Published OnlineFirst November 2, 2020; DOI: 10.1158/1078-0432.CCR-20-3184

International Consensus Definition of DNA Methylation Subgroups in Juvenile Myelomonocytic Leukemia

Maximilian Schönung[1,2], Julia Meyer[3], Peter Nöllke[4], Adam B. Olshen[5,6], Mark Hartmann[1], Norihiro Murakami[7], Manabu Wakamatsu[7], Yusuke Okuno[8], Christoph Plass[9], Mignon L. Loh[3,6], Charlotte M. Niemeyer[4,10], Hideki Muramatsu[7], Christian Flotho[4,10], Elliot Stieglitz[3,6], and Daniel B. Lipka[1,11]

ABSTRACT

Purpose: Known clinical and genetic markers have limitations in predicting disease course and outcome in juvenile myelomonocytic leukemia (JMML). DNA methylation patterns in JMML have correlated with outcome across multiple studies, suggesting it as a biomarker to improve patient stratification. However, standardized approaches to classify JMML on the basis of DNA methylation patterns are lacking. We, therefore, sought to define an international consensus for DNA methylation subgroups in JMML and develop classification methods for clinical implementation.

Experimental Design: Published DNA methylation data from 255 patients with JMML were used to develop and internally validate a classifier model. Accuracy across platforms (EPIC-arrays and MethylSeq) was tested using a technical validation cohort (32 patients). The suitability of both methods for single-patient classification was demonstrated using an independent cohort (47 patients).

Results: Analysis of pooled, published data established three DNA methylation subgroups as a *de facto* standard. Unfavorable prognostic parameters (*PTPN11* mutation, elevated fetal hemoglobin, and older age) were significantly enriched in the high methylation (HM) subgroup. A classifier was then developed that predicted subgroups with 98% accuracy across different technological platforms. Applying the classifier to an independent validation cohort confirmed an association of HM with secondary mutations, high relapse incidence, and inferior overall survival (OS), while the low methylation subgroup was associated with a favorable disease course. Multivariable analysis established DNA methylation subgroups as the only significant factor predicting OS.

Conclusions: This study provides an international consensus definition for DNA methylation subgroups in JMML. We developed and validated methods which will facilitate the design of risk-stratified clinical trials in JMML.

Introduction

Juvenile myelomonocytic leukemia (JMML) is a myeloproliferative/myelodysplastic neoplasm (1) with an incidence of 1.3 per million children and a median age at diagnosis of 2 years (2). Older age, elevated fetal hemoglobin (HbF) levels, and thrombocytopenia at diagnosis are established clinical risk factors correlating with poor clinical outcome (3).

More than 90% of patients with JMML harbor canonical mutations in *PTPN11, KRAS, NRAS, CBL*, or *NF1* that constitutively activate RAS signaling. Somatic *PTPN11* mutations are detected in approximately 35%–40% of patients, somatic *KRAS* and somatic *NRAS* mutations are found in about 15% each, while alterations of the *CBL* or *NF1* locus are observed in about 15% and 10% of patients, respectively (1, 4). Karyotypic abnormalities can be observed in 35% of patients, with monosomy 7 as the most frequent event occurring in 25% of children with JMML (5). Whole-exome or targeted sequencing studies identified recurrent secondary mutations in a number of patients with JMML, commonly affecting *JAK3, SETBP1*, and *SH2B3* (6–10). While older patients with *PTPN11*-, *NF1*-, or *NRAS*-driven disease have a high relapse incidence following allogeneic hematopoietic stem cell transplantation (HSCT), spontaneous disease regression is observed in some younger patients with *NRAS* and *CBL* mutations (11, 12). Interestingly, infants with Noonan syndrome, caused by germline *PTPN11* mutations, can experience a myeloproliferative disorder that can be indistinguishable from JMML, but generally has a self-limiting course.

In an attempt to resolve this striking clinical and biological heterogeneity, gene expression profiling identified patients with JMML with an acute myeloid leukemia (AML)-like expression signature that was associated with poor survival (13). However, the precise mechanisms regulating disease-specific gene expression patterns in JMML are still obscure. A link between aberrant RAS signaling and epigenetic remodeling has been suggested in the literature (14). Consistent with this, DNA hypermethylation of candidate gene promoters was described in JMML, including *CDKN2B, RASSF1A, CREBBP, RASA4*,

[1]Section Translational Cancer Epigenomics, Division Translational Medical Oncology, German Cancer Research Center (DKFZ) & National Center for Tumor Diseases (NCT), Heidelberg, Germany. [2]Faculty of Biosciences, Heidelberg University, Heidelberg, Germany. [3]Department of Pediatrics, Benioff Children's Hospital, University of California, San Francisco, California. [4]Division of Pediatric Hematology and Oncology, Department of Pediatrics and Adolescent Medicine, Medical Center, Faculty of Medicine, University of Freiburg, Freiburg, Germany. [5]Department of Epidemiology and Biostatistics, University of California, San Francisco, California. [6]Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, California. [7]Department of Pediatrics, Nagoya University Graduate School of Medicine, Nagoya, Aichi, Japan. [8]Medical Genomics Center, Nagoya University Hospital, Nagoya, Aichi, Japan. [9]Division Cancer Epigenomics, German Cancer Research Center (DKFZ), Heidelberg, Germany. [10]German Cancer Consortium (DKTK), partner site Freiburg, Germany. [11]Faculty of Medicine, Otto-von-Guericke-University, Magdeburg, Germany.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

M. Schönung and J. Meyer contributed equally as co-first authors of this article.

H. Muramatsu, C. Flotho, E. Stieglitz, and D.B. Lipka contributed equally as co-senior authors of this article.

Corresponding Author: Daniel B. Lipka, German Cancer Research Center (DKFZ) & National Center for Tumor Diseases (NCT) Heidelberg, INF 581, Heidelberg 69120, Germany. Phone: 49-6221-42-1603; Fax: 49-3212-8886999; E-mail: d.lipka@dkfz.de

Clin Cancer Res 2020;XX:XX–XX

doi: 10.1158/1078-0432.CCR-20-3184

Published OnlineFirst November 2, 2020; DOI: 10.1158/1078-0432.CCR-20-3184

Translational Relevance

Known clinical and genetic markers have limitations in predicting disease course and treatment outcome in juvenile myelomonocytic leukemia (JMML). Aberrant DNA methylation patterns correlate with clinical outcome in JMML across multiple studies, suggesting it as a biomarker to improve patient stratification. However, standardized approaches to identify DNA methylation subclasses are lacking. In this study, we defined an international consensus for DNA methylation subgroups in JMML, and developed and validated classification methods that are suitable in the single-patient setting. Our DNA methylation classifier is ready for clinical implementation and will enable the design of risk-stratified clinical trials in JMML.

the β-globin promoter, and others (15–21). The DNA methylation status of four candidate genes (*BMP4, CALCA, CDKN2B*, and *RARB*) was integrated into a prognostic model, with DNA methylation emerging as the strongest independent prognostic predictor in a multivariable analysis (22). More recently, three independent studies described DNA methylation subgroups in JMML based on genome-wide DNA methylation analysis (23–25). Study groups from Europe and United States identified three JMML subgroups with distinct clinical features, whereas a Japanese group suggested a binary classification into hyper- or hypomethylated disease. In all three studies, the hypermethylation subgroup was associated with *PTPN11* mutations and poor overall survival (OS). Together, these studies suggested DNA methylation as a biomarker which might help to improve patient stratification in JMML. However, standardized approaches to identify DNA methylation JMML subclasses are lacking.

In this study, we sought to define an international consensus for DNA methylation subgroups in JMML and to systematically describe the biological and clinical features associated with each subgroup. Furthermore, we developed and validated classification methods that are suitable for clinical implementation and will allow risk stratification in clinical trials.

Materials and Methods

Collection of data and patient samples

Illumina Infinium Human Methylation 450k Bead Chip raw data from 292 patients (203 male and 89 female) with JMML or patients with Noonan syndrome and myeloproliferative disease (NS/MPD) were collected from three recent publications and clinical annotations were obtained from the respective study groups (refs. 23–25; Supplementary Data S1; Supplementary Fig. S1). Healthy references were obtained from Gene Expression Omnibus (accession number, GSE36054). To ensure reproducibility on different technology platforms, 32 patients from this cohort were reanalyzed using targeted bisulfite sequencing (MethylSeq) and 31 of these 32 patients were reanalyzed by Infinium Human MethylationEPIC Bead Chip (EPIC) arrays ("technical validation cohort," Supplementary Data S1). An independent validation cohort of 47 patients from the participating study centers was collected for classifier validation. Written informed consent from parents or legal guardians of all patients was obtained according to the Declaration of Helsinki. Patients' material storage and collection were approved by institutional ethics committees.

Processing of DNA methylation array data

DNA methylation data were analyzed using the "RnBeads" Bioconductor package (26). Background correction ("methylumi.noob"; ref. 27) and beta-mixture quantile normalization were applied. Unreliable probes (Greedycut algorithm with detection $P < 0.01$), cross-reactive probes, and probes mapping to sex chromosomes were removed (Supplementary Fig. S1A). Samples with outlier intensities in 450k/EPIC array control probes were removed from the dataset as described in the RnBeads vignette (ref. 26; Supplementary Materials and Methods). Hierarchical clustering on the basis of SNP distances revealed study group–specific clusters that likely reflected ethnic differences (Supplementary Fig. S1B). In-line with this, initial exploratory analysis using principal component analysis separated EWOG-MDS (European Working Group of Myelodysplastic Syndromes in Childhood) from United States and Japanese patients in the first principal component (Supplementary Fig. S1C). We, therefore, applied: (i) SNPfiltering using dbSNP version 150 and (ii) Combat batch correction using the "sva" R package (ref. 28; Supplementary Fig. S1D). The final RnBeads dataset was generated from the binary exponential of the batch corrected value. Finally, all CpG dinucleotides (CpG) with variable DNA methylation across normal hematopoietic differentiation were excluded to account for potential differences in cell type composition of the samples (24). The preprocessing of EPIC array data was analogous to 450k data as described above. For the technical validation cohort, overlapping probes between 450k and EPIC arrays were determined. Patient-wise Pearson correlation was calculated across all probes and plotted as a correlation heatmap using the pheatmap R package (29).

Consensus clustering of DNA methylation array data

The 5,000 most variable CpGs (mvCpGs), defined as the probes with the highest SD between all samples, were identified and used for hierarchical consensus clustering using the ConsensusClusterPlus R package (30) with Ward linkage (ward.D2), Manhattan distance, and 500 bootstrap iterations. Within each DNA methylation consensus cluster, patients were ordered according to similarity by hierarchical clustering. Results were plotted using the pheatmap R package (29).

Copy-number variation analysis

Copy-number variation was inferred from 450k or EPIC array data using the Conumee Bioconductor R package (31). Focal amplifications and deletions were determined by "Genomic Identification of Significant Targets in Cancer" (GISTIC2; ref. 32). Data from healthy adult blood (EGAS00001002511; 450k array) and cord blood (GSE103189; EPIC array) were used as reference.

Extreme gradient boosting model

The "XGBoost" R package was used for a tree-based gradient boosting model (33). The patient cohort was split into a training ($n = 229$) and a testing cohort ($n = 55$), with random splits for each DNA methylation class to keep the class distribution balanced. CpGs which overlapped between 450k and EPIC methylation arrays were determined, and CpG sites with differentiation-dependent variation were removed. A multiclass classification model was trained for the prediction of DNA methylation subgroups using the 5,000 mvCpGs (softprob objective and 100 iterations).

Hyperparameter tuning was performed using a grid-search algorithm with 5-fold cross-validation (nrounds = 65, eta = 0.05, max_depth = 4, gamma = 1, colsample_bytree = 0.4, min_child_weight = 3, and subsample = 0.75). The model accuracy was determined by predicting DNA methylation classes in the testing cohort. The feature importance for the tuned XGBoost model was

Published OnlineFirst November 2, 2020; DOI: 10.1158/1078-0432.CCR-20-3184

calculated using the caret-wrapper function to determine the relative feature importance of the model CpGs (34). Of 5,000 model CpGs, 124 showed a feature importance of greater than zero (Supplementary Data S2). A refined JMML DNA methylation classifier was trained on the basis of these 124 CpGs and the hyperparameters from the grid-search algorithm. The model accuracy was tested by comparing the DNA methylation class predictions of the testing cohort with the meta-analysis DNA methylation group assignments. Furthermore, 31 of 32 patients from the technical validation cohort were reanalyzed on EPIC methylation arrays. The model prediction accuracy was then tested with this cohort to validate the model accuracy across different technology platforms. The robustness of the refined 124-CpG classifier was tested by repeated retraining using only a subset of CpG sites. This approach enabled us to determine the model accuracy for DNA methylation class prediction of the training and testing cohorts for models based on 5–120 CpGs (each subset randomly selected 100×).

DNA panel sequencing

DNA samples (32 from the technical cohort and 47 from the independent validation cohort) were sequenced using a custom Amplicon Panel (Paragon Genomics), targeting 26 genes that are recurrently mutated in JMML (Supplementary Materials and Methods; Supplementary Table S1). Ten nanograms of genomic DNA were used for each library and processed according to the manufacturer's protocol.

Library quality was assessed using the Bioanalyzer High Sensitivity DNA Analysis Kit (Agilent) and quantified using the Qubit HS Assay (Thermo Fisher Scientific). Dual-index libraries were sequenced on an Illumina MiSeq Instrument using the 150 bp PE mode. Mean on-target sequencing depth was 1,519× (median, 1,297×). A minimum mutant allele fraction (MAF) of 0.03 was required for reporting. Patient samples where the MAF of the putative driver mutation was <0.25 were removed from the analysis to ensure only clonal samples were being analyzed.

Human MethylationEPIC Bead Chip sample processing

Genomic DNA (100–250 ng) was subjected to the Genomics and Proteomics Core Facility at the German Cancer Research Center (Heidelberg, Germany) to perform genome-wide DNA methylation analysis using the Infinium Human MethylationEPIC Bead Chip Platform (Illumina).

Targeted MethylSeq library preparation and sequencing

A total of 3,000 CpG loci were submitted for custom-capture assay design, of which 2,992 CpGs were among the top ranked 5,000 CpGs discriminating the JMML DNA methylation subgroups identified in the meta-analysis. Genomic DNA (300 ng) was bisulfite converted using the TrueMethyl oxBS module according to the manufacturer's protocol (Tecan). Converted ssDNA was quantified using the Qubit ssDNA Assay (Thermo Fisher Scientific) and 100 ng was used as input for the custom-made Targeted MethylSeq Assay (Tecan). Final libraries were quality checked using the Bioanalyzer High Sensitivity DNA Analysis Kit (Agilent) and quantified using the Qubit HS Assay (Thermo Fisher Scientific) before pooling for next-generation sequencing. Single-index library pools were sequenced on the Illumina HiSeq 4000 with paired-end 150 bp mode.

Hierarchical clustering–based sample classification using MethylSeq data

Sequence reads were trimmed using Cutadapt (35) to remove adapters, and CpG methylation status was called using Bismark (36). Only 2,992 CpG sites were considered for analysis, and required to have a median 50× read coverage, resulting in removal of 150 CpG sites. From the remaining 2,842 CpG sites, 1,386 CpGs showed SDs greater than 0.25 across the 32 samples, which was the final dataset used to analyze the technical validation cohort (Supplementary Data S3). The beta-values of the 1,386 CpG sites of the 32 patients from the technical validation cohort were hierarchically clustered in both directions (samples and probes) utilizing an unsupervised approach using the hclust function and Ward method. The samples were classified into one of three classes on the basis of minimum distance to the centroid. The MethylSeq data from the independent validation cohort data were filtered accordingly and hierarchical clustering was performed using the list of 1,386 CpG sites determined on the basis of the analysis of the technical validation cohort. Methylation subgroup classification into one of three classes was based on the minimum distance to the centroid.

Statistical analysis

Clinical characteristics and mutation status were collected for all patients from the participating study groups. Pearson χ^2 test was used to test categorical variables for independence. One-way ANOVA was utilized to test for differences among continuous variables. Fisher exact test was used to test for enrichment of canonical JMML driver mutations for each DNA methylation subgroup. The primary endpoint, OS, was defined as the time between diagnosis and death or last follow-up. The Kaplan–Meier method was used to estimate survival rates and the two-sided log-rank test was employed to evaluate the equality of the survivorship functions in different subgroups. For univariable and multivariable analysis, the relationships between OS and the variables, *PTPN11* (mutation or not), gender, age (≥1 year or not), mutations (>1 mutation or not), HbF (normal vs. elevated), platelet count (≥50 or not), XGBoost subgroup [low DNA methylation (LM), intermediate DNA methylation (IM), or high DNA methylation (HM)], and MethylSeq subgroup (LM, IM, or HM), were examined. First, we tested all variables individually and then included significant variables in a multivariable model. $P \le 0.01$ was considered statistically significant, except in the multivariable analysis, where a cutoff of 0.05 was used.

Data sharing statement

The data used in this study are available for download from the European Genome-phenome archive (https://ega-archive.org/) under the accession number EGAS00001004682. All other relevant data are available from the authors upon request.

Results

Patients' characteristics of the meta-analysis cohort

DNA methylome data of 292 patients with JMML ($n = 263$) and NS/MPD ($n = 29$) were combined from three previously published studies for an integrative analysis (23–25). Data of 284 patients (143 EWOG-MDS; 104 Japan; and 37 United States) passed our quality control criteria (Supplementary Fig. S1A). Of these, 255 patients (90%) were diagnosed with JMML and 29 (10%) with NS/MPD (Supplementary Table S1). Patients with NS/MPD were significantly younger than patients with JMML, with a median age at diagnosis of 0.1 year (range, 0.0–0.3; $P = 8.2 \times 10^{-7}$); all were diagnosed before the age of 2 years (Supplementary Table S1).

For patients with JMML, the tissue source for the DNA methylome analysis was either bone marrow (BM; 69%, $n = 175$ patients) or peripheral blood (PB; 31%, $n = 80$ patients; Supplementary Table S2).

Published OnlineFirst November 2, 2020; DOI: 10.1158/1078-0432.CCR-20-3184

Median age at JMML diagnosis was 1.4 years (40% ≥2 years). A total of 69% of patients were male and 31% were female. HbF levels at diagnosis were available for 210 (82%) patients and elevated for age in 145 of 210 (69%) patients (37). Median leukocyte and platelet counts at diagnosis were 29.9 × 10^9/L (range, 2.9–563) and 65.5 × 10^9/L (range, 5–730), respectively. Canonical JMML driver mutations were detected in 90% of patients with JMML (38% *PTPN11*; 16% *NRAS*; 14% *KRAS*; 14% *CBL*; 9% *NF1*; 8% quintuple negative; and 1% NA). Monosomy 7 was detected in 14% of patients and was the only parameter tested that showed differential distribution between the study groups ($P = 2.4 \times 10^{-3}$): monosomy 7 was present in 22% of EWOG-MDS patients, but was only found in 9% and 3% of the Japanese and U.S. patients, respectively (Supplementary Table S2).

Consensus clustering confirms three DNA methylation subgroups in JMML

Unsupervised hierarchical consensus clustering using the 5,000 mvCpG sites identified three distinct DNA methylation subgroups, which were designated as LM, IM, and HM according to mean beta-values (**Fig. 1**; Supplementary Fig. S2A–S2J). When comparing our consensus approach with the assignments in the original study groups' publications, we identified the same methylation subgroup assignment for 96% of EWOG-MDS (120/125), 95% of United States (35/37), and 91% of Japanese (85/93) patients with JMML (**Fig. 2A–C**; refs. 23–25). A global concordance of 96% was calculated when combining IM and HM subgroups as hypermethylated JMML, which allowed a comparison with the two methylation subgroups identified in the Japanese study (ref. 25; Supplementary Table S3). All patients with NS/MPD were clustered in the LM subgroup and formed a subcluster that almost exclusively consisted of patients with NS/MPD. Of note, the DNA methylomes of patients with NS/MPD and LM were almost indistinguishable from those of healthy controls (Supplementary Fig. S3). There was no significant association between sample source (BM or PB) and DNA methylation subgroups ($P = 0.9012$, χ^2 test; Supplementary Table S3; Supplementary Fig. S4).

Consensus DNA methylation subgroups correlate with disease biology

The association of consensus DNA methylation subgroups with the canonical JMML subgroups and clinical hematologic parameters was investigated for all patients with JMML (Supplementary Table S3). Patients with somatic *PTPN11* mutations were predominantly assigned to the HM subgroup (61%) and comprised more than 70% of HM patients ($P = 3.2 \times 10^{-13}$; **Fig. 2D**; Supplementary Table S3). All HM patients had elevated HbF (100%) and were older at the time of diagnosis (median age: HM, 3; IM, 1.4; and LM, 0.7 years; $P < 2 \times 10^{-16}$, AOV-test; Supplementary Table S3). The IM subgroup was mainly characterized by the presence of *PTPN11* or *KRAS* mutations (42% and 27% of all IM patients). Of all monosomy 7 cases, 75% (27/36 patients) were found to cluster with the IM subgroup, whereas the remaining 25% of monosomy 7 cases clustered with the HM subgroup ($P = 1.5 \times 10^{-7}$, χ^2 test). In contrast, almost all patients with *CBL* mutations (35/36 patients, 97%; $P = 4.3 \times 10^{-15}$, χ^2 test) and the vast majority of patients with *NRAS* mutations (30/40 patients, 75%; $P = 1.2 \times 10^{-7}$, χ^2 test) clustered in the LM subgroup (**Figs. 1** and **2D**; Supplementary Table S3). The patients with quintuple-negative JMML comprised a small proportion of all patients with JMML (21/255 patients, 8%) and almost exclusively clustered with the LM (52%) or with the IM subgroup (38%; **Fig. 2D**; Supplementary Table S3).

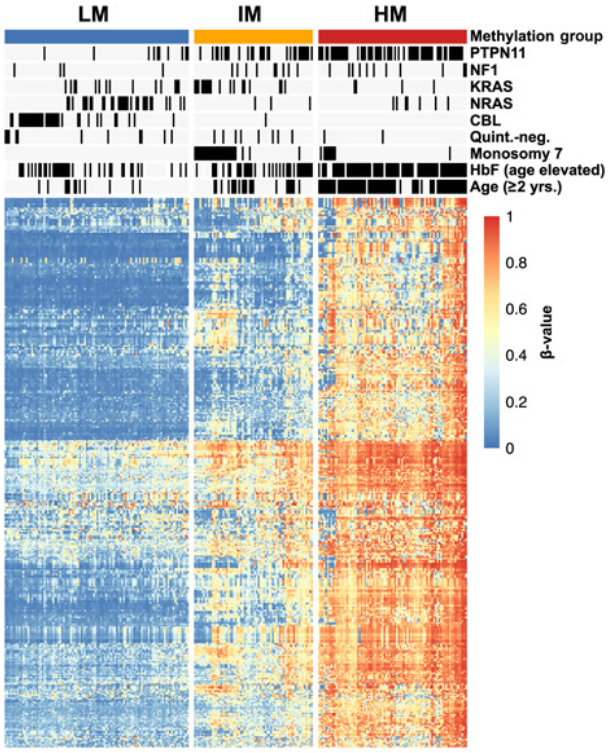

Figure 1.

DNA methylation patterns identify three biologically distinct JMML subgroups. The heatmap displays three DNA methylation subgroups among patients with JMML ($n = 255$; NS/MPD patients excluded). The patients (columns) were clustered by unsupervised consensus clustering ($k = 3$) using the 5,000 mvCpGs. Known clinical and biological features are annotated for each patient. The heatmap shows beta-values of 395 representative CpG sites (rows).

Genotype-specific enrichment analyses identified further associations between JMML driver mutations, clinical characteristics, and DNA methylation subgroups. For example, patients who had a *PTPN11* mutation and additional clinical high-risk factors (i.e., elevated HbF, thrombocytopenia, and higher age) were more likely assigned to the HM subgroup, whereas patients with *CBL* mutations were most commonly assigned to the LM subgroup, independent of the presence or absence of additional clinical risk factors (Supplementary Fig. S5A). In this retrospective analysis, all *NF1* patients who presented with thrombocytopenia were assigned to the HM subgroup and none of the *NF1* patients older than 2 years at the time of JMML diagnosis clustered with the LM subgroup (Supplementary Fig. S5A). Monosomy 7 cooccurred with *PTPN11* (16/36; 44%), *KRAS* (14/36; 39%), or *NF1* (2/36; 6%) mutations, but not with *NRAS* or *CBL* mutations. Patients with cooccurrence of monosomy 7 and a *KRAS* mutation were always assigned to the IM subgroup, while in the presence of *PTPN11* or *NF1* mutations, patients with monosomy 7 were assigned either to the IM or HM subgroups. Of note, monosomy 7 was absent in the LM subgroup (Supplementary Table S3). In the group of *KRAS*-mutant patients, a skewed distribution of molecular features became evident across patients from different study groups: while 9 of 14 (64%) Japanese patients with *KRAS* mutations were assigned to the LM subgroup ($P = 8.0 \times 10^{-2}$, Fisher exact test), 13 of 19 (68%) EWOG-MDS patients with *KRAS* mutations were categorized into the IM subgroup ($P = 5.5 \times 10^{-3}$, Fisher exact test). This

Published OnlineFirst November 2, 2020; DOI: 10.1158/1078-0432.CCR-20-3184

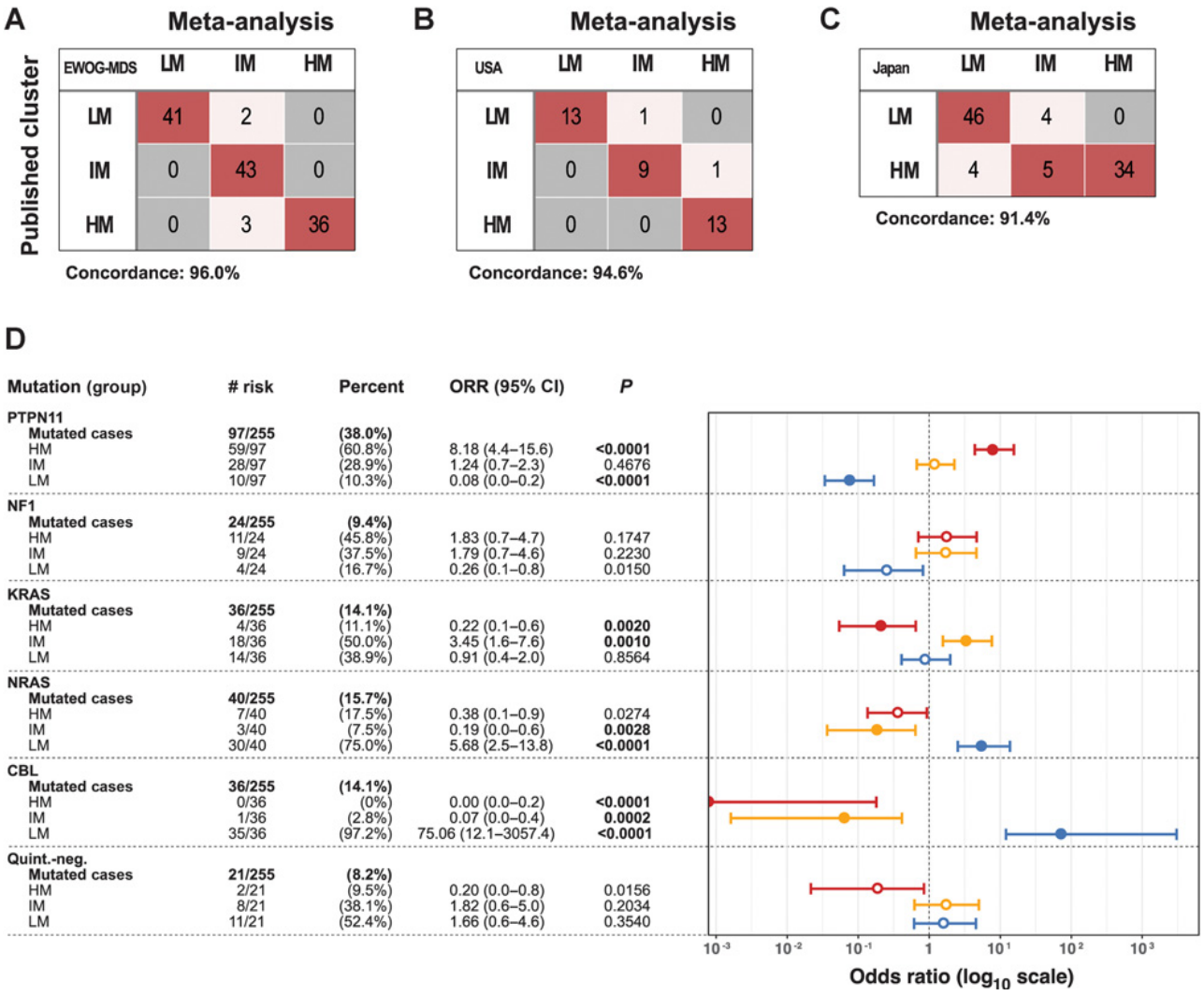

Figure 2.

Enrichment of JMML driver mutations in DNA methylation subgroups. **A–C,** Confusion table comparing the published methylation subgroups with the meta-analysis consensus clustering. EWOG-MDS patients (**A**) and U.S. patients (**B**) were classified in three methylation subgroups in the original publication, whereas the Japanese study (**C**) separated patients in two methylation subgroups. **D,** Absolute and relative numbers of patients with a specific JMML driver mutation and their distribution across the DNA methylation subgroups are depicted. ORs, CIs, and *P* values were calculated using Fisher exact test.

difference was paralleled by a more frequent cooccurrence of *KRAS* mutations with monosomy 7 in EWOG-MDS compared with Japanese patients (EWOG-MDS, $n = 12$; Japan, $n = 2$; $P = 6.0 \times 10^{-3}$, Fisher exact test; Supplementary Fig. S5B).

Consensus DNA methylation subgroups predict clinical outcome and identify patients at high risk

We next analyzed the impact of the consensus DNA methylation subgroups on clinical outcome. Because of differences in treatment regimens and follow-up periods, clinical outcome was analyzed separately for each study group. Kaplan–Meier curves for OS demonstrated that HM JMML had inferior 5-year OS in all study groups [EWOG-MDS, 62% (95% confidence interval (CI), 45–79); Japan, 46% (95% CI, 29–62); United States, 42% (95% CI, 22–79)], especially when compared with the LM subgroup (5-year OS EWOG-MDS, 87% (95% CI, 76–98); Japan, 74% (95% CI, 59–85); United States, 100%; Supplementary Fig. S6A–S6C). Long-term survival without HSCT [5-year treatment-free survival (TFS)] was only observed in patients assigned to the LM subgroup [EWOG-MDS, 24% (95% CI, 10–38); Japan, 39% (95% CI, 24–53); United States, 54% (95% CI, 33–89); Supplementary Fig. S6D–S6F].

Development and validation of a machine learning model for prospective JMML DNA methylation subgroup classification in the single-patient setting

To prospectively classify single patients with JMML into consensus DNA methylation subgroups as defined by this meta-analysis, we developed and validated a machine learning model. The combined JMML patient cohort was split into a training ($n = 229$) and a testing cohort ($n = 55$). A multiclass classification gradient boosting tree model (XGBoost) was developed on the basis of the 5,000 mvCpGs (**Fig. 3A**; Supplementary Fig. S1). To determine the model with the highest prediction accuracy, hyperparameter tuning was performed using a 5-fold cross-validation approach with subsequent feature

Published OnlineFirst November 2, 2020; DOI: 10.1158/1078-0432.CCR-20-3184

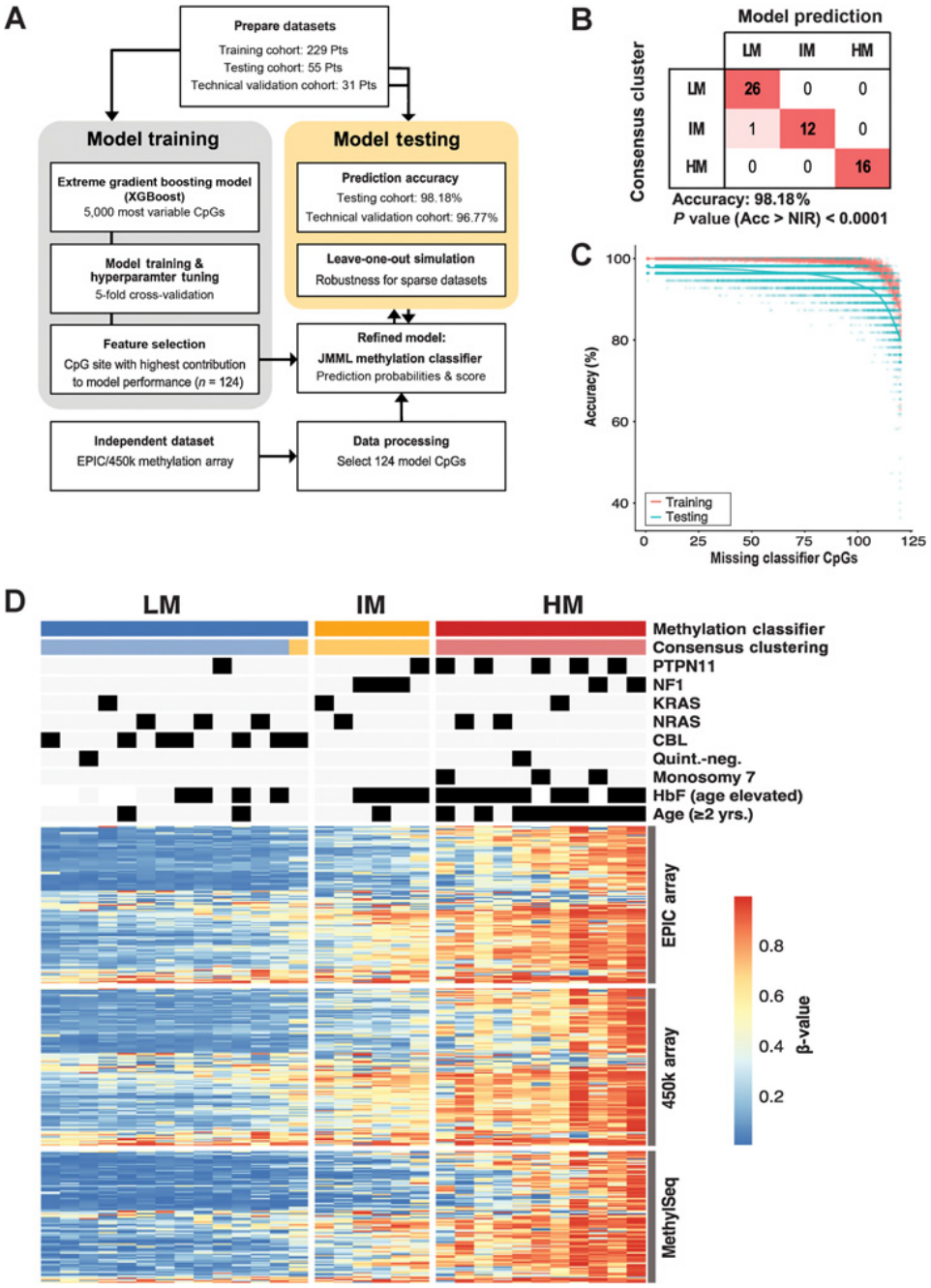

Figure 3.

Development of a DNA methylation classifier for JMML. **A,** The meta-analysis cohort was randomly split into a training ($n = 229$) and a testing cohort ($n = 55$). A machine learning classifier was built based on the 5,000 mvCpGs, which were covered by 450k and EPIC methylation arrays, and trained on the training cohort. The fractional contribution of each CpG to the overall model performance was calculated and the CpGs with the highest gain were selected ($n = 124$). This refined JMML methylation classifier was trained and the model performance evaluated on the testing cohort. In addition, 31 patients were reanalyzed on EPIC arrays to test the model accuracy across different technical assays ("technical validation cohort"). **B,** Confusion table for prediction of the testing cohort using the refined JMML methylation classifier (acc, accuracy; NIR, no-information rate). **C,** The JMML methylation classifier was retrained with subsets of the 124 model CpGs, by repeatedly leaving out between 5 and 120 CpG sites. The model accuracy of these sparse models for predicting the training or testing cohort was determined. **D,** Heatmap showing the DNA methylation beta-values for the 124 model CpG sites (rows) on 450k and EPIC methylation arrays for patients of the technical validation cohort ($n = 31$; columns). In addition, 104 of 124 CpGs assessed by the MethylSeq assay are shown. The predictions by the methylation classifier model are compared with the consensus clustering.

selection based on variable importance. The refined JMML DNA methylation classifier was trained on the basis of 124 CpG sites and validated by classifying the testing cohort (**Fig. 3B**; Supplementary Data S2). The model predicted consensus DNA methylation subgroups in the testing cohort with an accuracy of 98% (95% CI, 90.3%–99.9%); that is, only 1 of 55 patients was classified differently than by consensus clustering (**Fig. 3B**).

To test the power of the model for sparse datasets and the stability of the underlying JMML DNA methylation signature, a "leave-one-out" scenario was simulated (**Fig. 3C**). Repeated model training and accuracy assessments using increasing numbers of CpG sites (between 4 and 119 of the 124 model CpG sites) based on the training and testing cohorts were used for this simulation. We found that training the model with as few as 24 of 124 CpG sites still resulted in a prediction accuracy of about 85%. We confirmed the model performance across different technological platforms by reanalyzing 31 of 292 patients from the combined cohort (referred to as "technical validation cohort") on the EPIC array platform (one sample did not meet quality control cutoffs). The genome-wide DNA methylation measurements were highly correlated between the 450k and the EPIC array platforms (Supplementary Fig. S7A). Accordingly, the classifier model predicted the consensus DNA methylation classes from EPIC array data with an accuracy of 97% in the technical validation cohort (Supplementary Fig. S7B).

Published OnlineFirst November 2, 2020; DOI: 10.1158/1078-0432.CCR-20-3184

Targeted MethylSeq recapitulates consensus DNA methylation subgroup assignments

We also developed a targeted DNA methylation assay based on bisulfite sequencing (targeted MethylSeq) that could be incorporated into routine clinical workflows used for mutation testing. This assay encompassed 3,000 CpG sites, 2,992 of which overlapped with the 5,000 mvCpGs used for consensus clustering in this study (Supplementary Data S3). The 32 patients of the technical validation cohort (for MethylSeq 1 additional patient was available) were analyzed using this assay (Supplementary Data S4). After quality control filtering, 1,386 CpGs with an SD of greater than 0.25 between the samples were used for hierarchical clustering. Again, this resulted in three distinct DNA methylation subgroups that corresponded to the consensus assignments for all 32 samples analyzed (Supplementary Fig. S7C). Furthermore, 104 CpG sites of the 3,000 probes overlapped with the 124 CpG sites from the JMML methylation classifier, and their DNA methylation patterns were similar to the EPIC array data and the 450k array combined cohort analysis (**Fig. 3D**).

Validation of the JMML DNA methylation classifier in an independent patient cohort

An independent patient cohort consisting of 47 patients (9 from EWOG-MDS, 18 form Japan, and 20 from United States) without prior DNA methylation analysis was collected to validate the DNA methylation classification methods. All patients from this independent validation cohort were tested for known JMML mutations using targeted deep sequencing. The DNA methylation analysis was performed using both the EPIC array and the targeted MethylSeq platforms. In this cohort, 46 patients were diagnosed with JMML and 1 patient was diagnosed with NS/MPD. The median age at diagnosis was 1.7 years (range, 0.0–6.4). HbF was elevated for age in 77.8% of patients. Median leukocyte and platelet counts were 30.2×10^9/L (range, 5.0–166.4) and 67×10^9/L (range, 11–490), respectively (Supplementary Table S4). The median OS was 6.7 years. Canonical JMML driver mutations were detected in 96% of patients with JMML (44% *PTPN11*, 15% *KRAS*, 24% *NRAS*, 4% *CBL*, and 9% *NF1*). Accordingly, 4% of patients were annotated as "quintuple negative." DNA methylation subgroups were determined by the 124 CpG XGBoost classifier in a single-patient setting (**Fig. 4A**; Supplementary Table S4) and by minimum distance to the centroid based on the 1,386 CpGs sites that were previously used to cluster the technical validation cohort (**Fig. 4B**; Supplementary Data S2 and S3). On the basis of the 124 CpG XGBoost classifier, 21 patients were classified as HM, 10 as IM, and 16 as LM JMML (**Fig. 4A**). *PTPN11* mutations, higher age, and elevated HbF levels were significantly enriched in patients with HM JMML (Supplementary Table S4). When classifying the independent patient cohort based on the MethylSeq data, 20 patients were classified as HM, 8 as IM, and 19 as LM. Importantly, the concordance of HM prediction was 95% (**Fig. 4B**; Supplementary Fig. S8A). Of note, variant allele frequency measurements of canonical JMML driver mutations suggested that DNA methylation subgroups are not confounded by differences in tumor cell purity (Supplementary Fig. S8B and S8C). Previous studies have suggested an enrichment of secondary mutations predominately in HM JMML (23, 24). In the validation cohort, 71% of HM patients and 50% of IM patients, but none of the LM patients, showed recurrent secondary mutations (**Fig. 4C**). Interestingly, in patients with JMML with *PTPN11* mutations, most secondary mutations were detected in *JAK3* and *NF1* (79%; 11/14 patients with secondary mutation), whereas the most common secondary mutation in patients with *NRAS*-mutant JMML was *SETBP1* (100%; 3/3 patients with secondary mutation). In summary, we found a strong enrichment of secondary genetic hits in the HM subgroup.

The clinical relevance of the 124 CpG XGBoost DNA methylation classifier was highlighted by significant differences in prognosis (log-rank test, $P = 0.003$). OS at 2.5 years was significantly lower for patients with HM JMML as compared with patients with IM and LM JMML [2.5-year OS HM, 36% (95% CI, 19–69); IM, 79% (95% CI, 56–100); and LM, 88% (95% CI, 73–100); **Fig. 4D**]. Similarly, DNA methylation subgroups identified by the targeted MethylSeq assay were also associated with significant differences in prognosis [log-rank test, $P = 0.003$; 2.5-year OS HM, 45% (95% CI, 26–78); IM, 63% (95% CI, 37–100); and LM, 84% (95% CI, 69–100); **Fig. 4E**].

Methylation status is independently predictive of outcome

In univariable analyses (**Table 1**), the characteristics that reached significance at the 0.05 level for OS were age at diagnosis >12 months (HR, 2.75; 95% CI, 1.08–7.01; $P = 0.03$), platelets ≤50 (HR, 0.35; 95% CI, 0.14–0.89; $P = 0.026$), and DNA methylation group ($P = 0.0026$ for XGBoost and $P = 0.0015$ for MethylSeq). When a multivariable model was applied using age, platelets, and DNA methylation group, the DNA methylation grouping retained statistical significance for OS ($P = 0.046$ for XGBoost and $P = 0.039$ for MethylSeq; **Table 1**), whereas neither age ($P = 0.51$ for XGBoost and $P = 0.82$ for MethylSeq) nor platelets ($P = 0.13$ for XGBoost and $P = 0.16$ for MethylSeq) were significant.

Discussion

The diagnosis of JMML has so far been based on clinical and genetic criteria. The unifying feature is pathologic activation of the RAS signaling pathway, but the clinical course of JMML is highly heterogeneous. Patients with identical mutations may either experience spontaneous remission of the disease or relapse after allogeneic HSCT, indicating that genotype does not always dictate phenotype in this disease (38, 39). Outcomes can be partly predicted using a combination of clinical and genetic parameters (3, 6). However, well-established cutoffs for those parameters or even prognostic scores are missing. In addition, parameters, such as HbF and thrombocytopenia, may not be informative in cases where supportive treatment has been initiated before referral to a specialist center. Hence, robust, objective, and reproducible molecular assays are needed to advance the diagnosis, risk stratification, and treatment of JMML. The potential of DNA methylation to subclassify tumors such as glioma, AML, acute lymphoblastic leukemia, chronic lymphocytic leukemia (40, 41), and Waldenström macroglobulinemia (42) has been demonstrated in several publications (43–46). In recent years, a number of publications have highlighted the correlation of DNA methylation with disease biology and prognosis in JMML (18, 19, 22–25).

Combining published data from three JMML study groups, we analyzed the genetic and DNA methylation landscape in the largest JMML patient cohort studied to date.

This meta-analysis sheds some light on controversial topics that have been discussed in the JMML research community. It confirmed the presence of at least three JMML DNA methylation subgroups, with an obvious underrepresentation of Japanese patients in the IM subgroup. This might explain why a previous Japanese study identified two, instead of three, JMML DNA methylation subgroups (25). Similarly, we found that most *KRAS*-mutant EWOG-MDS patients clustered in the IM subgroup, while most *KRAS*-mutant Japanese patients clustered in the LM subgroup. Biological differences between the

Published OnlineFirst November 2, 2020; DOI: 10.1158/1078-0432.CCR-20-3184

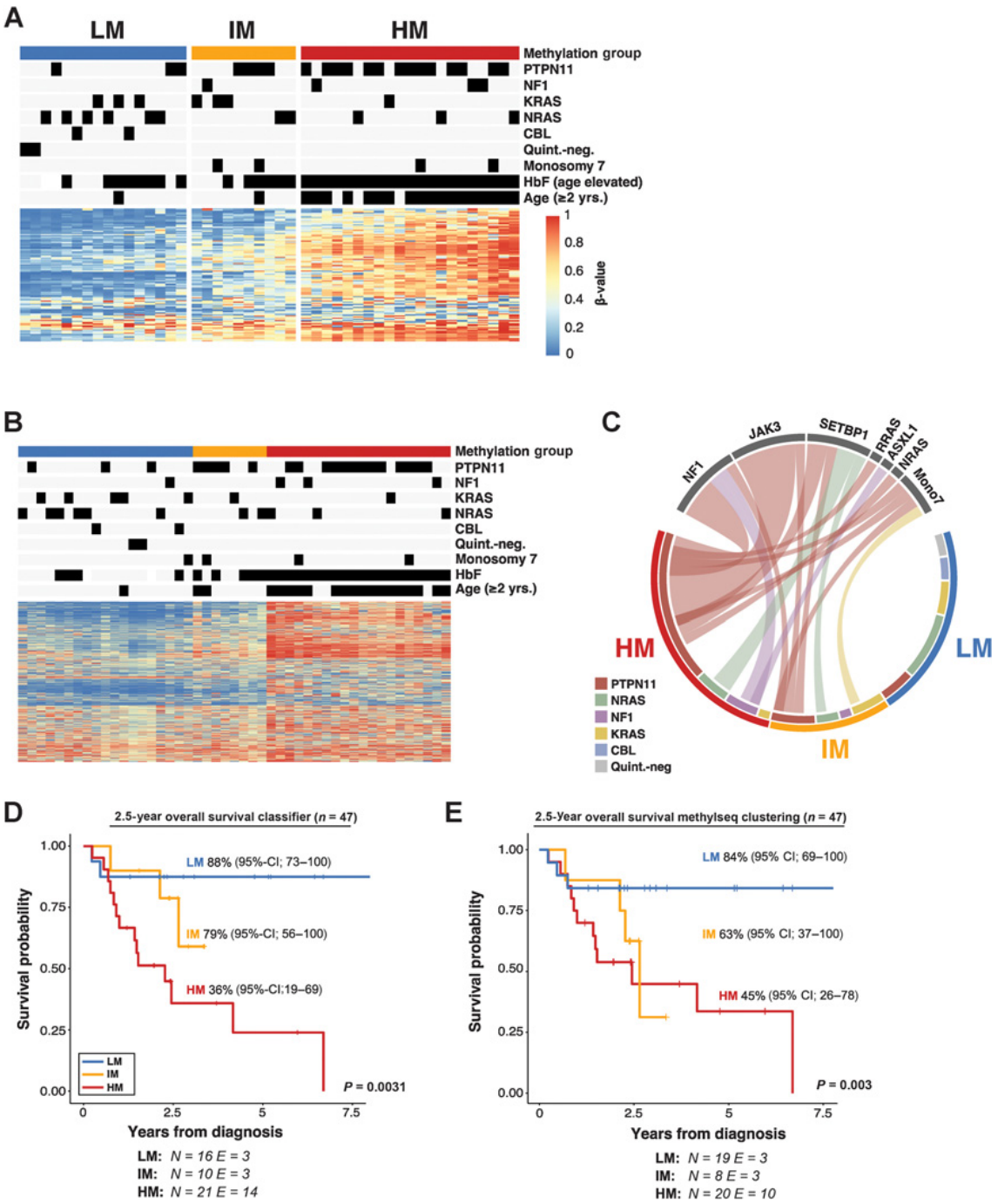

Figure 4.

Validation of the DNA methylation classifier in an independent patient cohort. **A,** Heatmap showing the DNA methylation beta-values of the 124 model CpG sites for an independent validation cohort ($n = 47$). Methylation subgroups were assigned by classifier predictions. **B,** Determination of the DNA methylation subgroups using targeted amplicon-bisulfite sequencing (MethylSeq). Patients were clustered using hierarchical clustering with Ward method. **C,** Circos plot displaying the association between known driver mutations and secondary mutations in the three DNA methylation subgroups as determined by the 124 CpG machine learning classifier. **D** and **E,** Kaplan–Meier curves showing the OS of patients with JMML stratified by DNA methylation subgroups. DNA methylation subgroups were assigned on the basis of the JMML methylation classifier (**D**) or MethylSeq clustering (**E**). Probabilities and CIs are indicated for each DNA methylation subgroup. The number of individuals at risk (*N*) and the number of events (*E*) are depicted at the bottom. Statistical significance was calculated using log-rank test.

Published OnlineFirst November 2, 2020; DOI: 10.1158/1078-0432.CCR-20-3184

Table 1. Univariable and multivariable analysis for OS in the biological validation cohort.

		OS from date of diagnosis	
Univariable analysis	***N***	**HR (95% CI)**	***P***
Age at diagnosis (months)			
≤12 months	27	1	
>12 months	20	2.75 (1.08–7.01)	**0.03**
Platelet count at diagnosis ($\times 10^9$)			
≤50	17	1	
>50	30	0.35 (0.14–0.89)	**0.026**
XGBoost cluster			
Low	16	1	
Intermediate	10	2.96 (0.48–18.22)	
High	21	8.07 (1.81–36.01)	**0.0026**
MethylSeq cluster			
Low	19	1	
Intermediate	8	6.76 (1.20–38.01)	
High	20	8.93 (1.99–40.0)	**0.0015**
Somatic *PTPN11* mutation			
No	26	1	
Yes	21	1.79 (0.72–44.45)	0.21
Sex			
Male	34	1	
Female	13	0.84 (0.28–2.57)	0.76
Somatic mutations at diagnosis			
≤1	29	1	
>1	18	2.33 (0.93–5.79)	0.068
HbF at diagnosis			
Not elevated for age	10	1	
Elevated for age	35	1.36 (0.39–4.73)	0.61
Multivariable analysis (XGBoost)			
Age at diagnosis (months)			
≤12 months	27	1	
>12 months	20	0.61 (0.15–2.47)	0.51
Platelet count at diagnosis ($\times 10^9$)			
≤50	17	1	
>50	30	0.49 (0.19–1.25)	0.13
XGBoost cluster			
Low	16	1	
Intermediate	10	3.09 (0.49–19.65)	
High	21	10.82 (1.56–74.84)	**0.046**
Multivariable analysis (MethylSeq)			
Age at diagnosis (months)			
≤12 months	27	1	
>12 months	20	0.85 (0.20–3.53)	0.82
Platelet count at diagnosis ($\times 10^9$)			
≤50	17	1	
>50	30	0.52 (0.20–1.31)	0.16
MethylSeq cluster			
Low	19	1	
Intermediate	8	6.17 (1.06–35.94)	
High	20	8.92 (1.25–63.48)	**0.039**

Note: A univariable analysis, including age at diagnosis, platelet count at diagnosis, XGBoost cluster, MethylSeq cluster, somatic *PTPN11* mutations, sex, number of somatic mutations, and HbF status, was conducted for predictors of OS. *P* values lower than 0.05 were considered as statistically significant and are depicted in bold. A multivariable analysis including the significant features was calculated thereafter including either the XGBoost or MethylSeq methylation cluster assignments.

Abbreviation: *N*, number of patients.

regional patient groups were further supported by the observation that monosomy 7 is rare in patients from Japan as compared with EWOG-MDS (monosomy 7 Japan, 9% and monosomy 7 EWOG-MDS, 22%). This difference cannot be explained by technical differences as the copy-number variation status was uniformly inferred from the 450k or EPIC array data for all patients. In conclusion, we observed differences in the genetic and epigenetic landscape of *KRAS* patients between Europe, United States, and Japan, which might, at least, in part, explain previously discussed differences in the clinical course of these patients (12, 47).

Our reanalysis established three DNA methylation subgroups as a *de facto* standard, which will allow the uniform classification of JMML into biologically and clinically meaningful risk groups. Furthermore, we developed and validated tools that allow for systematic and

reproducible classification of JMML in the clinical setting and demonstrated that DNA methylation classification in JMML is stable across orthogonal technology platforms. Analysis of an independent patient cohort validated these classification tools. Using multivariable analysis, we confirmed the prognostic relevance of DNA methylation subgroups and provide further evidence suggesting that DNA methylation subgroup is the strongest independent prognostic factor in JMML.

This meta-analysis provides a molecular rationale to define high-risk patients for whom allogeneic HSCT is not curative and who are, therefore, candidates for clinical trials testing innovative treatment options. On the other hand, there likely exists a subset of patients, characterized by an LM phenotype, for whom a watch-and-wait strategy plus supportive care might be the appropriate intervention.

In summary, this meta-analysis of 255 patients with JMML provides a consensus definition for DNA methylation subgroups in JMML. We have developed a DNA methylation classifier and validated its performance in an independent patient cohort. This classifier allows the prospective identification of DNA methylation subgroups for newly diagnosed patients based on the consensus definition described here. This work will support patient stratification and development of risk-adapted treatment strategies in the context of clinical trials and will improve the comparison of results obtained with different treatment strategies across study groups.

Authors' Disclosures

D.B. Lipka reports grants from German José Carreras Leukemia Foundation during the conduct of the study. No disclosures were reported by the other authors.

Authors' Contributions

M. Schönung: Conceptualization, data curation, formal analysis, validation, investigation, visualization, writing-original draft. **J. Meyer:** Conceptualization, data curation, formal analysis, validation, investigation, visualization, writing-original draft. **P. Nöllke:** Data curation, formal analysis. **A.B. Olshen:** Data curation, formal analysis. **M. Hartmann:** Data curation, formal analysis. **N. Murakami:** Data curation, formal analysis. **M. Wakamatsu:** Data curation, formal analysis. **Y. Okuno:** Data curation, formal analysis. **C. Plass:** Resources, funding acquisition. **M.L. Loh:** Resources, funding acquisition. **C.M. Niemeyer:** Resources, funding acquisition. **H. Muramatsu:** Conceptualization, resources, formal analysis, supervision, funding acquisition, project administration. **C. Flotho:** Conceptualization, resources, formal analysis, supervision, funding acquisition, writing-original draft, project administration. **E. Stieglitz:** Conceptualization, resources, formal analysis, supervision, funding acquisition, writing-original draft, project administration. **D.B. Lipka:** Conceptualization, resources, formal analysis, supervision, funding acquisition, writing-original draft, project administration.

Acknowledgments

We first want to thank the patients and their families for participating in research studies that allowed for this work to be accomplished. We thank the Microarray Unit of the Genomics and Proteomics Core Facility, German Cancer Research Center (DKFZ), for providing the Illumina Infinium MethylationEPIC arrays and related services. We also want to thank all members of the Division of Cancer Epigenomics (DKFZ) and the Division of Translational Medical Oncology (DKFZ and NCT), as well as Dr. Manuel Wiesenfarth (Division of Biostatistics, DKFZ) for helpful discussions related to this study. We want to thank Oliver Mücke for his excellent technical support and Jahan Parsa of Tecan for his assistance with pipeline development. We also acknowledge the Hilda Biobank at the University of Freiburg for specimen processing. This work was supported by the German José Carreras Leukemia Foundation (DJCLS) grant DJCLS R 15/01 (to C. Plass, D.B. Lipka, and C. Flotho); the German Research Foundation (DFG) grant CRC992-C05 (to C. Flotho); the German Federal Ministry of Education and Research (BMBF) grant "MyPred" (to C. Flotho and C.M. Niemeyer); the NIH, NCI grant 1U54CA196519 (to M.L. Loh and E. Stieglitz); NIH, National Heart, Lung, and Blood Institute grant K08HL135434 (to E. Stieglitz); the Pediatric Cancer Research Foundation (to E. Stieglitz); the V Foundation (to E. Stieglitz); the UCSF Catalyst Program (to E. Stieglitz); the California Cancer League (to E. Stieglitz); Dueling for Lincoln (to E. Stieglitz and M.L. Loh); Lemonade 4 Leukemia (to E. Stieglitz and M.L. Loh); the Frank A. Campini Foundation (to E. Stieglitz and M.L. Loh); the Leukemia and Lymphoma Society grant R6511-19 (to M.L. Loh); the Center for Advanced Technologies at UCSF; and the NCI Cancer Center Support grant 5P30CA082103.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received August 14, 2020; revised October 1, 2020; accepted October 21, 2020; published first November 2, 2020.

References

1. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. Blood 2016;127:2391–405.
2. Ries LAG, Smith MA, Gurney JG, Linet M, Tamra T, Young JL, et al. Cancer incidence and survival among children and adolescents: United States SEER Program 1975–1995. NCI, SEER Program. 1999. Available from: https://seer.cancer.gov/archive/publications/childhood/.
3. Locatelli F, Niemeyer CM. How I treat juvenile myelomonocytic leukemia. Blood 2015;125:1083–90.
4. Niemeyer CM, Flotho C. Juvenile myelomonocytic leukemia: who's the driver at the wheel? Blood 2019;133:1060–70.
5. Niemeyer CM, Arico M, Basso G, Biondi A, Cantu Rajnoldi A, Creutzig U, et al. Chronic myelomonocytic leukemia in childhood: a retrospective analysis of 110 cases. European Working Group on Myelodysplastic Syndromes in Childhood (EWOG-MDS). Blood 1997;89:3534–43.
6. Stieglitz E, Taylor-Weiner AN, Chang TY, Gelston LC, Wang Y-D, Mazor T, et al. The genomic landscape of juvenile myelomonocytic leukemia. Nat Genet 2015;47:1326–33.
7. Caye A, Strullu M, Guidez F, Cassinat B, Gazal S, Fenneteau O, et al. Juvenile myelomonocytic leukemia displays mutations in components of the RAS pathway and the PRC2 network. Nat Genet 2015;47:1334–40.
8. Pérez B, Kosmider O, Cassinat B, Renneville A, Lachenaud J, Kaltenbach S, 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.
9. Sakaguchi H, Okuno Y, Muramatsu H, Yoshida K, Shiraishi Y, Takahashi M, et al. Exome sequencing identifies secondary mutations of SETBP1 and JAK3 in juvenile myelomonocytic leukemia. Nat Genet 2013;45:937–41.
10. Stieglitz E, Troup CB, Gelston LC, Haliburton J, Chow ED, Yu KB, et al. Subclonal mutations in SETBP1 confer a poor prognosis in juvenile myelomonocytic leukemia. Blood 2015;125:516–24.
11. Niemeyer CM, Kang MW, Shin DH, Furlan I, Erlacher M, Bunin NJ, et al. Germline CBL mutations cause developmental abnormalities and predispose to juvenile myelomonocytic leukemia. Nat Genet 2010;42:794–800.
12. Matsuda K, Shimada A, Yoshida N, Ogawa A, Watanabe A, Yajima S, et al. Spontaneous improvement of hematologic abnormalities in patients having juvenile myelomonocytic leukemia with specific RAS mutations. Blood 2007;109:5477–80.
13. Bresolin S, Zecca M, Flotho C, Trentin L, Zangrando A, Sainati L, et al. Gene expression–based classification as an independent predictor of clinical outcome in juvenile myelomonocytic leukemia. J Clin Oncol 2010;28:1919–27.
14. MacLeod A, Rouleau J, Szyf M. Regulation of DNA methylation by the Ras signaling pathway. J Biol Chem 1995;270:11327–37.
15. Hasegawa D, Manabe A, Kubota T, Kawasaki H, Hirose I, Ohtsuka Y, et al. Methylation status of the p15 and p16 genes in paediatric myelodysplastic syndrome and juvenile myelomonocytic leukaemia. Br J Haematol 2005;128:805–12.

16. Johan MF, Bowen DT, Frew ME, Goodeve AC, Reilly JT. Aberrant methylation of the negative regulators RASSFIA, SHP-1 and SOCS-1 in myelodysplastic syndromes and acute myeloid leukaemia. Br J Haematol 2005;129:60–5.
17. Liu YL, Castleberry RP, Emanuel PD. PTEN deficiency is a common defect in juvenile myelomonocytic leukemia. Leuk Res 2009;33:671–7.
18. Poetsch AR, Lipka DB, Witte T, Claus R, Nöllke P, Zucknick M, et al. RASA4 undergoes DNA hypermethylation in resistant juvenile myelomonocytic leukemia. Epigenetics 2014;9:1252–60.
19. Wilhelm T, Lipka DB, Witte T, Wierzbinska JA, Fluhr S, Helf M, et al. Epigenetic silencing of AKAP12 in juvenile myelomonocytic leukemia. Epigenetics 2016;11:110–9.
20. Fluhr S, Boerries M, Busch H, Symeonidi A, Witte T, Lipka DB, et al. CREBBP is a target of epigenetic, but not genetic, modification in juvenile myelomonocytic leukemia. Clin Epigenetics 2016;8:50.
21. Fluhr S, Krombholz CF, Meier A, Epting T, Mücke O, Plass C, et al. Epigenetic dysregulation of the erythropoietic transcription factor KLF1 and the β-like globin locus in juvenile myelomonocytic leukemia. Epigenetics 2017;12:715–23.
22. Olk-Batz C, Poetsch AR, Nöllke P, Claus R, Zucknick M, Sandrock I, et al. Aberrant DNA methylation characterizes juvenile myelomonocytic leukemia with poor outcome. Blood 2011;117:4871–80.
23. Stieglitz E, Mazor T, Olshen AB, Geng H, Gelston LC, Akutagawa J, et al. Genome-wide DNA methylation is predictive of outcome in juvenile myelomonocytic leukemia. Nat Commun 2017;8:2127.
24. Lipka DB, Witte T, Toth R, Yang J, Wiesenfarth M, Nöllke P, et al. RAS-pathway mutation patterns define epigenetic subclasses in juvenile myelomonocytic leukemia. Nat Commun 2017;8:2126.
25. Murakami N, Okuno Y, Yoshida K, Shiraishi Y, Nagae G, Suzuki K, et al. Integrated molecular profiling of juvenile myelomonocytic leukemia. Blood 2018;131:1576–86.
26. Assenov Y, Müller F, Lutsik P, Walter J, Lengauer T, Bock C. Comprehensive analysis of DNA methylation data with RnBeads. Nat Methods 2014;11:1138–40.
27. Triche TJ, 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.
28. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 2012;28:882–3.
29. Kolde R. pheatmap: Pretty Heatmaps. CRAN R Package; 2015. Available from: https://cran.r-project.org/web/packages/pheatmap/.
30. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics 2010;26:1572–3.
31. Hovestadt V, Zapatka M. conumee: enhanced copy-number variation analysis using Illumina DNA methylation arrays. Bioconductor R Package; 2017. Available from: https://bioconductor.org/packages/release/bioc/html/conumee.html.
32. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol 2011;12:R41.
33. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. arXiv 2016. Available from: https://cran.r-project.org/web/packages/xgboost/.
34. Kuhn M. Building predictive models in R using the caret package. J Stat Softw 2008;28. Available from: https://www.jstatsoft.org/article/view/v028i05.
35. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 2011;17:10.
36. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for bisulfite-seq applications. Bioinformatics 2011;27:1571–2.
37. Huehns ER, Beaven G. Developmental changes in human haemoglobins. In: Benson P, editor.Biochem Dev. London, United Kingdom: William Heinemann Medical Books Ltd.; 1971. p.175–203.
38. Niemeyer CM, Strahm B, Dworzak M, De Moerloose B, Hasle H, Stary J, et al. JMML revisited: role und outcome of hematopoietic stem cell transplantation in subtypes of juvenile myelomonocytic leukemia (JMML). Blood 2012;120:955.
39. Locatelli F, Algeri M, Merli P, Strocchio L. Novel approaches to diagnosis and treatment of juvenile myelomonocytic leukemia. Expert Rev Hematol 2018;11:129–43.
40. Oakes CC, Seifert M, Assenov Y, Gu L, Przekopowitz M, Ruppert AS, et al. DNA methylation dynamics during B cell maturation underlie a continuum of disease phenotypes in chronic lymphocytic leukemia. Nat Genet 2016;48:253–64.
41. Wierzbinska JA, Toth R, Ishaque N, Rippe K, Mallm J-P, Klett LC, et al. Methylome-based cell-of-origin modeling (Methyl-COOM) identifies aberrant expression of immune regulatory molecules in CLL. Genome Med 2020;12:29.
42. Roos-Weil D, Giacopelli B, Armand M, Della-Valle V, Ghamlouch H, Decaudin C, et al. Identification of 2 DNA methylation subtypes of Waldenström macroglobulinemia with plasma and memory B-cell features. Blood 2020;136:585–95.
43. Bolouri H, Farrar JE, Triche T, Ries RE, Lim EL, Alonzo TA, et al. The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. Nat Med 2018;24:103–12.
44. Nordlund J, Bäcklin CL, Wahlberg P, Busche S, Berglund EC, Eloranta M-L, et al. Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia. Genome Biol 2013;14:r105.
45. Mackay A, Burford A, Carvalho D, Izquierdo E, Fazal-Salom J, Taylor KR, et al. Integrated molecular meta-analysis of 1,000 pediatric high-grade and diffuse intrinsic pontine glioma. Cancer Cell 2017;32:520–37.
46. Capper D, Jones DTW, Sill M, Hovestadt V, Schrimpf D, Sturm D, et al. DNA methylation-based classification of central nervous system tumours. Nature 2018;555:469–74.
47. Flotho C, Kratz CP, Bergsträsser E, Hasle H, Starý J, Trebo M, et al. Genotype-phenotype correlation in cases of juvenile myelomonocytic leukemia with clonal RAS mutations. Blood 2008;111:966–7.

Published OnlineFirst November 2, 2020; DOI: 10.1158/1078-0432.CCR-20-3184

Clinical Cancer Research

International Consensus Definition of DNA Methylation Subgroups in Juvenile Myelomonocytic Leukemia

Maximilian Schönung, Julia Meyer, Peter Nöllke, et al.

Clin Cancer Res Published OnlineFirst November 2, 2020.