HemaSphere
Powered by EHA

EUROPEAN HEMATOLOGY ASSOCIATION

Original Article
OPEN ACCESS

Integrated Genomic, Functional, and Prognostic Characterization of Atypical Chronic Myeloid Leukemia

Diletta Fontana[1], Daniele Ramazzotti[1], Andrea Aroldi[1,2], Sara Redaelli[1], Vera Magistroni[1], Alessandra Pirola[3], Antonio Niro[1], Luca Massimino[1], Cristina Mastini[1], Virginia Brambilla[4], Silvia Bombelli[1], Silvia Bungaro[5], Alessandro Morotti[6], Delphine Rea[7], Fabio Stagno[8], Bruno Martino[9], Leonardo Campiotti[10], Giovanni Caocci[11], Emilio Usala[12], Michele Merli[13], Francesco Onida[14], Marco Bregni[15], Elena Maria Elli[2], Monica Fumagalli[2], Fabio Ciceri[16], Roberto A. Perego[1], Fabio Pagni[4], Luca Mologni[1], Rocco Piazza[1,2], Carlo Gambacorti-Passerini[1,2]

Correspondence: Diletta Fontana (e-mail: diletta.fontana@unimib.it).

Abstract

Atypical chronic myeloid leukemia (aCML) is a *BCR-ABL1*-negative clonal disorder, which belongs to the myelodysplastic/myeloproliferative group. This disease is characterized by recurrent somatic mutations in *SETBP1*, *ASXL1* and *ETNK1* genes, as well as high genetic heterogeneity, thus posing a great therapeutic challenge. To provide a comprehensive genomic characterization of aCML we applied a high-throughput sequencing strategy to 43 aCML samples, including both whole-exome and RNA-sequencing data. Our dataset identifies *ASXL1*, *SETBP1*, and *ETNK1* as the most frequently mutated genes with a total of 43.2%, 29.7 and 16.2%, respectively. We characterized the clonal architecture of 7 aCML patients by means of colony assays and targeted resequencing. The results indicate that *ETNK1* variants occur early in the clonal evolution history of aCML, while *SETBP1* mutations often represent a late event. The presence of actionable mutations conferred both ex vivo and in vivo sensitivity to specific inhibitors with evidence of strong in vitro synergism in case of multiple targeting. In one patient, a clinical response was obtained. Stratification based on RNA-sequencing identified two different populations in terms of overall survival, and differential gene expression analysis identified 38 significantly overexpressed genes in the worse outcome group. Three genes correctly classified patients for overall survival.

Introduction

Atypical chronic myeloid leukemia (aCML) is a rare *BCR-ABL1*-negative clonal disorder belonging to the myelodysplastic/myeloproliferative group.[1] Its incidence is 1% to 2% of t(9;22) *BCR-ABL1*-positive CML.[2–5] This disorder affects elderly patients with a median age ranging between 60 and 76 years, with an apparent male predominance.[6,7] Patients present clinical features in common with *BCR-ABL1*-positive CML including splenomegaly, elevated white blood cells (WBC) count with a predominance of granulocytes and immature myeloid cells, and

[1]Department of Medicine and Surgery, University of Milano - Bicocca, Monza, Italy
[2]Hematology and Clinical Research Unit, San Gerardo Hospital, Monza, Italy
[3]GalSeq s.r.l., via Ludovico Ariosto, 21, 20091 Bresso (MI), Italy
[4]Department of Medicine and Surgery, Pathology, University of Milano - Bicocca, San Gerardo Hospital, Monza, Italy
[5]Centro Ricerca Tettamanti, Pediatria, University of Milano - Bicocca, Monza, Italy
[6]Department of Clinical and Biological Sciences, San Luigi Hospital, University of Turin, Turin, Italy
[7]Service d'Hématologie adulte, Hôpital Saint-Louis, Paris, France
[8]Division of Hematology and Bone Marrow Transplant, A.O.U. Policlinico - Vittorio Emanuele, Catania, Italy
[9]Division of Hematology, Azienda Ospedaliera 'Bianchi Melacrino Morelli', Reggio Calabria, Italy
[10]Department of Medicine and Surgery, Università degli Studi dell'Insubria, Varese, Italy
[11]Hematology Unit, Department of Medical Sciences and Public Health, University of Cagliari, Cagliari, Italy
[12]Hematology Unit, Ospedale Oncologico A. Businco, Cagliari, Italy
[13]Hematology, University Hospital Ospedale di Circolo e Fondazione Macchi, Varese, Italy
[14]Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, University of Milan, Milan, Italy
[15]Oncology-Hematology Unit, ASST Valle Olona, Busto Arsizio, Italy
[16]Unit of Hematology and Bone Marrow Transplantation, IRCCS San Raffaele Scientific Institute, Vita-Salute San Raffaele University, Milan, Italy
RP and CG-P are joint senior authors.
The authors declare no competing interests.
Supplemental Digital Content is available for this article.

HemaSphere (2020) 4:6(e497). http://dx.doi.org/10.1097/HS9.0000000000000497.
Received: 28 July 2020 / Accepted: 29 September 2020

moderate anemia. According to the 2016 revision of the WHO classification for myeloid neoplasms, the median overall survival for aCML patients is 24 months. Moreover, no established standards of care exist for its treatment.[8] Until 2012, the molecular lesions responsible for the onset of this leukemia remained unknown. The use of Next Generation Sequencing techniques (NGS) allowed our and other groups to identify recurrent somatic mutations occurring in *SETBP1* and *ETNK1* genes,[6,9,10] later confirmed by several other independent studies.[11–18] The application of NGS technologies demonstrated the presence of several other mutations involving *ASXL1*, *CBL*, *EZH2*, *NRAS*, *TET2*, *CSFR3R*, and *U2AF1* genes. The identification of somatic variants occurring in a large number of genes clearly indicates that the genetic basis of aCML is heterogeneous, in striking contrast with classical CML. This heterogeneity poses a great challenge to the dissection of the molecular steps required for aCML leukemogenesis. Here we report a comprehensive analysis including mutation profiling, gene expression analysis and clinical outcome in a cohort of 43 aCML patients. We experimentally validated actionable mutations and identified the clonal hierarchy of multiple mutated genes. For a patient carrying both *ETNK1* G245 V and *NRAS* G12D mutations, in addition to targeting them *ex vivo*, we established a patient-derived xenograft in order to test the activity of the MEK inhibitor trametinib in vivo. Finally, we found 3 differentially expressed genes, allowing the clustering of our patients' cohort into 2 groups based on their survival.

Materials and methods

Patients

Diagnosis of aCML was performed according to the World Health Organization 2016 classification. All patients provided written informed consent, which was approved by the institutional ethics committee. This study was conducted in accordance with the Declaration of Helsinki. Bone marrow (BM) samples were collected at diagnosis, and leukemic cells were obtained by separation on a Ficoll-Paque Plus gradient (GE Healthcare, Milan, Italy). Surface markers were evaluated by fluorescence-activated cell sorting (FACS) analysis, and myeloid cells (positive for CD33, CD13 or CD117 staining) made up $>80\%$ of the total cells.

Whole exome sequencing

Ten million cells were used for genomic DNA extraction by using PureLink Genomic DNA kit (Thermo Fisher, Milan, Italy) according to manufacturer's instructions. 1 μg of gDNA was used to generate exome libraries (Galseq, Monza, Italy). Mean exon coverage was 80×. To identify somatically acquired mutations we compared DNA from leukocytes and constitutive DNA extracted from lymphocytes or buccal swabs. Bioinformatic analysis was performed as already described in.[6]

MethoCult™ colonies assay and combined treatment

One million peripheral blood (PB) or 2×10^5 BM cells were seeded in methylcellulose-based medium Methocult H4034 (StemCell Technologies, Meda, Italy), according to manufacturer's instructions, and plated in 6-well dishes. After 2 weeks of incubation at 37°C, 5% CO_2, individual colonies were picked, washed in PBS, and lysed in 20 μL of the following buffer: 10 mM Tris-HCl, 50 mM NaCl, 6.25 mM $MgCl_2$, 0.045% NP40, 0.45% Tween-20; pH 7.6. On average 50 colonies per sample were isolated. After adding 1 μL of 20 μg/mL proteinase K, the lysate was incubated at 56°C for 1 hour and at 95°C for 15 minutes. Subsequently, the sample was amplified using dedicated barcoded primers by PCR, and underwent deep-sequencing. For combined treatment, 2×10^5 BM-derived cells were seeded in methylcellulose-based medium in presence of phosphoethanolamine 1 mM (Merck Life Science, Milan, Italy), trametinib 10 nM (Selleck Chemicals, Rome, Italy), trametinib 100 nM and combination of them. After 2 weeks of incubation, colonies were counted. Expected additive effect of the combination viability is the product of the 2 singlet viabilities. For actionable mutations targeting, all the inhibitors used (crizotinib, dasatinib, imatinib, and ruxolitinib) were purchased from Selleck Chemicals.

Deep-sequencing

Amplicon libraries were generated starting from 500 ng PCR product, purified on agarose gel, end-repaired and adenylated at 3' ends before ligation of Truseq DNA Adapter Indices, and then amplified with 6-cycles PCR. Libraries were sequenced on an Illumina HiSeq 2500 instrument with paired-end reads 150 bp long. Paired fastq were initially deindexed using a custom, home-made tool and subsequently aligned to the reference human genome (hg38) using BWA.[19] Coverage was $>2000\times$ for all samples. Bam alignment files were generated from SAM using Samtools.[20] Variant calls were performed using CEQer2.[21]

Clonal architecture analysis

Each methylcellulose colony was sequenced to reconstruct the clonal architecture of the corresponding sample. The indexed (barcoded) amplicons underwent NGS sequencing and were analyzed after deindexing using dedicated in-house bioinformatics tools. A generic mutation A was considered to be an earlier event compared to mutation B if A was identified in individual colonies in absence of mutation B.

aCML patient-derived xenograft (PDX) establishment

Ten 6 weeks-old NOD.Cg-Prkdc^{SCID}Il2rg^{tm1Wjl}/SzJ (NSG) mice were purchased from Charles River (Milan, Italy), kept under standard conditions following the guidelines of the University of Milano-Bicocca ethical committee for animal welfare, and treated in accordance with European Community guidelines as approved by the Italian Ministry of Health. The protocol was approved by the Italian Ministry of Health and by the Institutional Committee for Animal Welfare. Mice were sublethally irradiated with 200 cGy, and after 24 hours 10^7 CMLPh-042 ficoll-purified bone marrow cells were i.v. transplanted through tail vein injection. 80ug/ml gentamicin was added to drinking water to prevent infections. Mice body weight was evaluated three times a week for the whole experiment duration. Peripheral blood collection was carried out every 15 days after transplantation to check human bone marrow cells engraftment. Human CD45+ cells engraftment occurred after 45 days from transplantation, then mice were randomized in 2 groups to

receive vehicle alone (3 mice) or 1 mg/kg trametinib suspended in 0.5% carboxymethylcellulose/0.1% Tween80 by oral gavage once a day for 66 days (4 mice). Animals showing signs of morbidity (weight loss, hunched posture, unsteady gait, respiratory distress) were sacrificed before the end of the experiment. Sternum, spleen, lung, heart, liver, bowel, kidney, femur, tibia, and vertebrae were surgically extracted and paraffin fixed in buffer neutral formalin for immunohistochemistry experiments.

Analysis of human cells in peripheral blood by flow cytometry

Human cells engraftment was check every 15 days by flow cytometry. For the blood collection, mice were placed under a heat lamp to promote peripheral vasodilatation and were mechanically restrained using a plexiglass chamber. Then, a small transverse cut was performed in the lateral tail vein with a sterile lancet. Blood drops were collected with a micropipette and mixed with 10 µl EDTA 0.5 M. A maximum volume of 100 µl of blood was collected. After blood collection, a slight pressure was applied to ensure the bleeding stop. 100 µl EDTA-anticoagulated peripheral blood were lysed twice in red blood cells lysis buffer (140 mM NH_4Cl, 8 mM Tris, pH 7.2) at RT for 5 minutes, and then cells were washed and suspended in PBS. Subsequently, cells were stained with the Alexa Fluor® 700-conjugated anti-human CD45 (HI30; Biolegend, San Diego, CA, USA) and PE-conjugated anti-mouse CD45 (30-F11; Biolegend) antibodies, at RT for 30 minutes. Dual-color flow cytometry was performed on MoFlo Astrios cell sorter equipped with Summit 6.3 software (both from Beckman Coulter, Miami). The acquisition process was stopped when at least 5000 events were collected in the population gate. Off-line analysis was performed using Kaluza 1.3 software (Beckman Coulter). Engraftment occurred in 70% of mice.

Immunohistochemistry staining

Formalin fixed sections were processed using automated tissue processors and embedded in paraffin. Then we obtained 2 µm sections for haematoxylin and eosin staining and immunohistochemical staining with an anti-CD45 antibody (Anti-CD45 ab10559, Abcam). Immunohistochemistry was performed on a Dako Omnis platform (Glostrup, Denmark) with a dilution of 1:1500 of the antibody with an initial concentration of 1 mg/ml. Slides were digitally scanned using a ScanScope CS digital scanner (Aperio, Park Center Dr., Vista, CA, USA).

Cytogenetic analysis and western blot

For cytogenetic analysis and western blot details, please see Supplemental Digital Content, http://links.lww.com/HS/A105.

RNA-sequencing

Ten million cells were lysed in TRIzol (Thermo Fisher Scientific) and RNA was extracted according to manufacturer's instructions. 2 µg of RNA (concentration 400ng/µl) were used for library preparation (Galseq, Monza, Italy); the average per-sample read count was 35 M. For batch effect correction, please see Supplemental Digital Content, http://links.lww.com/HS/A105.

Patients' stratification

We analyzed batch corrected RNA-sequencing expression data to assess the presence of any clinically relevant subtype within our cohort. To do so, we first removed genes at low variance (variance < 0.01) and mitochondrial gene counts and normalized the remaining counts across patients. We then performed clustering analysis using CIMLR.[22] In short, this method constructs a set of multiple Gaussian Kernels from RNA-sequencing data and uses them to effectively reduce noise and separate patients presenting different profiles. In our cohort, CIMLR discovered two distinct subtypes (Supplementary Fig. 3, http://links.lww.com/HS/A105).

Features selection and pathway enrichment

To achieve an understanding of the genes that are differentiating the two discovered subtypes, we first consider a list of known cancer-related genes[23–26] (Supplementary Table 1, http://links.lww.com/HS/A106) and performed t-test for each gene in the list to assess whether a significant difference in expression was present (t-test p value adjusted for false discovery rate $p < 0.01$). We then selected genes among the significant ones that were differentially expressed in more than 75% of the patients of a cluster and conjunctively less than 25% of the patients in the other cluster to obtain a final list of 38 genes. This further filter aims at ensuring the biological relevance of the selected genes in terms of differential gene expression between the two clusters. We finally performed pathway enrichment considering these 38 genes by using the Max Plank Institute for Molecular Genetics ConsensusPathDB-human CPDB tool, setting the original list of known cancer-related genes as a background reference.

Patients' classification

We further reduced the list of 38 differentially expressed genes by considering only the top 3 genes. Briefly, we selected all the genes differentially expressed in more than 80% of the patients of a cluster and conjunctively less than 20% of the patients in the other cluster. We constructed a random forest classifier[27] using these genes. This method is used for classification and specifically performs the construction of multiple decision trees at training time and estimate the resulting class as the mode of the classes inferred by the individual trees. These analyses were implemented using the caret R package (version 6.0–84).

Next generation sequencing data

The NGS data discussed in this publication have been deposited in NCBI's Sequence Read Archive (SRA) and are accessible through accession number PRJNA60458.

Results

Clinical characteristics of patients

The clinical characteristics of aCML patients are summarized in Supplementary Table 2, http://links.lww.com/HS/A107. The median age at diagnosis was 65 years (range 38–85). The average blast percentage was 2.8% (range 0–17.0; SD 4.1) in peripheral blood and 3.4% (range 0–19.0; SD 3.8) in bone marrow. Splenomegaly was observed in 69.8% and bone marrow fibrosis (mostly MF1) in 37.2% of cases.

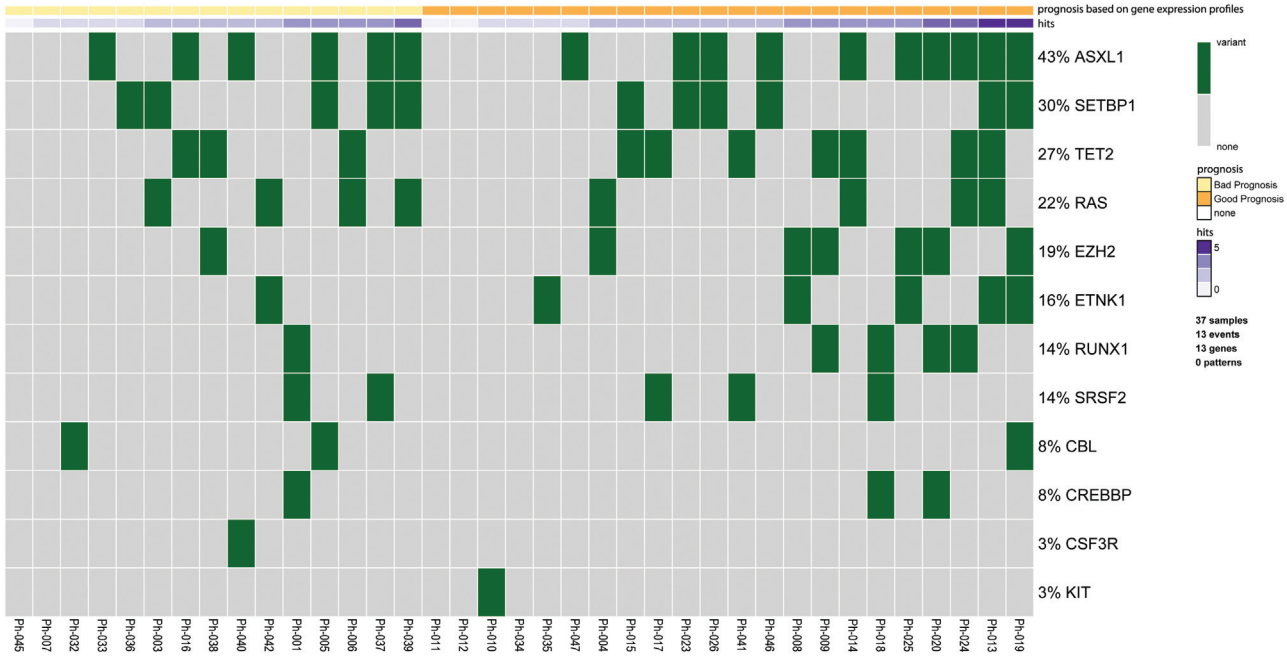

Figure 1. Oncoprint. Oncoprint showing somatic mutations for a panel of 12 genes in 37 patients.

Mutation profiling

Whole Exome Sequencing (WES) was performed on 37 aCML samples. All the somatic mutations identified were scored according to OncoScore.[28] Six patients revealed the presence of a single somatic mutation, 14 patients carried 2 mutations, 8 patients showed the co-presence of 3 mutations, while 4 or 5 mutations were present in 4 patients. Interestingly, 5 patients showed no mutations in known oncogenes, splicing factors, epigenetic factors, and cancer-related genes, and they carried a normal karyotype (Fig. 1; Supplementary Table 3, http://links.lww.com/HS/A108; Supplementary Table 4, http://links.lww.com/HS/A109); however, in all cases we identified somatic, likely passenger mutations, which supports the existence of clonal hematopoiesis also for these patients (Supplementary Tables 5, http://links.lww.com/HS/A110, 6, http://links.lww.com/HS/A111, 7, http://links.lww.com/HS/A112, 8, http://links.lww.com/HS/A113, and 9, http://links.lww.com/HS/A114). The most frequent mutation in our cohort was represented by *ASXL1* (43.2%). Other frequently mutated genes were *SETBP1* (29.7%), *TET2* (27.0%), and *KRAS/NRAS* (21.6%), confirming our previous results.[6] Mutations in *EZH2* occurred in 18.9% of patients, while *ETNK1* mutations were present in 16.2% of cases, again supporting our previous findings.[9] In 13.5% of patients *RUNX1* or *SRSF2* mutations were detected. Furthermore, 3 patients showed mutation of *CBL* or *CREBBP* (8.1% of cases), while *CSF3R* and *KIT* mutations occurred in one patient each.

Targeting of actionable mutations

In 27% of cases (10/37 patients) actionable mutations occurring in *KIT*, *NRAS*, *KRAS*, and *CSF3R* genes were found. All these mutated genes could be targeted with clinically available drugs, such as dasatinib, trametinib or ruxolitinib[29] (Supplementary Table 4, http://links.lww.com/HS/A109). In order to test whether these drugs were able to affect the growth of the leukemic clones in aCML cases, we performed colony assays in presence or absence of selected inhibitors on bone marrow-derived cells from four *RAS* mutated patients (CMLPh-003, CMLPh-039, CMLPh-006 and CMLPh-042, carrying *NRAS* G12R, *NRAS* G12D, *KRAS* A146 V, and *NRAS* G12D mutations, respectively), from a KIT D816 V positive patient (CMLPh-010), and from a *CSF3R* T618I mutated patient (CMLPh-040) (Figs. 2 and 4A, C). Patients with *RAS* mutations showed sensitivity to the MEK inhibitor trametinib, with 50% growth inhibition around 10 nM for *NRAS* mutated patients and around 50 nM for the patients with *KRAS* mutation (Figs. 2A–C and 4A, C), while unrelated inhibitors used as negative controls, such as imatinib, dasatinib and crizotinib, were inactive. CMLPh-010 patient presented a *KIT* D816 V mutation which is known to be highly sensitive to dasatinib, but resistant to imatinib.[30] Colony assays showed high sensitivity to dasatinib starting at 10 nM, while growth was almost completely abrogated at 0.1 μM; in contrast, imatinib was ineffective even at 3 μM (Fig. 2D). CMLPh-040 cells, bearing a *CSF3R* somatic mutation, were grown in presence of increasing concentrations of dasatinib, ruxolitinib, or crizotinib as negative control (Fig. 2E). Dasatinib completely inhibited cell growth already at 0.1 μM, while ruxolitinib showed a 50% inhibition at 0.1 μM and completely inhibited cell growth starting from 0.3 μM. Crizotinib did not affect colonies formation at any concentration. Taken together, these results indicate that the presence of *KIT*, *RAS*, and *CSF3R* mutations in aCML cells predicts sensitivity to clinically available inhibitors, at least ex vivo. Given the very poor prognosis of this disorder, these findings suggest a possible targeted treatment for a subset of aCML patients.

Hierarchical architecture of aCML patients

The temporal reconstruction of the different mutations occurring in a clonal disorder can have important biological, prognostic and therapeutic repercussions. For these reasons, in

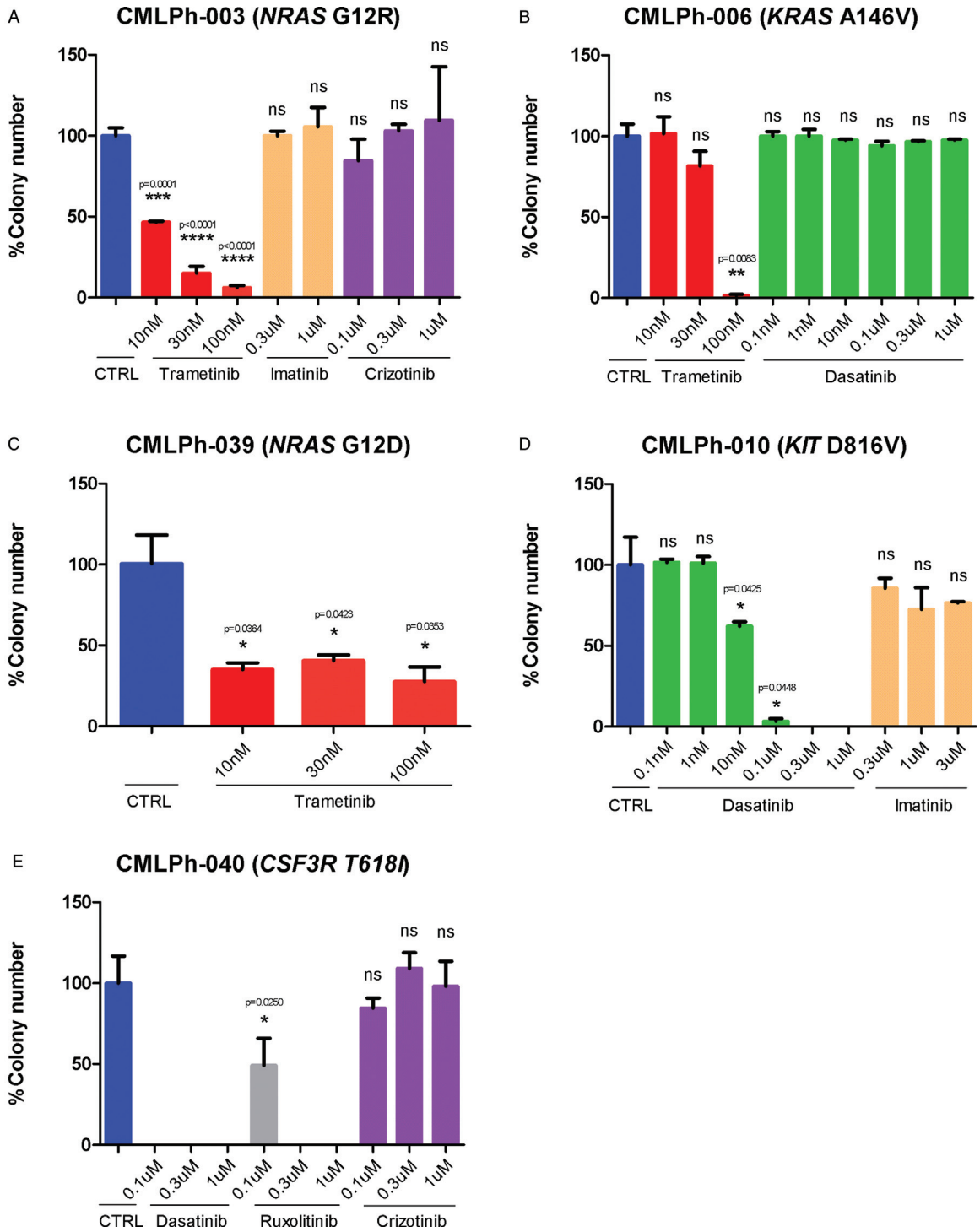

Figure 2. Colony formation assay. X axis represents different treatments and Y axis represents total number of colonies formed, normalized to 100 (colony counts in control conditions). Results are shown as the mean±s.d. (n=2). (A) Patient CMLPh-003 carrying *NRAS* G12R mutation. (B) Patient CMLPh-006 carrying *KRAS* A146V mutation. (C) Patient CMLPh-039 carrying *NRAS* G12R mutation. (D) Patient CMLPh-010 carrying *KIT* D816V mutation. (E) Patient CMLPh-040 carrying *CSF3R* T618I mutation.

patients whose bone marrow cells were available, we studied clonal evolution through the analysis of individual leukemic clones by methylcellulose assays (Fig. 3). The clonal architecture could be reconstructed in 7 patients. According to exome sequencing, patient CMLPh-003 was mutated in both *SETBP1* and *NRAS*. Clonal analysis confirmed the presence of *SETBP1* G870S in all the tested clones, while heterozygous *NRAS* G12R mutation was detected in 67% (Fig. 3A). Notably in the remaining 33% another heterozygous *NRAS* variant, G12D, was detected. Retrospective reanalysis of exome data confirmed the presence of the newly identified variant, which had been previously filtered-out from exome data because of the low

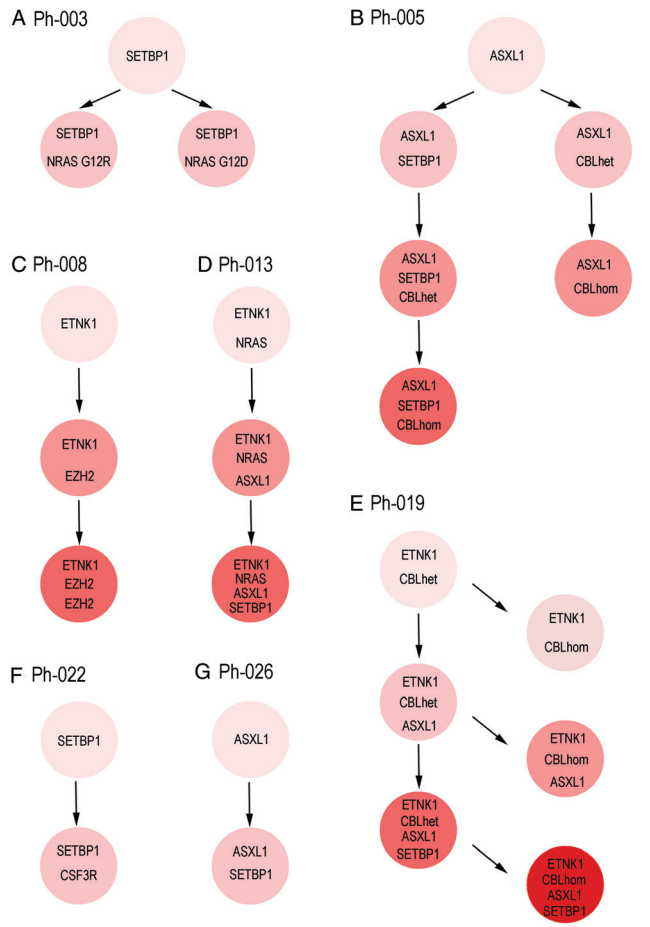

Figure 3. Clonal architecture of aCML patients. Schematic representation of the clonal architecture of 7 aCML patients whose bone marrow mononuclear cells were grown in semisolid medium and underwent targeted resequencing based on previously identified somatic mutations. CBLhet indicates a heterozygous somatic mutation; CBLhom indicates a homozygous somatic *CBL* mutation.

frequency. Patient CMLPh-005 carried mutations in *ASXL1*, *CBL* and *SETBP1* genes. Targeted analysis performed on 68 clones revealed a complex, branching evolution, with 63 clones carrying all the 3 variants. Of them, 47 (74.6%) had a heterozygous and 16 (25.4%) a homozygous *CBL* variant. Four clones (4.2%) carried *ASXL1* and *SETBP1* but not *CBL* mutations, while only a single clone was mutated in *ASXL1* and *CBL* in absence of *SETBP1* mutations, indicating that *CBL* mutations occurred independently in two different subclones (Fig. 3B). Allelic imbalance analysis of exome data using CEQer[21] revealed that *CBL* homozygosity was caused by a telomeric somatic uniparental disomy. Patient CMLPh-008 was mutated in *ETNK1* and *EZH2*, carrying both *EZH2* D608G and R634H variants. All the 27 clones analyzed carried *ETNK1* mutation, while one was WT for both *EZH2* D608G and R634H, indicating that both mutations occurred late in the clonal evolution of this patient (Fig. 3C). Patient CMLPh-013 carried *ASXL1*, *ETNK1*, *NRAS* and *SETBP1* mutations. Of the 39 clones analyzed, 34 (82.9%) showed the coexistence of all mutations, 4 were mutated in *ASXL1*, *ETNK1* and *NRAS* and 1 in *ETNK1* and *NRAS*, suggesting that *ETNK1* and *NRAS* were early events, *ASXL1* an intermediate one and *SETBP1* a late occurring mutation (Fig. 3D). Patient CMLPh-019 was characterized by the presence of a complex mutational status, with mutations occurring in *SETBP1*, *ETNK1*, *ASXL1* and *CBL* genes. Targeted resequencing analysis revealed the presence of all the 4 variants in 44/60 (73.3%) clones; in 15/60 (25%) we detected the presence of mutated *ETNK1*, *ASXL1* and *CBL*. Of these 15 clones, 33% carried heterozygous and 67% homozygous *CBL* mutations. In one clone (1.7%) we detected heterozygous *ETNK1* and homozygous *CBL*, indicating a strong selective pressure towards the acquisition of homozygous *CBL* mutations (Fig. 3E). Allelic imbalance analysis of CMLPh-019 exome revealed that *CBL* homozygosity was caused by a somatic uniparental disomy event occurring in the telomeric region of the long arm of chromosome 11 (Supplementary Fig. 4, http://links.lww.com/HS/A105). Patient CMLPh-022 was mutated in both *SETBP1* and *CSF3R*. Clonal analysis found the presence of *SETBP1* in all the 40 clones analyzed, while *CSF3R* mutation was detected in 47.5% of colonies, suggesting that *CSF3R* mutations occurred later than *SETBP1* (Fig. 3F). Patient CMLPh-026 was mutated in *ASXL1* and *SETBP1*. Targeted analysis revealed the presence of mutated *ASXL1* in all the 58 clones analyzed, while heterozygous *SETBP1* mutation was detected in 50 clones (86.2%), indicating that *ASXL1* mutations occurred earlier than *SETBP1* ones (Fig. 3G).

Altogether, these results indicate that, when *ETNK1* and *SETBP1* mutations are co-present, *ETNK1* variants occur earlier in the clonal evolution history of aCML, while *SETBP1* mutations generally represent late events; interestingly, in two cases where *ASXL1* was mutated together with *SETBP1*, *ASXL1* mutations occupied an intermediate hierarchical position. *CBL* mutations, when present, showed a strong tendency toward reaching homozygosity through somatic uniparental disomy.

Combined targeting treatment of ETNK1 and NRAS mutations

In patient CMLPh-042, *ETNK1* G245 V and *NRAS* G12D were found to be early mutational events (Supplementary Table 10, http://links.lww.com/HS/A115). From the available sequencing data, we could not determine which of the 2 mutations arose first. We decided to target both mutations ex vivo: trametinib was used to block the NRAS pathway: at 10 nM it was able to reduce ERK phosphorylation (Supplementary Fig. 5, http://links.lww.com/HS/A105). ETNK1 was targeted with the use of phosphoethanolamine (P-Et), which was found to abrogate the effects of *ETNK1* mutations (Fontana et al Nat Commun., in press). Colonies grown from the patient showed a strong synergistic effect of the combination treatment of trametinib 10 nM and P-Et 1 mM, according to the Bliss Independence Principle[31] (expected additive effect of P-Et 1 mM + trametinib 10 nM combination: 146 colonies; observed colonies: 0; Fig. 4A, C).

Subsequently, a patient-derived xenograft (PDX) model was established. Treatment with 1 mg/kg trametinib allowed animals to live until the end of the experiment, while untreated animals showed signs of leukemia and were sacrificed (Fig. 5). Due to the limited number of animals transplanted, statistical significance was borderline ($p=0.07$). However, analysis of human cells in peripheral blood assessed by flow cytometry showed a complete reduction of human CD45+ cells in the treated animals compared to untreated ones (Fig. 5), and immunohistochemistry revealed the absence of human CD45 cells in the treated animals in both spleen and bone marrow, at difference from untreated ones (Fig. 5).

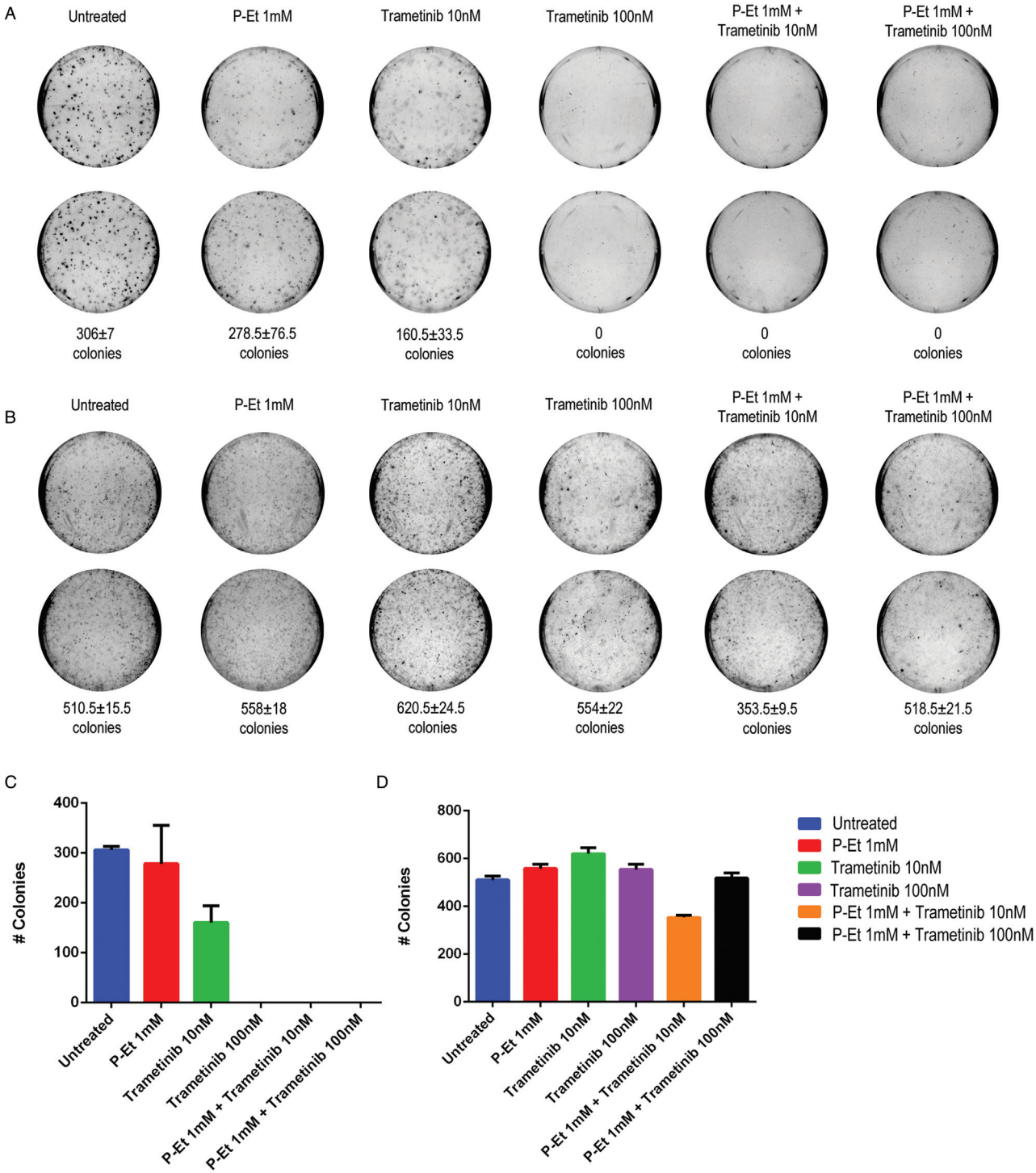

Figure 4. Effects of trametinib and phosphoethanolamine in patient CMLPh-042. (A) Colony-forming assay at the onset: bone marrow derived cells were left untreated or treated with phosphoethanolamine 1 mm, trametinib 10 nM, trametinib 100 nM, or combination of the two drugs. Colonies were counted after 15 days. (B) Colony-forming assay performed at relapse after treatment with trametinib 1 mg/day.(C) Number of colonies at the onset. (D) Number of colonies at relapse.

CMLPh-042 was treated on a named patient protocol with trametinib, at 1 mg/day. Unfortunately, no clinical grade P-Et was available. A hematological response was obtained with normalization of WBC, reduction of approximately 50% in the volume of the spleen and discontinuation of hydroxyurea (2 g/day). After 3 months of treatment the patient became severely anemic, transfusion dependent, the spleen volume and leukocyte numbers started to grow again, and trametinib was discontinued at month 4. Colonies grown at this time point showed complete resistance to trametinib (Fig. 4B, D), a testimony to the selection process which developed over 3 months, in spite of only a transient clinical response. Conventional cytogenetic analysis performed at

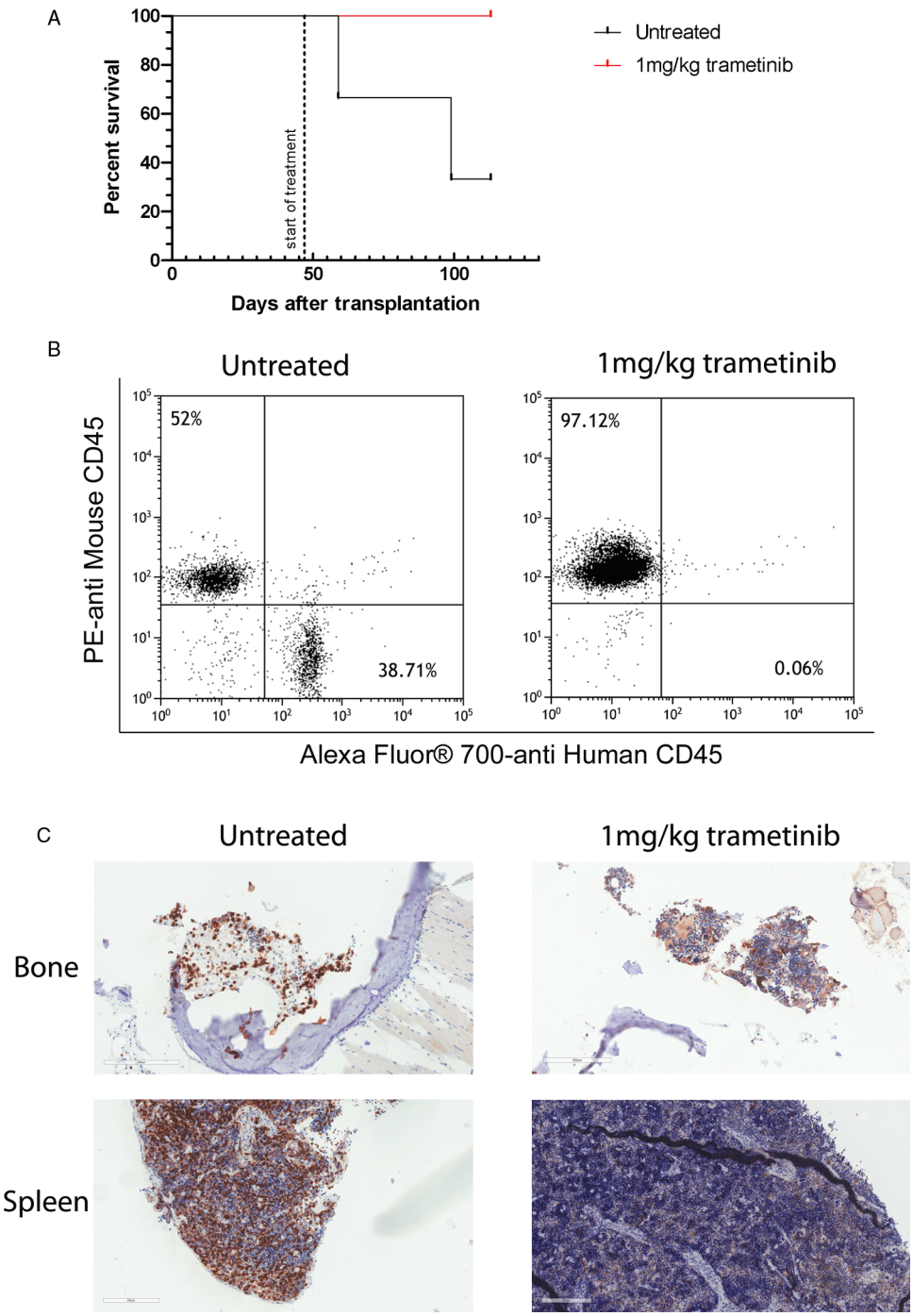

Figure 5. In-vivo experiments. (A) Overall survival (OS) of mice treated with 1 mg/kg trametinib by oral gavage once a day (red line) as compared with controls (black line). OS were analyzed using Kaplan-Meier plot and the log-rank test. (B) Analysis of human CD45 cells in peripheral blood assessed by flow cytometry in PDX models treated with 1 mg/kg trametinib compared to controls. Representative plots are shown. (C) Immunohistochemistry (human CD45 expression) of bone and spleen PDX models treated with 1 mg/kg trametinib by oral gavage once a day compared to controls. Representative images are shown. Scale bar: 200 μm.

the time of relapse revealed the presence of isocromosome 17q (karyotype: 46,Y,i(17)(q10)[16]/46,XY[4]) with the loss of one copy of 17p, which includes the p53 locus, suggesting a possible mechanism of acquired resistance to MEK inhibition (Supplementary Fig. 6, http://links.lww.com/HS/A105).[32] These data confirm that targeting the effects of a single mutation event can result in a clinical response, but of limited duration.

Patients stratification based on RNA-sequencing data

RNA-sequencing was performed on all 43 patients; no fusions were detected.[33] Stratification based on whole-transcriptome data[22] identified two clearly different populations (26 and 17 patients) in terms of Overall Survival (OS), with 2 year OS of

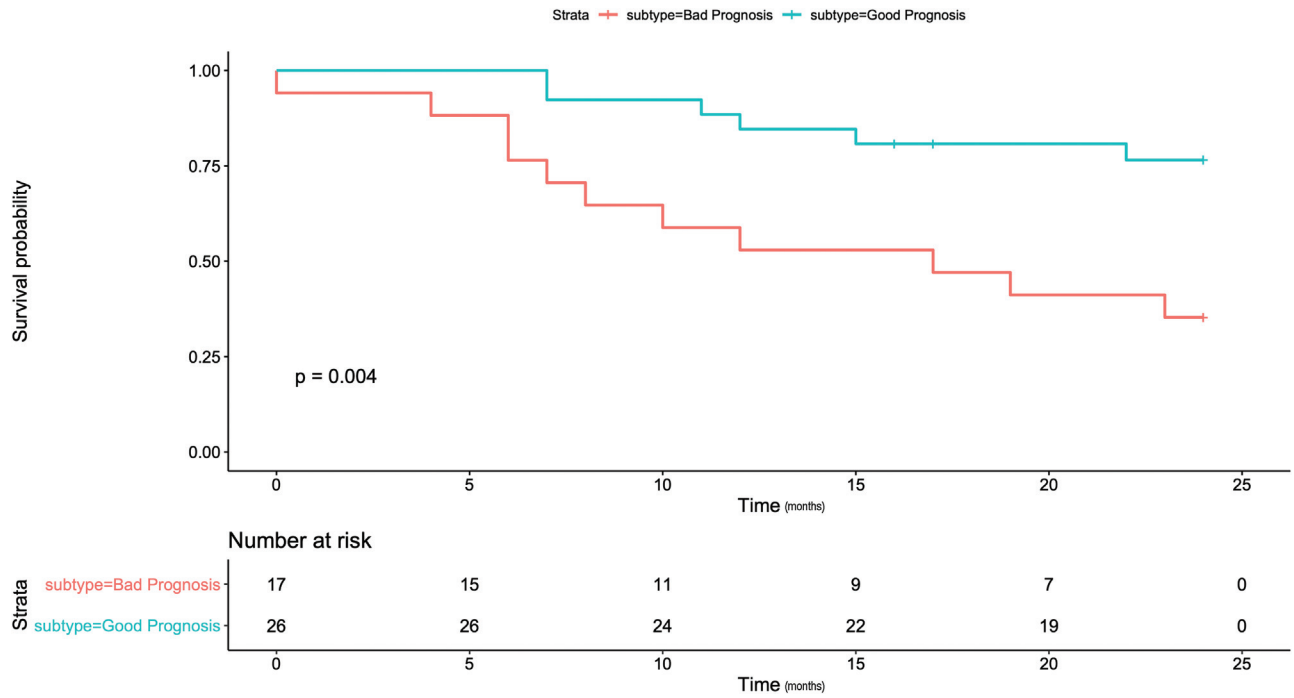

Figure 6. Overall survival curve (Kaplan-Meier curve). Overall survival curve censored at 24 months shows significantly different outcomes (low-rank p = 0.004).

69.23% [95% IC: 48.21%-86.67%] and 35.29% [95% IC: 14.21%-61.67%] respectively (log-rank test for trend: p = 0.004, Fig. 6; Supplementary Table 11, http://links.lww.com/HS/A116). The group with better prognosis showed a higher frequency of *ETNK1* mutations (22.7% vs 6.7%; hypergeometric test: p = 0.032). We performed differential gene expression analysis to detect genes that were differentially expressed between the two patients' populations. Functional enrichment annotation of the differentially expressed genes revealed several biological processes involved (Supplementary Tables 12, http://links.lww.com/HS/A117–13, http://links.lww.com/HS/A118), including gene transcription and cell differentiation, cell cycle regulation, mitochondrial activity, DNA repair. From these lists, we selected cancer-related GO terms (Fig. 7A) as well as cancer-related pathways (Fig. 7B) to perform clustering analysis. These analyses clearly showed 2 distinct clusters based on patients' outcome: cancer-related terms were highly enriched in patients with poor prognosis. Further analysis revealed 38 overexpressed genes in the group with negative clinical outcome (t-test p value adjusted for false discovery rate,[34] $p_{adj} < 0.01$; Table 1). To further reduce the number of classifier genes, we then considered expression data for the 3 most significant genes from the previous list (namely *DNPH1*, *GFI1B*, and *PARP1*). Using these 3 genes, we built a classifier model that is able to separate patients according to the respective subtype (better vs worse prognosis). The results showed that overexpression of these 3 genes is highly predictive of poor prognosis, and a random forest algorithm[27] applied to the 3 most significant genes achieves a 95.03% accuracy (out-of-bag error rate of 4.65%) assessed by means of 10 fold cross validation (Fig. 8).

Discussion

In the last decade, the application of NGS techniques dramatically improved our understanding of several issues in the biology of neoplasias. However, this knowledge has rarely been translated into better prognosis or treatment tools. In the present work, we analyzed both mutation profiles and RNA-sequencing expression data from a large cohort of aCML patients. aCML is a highly heterogeneous disorder characterized by both myelodysplastic and myeloproliferative features.[1] Several mutations in different genes are responsible for the onset of the disease. In our cohort, mutations in *ASXL1* gene were the most frequent alteration (16/37 cases), as already reported by other recent studies.[35–37] Additional frequent mutations involved *SETBP1*, *ETNK1*, *TET2* and *RAS* genes. This mutation profile is very close to that of a recent French report from Julien et al,[38] where *SETBP1* mutations occur in 30.3% of the aCML cases, *ETNK1* in 7.4% of patients, while *ASXL1* was found mutated in 68.8% of cases. Surprisingly, however, our profile is different from a recently published study,[36] in which Zhang and colleagues report the frequency of *SETBP1* mutations to be as low as 7.4% (2/27). The explanation for this difference is at present unclear; however, it may be caused by a different genetic background for the European/US aCML populations. In the context of the differential diagnosis of aCML and related clonal disorders such as CNL, *CSF3R* mutations are particular noteworthy.[39–41] In our study, we found only one patient out of 37 carrying a *CSF3R* mutation. This result confirms the rarity of such event in the mutational landscape of aCML. In addition to the high and distinctive heterogeneity of this disease, the lack of detailed information regarding the clonal hierarchy of the mutations contributes to make treatment approaches to this disorder even more difficult. Given the genetic complexity of aCML, we tried to reconstruct the clonal hierarchy of the observed mutations in order to identify early versus late occurring mutations. The clonal architecture of 7 aCML patients was characterized by colony assays and

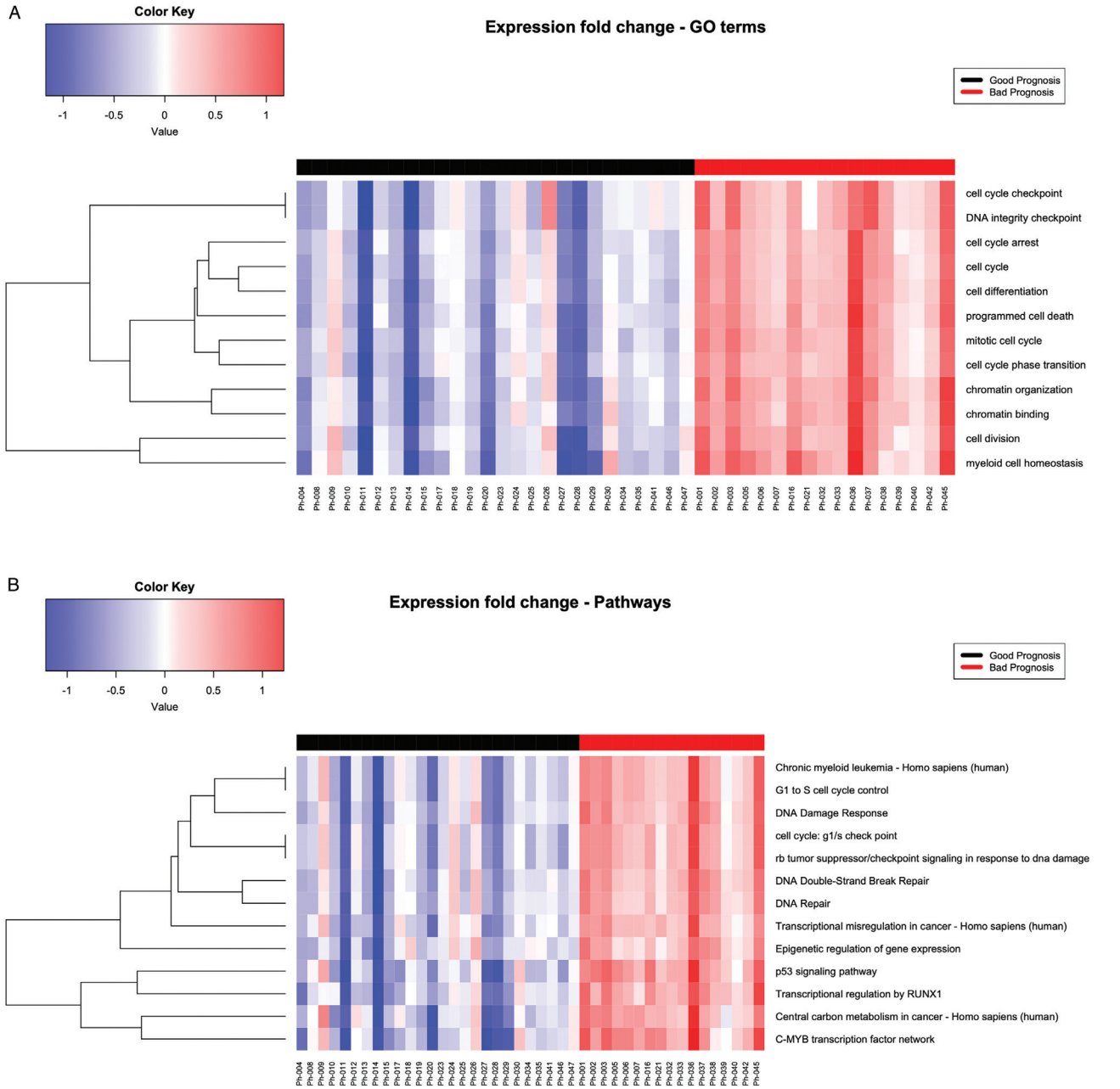

Figure 7. Gene ontology and pathway heatmaps. (A) Heatmap showing expression for a set of selected cancer-related GO terms is presented. (B) Heatmap showing expression for a set of selected cancer-related pathways is presented.

targeted resequencing. Our findings suggest that, when present, *ETNK1* variants occur at the initial stages of clonal evolution history of aCML. These results are in line with our recent findings regarding the role of *ETNK1* in the induction of a mutator phenotype (Fontana et al Nat Commun., in press).

The knowledge of the clonal progression can be useful to suggest treatment priorities. The inclusion of early occurring mutations in the targeting strategy carries the advantage of hitting all leukemic cells. Surprisingly, it is important to note that, in contrast to gene expression results, the presence of mutations, either as single or combined events, and weighted by their respective OncoScore, failed to predict clinical outcome. This finding would seem to contradict what previously asserted by our group.[6] Unfortunately, the rarity of the disease and the complexity and heterogeneity of the somatic drivers involved in this disorder, where *SETBP1* is often present together with several other driver mutations such as *ASXL1*, *ETNK1*, *TET2* or others, makes a clear dissection of the prognostic effect of individual variants extremely challenging. Probably, a larger cohort of aCML patients will be required to thoroughly analyze this point. In contrast, transcriptomic data demonstrated the presence of two distinct populations which significantly differed in terms of overall survival. Based

Table 1

Differentially expressed genes between the two clusters of patients.

Gene Name	Median Good Prognosis	Median Bad Prognosis	Log2 fold change Good prognosis/Bad prognosis	p value Two-Sided	OncoScore	Is Oncogene?
AIMP2	54.50	185.00	−1.76	0.00004	50.49	1
AURKA	60.50	184.00	−1.60	0.00407	75.65	1
CDK4	203.00	802.00	−1.98	0.00074	74.80	1
CHST10	20.50	78.00	−1.93	0.00317	30.00	1
DNAJC11	199.50	532.00	−1.42	0.00006	33.33	1
DNPH1	18.00	95.00	−2.40	0.00003	45.99	1
EI24	138.00	430.00	−1.64	0.00200	45.75	1
ERCC2	119.00	304.00	−1.35	0.00048	72.57	1
EXTL2	24.50	99.00	−2.01	0.03693	42.60	1
FAM189B	78.50	227.00	−1.53	0.00029	0.00	0
GADD45GIP1	57.50	274.00	−2.25	0.00010	52.13	1
GATA1	26.50	318.00	−3.58	0.00188	44.21	1
GFI1B	31.00	348.00	−3.49	0.00203	44.22	1
GNA12	175.50	502.00	−1.52	0.00103	34.94	1
H2AFX	84.50	376.00	−2.15	0.00225	70.66	1
HSPD1	825.00	2040.00	−1.31	0.00686	31.55	1
IDH2	614.00	1830.00	−1.58	0.00136	78.73	1
KEAP1	160.50	445.00	−1.47	0.00005	41.29	1
KIT	136.50	722.00	−2.40	0.00880	31.28	1
MEN1	216.00	598.00	−1.47	0.00053	88.03	1
MYBBP1A	212.50	599.00	−1.50	0.00017	55.23	1
MYC	177.50	1031.00	−2.54	0.00095	69.59	1
NME1	13.00	79.00	−2.60	0.00165	82.83	1
PA2G4	323.50	960.00	−1.57	0.00224	61.57	1
PARP1	329.50	1437.00	−2.12	0.00227	56.94	1
PBX1	34.50	186.00	−2.43	0.00173	79.54	1
PDZK1IP1	3.00	27.00	−3.17	0.00371	50.88	1
PIR	3.50	23.00	−2.72	0.00122	12.07	0
POLRMT	196.00	485.00	−1.31	0.00005	14.73	0
RITA1	34.00	126.00	−1.89	0.00004	12.30	0
RNF43	22.50	58.00	−1.37	0.00460	80.49	1
SALL2	21.00	50.00	−1.25	0.00734	58.62	1
SLC39A4	29.50	101.00	−1.78	0.00059	18.08	0
SMARCB1	264.50	608.00	−1.20	0.00172	80.79	1
TAL1	62.50	525.00	−3.07	0.00057	74.33	1
TCF3	223.00	769.00	−1.79	0.00005	47.38	1
TIMP3	16.00	190.00	−3.57	0.00019	50.28	1
TP53	159.00	608.00	−1.94	0.00006	90.62	1

The table reports a list of 38 genes significantly higher expressed in the cluster with bad prognosis. Median expression values for the 2 clusters, Log2 fold change Good vs Bad prognosis, and t-tests to assess their differences are also reported. Their OncoScore as well as their classification as oncogenes (marked as 1 in the table) are presented.

on these findings, we identify a 3-genes signature capable of stratifying aCML patients according to their prognoses with high accuracy. The 3 identified genes are known oncogenes and play roles in cellular proliferation, DNA repair, tumor transformation (Supplementary Table 14, http://links.lww.com/HS/A119). A limitation of these results resides in the lack of a validation cohort which will need future work, and in the limited number of aCML patients under study, due to the rarity of the disease. The fact that mutation analysis did not prove clinically informative but gene expression profiles did, raises important considerations. While mutations certainly affect gene functions, several non-mutational mechanisms can also affect gene activity (eg, promoter methylation, histone modifications, gene amplifications and deletions). All these mechanisms result in altered gene expression, the analysis of which could provide clinical insights into the disease behavior. In addition, gene expression profiles also reflect the impact of a certain mutation on the genetic background of each individual patient.

Our ex vivo and in vivo results also indicate that the presence of actionable mutations could indeed inform therapy and drive synergistic combinations; in one case where *NRAS* and *ETNK1* mutations were simultaneously present, trametinib gave evidence of clinical, albeit limited, therapeutic activity.

In conclusion, our work provides novel insights on aCML clonal evolution and suggests the presence of 2 subtypes of aCML characterized by distinct RNA expression profiles, showing different clinical outcomes. Further studies will be required to confirm our findings. In general, it will be important to obtain more insights into the molecular mechanisms governing aCML development and progression, and to convert them into better treatment strategy modalities, since no effective therapies are available to date for aCML and the outcome is almost invariably fatal.

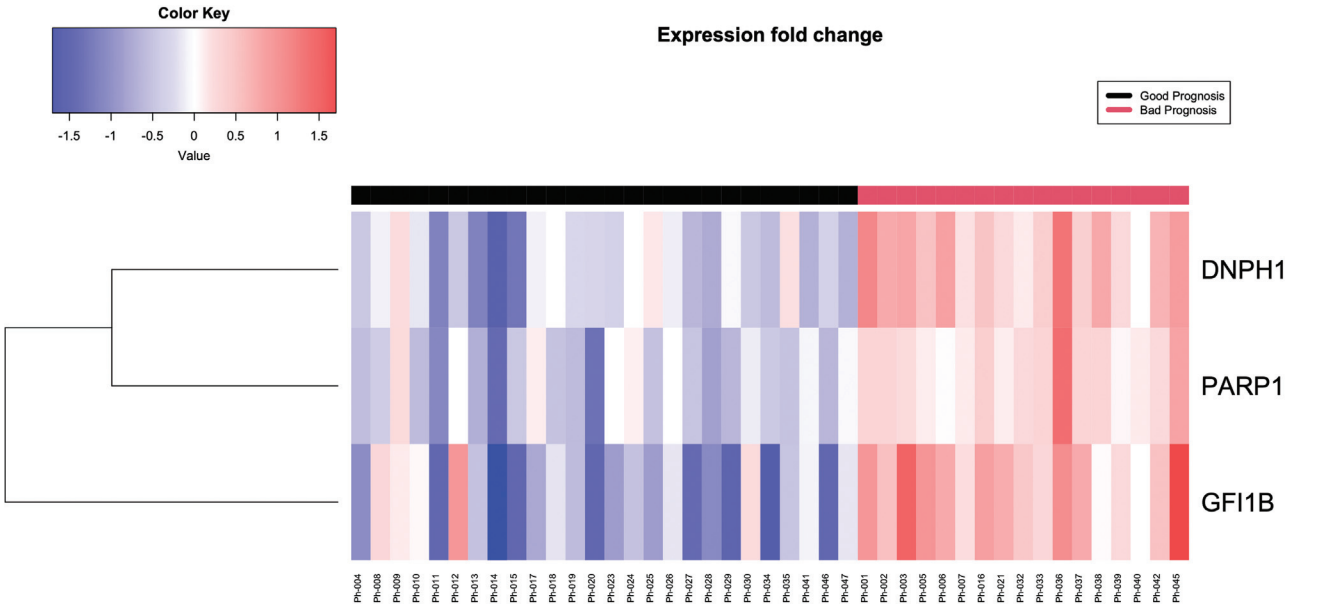

Figure 8. **Heatmap showing fold change for the threetop differentially expressed genes used to classify good vs bad prognosis subtypes is presented.**

Sources of Funding

This work was supported by Fondazione AIRC per la Ricerca sul Cancro 2018 (IG-22082) to RP, Fondazione AIRC per la Ricerca sul Cancro 2015 (IG-17727) to RP, Fondazione AIRC per la Ricerca sul Cancro 2017 (IG-20112) to CGP.

Acknowledgments

We kindly acknowledge the contributions of Michela Viltadi for technical help.

References

1. 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;127:2391–2405.
2. Vardiman JW, Bain B, Inbert M. Jaffe E, Harris NL, Stein H, Vardiman J, et al. Atypical chronic myeloid leukemia. *WHO classification of tumors: pathology and genetics of tumours of haematopoietic and lymphoid tissues.* Lyon: IARC Press; 2001;53–57.
3. Oscier D. Atypical chronic myeloid leukemias. *Pathologie-biologie.* 1997;45:587–593.
4. Breccia M, Biondo F, Latagliata R, et al. Identification of risk factors in atypical chronic myeloid leukemia. *Haematologica.* 2006;91:1566–1568.
5. Cazzola M, Malcovati L, Invernizzi R. Myelodysplastic/myeloproliferative neoplasms. *Hematol Am Soc Hematol Educ Program.* 2011;1:264–272.
6. Piazza R, Valletta S, Winkelmann N, et al. Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. *Nat Genet.* 2013;45:18–24.
7. Gotlib J, Maxson JE, George TI, et al. The new genetics of chronic neutrophilic leukemia and atypical CML: implications for diagnosis and treatment. *Blood.* 2013;122:1707–1711.
8. Gotlib J. How I treat atypical chronic myeloid leukemia. *Blood.* 2017;129:838–845.
9. Gambacorti-Passerini CB, Donadoni C, Parmiani A, et al. Recurrent ETNK1 mutations in atypical chronic myeloid leukemia. *Blood.* 2015;125:499–503.
10. Makishima H, Yoshida K, Nguyen N, et al. Somatic SETBP1 mutations in myeloid malignancies. *Nat Genet.* 2013;45:942–946.
11. Lasho TL, Finke CM, Zblewski D, et al. Novel recurrent mutations in ethanolamine kinase 1 (ETNK1) gene in systemic mastocytosis with eosinophilia and chronic myelomonocytic leukemia. *Blood Cancer J.* 2015;5:e275.
12. Meggendorfer M, Bacher U, Alpermann T, et al. SETBP1 mutations occur in 9% of MDS/MPN and in 4% of MPN cases and are strongly associated with atypical CML, monosomy 7, isochromosome i(17)(q10), ASXL1 and CBL mutations. *Leukemia.* 2013;27:1852–1860.
13. Sakaguchi H, Okuno Y, Muramatsu H, et al. Exome sequencing identifies secondary mutations of SETBP1 and JAK3 in juvenile myelomonocytic leukemia. *Nat Genet.* 2013;45:937–941.
14. Inoue D, Kitaura J, Matsui H, et al. SETBP1 mutations drive leukemic transformation in ASXL1-mutated MDS. *Leukemia.* 2015;29:847–857.
15. Fabiani E, Falconi G, Fianchi L, et al. SETBP1 mutations in 106 patients with therapy-related myeloid neoplasms. *Haematologica.* 2014;99:e152–e153.
16. Patnaik MM, Itzykson R, Lasho T, et al. ASXL1 and SETBP1 mutations and their prognostic contribution in chronic myelomonocytic leukemia: a two-center study of 466 patients. *Leukemia.* 2014;28:2206–2212.
17. Thol F, Suchanek KJ, Koenecke C, et al. SETBP1 mutation analysis in 944 patients with MDS and AML. *Leukemia.* 2013;27:2072–2075.
18. Wang XA, Muramatsu H, Okuno Y, et al. GATA2 and secondary mutations in familial myelodysplastic syndromes and pediatric myeloid malignancies. *Haematologica.* 2015;100:E398–E401.
19. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25:1754–1760.
20. Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools. *Bioinformatics.* 2009;25:2078–2079.
21. Piazza R, Magistroni V, Pirola A, et al. CEQer: a graphical tool for copy number and allelic imbalance detection from whole-exome sequencing data. *PloS One.* 2013;8:e74825.
22. Ramazzotti D, Lal A, Wang B, et al. Multi-omic tumor data reveal diversity of molecular mechanisms that correlate with survival. *Nat Commun.* 2018;9:4453.
23. Bailey MH, Tokheim C, Porta-Pardo E, et al. Comprehensive characterization of cancer driver genes and mutations. *Cell.* 2018;173:371–385. e318.
24. Liu Y, Sun J, Zhao M. ONGene: a literature-based database for human oncogenes. *J Genet Genomics.* 2017;44:119–121.
25. Zhao M, Sun J, Zhao Z. TSGene: a web resource for tumor suppressor genes. *Nucleic Acids Res.* 2013;41:D970–976.

26. Zhao M, Kim P, Mitra R, et al. TSGene 2.0: an updated literature-based knowledgebase for tumor suppressor genes. *Nucleic Acids Res 2016.* 2016;44:D1023–1031.
27. Ho TK. Random Decision Forests. *Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC*, 14–16 August 1995, IEEE. 1995;1:278–282.
28. Piazza R, Mologni L, Ramazzotti D, et al. Oncoscore, a Novel, internet-based tool to assess the oncogenic potential of genes can differentiate Between CP-CML and BC-CML associated genes, and between CP-CML patients with good and bad prognosis. *Blood.* 2016;128:
29. Schwartz LC, Mascarenhas J. Current and evolving understanding of atypical chronic myeloid leukemia. *Blood Rev.* 2019;33:74–81.
30. Schittenhelm MM, Shiraga S, Schroeder A, et al. Dasatinib (BMS-354825), a dual SRC/ABL kinase inhibitor, inhibits the kinase activity of wild-type, juxtamembrane, and activation loop mutant KIT isoforms associated with human malignancies. *Cancer Res.* 2006;66:473–481.
31. Bliss CI. The calculation of microbial assays. *Bacteriol Rev.* 1956;20:243–258.
32. Najem A, Krayem M, Sales F, et al. P53 and MITF/Bcl-2 identified as key pathways in the acquired resistance of NRAS-mutant melanoma to MEK inhibition. *Eur J Cancer.* 2017;83:154–165.
33. Piazza R, Pirola A, Spinelli R, et al. FusionAnalyser: a new graphical, event-driven tool for fusion rearrangements discovery. *Nucleic Acids Res.* 2012;40:e123.
34. Yoav B, Daniel Y. The control of the false discovery rate in multiple testing under dependency. *Ann Stat.* 2001;29:1165–1188.
35. Meggendorfer M, Haferlach T, Alpermann T, et al. Specific molecular mutation patterns delineate chronic neutrophilic leukemia, atypical chronic myeloid leukemia, and chronic myelomonocytic leukemia. *Haematologica.* 2014;99:
36. Zhang H, Wilmot B, Bottomly D, et al. Genomic landscape of neutrophilic leukemias of ambiguous diagnosis. *Blood.* 2019;134:867–879.
37. Faisal M, Stark H, Busche G, et al. Comprehensive mutation profiling and mRNA expression analysis in atypical chronic myeloid leukemia in comparison with chronic myelomonocytic leukemia. *Cancer Med.* 2019;8:742–750.
38. Julien V, Rea D, Thepot S, et al. Current treatments do not improve the prognosis of patients with atypical CML and unclassified MDS/MPN. A joint report from Fi-LMC, FIM, Gfch and GFM. *Blood.* 2019;134 (Supplement_1):2954.
39. Pardanani A, Lasho TL, Laborde RR, et al. CSF3R T618I is a highly prevalent and specific mutation in chronic neutrophilic leukemia. *Leukemia.* 2013;27:1870–1873.
40. Maxson JE, Tyner JW. Genomics of chronic neutrophilic leukemia. *Blood.* 2017;129:715–722.
41. Wang SA, Hasserjian RP, Fox PS, et al. Atypical chronic myeloid leukemia is clinically distinct from unclassifiable myelodysplastic/myeloproliferative neoplasms. *Blood.* 2014;123:2645–2651.