Comprehensive characterization of pediatric AML reveals diverse fusion oncoproteins and age-specific mutational interactions

Hamid Bolouri[1#*], Jason E Farrar[2#], Timothy Triche Jr[3#], Rhonda E Ries[4#], Emilia L Lim[5]; Todd A Alonzo[6,7]; Yussanne Ma[5]; Richard Moore[5]; Andrew Mungall[5]; Marco A Marra[5]; Jinghui Zhang[8]; Xiaotu Ma[8]; Yu Liu[8]; Yanling Liu[8]; Jaime M Guidry Auvil[9]; Tanja M Davidsen[9]; Patee Gesuwan[9]; Leandro C Hermida[9]; Bodour Salhia[10]; Stephen Capone[3]; Giridharan Ramsingh[3]; Christian Michel Zwaan[11]; Sanne Noort[11]; Stephen R Piccolo[12,13]; E Anders Kolb[14]; Alan S Gamis[15]; Malcolm A Smith[16]; Daniela S Gerhard[9]; and Soheil Meshinchi[4*]

[1]Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, WA; [2]Winthrop P Rockefeller Cancer Institute, University of Arkansas for Medical Sciences and Arkansas Children's Research Institute, Little Rock, AR; [3]Jane Anne Nohl Division of Hematology, USC/Norris Comprehensive Cancer Center, Los Angeles, CA; [4]Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA; [5]Michael Smith Genome Sciences Centre, British Columbia Cancer Agency, Vancouver, BC, Canada; [6]Keck School of Medicine, University of Southern California, Los Angeles, CA; [7]Children's Oncology Group, Monrovia, CA; [8]Division of Computational Biology, St Jude Children's Research Hospital, Memphis, TN; [9]Office of Cancer Genomics, National Cancer Institute, Bethesda, MD; [10]Department of Translational Genomics, Keck School of Medicine, University of Southern California, Los Angeles, CA; [11]Dept of Pediatric Oncology, Erasmus MC-Sophia Children's Hospital, Rotterdam, [12]Department of Biology, Brigham Young University, Provo, UT; [13]Department of Biomedical Informatics, University of Utah, Salt Lake City, UT; [14]Nemours Center for Cancer and Blood Disorders, Alfred I. DuPont Hospital for Children, Wilmington, DE; [15]Division of Hematology/Oncology/Bone Marrow Transplantation, Children's Mercy Hospitals and Clinics, Kansas City, MO; [16]Cancer Therapy Evaluation Program, National Cancer Institute, Bethesda, MD

# Equal contribution
* corresponding author

Abstract

We present the molecular landscape of pediatric acute myeloid leukemia (AML), characterizing nearly 1,000 participants in Children's Oncology Group (COG) AML trials. The COG/NCI TARGET AML initiative assessed cases by whole-genome, targeted DNA, mRNA, miRNA sequencing and CpG methylation profiling. Validated DNA variants revealed diverse, infrequent mutations. Fewer than 40 genes are mutated in >2% of cases. In contrast, somatic structural variants, including novel gene fusions and focal *MBNL1*, *ZEB2*, and *ELF1* deletions, disproportionately impact younger patients. *DNMT3A* and *TP53* mutations, common in adults, are conspicuously absent from virtually all pediatric cases. Novel *GATA2*, *FLT3*, and *CBL* mutations, recurrent *MYC*-ITD, *NRAS*, *KRAS*, and *WT1* mutations are frequent. Deletions, mutations, and promoter hypermethylation convergently impact Wnt signaling, Polycomb repression, innate immune cell interactions, and a cluster of zinc finger genes associated with *KMT2A* rearrangements. These results highlight the need for, and facilitate the development of, age-appropriate targeted therapies in pediatric AML.

Acute leukemia is the most common form of childhood cancer[1], and its incidence is increasing. Despite constituting only 20% of pediatric acute leukemia, AML is overtaking acute lymphoblastic leukemia as the leading cause of childhood leukemic mortality, in part because current prognostic schemas classify many children who will ultimately succumb to their disease as low- or intermediate-risk. Additionally, aside from investigational tyrosine kinase inhibitors for *FLT3*-activated AML, targeted therapies are not used in pediatric AML. Both problems stem from an inadequate understanding of the biology of childhood AML.

AML is a molecularly heterogeneous group of diseases affecting patients of all ages. Recent genome-scale studies have revealed novel, potentially targetable mutations prevalent in adult *de novo* AML[2-4]. However, the relevance of these findings to childhood AML remains unclear, since several of the most common adult mutations appear far less prevalent in pediatric AML[5,6].

To date, no comprehensive characterization of pediatric AML has been described. Here, we report the initial results of the TARGET (Therapeutically Applicable Research to Generate Effective Treatments) AML initiative, a collaborative COG/NCI project to comprehensively characterize the mutational, transcriptional, and epigenetic landscapes of a large, well-annotated cohort of pediatric AML.

Results

Overview of cohort characteristics

A total of 1023 children enrolled in COG studies are included in the TARGET AML dataset. Comprehensive clinical data, including clinical outcomes and test results for common sequence aberrations (outlined in **Table S1**), are available for 993 patients. Of these, 815 subjects were profiled for somatic mutations at presentation: 197 by whole-genome sequencing (WGS), and 800 by targeted capture sequencing (TCS), at read depths averaging 500x, for validation of mutations identified by WGS. The WGS discovery cohort of diagnostic and remission (germline comparison) specimens were selected from patients treated on recent COG studies who achieved an initial remission to induction chemotherapy. These trials randomized type or timing of induction therapy (CCG-2961)[7] and the

addition of gemtuzumab ozogamicin in a single arm pilot (AAML03P1)[8] or randomized fashion (AAML0531)[9]. Specimens for TCS validation were obtained from 800 patients, including 182 from the WGS discovery cohort (153 with matched remission samples). A complete listing of cases and their characterization is available in the TARGET sample matrix (https://ocg.cancer.gov/programs/target/data-matrix). The age at presentation of TARGET AML participants ranged from 8 days to 29 years (median 10 years, **Fig. 1a**). Infants, children and adolescents/young adults (AYA) differ broadly by cytogenetic and clinical risk-group classifications (**Fig. 1a**, multivariate Chi-squared $p<10^{-22}$), consistent with observed differences in clinically-evaluated structural abnormalities and mutations (summarized in **Fig. 1b**). Notably, among these clinically detected abnormalities, only 5 mutations and 5 structural aberrations occur in more than 5% of patients.

We validated each class of somatic DNA sequence alteration discovered by WGS through secondary assays (**Figs. 1c** and **S1**). Single nucleotide variants (SNVs) and short insertions and deletions (indels) were confirmed by TCS of the coding sequences of genes identified as recurrently altered in the WGS studies. WGS-detected copy number alterations were confirmed by GISTIC2.0 scores from SNP arrays; WGS-detected structural changes (such as translocations and inversions) were confirmed by RNA-seq and clinical data. Across variant types, we find >70% concordance between at least two assays. These variants are referred to as verified variants hereon. An overview of the multiplatform-verified somatic DNA variants in 684 patients is presented in **Fig. 2a**. Roughly a quarter of patients possess normal karyotype, yet nearly all revealed at least one recurrent verified somatic DNA alteration, and at least 12 common cancer-associated cellular processes are recurrently impacted (**Fig. S2, Table S2a, b**).

We carried out analyses of microRNA, mRNA, and/or DNA methylation in 412 subjects. A summary of the assays performed and case-assay overlap is presented in **Fig. S3**. We compared our verified variants to those of 177 adult AML cases from the Cancer Genome Atlas (TCGA) project[2], stratified by the age groupings outlined in **Fig. 1a**. The TARGET and TCGA discovery cohorts both contained numerous AYA patients (**Table S3**). Importantly, our conclusions regarding the molecular characteristics of this age group are identical when analyzing either or both cohorts (**Fig. S4**).

Somatic gene mutations in pediatric AML

Like adult AML, pediatric AML has one of the lowest rates of mutation among molecularly well-characterized cancers (**Fig. S5**), with < 1 somatic, protein-coding change per megabase in most cases. However, the landscape of somatic variants in pediatric AML is markedly different from that reported in adults[2,3] (**Figs. 2b, S6-S7, Table S4**). *RAS*, *KIT*, and *FLT3* alterations, including novel, pediatric-specific *FLT3* mutations (FLT3.N), are more common in children. Mutational burden increases with age, yet older patients have relatively fewer recurrent cytogenetic alterations. Indeed, the number of coding SNVs, within and across cohorts, is best predicted by age (**Fig. 2c**, $p<10^{-15}$) and by cytogenetic subgroup. In contradistinction to the higher prevalence of small sequence variants in older patients, recurrent structural alterations, fusions, and focal copy number aberrations are more common in younger patients (**Fig. 2d-e**, $p<10^{-3}$, see below). Patients with *CBFA2T3-GLIS2*, *KMT2A,* or *NUP98* fusions tend to have fewer mutations ($p<10^{-9}$), with subgroups demonstrating inferior clinical outcome (**Fig. S8**). Patients with core binding factor rearrangements tend to have more mutations than expected for their age ($p<10^{-15}$), yet more favorable outcomes. The mutational spectrum of coding SNVs (**Fig. S5**) accumulates C→T transitions with age ($p<10^{-3}$), with additional C→A transversions in t(8;21) ($p<10^{-2}$) and aberrant karyotype ($p<10^{-2}$) patients.

After adjustment for cytogenetics and multiple comparisons, *NRAS* ($p<10^{-3}$) and *WT1* ($p<10^{-3}$) are mutated significantly more often in younger patients, while *DNMT3A* ($p<10^{-23}$), *IDH1/2* ($p<10^{-4}$), *RUNX1* ($p<10^{-4}$), *TP53* ($p<10^{-4}$), and *NPM1* ($p<0.03$) are mutated significantly more frequently in older patients. *KRAS, CBL, GATA2, SETD2,* and *PTPN11* mutations appear to be more common in younger patients ($0.05<p<0.1$, adjusted, **Figs. S7** and **S9**). We identified a prominent hotspot of *MYC* alterations[10] and previously unreported internal tandem duplications appearing exclusively in children (**Fig. S7**). These observations are replicated in an independent ECOG cohort (**Fig. S10a**) of 384 adult AML patients[4]. Since gene fusions have characteristic cooperating mutations[11], we devised a weighted resampling scheme to compare mutation frequencies in 584 TARGET and 131 TCGA AML cases while controlling for karyotypic associations. The results (**Fig. S10b**) confirm the generality of the pediatric-adult differences identified above.

For genes such as *CBL, GATA2, WT1, MYC* and *FLT3*, both the frequency and the sites of mutation often differ between children and adults (**Figs. 3a** and **S7**), with multiple, frequently recurrent, alterations distinct from those identified in adult AML. *RAS*-related mutations (mutant *KRAS*, *NRAS*, *PTPN11*, or *NF1*) are common, particularly with *KMT2A* fusions (**Fig. S11, Tables S4-S6**). In addition to being more common and varied, *WT1* mutations appear more likely to be of clonal origin in younger patients (**Fig. 3b**) despite the majority of pediatric patients presenting with multiple detectable sub-clones (**Fig. S12**).

These differences are clinically significant: we have previously shown that novel *FLT3* mutations are functional, and yield poor responses to standard therapy[12]. The established adverse impact of *FLT3*-ITDs on survival is significantly modulated by co-occurring variants, including *WT1* and *NPM1* mutations and *NUP98* translocations. As shown in **Figs. 3c and S13-S14**, three independent, large-scale studies demonstrate that *FLT3*-ITD accompanied by *NPM1* mutations is associated with relatively favorable outcomes in pediatric patients, while *FLT3*-ITD with *WT1* mutations and/or *NUP98-NSD1* fusions yields poorer outcomes than *FLT3*-ITD alone.

We found no coding mutations of *DNMT3A* in pediatric AML, despite its frequency in adults. Spontaneous deamination of 5-methylcytosine is strongly associated with aging, and *DNMT3A* contains a CpG dinucleotide yielding hotspot R882 mutations by C-to-T deamination[13]. *DNMT3A* also directly interacts with *TP53*[14], itself impacted far more frequently in adults. Mutations of *DNMT3A* or *TP53* drive clonal hematopoiesis in many apparently healthy adults[15] but are rare in children, as are the *IDH1* and *IDH2* mutations with which they often co-occur.

The spectrum of somatic structural DNA changes in pediatric AML

Many pediatric AML cases harbor chromosomal copy number changes distinct from those reported in adults (**Fig. 4a**). Among the 197 cases assayed by WGS, we identified 14 novel focal deletions involving *MBNL1*, a splicing regulator, or *ZEB2*, a key regulator of normal[16] and leukemic[17] hematopoiesis (**Fig. S15**). Despite occurring on separate chromosomes, in regions devoid of other deletions, *MBNL1*:*ZEB2* co-deletions occur far more often than expected ($p<10^{-13}$). Half of these accompany *KMT2A-MLLT3* fusions ($p=0.035$, **Fig. S11, Tables S5-S6**). Samples with *MBNL1*:*ZEB2* co-deletions carry a larger number of recurrent mutations ($p=0.015$), and *KMT2A*-fusion samples with del(MBNL1) or del(*ZEB2*) have a larger number of additional cytogenetic abnormalities ($p<0.0005$). Another 15 novel, validated focal deletions specifically impact *ELF1*, an ETS-family transcriptional regulator of *MEIS1*[18]. A statistically significant difference in *ELF1* mRNA expression exists between *ELF1*-deleted and intact samples ($p<0.01$), with 63 genes differentially expressed between the two groups ($p<0.01$, **Fig. S16**). Among

other novel recurrent copy losses, we note five heterozygous deletions of a region containing the *IL9R* gene (**Table S5**) co-occurring with *KIT* mutations and t(8;21).

Consistent with our previous findings regarding *NUP98-NSD1* fusions[19], an expansive catalog of gene fusions, many observed primarily or exclusively in pediatric cases, underscores the disproportionate impact of structural variants in younger patients (**Figs. 4b** and **S17-S18**). But patterns of exclusion and cooperation are not limited to patients with recurrent structural alteration*:* mutant *GATA2* is frequently seen in children with normal karyotype (NK) AML, and both *GATA2* ($p<10^{-9}$) and *CSF3R* ($p<10^{-6}$, **Fig. S19**) mutations co-occur with mutations of *CEBPA*[20]. *GATA2* and *CEBPA* are key regulators of hematopoiesis[21,22], both interacting with *RUNX1* in normal hematopoiesis and leukemogenesis[23]. As with *FLT3*/*NUP98-NSD1*/*WT1* interactions, these findings show prognostic interactions in pediatric AML outcome (**Fig. S19b**). *RUNX1* mutations and *RUNX1-RUNX1T1* gene fusions are significantly exclusive of *GATA2* and *CEBPA* mutations (p=0.006, **Fig. S20, Table S7**). All four are significantly exclusive of *KMT2A* rearrangements ($p<10^{-15}$), *CBFB-MYH11* gene fusions ($p<10^{-11}$), and *ETV6* aberrations (p=0.01).

DNA methylation subtypes in pediatric AML

As summarized in **Fig. 4c**, aberrations affecting epigenetic regulators are widespread and rarely overlap in AML, but their origin (structural vs. mutational) and frequency differs between children and adults. Combining DNA methylation and mRNA expression results in 456 TARGET and TCGA AML cases, we identified dozens of genes with recurrent transcriptional silencing via promoter hypermethylation across TARGET and TCGA AML patients (**Figs. 5a and 5c, Tables S8-S9,** details in **Figs. S21-S22**). A number of samples exhibited widespread silencing of genes by aberrant promoter hypermethylation, and this group is enriched for younger patients with *WT1* mutations (p=0.0012, **Fig. 5a**, hyper-silenced group). Aberrant Wnt/β-catenin signaling is required for the development of leukemic stem cells[24], and one or more of the Wnt pathway regulators *DKK1*, *SMAD1, SMAD5, SFRP4, SFRP5, AXIN2, WIF1, FZD3, HES1,* or *TLE1* is deleted or aberrantly methylated in most AML cases[25]. Repression of activating NK cell ligands (particularly *ULBP1/2/3*) appears to be common in pediatric patients, which may represent a therapeutic target[26]. In *KMT2A*-rearranged patients, a cluster of poorly characterized zinc finger genes on chromosome 19 is recurrently silenced.

We applied non-negative matrix factorization (NMF) to CpG methylation data from 284 TARGET and TCGA AML patients with DNA methylation data. By cross-validation, we identified 31 signatures (**Table S10**) that best captured methylation differences across samples, after controlling via *in silico* purification for differences in cellularity. Unsupervised clustering of the resulting DNA methylation signatures largely separated patients by age and karyotypic subtypes (**Figs. 5b and S23**), but also revealed a signature which did not associate strongly with age or established prognostic factors (Signature 13, **Fig. 5b**). Two signatures (signatures 2 and 13) predicted significantly ($p < 0.05$) poorer event-free survival in both pediatric and adult patients with above-median scores, after stratifying by cohort and adjusting for *TP53* mutation status and white blood cell count (**Fig. S24**). Larger sample sizes are needed to evaluate the clinical significance of these findings.

The pediatric AML transcriptome is shaped by diverse miRNAs

We performed miRNA sequencing of 152 cases to characterize miRNA expression patterns in pediatric AML. Unsupervised clustering of the data revealed 4 discrete subgroups that were correlated with specific genomic alterations (**Figs. 6a** and **S25**), including high miR-10a expression in samples with *NPM1* mutations, consistent with previous reports[27]. Further, Cox proportional hazards analyses identified

multiple miRNAs associated with clinical outcome (**Figs. S26-S28, Table S11**), including miR-155, which we previously reported to predict poor survival[28].

Differential expression analyses using Wilcoxon tests revealed miRNAs that are differentially expressed between pediatric and adult AML (**Fig. 6b**). Of note, miR-330 was the most over-expressed in pediatric samples, and has previously been shown to have oncogenic potential in AML[29].

Several age-associated miRNAs harbor binding sites within, and have expression levels anti-correlated with, putative target genes that may be involved in RNA and protein processing suggesting that miRNAs could contribute to leukemogenesis through the dysregulation of transcripts and proteins[30]. Of note, *let-7b,* which is a potential regulator of protein synthesis via *EIF2S3* (**Fig. 6c**), is typically less abundantly expressed in pediatric AML (**Fig. 6d**). However, high *let-7b* expression in pediatric AML is associated with shorter time to relapse (log-rank $p<0.05$, **Fig. 6e**).

Discussion

Using a large cohort of patients, this study establishes the prevalence of, and coincident relationships among, recurrent somatic genetic and epigenetic abnormalities in pediatric AML. We observe several features in common between pediatric and adult AML: a low overall mutation rate in comparison to other cancers, a long tail of infrequently affected genes, and overlap among recurrently impacted genes. However, pediatric AML exhibits distinctive and critically important characteristics. We and others have previously reported on the presence and clinical impact of novel fusion genes in pediatric AML[19,31]. As this study illustrates, the impact of fusion transcripts in AML is both broad and age-dependent. Recognition and comprehensive testing for these alterations are key first steps in the development of new and potentially novel modes of targeted therapy[32].

Recurrent focal deletions represent a unique aspect of pediatric AML. Regional (e.g. chromosomal arm- and band-level) copy loss differs substantially by age, but surprisingly, focal areas of copy loss are also more common in children, specifically impacting *ZEB2*, *MBNL1*, and *ELF1*. *MBNL1* is upregulated by the *KMT2A-AF9* fusion protein[33], and genes involved in post-transcriptional processing (*SETD2, U2AF1, DICER1*) harbor the sole recurrent mutation in several *KMT2A*-rearranged cases, suggesting a functional role for altered splicing in pediatric leukemogenesis. Alterations in *ZEB2* have been identified as cooperating events in murine *CALM-AF10* leukemia models[34] while *ZEB2* knockout mice develop myelofibrosis[35], suggesting a fundamental role for this gene in the pathogenesis of AML.

Many of the genes characteristically mutated in AML are altered at widely variable frequencies across age groups; several (including *FLT3* and *WT1*) are impacted by pediatric specific variants and hotspots. Clinical tests for a handful of genomic alterations are widely used to risk-stratify patients and determine treatment regimens. However, the current practice of considering the effect of each somatic alteration in isolation is inadequate. As we illustrate for *FLT3*-ITD, interactions among sequence variants can have dramatic clinical consequences. Moreover, some interactions appear to be age-specific. In pediatric AML, *FLT3*-ITD and *NPM1* mutations co-occur in the absence of *DNMT3A* mutations in a group of patients with superior outcomes (**Figs. 3c, S13 and S14**), in contrast to inferior outcomes reported in adults where *FLT3*-ITD and *NPM1* mutations frequently co-occur with mutations in *DNMT3A*[3]. In the

TCGA adult AML cohort, over half the subjects with somatic *FLT3* and *NPM1* mutations also possessed somatic *DNMT3A* mutations[2]. Subsequent studies established the generality of this result[3], and revealed that *DNMT3A* mutations are early clonal events[36], which often co-operate with later *NPM1* and *FLT3* mutations to promote chemoresistance, mutagenesis,[37] and inferior outcomes[38]. Similarly, the co-occurrence of *FLT3*-ITD with *WT1* mutations or *NUP98-NSD1* fusions accompanies frequent induction failure and dismal outcomes in children with AML (multivariate $p<10^{-4}$, **Figs. 3c, S13 and S14**).

In TARGET, TCGA, and ECOG AML cases, *WT1* mutations were mutually exclusive with those of *ASXL1* and *EZH2* ($p < 10^{-3}$). *WT1* recruits *EZH2* to specific targets[39], and *WT1* mutations have been linked to hypermethylation of *EZH2* target genes[40]. Mutant *ASXL1* abolishes *EZH2*-mediated silencing of HOX genes[41]. *EZH2* resides on a recurrently deleted region of chromosome 7, and decreased *EZH2* activity is associated with treatment resistant AML[42]. In pediatric AML, mutant *WT1* and *EZH2* appear to be of exclusively clonal or near-clonal origin, with nearly a quarter of TARGET cases harboring mutations affecting one or the other. Aberrant *WT1*, *EZH2*, or *ASXL1* predicted induction failure in TARGET AML cases (multivariate $p<0.05$, adjusted for interactions with *FLT3* alterations*, NUP98-NSD1,* and *KMT2A* fusions) and were largely mutually exclusive with *KMT2A* rearrangements ($p < 10^{-5}$). Many of these patients present without apparent chromosomal abnormalities at diagnosis, yet less than 20% achieve long-term remission with standard treatment, highlighting the importance of molecular stratification to achieve better outcomes. It is possible that early events such as *WT1* mutations and *NUP98-NSD1* fusions in children may play a similar role to that observed for *DNMT3A* mutations[13] in adults, with significant implications for risk stratification in AML across age groups.

Our data also demonstrate that DNA-methylation and miRNA expression profiles both accompany and complement DNA alterations, and can stratify pediatric AML patients in terms of both overall and progression-free survival. These findings suggest a need to update pediatric AML clinical risk categories beyond current classifications, with important implications for clinical practice.

Despite incremental improvements with increasingly intensified regimens, modern outcomes in pediatric AML have plateaued, with only ~60% of patients achieving long term survival. As many as 10% of children will die from direct complications of treatment. Survivors suffer unacceptably high rates of long-term morbidities resulting from anthracycline exposure or sequelae of hematopoietic stem cell transplantation. As illustrated herein, pediatric AML is a collection of molecularly diverse diseases with similar phenotypes. No single treatment strategy is likely to be effective for all pediatric AML subtypes, which may explain repeated failures of randomized clinical trials to improve outcomes in recent years. In keeping with the shift towards comprehensive, molecularly based classification schemas in AML[3], the time has come to develop targeted therapies that address specific vulnerabilities of pediatric subtypes. The TARGET AML dataset will serve as a foundation for development of pediatric-specific classification schemas and the development of personalized treatment strategies.

Acknowledgements

Dedicated to the memory of our colleague, mentor and friend, Dr. Robert Arceci, whose vision and perseverance set this effort in motion: “I may not have gone where I intended to go, but I think I have ended up where I needed to be.” (Douglas Adams, *The Long Dark Tea-Time of the Soul*)

The results published here are based upon data generated by the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) initiative and The Cancer Genome Atlas. Data used for this analysis are available under dbGaP accession numbers phs000465 and phs000178. The TARGET initiative is supported by NCI Grant U10CA98543. Work performed under contracts from the National Cancer Institute, US National Institutes of Health within HHSN261200800001E includes specimen processing (the Children's Oncology Group Biopathology Center), whole genomic sequencing (Complete Genomics) and RNA and targeted capture sequencing (British Columbia Cancer Agency). The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. Computation for the work described in this paper was supported in part by Fred Hutchinson Scientific Computing, the University of Southern California's Center for High-Performance Computing, and NSF award ACI-1341935. This work was additionally supported by COG Chairs U10CA180886 and U10CA98543; COG Statistics and Data Center U10CA098413 and U10CA180899; COG Specimen Banking U24CA114766; R01CA114563 (SM); St. Baldrick's Foundation (JEF, TT, SM); P20GM121293 (JEF); Arkansas Biosciences Institute (JEF), and the Jane Anne Nohl Hematology Research Fund (TT).

Figure Legends

Figure 1. An overview of the TARGET AML study. (**a**) The distribution of subjects by clinical risk category and cytogenetic classification is shown adjacent to each age group analyzed (Infant, <3 years; Child, 3 to <15 years; Adolescent/Young Adult (AYA), 15 to <40 years). (**b**) A summary of the clinically established molecular aberrations in the cohort (n=993) is illustrated. FLT3.ITD, *FLT3* internal tandem duplications, FLT3PM, *FLT3* D835 point mutations. (**c**) Overview of the genomic variant discovery, verification, and validation process. We characterized diagnostic and remission (taken as germline) samples from 197 patients using whole genome sequencing (WGS) and verified 153 diagnostic/remission case pairs by targeted capture sequencing (TCS) of genes recurrently impacted in the WGS samples (an additional 29 WGS cases were verified by TCS of diagnostic cases only, see **Fig. S1**). 72% of WGS SNVs, and 76% of WGS indels were confirmed by TCS (red & green text in figures). For focal copy number (CN) alterations spanning fewer than 7 genes, 75% of recurrent WGS deletion/loss and 85% gain/amplification calls matched recurrent alterations discovered by SNP6 arrays in 96 matching samples. For chromosomal junctions, we integrated WGS, clinical and RNA-seq data by majority vote, confirming 89% of WGS junction calls.

Figure 2. Age-related differences in mutational and structural alterations in AML. (a) Distribution of variants per sample. At least one variant impacting a gene recurrently altered in pediatric AML was identified by multi-platform validated variants in 684 patients. Junction, protein fusions (see methods); chromCNV, chromosomal arm/band level copy variant; focalCNV, gene level copy variant; indel, small insertion/deletion; SNV, single nucleotide variant. (**b**) Age-dependent differences in the prevalence of mutations. *FLT3* mutations are plotted in 3 categories: internal tandem duplication (ITD; FLT3.ITD), activation loop domain (FLT3.C), and novel, childhood-specific changes (FLT3.N). (**b, inset**) A pattern of waxing or waning mutation rates across age groups is evident in selected genes (*KRAS* and *NPM1* illustrated). (**c**) Childhood AML, like adult AML, has a low somatic mutational burden (top and **Fig. S5**), but is more frequently impacted by common cytogenetic alterations (lower section). For color key, see legend at bottom-right. (**d**) The ratio of the burden of structural variation to SNVs/indels is high in infancy and early childhood and declines with age. For color key, see legend at bottom-right. (**e**) Using a sliding-window approach to account for uneven sampling by age, the incidence of common translocations in AML is age-specific (multi-variate Chi-squared $p < 10^{-30}$), and greatest in infants compared to all other ages (Chi-squared $p < 10^{-22}$). *KMT2A* fusions are most common in infants (Chi-squared $p < 10^{-20}$), while core binding factor fusions tend to affect older children (Chi-squared $p < 10^{-7}$).

Figure 3. Biological and prognostic interactions between alterations of *WT1*, *NPM1*, *FLT3*-ITD and *NUP98-NSD1*. (**a**) *WT1* mutations appear more frequently (TCGA: 7.3%, 13 alterations among 177 patients; TARGET: 18.4%, 150 alterations among 815 patients, Fisher's exact p = 0.0002) and impact novel sites in childhood AML. (**b**) Inference of the clonal origin of selected mutations in 197 TARGET AML cases with WGS and 177 TCGA AML cases. (**c**) The clinical impact of *FLT3*-ITD is modulated by other sequence aberrations. 963 TARGET patients had complete data for *FLT3*-ITD, *NPM1* and *WT1* mutation and *NUP98-NSD1* fusions. Patients with *FLT3*-ITD plus *WT1* and/or *NUP98-NSD1* fusion (n=73) exhibit markedly inferior event-free (multivariate p<0.001) and overall survival (see **Fig. S13**), while co-occurrence of *NPM1* mutations with *FLT3*-ITD associates with improved survival. These findings are confirmed by two separate studies from which TARGET cases were selected (AAML0531 and CCG-2961)

as well as an independent cohort of patients treated on European cooperative group trials (DCOG, see online methods).

Figure 4. Chromosomal alterations in AML. (**a**) Patterns of regional and chromosomal gain (outward projection) and loss (inward projection) in the TARGET and TCGA AML cohorts. Losses of 5q, 7, and 17 predominate in adults, while gains of 4, 6, 19, and losses of 9, X, and Y are more common in younger patients. Chromosome numbers are printed on the outside and inside of the circle plot, and colored where there are large pediatric-adult differences. (**b**) Age-specific distributions of validated gene fusions. The fraction of events within an age group for each fusion pair is indicated by white-red shading, while the color of the fusion labels indicates the primary cytogenetic group (colors same as in **Fig. 1a**, see also **Figs. S17-S18**). The number in each box indicates the number of patients carrying the indicated translocation (labels at left). (**c**) Structural and mutational aberrations affecting epigenetic regulators in TARGET (WGS) and TCGA AML cohorts.

Figure 5. Aberrant DNA methylation in adult and pediatric AML. (**a**) Integrative analysis of genes with recurrent mutations, deletions, or transcriptional silencing by promoter hypermethylation (rows) in TARGET and TCGA AML cases (columns). Cluster associations are labeled at the top, including a prominent group enriched for younger patients with *WT1* mutations (p=0.0012) that shows extensive transcriptional silencing across dozens of genes (blue boxed region, Hypersilenced). The cytogenetic group, *IDH1/2* mutation status (gray, mutated; white, wild-type or unknown) and TARGET/TCGA cohort membership for each sample is indicated below the main figure. The top marginal histogram indicates the total number of genes impacted for each patient. Gene/cytogenetic associations are shown to the right of the main figure, where per gene-rate of involvement by cytogenetic class is indicated by color and shading (unfilled = no involvement; full shading = maximum observed involvement of any gene within patients of the indicated cytogenetic grouping). Wnt regulators and activating NK cell ligands (e.g. *DKK1, WIF1* and *ULBP1, ULBP2, ULBP3,* respectively) are silenced across cytogenetic subtypes (labeled at far right). Distinct groups of silenced genes are also associated with *IDH1* or *IDH2* mutant patients and in *KMT2A*-rearranged patients. A subset of genes (56 of 119) altered in >3 patients and of patients (n=310; 168 TARGET, 142 TCGA subjects) with one or more genes silenced by promoter methylation is illustrated (see **Figs. S21-S22** and **Tables S8-S9** for enumeration of all 119 genes in all 456 evaluable subjects.). (**b**) A subset (16 of 31) of DNA methylation signatures derived by non-negative matrix factorization (NMF) and *in silico* purification, with samples ordered by hierarchical clustering of signatures (labeled at right). Genomic associations are indicated to the left of the main panel. Signature 13 does not correspond directly to known recurrent alterations, however, along with signature 2 displays potential prognostic significance (see **Fig. S24**). The patient-specific score matrix and display of all 31 signatures are provided in **Table S10** and **Fig. S23**. (**c**) Examples of expression/promoter methylation relationships for *IL2RA* and *SFRP5*, 2 genes identified as recurrently silenced (panel **a**) which also contribute to NMF signatures (panel **b**) are shown. Y-axis: transformed expression (asinh(TPM)), x-axis: promoter CpG methylation. The vertical red line indicates the empirically established silencing threshold.

Figure 6. miRNAs differentially regulate distinct molecular and age sub-groups in AML (**a**) Unsupervised clustering of miRNA expression patterns in 152 childhood AML cases identifies four patient subgroups (colored bands at top) with correlation to somatic alterations as indicated (blue bars on gray background), and subgroup-specific miRNA expression (miR-10 and miR-21 are highlighted as examples). **(b)** Age-related differences in miRNA expression are evident between adult (n=162) and pediatric AML (n=152). Volcano plot indicates differentially expressed miRNAs between adult and

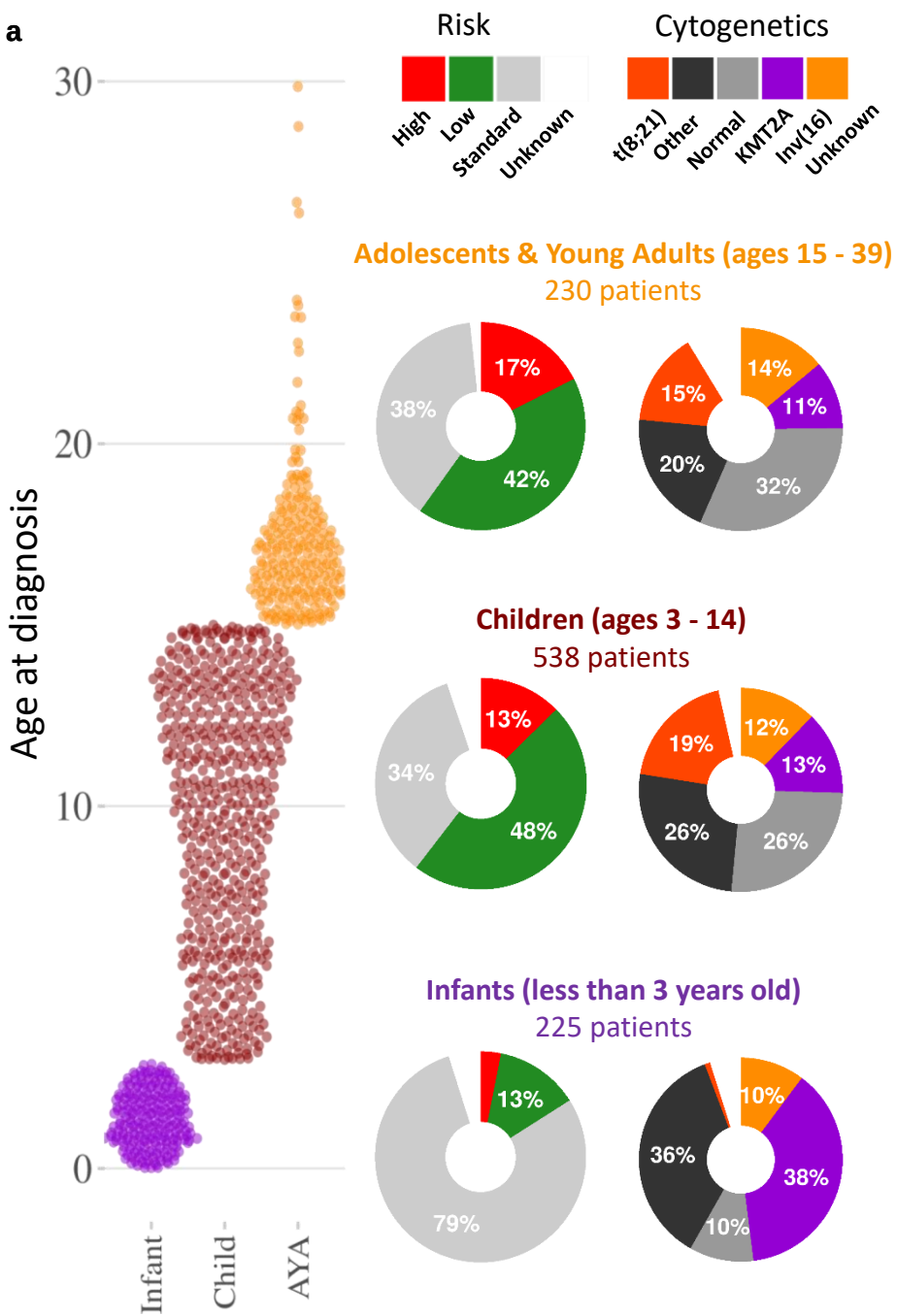
pediatric cases. Red-green point shading indicates relative under- or over-expression in TARGET, respectively. (Wilcoxon test, Benjamini-Hochberg adjusted $P<0.05$; Threshold indicated by dashed red line). **(c)** A predicted miRNA:mRNA target relationship involving *let-7b*, which is (**d**) less abundant in most pediatric cases than in adult cases. **(e)** High expression of *let-7b* occurs in a minority of pediatric AML and is associated with shorter time to relapse.

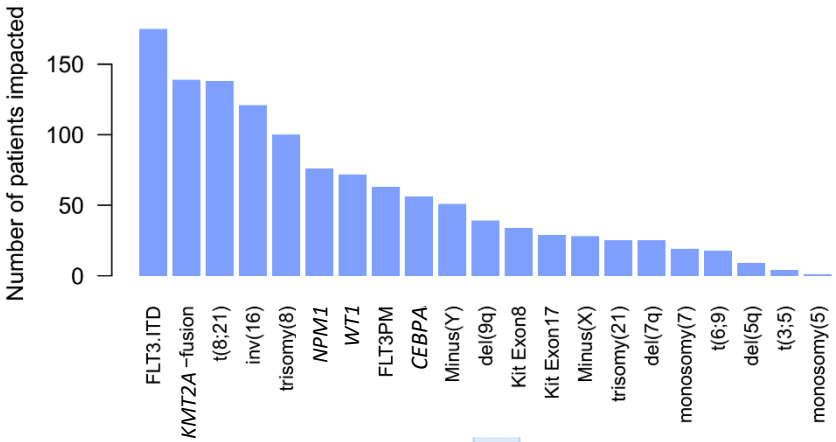
References

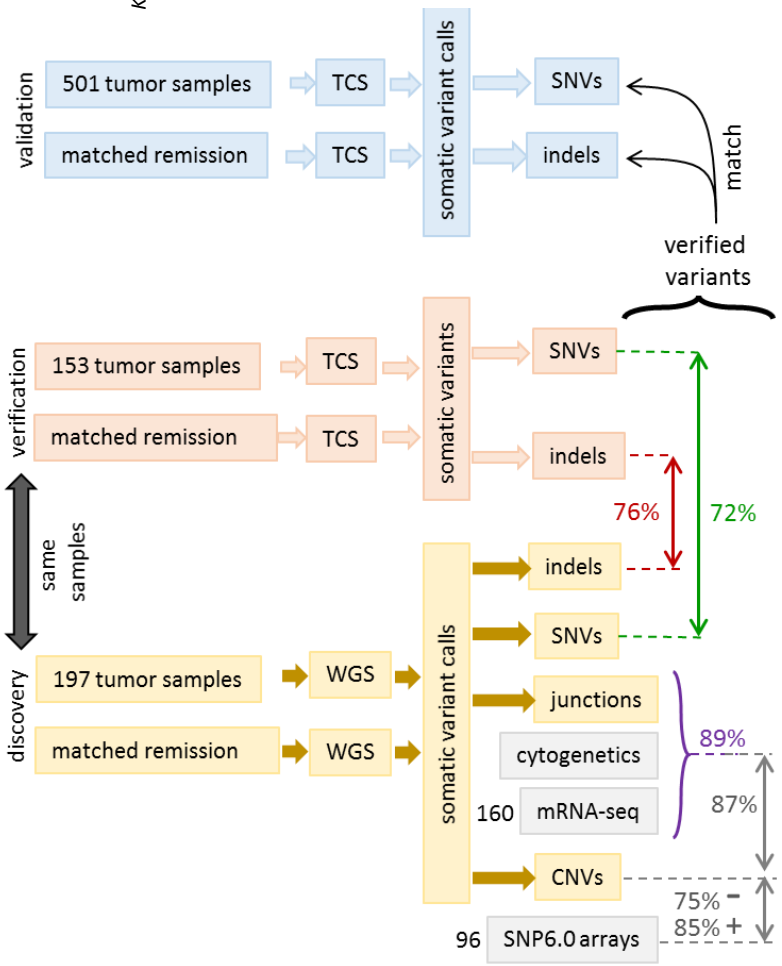
1. Steliarova-Foucher, E.*, et al.* International incidence of childhood cancer, 2001-10: a population-based registry study. *Lancet Oncol* **18**, 719-731 (2017).
2. Cancer Genome Atlas Research, N. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med* **368**, 2059-2074 (2013).
3. Papaemmanuil, E.*, et al.* Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N Engl J Med* **374**, 2209-2221 (2016).
4. Patel, J.P.*, et al.* Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. *N Engl J Med* **366**, 1079-1089 (2012).
5. Ho, P.A.*, et al.* Leukemic mutations in the methylation-associated genes DNMT3A and IDH2 are rare events in pediatric AML: a report from the Children's Oncology Group. *Pediatr Blood Cancer* **57**, 204-209 (2011).
6. Farrar, J.E.*, et al.* Genomic Profiling of Pediatric Acute Myeloid Leukemia Reveals a Changing Mutational Landscape from Disease Diagnosis to Relapse. *Cancer Res* **76**, 2197-2205 (2016).
7. Lange, B.J.*, et al.* Outcomes in CCG-2961, a children's oncology group phase 3 trial for untreated pediatric acute myeloid leukemia: a report from the children's oncology group. *Blood* **111**, 1044-1053 (2008).
8. Cooper, T.M.*, et al.* AAML03P1, a pilot study of the safety of gemtuzumab ozogamicin in combination with chemotherapy for newly diagnosed childhood acute myeloid leukemia: a report from the Children's Oncology Group. *Cancer* **118**, 761-769 (2012).
9. Gamis, A.S.*, et al.* Gemtuzumab ozogamicin in children and adolescents with de novo acute myeloid leukemia improves event-free survival by reducing relapse risk: results from the randomized phase III Children's Oncology Group trial AAML0531. *J Clin Oncol* **32**, 3021-3032 (2014).
10. Lavallee, V.P.*, et al.* Identification of MYC mutations in acute myeloid leukemias with NUP98-NSD1 translocations. *Leukemia* **30**, 1621-1624 (2016).
11. Faber, Z.J.*, et al.* The genomic landscape of core-binding factor acute myeloid leukemias. *Nat Genet* **48**, 1551-1556 (2016).
12. Tarlock, K.*, et al.* Discovery and Functional Validation of Novel Pediatric Specific FLT3 Activating Mutations in Acute Myeloid Leukemia: Results from the COG/NCI Target Initiative. *Blood* **126**, 87-87 (2015).
13. Ley, T.J.*, et al.* DNMT3A mutations in acute myeloid leukemia. *N Engl J Med* **363**, 2424-2433 (2010).
14. Wang, Y.A.*, et al.* DNA methyltransferase-3a interacts with p53 and represses p53-mediated gene expression. *Cancer Biol Ther* **4**, 1138-1143 (2005).
15. Genovese, G.*, et al.* Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med* **371**, 2477-2487 (2014).
16. Goossens, S.*, et al.* The EMT regulator Zeb2/Sip1 is essential for murine embryonic hematopoietic stem/progenitor cell differentiation and mobilization. *Blood* **117**, 5620-5630 (2011).
17. Goossens, S.*, et al.* ZEB2 drives immature T-cell lymphoblastic leukaemia development via enhanced tumour-initiating potential and IL-7 receptor signalling. *Nat Commun* **6**, 5794 (2015).
18. Xiang, P.*, et al.* Identification of E74-like factor 1 (ELF1) as a transcriptional regulator of the Hox cofactor MEIS1. *Exp Hematol* **38**, 798-798, 808 e791-792 (2010).
19. Ostronoff, F.*, et al.* NUP98/NSD1 and FLT3/ITD coexpression is more prevalent in younger AML patients and leads to induction failure: a COG and SWOG report. *Blood* **124**, 2400-2407 (2014).

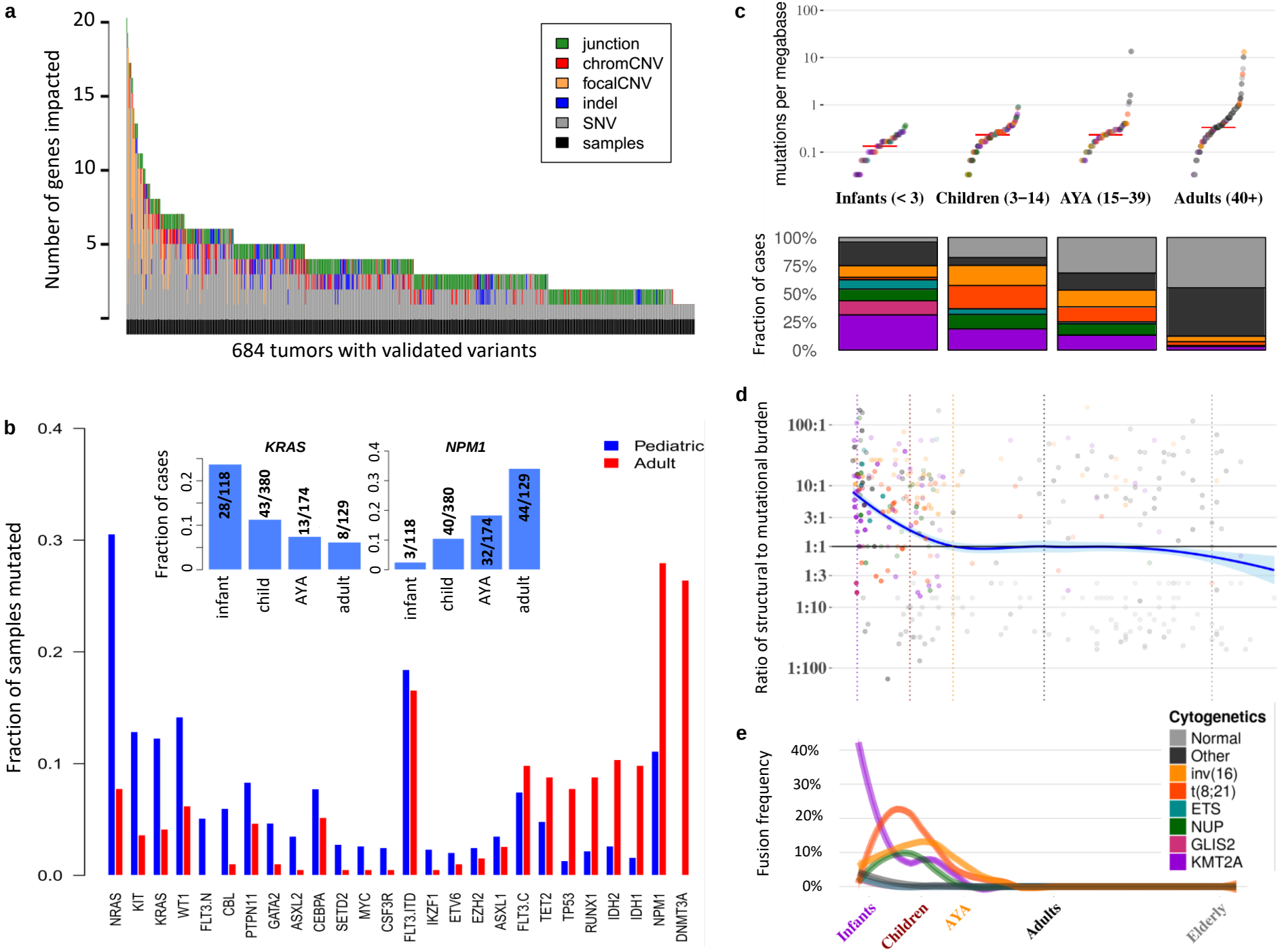
20. Maxson, J.E.*, et al.* CSF3R mutations have a high degree of overlap with CEBPA mutations in pediatric AML. *Blood* **127**, 3094-3098 (2016).
21. Quintana-Bustamante, O.*, et al.* Overexpression of wild-type or mutants forms of CEBPA alter normal human hematopoiesis. *Leukemia* **26**, 1537-1546 (2012).
22. Vicente, C., Conchillo, A., Garcia-Sanchez, M.A. & Odero, M.D. The role of the GATA2 transcription factor in normal and malignant hematopoiesis. *Crit Rev Oncol Hematol* **82**, 1-17 (2012).
23. Ng, K.P.*, et al.* Runx1 deficiency permits granulocyte lineage commitment but impairs subsequent maturation. *Oncogenesis* **2**, e78 (2013).
24. Wang, Y.*, et al.* The Wnt/beta-catenin pathway is required for the development of leukemia stem cells in AML. *Science* **327**, 1650-1653 (2010).
25. Valencia, A.*, et al.* Wnt signaling pathway is epigenetically regulated by methylation of Wnt antagonists in acute myeloid leukemia. *Leukemia* **23**, 1658-1666 (2009).
26. Nanbakhsh, A.*, et al.* c-Myc regulates expression of NKG2D ligands ULBP1/2/3 in AML and modulates their susceptibility to NK-mediated lysis. *Blood* **123**, 3585-3595 (2014).
27. Marcucci, G.*, et al.* MicroRNA expression in cytogenetically normal acute myeloid leukemia. *N Engl J Med* **358**, 1919-1928 (2008).
28. Ramamurthy, R.*, et al.* miR-155 expression and correlation with clinical outcome in pediatric AML: A report from Children's Oncology Group. *Pediatr Blood Cancer* **63**, 2096-2103 (2016).
29. Fooladinezhad, H., Khanahmad, H., Ganjalikhani-Hakemi, M. & Doosti, A. Negative regulation of TIM-3 expression in AML cell line (HL-60) using miR-330-5p. *Br J Biomed Sci* **73**, 129-133 (2016).
30. Lim, E.L.*, et al.* Comprehensive Sequence Analysis of Relapse and Refractory Pediatric Acute Myeloid Leukemia Identifies miRNA and mRNA Transcripts Associated with Treatment Resistance - a Report from the COG/NCI-Target AML Initiative. *Blood* **126**, 687-687 (2015).
31. Gruber, T.A.*, et al.* An Inv(16)(p13.3q24.3)-encoded CBFA2T3-GLIS2 fusion protein defines an aggressive subtype of pediatric acute megakaryoblastic leukemia. *Cancer Cell* **22**, 683-697 (2012).
32. Liang, K.*, et al.* Therapeutic Targeting of MLL Degradation Pathways in MLL-Rearranged Leukemia. *Cell* **168**, 59-72 e13 (2017).
33. Itskovich SS, C.J., Mulloy JC, Disney MD, Kumar AR. MBNL1 As a New Therapeutic Target in MLL-Fusion Gene Leukemia. in *Americal Society of Hematology Annual Conference* 462 (Orlando, Fl., 2015).
34. Caudell, D.*, et al.* Retroviral insertional mutagenesis identifies Zeb2 activation as a novel leukemogenic collaborating event in CALM-AF10 transgenic mice. *Blood* **115**, 1194-1203 (2010).
35. Li, J.*, et al.* The EMT transcription factor Zeb2 controls adult murine hematopoietic differentiation by regulating cytokine signaling. *Blood* **129**, 460-472 (2017).
36. Shlush, L.I.*, et al.* Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature* **506**, 328-333 (2014).
37. Guryanova, O.A.*, et al.* DNMT3A mutations promote anthracycline resistance in acute myeloid leukemia via impaired nucleosome remodeling. *Nat Med* **22**, 1488-1495 (2016).
38. Loghavi, S.*, et al.* Clinical features of de novo acute myeloid leukemia with concurrent DNMT3A, FLT3 and NPM1 mutations. *J Hematol Oncol* **7**, 74 (2014).
39. Xu, B.*, et al.* Tumor suppressor menin represses paired box gene 2 expression via Wilms tumor suppressor protein-polycomb group complex. *J Biol Chem* **286**, 13937-13944 (2011).
40. Sinha, S.*, et al.* Mutant WT1 is associated with DNA hypermethylation of PRC2 targets in AML and responds to EZH2 inhibition. *Blood* **125**, 316-326 (2015).
41. Abdel-Wahab, O.*, et al.* ASXL1 mutations promote myeloid transformation through loss of PRC2-mediated gene repression. *Cancer Cell* **22**, 180-193 (2012).

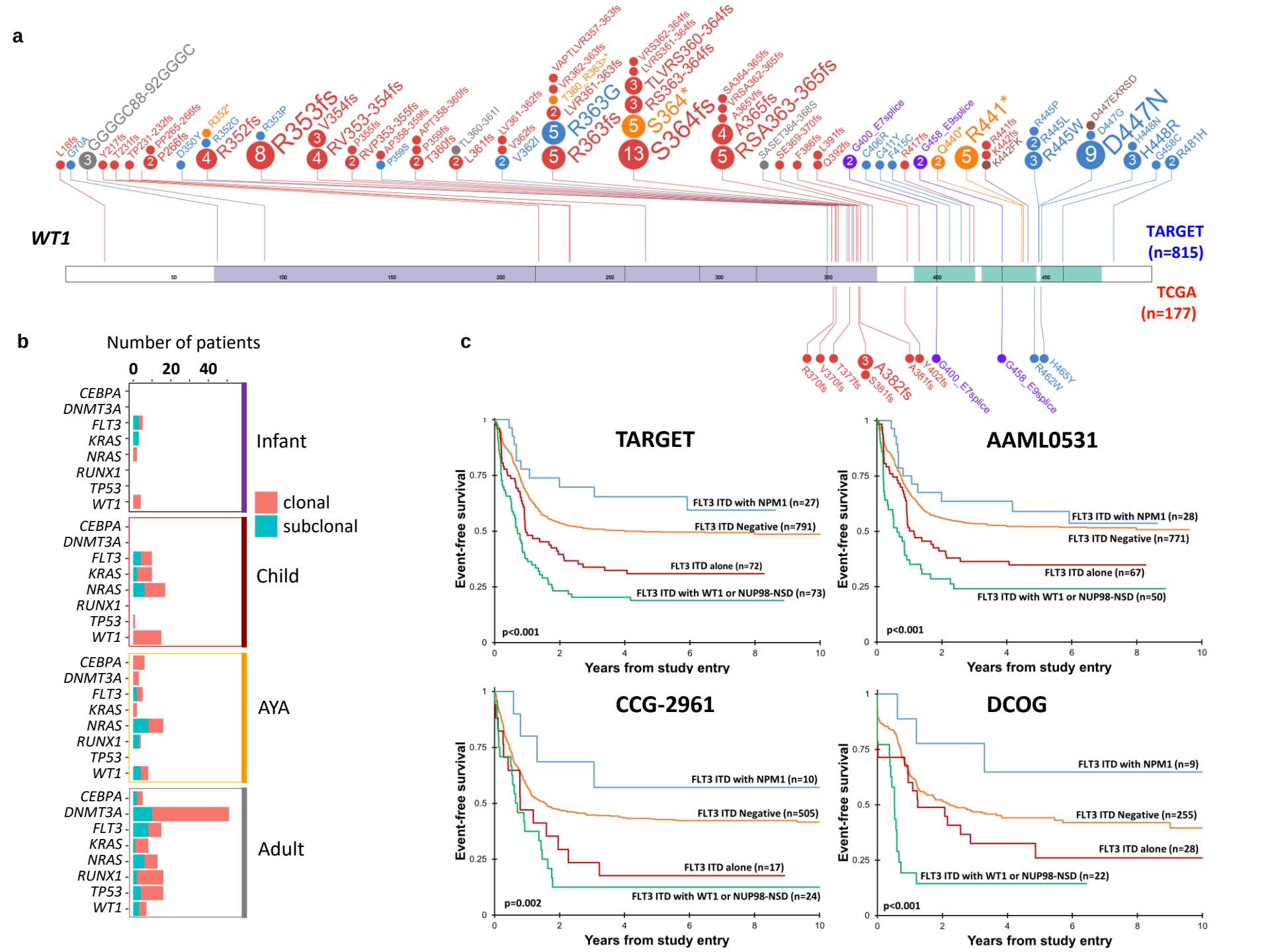
42. Gollner, S.*, et al.* Loss of the histone methyltransferase EZH2 induces resistance to multiple drugs in acute myeloid leukemia. *Nat Med* **23**, 69-78 (2017).

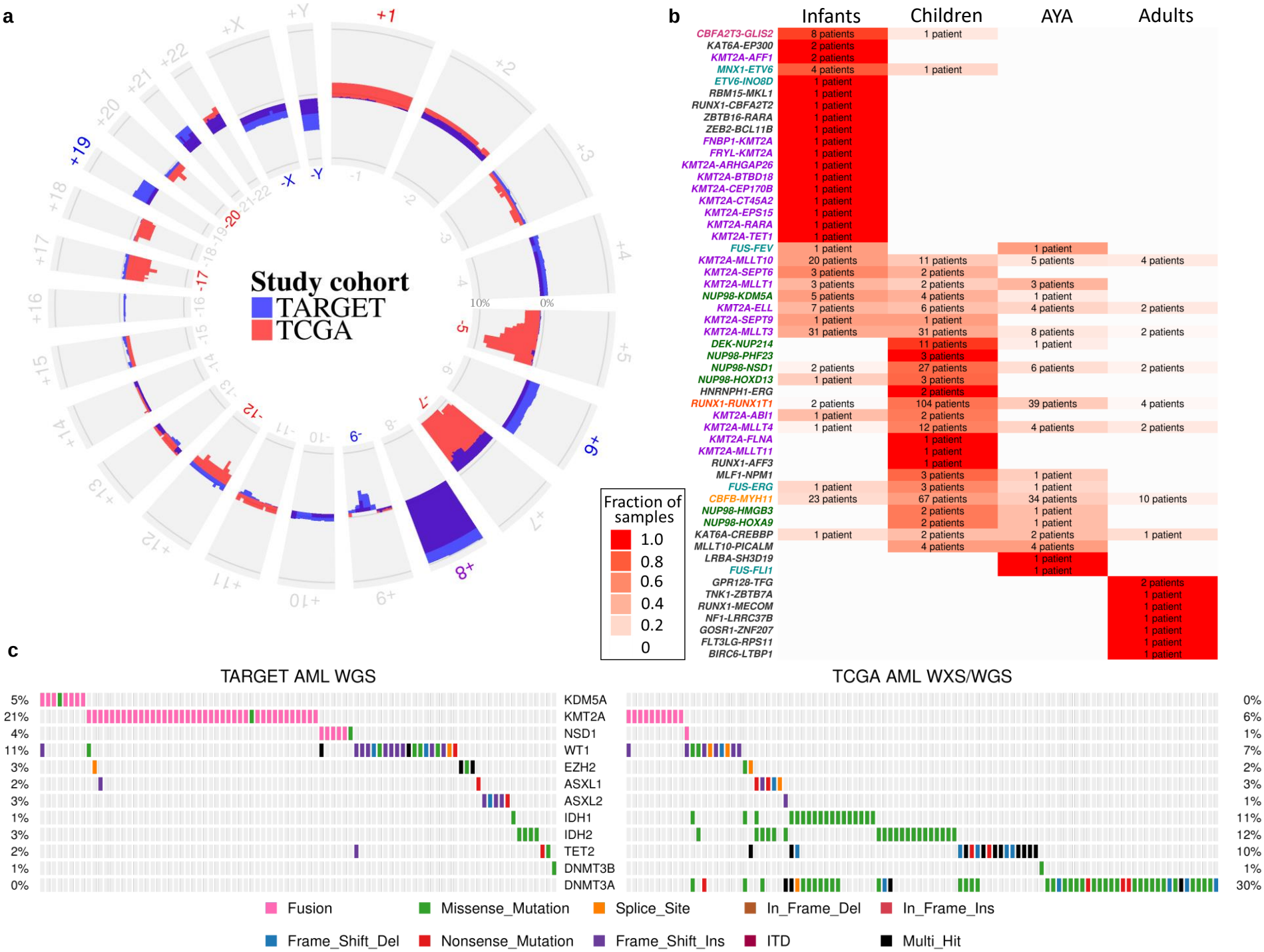

a
Risk
High
Low
Standard
Unknown
Cytogenetics
t(8;21)
Other
Normal
KMT2A
Inv(16)
Unknown
Age at diagnosis
30
20
10
0
Infant
Child
AYA
Adolescents & Young Adults (ages 15 - 39)
230 patients
17%
42%
38%
15%
14%
11%
20%
32%
Children (ages 3 - 14)
538 patients
13%
48%
34%
19%
12%
13%
26%
26%
Infants (less than 3 years old)
225 patients
13%
79%
10%
36%
38%
10%

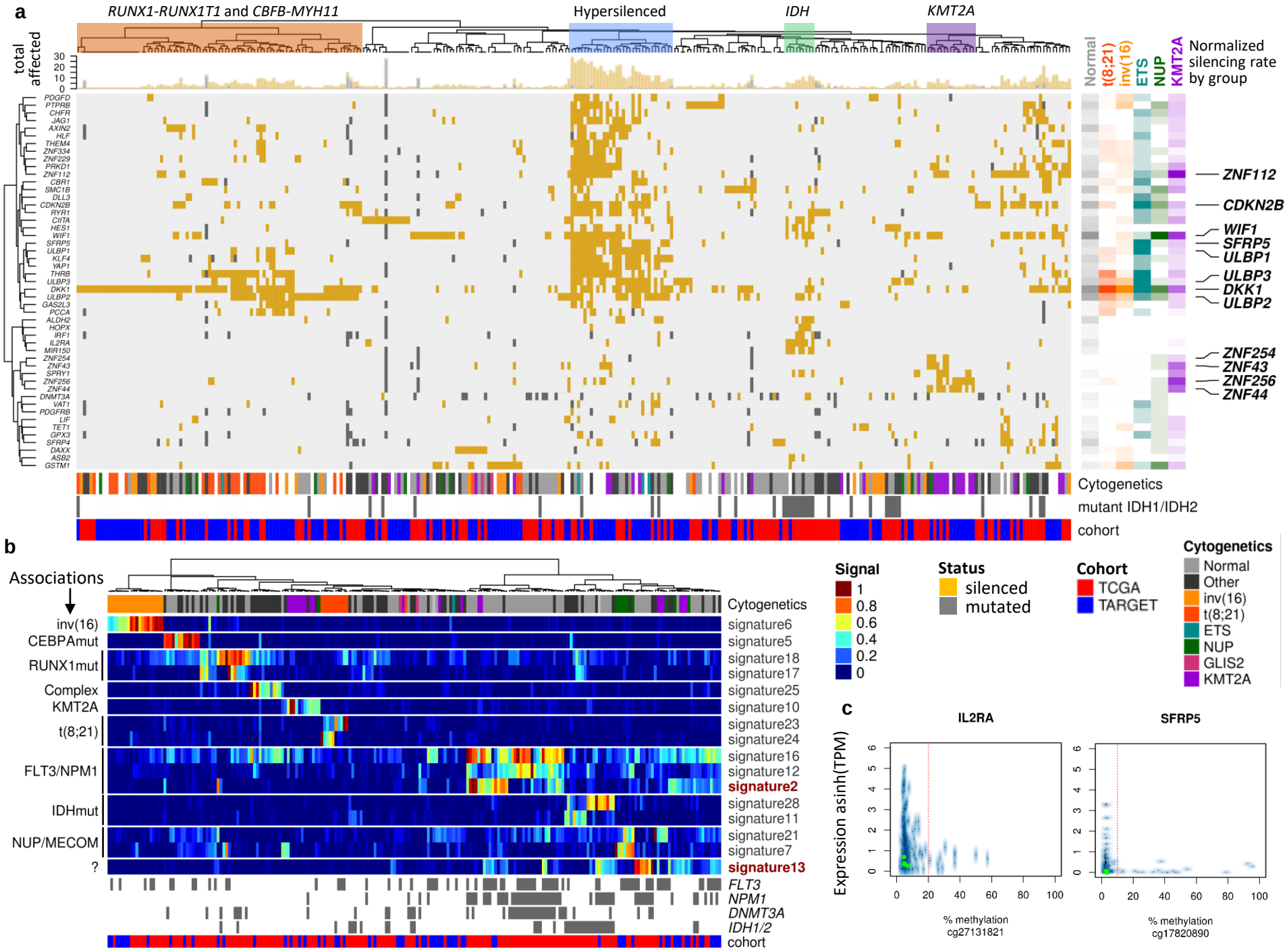

b
Number of patients impacted
150
100
50
0
FLT3.ITD
KMT2A −fusion
t(8;21)
inv(16)
trisomy(8)
NPM1
WT1
FLT3PM
CEBPA
Minus(Y)
del(9q)
Kit Exon8
Kit Exon17
Minus(X)
trisomy(21)
del(7q)
monosomy(7)
t(6;9)
del(5q)
t(3;5)
monosomy(5)

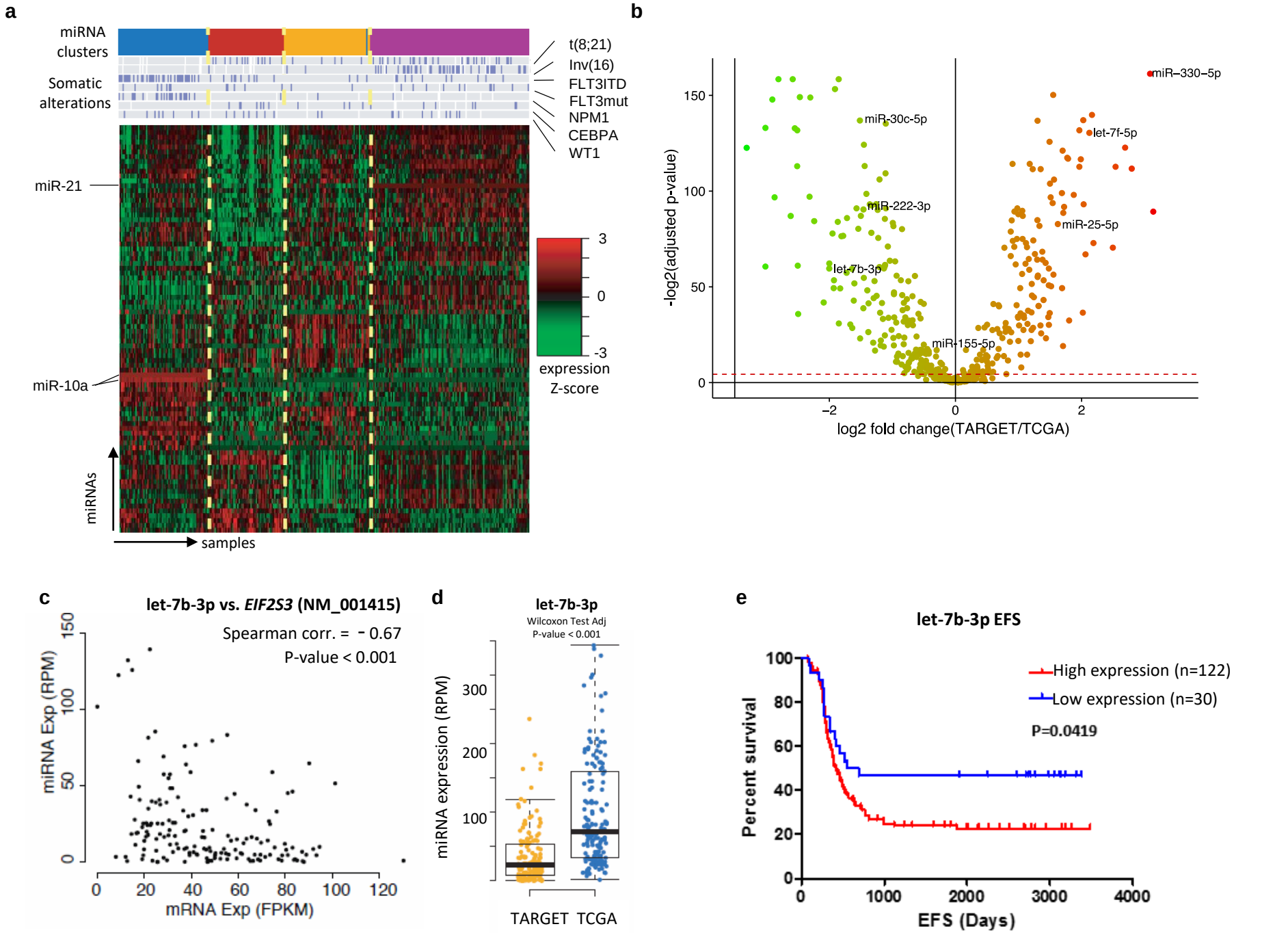

c
validation
501 tumor samples
matched remission
TCS
TCS
somatic variant calls
SNVs
indels
match
verified variants
verification
153 tumor samples
matched remission
TCS
TCS
somatic variants
SNVs
indels
76%
72%
same samples
discovery
197 tumor samples
matched remission
WGS
WGS
somatic variant calls
indels
SNVs
junctions
cytogenetics
160
mRNA-seq
CNVs
96
SNP6.0 arrays
89%
87%
75% −
85% +


a
Number of genes impacted
20
15
10
5
junction
chromCNV
focalCNV
indel
SNV
samples
684 tumors with validated variants
c
mutations per megabase
100
10
1
0.1
Infants (< 3)
Children (3−14)
AYA (15−39)
Adults (40+)
Fraction of cases
100%
75%
50%
25%
0%
b
Fraction of samples mutated
0.4
0.3
0.2
0.1
0.0
Pediatric
Adult
KRAS
Fraction of cases
28/118
43/380
13/174
8/129
infant
child
AYA
adult
NPM1
3/118
40/380
32/174
44/129
NRAS
KIT
KRAS
WT1
FLT3.N
CBL
PTPN11
GATA2
ASXL2
CEBPA
SETD2
MYC
CSF3R
FLT3.ITD
IKZF1
ETV6
EZH2
ASXL1
FLT3.C
TET2
TP53
RUNX1
IDH2
IDH1
NPM1
DNMT3A
d
Ratio of structural to mutational burden
100:1
10:1
3:1
1:1
1:3
1:10
1:100
e
Fusion frequency
40%
30%
20%
10%
0%
Cytogenetics
Normal
Other
inv(16)
t(8;21)
ETS
NUP
GLIS2
KMT2A
Infants
Children
AYA
Adults
Elderly


a
WT1
TARGET
(n=815)
TCGA
(n=177)
b
Number of patients
0
20
40
CEBPA
DNMT3A
FLT3
KRAS
NRAS
RUNX1
TP53
WT1
Infant
Child
AYA
Adult
clonal
subclonal
c
TARGET
FLT3 ITD with NPM1 (n=27)
FLT3 ITD Negative (n=791)
FLT3 ITD alone (n=72)
FLT3 ITD with WT1 or NUP98-NSD (n=73)
p<0.001
AAML0531
FLT3 ITD with NPM1 (n=28)
FLT3 ITD Negative (n=771)
FLT3 ITD alone (n=67)
FLT3 ITD with WT1 or NUP98-NSD (n=50)
p<0.001
CCG-2961
FLT3 ITD with NPM1 (n=10)
FLT3 ITD Negative (n=505)
FLT3 ITD alone (n=17)
FLT3 ITD with WT1 or NUP98-NSD (n=24)
p=0.002
DCOG
FLT3 ITD with NPM1 (n=9)
FLT3 ITD Negative (n=255)
FLT3 ITD alone (n=28)
FLT3 ITD with WT1 or NUP98-NSD (n=22)
p<0.001
Event-free survival
Years from study entry


a
Study cohort
TARGET
TCGA
10%
0%
b
Infants
Children
AYA
Adults
CBFA2T3-GLIS2
KAT6A-EP300
KMT2A-AFF1
MNX1-ETV6
ETV6-INO8D
RBM15-MKL1
RUNX1-CBFA2T2
ZBTB16-RARA
ZEB2-BCL11B
FNBP1-KMT2A
FRYL-KMT2A
KMT2A-ARHGAP26
KMT2A-BTBD18
KMT2A-CEP170B
KMT2A-CT45A2
KMT2A-EPS15
KMT2A-RARA
KMT2A-TET1
FUS-FEV
KMT2A-MLLT10
KMT2A-SEPT6
KMT2A-MLLT1
NUP98-KDM5A
KMT2A-ELL
KMT2A-SEPT9
KMT2A-MLLT3
DEK-NUP214
NUP98-PHF23
NUP98-NSD1
NUP98-HOXD13
HNRNPH1-ERG
RUNX1-RUNX1T1
KMT2A-ABI1
KMT2A-MLLT4
KMT2A-FLNA
KMT2A-MLLT11
RUNX1-AFF3
MLF1-NPM1
FUS-ERG
CBFB-MYH11
NUP98-HMGB3
NUP98-HOXA9
KAT6A-CREBBP
MLLT10-PICALM
LRBA-SH3D19
FUS-FLI1
GPR128-TFG
TNK1-ZBTB7A
RUNX1-MECOM
NF1-LRRC37B
GOSR1-ZNF207
FLT3LG-RPS11
BIRC6-LTBP1
Fraction of samples
1.0
0.8
0.6
0.4
0.2
0
c
TARGET AML WGS
TCGA AML WXS/WGS
KDM5A
KMT2A
NSD1
WT1
EZH2
ASXL1
ASXL2
IDH1
IDH2
TET2
DNMT3B
DNMT3A
Fusion
Missense_Mutation
Splice_Site
In_Frame_Del
In_Frame_Ins
Frame_Shift_Del
Nonsense_Mutation
Frame_Shift_Ins
ITD
Multi_Hit


a
RUNX1-RUNX1T1 and CBFB-MYH11
Hypersilenced
IDH
KMT2A
total affected
30
20
10
0
Normal
t(8;21)
inv(16)
ETS
NUP
KMT2A
Normalized silencing rate by group
PDGFD
PTPRB
CHFR
JAG1
AXIN2
HLF
THEM4
ZNF334
ZNF229
PRKD1
ZNF112
CBR1
SMC1B
DLL3
CDKN2B
RYR1
CIITA
HES1
WIF1
SFRP5
ULBP1
KLF4
YAP1
THRB
ULBP3
DKK1
ULBP2
GAS2L3
POC4
ALDH2
HOPX
IRF1
IL2RA
MIR150
ZNF254
ZNF43
SPRY1
ZNF256
ZNF44
DNMT3A
VAT1
PDGFRB
LIF
TET1
GPX3
SFRP4
DAXX
ASB2
GSTM1
ZNF112
CDKN2B
WIF1
SFRP5
ULBP1
ULBP3
DKK1
ULBP2
ZNF254
ZNF43
ZNF256
ZNF44
Cytogenetics
mutant IDH1/IDH2
cohort
b
Associations
Cytogenetics
inv(16)
CEBPAmut
RUNX1mut
Complex
KMT2A
t(8;21)
FLT3/NPM1
IDHmut
NUP/MECOM
?
signature6
signature5
signature18
signature17
signature25
signature10
signature23
signature24
signature16
signature12
signature2
signature28
signature11
signature21
signature7
signature13
FLT3
NPM1
DNMT3A
IDH1/2
cohort
Signal
1
0.8
0.6
0.4
0.2
0
Status
silenced
mutated
Cohort
TCGA
TARGET
Cytogenetics
Normal
Other
inv(16)
t(8;21)
ETS
NUP
GLIS2
KMT2A
c
IL2RA
SFRP5
Expression asinh(TPM)
% methylation
cg27131821
% methylation
cg17820890


a
miRNA clusters
Somatic alterations
t(8;21)
Inv(16)
FLT3ITD
FLT3mut
NPM1
CEBPA
WT1
miR-21
miR-10a
miRNAs
samples
3
0
-3
expression Z-score
b
miR–330–5p
miR-30c-5p
let-7f-5p
miR-222-3p
miR-25-5p
let-7b-3p
miR-155-5p
-log2(adjusted p-value)
log2 fold change(TARGET/TCGA)
c
let-7b-3p vs. EIF2S3 (NM_001415)
Spearman corr. = - 0.67
P-value < 0.001
miRNA Exp (RPM)
mRNA Exp (FPKM)
d
let-7b-3p
Wilcoxon Test Adj P-value < 0.001
miRNA expression (RPM)
TARGET
TCGA
e
let-7b-3p EFS
High expression (n=122)
Low expression (n=30)
P=0.0419
Percent survival
EFS (Days)