The Genomic Landscape of Juvenile Myelomonocytic Leukemia

HHS Public Access

Author manuscript

Nat Genet. Author manuscript; available in PMC 2016 May 01.

Published in final edited form as:

Nat Genet. 2015 November ; 47(11): 1326–1333. doi:10.1038/ng.3400.

The Genomic Landscape of Juvenile Myelomonocytic Leukemia

Elliot Stieglitz[#1], **Amaro N. Taylor-Weiner**[#2], **Tiffany Y. Chang**[1], **Laura C. Gelston**[1], **Yong-Dong Wang**[3], **Tali Mazor**[4], **Emilio Esquivel**[1], **Ariel Yu**[1], **Sara Seepo**[2], **Scott Olsen**[5], **Mara Rosenberg**[2], **Sophie L. Archambeault**[1], **Ghada Abusin**[6], **Kyle Beckman**[1], **Patrick A. Brown**[7], **Michael Briones**[8], **Benjamin Carcamo**[9], **Todd Cooper**[10], **Gary V. Dahl**[11], **Peter D. Emanuel**[12], **Mark N. Fluchel**[13], **Rakesh K. Goyal**[14], **Robert J. Hayashi**[15], **Johann Hitzler**[16], **Christopher Hugge**[17], **Y. Lucy Liu**[12], **Yoav H. Messinger**[18], **Donald H. Mahoney Jr**[19], **Philip Monteleone**[20], **Eneida R. Nemecek**[21], **Philip A. Roehrs**[22], **Reuven J. Schore**[23], **Kimo C. Stine**[24], **Clifford M. Takemoto**[7], **Jeffrey A. Toretsky**[25,26], **Joseph F. Costello**[4], **Adam B. Olshen**[27,28], **Chip Stewart**[2], **Yongjin Li**[3], **Jing Ma**[29], **Robert B. Gerbing**[30], **Todd A. Alonzo**[31], **Gad Getz**[2,32,33], **Tanja Gruber**[29,34], **Todd Golub**[2,35,36], **Kimberly Stegmaier**[2,35,36], and **Mignon L. Loh**[1,37]

[1]Department of Pediatrics, Benioff Children's Hospital, University of California, San Francisco, San Francisco, CA

[2]Broad Institute of MIT and Harvard, Cambridge, MA

[3]Department of Computational Biology, St. Jude Children's Research Hospital, Memphis, TN

[4]Department of Neurological Surgery, University of California, San Francisco, CA

[5]Hartwell Center for Bioinformatics and Biotechnology, St. Jude Children's Research Hospital, Memphis, TN

[6]Stead Family Department of Pediatrics, University of Iowa Carver College of Medicine, Iowa City, IA

[7]Department of Pediatrics, The Johns Hopkins Hospital, Baltimore, MA

Corresponding Authors: Elliot Stieglitz and Mignon Loh. Loh Laboratory, Helen Diller Family Comprehensive Cancer Center, Box 3112 1450 3rd Street, Room 230 San Francisco, CA 94158 Tel: (415) 514-0853 Fax: (415) 502-5127. E.S. (elliot.stieglitz@ucsf.edu) and M.L.L (mignon.loh@ucsf.edu)..

Author Contributions

E.S., L.C.G, T.M., E.E., A.Y., K.B., and S.L.A. performed the experiments. E.S., A.T.-W., Y.D.W., T.M., M.R., A.B.O., Y.L., J.M., R.B.G., and T.A.A. performed data analysis. G.A., M.B., B.C., T.C., G.V.D., P.D.E., M.N.F., R.K.G., R.J.H., J.H., C.H., Y.L.L., Y.H.M., D.H.M., E.R.N., P.A.R., R.J.S., K.C.S., C.M.T., J.T., contributed reagents, materials and analysis tools. E.S., A.T.-W., and M.L.L. wrote the first draft of the manuscript. T.Y.C. performed statistical analysis. J.F.C., C.S., G.G., T.A.G., T.R.G., K.S., M.L.L. supervised research. S.S., and S.R.O managed the project. All coauthors contributed to the final version of the manuscript.

Competing Financial Interests

The authors declare no competing financial interests.

URLs.

Exome Aggregation Consortium (ExAC), http://exac.broadinstitute.org; DesingStudio, http://desgnstudio.illumina.com.

Accession codes.

High-throughput sequencing data have been deposited in the database of Genotypes and Phenotypes (dbGaP) under accession phs000973.v1.p1.

[8]Department of Pediatrics, Emory University School of Medicine, Aflac Cancer and Blood Disorder Center, Atlanta, GA

[9]Department of Pediatrics, Texas Tech University, El Paso, TX

[10]Department of Pediatrics, Seattle Children's Hospital, Seattle, WA

[11]Department of Pediatrics, Stanford School of Medicine, Stanford, CA

[12]Winthrop P. Rockefeller Cancer Institute, University of Arkansas for Medical Sciences, Little Rock, AR

[13]Department of Pediatric Hematology Oncology, University of Utah, Salt Lake City, UT

[14]Department of Pediatrics, Children's Hospital of Pittsburgh of UPMC, Pittsburgh, PA

[15]Department of Pediatrics, Washington University School of Medicine, St. Louis, MO

[16]Division of Hematology/Oncology, The Hospital for Sick Children, Toronto, Ontario, Canada

[17]Pediatric Hematology Oncology, SSM Cardinal Glennon Children's Medical Center, Saint Louis, MO

[18]Division of Pediatric Hematology Oncology, Children's Hospitals and Clinics of Minnesota, Minneapolis, MN

[19]Department of Pediatrics, Texas Children's Hospital, Baylor College of Medicine, Houston, TX

[20]Pediatric Hematology Oncology, Pediatric Specialists of Lehigh Valley Hospital, Bethlehem, PA

[21]Pediatric Bone Marrow Transplant Program, Oregon Health & Science University, Portland, OR

[22]Department of Pediatrics, University of North Carolina at Chapel Hill, NC

[23]Division of Pediatric Oncology, Children's National Medical Center, Washington, DC

[24]Department of Pediatrics, University of Arkansas for Medical Sciences, Little Rock, AR

[25]Department of Pediatrics, Georgetown University, Washington, DC

[26]Department of Oncology, Georgetown University, Washington, DC

[27]Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, CA

[28]Department of Epidemiology and Biostatistics, University of California, San Francisco, CA

[29]Department of Pathology, St. Jude Children's Research Hospital, Memphis, TN

[30]Department of Statistics, Children's Oncology Group, Monrovia, CA

[31]Keck School of Medicine, University of Southern California, Los Angeles, CA

[32]Harvard Medical School, Boston, MA

[33]Department of Pathology and Cancer Center, Massachusetts General Hospital, Boston, MA

[34]Department of Oncology, St. Jude Children's Research Hospital, Memphis, TN

[35]Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, MA

[36]Division of Hematology/Oncology, Boston Children's Hospital and Harvard Medical School, Boston, MA

[37]Department of Pediatrics, Benioff Children's Hospital, Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, CA

[#] These authors contributed equally to this work.

Abstract

Juvenile myelomonocytic leukemia (JMML) is a myeloproliferative neoplasm (MPN) of childhood with a poor prognosis. Mutations in *NF1, NRAS, KRAS, PTPN11* and *CBL* occur in 85% of patients, yet there are currently no risk stratification algorithms capable of predicting which patients will be refractory to conventional treatment and therefore be candidates for experimental therapies. In addition, there have been few other molecular pathways identified aside from the Ras/MAPK pathway to serve as the basis for such novel therapeutic strategies. We therefore sought to genomically characterize serial samples from patients at diagnosis through relapse and transformation to acute myeloid leukemia in order to expand our knowledge of the mutational spectrum in JMML. We identified recurrent mutations in genes involved in signal transduction, gene splicing, the polycomb repressive complex 2 (PRC2) and transcription. Importantly, the number of somatic alterations present at diagnosis appears to be the major determinant of outcome.

INTRODUCTION

Juvenile myelomonocytic leukemia (JMML) is a rare but aggressive form of childhood leukemia that exhibits both myelodysplastic and myeloproliferative properties[1]. The only curative therapy is hematopoietic stem cell transplant (HSCT)[2]. However, some patients exhibit highly aggressive disease despite HSCT, while spontaneous remissions are occasionally observed in others with minimal therapy[3,4]. The lack of current laboratory, genetic, and clinical features to distinguish these patients[5,6] presents a clinical dilemma for physicians and parents. We hypothesized that complete genomic characterization of JMML would aid in distinguishing these cases and further identify relevant molecular targets for the development of novel therapies in patients with the most aggressive disease phenotypes.

Mutations in *NF1, NRAS, KRAS, PTPN11* and *CBL* ("Ras pathway") currently allow for a molecular diagnosis in 85% of patients[7-11]. Recently, secondary mutations in *SETBP1* and *JAK3* were identified by whole exome sequencing in a small number of patients with JMML at diagnosis[12]. We subsequently identified several patients who had an increase in allele frequency of *SETBP1* mutations at relapse. We then harnessed droplet digital (dd) PCR to show that subclonal *SETBP1* mutations were present in nearly a third of patients with JMML at diagnosis and independently predicted relapse[13]. These findings indicated a level of genetic complexity previously unrecognized in JMML, and given the limited numbers of patients with non-syndromic, *de novo* JMML who have had exome sequencing performed, we set out to assess the genomic landscape of JMML.

We sequenced samples from patients (n=29) with matched tumor/normal pairs. Seven of these patients also had serially acquired relapse and/or transformation to AML samples available for sequencing. We then validated our findings in an independent cohort of 71 patients (Supplementary Figure 1), of whom nine had paired diagnostic-relapse samples available. Two of the 29 patients that had exome sequencing were suspected of having Noonan syndrome. Upon confirmation, they were removed from all outcome analyses which were specific to somatically mutated JMML.

RESULTS

Sequencing of JMML samples using optimized algorithms

We performed whole exome sequencing (WES) at a mean coverage of 95x (Supplementary Table 1) on 22 patients with paired germline-diagnosis samples and an additional seven patients with germline-diagnosis-relapse samples (Figure 1). Due to the frequent contribution of germline mutations in the development of JMML[7,11], we optimized an algorithm to detect tumor in normal content (deTiN) to retrieve mutations that would otherwise have been missed using a traditional tumor-normal bioinformatics approach. Four tissue types of germline material were used to serve as normal controls, including buccal cells, cord blood, Epstein Barr virus (EBV) immortalized lymphoblasts, and fibroblasts. However, by comparing several intra-patient germline sources that contained varying degrees of tumor content, it became evident that each tissue type had different amounts of tumor contamination in the normal. For example, in patient UPN2026, we first detected a heterozygous mutation in *RRAS2* from a buccal swab but repeat sequencing of EBV immortalized B cells was wild type (Supplementary Figure 2). We therefore implemented deTiN, to both assess and correct for the purity of each germline source.

In total, we identified 10 genes that were mutated outside of the previously documented five Ras pathway lesions (Supplementary Table 2). These mutations occurred in known oncogenes and tumor suppressors that broadly fell into categories of genes affecting the Ras pathway, signal transduction, transcription factors, epigenetic regulation, and the spliceosome complex (Figure 1 and Table 1).

We next carried out targeted deep sequencing (mean coverage 1380X) on 71 diagnostic tumors as well as nine patients who had diagnostic/relapse pairs for the 15 genes identified from exome sequencing in order to determine the frequency of these mutations. In addition, we sequenced *JAK3* as prior reports had identified recurrent subclonal mutations in some patients. We also included eleven patients from the discovery cohort in the resequencing cohort for validation purposes. All pathogenic mutations identified from exome sequencing were identified by targeted resequencing. Recurrent mutations were detected in 14 of the 15 genes, excluding *RRAS2* which was only identified in the discovery cohort (Supplementary Table 3). The incidence of mutations in these genes is presented for a total of 100 patients, of whom 29 were in the discovery cohort and 71 in the validation cohort (Supplementary Figure 3). RNA-sequencing (RNA-seq) data was available in 14 patients (15 samples), and confirmed that each of the pathogenic variants detected on exome and deep sequencing were transcribed and detected in the RNA (Supplementary Table 4).

Somatic copy number alterations

The most common karyotypic abnormality identified in WES was deletion of one copy of chromosome 7, which occurred in 5 of the exome cases (Supplementary Figure 4). One patient with a t(11;17) detected on diagnostic cytogenetics had additional evidence of disruption of the *NF1* locus on fluorescent in-situ hybridization (data not shown), and was found to have a germline mutation in WES. The *NF1* c.2041C>T allele had a fractional abundance of 33% in RNA transcriptome sequencing (15/45 reads). However, RNA-seq failed to identify a fusion RNA in this patient, suggesting that the translocation breakpoints might be intergenic. RNA-seq in 14 additional patients did not reveal any structural rearrangements resulting in chimeric transcripts. Consistent with previous reports[14,15] the only additional recurrent structural variants were copy neutral isodisomy at 11q23.3 and 17q11.2 where *CBL* and *NF1* are located, respectively (Supplementary Figure 5).

Expanding the spectrum of Ras pathway mutations

In addition to canonical *NF1*, *KRAS*, *NRAS*, *PTPN11* and *CBL* alterations, we identified a previously described *RRAS* mutation in one patient[16], as well as an *RRAS2* alteration in another patient, neither of which harbored other driver mutations at diagnosis. Mutations in *RRAS2*, a member of the Ras GTPase superfamily encoding the TC21 oncoprotein, have previously been limited to solid malignancies including ovarian, endometrial and squamous cell carcinomas[17] and regulate TGFβ signaling through the loss of *NF1*[18]. The *RRAS2* point mutation in our patient results in the p.Q72L amino acid substitution that was previously shown to confer transformative properties as evidenced by increased colony formation in vitro and increased tumorgenicity after inoculation in mice compared to wild type *RRAS2*[19].

While Ras pathway lesions have traditionally been thought to represent largely mutually exclusive events[20], co-existing mutations in *NRAS*, *KRAS*, *PTPN11*, *CBL*, and *NF1* were found in 11/100 (11.0%) of patients (Figure 2). Analysis of single colonies for two patients with compound Ras pathway mutations indicate that both mutations occurred in the same colony (Supplementary Table 5). *PTPN11* and *NF1* lesions were the most frequent of these cooperative events. In addition, one patient harbored two *NRAS* lesions (p.G13D and p.Q61K).

NF1 mutations in patients without clinical NF1

Two patients in our discovery cohort who were not suspected of having neurofibromatosis type 1 (NF1) nevertheless harbored germline *NF1* mutations. One of these patients was heterozygous for the mutation in both the tumor and matched normal (buccal cells) with no loss of heterozygosity noted in the tumor. Another patient had an *NF1* mutation with an allelic fraction of 10.1% in the germline tissue (buccal cells). The allelic fraction in the tumor increased to 45.3% after a copy neutral duplication at 17q. While a low allelic event in the germline tissue could be explained as a somatic only variant with contamination of tumor into normal, application of the deTIN algorithm estimated that no such contamination (95% CI=0-3.35%) occurred, based on the absence of copy number alteration in the normal tissue. This raises the possibility of somatic mosaicism, which occurs in clinical NF1 and often presents with subtle clinical features[21]. In our validation cohort, an additional six patients harbored *NF1* mutations but only three were suspected of having clinical NF1. It is

not possible to determine the somatic versus germline origin of these lesions, as germline tissue was not available on these patients. These observations are consistent with a previous study indicating that JMML can be the first manifestation of NF1 in some affected infants and young children[22].

Activated JAK/STAT corresponded with poor clinical outcome

We previously found that increased STAT5 phosphorylation in a subset of myeloid cells is a general feature of JMML[23]. The tumor suppressor LNK is encoded by *SH2B3* and negatively regulates JAK/STAT signaling[24]. Mutations in *SH2B3* were first reported in a variety of adult myeloproliferative neoplasms (MPNs) including primary myelofibrosis and essential thrombocythemia, as well as in isolated erythrocytosis[25-27]. The alterations included nonsense and missense mutations affecting the pleckstrin homology (PH) and/or the Src homology 2 (SH2) domains. Mutations in *SH2B3* have also been identified in lymphoid malignancies including acute lymphoblastic leukemia as both germline and somatic events[28]. Two patients in the discovery cohort possessed compound heterozygous mutations in *SH2B3* in addition to a known JMML driver mutation. Each patient had a nonsense mutation in both the PH domain as well as the SH2 domain (Figure 3A). While one patient harbored these mutations at diagnosis, the other only had detectable *SH2B3* mutations upon transformation to acute myeloid leukemia (AML). The latter patient relapsed with AML three years after his initial diagnosis. Protein lysates collected from unsorted primary mononuclear cells from these patients showed a decrease in LNK by immunoblot, commensurate with the allelic fraction of the patients nonsense mutations, respectively (Figure 3B). One additional patient without any other identifiable mutations was found via exome sequencing to harbor a germline *SH2B3* p.E400K alteration, which is a non-synonymous SNP that has been previously associated with idiopathic erythrocytosis[29]. In the validation cohort, an additional four patients were found to harbor *SH2B3* mutations for a cumulative incidence of 7.0% (7/100) (Supplementary Table 3).

We did not identify any activating *JAK3* mutations in our exome discovery cohort, but performed resequencing with improved coverage (680X) in our validation cohort given prior reports[12]. These studies uncovered *JAK3* mutations in 6 of 83 cases (7.3%). Of note, 3 of 4 patients with *JAK3* mutations at diagnosis went on to relapse. Two other patients had lesions detected only at relapse (allelic fractions of 8.6% and 37.4%) and presumably had subclonal mutations at diagnosis below the threshold of identification by our exome sequencing.

SETBP1 mutations are frequently subclonal at diagnosis

We detected a similar incidence of *SETBP1* (7/100, 7.0%) mutations using deep sequencing as was previously reported by Sakaguchi et al[12]. An additional 17 patients were found to harbor subclonal *SETBP1* mutations using ddPCR [13]. Combined, subclonal and clonal *SETBP1* mutations occurred in 24/100 patients (Supplementary Table 6). No mutations were found on exome sequencing in *ETNK1*, a gene that was recently reported to co-occur with *SETBP1* in atypical CML[30].

Transcription Factor mutations in JMML

We identified mutations in *GATA2*[31], a transcription factor broadly involved in hematopoiesis. Recent work has shown that germline *GATA2* mutations are responsible for several syndromes leading to a predisposition to myeloid malignancies and opportunistic infections[32-34]. Two patients harbored three heterozygous mutations, all of which occur in the regulatory zinc finger domain (ZF1). The overwhelming majority of mutations occurring in prior reports are in either ZF1 or ZF2[34-36]. One of these two patients harbored both a germline and a somatic *GATA2* mutation as previously described[37]. The other patient was found to have a somatic p.N317S variant on one allele which has been previously detected in patients with both MDS and AML[38,39].

A heterozygous nonsense *RUNX1* mutation was also found in one patient at diagnosis and in another at relapse, representing the first reports of haploinsufficient *RUNX1* mutations in JMML[40]. Haploinsufficient missense mutations in *RUNX1* have been previously reported in other malignancies such as MDS and SML, and are predicted to be inactivating[41].

Wiskott-Aldrich syndrome mutations (WAS) in JMML

Previous reports have documented germline *WAS* mutations in patients suspected of having JMML[42,43]. Similarly, we detected one male patient with a germline *WAS* p.H180N mutation which has been described in several patients with Wiskott-Aldrich syndrome[44]. Interestingly, this patient also had a somatic *NRAS* p.G12S mutation. To the best of our knowledge there are no prior reports in the literature of patients with concomitant germline mutations in *WAS* and somatic mutations in Ras pathway genes.

Mutations in Epigenetic modifiers are frequent in JMML

In contrast to previous reports that emphasized the rarity of genetic mutations in epigenetic modifying genes in JMML[45-47], we identified mutations in 14/100 patients for a combined incidence in our cohort of 14.0%. Components of the polycomb repressive complex 2 (PRC2), including *EZH2* (4/100, 4.0%) and *ASXL1* (8/100, 8.0%) were mutated at diagnosis and clonally expanded in several patients at relapse (Supplementary Tables 2 and 3). Interestingly, all four patients with mutations in *EZH2* (7q36.1) also had monosomy 7 (Figure 2). We also found *DNMT3A* mutations in 3 patients, including one at the previously reported R882 hotspot and two other frame-shift mutations. Genome wide DNA methylation analysis demonstrated that four of five patients with *ASXL1* mutations had a globally hypermethylated profile compared to those with wild type *ASXL1* (Supplementary Figure 6), while one patient with a *DNMT3A* mutation had a globally hypomethylated profile consistent with previous reports[48,49].

The spliceosome complex is implicated in JMML

A member of the spliceosome complex, *ZRSR2* was mutated in one patient in our discovery cohort as well as two additional patients upon targeted resequencing. A pair wise analysis was carried out comparing RNA-seq data from one patient with a *ZRSR2* mutation and 14 patients wild type for *ZRSR2*. The patient with mutant *ZRSR2* had significantly increased retention of U12-type introns compared to those with wildtype *ZRSR2*, consistent with prior

reports in adults with MDS[50] (Supplementary Figure 7). These are the first reported *ZRSR2* mutations in JMML[47,51,52]. No mutations in *SF3B1*, *U2AF1* or *SRSF2* were found during our discovery phase.

Phylogenetic evolution at relapse

Focusing on patients with both diagnostic and relapsed material allowed us to plot the acquisition of mutations over time (Supplementary Figure 8 and Supplementary Tables 2 and 3). All of the pathogenic mutations at diagnosis were present at relapse. However, it should be noted that due to the low number of mutations per sample (0.468 mutations per megabase), these analyses lacked the power to definitively track these clones over time. To further investigate the clonality of the mutations at diagnosis, we analyzed individual colonies obtained from diagnostic colony forming unit-granulocyte-macrophage assays (Supplementary Table 7). Patients UPN1420 and UPN2531 were particularly interesting because of the pattern of *SH2B3* mutations detected. UPN2531 possessed a *PTPN11* driver mutation along with compound heterozygous mutations in *SH2B3* at diagnosis. By contrast, UPN1420 only acquired mutations in *SH2B3* upon transformation to AML (Supplementary Figure 9A). More strikingly, colonies were homozygous for either the p.Q258* (19/24) or p.F390fs mutations (5/24). These clones were likely the result of copy number alterations occurring on 12q in individual cells that contained heterozygous *SH2B3* mutations in a model of convergent evolution. Copy number variation analysis of this sample at relapse shows distinct duplication and loss at 12q where *SH2B3* resides (Supplementary Figure 9B). We hypothesize that the allele containing the p.Q258* mutation duplicated in one clone while the p.F390fs achieved homozygosity in the other.

Clinical and Biologic Features at Diagnosis

A total of 98 of 100 children with JMML diagnosed between 2001 and 2013 at North American institutions were included in our analysis, with two patients excluded due to insufficient follow-up data (Supplementary Tables 8). Zero or 1 somatic alterations (pathogenic mutations or monosomy 7) were identified in 64 (65.3%) patients at diagnosis compared to 2 or more alterations in 34 (34.7%) (mean number of alterations = 1.32; Range = 0 to 4). Several clinical and biologic characteristics showed statistically significant differences when compared between these groups (Supplementary Table 9). A higher proportion of patients diagnosed with 2 or more alterations were older ($p<0.001$) and male ($p=0.001$). These patients also demonstrated a higher rate of monosomy 7 ($p<0.001$) or somatic *NF1* mutations ($p<0.001$).

Somatic Alterations at Diagnosis Predicts Outcome

Using the number of somatic events at diagnosis, we evaluated differences in event free survival (EFS) and overall survival (OS). We initially identified a trend towards improved outcome in patients with 0-1 somatic alterations compared to those with 2 or more alterations in our 27 patient exome cohort ($p= 0.12$)(Supplementary Figure 10a). This was later validated in our independent cohort of 71 patients ($p=0.0004$)(Supplementary Figure 10b). Combining both cohorts (98), patients with 0-1 somatic alterations at diagnosis had improved outcomes compared with patients who had 2 or more alterations (EFS HR 2.09,

p=0.009, OS HR 2.02 p=0.018). The 10-year EFS and OS rates according to number of somatic events at diagnosis were as follows (± SE): 0-1 alterations, 60.6% ±6.2% and 65.1% ±6.0% and 2 or more alterations, 22.7% ± 7.4% and 29.0% ± 8.3% (Figures 4A and 4B; p=0.003 and p=0.002). Of note, log-rank tests detected statistically significant differences in times to both outcomes based on this predictor variable. Interestingly, EFS and OS based on canonical Ras pathway driver alterations including *PTPN11, NF1, NRAS, KRAS* and *CBL* were not statistically different from each other in our cohort (p=0.834)(Supplementary Figure 11). Of note, EFS for patients with double Ras pathway mutations (14%, n=7) in our cohort was significantly inferior compared to patients with a single Ras pathway mutation (62%, n=55)(p=0.017). This difference in outcome implies a dosage effect based on the number of Ras mutations and suggests that both compound Ras mutations have biological activity.

Prior studies have assigned a poorer prognostic significance to several clinical and laboratory characteristics in patients with JMML including; older age, male gender, lower platelet count, higher fetal hemoglobin, monosomy 7, *PTPN11* status, and subclonal *SETBP1* status. In univariate Cox analyses (Supplementary Table 10), characteristics that reached significance for EFS and OS were: age ≥24 months (HR 2.09 CI 1.20-3.62, p=0.009 and HR 2.03, CI 1.13-3.66, p=0.018) and monosomy 7 (HR 1.24, CI 1.01-1.52, p=0.040 and HR 1.29, CI 1.04-1.60, p=0.019). Importantly, the most statistically significant covariate in this analysis was the number of somatic alterations at diagnosis (EFS HR 2.65, CI 1.27-5.54, p=0.01 and OS HR 3.13, CI 1.39-7.01, p=0.006). Furthermore, when modeled in a Cox multivariate regression, only the number of somatic alterations at diagnosis retained statistical significance for both EFS and OS. Two or more somatic events remained independently prognostic of poor outcome (HR 2.86 CI 1.35-6.05, p=0.0006 and HR 3.00, CI 1.27-8.40, p=0.0008) after adjusting for age, platelet count, fetal hemoglobin, and NF1 status at diagnosis (Table 2).

We assessed the relapse incidence in patients who were treated with HSCT. Of the 79 patients in our cohort who received HSCT and had follow-up data, 49 (62%) harbored 0-1 alterations at diagnosis compared to 30 (38%) with 2 or more. As expected, the cumulative incidence of relapse from time of HSCT was significantly higher in the cohort with 2 or more events at diagnosis (77.7% ± 8.56) compared to those with 0-1 events (38.4% ± 8.33, log-rank p= 0.0002) (Supplementary Figure 12).

DISCUSSION

While previous genomic and biochemical analyses in JMML have emphasized the dependence on the Ras pathway for initiation of disease[12], we hypothesized that additional pathways may be important for both disease initiation and progression. By studying patients with progressive disease and by optimizing the use of germline controls, the mutational spectrum in JMML is now expanded to include some molecules with “off the shelf” therapeutic implications while others will require further investigation. In addition to identifying another member in the Ras family of oncogenes mutated in JMML (*RRAS2*), we also uncovered additional genetic pathways contributing to leukemogenesis, including

upregulated JAK/STAT signaling, and epigenetic modification through the PRC2 and spliceosome complexes.

This report is the first to identify *SH2B3* mutations in JMML. Combined with a previous study that uncovered missense JAK3 alterations[12], these data raise the possibility of using JAK inhibitor therapy in the subset of patients harboring these mutations. This is also the first report of *ZRSR2* mutations in this disease.

In contrast to chronic myelomonocytic leukemia (CMML) where epigenetic mutations are frequent[53], previous reports in JMML have focused on the rarity of these events. However, 14% of patients in our cohort harbored genetic alterations in epigenetic modifying genes. We show that *ASXL1* mutations result in globally hypermethylated profiles. Patients exhibiting aberrant methylation would be predicted to respond to DNA hypomethylating agents and prior case reports support that these agents have clinical utility in JMML[54]. Currently, a phase I/II trial using 5-azacytidine in children with JMML is enrolling patients in Europe. Mutations in *DNMT3A* typically cause decreased methyltransferase activity and result in focal areas of CpG hypomethylation[55].

Obtaining germline tissue in children was of critical importance in our study because approximately 25% of patients have inherited syndromes that predispose to the development of JMML. In addition, unlike other hematologic malignancies where one can use remission marrow samples for germline tissue, JMML patients rarely respond to traditional chemotherapy prior to transplant[56]. By developing innovative algorithms to identify TiN content with retrieval of minor differences in tumor-normal comparisons, we identified several variants that would have otherwise been overlooked using standard approaches. We showed that sources such as buccal swabs, cord blood and EBV immortalized lymphocytes were all suboptimal compared to bone marrow fibroblasts. While buccal samples are frequently “contaminated” with tumor by infiltrating leukemic cells, cord blood and EBV samples revealed biologic insights into this disease.

In the majority of cord blood samples, the canonical mutation was retrievable, thus revealing the *in utero* nature of leukemic initiation[57,58]. Of note, the mean age at diagnosis of the five patients who had cord blood available was 6.6 months. Interestingly, no secondary mutations were identified at diagnosis in any of these children, which may explain why younger age at diagnosis has traditionally been a more favorable clinical risk factor for outcome. B cells immortalized with EBV also provide insights into the range of hematopoietic lineages affected by JMML mutations, with many patients appearing to have clonal involvement of their mutations in the B cell compartment, while others only had myeloid cell involvement. There were insufficient patients available to determine how frequently this occurs.

Several mutated genes novel to JMML were identified in relapsed samples, including *RUNX1* and *JAK3*. However, in contrast to most malignancies where a branching model of disease evolution with mutation acquisition and dropout is common[59], JMML appears to have a unique model of progression. Specifically, mutations identified at diagnosis invariably persisted at relapse in all 16 patients analyzed over time. In addition, the cancer

cell fraction for mutations identified at diagnosis on exome sequencing approached 100%, implying that all mutations occurred within a single dominant clone. Single colony analysis in three patients with available cryopreserved cells was consistent with a linear model of the acquisition of additional genetic features (Supplementary Figure 9 and Supplementary Tables 5 and 7). When developing new therapeutic strategies for these patients, it will be critical to understand how secondary mutations alter the behavior of these cells in contrast to cells harboring only the primary lesion. Indeed, combinatorial therapies with agents that target the Ras pathway as well as the secondary genetic event may be necessary, as our work demonstrates that the clone harboring the secondary event frequently expands at disease progression. For example, combining a MEK inhibitor with DNA hypomethylating agents could prove more efficacious in the correct genetic context than monotherapy with a MEK inhibitor alone[60,61].

Previous reports have reached different conclusions regarding the prognostic significance of canonical Ras pathway mutations in JMML[5,6]. We show that the five Ras pathway genes are not independently prognostic of outcome, precluding their sole use in a risk stratification algorithm. Consistent with our hypothesis, we found that the number of mutations at diagnosis, rather than the type of mutations, possesses prognostic relevance.

In conclusion, we demonstrate that patients with JMML who harbor two or more somatic alterations at diagnosis had a significantly worse EFS and OS compared to those with one or fewer events. Prior reports which identified older age, male gender, and lower platelet count as markers of poor outcome were likely descriptors of patients with two or more underlying somatic alterations. Importantly, in addition to identifying patients with aggressive disease, our expanded mutational spectrum will serve as a platform for studying the interactions of these secondary events in the context of hyperactive Ras/MAPK signaling. These studies will allow further interrogation of multiple molecular targets that can be exploited in patients predicted to have poor outcomes using conventional HSCT.

Online Methods

Patients

Twenty-nine patients diagnosed with JMML were included in the discovery cohort and had germline-tumor pairs evaluated by whole exome sequencing. Seven of these patients also had relapse samples included in the exome analysis (Supplementary Table 1). A validation cohort comprised of 71 distinct patients was tested for the 16 genes found to be mutated (including *JAK3*, for which we did not identify lesions on exome) during the discovery phase. Ninety-eight patients with confirmed JMML based on the internationally accepted criteria[62] were included in the survival analysis. In addition to the date of diagnosis, follow-up, and sequencing data, 12 potential clinical and laboratory prognostic factors were analyzed (Supplementary Tables 8). Ten of these patients had relapsed tissue included in the targeted resequencing. Fifty of the validation cohort samples were from patients that were enrolled in clinical trial AAML0122, which evaluated the safety and efficacy of a farnesyl transferase inhibitor as a window treatment before HSCT[63]. Eleven of the 81 validation samples were previously included in the discovery cohort. Fourteen patients from the combined cohorts also had RNA-seq performed. Twelve patients had genome wide DNA

methylation analysis performed. Approvals for these studies were obtained from the University of California San Francisco (UCSF) Committee on Human Research. All participants/guardians provided informed consent in accordance with the Declaration of Helsinki.

Tissue

Germline DNA was extracted using standard methods from buccal swabs, cord blood, skin fibroblasts, bone marrow fibroblasts, or EBV immortalized cell lines as available. Tumor DNA was extracted using standard methods from bone marrow or peripheral blood mononuclear cells obtained at diagnosis and relapse when available. RNA was extracted using a TRIzol chloroform method.

Discovery

Whole exome sequencing was performed utilizing the Illumina platform as described previously. In brief, DNA was fragmented by sonication (Covaris Inc., Woburn, MA) to 150 bp and further purified using Agencourt AMPure XP beads. Fifty ng of size-selected DNA was then ligated to specific adaptors during library preparation (Illumina TruSeq, Illumina Inc., San Diego, CA). Each library was made with sample-specific barcodes and quantified by quantitative PCR (Kapa Biosystems, Inc., Woburn, MA), and two libraries were pooled to a total of 500 ng for exome enrichment using the Agilent SureSelect hybrid capture kit (Whole Exome_v1.1; Agilent Technologies, Santa Clara, CA). Several captures were pooled further and sequenced in one or more lanes to a final equivalent of two exomes per lane on a HiSeq 2500 system (Illumina Inc., San Diego, CA).

Somatic variant, germline variant, and small deletions and insertion calling was performed within the Firehose environment at the Broad Institute with the previously published MuTect[64], Haplotype[65,66] and Indelocator[67]. In addition, we manually reviewed all candidate driver events using the Integrative Genome Viewer Genome-wide copy-ratio profiles were inferred using CAPSEG[68]. Read-depth at capture targets in tumor samples was calibrated to estimate copy-ratio using depths observed in a panel of normal genomes. We then segmented the copy-ratio profiles using the circular binary segmentation (CBS) algorithm[68]. Next we performed allelic copy analysis using reference and alternate counts at germline heterozygous SNP sites. This data allows for identification of copy neutral LOH events and inference of the contribution of each chromosome to observed copy profiles. We then combined the two data types to derive allelic copy ratio profiles (Supplementary Figure 4).

To determine which observed somatic variants were pathogenic, we first identified events that had a minor allele frequency less than 0.001 in the ExAC database and were marked as deleterious by Mutation Assessor. Next, we enriched for mutated genes that were previously implicated in myeloid malignancies. Finally, we compared the mutation's variant classification with the expected gene function and selected inactivating mutations for tumor suppressors and activating mutations in regulatory domain hotspots for oncogenes as annotated in the COSMIC database[70,71].

To assess tumor in normal (TiN) contamination we applied deTiN (manuscript in preparation) using both copy number events and mutations when available. Briefly, deTiN estimates tumor in normal contamination (TiN) using evidence for putative somatic events observed in the normal (Supplementary Table 11). The contamination estimates were then applied to our SNV and InDel calls to recover events previously rejected because of an impure germline source. We were then able to ensure higher sensitivity even with non-optimal germline samples. We observed distinct differences of tumor in normal content specific to each tissue type (Supplementary Figure 13). Fibroblasts from either bone marrow or skin (n=4) had the lowest TiN content with a mean of less than 1%. Sixty percent of the buccal samples (n=10) collected at the time of diagnosis were infiltrated by leukemia cells and had a mean TiN content of 11%, ranging from 0% to 25%. A previous study cast doubt on the use of EBV-immortalized cell lines as a source of germline material in JMML[13]. Consistent with this observation, EBV-transformed lymphoblasts from two of six JMML patients had a TiN of <1% while four others were clearly involved by the mutations in a clonal fashion. Similarly, cord blood (n=5) from patients contained greater than 60% TiN in 3 samples reflecting the presence of leukemia cells at birth[72].

To estimate TiN using mutations, deTiN fits the fraction of alternate reads in the normal sample at candidate somatic sites. When samples had fewer than three somatic single nucleotide variants identified with our standard mutation calling pipeline, we used alternate and reference allele counts at previously identified canonical events in those samples to build the contamination model. To construct this model, deTiN first calculates the probability that a given candidate mutation is a somatic or germline event over a range of potential TiN levels. Then we iterate over TiN levels, each time constraining the fit using mutations. We repeat the same process using heterozygous SNPs in regions of allelic imbalance and constrain the estimate using the amount of allele imbalance observed in the normal in relation to the tumor. We then combine each of these estimates by adding the likelihood curves for each.

Next we use the combined TiN estimate to keep mutation calls from variants that were originally rejected as germline, but are actually 1000 times more likely to be explained by TiN.

Validation

A customized TruSeq amplicon kit was designed utilizing the online DesignStudio pipeline (llumina) that targeted the entire coding regions of 16 genes (ASXL1, CBL, DNMT3A, EZH2, GATA2, JAK3, KRAS, NF1, NRAS, PTPN11, RRAS, RRAS2, RUNX1, SETBP1, SH2B3, and *ZRSR2*), and covered a total of 49,710 bp with 184 amplicon regions. Libraries prepared from DNA of each of 91 samples were indexed and subjected to 250 bp paired-end sequencing on Illumina MiSeq, and data of three independent MiSeq runs were combined to generate an average of 761,678 reads (median 383,391 reads, Supplementary Figure 14A) per sample. While 95.6% of total reads passed quality filtering, more than 97.4% of mapped reads (99.8% mapping rate) were in the targeted regions, which provided an average of depth coverage at 1380X across the targeted regions (Supplementary Figure 14B) (Supplementary Table 12). Hotspots not adequately covered on the MiSeq were analyzed by

Sanger. The targeted sequencing data was analyzed as described previously[73]. In brief, after quality trimming, reads of each sample were aligned to the human genome reference sequence (GRC37/hg19), and CLC Genomic Workbench (v7.5, CLC Bio, Denmark) was used for mapping and variant calling. An in-house tool was then used to filter through the variant list, and the consequence of protein sequence changes resulting from detected DNA-sequence changes were predicted using the functional annotation tool ANNOVAR. Only variants with greater than 5% allelic fraction and at least 8x depth coverage of the mutation allele were included as a somatic alteration, while additional detectable variants were reported for completeness (Supplementary Table 3).

DNA methylation analysis

Genomic DNA was bisulfite converted using the EZ DNA Methylation Kit (Zymo Research) and processed on Infinium HumanMethylation450 bead arrays (Illumina Inc.) according to the manufacturer's protocol. Probe-level signals for individual CpG sites were subject to both background and global dye-bias correction[74]. Probes that map to regions with known germline polymorphisms, to multiple genomic loci, or to either sex chromosome were filtered out. Two-way unsupervised hierarchical clustering was performed using Euclidean distance and Ward linkage on the 1,486 most variable CpG sites across the cohort, with variability ranked by standard deviation. Four subjects with clonal *ASXL1* mutations at diagnosis were included for analysis. An additional seven patients were chosen as controls as they had similar mutational profiles compared to the subjects except they did not possess any known genetic mutations in epigenetic regulating genes (Supplementary Figure 6).

RNA sequencing

RNA was extracted, prepared into mRNA libraries, and sequenced by Illumina HiSeq resulting in paired 50nt reads, and subjected to quality control and previously described[75,76]. RNA reads were aligned using TopHat. Gene expression was quantified for transcripts corresponding to Gencode v12 using RPKM. Rearrangements detection was performed using dRanger as described previously[77]. Fusion detection was performed using TopHat-Fusion.

Analysis of RNA sequencing data

For RNA-seq, paired-end sequencing was performed using the HiSeq platform with 100bp read length. Paired-end reads from RNA-seq were aligned to the following 4 database files using BWA (0.5.10) aligner: (1) the human GRCh37-lite reference sequence, (2) RefSeq, (3) a sequence file representing all possible combinations of non-sequential pairs in RefSeq exons, (4) AceView database flat file downloaded from UCSC representing transcripts constructed from human ESTs. The mapping results from (2) to (4) were translated to human reference genome coordinates. In addition, they were aligned using STAR 2.3.0 to the human GRCh37-lite reference sequence without annotations. A BAM file was constructed by selecting the best alignment among the five mappings. Poor quality mappings were improved using SIM4 when possible to generate the final BAM. The coverage was calculated using an in-house pipeline. SV detection was carried out using CICERO, a novel

algorithm that uses de novo assembly to identify structural variation in RNA-seq (Li et al, manuscript in preparation).

Digital gene expression profiling

The transcript expression levels were estimated as Fragments Per Kilobase of transcript per Million fragments (FPKM); gene FPKMs were computed by summing the transcript FPKMs for each gene using Cuffdiff2[78]. A gene was considered "expressed" if the FPKM value >= 0.5 based on the distribution of FPKM gene expression levels. Genes that were not expressed in any sample were excluded from the final data matrix for downstream analysis.

Sanger sequencing

Sanger sequencing of the discovery cohort including *ASXL1* (p.Y591* and p.H630fs), *DNMT3A* (p.R882C), *EZH2* (p.V674L), *GATA2* (p.N317S), *RRAS* (p.Q87L), *RRAS2* (p.Q72L), *RUNX1* (p.R349fs), *SETBP1* (p.D868N, p.G870S and p.G870S), *SH2B3* (p.Q258*, p.W262*, p.F390fs, p.E400K and p.H414_splice) and *ZRSR2* (p.G179E) was performed using the primers listed in Supplementary Table 13. All mutations were tested in both germline and tumor tissue in order to determine germline versus somatic acquisition (Supplementary Figure 15). Sequencing was carried out using Hot Start Polymerase (Promega) and the following PCR conditions: 95°C for 2 minutes; 35 cycles of 94°C for 45 seconds, 60°C for 30 seconds, and 72°C for 40 seconds; and 1 cycle at 75°C for 5 minutes. Sequences were aligned using CLC software (CLC Bio, Aarhus, Denmark).

Single colony analysis

Mononuclear cells were isolated from fresh bone marrow or peripheral blood samples and resuspended in Iscove's Modified Dulbecco's Medium (IMDM) + 2% FBS. Cells were suspended at a concentration of 200,000 cells/ml. 154 μl of the cell suspension was added to a tube with the following: 1.2 ml Human Methylcellulose (R&D Systems Human Methylcellulose Complete Media, Cat. HSC003), 15 μl of 100X penicillin/streptomycin, diluted in water and IMDM to complete the tube volume to 1.54 ml. Cytokines included erythropoietin, granulocyte macrophage colony-stimulating factor (GM-CSF), interleukin 3 (IL-3), and stem cell factor (SCF). The solution was vortexed for 15 s and rested for 15 min. 1.1 ml was plated into a 35 × 10 mm Petri dish (BD Falcon, Cat. 351008) placed into a 150 × 15 mm Petri dish (BD Falcon, Cat. 351058) with another dish containing sterile water, and placed into an incubator at 37°C, 5% CO2. After 14 days, plates were removed from the incubator, and 2 μl of colonies (clusters of 50 cells or more) were plucked under a microscope at 40X magnification and resuspended in 20 μl of TE buffer. PCR amplification was then carried out for each of the mutations using 1 mL of the vortexed colony suspension and the primers listed in Supplementary Table 13.

Fibroblast cultures

Bone marrow or skin cells were cultured in Chang media (Irvine Scientific, #T105) which was changed every three days. Cells were split when the fibroblasts became 70% confluent and were harvested for DNA extraction using standard methods when the fibroblasts began to have a flattened morphologic appearance.

Immunoblotting

Whole cell-extracts were prepared from previously cryopreserved samples in NP-40 lysis buffer supplemented with protease inhibitors. Blots were blocked with 5% BSA in TBS with .1% Tween 20 for 1 hour at room temperature and probed with primary anti-LNK (H-129, Santa Cruz) and secondary HRP-conjugated antibody (GE Healthcare) for 1 hour each. Protein quantification for westerns was performed using Image Lab Software (v.5.2, Biorad).

Statistical analyses

Data from AAML0122 were current as of June 8, 2010. Data from the remaining patients were current as of October 1st 2014. The number of somatic alterations at diagnosis, categorized as 0-1 vs. 2 or more, was the predictor variable of interest in this analysis. Somatic alterations were defined as a gene containing a pathogenic mutation or monosomy 7. Clinical and biologic features were compared between these groups. The significance of observed differences in proportions was tested using the χ^2 test (for categorical variables) or *t* test (for continuous variables), and Fisher's exact test when data were sparse. The Mann-Whitney test was used to determine the differences in medians. The Kaplan-Meier method[79] was used to estimate probabilities of event-free survival (EFS) and overall survival (OS). EFS was defined as time from diagnosis to first occurrence of relapse, treatment related mortality, or death; OS was defined as time from diagnosis until death. EFS and OS survival distributions were compared according to number of somatic alterations at diagnosis using a two-sided log-rank test. Cox proportional hazards models[62] were used to estimate the hazard ratio (HR) for defined groups of patients in univariate and multivariate analyses of OS and EFS. Relapse incidence (RI) was defined as the probability of having a relapse before time *t;* death without experiencing a relapse was considered a competing event. Cumulative incidence was compared using the Gray's test[63]. P values less than 0.05 were considered statistically significant. All statistical analyses were performed using STATA, version (STATA, College Station, TX).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank the patients and their families for participating in this research, without which this work would not be possible. In addition, Drs. Alan Ikeda, Elizabeth Raetz, Nancy Bunin, and Jerry Finkelstein have all assisted in this research by providing invaluable patient samples.

This work was supported by Carlos Slim Foundation in Mexico as part of the Slim Initiative for Genomic Medicine; the St. Baldrick's Foundation (E.S.); Alex's Lemonade Stand Foundation (E.S.); the Leukemia and Lymphoma Society (grant numbers 6059-09 and 6466-15) (M.L.L.) and (K.S.)(LLS Scholar); the National Institutes of Health, National Cancer Institute grants T32CA128583 (E.S.), R01CA173085 (M.L.L.), the National Cancer Institute Cancer Center Support grant 5P30CA082103, COG Statistics and Data Center 1U10CA180899 (T.A.), COG Chair's Grant 5U10CA098543 (T.A.), NIH P30CA82103 (A.B.O.); the Frank A. Campini Foundation (E.S. and M.L.L.); Hyundai Hope on Wheels (M.L.L.); and a University of California San Francisco Dean's Commitment to the Center for Advanced Technologies facility.

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; 124:2487–97. [PubMed: 25163700]
2. Dvorak CC, Loh ML. Juvenile Myelomonocytic Leukemia: Molecular Pathogenesis Informs Current Approaches to Therapy and Hematopoietic Cell Transplantation. Front Pediatr. 2014; 2:25. [PubMed: 24734223]
3. Matsuda K, et al. Spontaneous improvement of hematologic abnormalities in patients having juvenile myelomonocytic leukemia with specific RAS mutations. Blood. 2007; 109:5477–80. [PubMed: 17332249]
4. Matsuda K, et al. Long-term survival after nonintensive chemotherapy in some juvenile myelomonocytic leukemia patients with CBL mutations, and the possible presence of healthy persons with the mutations. Blood. 2010; 115:5429–31. [PubMed: 20595524]
5. Bresolin S, et al. Gene expression-based classification as an independent predictor of clinical outcome in juvenile myelomonocytic leukemia. J Clin Oncol. 2010; 28:1919–27. [PubMed: 20231685]
6. Flotho C, et al. Genotype-phenotype correlation in cases of juvenile myelomonocytic leukemia with clonal RAS mutations. Blood. 2008; 111:966–7. author reply 967-8. [PubMed: 18182584]
7. Side L, et al. Homozygous inactivation of the NF1 gene in bone marrow cells from children with neurofibromatosis type 1 and malignant myeloid disorders. N Engl J Med. 1997; 336:1713–20. [PubMed: 9180088]
8. Flotho C, et al. RAS mutations and clonality analysis in children with juvenile myelomonocytic leukemia (JMML). Leukemia. 1999; 13:32–7. [PubMed: 10049057]
9. Tartaglia M, et al. Somatic mutations in PTPN11 in juvenile myelomonocytic leukemia, myelodysplastic syndromes and acute myeloid leukemia. Nat Genet. 2003; 34:148–50. [PubMed: 12717436]
10. Loh ML, et al. Mutations in PTPN11 implicate the SHP-2 phosphatase in leukemogenesis. Blood. 2004; 103:2325–31. [PubMed: 14644997]
11. Loh ML, et al. Mutations in CBL occur frequently in juvenile myelomonocytic leukemia. Blood. 2009; 114:1859–63. [PubMed: 19571318]
12. Sakaguchi H, et al. Exome sequencing identifies secondary mutations of SETBP1 and JAK3 in juvenile myelomonocytic leukemia. Nat Genet. 2013; 45:937–41. [PubMed: 23832011]
13. Stieglitz E, et al. Subclonal mutations in SETBP1 confer a poor prognosis in juvenile myelomonocytic leukemia. Blood. 2015; 125:516–24. [PubMed: 25395418]
14. Niemeyer CM, et al. Germline CBL mutations cause developmental abnormalities and predispose to juvenile myelomonocytic leukemia. Nat Genet. 2010; 42:794–800. [PubMed: 20694012]
15. Stephens K, et al. Interstitial uniparental isodisomy at clustered breakpoint intervals is a frequent mechanism of NF1 inactivation in myeloid malignancies. Blood. 2006; 108:1684–9. [PubMed: 16690971]
16. Flex E, et al. Activating mutations in RRAS underlie a phenotype within the RASopathy spectrum and contribute to leukaemogenesis. Hum Mol Genet. 2014; 23:4315–27. [PubMed: 24705357]
17. Graham SM, et al. TC21 and Ras share indistinguishable transforming and differentiating activities. Oncogene. 1999; 18:2107–16. [PubMed: 10321735]
18. Patmore DM, et al. In vivo regulation of TGF-beta by R-Ras2 revealed through loss of the RasGAP protein NF1. Cancer Res. 2012; 72:5317–27. [PubMed: 22918885]
19. Chan AM, Miki T, Meyers KA, Aaronson SA. A human oncogene of the RAS superfamily unmasked by expression cDNA cloning. Proc Natl Acad Sci U S A. 1994; 91:7558–62. [PubMed: 8052619]
20. Locatelli F, Niemeyer CM. How I treat juvenile myelomonocytic leukemia (JMML). Blood. 2015
21. Tinschert, S., et al. Segmental neurofibromatosis is caused by somatic mutation of the neurofibromatosis type 1 (NF1) gene. 2000.

22. Side LE, et al. Mutations of the NF1 gene in children with juvenile myelomonocytic leukemia without clinical evidence of neurofibromatosis, type 1. Blood. 1998; 92:267–272. [PubMed: 9639526]
23. Kotecha N, et al. Single-cell profiling identifies aberrant STAT5 activation in myeloid malignancies with specific clinical and biologic correlates. Cancer Cell. 2008; 14:335–43. [PubMed: 18835035]
24. Rudd CE. Lnk adaptor: novel negative regulator of B cell lymphopoiesis. Sci STKE 2001. 2001:e1.
25. Oh ST, et al. Novel mutations in the inhibitory adaptor protein LNK drive JAK-STAT signaling in patients with myeloproliferative neoplasms. Blood. 2010; 116:988–92. [PubMed: 20404132]
26. Lasho TL, Pardanani A, Tefferi A. LNK mutations in JAK2 mutation-negative erythrocytosis. N Engl J Med. 2010; 363:1189–90. [PubMed: 20843259]
27. Hurtado C, et al. LNK can also be mutated outside PH and SH2 domains in myeloproliferative neoplasms with and without V617FJAK2 mutation. Leuk Res. 2011; 35:1537–9. [PubMed: 21794913]
28. Perez-Garcia A, et al. Genetic loss of SH2B3 in acute lymphoblastic leukemia. Blood. 2013; 122:2425–32. [PubMed: 23908464]
29. McMullin MF, Wu C, Percy MJ, Tong W. A nonsynonymous LNK polymorphism associated with idiopathic erythrocytosis. Am J Hematol. 2011; 86:962–4. [PubMed: 21990094]
30. Gambacorti-Passerini CB, et al. Recurrent ETNK1 mutations in atypical chronic myeloid leukemia. Blood. 2015; 125:499–503. [PubMed: 25343957]
31. Yang Z, et al. Increased c-Jun expression and reduced GATA2 expression promote aberrant monocytic differentiation induced by activating PTPN11 mutants. Mol Cell Biol. 2009; 29:4376–93. [PubMed: 19528235]
32. Pasquet M, et al. High frequency of GATA2 mutations in patients with mild chronic neutropenia evolving to MonoMac syndrome, myelodysplasia, and acute myeloid leukemia. Blood. 2013; 121:822–9. [PubMed: 23223431]
33. Kazenwadel J, et al. Loss-of-function germline GATA2 mutations in patients with MDS/AML or MonoMAC syndrome and primary lymphedema reveal a key role for GATA2 in the lymphatic vasculature. Blood. 2012; 119:1283–91. [PubMed: 22147895]
34. Hsu AP, et al. Mutations in GATA2 are associated with the autosomal dominant and sporadic monocytopenia and mycobacterial infection (MonoMAC) syndrome. Blood. 2011; 118:2653–5. [PubMed: 21670465]
35. Hahn CN, et al. Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. Nat Genet. 2011; 43:1012–7. [PubMed: 21892162]
36. Ostergaard P, et al. Mutations in GATA2 cause primary lymphedema associated with a predisposition to acute myeloid leukemia (Emberger syndrome). Nat Genet. 2011; 43:929–31. [PubMed: 21892158]
37. Stieglitz E, et al. Mutations in GATA2 are rare in juvenile myelomonocytic leukemia. Blood. 2014; 123:1426–7. [PubMed: 24578498]
38. Fasan A, et al. GATA2 mutations are frequent in intermediate-risk karyotype AML with biallelic CEBPA mutations and are associated with favorable prognosis. Leukemia. 2013; 27:482–5. [PubMed: 22814295]
39. Papaemmanuil E, et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood. 2013; 122:3616–27. quiz 3699. [PubMed: 24030381]
40. Bauer DE, Loh ML, Bhagat G, Cantor AB, Kung AL. Potential Role Of RUNX1 In The Pathogenesis Of Juvenile Myelomonocytic Leukemia (JMML). Blood. 2013; 122:45–45.
41. Osato M. Point mutations in the RUNX1/AML1 gene: another actor in RUNX leukemia. Oncogene. 2004; 23:4284–96. [PubMed: 15156185]
42. Sano H, et al. Wiskott-Aldrich syndrome with unusual clinical features similar to juvenile myelomonocytic leukemia. Int J Hematol. 2012; 96:279–83. [PubMed: 22736231]
43. Yoshimi A, et al. Wiskott-Aldrich syndrome presenting with a clinical picture mimicking juvenile myelomonocytic leukaemia. Pediatr Blood Cancer. 2013; 60:836–41. [PubMed: 23023736]

44. Albert MH, et al. X-linked thrombocytopenia (XLT) due to WAS mutations: clinical characteristics, long-term outcome, and treatment options. Blood. 2010; 115:3231–8. [PubMed: 20173115]
45. Perez B, et al. Genetic typing of CBL, ASXL1, RUNX1, TET2 and JAK2 in juvenile myelomonocytic leukaemia reveals a genetic profile distinct from chronic myelomonocytic leukaemia. Br J Haematol. 2010; 151:460–8. [PubMed: 20955399]
46. Sugimoto Y, et al. Spectrum of molecular defects in juvenile myelomonocytic leukaemia includes ASXL1 mutations. Br J Haematol. 2010; 150:83–7. [PubMed: 20408841]
47. Kar SA, et al. Spliceosomal gene mutations are frequent events in the diverse mutational spectrum of chronic myelomonocytic leukemia but largely absent in juvenile myelomonocytic leukemia. Haematologica. 2013; 98:107–13. [PubMed: 22773603]
48. Ribeiro AF, et al. Mutant DNMT3A: a marker of poor prognosis in acute myeloid leukemia. Blood. 2012; 119:5824–31. [PubMed: 22490330]
49. Ley TJ, et al. DNMT3A mutations in acute myeloid leukemia. N Engl J Med. 2010; 363:2424–33. [PubMed: 21067377]
50. Madan V, et al. Aberrant splicing of U12-type introns is the hallmark of ZRSR2 mutant myelodysplastic syndrome. Nat Commun. 2015; 6:6042. [PubMed: 25586593]
51. Hirabayashi S, et al. Spliceosomal gene aberrations are rare, coexist with oncogenic mutations, and are unlikely to exert a driver effect in childhood MDS and JMML. Blood. 2012; 119:e96–9. [PubMed: 22238327]
52. Takita J, et al. Novel splicing-factor mutations in juvenile myelomonocytic leukemia. Leukemia. 2012; 26:1879–81. [PubMed: 22343734]
53. Jankowska AM, et al. Mutational spectrum analysis of chronic myelomonocytic leukemia includes genes associated with epigenetic regulation: UTX, EZH2, and DNMT3A. Blood. 2011; 118:3932–41. [PubMed: 21828135]
54. Furlan I, et al. Intriguing response to azacitidine in a patient with juvenile myelomonocytic leukemia and monosomy 7. Blood. 2009; 113:2867–8. [PubMed: 19299654]
55. Russler-Germain DA, et al. The R882H DNMT3A mutation associated with AML dominantly inhibits wild-type DNMT3A by blocking its ability to form active tetramers. Cancer Cell. 2014; 25:442–54. [PubMed: 24656771]
56. Archambeault S, et al. Development of an allele-specific minimal residual disease assay for patients with juvenile myelomonocytic leukemia. Blood. 2008; 111:1124–1127. [PubMed: 18000165]
57. Kothari A, Hulbert ML, Cottrell CE, Nguyen TT. Aggressive congenital juvenile myelomonocytic leukemia associated with somatic KRAS p.G13D mutation and concurrent germline IGF1R duplication. Leuk Lymphoma. 2014:1–4.
58. Lee ML, et al. Juvenile Myelomonocytic Leukemia in a Premature Neonate Mimicking Neonatal Sepsis. Pediatr Neonatol. 2013:S1875–9572(13).
59. Greaves M, Maley CC. Clonal evolution in cancer. Nature. 2012; 481:306–13. [PubMed: 22258609]
60. Chang T, et al. Sustained MEK inhibition abrogates myeloproliferative disease in Nf1 mutant mice. J Clin Invest. 2013; 123:335–9. [PubMed: 23221337]
61. Lyubynska N, et al. A MEK inhibitor abrogates myeloproliferative disease in Kras mutant mice. Sci Transl Med. 2011; 3:76ra27.
62. Chan RJ, Cooper T, Kratz CP, Weiss B, Loh ML. Juvenile myelomonocytic leukemia: a report from the 2nd International JMML Symposium. Leuk Res. 2009; 33:355–62. [PubMed: 18954903]
63. Stieglitz E, et al. Phase II/III trial of a pre-transplant farnesyl transferase inhibitor in juvenile myelomonocytic leukemia: A report from the Children's Oncology Group. Pediatric Blood & Cancer. 2015; 62:629–636. [PubMed: 25704135]
64. Cibulskis K, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol. 2013; 31:213–9. [PubMed: 23396013]
65. DePristo MA, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011; 43:491–8. [PubMed: 21478889]

66. McKenna A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010; 20:1297–303. [PubMed: 20644199]
67. Chapman MA, et al. Initial genome sequencing and analysis of multiple myeloma. Nature. 2011; 471:467–72. [PubMed: 21430775]
68. Venkatraman ES, Olshen AB. A faster circular binary segmentation algorithm for the analysis of array CGH data. Bioinformatics. 2007; 23:657–63. [PubMed: 17234643]
69. Exome Aggregation Consortium (ExAC).
70. Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic Acids Res. 2011; 39:e118. [PubMed: 21727090]
71. Forbes SA, et al. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic Acids Res. 2015; 43:D805–11. [PubMed: 25355519]
72. Triche TJ Jr. Weisenberger DJ, Van Den Berg D, Laird PW, Siegmund KD. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. Nucleic Acids Res. 2013; 41:e90. [PubMed: 23476028]
73. Cancer Genome Atlas Research, N. Comprehensive genomic characterization of squamous cell lung cancers. Nature. 2012; 489:519–25. [PubMed: 22960745]
74. Dulak AM, et al. Exome and whole-genome sequencing of esophageal adenocarcinoma identifies recurrent driver events and mutational complexity. Nat Genet. 2013; 45:478–86. [PubMed: 23525077]
75. Trapnell C, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 2010; 28:511–5. [PubMed: 20436464]
76. Trapnell C, et al. Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotechnol. 2013; 31:46–53. [PubMed: 23222703]
77. Kaplan EL, Meier P. Nonparametric-Estimation from Incomplete Observations. Journal of the American Statistical Association. 1958; 53:457–481.
78. Cox DR. Regression Models and Life-Tables. J. Roy. Statist. Soc. Ser. B. 1972; 34:187–220.
79. Gray RJ. A Class of K-Sample Tests for Comparing the Cumulative Incidence of a Competing Risk. Annals of Statistics. 1988; 16:1141–1154.

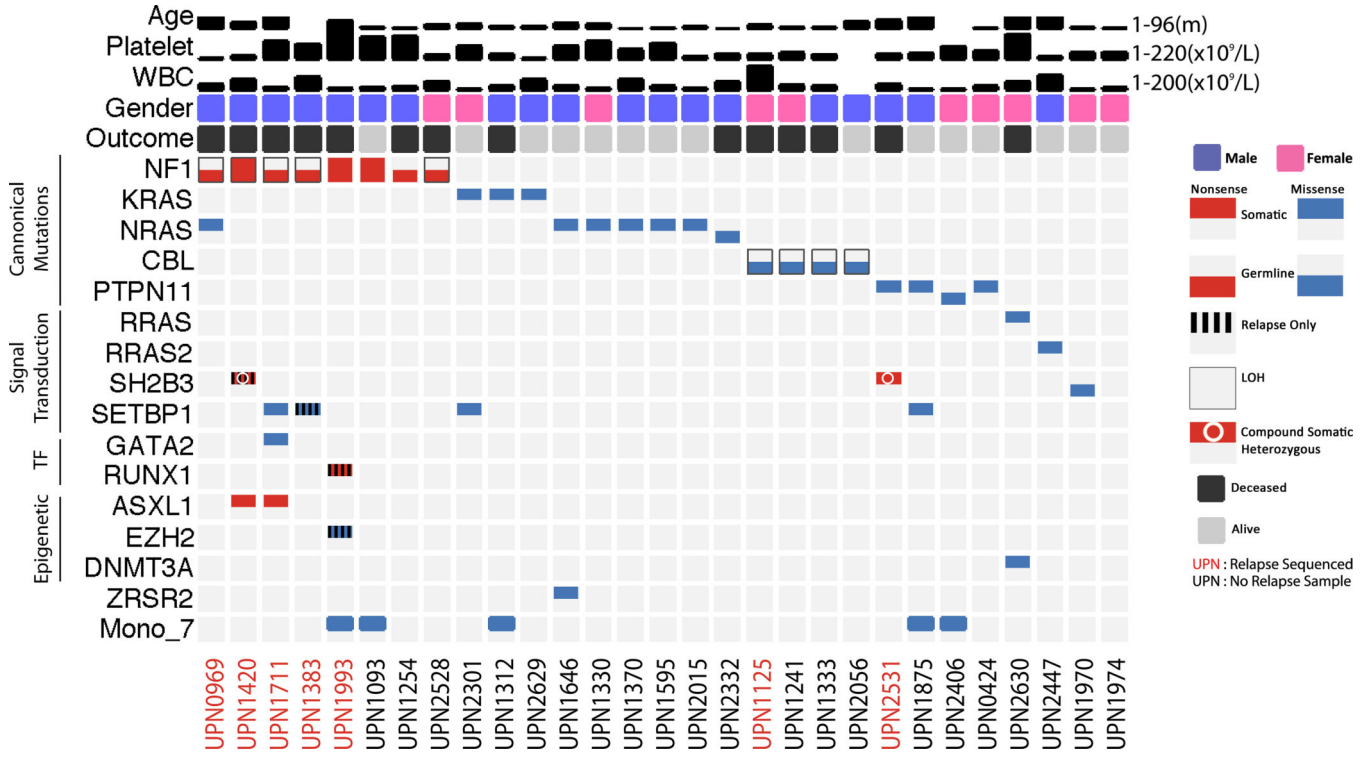

Figure 1.

Mutations identified by exome sequencing. Twenty-nine patients who underwent whole exome sequencing are displayed. Each patient is presented in a single condensed column including mutations identified at germline, diagnostic (noted in black) and relapse (noted in red) timepoints. Germline mutations are presented in colors in the bottom half of the box of any given gene and somatic mutations in the top half. Mutations only present at relapse are denoted with vertical striped bars. Loss of heterozygosity in a single gene is annotated with a thin black rectangle surrounding the mutation. Somatic compound heterozygous mutations are noted with a white circle.

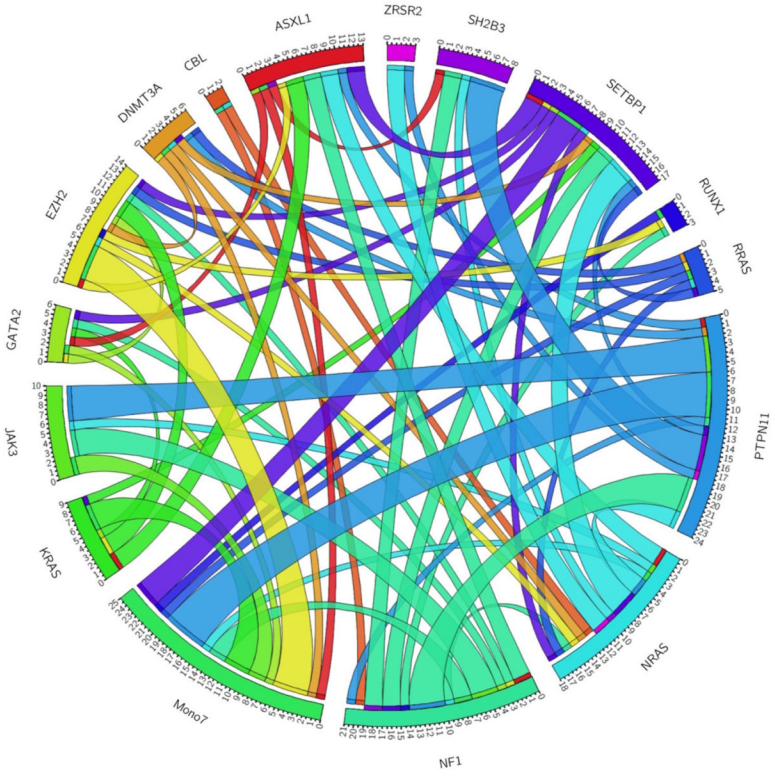

Figure 2.

Circos plot of samples with at least two mutations. Using data from whole exome and targeted resequencing, patients with at least two mutations are depicted. Associations between genomic alterations in the same patient are marked by connecting bands, with the width of the band proportional to the frequency of the association.

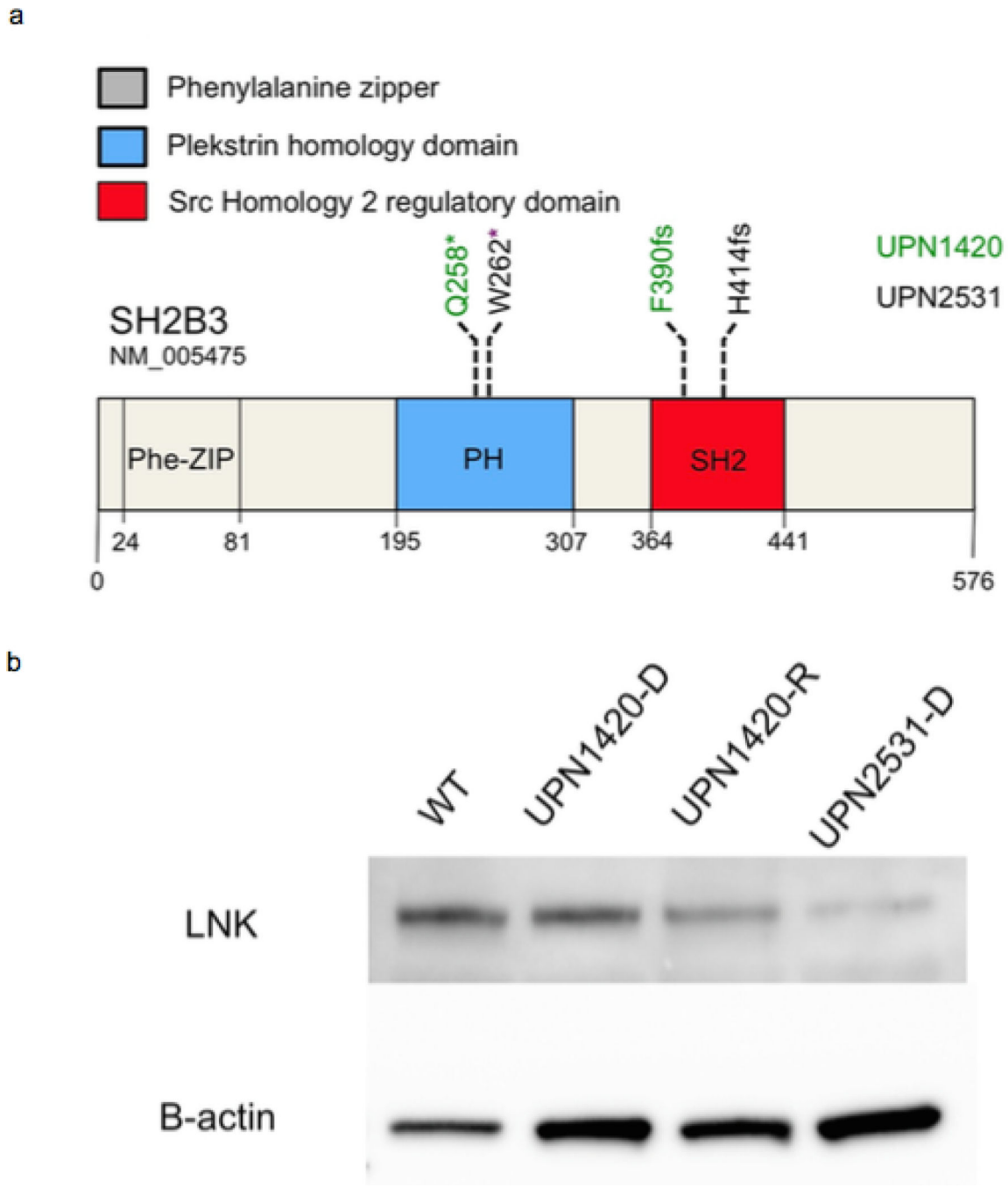

Figure 3.

Mutations in *SH2B3* decrease expression of LNK. **(a)** Compound mutations in the Plekstrin and Src Homology 2 domains are presented for each patient found to harbor *SH2B3* lesions on whole exome sequencing. **(b)** Western blot analysis from whole cell lysates using anti-LNK and anti-B-actin antibodies. Commensurate with the allelic fraction of each mutation (UPN1420-Relapse, 31%, UPN2531-Diagnosis, 37%), the expression of LNK is decreased.

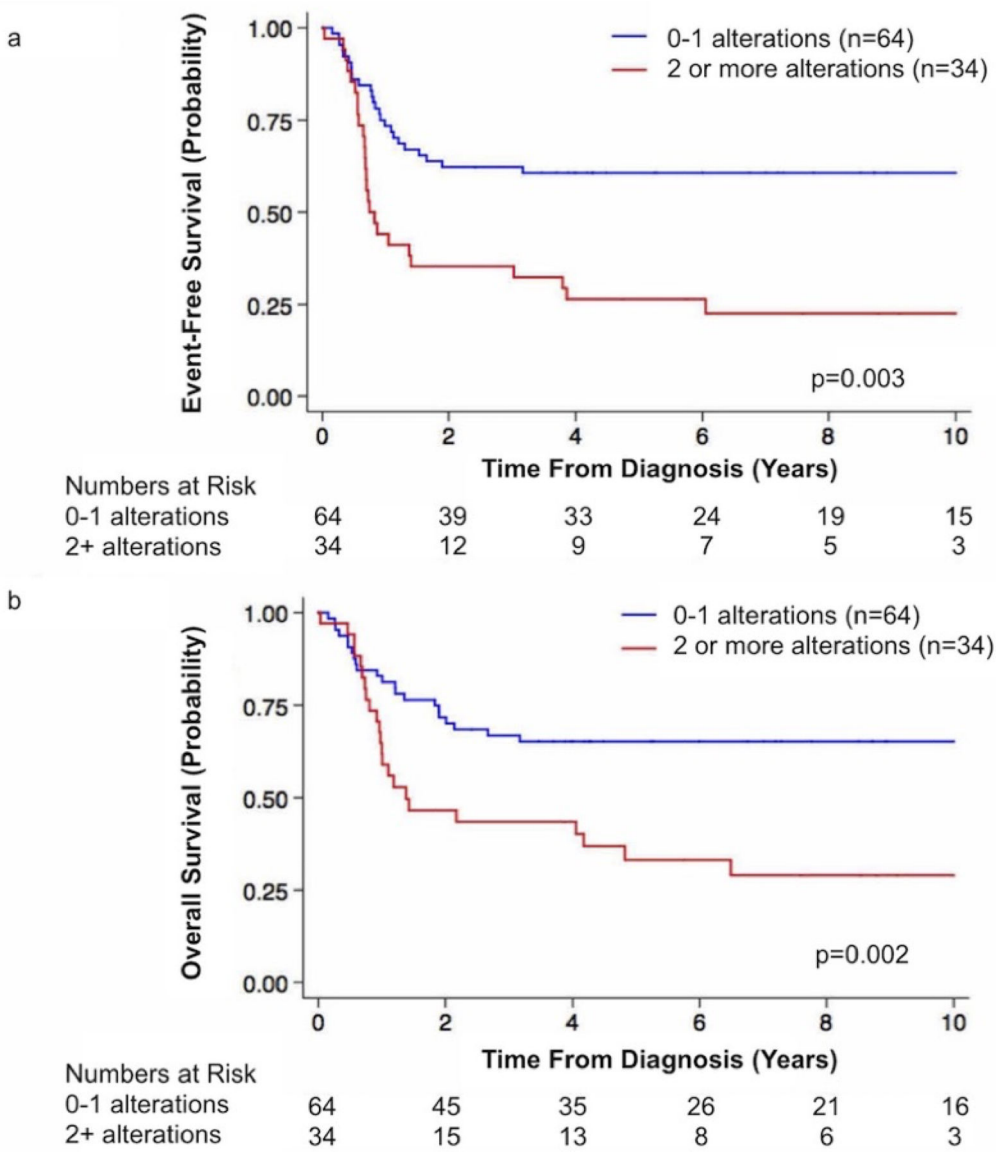

Figure 4.

Event-free and overall survival of patients stratified by the number of somatic alterations. Kaplan-Meier estimated **(a)** event-free survival (log-rank p=0.002) and **(b)** overall survival (log-rank p=0.002) according to the number of somatic alterations at diagnosis.

Table 1

Mutations detected on WES excluding known Ras pathway genes.

Gene	Mutation type	RefSeq	Amino acid change	Nucleotide change	UPN
ASXL1	Nonsense	NM_015338	p.Y591*	c.1773C>G	UPN1420
ASXL1	Frameshift	NM_015338	p.H630fs	c.1888_1910del	UPN1711
DNMT3A	Missense	NM_022552	p.R882C	c.2644C>T	UPN2630
EZH2	Missense	NM_001203247	p.V674L	c.2020G>C	UPN1993
GATA2	Missense	NM_001145661	p.N317S	c.950A>G	UPN1711
RRAS	Missense	NM_006270	p.Q87L	c.260A>T	UPN2630
RRAS2	Missense	NM_012250	p.Q72L	c.215A>T	UPN2447
RUNX1	Frameshift	NM_001001890	p.R349fs	c.1047_1048insC	UPN1993
SETBP1	Missense	NM_015559	p.G870S	c.2608G>A	UPN1711
SETBP1	Missense	NM_015559	p.D868N	c.2602G>A	UPN2301
SETBP1	Missense	NM_015559	p.I871T	c.2612T>C	UPN1875
SETBP1	Missense	NM_015559	p.G870S	c.2608G>A	UPN1383
SH2B3	Frameshift	NM_005475	p.F390fs	c.1170delC	UPN1420
SH2B3	Nonsense	NM_005475	p.Q258*	c.772C>T	UPN1420
SH2B3	Nonsense	NM_005475	p.W262*	c.785G>A	UPN2531
SH2B3	Splice site	NM_005475	p.H414_splice	c.1240_splice	UPN2531
SH2B3	Missense	NM_005475	p.E400K	c.1198G>A	UPN1970
ZRSR2	Missense	NM_005089	p.G179E	c.536G>A	UPN1646

Table 2

Multivariate analysis of prognostic variables in JMML.

Multivariate Cox Analysis	EFS from date of diagnosis			OS from date of diagnosis		
Somatic Alterations at Diagnosis	**HR**	**95% CI**	**p**	**HR**	**95% CI**	**p**
0-1	1			1		
2 or more	2.65	1.27-5.54	**0.01**	3.13	1.39-7.01	**0.006**
Age at diagnosis (months)						
<24	1			1		
>24	1.62	0.82-3.21	0.167	1.55	0.73-3.28	0.25
Platelet count at diagnosis $\times 10^9$						
≥40	1			1		
<40	1.64	0.86-3.08	0.128	2.19	1.09-4.39	**0.028**
Fetal hemoglobin at diagnosis						
Not elevated for age	1			1		
Elevated for age	1.56	0.81 -3.02	0.184	0.93	0.69-2.80	0.355
NF1 status						
No	1			1		
Yes	0.85	0.38 1.93	0.7	0.5	0.19-1.29	0.15