cases of familial JMML of unknown genetics and analysis of the French JMML cohort identified 11 patients with variants in *SH2B3*, encoding LNK, a negative regulator of the JAK-STAT pathway. All variants were absent from healthy population databases, and mutation spectrum was consistent with a loss of function of the LNK protein. A stoploss variant was shown to affect both protein synthesis and stability. The other variants were either truncating or missense, the latter affecting the SH2 domain that interacts with activated JAK. Of the 11 patients, 8 from 5 families inherited pathogenic bi-allelic *SH2B3* germline variants from their unaffected heterozygous parents. These children represent half of the cases with no identified causal mutation in the French cohort. They displayed typical clinical and hematological JMML features with neonatal onset and marked thrombocytopenia. They were characterized by absence of additional genetic alterations and a hypomethylated DNA profile with fetal characteristics. All patients showed partial or complete spontaneous clinical resolution. However, progression to thrombocythemia and immunity-related pathologies may be of concern later in life. Bi-allelic *SH2B3* germline mutations thus define a new condition predisposing to a JMML-like disorder, suggesting that the JAK pathway deregulation is capable of initiating JMML, and opening new therapeutic options.

Keywords: SH2B3, LNK, JMML, childhood leukemia, genetic predisposition.

Introduction

Juvenile myelomonocytic leukemia (JMML) is a rare, aggressive myeloproliferative neoplasm affecting infants and young children. It is characterized by excessive granulomonocytic proliferation in bone marrow (BM) and peripheral blood (PB) leading to splenomegaly, leukocytosis with precursors in peripheral blood, monocytosis, infiltration of peripheral tissues with histiocytes and normal or moderately increased blast count[1,2].

The natural course of JMML is generally rapidly fatal. The only potentially curative treatment is BM transplantation (BMT). However, the disease is characterized by a highly heterogeneous course, with a third of patients progressing to acute leukemia, while about 10% have indolent forms or even spontaneous resolutions[3].

JMML arises from the hematopoietic stem or myeloid progenitor cells[4–6]. It is initiated by abnormal activation of RAS signaling, leading to *in vitro* hypersensitivity of myeloid progenitors to GM-CSF[7]. RAS pathway hyperactivation is due to mutations in genes encoding RAS proteins (*NRAS, KRAS*) or regulators (*PTPN11*, *NF1* or *CBL*)[8,9], that define genetic and clinical subgroups. More rarely, activating mutations affecting other small G proteins (RRAS*,* RRAS2*,* RIT1)[10] have also been described, as well as fusions causing activation of genes coding for transducers upstream of the RAS pathway[11,12].

A particular feature of JMML is their frequent occurrence in the context of a predisposing genetic syndrome, including Noonan syndrome (NS), neurofibromatosis type 1 and CBL syndrome, which are due to a constitutional upregulation of the RAS-MAPK pathway (the so-called RASopathies)[13–16].

The identification of an activating mutation affecting the RAS pathway confirms the diagnosis of JMML in over 90% of cases[17]. However, there are still a small number of cases in whom no mutation has been identified (e.g., 5% in the French JMML cohort).

SH2B3 (OMIM 605093) encodes LNK (lymphocyte adaptor protein), an SH2-domain adaptor protein acting as a negative regulator of intracellular signaling promoted by cytokines and growth factors activating the JAK-STAT pathway[18]. Earlier work suggested that Lnk

negatively regulates normal hematopoietic stem and progenitor cell (HSPC) expansion and self-renewal[19,20]. Somatic *SH2B3* variants have been reported in some JMML cases[21] and *Sh2b3*$^{-/-}$ mice develop a myeloproliferative neoplasm (MPN)[22,23]. Germline mono-allelic[24,25,26,27], or more rarely bi-allelic[28,29] LNK loss of function (LoF) mutations have been previously identified in various hematological diseases.

Here we provide evidence that bi-allelic germline inactivating variants in *SH2B3* underlie a disease predisposing to a JMML-like disorder. We report 11 patients with *SH2B3*-associated JMML, including 8 from 5 families who inherited bi-allelic pathogenic germline *SH2B3* variants. *SH2B3*-associated JMML represents a new specific entity characterized by extremely early onset, probable persistence of fetal features, and partial or complete spontaneous clinical resolution. JMML in this group of patients is clearly distinct from other JMMLs as a whole, but also from the disease associated with germline mono-allelic or somatic mutations of *SH2B3* that have previously been reported.

Methods

Patients

The study included 234 patients with JMML referred to the French reference laboratory (1995-2023), and one family with 2 probands from the Bambino-Gesù Children's Hospital (OPBG), Rome, Italy.

Diagnosis of JMML was based on clinical and hematological findings, centrally reviewed cytomorphological examination of PB and BM smears, and genetic screening of genes known to initiate JMML or differential diagnoses (Wiskott-Aldrich syndrome, *GATA2*, osteopetrosis). Absence of *BCR*::*ABL1* and *KMT2A* rearrangement was checked. In patients with no RAS-related variant, the presence of rare fusions involving *FLT3, ALK, PDGFRA/B*, was assessed by RNAseq analysis. Karyotyping was performed using standard procedures.

All patients fulfilled the WHO JMML criteria.

Ethical approval

The study was approved by the institutional review board (IRB) of "Hôpitaux Universitaires Paris Nord Val-de-Seine", Paris University, AP-HP (IRB: 00006477) and "Ospedale Pediatrico Bambino Gesù (Ref: 1702_OPBG_2018), in accordance with the Helsinki declaration.

Samples

BM or PB samples were obtained at JMML diagnosis and in patients' parents. Mononuclear cells were isolated on a Ficoll gradient (Eurobio). Constitutional origin of the variant(s) was assessed using skin derived-fibroblasts. Healthy control used for methylation and RNAseq assays were males of 9, 10 and 12 year old. The fetal BM sample was obtained from a spontaneous abortion at 15 weeks of gestation. Supplementary Figure 1 shows the analyses performed in patients.

Genome-wide DNA array analysis

Genomic DNAs samples were analyzed by single-nucleotide polymorphism (SNP) array technologies using the Genome-Wide GeneChip Human SNP Array 6.0 (Affymetrix), and/or high-density array comparative genomic hybridization (CGH)[12].

Whole-exome sequencing (WES) and whole genome sequencing (WGS)

Target enrichment was performed using the SureSelect Human All Exon V4+UTRs or SureSelect AllExon V5 (Agilent Technologies, Santa Clara, CA, USA), and captured regions were sequenced with a HiSeq2000 or NextSeq500 instrument (Illumina, San Diego, CA, USA), as previously described[12,33]. See details in Supplementary Material.

Sequencing of mononucleated cells mRNA (mRNAseq)

Libraries were prepared with NEBNext Ultra II Directional RNA Library Prep Kit, according supplier recommendations. Paired-end 100-bp reads sequencing was performed on a NovaSeq platform. Image analysis and base calling was performed using Illumina Real Time Analysis (3.4.4) with default parameters. The bioinformatic pipeline is described in supplementary information. Unsupervised analyses of gene expression and differential

expression were conducted using the Galileo tool (IntegraGen) (see supplementary information).

DNA methylation (capture EM-seq)

Libraries were prepared using the Twist Targeted Methylation sequencing protocol system (Twist Bioscience) according to the guidelines and subjected to paired-end sequencing with 100-bp reads on a Novaseq 6000 platform (see supplementary information). Unsupervised classification based on the 1000 most variant 100bp tiles was generated using hierarchical clustering (cosine distance, Ward method) in R software.

More methods details are provided in the Supplementary material

Results

Exome sequencing in two families segregating JMML identifies homozygous mutations in *SH2B3*

Trio exome sequencing was undertaken in proband #79.1 from the French JMML cohort, who suffered a syndromic neonatal JMML with no identified mutation. The analysis revealed the presence of a homozygous variant (c.1160G>C, p.Gly387Ala) in *SH2B3* in the affected child. A review of the child's records confirmed that he met the consensus clinical and hematological criteria for the diagnosis of JMML, with hepatosplenomegaly, massive monocytosis (14×10^9/L), blasts, circulating myeloid precursors and thrombocytopenia (21×10^9/L) (Figure 1 A; Supplementary Table 1). The BM smear showed hypercellularity associated with a decreased number of megakaryocytes. Spontaneous *in vitro* growth of myeloid progenitors was positive. His brother (#79.2), born a few years later, also presented with neonatal JMML (Figure 1). Genetic analysis showed that he had inherited the familial *SH2B3* mutation in a bi-allelic pattern.

Independently, WES was performed in a family having two siblings diagnosed with an unclassified neonatal-onset syndromic JMML at the OPBG (Rome, Italy). The proband,

#OPBG-2, was diagnosed at the age of one month, with elevated leukocyte counts, monocytosis (4.2×10^9/L), thrombocytopenia (41×10^9/L) and 3% myeloblasts (Figure 1 B; Supplementary Table 1). Immature monocytes and mild dysgranulopoiesis were observed on PB. Hypercellular BM, with granulocytic expansion, and 7% blasts, was consistent with diagnosis of JMML. Similarly, his 5 year-old sister, #OPBG-1, had been diagnosed with JMML in the first days of life. Following delivery, she had splenomegaly and respiratory difficulties, which necessitated 24h intubation. Blood count showed leukocytosis (56×10^9/L), anemia (11.8 g/dL), and thrombocytopenia (35×10^9/L) (Supplementary Table 1). She experienced spontaneous resolution under active surveillance. Hypersensitivity of myeloid progenitor cells to GM-CSF was observed in both children. WES analysis revealed the presence of a homozygous frameshift variant (c.1709dupA, p.Asn570LysfsTer82) in *SH2B3* in both siblings.

Targeted sequencing of *SH2B3* in the whole French JMML cohort

SH2B3 sequencing was then extended to the 234 patients of the French JMML cohort. Filtering out variants with a frequency $>10^{-4}$ in the general population (GnomAD Non cancer v.2.1.1 database), at least one variant was identified in 9/234 (3.8%) patients, including 6 of the 12 (50%) mutation negative cases (Supplementary Figure 2). Searching for mutations in patients' fibroblasts and/or in their parents revealed that, in 8 of the 9 patients, the *SH2B3* variants identified were germline. Of these 8 patients, 6 (from 4 families) carried bi-allelic *SH2B3* variants (Figure 1C). In the remaining 2 patients, the germline *SH2B3* variant was mono-allelic. WGS was performed that ruled out an alteration of the second *SH2B3* allele.

Variant segregation in 6 families showed that germline variants were always inherited from unaffected heterozygous parents (Figure 1C). In 4 families, all affected children inherited biallelic *SH2B3* variants, consistent with an autosomal recessive transmission of JMML predisposition. Three of these families were consanguineous. In family #48, the mono-allelic germline *SH2B3* variant in the affected child was inherited from the unaffected father.

Only one of the 234 patients (#209) had exclusively a somatic *SH2B3* variant.

Assessing variant pathogenicity

Including the Italian family, a total of 7 different *SH2B3* variants were identified. Among them, 3 were truncating (frameshift indels or stop codon), 3 were missense, and 1 caused protein elongation with a divergent C-terminus (Table 1; Figure 2A). Each variant was found on a single occurrence except p.Asp231Glyfs*39, which was found in 2 unrelated patients. All changes were predicted to be deleterious; none was referenced in the general population (Table 1), and importantly no homozygous truncating variant was referenced in gnomAD.

Based on the bi-allelic occurrence of variants, a LoF effect was postulated as functional consequence of the identified mutations. In line with this hypothesis, transient transfection experiments performed in 293T cells documented a dramatically reduced stability of the p.Asn570Lys*82 LNK mutant as compared to the wild-type protein (Figure 2B). Besides this effect at the protein level, assessment of the *SH2B3* mRNA levels by real-time PCR in primary fibroblasts obtained from patient #OPBG-2 and an unaffected control provided evidence of a significantly reduced amount in the former, which was suggestive of a reduced stability of the transcript carrying the variant (Figure 2C). Overall, these findings indicate a disruptive impact in terms of transcript stability/processing and protein stability. Similarly, all missense variants were found to target the SH2 domain, which mediates binding to phosphorylated JAK and contains only rare polymorphisms in the general population (Supplementary Figure 3). All 3 variants were predicted to damage the structure of the SH2 domain of the LNK protein (Figure 2D), likely resulting in a decreased capacity to bind its JAK2 target.

Altogether, these findings are consistent with a LoF of *SH2B3* as a driver mechanism underlying JMML.

Hematologic and clinical features of patients with *SH2B3* mutations

All patients presented with splenomegaly, high monocyte counts and hematological data meeting the updated WHO consensus criteria for JMML diagnosis[31] (Supp Table 1). Patients with bi-allelic germline mutations (n=8) had higher WBC counts ($p<0.001$), monocyte ($p=0.045$) and lymphocyte counts ($p=0.004$) than those with mono-allelic/somatic (n=3)

variants or patients from the whole JMML cohort (Table 2). This also holds true when compared to patients with NS-associated JMML (Table 2) or with any of the genetic subgroups of JMML taken separately[9]. Notably, all patients with *SH2B3* alterations but one displayed marked thrombocytopenia and the median platelet count was significantly lower than in the whole JMML cohort (28×10^9/L vs 64×10^9/L; p=0.02), with frequent lack of megakaryocytes in patients' BM (Supp Table 1). BM cellularity was normal or increased, with predominant granulocytic proliferation. Mild dysplastic features were observed in about half of the patients. Blast cells were moderately elevated but lower than 20% in blood and BM.

In patients with bi-allelic germline mutations, JMML onset was neonatal. Most patients (5/6) were not transplanted and underwent spontaneous resolutions.

In striking contrast, the median age of onset of JMML with mono-allelic germline or somatic *SH2B3* variants was 3.8 years, which is later than patients with bi-allelic germline *SH2B3* variants. The former had a markedly more severe course compared to the latter, and have all undergone transplantation.

Longitudinal follow-up of patients with bi-allelic *SH2B3* germline variants shows that after an initial stage of thrombocytopenia associated with profound impairment of the megakaryocytic lineage, the children's platelet counts increased rapidly, reaching above-normal levels in several of them, with even thrombocytosis in patient #79.2 (Supplementary Figure 4).

Genetic makeup of *SH2B3* mutated JMML

To determine the mutational makeup of JMML associated with *SH2B3* variants, genome-wide DNA array analysis and high-depth sequencing of recurrently mutated genes in JMML were carried out. WGS was also performed on patient with available germline sample (Supplementary Figure 1). All patients with bi-allelic *SH2B3* variants had normal karyotype and none had deleterious variant in JMML known drivers. Extensive genetic analysis did not evidence any other germline or somatic pathogenic or likely pathogenic variant or CNA (Figure 3).

In contrast, patients with a mono-allelic germline *SH2B3* variant acquired several additional somatic variants. Regardless their germline or somatic status, monoallelic *SH2B3* mutations were systematically associated with a *PTPN11* mutation (Figure 3).

Notably, 2 patients (#53; #209) with a mono-allelic *SH2B3* truncating variant and a *PTPN11* mutation subsequently acquired chromosome 12q aUPD (12q21.1-q24.31 and 12q13.2-q24.31, respectively), resulting in copy-neutral loss of heterozygosity (LOH) of the *SH2B3* variants, but also of *PTPN11*, which is located on chromosome 12q24.13, close to *SH2B3* (12q24.12). Such a pattern of events is rarely found in patients with *PTPN11* mutations, which have a gain-of-function behavior, but is a typical mechanism for the complete inactivation of tumour suppressor genes. In our patients, it leads to the bi-allelic loss of function of LNK. Taken together, this supports the view that the loss of function of *SH2B3* is a driving event in JMML in these cases as well. It is likely that the proximity to *PTPN11* is the reason why a copy-neutral LOH is privileged over a simple deletion of *SH2B3* in these cases.

DNA methylation profiling

DNA methylation was also analyzed. A reference group of 54 JMML of known genetic groups and 2 healthy samples (one postnatal BM, one prenatal BM) were used for comparison. As expected, hierarchical clustering of reference samples delineated three main groups, corresponding to methylation-low, intermediate and high JMML subtypes. The patients with mono-allelic germline (#53) or somatic (#209) *SH2B3* mutation clustered within the intermediate and hypermethylated groups, respectively. In contrast, samples with bi-allelic germline *SH2B3* variants defined a distinct cluster (cluster 1) with the strongest hypomethylation pattern (Figure 4A). Interestingly, the prenatal BM clustered with the SH2B3 cluster whereas the postnatal BM clustered in the cluster 3 (Figure 4A).

Altogether, methylation profiling is consistent with the idea that JMML with bi-allelic germline *SH2B3* mutations represents a specific entity that may be linked to prenatal features, in line with the neonatal presentation.

RNAseq of JMML mononucleated cells

Total RNA sequencing was performed on mononuclear cells obtained at diagnosis in 4 patients with bi-allelic germline *SH2B3* mutations. The gene expression profile was compared to that of 17 reference JMML samples and 2 healthy postnatal BM samples. PCA analysis showed that the 4 *SH2B3* cases clustered within the group of JMML samples (Figure 4B). Analysis of differential gene expression between *SH2B3* JMML and postnatal BM samples evidenced 447 genes overexpressed compared with 717 genes under expressed in *SH2B3* JMML cases (Figure 4C). Interestingly, the oncofetal transcript *LIN28B* and its two targets *HBG2* and *IGF2BP1* were among the most overexpressed genes in *SH2B3* JMML samples. When compared with other JMML cases, the number of over-expressed genes fell to only 13, with the oncofetal transcript *IGF2BP1* ranking first (Figure 4D).

Discussion

SH2B3 encodes the lymphocyte adaptor protein LNK, a member of the SH2B adaptor family of proteins. LNK is predominantly expressed in hematopoietic stem and progenitor cells and is a key negative regulator of numerous cytokine and growth factor receptors[36].

We show here that bi-allelic germline LoF *SH2B3* mutations represent a new disease predisposing to a JMML-like disorder with autosomal recessive inheritance. In our cohort, it accounted for 50% of JMML cases with unsolved genetics, and all cases experienced spontaneous resolution, in line with a hypomethylated DNA pattern indicative of a favorable prognostic[37].

JMML associated germline *SH2B3* mutations were frameshift or missense targeting the SH2 domain. The SH2 domain is required for LNK to bind to the phosphorylated tyrosine (pY813) of JAK2 or JAK3 when activated in response to cytokines. Variants targeting the SH2 domain have been reported to cause severe disruption to LNK function[20]. All variants were thus consistent with LoF. The deficiency in LNK leads to increased JAK/STAT signaling in a

cytokine-independent manner[36], and mice deficient in *Lnk* were demonstrated to have a variety of hematopoietic phenotypes, including an expansion of HSPC compartment[19,38], which is consistent with JMML[5,6]. It is now known that LNK controls the expansion of HSC and myeloid progenitors predominantly by modulating the thrombopoietin (TPO) signaling pathway[20]. Indeed, TPO which was initially identified as the cytokine that stimulates megakaryopoiesis and platelet production has also been shown to be a major regulator of HSC self-renewal and quiescence[20].

Both germline and somatic LNK LoF mutations have been previously identified in a range of hematological diseases[24]. Heterozygous *SH2B3* germline variants, some of which are relatively common in the general population, have been associated with a low penetrance increased risk of MPN[25] and idiopathic erythrocytosis (IE)[26,27]. Homozygous pathogenic variants, by contrast, are absent from reference population databases and have scarcely been described in patients. The first case reported is that of a consanguineous family in which two siblings carrying D231Profs*38 presented with growth retardation, mild developmental delay and autoimmune disorders[28]. Though this was not explicitly presented as a JMML, the proband developed a phenotype very similar to our patients with high WBC count, <20% blasts, thrombocytopenia, and no megakaryocytes in BM at 4 weeks old. Blood cell counts normalized without intervention by age 14 months; however, he then developed liver cirrhosis alike one of our patients.

More recently, two unrelated patients with homozygous germline LoF *SH2B3* variants R148Profs*40 and V402M were described with developmental delay, hepatosplenomegaly, myeloproliferation and autoimmune disorders[29]. Paradoxically, unlike other patients, a major thrombocytosis with megakaryocytic hyperplasia was noted in the BM of these patients. Similarly, even though our patients initially present with deep thrombocytopenia they progress after JMML spontaneous resolution towards thrombocytosis. This could therefore represent two successive phases of the same disease.

How LNK-deficient patients progress from a major defect in platelet production to excess production is unclear. The effect of LNK loss on megakaryopoiesis may be age-

dependent[39,40]. Indeed, the role of TPO appears to be different in fetal HSCs, which have a high capacity for proliferation and self-renewal for up to 6 months after birth, and in adult HSCs which are maintained quiescent[41]. In this respect, it is interesting to note that JMML in patients with bi-allelic germline *SH2B3* mutations all had neonatal onset and retained fetal characteristics, such as a DNA methylation pattern resembling that of fetal BM and expression of the oncofetal transcripts *LIN28B* and *IGF2BP1*[42] . The progressive loss of this fetal context may then explain the spontaneous regression of JMML and the switch to thrombocytosis.

RASopathies may be associated with immune disease[43], and the combination of spontaneously resolving forms of JMML associated with a propensity to subsequently develop immune-related disorders is particularly reminiscent of what is observed in patients with CBL syndrome[16] . Interestingly, LNK recruits the CBL E3 ubiquitin ligase via its tyrosine residue at position 572 and thus inactivates JAK2 by degradation via the proteasome[44]. This functional link has recently been extended to RAS activation. Indeed, JAK2 stabilization induced by CBL LoF dramatically enhances palmitoylation of NRAS thereby activating the RAS-MAPK signaling pathway and promoting myeloid leukemogenesis[45]. Such interplay between the JAK-STAT and RAS-MAPK signalling pathways highlights how the loss of LNK could induce JMML whose pathogenesis is considered to be strictly linked to activation of the RAS pathway[46].

Besides immune-hematological defects, several of our patients, as well as those reported in the literature[28,29], had multiple extra-hematological disorders suggestive of a syndromic presentation. However, these signs were inconstant and differed widely from one patient to another, even within a family. It is therefore difficult to incriminate loss of *SH2B3* function in these associated features, which could simply result from the increased risk of autosomal recessive disorders in a context of consanguinity.

Our analysis also identified JMML patients with mono-allelic *SH2B3* variants. Regardless of their germline or somatic status, these patients had a strikingly different presentation than those with bi-allelic germline variants, with a later age of onset and a much more severe

clinical course requiring BMT. In line with clinical presentation, the mutational make-up and DNA methylation profile of patients with *SH2B3* mutations highly differed depending on the allelic pattern. Unlike patients with bi-allelic germline *SH2B3* mutation, for whom this alteration seems to represent the only initiating driver of the JMML-like disorder, patients with mono-allelic germline or somatic *SH2B3* variants presented with multiple somatic alterations, systematically including a somatic variant in *PTPN11*, though we could not determine which lesion was the initiating driver.

In conclusion, our findings show that *SH2B3* can act as a driver in JMML in 2 ways, highlighting the importance of deregulation of the JAK-STAT pathway in this disease. On the one hand, mono-allelic *SH2B3* alteration, either germline or somatic, drive a severe 'classical' JMML associated with frequent loss of heterozygosity and additional somatic mutation. On the other hand, bi-allelic germline variants of *SH2B3* define a singular recessive clinical entity associated with a predisposition to develop a neonatal JMML-like disorder. This new entity is estimated to account for half of the cases of JMML that remained unresolved on the genetic level. Its biology and clinical course distinguish it clinically from cases of JMML associated with somatic or constitutional mono-allelic pathogenic *SH2B3* variants. The absence of secondary somatic changes is reminiscent of the JMML-like disorder seen in some patients with NS. However, apart from the neonatal presentation, none of the unique hematological features of patients with germline *SH2B3* bi-allelic mutations are shared with NS-JMML[47]. Furthermore, all of our cases had a favorable outcome, which is less common in NS-JMML[47] . This, together with the immunological landscape, makes them clinically more similar to patients with CBL syndrome. Indeed, strikingly, all previously reported patients were also described to develop autoimmune disorders (thyroiditis, hepatitis) at a later age[28,29]. Similarly, although their follow-up is still short, one of the present patients developed autoimmune thyroiditis and another presented with a severe liver disorder requiring transplantation. Thus, in addition to JMML-like disorder, this disease seems to cause serious immune disorders, which should be monitored to favor a prompt management and effective care.

References

1. Chang TY, Dvorak CC, Loh ML. Bedside to bench in juvenile myelomonocytic leukemia: insights into leukemogenesis from a rare pediatric leukemia. Blood. 2014;124(16):2487-2497.

2. Locatelli F, Niemeyer CM. How I treat juvenile myelomonocytic leukemia. Blood. 2015;125(7):1083-1090.

3. Locatelli F, Nöllke P, Zecca M, et al. Hematopoietic stem cell transplantation (HSCT) in children with juvenile myelomonocytic leukemia (JMML): results of the EWOG-MDS/EBMT trial. Blood. 2005;105(1):410-419.

4. Lapidot T, Grunberger T, Vormoor J, et al. Identification of human juvenile chronic myelogenous leukemia stem cells capable of initiating the disease in primary and secondary SCID mice. Blood. 1996;88(7):2655-2664.

5. Caye A, Rouault-Pierre K, Strullu M, et al. Despite mutation acquisition in hematopoietic stem cells, JMML-propagating cells are not always restricted to this compartment. Leukemia. 2020;34(6):1658-1668.

6. Louka E, Povinelli B, Rodriguez-Meira A, et al. Heterogeneous disease-propagating stem cells in juvenile myelomonocytic leukemia. J Exp Med. 2021;218(2):e20180853.

7. Emanuel PD, Bates LJ, Castleberry RP, Gualtieri RJ, Zuckerman KS. Selective hypersensitivity to granulocyte-macrophage colony-stimulating factor by juvenile chronic myeloid leukemia hematopoietic progenitors. Blood. 1991;77(5):925-929.

8. Lasho T, Patnaik MM. Juvenile myelomonocytic leukemia - A bona fide RASopathy syndrome. Best Pract Res Clin Haematol. 2020;33(2):101171.

9. Niemeyer CM. JMML genomics and decisions. Hematology. 2018;2018(1):307-312.

10. Flex E, Jaiswal M, Pantaleoni F, et al. Activating mutations in RRAS underlie a phenotype within the RASopathy spectrum and contribute to leukaemogenesis. Hum Mol Genet. 2014;23(16):4315.4327.

11. Niemeyer CM, Flotho C. Juvenile myelomonocytic leukemia: who's the driver at the wheel? Blood. 2019;133(10):1060-1070.

12. Caye A, Strullu M, Guidez F, et al. Juvenile myelomonocytic leukemia displays mutations in components of the RAS pathway and the PRC2 network. Nat Genet. 2015;47(11):1334-1340.

13. Kratz CP, Niemeyer CM, Castleberry RP, et al. The mutational spectrum of PTPN11 in juvenile myelomonocytic leukemia and Noonan syndrome/myeloproliferative disease. Blood. 2005;106(6):2183-2185.

14. Tartaglia M, Niemeyer CM, Fragale A, et al. Somatic mutations in PTPN11 in juvenile myelomonocytic leukemia, myelodysplastic syndromes and acute myeloid leukemia. Nat Genet. 2003;34(2):148-150.

15. Pérez B, Mechinaud F, Galambrun C, et al. Germline mutations of the CBL gene define a new genetic syndrome with predisposition to juvenile myelomonocytic leukaemia. J Med Genet. 2010;47(10):686-691.

16. Niemeyer CM. RAS diseases in children. Haematologica. 2014;99(11):1653-1662.

17. Stieglitz E, Taylor-Weiner AN, Chang TY, et al. The genomic landscape of juvenile myelomonocytic leukemia. Nat Genet. 2015;47(11):1326-1333.

18. Oh ST, Simonds EF, Jones C, et al. Novel mutations in the inhibitory adaptor protein LNK drive JAK-STAT signaling in patients with myeloproliferative neoplasms. Blood. 2010;116(6):988-992.

19. Takaki S, Morita H, Tezuka Y, Takatsu K. Enhanced hematopoiesis by hematopoietic progenitor cells lacking intracellular adaptor protein, Lnk. J Exp Med. 2002;195(2):151-160.

20. Bersenev A, Wu C, Balcerek J, Tong W. Lnk controls mouse hematopoietic stem cell self-renewal and quiescence through direct interactions with JAK2. J Clin Invest. 2008;118(8):2832-2844.

21. Morales CE, Stieglitz E, Kogan SC, Loh ML, Braun BS. Nf1 and Sh2b3 mutations cooperate in vivo in a mouse model of juvenile myelomonocytic leukemia. Blood Adv. 2021;5(18):3587-3591.

22. Bersenev A, Wu C, Balcerek J, et al. Lnk constrains myeloproliferative diseases in mice. J Clin Invest. 2010;120(6):2058-2069.

23. Oh ST. When the Brakes are Lost: LNK Dysfunction in Mice, Men, and Myeloproliferative Neoplasms. Ther Adv Hematol, 2011;2(1):11-19.

24. Maslah N, Cassinat B, Verger E, Kiladjian J-J, Velazquez L. The role of LNK/SH2B3 genetic alterations in myeloproliferative neoplasms and other hematological disorders. Leukemia. 2017;31(8):1661-1670.

25. Coltro G, Lasho TL, Finke CM, et al. Germline SH2B3 pathogenic variant associated with myelodysplastic syndrome/myeloproliferative neoplasm with ring sideroblasts and thrombocytosis. Am J Hematol. 2019;94(9):E231-E234.

26. McMullin MF, Wu C, Percy MJ, Tong W. A nonsynonymous LNK polymorphism associated with idiopathic erythrocytosis. Am J Hematol. 2011;86(11):962-964.

27. Spolverini A, Pieri L, Guglielmelli P, et al. Infrequent occurrence of mutations in the PH domain of LNK in patients with JAK2 mutation-negative “idiopathic” erythrocytosis. Haematologica. 2013;98(9):e101-102.

28. Perez-Garcia A, Ambesi-Impiombato A, Hadler M, et al. Genetic loss of SH2B3 in acute lymphoblastic leukemia. Blood. 2013;122(14):2425-2432.

29. Blombery P, Pazhakh V, Albuquerque AS, et al. Biallelic deleterious germline SH2B3 variants cause a novel syndrome of myeloproliferation and multi-organ autoimmunity. EJHaem. 2023;4(2):463-469.

30. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. Blood. 2016;127(20):2391-2405.

31. Khoury JD, Solary E, Abla O, et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Myeloid and Histiocytic/Dendritic Neoplasms. Leukemia. 2022;36(7):1703-1719.

32. Elghetany MT, Cavé H, De Vito R, Patnaik MM, Solary E, Khoury JD. Juvenile myelomonocytic leukemia; moving forward. Leukemia. 2023;37(3):720-722.

33. Motta M, Pannone L, Pantaleoni F, et al. Enhanced MAPK1 Function Causes a Neurodevelopmental Disorder within the RASopathy Clinical Spectrum. Am J Hum Genet. 2020;107(3):499-513.

34. Rentzsch P, Schubach M, Shendure J, Kircher M. CADD-Splice—improving genome-wide variant effect prediction using deep learning-derived splice scores. Genome Med. 2021;13(1):31.

35. Ioannidis NM, Rothstein JH, Pejaver V, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet. 2016;99(4):877-885.

36. Morris R, Butler L, Perkins A, Kershaw NJ, Babon JJ. The Role of LNK (SH2B3) in the Regulation of JAK-STAT Signalling in Haematopoiesis. Pharmaceuticals (Basel). 2021;15(1):24.

37. Schönung M, Meyer J, Nöllke P, et al. International Consensus Definition of DNA Methylation Subgroups in Juvenile Myelomonocytic Leukemia. Clin Cancer Res. 2021;27(1):158-168.

38. Velazquez L, Cheng AM, Fleming HE, et al. Cytokine signaling and hematopoietic homeostasis are disrupted in Lnk-deficient mice. J Exp Med. 2002;195(12):1599-1611.

39. Nakamura-Ishizu A, Suda T. Multifaceted roles of thrombopoietin in hematopoietic stem cell regulation. Ann N Y Acad Sci. 2020;1466(1):51-58.

40. Buza-Vidas N, Antonchuk J, Qian H, et al. Cytokines regulate postnatal hematopoietic stem cell expansion: opposing roles of thrombopoietin and LNK. Genes Dev. 2006;20(15):2018-2023.

41. Pietras EM, Passegué E. Linking HSCs to their youth. Nat Cell Biol. 2013;15(8):885-887.

42. Helsmoortel HH, Bresolin S, Lammens T, et al. LIN28B overexpression defines a novel fetal-like subgroup of juvenile myelomonocytic leukemia. Blood. 2016;127(9):1163-1172.

43. Bader-Meunier B, Cavé H, Jeremiah N, et al. Are RASopathies new monogenic predisposing conditions to the development of systemic lupus erythematosus? Case report and systematic review of the literature. Semin Arthritis Rheum. 2013;43(2):217-219.

44. Lv K, Jiang J, Donaghy R, et al. CBL family E3 ubiquitin ligases control JAK2 ubiquitination and stability in hematopoietic stem cells and myeloid malignancies. Genes Dev. 2017;31(10):1007-1023.

45. Ren J-G, Xing B, Lv K, et al. RAB27B controls palmitoylation-dependent NRAS trafficking and signaling in myeloid leukemia. J Clin Invest. 2023;133(12):e165510.

46. Kotecha N, Flores NJ, Irish JM, et al. Single-cell profiling identifies aberrant STAT5 activation in myeloid malignancies with specific clinical and biologic correlates. Cancer Cell. 2008;14(4):335-343.

47. Strullu M, Caye A, Lachenaud J, et al. Juvenile myelomonocytic leukaemia and Noonan syndrome. J Med Genet. 2014;51(10):689-697.

Table 1: Germline and somatic *SH2B3* variants identified in JMML patients. (NM_005475.3)

Genomic coordinate (hg38)	Nucleotide change	Protein change	rs number	COSMIC^	description	MAF% (GnomAD	REVEL	CADD PHRED	Predicted consequence / functional behavior
chr12:g.111856451G>T	c.502G>T	p.Glu168*	-	-	truncating	0	-	33	Protein loss
chr12:g.111418830_111418836dup	c.685_691dupGGCCCCG	p.Asp231Glyfs*39	-	COSM9361236	truncating	0	-	NA	Protein loss
chr12:g.111447346dup	c.1038dupG	p.Leu347Alafs*38	-	COSM4170093	truncating	0	-	NA	Protein loss
chr12g.111447468G>C	c.1160G>C	p.Gly387Ala	-	-	Missense (SH2)	0	0,884	28,5	Structural damage
chr12:g.111447509T>C	c.1201T>C	p.Tyr401His	rs1426386395	-	Missense (SH2)	0	0,958	28,5	Structural damage
chr12:g.111447663G>A	c.1244G>A	p.Arg415His	rs918140013	COSM6228211	Missense (SH2)	0	0,542	28,4	Structural damage / impaired target binding and reduced ability to inhibit signalling (Morris, Nat Com 2021)[30]
chr12:g111886085 C>CA	c.1709dupA	p.Asn570LysfsTer82	-	-	Stoploss	0	-	33	Decreased transcript stability/processing and protein stability (this study)

Table 2: Clinical and hematological features of patients with *SH2B3* mutated JMML in comparison with patients with NS-JMML[47] (Strullu *et al*., updated) and with the whole French JMML cohort (excluding patients with NS).

Genetic group	SH2B3 GL biallelic	SH2B3 GL/somatic	Noonan syndrome	Total cohort w/o NS
Number of cases	n=8	n=3	n=33	n=161
M/F sex ratio	1	3	1.8	1.5
Age (months) at onset of hematological anomalies median [range]	0.3	45.6	1.3	20.9
	[0-2]	[26.4-52.8]	[0-14.9]	0.7 – 188.4
Peripheral blood				
Hemoglobin count (g/dL)	11.3	10.50	12.1	9.30
	[6.7-19.9]	[10.1-11.2]	[6.6 – 21.9]	3.5 – 13.8
Platelets counts. median [range]. $\times 10^9$/L	28	23	126	64
	[13-289]	[18-93]	[19-430]	[5 - 428]
WBC counts. median [range]. $\times 10^9$/L	61.5	29.4	23.6	25.0
	[46-81.6]	[11.7-71.2]	[3.2-112]	[4.6 – 168]
Monocytes counts. median [range]. $\times 10^9$/L	7.58	2.6	4.9	4.9
	[2.8-19.6]	[2.06-17]	[0.6-24.6]	[0.9 - 38]
Lymphocyte count . median [range]. $\times 10^9$/L	16.9	5.0	7.4	7.8
	[6.9-40]	[3.16-12.8]	[2-37]	[0.8-42]
Myeloid precursors in peripheral blood. no. of patients (%)	8/8	3/3	21/24	118/137 (86%)
Blasts in peripheral blood. median (%)	4.0	3.0	0.0	1.0
	[0-6]	[1-17]	[0-8.5]	[0-27]
Bone marrow				
Blast in bone marrow. median (%)	4.5	6.0	3.0	4.0
	[2-11]	[2-12]	[0-9]	[0-73]
Dysplastic features. no. of patients (%)	5/8	3/3	16/26	36/49 (73%)
Decreased erythroid lineage. no. of patients (%)	7/8	1/3	9/26	9/49 (18%)
Decreased megakaryocytes. no. of patients (%)	5/8	1/3	7/26	27/49 (55%)
Clinical presentation				
Splenomegaly. no. of patients (%)	8/8	3/3	23/33	149 (94%)
Hepatomegaly. no. of patients (%)	7/8	2/3	20/33	120 (75%)
Adenopathy. no. of patients (%)	4/4	1/3	1/29	87 (55%)
Skin lesions no. of patients (%)	0/8	1/3	4/29	55 (38%)
Respiratory symptoms no. of patients (%)	1/8	1/3	15/29	22 (16%)
Bleeding no. of patients (%)	2/8	1/3	1/29	32 (24%)
Secondary genetic alteration. no. of patients (%)	0/8	3/3	0/31	74/152 (49%)
Treatment and follow up				
Chemotherapy. no. of patients (%)	4/8	3/3	6/28	-
Bone marrow transplantation. no. of patients (%)	1/8	3/3	2/31	118/161 (73%)
Clinical outcome (alive) no. of patients (%)	8/8	2/3	20/31	111/161 (68%)

GL: germline ; NS: Noonan syndrome ; no: number ; WBC: white blood cell

Figure legends

Figure 1: Patients cytomorphological data and pedigrees (A-B) May-Grümwald-Giemsa stained smears showing cytomorphological features at diagnosis of JMML. **(A)** Patient 79.1 (A1-4), 79.2 (A5-8), **(B)** Patient OPBG (B1-2). Peripheral blood smears (A1-3; A5-8; B1) show monocytes (*Mo), myeloid precursors (*My), dysmorphic basophiles (*Ba) and undifferentiated myeloid blasts (*Bl). Bone marrow smear (A4; B2) show hypercellularity with myeloid blasts and discrete signs of dysgranulopoiesis. The erythrocyte lineage is decreased and the megakaryocytic lineage is absent. **(C)** Pedigrees of patients with germline *SH2B3* variants. All children but none of the heterozygous parents had JMML. Germline variants (dark blue) were either bi-allelic (full circle/square; families 1 to 4) or mono-allelic (half cicle/square; families 5-6) in the affected child. Somatic alteration is indicated in light blue. aUPD: acquired uniparental isodisomy. JMML: juvenile myelomonocytic leukemia. ND: not done

Figure 2: *SH2B3* variants and consequences on the LNK protein. **(A)** Lollipop plot showing the distribution of *SH2B3/LNK* variants over a schematic representation of the LNK protein. The PH domain enables interaction with cell membrane phospholipids and the SH2 domain binds to the phosphorylated tyrosines of target proteins. **(B-C)** Disease-causing *SH2B3* stoploss frameshift mutation affect RNA and protein stability. **(B)** Protein stability was assessed in 293T cells transiently transfected with wild-type Xpress-tagged SH2B3 and mutant carrying the c.1709dupA (p.Asp570Lys*fs*82) mutation. After transfection (48h), cells were treated with cycloheximide (CHX, 20 µg/ml) for the indicated time or left untreated. Protein levels was assessed by immunoblotting, using an anti-Xpress monoclonal antibody. GAPDH levels are shown to document equal loading of total proteins from cell lysates. Western blot from a representative experiment of three performed is shown. **(C)** qPCR analyses performed on RNA extracted from patient and control fibroblasts show a significant decrease of *SH2B3* expression level indicating RNA decay of the allele carrying the *SH2B3*

variant. Data are graphed indicating fold change of *SH2B3* in patients' fibroblasts over control cells (WT), set as 1. *GAPDH* was used as endogenous control. Histograms show mean values ±SD of three independent experiments, each performed in triplicate. The analysis of the expression was performed calculating the fold change using the 2−ΔΔCt formula and the results were statistically analysed by PRISM7, using two-tailed unpaired t-test with Bonferroni correction. *** p<0.001. **(D)** 3D structure of the SH2 domain of the LNK protein with the JAK2 phosphopeptide pY813 modeled from crystallographic data of the Morris *et al.* study[36] (PDB:7R8W). The position of variant amino acids was determined by structural homology between mouse and human SH2 domains.

aa: amino acids. CHX: Cycloheximide. WT: wild type

Figure 3: Clonal architecture of *SH2B3* mutated JMML. Patients with a biallelic germline mutation of *SH2B3* (red clones) (#79, #201, #216, #244) all have a monomorphic profile with no additional genetic alterations. Only patient 79.1 received bone marrow transplant (BMT) (upper panel). Patients with a germline monoallelic *SH2B3* mutation (#53, #48) and/or somatic *SH2B3* mutation (orange clones) acquired several additional somatic mutations. Patient #53 had a somatic *PTPN11* variant and patient #48 had somatic variants of *PTPN11* and *NF1*, whose order of appearance could not be inferred from the allelic frequencies. Both patients underwent BMT (medium panel). In patient #209, allelic frequency of the somatic *SH2B3* variant was consistent with its presence in the full leukemia clone. Multiple other somatic alterations were acquired, including a driver somatic mutation in *PTPN11*. Here again, all potentially driver mutations were present at variant allele frequency (VAF) consistent with early co-occurrence during the course of the disease and order or appearance could not be determined (lower panel). aUPD: acquired uniparental isodisomy. BMT: bone marrow transplant. JMML: juvenile myelomonocytic leukemia.

Figure 4: DNA methylation pattern and gene expression profiling suggest persistence of fetal cues in JMML with bi-allelic germline *SH2B3* variants (A) DNA methylation

analysis of patients with *SH2B3* variants (n=8) and reference JMML of known genetic groups (i.e. PTPN11 somatic, NRAS somatic, KRAS somatic, NF1 or CBL) (n=54). Unsupervised hierarchical clustering based on the 1000 most variant 100 bp tiles identified three main DNA methylation subgroups (low, intermediate, high). JMML with bi-allelic *SH2B3* germline mutations defines a subgroup of hypomethylated JMML (cluster 1) whereas JMML with either mono-allelic or somatic *SH2B3* mutations cluster in the intermediate (cluster 4) or high (cluster 5) methylation groups. **(B-D)** Gene expression profile analyses of mononucleated cells of JMML with bi-allelic *SH2B3* germline alteration (n=4) in comparison with other types of JMML (n=17) and healthy bone marrow samples. Principal component analysis (PCA) (B) and vulcanoplots showing differential gene expression between JMML with bi-allelic *SH2B3* germline alteration and healthy bone marrow (C) or other types of JMML (D). A qval threshold of ≤ 0.05 and a minimum fold change of 1.5 were used to define differentially expressed genes. BM: bone marrow. FC: fold change. JMML: juvenile myelomonocytic leukemia. PC: principal component. Meth: methylation

Figure 1

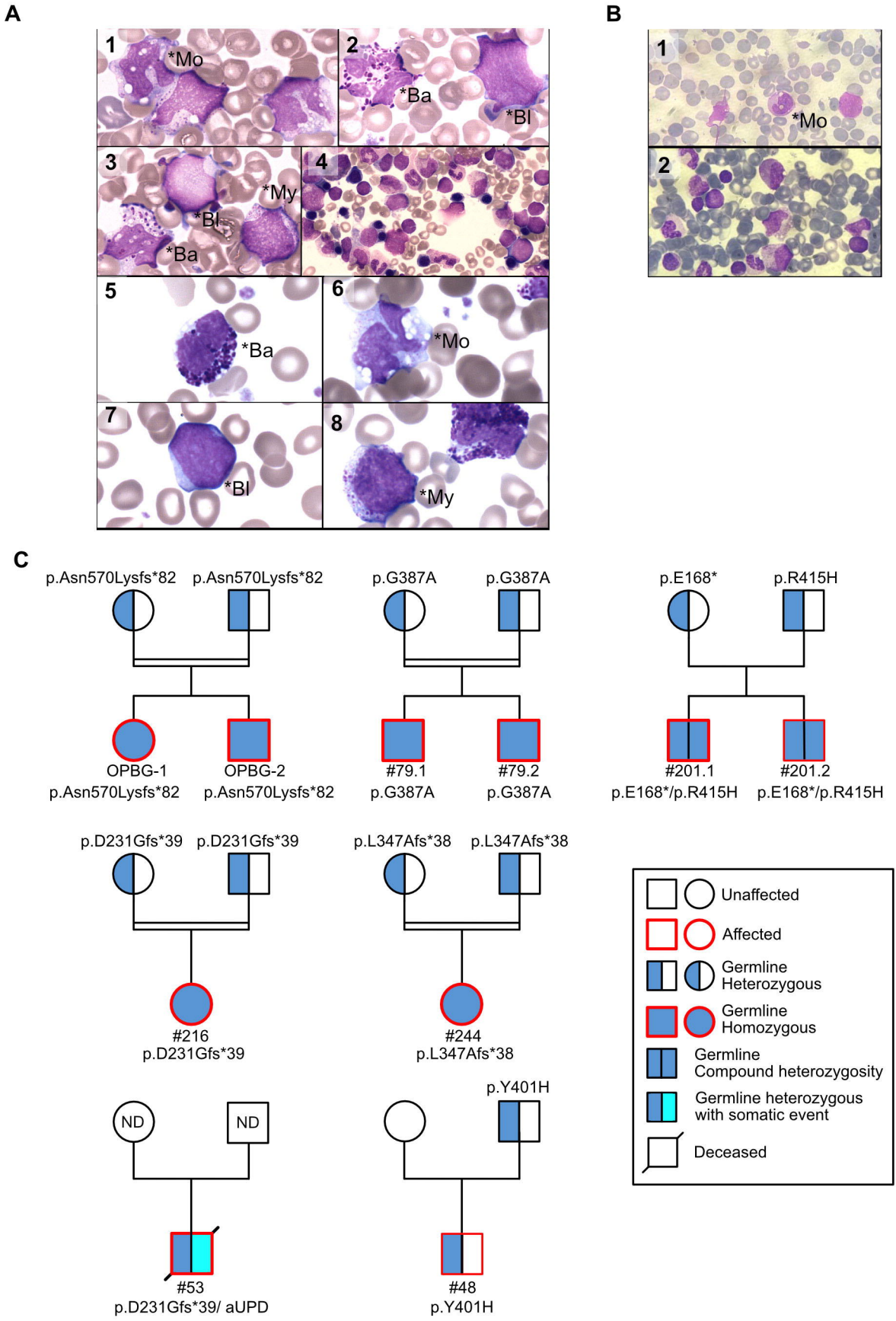

A
1
*Mo
2
*Ba
*Bl
3
*My
*Bl
*Ba
4
5
*Ba
6
*Mo
7
*Bl
8
*My
B
1
*Mo
2
C
p.Asn570Lysfs*82
p.Asn570Lysfs*82
OPBG-1
p.Asn570Lysfs*82
OPBG-2
p.Asn570Lysfs*82
p.G387A
p.G387A
#79.1
p.G387A
#79.2
p.G387A
p.E168*
p.R415H
#201.1
p.E168*/p.R415H
#201.2
p.E168*/p.R415H
p.D231Gfs*39
p.D231Gfs*39
#216
p.D231Gfs*39
p.L347Afs*38
p.L347Afs*38
#244
p.L347Afs*38
ND
ND
#53
p.D231Gfs*39/ aUPD
p.Y401H
#48
p.Y401H
Unaffected
Affected
Germline Heterozygous
Germline Homozygous
Germline Compound heterozygosity
Germline heterozygous with somatic event
Deceased

Figure 2

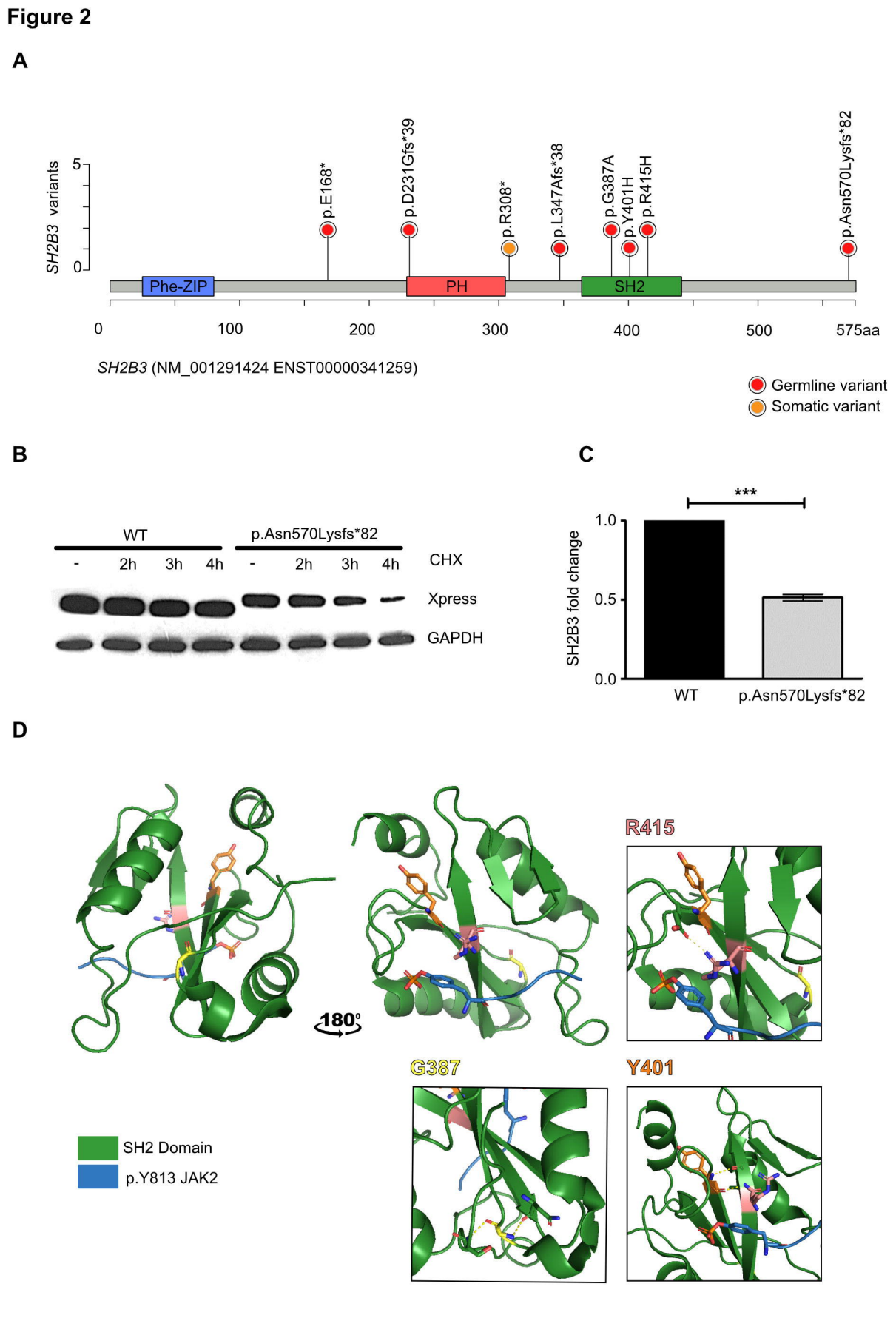

A
SH2B3 variants
5
0
p.E168*
p.D231Gfs*39
p.R308*
p.L347Afs*38
p.G387A
p.Y401H
p.R415H
p.Asn570Lysfs*82
Phe-ZIP
PH
SH2
0
100
200
300
400
500
575aa
SH2B3 (NM_001291424 ENST00000341259)
Germline variant
Somatic variant
B
WT
p.Asn570Lysfs*82
- 2h 3h 4h - 2h 3h 4h
CHX
Xpress
GAPDH
C
***
SH2B3 fold change
1.0
0.5
0.0
WT
p.Asn570Lysfs*82
D
180°
R415
G387
Y401
SH2 Domain
p.Y813 JAK2

Figure 3

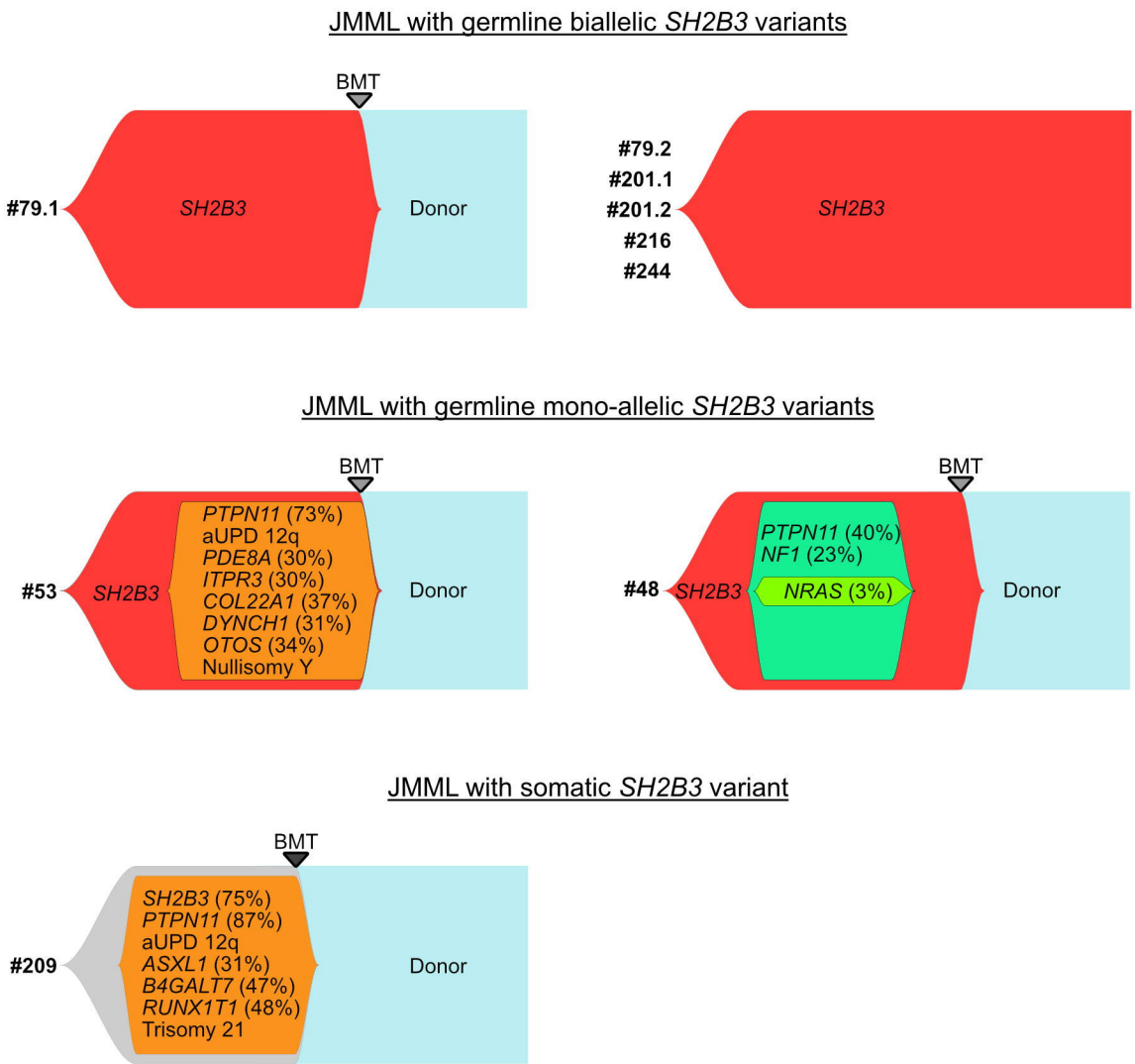
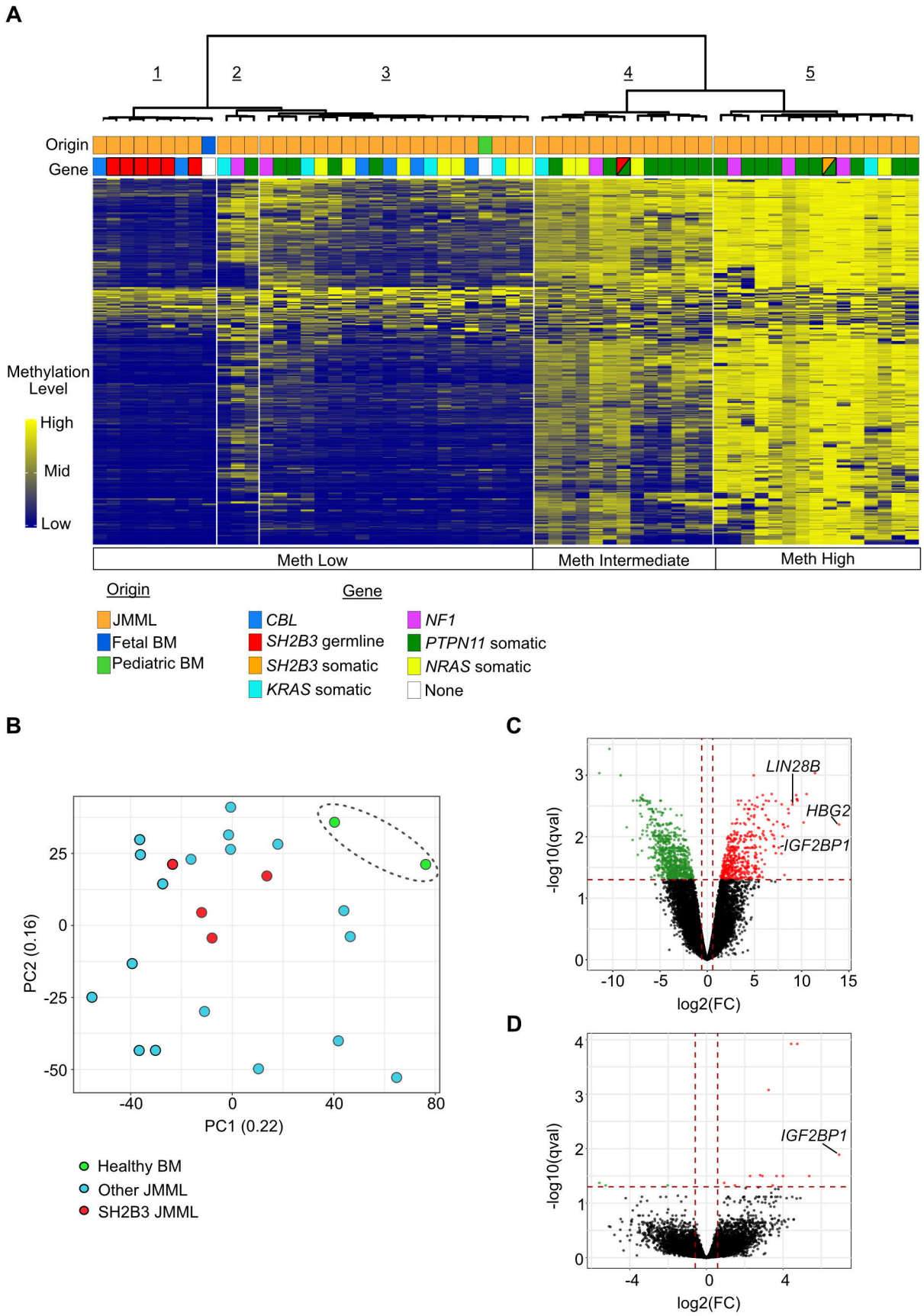

JMML with germline biallelic SH2B3 variants
BMT
#79.1
SH2B3
Donor
#79.2
#201.1
#201.2
#216
#244
SH2B3
JMML with germline mono-allelic SH2B3 variants
BMT
#53
SH2B3
PTPN11 (73%)
aUPD 12q
PDE8A (30%)
ITPR3 (30%)
COL22A1 (37%)
DYNCH1 (31%)
OTOS (34%)
Nullisomy Y
Donor
BMT
#48
SH2B3
PTPN11 (40%)
NF1 (23%)
NRAS (3%)
Donor
JMML with somatic SH2B3 variant
BMT
#209
SH2B3 (75%)
PTPN11 (87%)
aUPD 12q
ASXL1 (31%)
B4GALT7 (47%)
RUNX1T1 (48%)
Trisomy 21
Donor

Figure 4


A
1
2
3
4
5
Origin
Gene
Methylation Level
High
Mid
Low
Meth Low
Meth Intermediate
Meth High
Origin
JMML
Fetal BM
Pediatric BM
Gene
CBL
SH2B3 germline
SH2B3 somatic
KRAS somatic
NF1
PTPN11 somatic
NRAS somatic
None
B
PC2 (0.16)
PC1 (0.22)
Healthy BM
Other JMML
SH2B3 JMML
C
LIN28B
HBG2
IGF2BP1
-log10(qval)
log2(FC)
D
IGF2BP1
-log10(qval)
log2(FC)

SUPPLEMENTARY MATERIAL

Germline bi-allelic *SH2B3* alteration predisposes to a neonatal juvenile myelomonocytic leukemia-like disorder

Chloé Arfeuille[1,2], Yoann Vial[1,2], Margaux Cadenet[1,2], Aurélie Caye-Eude[1,2], Odile Fenneteau[3], Quentin Neven[4], Adeline A Bonnard[1,2], Simone Pizzi[5], Giovanna Carpentieri[5], Yline Capri[6], Katia Girardi[7], Lucia Pedace[7], Marina Macchiaiolo[8], Kamel Boudhar[9], Monia ben Khaled[10], Wadih Abou Chahla[11], Anne Lutun[12], Mony Fahd[4], Séverine Drunat[1], Elisabetta Flex[13], Jean-Hugues Dalle[4], Marion Strullu[2,4], Franco Locatelli[11,14], Marco Tartaglia[5], Hélène Cavé[1,2]

Supplementary Material and Methods

Genomic DNA and RNA extraction

Genomic DNA was extracted using a QIAamp DNA Mini Kit (Qiagen Gmbh). DNA concentration was measured using a Varioskan LUX (Thermo Scientific). RNA was extracted using the RNeasy Mini or Micro Kit (Qiagen); cDNA was obtained by reverse transcription of 1 µg RNA with random hexamers.

Targeted sequencing

A panel of genes implicated in JMML (*AEBP2, ASXL1, BCORL1, BRAF, CBL, CDC42 .2, CDCA5, CDKN2A, CDKN2B, CDYL, CREBBP, DNMT3A, DOCK2, EED, EP300, ESCO1, ESCO2, EZH2, GATA2, HDAC8, HRAS, IKZF1, JAK3, JARID2, KRAS, LZTR1, MAP2K1, MAP2K2, MAU2, NF1, NIPBL, NPM1, NRAS, PDE8A, PDS5A, PDS5B, PLXNB2, PPP1CB, PTPN11, RAC2, RAD21, RAF1, RASA2, RBBP4, RBBP7, RIT1, RRAS, RRAS2, SETBP1, SF3B1, SH2B3, SH3BP1, SHOC2, SMC1A, SMC3, SOS1, SOS2, SPRED1, SRSF2, STAG1, STAG2, SUZ12, TP53, U2AF1, WAPAL, WT1, ZRSR2)* was analyzed by parallel sequencing (supplementary information), using capture-based target enrichment (Custom SureSelect XT-HS2, Agilent) and sequencing on a NextSeq 500 (High Output Kit v2, 2*150bp) (Illumina). Bioinformatical alignment was performed using the Pipeline Local Run Manager v.2.4.0 (Illumina). Variant calling was performed using VarScan v.2.3.5. Variant classification was conducted using Alissa Interpret (Agilent technologies). Average sequencing depth was 1500X. Variants were filtered based on >10 reads and an allelic frequency (VAF) >5%.

Whole genome sequencing analysis:

Laboratory: PCR-free libraries were prepared with NEBNext Ultra II DNA Library Prep Kit (Illumina) according to supplier recommendations and sequenced on an Illumina NovaSeq6000 as paired reads of 150 nucleotides.

Base calling was performed using Illumina Real Time Analysis with default parameters. Reads were mapped to the human genome build (hg38) using the Burrows-Wheeler Aligner (BWA) tool. Variant and copy number alteration (CNA) calling is described in supplementary data. Prediction of the functional and clinical consequences of variants was performed by CADD (GRCh38-v1.6)[34] and REVEL[35]. Previous occurrence in the ClinVar (https://www-ncbi-nlm-nih-gov/clinvar) and COSMIC (v98) (https://cancer.sanger.ac.uk/cosmic) database were checked. Minimal allele frequencies (MAF) were obtained from GnomAD version 2.1.1 (non cancer).

Variant calling: For constitutional DNA, identification of Single Nucleotide Variations (SNV) and small insertions/deletions (up to 20bp), was performed via the Broad Institute's GATK Haplotype Caller GVCF tool. Detection of germline SNV was performed on paired analysis enabling mutation enrichment determination. For somatic DNA, Broad Institute's MuTect tool (2.0, max_alt_alleles_in_normal_count=2; max_alt_allele_in_normal_fraction=0.04) was used for somatic DNA. An IntegraGen (Evry, France) post-processing, developed to filter out candidate somatic mutations which are more consistent with artifacts or germline mutations, is applied. Only somatic mutations considered as PASS, t_lod_fstar, alt_allele_in_normal, germline_risk and panel_of_normals were retained. A somatic score ranging from 1 to 30 was calculated for each mutation, when a normal sample was available, with a score of 30 translating to the highest confidence index. This score takes into account both counts and the frequencies of mutated allele in both samples in order to minimize false positive variations. Also, the mutations with a QSS score below 20 and a VAF of tumor <0.02 are removed.

CNA calling: GATK (V4.1.4.1) was used to investigate genomic CNA by comparing tumor DNA exome data to a reference sample pool and by comparing B-allele frequency (BAF) with the distribution of variant allele frequencies of SNPs reported in GnomAD. All changes were annotated with AnnotSV to provide a comprehensive summary of structural variation in the human genome.

Bioinformatics: Quality of reads was assessed for each sample using FastQC (v.0.11.4; http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). RNA-SeQC provided key measures of data quality. These metrics are shown within Reporting and include yield, alignment and duplication rates, rRNA content, regions of alignment (exon, intron and intragenic). Alignment is performed by STAR (https://github.com/alexdobin/STAR). The duplicate reads (e.g., paired-end reads in which the insert DNA molecules have identical start and end locations in the Human genome) are removed (sambamba tools). To detect fusion-genes candidates in RNA-seq data, FusionCatcher (Start with fastq files) and Star-Fusion (Start with alignment files) are used in order to achieve higher detection efficiency. These were run using defaults configurations.

Constructs

The c.1709dupA (p.Asp570Lysfs82) was introduced in an Xpress-tagged *SH2B3* cDNA (NM_005475.3) cloned in pcDNA6/HisC (Invitrogen) by site-directed mutagenesis (QuikChange XL kit, Agilent Technologies). Constructs were verified by direct sequencing.

Evaluation of protein stability

293T cell lines were cultured in Dulbecco's modified Eagle's medium supplemented with 10% heat-inactivated fetal bovine serum (FBS) (EuroClone) and 1% penicillin-streptomycin, at 37 °C, 5% CO_2. Cells were seeded in 6-well plates the day before transfection and then transfected at 70% confluency with Fugene 6 (Promega) with constructs of interest. 48 hours after transfection, cells were treated with cycloheximide (20 mg/mL) for 2, 3 and 4 hours or left untreated before lysis. Immunoblotting was performed so that Xpress-tagged *SH2B3* levels could be assessed using a monoclonal Xpress antibody (Invitrogen). Probing membranes with an anti-GAPDH antibody (Santa Cruz) allowed normalization of protein content.

Quantitative PCR (q-PCR)

Five hundred nanograms of total RNA were reverse transcribed using SuperScript-III (Invitrogen), following the manufacturer's instructions. Detection of *SH2B3* and *GAPDH* was performed using a qPCR system with SensiFAST Probe Lo-ROX Kit (Bioline), according to the manufacturer's instructions. Specific TaqMan Gene Expression Assays probes (Applied Biosystems), were used to analyze *SH2B3* mRNA expression. Samples were normalized according to *GAPDH* transcript levels. Each sample was run in triplicate, in at least three independent experiments. qPCR assays were performed in a 7500 Fast qPCR machine (Applied Biosystems). The fold change was calculated using the $2^{-\Delta\Delta Ct}$ method.

Gene expression analyses by mRNAseq

Quantification of gene expression: STAR was used to obtain the number of reads associated to each gene in the Gencode v.31 annotation (restricted to protein-coding genes, antisense and lincRNAs). Raw counts for each sample were imported into R statistical software. Extracted count matrix was normalized for library size and coding length of genes to compute FPKM expression levels.

Unsupervised analysis: The Bioconductor edgeR package was used to import raw counts into R statistical software, and compute normalized log2 CPM (counts per millions of mapped reads) using the TMM (weighted trimmed mean of M-values) as normalization procedure. The normalized expression matrix from the 1000 most variant genes (based on standard deviation) was used to classify the samples according to their gene expression patterns using principal component analysis (PCA). PCA was performed by FactoMineR::PCA function with “ncp = 10, scale.unit = FALSE” parameters.

Differential expression analysis: Differential expression analysis was performed using the Bioconductor limma package and the voom transformation. To improve the statistical power of the analysis, only genes expressed in at least one sample (FPKM ≥ 0.1) were considered. A qval threshold of ≤ 0.05 and a minimum fold change of 1.5 were used to define differentially expressed genes.

DNA methylation (capture EM-seq)

Laboratory: For the panel capture part, a total of 1500 ng from 8 libraries, 187.5 ng each, were pooled and incubated 65°C for 16h with blockers, methylation enhancer, and Human methylome panel probes in a final volume of 40 µL. Washing steps were performed following the recommendations provided by Twist. Subsequently, the pooled sample was amplified with 8 cycles of PCR, quantified using qPCR and subjected to paired-end sequencing with 100-bp reads on an Illumina Novaseq 6000 platform.

Sequence alignment and quantification of DNA methylation: Quality of reads was assessed for each sample using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). BS-Seeker2[1] was used to map capture EM-seq data to the human hg38 genome and retrieve the number of methylated and unmethylated cytosines at each covered CpG site. Methylation rates were then integrated across CpG island (CGI)-based and gene-based features. CGI-based features were defined as follows: CpG islands (from UCSC database hg38), shores (2 kb on each side of the island) and shelves (2 kb on each side of the shores). DNA methylation outside CpG islands was analyzed by grouping CpG sites not located in CGI-based features every 100kb window. Gene-based features were defined based on Ensembl *homo sapiens* hg38 genes. The methylation rate was calculated for each gene across the promoter region (TSS +/- 500bp) and the gene body. Finally, methylation was analyzed by grouping CpGs in 100 bp tiles.

Unsupervised classification: Unsupervised classifications based on the 1000 most variant 100 bp tiles (based on standard deviation) were generated using principal component analysis (PCA) and hierarchical clustering (cosine distance, Ward method) in R software. We used consensus clustering (Bioconductor ConsensusClusterPlus package[2]) to examine the stability of the clusters. We established consensus partitions of the data set in K clusters (for K = 2, 3, ..., 10), on the basis of 1,000 resampling iterations (80% of genes, 80% of sample) of hierarchical clustering, with Pearson's dissimilarity as the distance metric and Ward's method for linkage analysis. We used the cumulative distribution functions (CDFs) of the consensus matrices to determine the optimal number of clusters (K = 5), considering both the shape of the functions and the area under the CDF curves, as previously described.

Statistical analysis and graphical representation

Protein *in silico* modelisation was conducted with Uniprot database (https://www.uniprot.org/uniprotkb/Q9UQQ2/entry) and AlphaFold (https://alphafold.ebi.ac.uk/entry/Q9UQQ2)

Statistical analyses were performed using GraphPad Prism (v.8.1.2). Medians were compared using the Mann-Whitney test.

Graphical representations were performed using ggplot2 (v.3.3.5) with RStudio (2022.02.0+443). Lollipop plot was generated using cBioportal Mutationmapper tool (https://www.cbioportal.org/mutation_mapper). Fishplots were created with the chrisamiller/fishplot (v.0.5) package using RStudio (2022.02.0+443). Illustrations were created using Affinity Designer (V1.10.6.1665).

References

1. Guo W, Fiziev P, Yan W, et al. BS-Seeker2: a versatile aligning pipeline for bisulfite sequencing data. BMC Genomics 2013;14(1):774.
2. Monti S. Consensus Clustering: A Resampling-Based Method for Class Discovery and Visualization of Gene Expression Microarray Data. 28.

Supplementary Table 1

	Patient	79.1	79.2	201.1	201.2
	Gender	M	M	F	M
	Age at JMML onset	15 days	36 days	22 days	11 days
	geographic origin	Morocco	Morocco	Algeria	Algeria
	consanguinity	yes	yes	no	no
At initial presentation	**Hematological counts in peripheral blood**				
	Hemoglobin count (g/dL)	12,8	10,8	6,7	19,9
	Platelet count (×10^9/l)	21	239	37	21
	White blood cells count (×10^9/l)	57,5	65,5	46	114
	Monocytes count (×10^9/l)	14,38	7,86	2,8	11,6
	Lymphocyte count (×10^9/l)	16,7	25,5	21	40
	Myeloid precursors (%)	15	10	9	23
	Blasts in peripheral blood (%)	6	4	6	0
	Hematological counts in bone marrow				
	Blast in bone marrow (%)	11	2	4	4
	Cellularity	Rich	Rich	Moderate	Moderate
	Dysplastic features	moderate	no	no	moderate
	Erythroid lineage	decreased (7%)	Normal (22%)	decreased (4%)	decreased (13%)
	Megakaryocytic lineage	absent	fairly abundant	absent	absent
	Clinical presentation				
	Hepatomegaly	yes	no	yes	yes
	Splenomegaly	yes	yes	yes	yes
	Adenopathy	no	no	yes	no
	Pallor	yes	no	yes	no
	Fever	no	no	no	no
	Skin lesions	no	no	no	no
	Respiratory symptoms	no	no	no	no
	Cardiac symptoms	no	no	no	no
	Bleeding	no	no	no	no
	Other clinical signs	RCIU, microcephaly, cardiomegaly, tracheomalacia, short stature	no	no	no
	Familial history	Thyroiditis (mother and two maternal aunts). insulin-dependent diabetes at 25 yrs (maternal uncle). learning and language difficulties with deafness (paternal aunt)		no	
	***SH2B3* variants**				
	Nucleotidic change	c.1160G>C	c.1160G>C	c.502G>T; c.1244 G>A	c.502G>T; c.1244 G>A
	Protein change	p.Gly387Ala	p.Gly387Ala	p.Glu168*; p.Arg415His	p.Glu168*; p.Arg415His
	Allelic status	homozygous	homozygous	compound heterozygous	compound heterozygous
	Status (germline/somatic)	germline	germline	germline	germline
	inheritance	AR	AR	AR	AR
	Other JMML features				
	Karyotype	46, XY	46, XY	46, XX	ND
	Secondary genetic alteration (*SH2B3*)	no	no	no	no
	Secondary genetic alteration (*PTPN11*)	no	no	no	no
	Fetal hemoglobin elevated for age	no	no	no	no
	Spontaneous growth of myeloid progenitors	yes	yes	ND	ND
	Treatment and follow up				
	Chemotherapy	6-MP	no	6-MP	no
	Bone marrow transplantation	yes	no	no	no
	Clinical outcome	Alive. Statural delay Very large splenomegaly Abdominal distension	Alive. Slight statural delay Autoimmune hypothyroidism	Alive. Persistent splenomegaly. Neutropenia	Alive. liver transplant after cirrhosis. Persistent neutropenia
	Age (years) as of june 2023	13	7	9	5

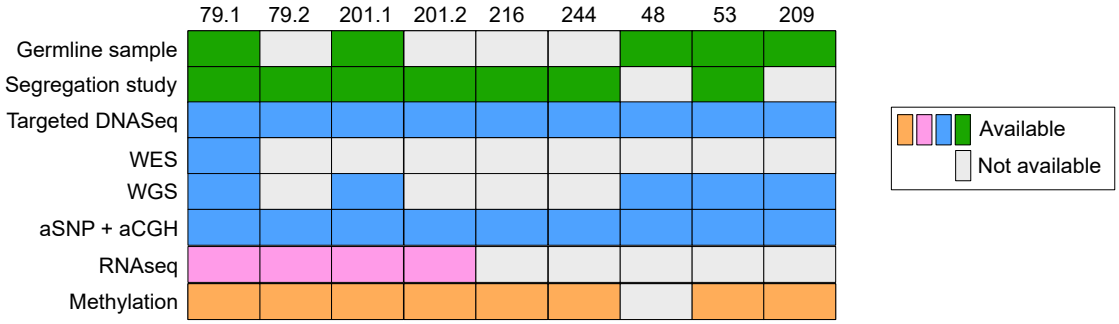
Supplementary Table 1 – continued

	Patient	216	244	OPBG_1	OPBG_2
	Gender	F	F	F	M
	Age at JMML onset	1 day	22 day	30 days	30 days
	geographic origin	Tunisia	North Africa	Romany	Romany
	consanguinity	yes	yes	yes	yes
At initial presentation	**Hematological counts in peripheral blood**				
	Hemoglobin count (g/dL)	17,8	9,4	11,8	8
	Platelet count (×10^9/l)	13	15	35	41
	White blood cells count (×10^9/l)	81	81,57	56,1	55,6
	Monocytes count (×10^9/l)	7,3	19,6	6,1	4,2
	Lymphocyte count (×10^9/l)	12,9	17,1	6,9	11,01
	Myeloid precursors (%)	9,5	11	25	3
	Blasts in peripheral blood (%)	1,5	4	5	4
	Hematological counts in bone marrow				
	Blast in bone marrow (%)	11	2	5	6,5
	Cellularity	Normal	Moderate	Moderate	Rich
	Dysplastic features	no	yes	yes	yes
	Erythroid lineage	decreased	decreased	decreased	decreased
	Megakaryocytic lineage	Present	absent	absent	dysplastic
	Clinical presentation				
	Hepatomegaly	yes	yes	yes	yes
	Splenomegaly	yes	yes	yes	yes
	Adenopathy	no	yes	yes	yes
	Pallor	yes	yes	yes	yes
	Fever	no	yes	no	no
	Skin lesions	no	no	no	no
	Respiratory symptoms	no	yes	no	no
	Cardiac symptoms	no	no	no	no
	Bleeding	Generalized petechial purpura	Meningeal haemorrhage	no	no
	Other clinical signs	no	RCIU; Facial dysmorphic features; vermis haematoma and haemoperitoneum	Facial dysmorphic features; coronary heart disease; Botallo's Duct	Facial dysmorphic features; Coronary heart disease; complex limb reduction defect
	Familial history	no	no	no	
	***SH2B3* variants**				
	Nucleotidic change	c.685_691dupGGCCCCG	c.1038dupG	c.1709dupA	c.1709dupA
	Protein change	p.Asp231Glyfs*39	p.L347Afs*38	p.Asp570Lysfs82	p.Asp570Lysfs82
	Allelic status	homozygous	homozygous	hom	hom
	Status (germline/somatic)	germline	germline	germline	germline
	inheritance	AR	AR	AR	AR
	Other JMML features				
	Karyotype	46, XX	46, XX	46 XX	46XY
	Secondary genetic alteration (*SH2B3*)	no	no	na	na
	Secondary genetic alteration (*PTPN11*)	no	no	na	na
	Fetal hemoglobin elevated for age	elevated		no	no
	Spontaneous growth of myeloid progenitors	ND	ND	Yes	Yes
	Treatment and follow up				
	Chemotherapy	no	6-MP	6-MP and cis-retinoic acid	no
	Bone marrow transplantation	no	no	no	no
	Clinical outcome	Alive. Spontaneous resolution	Alive. Persistent hepato-splenomegaly and hepatomegaly	Alive. Spontaneous resolution	Alive. Spontaneous resolution
	Age (years) as of june 2023	*5*	*2*	*12*	*8*

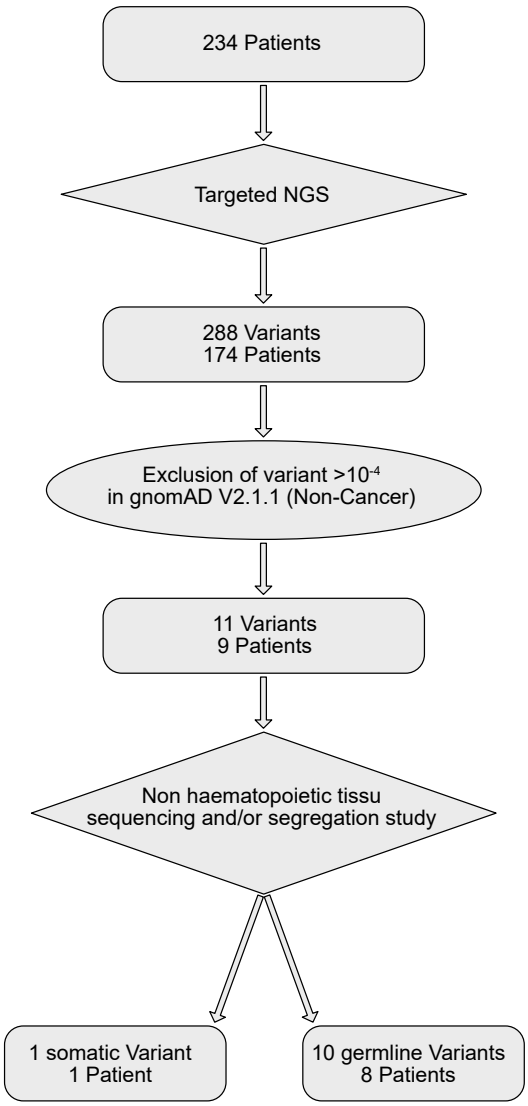
Supplementary Table 1 – continued

	Patient	48	53	209
	Gender	M	M	M
	Age at JMML onset	3,8 years	2,2 years	4,4 years
	geographic origin	unknown	unknown	unknown
	consanguinity	no	no	yes
At initial presentation	**Hematological counts in peripheral blood**			
	Hemoglobin count (g/dL)	10,1	11,2	10,5
	Platelet count (×10^9/l)	18	23	93
	White blood cells count (×10^9/l)	11,7	71,2	29,4
	Monocytes count (×10^9/l)	2,6	17	2,06
	Lymphocyte count (×10^9/l)	3,16	12,8	5
	Myeloid precursors (%)	8	14	4
	Blasts in peripheral blood (%)	17	3	1
	Hematological counts in bone marrow			
	Blast in bone marrow (%)	12	6	2
	Cellularity	Rich	Rich	Very rich
	Dysplastic features	moderate	no	no
	Erythroid lineage	decreased	normal	normal
	Megakaryocytic lineage	absent	present	decreased
	Clinical presentation			
	Hepatomegaly	yes	yes	no
	Splenomegaly	yes	yes	yes
	Adenopathy	no	yes	no
	Pallor	yes	no	no
	Fever	yes	no	no
	Skin lesions	no	xanthogranuloma	no
	Respiratory symptoms	no	no	pulmonary emphysema
	Cardiac symptoms	no	no	no
	Bleeding	yes		
	Other clinical signs	no	no	Intermittent left lameness Thyroid nodules
	Familial history		multiple sclerosis (father); Splenectomy at 23 yrs (unknown reason) (maternal grandfather)	Asthma (brother) Childhood leukemia (maternal aunt) Type 2 diabetes (paternal grandmother)
	***SH2B3* variants**			
	Nucleotidic change	c.1201T>C	c.685_691dupGGCCCCG	c.922C>T
	Protein change	p.Tyr401His	p.Asp231Glyfs*39	p.Arg308*
	Status (homozygous/heterozygous)	heterozygous	heterozygous	n/a
	Status (germline/somatic)	germline	germline	somatic
	inheritance	AD	AD	n/a
	Other JMML features			
	Karyotype	46XY	46,XY[15]/45,X,-Y[12]	46, XY with secondary +21
	Secondary genetic alteration (*SH2B3*)	no	aUPD	aUPD
	Secondary genetic alteration (*PTPN11*)	*PTPN11:* p.Glu76Gly	*PTPN11*: p.Gly503Ala	*PTPN11*: p.Asp61Tyr
	Fetal hemoglobin elevated for age	UK	yes	yes
	Spontaneous growth of myeloid progenitors	yes	yes	yes
	Treatment and follow up			
	Chemotherapy	Chemotherapy	6-MP	6-MP
	Bone marrow transplantation	yes	yes	yes
	Clinical outcome	Alive	Death following veno-occlusive disease	Alive. Post-transplant complications
	Age (years) as of june 2023	*19*	*17*	*9*

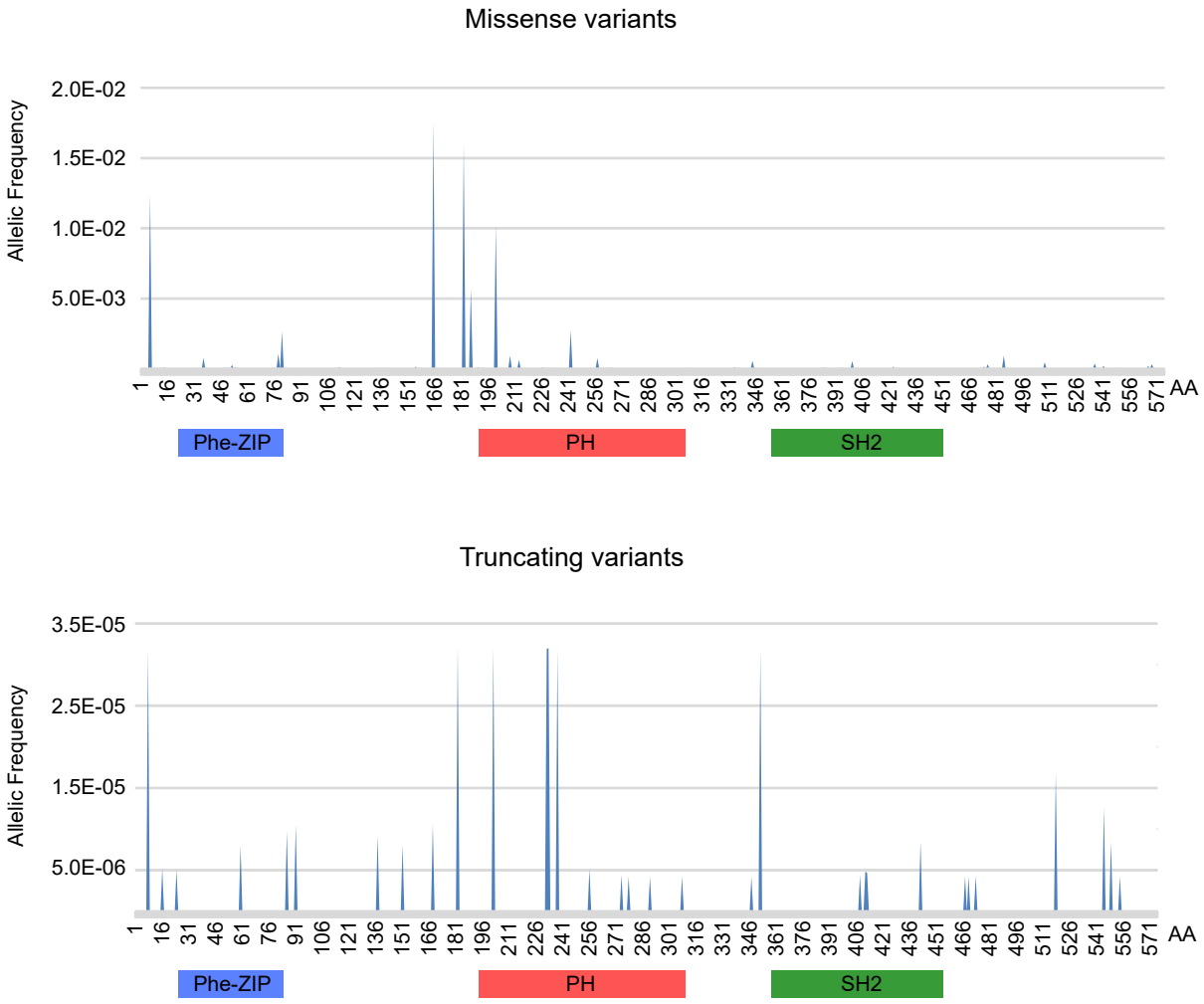
Supplementary Figure 1. Graphical representation of the analyses performed in each patient with *SH2B3* variants



Supplementary Figure 2. Flow chart of the *SH2B3* study in the French cohort of patients with JMML



Supplemental Figure 3. Frequency of *SH2B3* missense (upper panel) and truncating (bottom panel) variants in the GnomAD version 2.1.1 (non cancer) database. Notably, missense variations are mostly found in the PH domain whereas they are rare in the SH2 domain.



Supplementary Figure 4. Evolution of platelet counts with age in patients with *SH2B3* bi-allelic germline variants.