Leukemia (2020) 34:2621–2634
https://doi.org/10.1038/s41375-020-0839-4

ARTICLE

Acute myeloid leukemia

Clinical presentation and differential splicing of *SRSF2*, *U2AF1* and *SF3B1* mutations in patients with acute myeloid leukemia

Stefanos A. Bamopoulos[1,2] · Aarif M. N. Batcha[3,4] · Vindi Jurinovic[1,3] · Maja Rothenberg-Thurley[1] · Hanna Janke[1] · Bianka Ksienzyk[1] · Julia Philippou-Massier[5] · Alexander Graf[5] · Stefan Krebs[5] · Helmut Blum[5] · Stephanie Schneider[1,6] · Nikola Konstandin[1] · Maria Cristina Sauerland[7] · Dennis Görlich[7] · Wolfgang E. Berdel[8] · Bernhard J. Woermann[9] · Stefan K. Bohlander[10] · Stefan Canzar[11] · Ulrich Mansmann[3,4,12,13] · Wolfgang Hiddemann[1,12,13] · Jan Braess[14] · Karsten Spiekermann[1,12,13] · Klaus H. Metzeler[1,12,13] · Tobias Herold[1,12,15]

Received: 2 January 2020 / Revised: 6 April 2020 / Accepted: 14 April 2020 / Published online: 1 May 2020

Abstract

Previous studies demonstrated that splicing factor mutations are recurrent events in hematopoietic malignancies with both clinical and functional implications. However, their aberrant splicing patterns in acute myeloid leukemia remain largely unexplored. In this study, we characterized mutations in *SRSF2, U2AF1*, and *SF3B1*, the most commonly mutated splicing factors. In our clinical analysis of 2678 patients, splicing factor mutations showed inferior relapse-free and overall survival, however, these mutations did not represent independent prognostic markers. RNA-sequencing of 246 and independent validation in 177 patients revealed an isoform expression profile which is highly characteristic for each individual mutation, with several isoforms showing a strong dysregulation. By establishing a custom differential splice junction usage pipeline, we accurately detected aberrant splicing in splicing factor mutated samples. A large proportion of differentially used junctions were novel, including several junctions in leukemia-associated genes. In *SRSF2*(P95H) mutants, we further explored the possibility of a cascading effect through the dysregulation of the splicing pathway. Furthermore, we observed a validated impact on overall survival for two junctions overused in *SRSF2*(P95H) mutants. We conclude that splicing factor mutations do not represent independent prognostic markers. However, they do have genome-wide consequences on gene splicing leading to dysregulated isoform expression of several genes.

Introduction

The discovery of recurring somatic mutations within splicing factor genes in a large spectrum of human malignancies has brought attention to the critical role of splicing and its complex participation in carcinogenesis [1–3]. The spliceosome is a molecular machine assembled from small nuclear RNA (snRNA) and proteins and is responsible for intron removal (splicing) in pre-messenger RNA. In acute myeloid leukemia (AML), splicing factor mutations occur most frequently in *SRSF2, U2AF1*, and *SF3B1*. The splicing factors encoded by these genes are all involved in the recognition of the 3′-splice site during pre-mRNA processing [4]. Splicing factor (SF) mutations are especially common in haematopoietic malignancies, where they occur early on and remain stable throughout the disease evolution of myelodysplastic syndromes (MDS) [1, 5–9]. SF mutations are also prevalent in AML, which is often the result of myeloid dysplasia progression, with reported frequencies of 6–10%, 4–8%, and 3% for *SRSF2, U2AF1*, and *SF3B1* mutations, respectively [2, 4, 10, 11].

SF mutations rarely co-occur within the same patient, implying the lack of a synergistic effect or synthetic lethality [1, 2, 6]. They are typically heterozygous point mutations, frequently coincide with other recurrent mutations in haematopoietic malignancies and are associated

Supplementary information The online version of this article (https://doi.org/10.1038/s41375-020-0839-4) contains supplementary material, which is available to authorized users.

✉ Stefanos A. Bamopoulos
stefanos.bamopoulos@charite.de

✉ Tobias Herold
tobias.herold@med.uni-muenchen.de

Extended author information available on the last page of the article

with aberrant splicing in genes recurrently mutated in AML [2, 4, 8]. Notably, the aberrant splicing patterns are distinct for each SF mutation, suggesting that SF mutations do not share the same mechanism of action and should be recognized as individual alterations [4, 9, 12–17].

The clinical characteristics and outcome of patients with SF mutations are well defined in MDS [1, 3, 8, 9]. Meanwhile, attempts at determining the role of SF mutations as independent prognostic markers in AML have often been limited to specific subgroups and it remains unclear, whether the inferior survival associated with SF mutations is confounded by their association with older age or accompanying mutations [10, 18]. Additionally, while evidence of aberrant splicing due to SF mutations has emerged for many genes relevant in AML, it is yet uncertain whether and how these changes directly influence disease initiation or evolution.

The aim of this study was a comprehensive analysis of the prognostic implications of SF mutations in two well-characterized and intensively treated adult AML patient cohorts amounting to a total of 2678 patients. In addition, the core functional consequences of SF mutations were explored using targeted amplicon sequencing in conjunction with RNA-sequencing on two large datasets.

Patients and methods

Patients

Our primary cohort included a total of 1138 AML patients treated with intensive chemotherapy in two randomized multicenter phase 3 trials of the German AML Cooperative Group (AMLCG). Treatment regimens and inclusion criteria are described elsewhere [2]. A cohort of 1540 AML patients participating in multicenter clinical trials of the German-Austrian AML Study Group (AMLSG) were used for validation [19]. Cohort composition and filtering criteria are outlined in the supplement.

Molecular workup

All participants of the AMLCG cohort received cytogenetic analysis, as well as targeted DNA-sequencing as described previously [2]. The subset of the AMLSG cohort included in this study received a corresponding molecular workup, described elsewhere [19].

RNA-sequencing and data processing

Using the Sense mRNA Seq Library Prep Kit V2 (Lexogen; Vienna, Austria) 246 samples underwent, strand-specific, paired-end sequencing on a HiSeq 1500 instrument (Illumina; San Diego, CA, USA). A subset of the Beat AML cohort ($n = 177$) was used for validation [20]. The same bioinformatics analysis was used for both datasets and is described in the supplementary. The samples were aligned to the reference genome (Ensembl GRCh37 release 87) using the STAR [21] aligner with default parameters. Splice junctions from all samples were pooled, filtered, and used to create a new genomic index. Multi-sample 2-pass alignments to the re-generated genome index were followed, using the STAR recommended parameters for gene-fusion detection. Read counts of transcripts and genes were measured with salmon [22]. Read counts of splice junctions were extracted from the STAR output.

Differential expression analyses and differential splice junction usage (DSJU)

Differentially expressed isoforms were identified with the limma [23] package after TMM-normalization [24] with edgeR [25] and weighting with voom [26, 27]. DSJU was quantified similarly using the diffSplice function of the limma package. Differentially expressed exonic and intronic segments were also quantified with the diffSplice function after counting with DEXSeq [28]. All analyses are described in the supplementary. Raw read counts for all analyses are available in the GEO database (GSE146173).

Nanopore cDNA sequencing and analysis

Total RNA was transcribed into cDNA using the TeloPrime Full-Length cDNA Amplification Kit (Lexogen), which is highly selective for polyadenylated full-length RNA molecules with 5′-cap structures. Two barcoded samples for multiplexed analysis were sequenced on the Oxford Nanopore Technologies MinION platform. Alternative isoform analysis was performed with FLAIR [29].

Statistics

Statistical analysis was performed using the R-3.4.1 [30] software package. Correlations between categorical and continuous variables were performed using the Mann–Whitney *U*-test while the Pearson's chi-squared test was used for comparisons between categorical variables. In case of multiple testing, *p*-value adjustment was performed as described in the supplement. Survival analysis was performed and visualized using the Kaplan–Meier method and the log-rank test was utilized to capture differences in relapse-free survival (RFS) and overall survival (OS). Patients receiving an allogeneic stem cell transplant were censored at the day of the transplant, for both RFS and OS. Additionally, Cox regressions were performed for all available clinical parameters and recurrent aberrations. Cox

multiple regression models were then built separately for RFS and OS, using all variables with an unadjusted p-value < 0.1 in the single Cox regression models.

Results

Clinical features of AML patients with SF mutations

We characterized SF mutations in two independent patient cohorts (the AMLCG and AMLSG cohorts). Our primary cohort (AMLCG) consisted of 1119 AML patients (Fig. S1), 232 (20.7%) of which presented with SF mutations. The three most commonly affected SF genes, *SRSF2*, *U2AF1*, and *SF3B1* were mutated in 11.9% ($n = 133$), 3.4% ($n = 38$), and 4.0% ($n = 45$) of the patients (Fig. 1a). In agreement with previous findings [19], SF mutations were in their majority mutually exclusive, heterozygous hotspot mutations (Fig. 1a, b). The four most common point mutations were *SRSF2* (P95H) ($n = 69$), *SRSF2*(P95L) ($n = 27$), *U2AF1*(S34F) ($n = 18$), and *SF3B1*(K700E) ($n = 18$) mutations (Fig. 1c). The clinical characteristics of patients harboring SF mutations are summarized in Tables 1 and S1 (AMLSG cohort), along with a statistical assessment between cohorts (Table S2). We observed a high overall degree of similarity regarding clinical features of SF mutated patients between the AMLCG and AMLSG cohorts, despite their large median age difference. Mutations in *SRSF2*, *U2AF1*, and *SF3B1* occurred more frequently in secondary AML (44.7% compared to 18.2% in de novo AML, $p < 0.001$ for *SRSF2* and *U2AF1* mutations and $p = 0.021$ for *SF3B1* mutations) and were all associated with older age (*SRSF2*: $p < 0.001$, *U2AF1*: $p = 0.007$, *SF3B1*: $p = 0.001$). As reported previously [1], *SRSF2* and *U2AF1* mutated patients were predominantly male (76.7%; $p < 0.001$ and 76.3%; $p = 0.003$, respectively). Furthermore, patients harboring *SRSF2* mutations presented with a lower white blood cell count (WBC; median 13.3 10^9/L vs. 22.4 10^9/L; $p = 0.002$) while *U2AF1* mutated patients presented with a reduced blast percentage in their bone marrow when compared to SF wildtype patients (median 60% vs. 80%; $p = 0.008$).

Associations of SF mutations and other recurrent alterations in AML

In second step, we investigated associations between SF mutations and recurrent cytogenetic abnormalities and gene mutations in AML (Fig. 2). Notably, SF mutations were not found in inv(16)/*t*(16;16) patients ($n = 124$), with the exception of one inv(16)/*t*(16;16) patient harboring a *U2AF1*(R35Q) mutation. The same held true for *t*(8;21) patients ($n = 98$), where only one patient had a rare deletion in *SRSF2*. Additionally, all patients in the AMLCG cohort presenting with an isolated trisomy 13 ($n = 9$) also harbored an *SRSF2* mutation ($p < 0.001$), as described previously [31].

Mutations in all SF genes correlated positively with mutations in *BCOR* (all $p < 0.001$) and *RUNX1* (all $p < 0.001$) and negatively with mutations in *NPM1* (*SRSF2* and *U2AF1*: $p < 0.001$, *SF3B1*: $p = 0.006$). Expectedly, *SRSF2*(P95H) and *SRSF2*(P95L) mutations shared a similar pattern of co-expression including significant pairwise associations with mutations in *ASXL1*, *IDH2*, *RUNX1* (both $p < 0.001$) and *STAG2* ($p < 0.001$ and $p = 0.002$, respectively). However, apart from *IDH2* mutations where co-occurrence was comparable (OR: 3.4 vs. 5.1), mutations in *ASXL1*, *RUNX1*, and *STAG2* coincided more frequently with *SRSF2*(P95H) mutations. Despite this, *SRSF2*(P95L) mutations showed a slightly increased co-occurrence with other recurrent AML mutations (median 5 vs. 4 mutations, $p = 0.046$).

Prognostic relevance of SF mutations for relapse-free survival and overall survival

The prognostic impact of *SRSF2*, *U2AF1*, and *SF3B1* mutations was initially assessed using Kaplan–Meier graphs and log-rank testing. All SF mutations presented with both inferior relapse-free survival (RFS) and overall survival (OS) when compared to SF wildtype patients (Figs. S2.1, S2.2; Table S3). The effect was most pronounced in *U2AF1* mutated patients with an one-year survival rate of only 29.1%, followed by *SF3B1* (40.6%) and *SRSF2* mutated patients (49.2%). Different point mutations inside the same SF gene did not differ significantly in their effect on OS.

To confirm the observed prognostic impact of SF mutations, we performed single Cox regressions on all available clinical and genetic parameters. In agreement with the Kaplan–Meier estimates, patients harboring *SRSF2* (P95H), *SRSF2*(P95L), *U2AF1*(S34F), and *SF3B1*(K700E) mutations had significantly reduced RFS and OS (Fig. S3.1). To test whether any SF mutation was an independent prognostic marker, multiple Cox regression models (Figs. 3 and S3.1, 3.2) were built by integrating all parameters significantly associated ($p < 0.1$) with RFS and OS in the single Cox regression models. Along with several known predictors, only *U2AF1*(S34F) mutations presented with prognostic relevance for both RFS (Hazard ratio = 2.81, $p = 0.012$) and OS (HR = 1.90, $p = 0.034$) in the AMLCG cohort, but not in the AMLSG cohort (OS: HR = 1.39, $p = 0.416$; RFS: HR = 1.38, $p = 0.419$). However, when aggregating mutations at the gene level, mutations in *SRSF2* and *SF3B1* presented with prognostic relevance for RFS in the AMLSG cohort (HR = 1.77, $p = 0.008$; HR = 2.15, $p = 0.014$; respectively), while not reaching significance in the AMLCG cohort ($p = 0.586$ and $p = 0.060$, respectively). When looking only at de novo AML patients, the prognostic impact of *U2AF1*(S34F) mutations

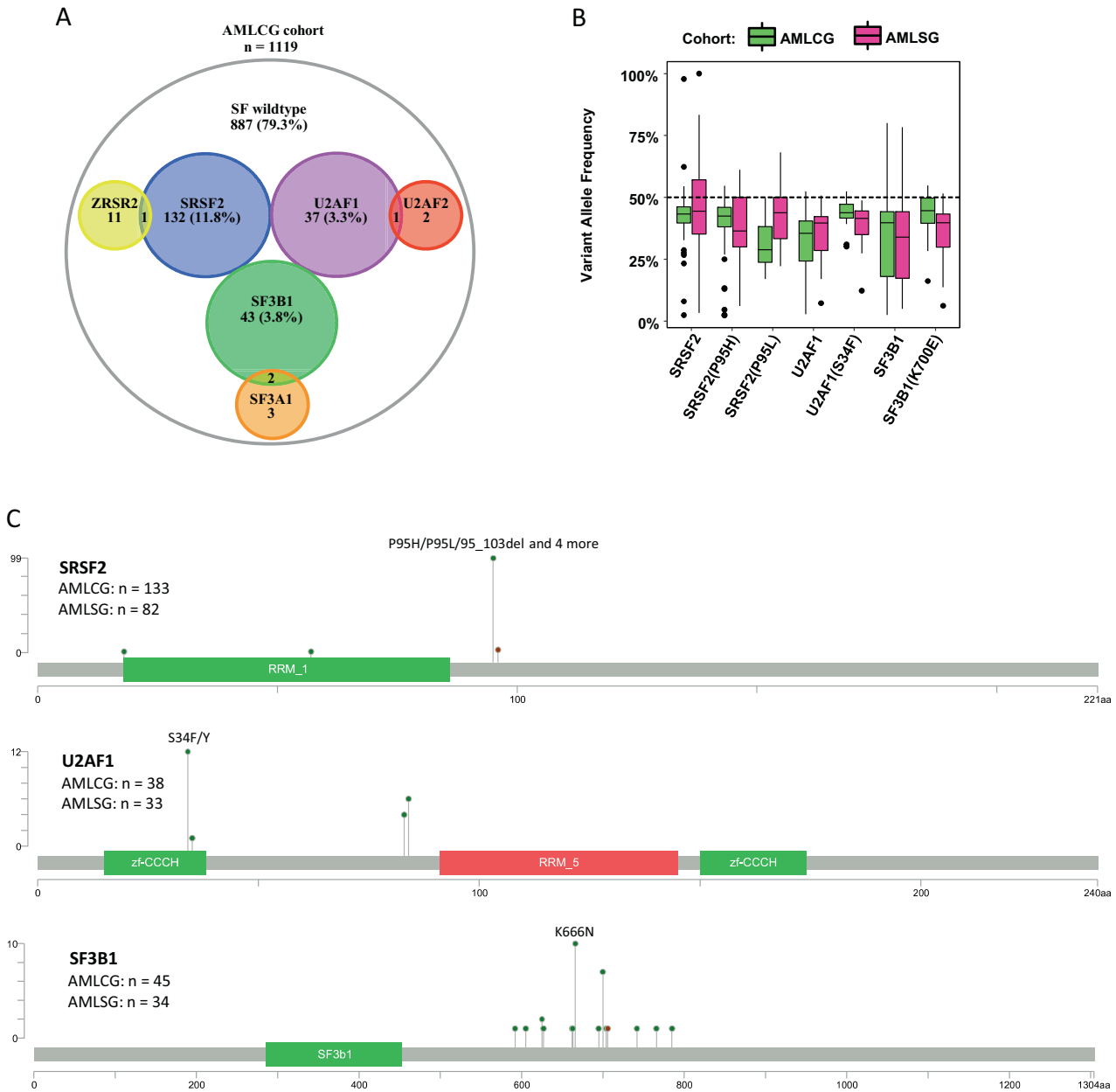

Fig. 1 Frequency and location of SF mutations. a Euler diagram depicting the frequency of SF mutations in the AMLCG cohort. *SRSF2*, *U2AF1* and *SF3B1* mutations were mutually exclusive. **b** Variant allele frequency of SF mutations in both study cohorts. **c** Mutation plots showing the protein location of all SF mutations in the genes *SRSF2*, *U2AF1*, and *SF3B1* for all patients of the AMLCG cohort. Number of patients harboring SF mutations are additionally provided for both cohorts.

diminished ($p = 0.075$), yet the prognostic impact observed for *SRSF2* and *SF3B1* remained significant in the AMLSG cohort (HR = 1.84, $p = 0.009$; HR = 2.43, $p = 0.015$; respectively) (Tables S4.1, S4.2).

Differential isoform expression in SF mutated patients

We next assessed the impact of SF mutations on mRNA expression. To this end, whole-transcriptome RNA-sequencing was performed on 246 AML patients, 29 of which harbored a mutation in the SF genes of interest, while 199 SF wildtype patients were used as a control (Figure S4 and Table S5). In addition, a subset of the Beat AML cohort ($n = 177$) with matched DNA-sequencing and RNA-sequencing data was used for validation [20].

After low-coverage filtering, we performed a differential isoform expression analysis for ~90,000 isoforms. Differential expression was restricted to a small fraction of all expressed isoforms (<0.5%; Fig. 4a and Table S6). Little

Table 1 Clinical characteristics of SF mutations in the AMLCG cohort.

Variables	SF wildtype	*SRSF2*	*p*	*U2AF1*	*p*	*SF3B1*	*p*
No. of patients	903	133	–	38	–	45	–
Age, years, and median (range)	55 (18–86)	65 (25–80)	**<0.001**	64 (23–74)	**0.007**	65 (31–78)	**0.001**
Female sex, no. (%)	490 (54.3)	31 (23.3)	**<0.001**	9 (23.7)	**0.003**	19 (42.2)	0.387
Hemoglobin, g/dL, median (range)	9 (3.5–16)	8.9 (3.8–14.7)	0.466	9.2 (6–13.6)	1.000	8.9 (6.8–13.4)	1.000
WBC count, 10^9/L, median (range)	22.4 (0.1–798.2)	13.3 (0.5–406)	**0.008**	7 (0.7–666)	0.079	22.4 (0.9–269.5)	1.000
Platelets, 10^9/L, median (range)	55 (0–1760)	49.5 (0–643)	0.736	47 (11–132)	0.466	67 (5–585)	0.744
LDH, U/L, median (range)	448 (76–19624)	362 (150–14332)	0.118	346 (128–3085)	0.313	472 (142–7434)	0.950
BM blasts, %, median (range)	80 (6–100)	76 (15–100)	0.455	60 (10–95)	**0.002**	70 (13–95)	0.206
Performance status (ECOG) > 1, no. (%)	157 (25.9)	18 (26.1)	1.000	8 (27.6)	1.000	4 (16)	0.941
Primary AML, no. (%)	786 (87)	100 (75.2)	**0.016**	24 (63.2)	**0.003**	29 (64.4)	**0.002**
Secondary AML, no (%)	69 (7.6)	30 (22.6)	**<0.001**	13 (34.2)	**<0.001**	11 (24.4)	**0.021**
Therapy-related AML, no (%)	48 (5.3)	3 (2.3)	0.507	1 (2.6)	1.000	5 (11.1)	0.380
Allogeneic transplant, no. (%)	296 (32.8)	27 (20.3)	**0.022**	7 (18.4)	0.273	12 (26.7)	0.797
Complete Remission, no. (%)	641 (71)	71 (53.4)	**<0.001**	18 (47.4)	**0.015**	19 (42.2)	**0.001**
Relapse, no (%)	366 (63.5)	46 (82.1)	**0.034**	11 (78.6)	0.713	14 (87.5)	0.266
Deceased, no (%)	575 (63.7)	112 (84.2)	**<0.001**	34 (89.5)	**0.010**	41 (91.1)	**0.002**

WBC white blood cells, *LDH* lactate dehydrogenase, *BM* bone marrow, *ECOG* Eastern Co-operative Oncology Group performance score.

overlap of differentially expressed (DE) isoforms was found when different SF mutation groups were compared to the control, consistent with previous observations [32]. However, ten isoforms were reported as DE in both *SRSF2* (P95H) and *SRSF2*(P95L) mutated samples, all with the same fold-change direction (Fig. 4b). Out of those, the isoforms in *GTF2I*, *H1F0*, *INHBC*, *LAMC1*, and one of the isoforms of *METTL22* (*ENST00000562151*) were also significant in the validation cohort for both *SRSF2*(P95H) and *SRSF2*(P95L). Additionally, the isoform of *H1F0* was also reported as DE for *U2AF1*(S34F) mutants in both cohorts. For *SRSF2*(P95H) mutants 107 of all DE isoforms also reached significance in the validation cohort (40.1%), while for the other SF mutation subgroups validation rates ranged from 15.1 to 27.3% increasing with larger mutant sample sizes. Notably, mutated and wildtype samples showed large differences in the expression levels of several isoforms (Fig. 4c and Fig. S5). The top two overexpressed isoforms in *SRSF2*(P95H) both corresponded to *INTS3*, which was recently reported as dysregulated in *SRSF2*(P95H) mutants co-expressing *IDH2* mutations [33]. Several DE isoforms identified in SF mutated patients correspond to cancer-related genes, many of which have a known role in AML. Specifically, genes with DE isoforms included, but were not limited to *BRD4* [34], *EWSR1* [35], and *YBX1* [36] in *SRSF2*(P95H) mutated samples, *CUX1* [37], *DEK* [15, 38], and *EZH1* [39] in *U2AF1*(S34F) mutated samples, as well as *PTK2* [40] in *SF3B1*(K700E) mutated patients (Tables S7.1, S7.2).

Hierarchical clustering using DE isoforms was performed on all samples to assess the expression homogeneity of SF mutations. A tight clustering of samples harboring identical SF point mutations was observed, indicating an isoform expression profile highly characteristic for each individual SF mutation (Figs. S6.1–S6.3). When using DE isoforms resulting from the comparison of all *SRSF2* mutated samples against SF wildtype samples, the samples did not cluster as well. This stands in agreement with the limited overlap of differentially expressed isoforms found between the two *SRSF2* point mutations examined and suggests at least some heterogeneity among them. The same also held true for *U2AF1* mutated samples, however all *SF3B1* mutated samples still clustered together when compared as a single group to the control.

Differential splicing in SF mutants

Previous studies have reported differential splicing as causal for isoform dysregulation in SF mutants [41, 42]. To detect aberrant splicing in our dataset, we quantified the usage of all unique splice junctions (Fig. S7). After filtering out junctions with low expression, 221,249 unique junctions (19.3% novel) remained across 15,526 annotated genes (Table S8). Applying the same workflow to the Beat AML cohort yielded 194,158 junctions (8.3% novel). Notably, of the 172,518 junctions shared across both datasets, 10,029 (5.8%) were novel. The novel junctions passing our filtering criteria were supported by a high amount of reads and samples with a distribution comparable to that of annotated junctions (Fig. 5a). Neither the number of novel junctions nor the number of reads supporting them correlated with the presence of SF mutations.

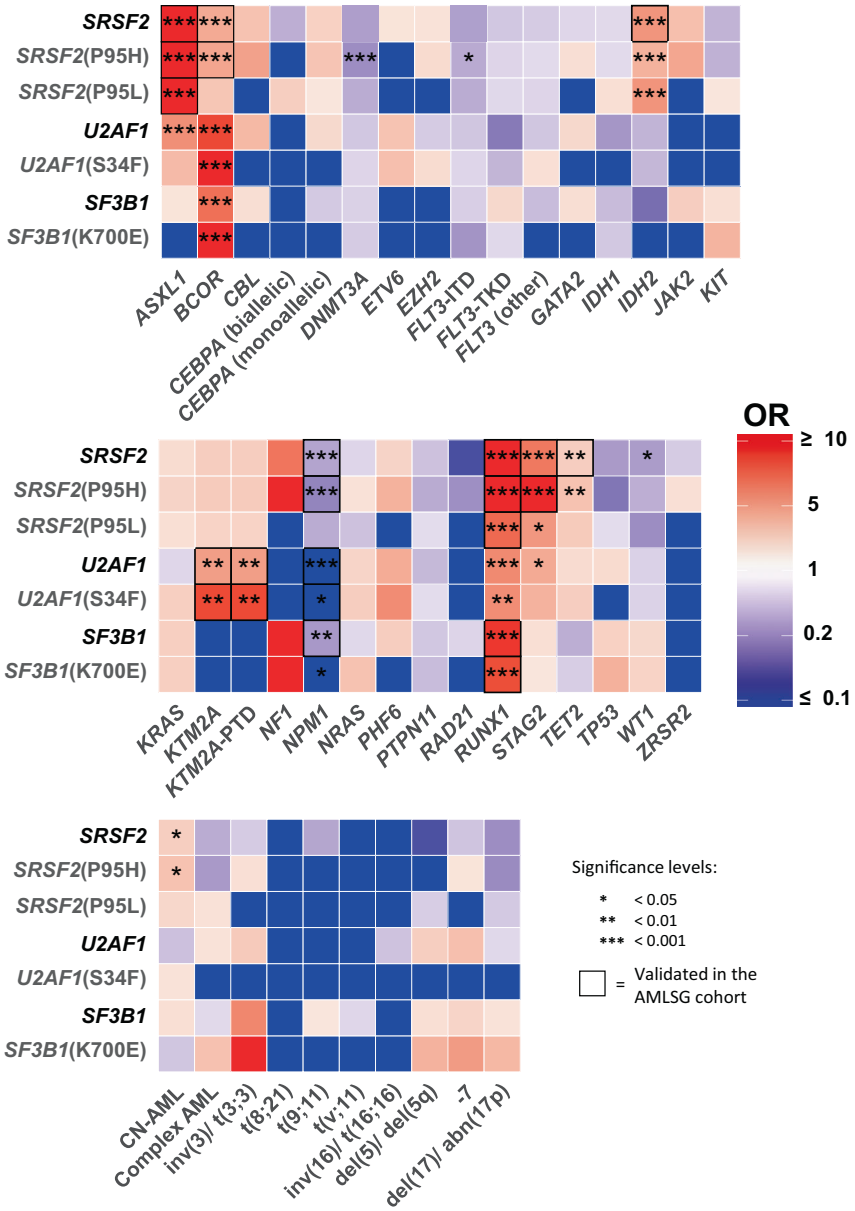

Fig. 2 Correlations between SF mutations and recurrent abnormalities. Correlation matrix depicting the co-occurrence of SF mutations and recurrent mutations in AML, as well as cytogenetic groups, as defined in the ELN 2017 classification. Only variables with a frequency >1% in the AMLCG cohort are shown. FLT3-ITD *FLT3* internal tandem duplication mutation, FLT3-TKD *FLT3* tyrosine kinase domain mutation, CN-AML cytogenetically normal AML.

In consideration of the high proportion of novel junctions in both datasets, we employed a customized pipeline that can quantify the differential splice junction usage (DSJU) of each individual junction. Of the several hundred junctions reported as differentially used in our primary cohort ($p < 0.05$, $\log_2$(fold change) > 1), 20.2–45.9% constituted novel junctions (Tables S9.1–S10.4) and were classified as described previously (Fig. 5b) [15]. Unsurprisingly, validation rates increased with larger mutant sample sizes, ranging from 9.3% (*SF3B1*(K700E); $n = 3$) to 74.0% (all *SRSF2* mutants; $n = 26$). Furthermore, validation rates were higher for novel junctions (mean 39.3% vs. 21.5% known junctions), likely due to the stricter initial filtering criteria applied. By performing nanopore sequencing of one *SRSF2* (P95H) mutant and one SF wildtype sample we were able to confirm the usage of several novel junctions and detect resulting novel isoforms as exemplified for *IDH3G* in Fig. 6a, b. A tendency towards decreased junction usage was observed for all SF point mutations and was most evident in *SF3B1*(K700E) mutants (1423/1927; 73.9% of differentially used junctions). The total number of splicing events, however, was not reduced in SF mutants (mean 9,275,359 events vs. 9,192,697 in wildtype patients).

A quantification of all non-overlapping exonic and intronic segments showed a limited amount of differentially expressed segments (0.2–1.3% of all filtered segments, Tables S11–S13.4) in line with the modest effect on splice junction usage observed in SF mutants. Notably,

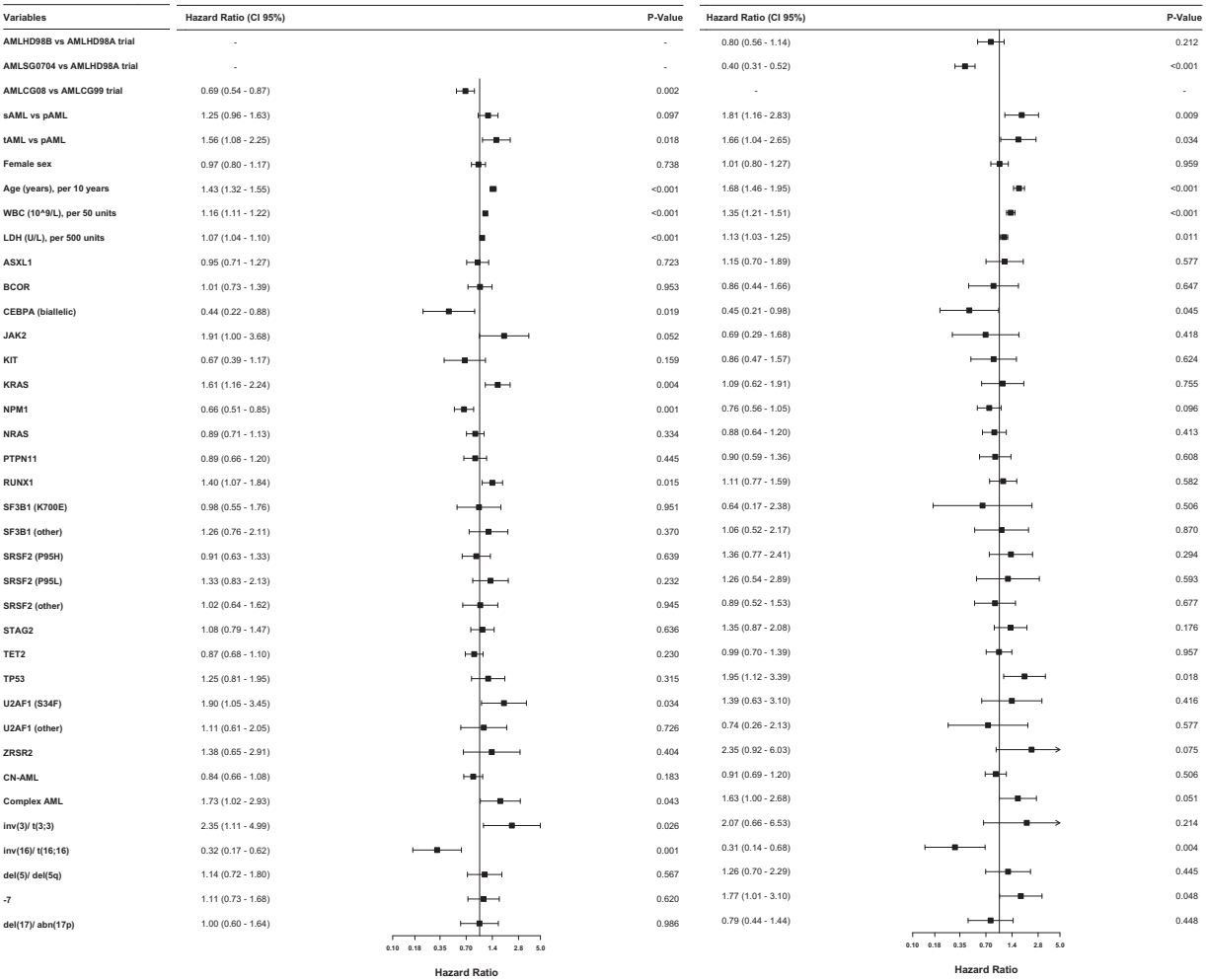

Fig. 3 Multiple Cox regression models for overall survival. Multiple Cox regression models (for OS) were performed separately for patients in the AMLCG cohort (on the left) and AMLSG cohort (on the right). The models include all variables with $p < 0.1$ in the single Cox regression models of the primary cohort (AMLCG cohort). CN-AML cytogenetically normal AML, LDH lactate dehydrogenase, pAML primary AML, sAML secondary AML, tAML therapy-related AML, WBC white blood cells.

all SF mutant populations presented with decreased expression of both exonic (67.7–81.3%) and intronic (56.2–81.0%) segments. Both the number of differentially expressed segments and the amount of downregulated segments was most modest in *SRSF2*(P95H) mutants (1464 total segments, 61.8% downregulated) and most extreme in *SF3B1*(K700E) mutants (9853 total segments, 81.2% downregulated) following the trend observed in the DSJU analysis.

In an additional step, the splice junction counts reported by Okeyo-Owuor et al. were used to detect DSJU between CD34+ cells with U2AF1(S34) mutations ($n = 3$) and SF wildtype ($n = 3$) via the same pipeline applied to the AMLCG and Beat AML cohorts. While no identical junctions were differentially used in all three datasets, 16 genes were reported as differentially spliced in all, including leukemia or cancer-associated genes (*ABI1*, *DEK*, *HP1BP3*, *MCM3*, and *SET*), as well as *HNRNPK* (a major pre-mRNA binding protein), thereby further refining our list of genes with strong evidence of differential splicing between *U2AF1*(S34F) mutants and SF wildtype samples (Table S14).

Pathway analysis of genes dysregulated in SF mutants

We systematically compared genes with at least one DE isoform and those reported as differentially spliced in all SF mutation subgroups. For *SRSF2* mutants, genes significant in both analyses included *EWSR1*, *H1F0*, *INTS3*, and *YBX1*. In general, out of the genes examined in both analyses only 9.8–23.3% (depending on the SF mutation) of genes reported as having a DE isoform were also reported as being differentially spliced. Conversely, 3.3–28.5% of differentially spliced genes were also reported as having a DE isoform. These findings suggest

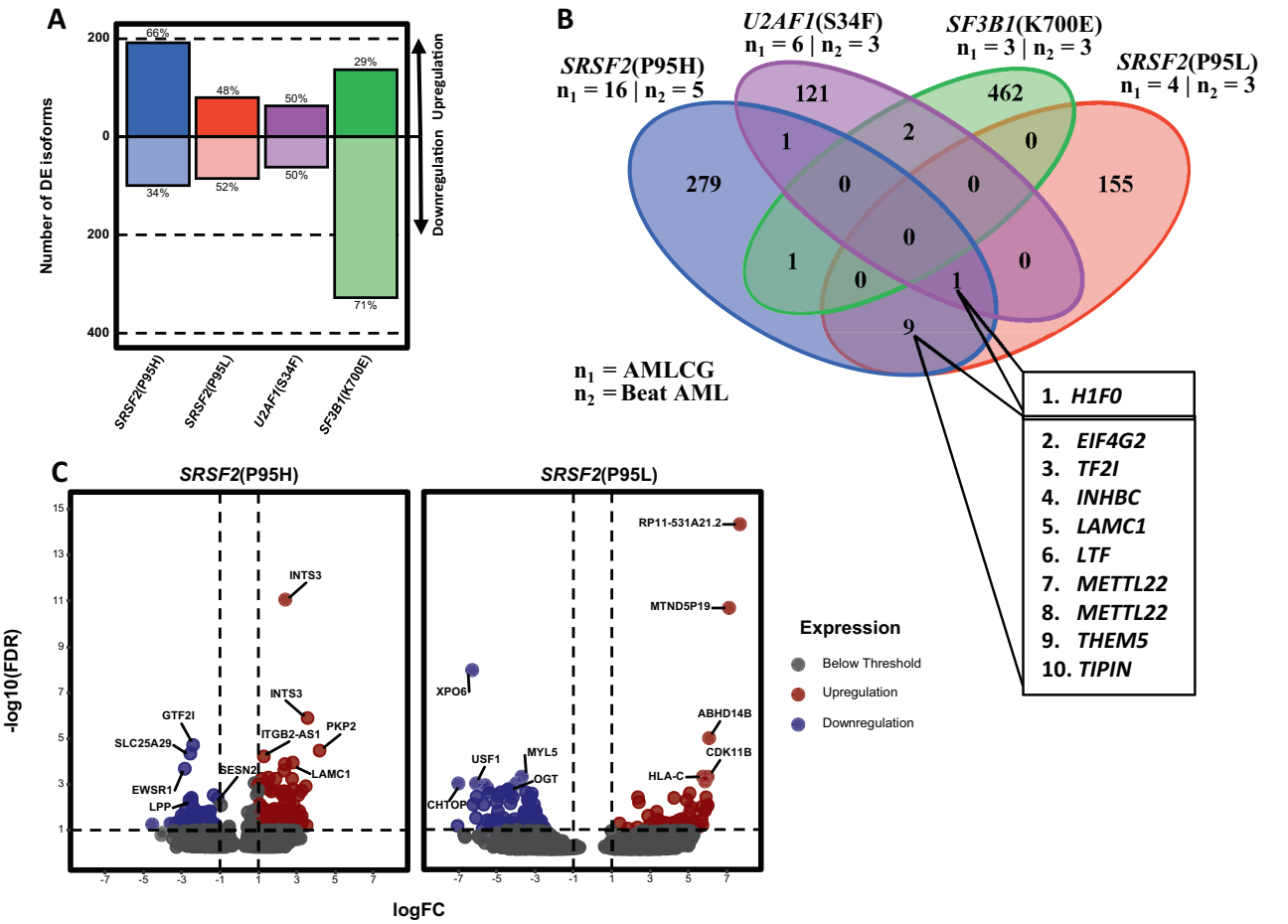

Fig. 4 Differential isoform expression analysis in the AMLCG cohort. a Number of differentially expressed isoforms for each SF point mutation. **b** Isoforms reported as differentially expressed in the same direction among patients with different SF mutations vs. SF wildtype patients. HGNC symbols of the genes in which the common isoforms are located are shown. The corresponding Ensembl isoform identifiers are: 1. ENST00000340857, 2. ENST00000481621, 3. ENST00000309668, 4. ENST00000258341, 5. ENST00000426532, 6. ENST00000562151, 7. ENST00000564133, 8. ENST00000368817, 9. ENST00000566524, 10. ENST00000530211. **c** Volcano plots showing the magnitude of differential isoform expression for *SRSF2*(P95H) and *SRSF2*(P95L). The *x*-axis corresponds to the $\log_2$(fold change) of each isoform between mutated and wildtype samples, while the y-axis corresponds to –log10(FDR), where FDR represents the adjusted *p*-value for each isoform (False Discovery Rate).

that differential gene splicing does not always lead to altered isoform expression while at the same time differential isoform expression cannot always be attributed to an explicit splicing alteration. Considering the complementary nature of the analyses, we performed gene ontology (GO) analysis by combining the genes with evidence of differential isoform usage or differential splicing. Interestingly, GO terms enriched for both *SRSF2* mutants included "mRNA splicing, via spliceosome" ($p < 0.001$ and $p = 0.046$, respectively) and "mRNA splice site selection" ($p = 0.022$ and $p = 0.019$, respectively; Fig. 6c and Tables S15.1–S15.7).

Since the splicing pathway was enriched in the genes dysregulated in both *SRSF2* mutants, we cross-referenced our differential expression and differential splicing analysis results with a list of all genes involved in splicing. Of the 317 splicing-related genes expressed in our dataset, 101 were dysregulated in at least one SF mutant group. On average 30.5 (range 6–52) splicing-related genes were dysregulated per SF point mutation. Of note, both *SRSF2* point mutations associated with differential splicing of *HNRNPA1* and *HNRNPUL1*, as well as *PCF11* and *TRA2A*. Interestingly, one of the differential splicing events reported in both *SRSF2* mutants involved the under-usage of the same novel splice junction in *TRA2A* (Fig. S8). *TRA2A* has previously been shown to be differentially spliced in mouse embryo fibroblasts upon *SRSF2* knockout [43]. Furthermore, it has been shown that both HNRNPA1 and SRSF2 interact with the loop 3 region of 7SK RNA and by favoring the dissociation of SRSF2, HNRNPA1 may lead to the release of active P-TEFb [44]. Taken together, our results indicate a strong dysregulation of the splicing pathway in SF mutants including several genes closely associated with SRSF2.

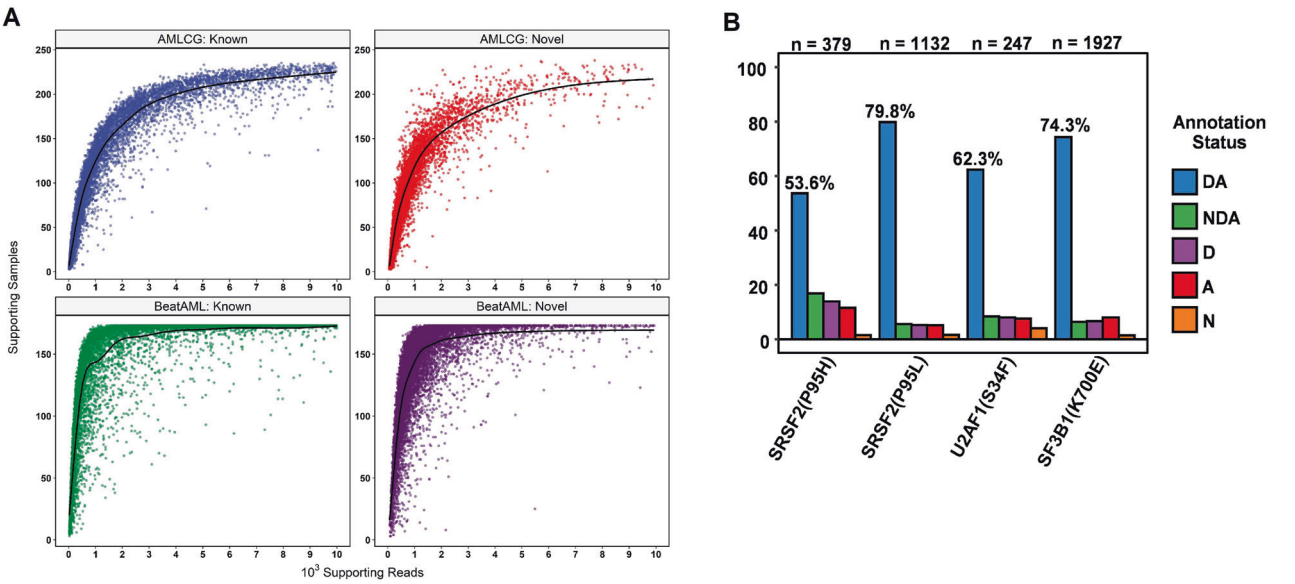

Fig. 5 Differential splice junction usage. a Scatterplot displaying the number of samples as well as the total number of reads supporting each splice junction, separately for known and novel splice junctions in both RNA-Seq datasets. To preserve visibility 20,000 random junctions are shown for each group. **b** Barchart showing the annotation status of splice junctions reported as differentially used. Novel splice junctions where classified into five groups based on their annotation status as described previously [15] ("DA": annotated junctions, "NDA": unknown combination of known donor and acceptor sites, "D": known donor, but novel acceptor site, "A": novel donor, but known acceptor site, "N": previously unknown donor and acceptor site).

Clinical relevance of differential splice junction usage

We examined the potential clinical relevance of DSJU by constructing single Cox regression models to predict OS using splice junction usage as the predictor variable. All junctions with validated differential usage in at least one SF mutant population were considered ($n = 299$). Out of these, 12 significantly impacted OS (adjusted $p < 0.1$). This subset of junctions was used to construct identical models in the Beat AML cohort. Two annotated splice junctions in the genes *EVL* and *NBEAL2* remained significant after *p*-value adjustment in both cohorts ($p < 0.1$, Fig. 7a, b), both of which were overused in *SRSF2*(P95H) mutants. Interestingly, the junction in *EVL* was used in only 42.7% of the SF wildtype samples in the AMLCG cohort (49.4% in the Beat AML cohort) but was used in most SF mutant samples (AMLCG: 73.2%, Beat AML: 89.5%). In contrast, the junction in *NBEAL2* did not present with significantly increased usage in SF mutated samples. A subsequent analysis using Kaplan–Meier curves and log-rank testing confirmed the significant impact on OS (junction in *EVL*: $p < 0.001$; junction in *NBEAL2*: $p = 0.020$).

Discussion

The clinical relevance of SF mutations and their aberrant splicing patterns have been explored in myelodysplasia, while comparable data for AML is lacking. In this study, we examined two AML patient cohorts, encompassing a total of 2678 patients from randomized prospective trials, to characterize SF mutations clinically. This analysis was complemented by RNA-sequencing analysis of two large datasets to reveal targets of aberrant splicing in AML.

We show that SF mutations are frequent alterations in AML, identified in 21.4% of our primary patient cohort, especially in elderly patients and in secondary AML. SF mutations are associated with other recurrent mutations in AML, such as *BCOR* and *RUNX1* mutations, however *SRSF2*(P95L) mutations co-occur less often with those mutations when compared to *SRSF2*(P95H) mutations, albeit showing a slightly increased mutational load. This suggests a more diverse co-expression profile of *SRSF2* (P95L).

Previous studies have demonstrated the predictive value of SF mutations in clonal haematopoiesis of indeterminate potential (CHIP) [45], MDS [6, 8, 46–48], and AML [10, 18, 19, 49]. However, survival analyses in AML were, in their majority, hampered by small sample sizes and limited availability of further risk factors. Therefore, we examined whether SF mutations impact survival while accounting for recently proposed risk parameters included in the ELN 2017 classification [50]. In our analysis, *SRSF2* and *SF3B1* mutations were no independent prognostic markers for OS in AML. *U2AF1*(S34F) mutations displayed poor OS in the AMLCG cohort, which we were unable to validate in the AMLSG cohort. The discrepancy in survival

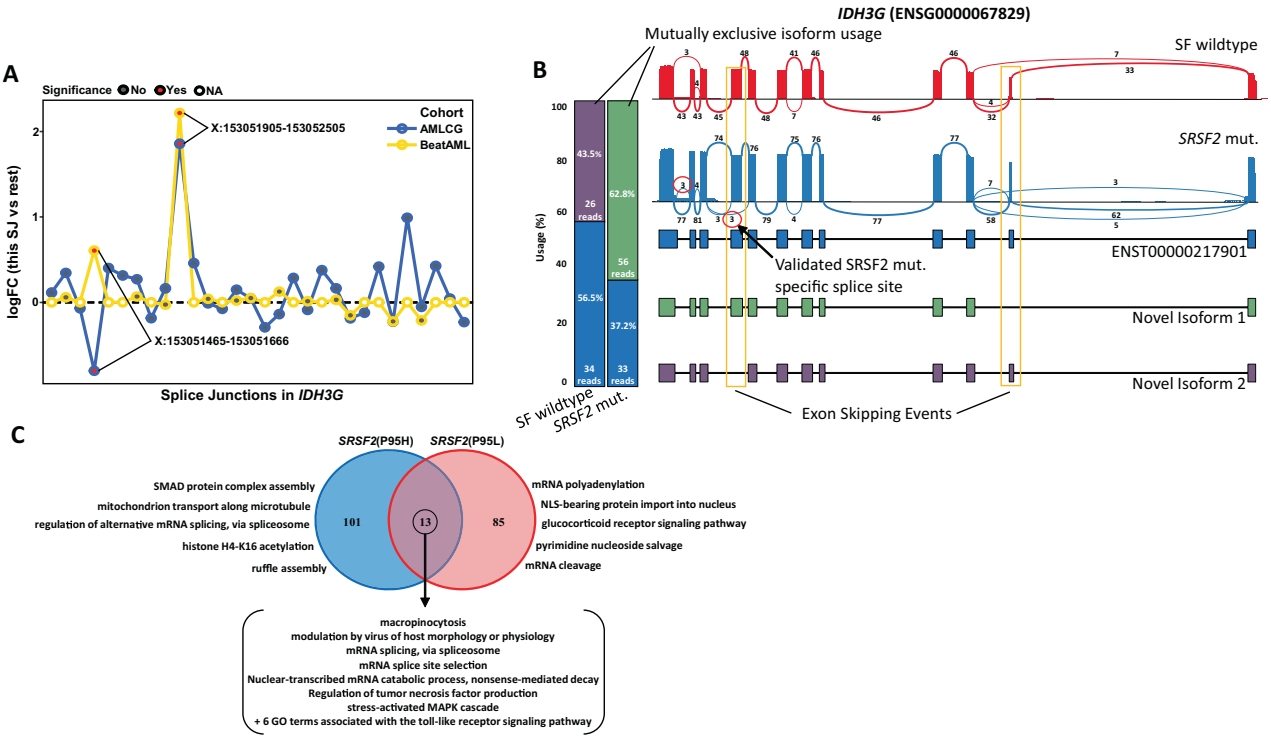

Fig. 6 Splicing dysregulation in *SRSF2* mutants. a DSJU of all splice junctions inside *IDH3G* (ENSG00000067829) are shown for *SRSF2*(P95H) mutants compared to SF wildtype patients in the AMLCG and Beat AML cohort. The *y*-axis denotes the $\log_2$(fold-change) (logFC) of the usage of each splice junction. To determine significance the logFC of each splice junction (SJ) is compared to the logFC of all other junctions inside the same gene. The *x*-axis denotes individual splice junctions. Two junctions (defined by their coordinates) in *IDH3G* were significant in both cohorts, however the first one (X:153051465–153051666) showed a differential usage in opposing directions among the two cohorts. **b** Nanopore sequencing results (for the gene *IDH3G*) of one *SRSF2*(P95H) mutated sample and one SF wildtype sample validating the differential usage of novel splice junctions. FLAIR distribution of isoforms is shown on the left. The two samples shared the usage of one known isoform and in addition showed usage of two mutually exclusive novel isoforms (exon composition of the isoforms is shown in the lower section). Above a sashimi plot is shown, depicting the read coverage of the exons and splice junctions in each sample. The yellow boxes highlight examples of exon skipping. Circled in red are the splice junctions that were found to be differentially used between *SRSF2*(P95H) mutants and SF wildtype patients in both cohorts, as shown in **a**. The black arrow indicates the validated novel splice junction with a congruent logFC for both cohorts. **c** Overlap of GO terms between *SRSF2*(P95H) and *SRSF2*(P95L) mutants for the "biogical process" domain. Additional terms associated with the toll-like receptor signaling pathway: GO:0034142, GO:0034146, GO:0034162, GO:0034166, GO:0038123, and GO:0038124. Additionally, the top five most representative (≤25 genes) terms for each mutation are shown.

of SF mutated patients between the two cohorts lied most likely in the large age difference of the participants (median age difference of 8 years), which also led to a higher percentage of patients receiving allogeneic transplants in the AMLSG cohort (56.5% vs. 30.6% in the AMLCG cohort). In summary, SF mutations are early evolutionary events and define prognosis and transformation risk in CHIP and MDS patients, yet there is no clear independent prognostic value of SF mutations in AML.

Two large RNA-sequencing studies have been performed previously, to detect aberrantly spliced genes in SF mutants, both of which focused on MDS patients [41, 42]. In this study we described a distinct differential isoform expression profile for each SF point mutation. Furthermore, we evaluated differential splicing for the four most common SF point mutations via a customized pipeline to determine differential usage of both known and novel splice junctions. Our pipeline enables the differential quantification of individual splice junctions without restricting the analysis to annotated alternative splicing events. We argue that the strength of our analysis lies in the accurate detection of single dysregulated junctions (especially in cases where splice sites are shared by multiple junctions) in an annotation-independent manner achieving validation rates up to 74.0% in our largest mutant sample group (*SRSF2*, $n = 19$). Limitations of the analysis include the restriction to junctions with both splice sites within the same gene (a restriction shared by most differential splicing algorithms) and genes with at least two junctions. However, the reduced requirements of our analysis could prove valuable in the study of differential splicing in organisms with lacking annotation.

All SF point mutations shared a tendency towards decreased splice junction usage, which did not affect the

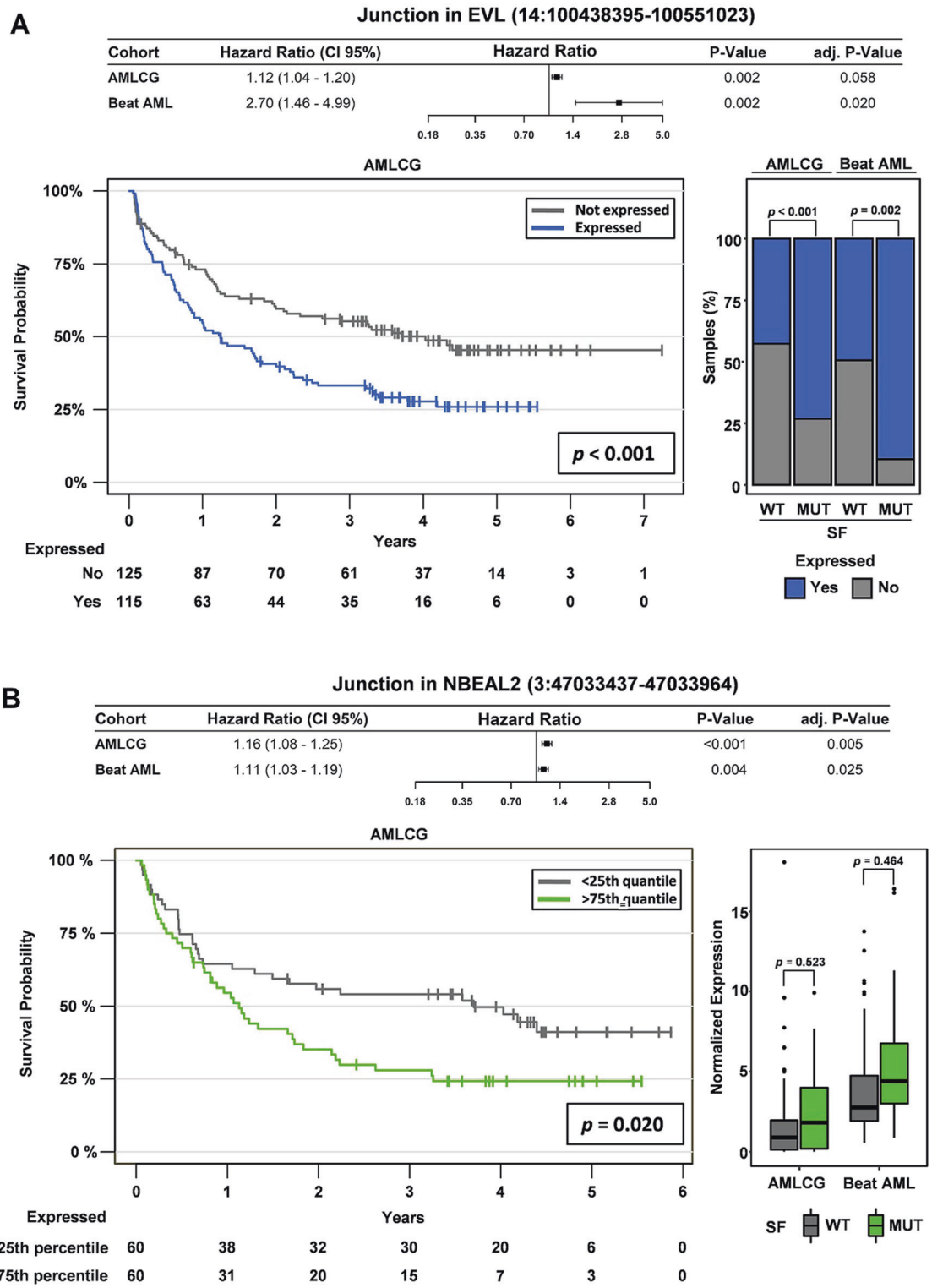

Fig. 7 Impact of differential splice junction usage on overall survival. Cox regression analysis on all significantly differentially used splice junctions (in any SF mutant) revealed that the usage of one splice junction in *EVL* (genomic coordinates: 14:100438395–100551023, **a** and one junction in *NBEAL2* (3:47033437–46933964, **b** associated with inferior OS. Hazard ratios (per normalized expression unit) shown in the upper panels. KM-graphs with *p*-values corresponding to log-rank tests depict the association with OS. Right panels visualize the junction usage in SF wildtype and SF mutated samples.

global number of splicing events in SF mutants. For two junctions in the genes *EVL* and *NBEAL2*, which are significantly overused in *SRSF2*(P95H) mutants, we were able to show a robust association with OS in both datasets. Furthermore, usage of the junction in *EVL* was clearly associated with the presence of SF mutations. While no confounding variables were considered for this analysis, it justifies the study of dysregulated splicing patterns as a means of identifying patients with poor prognosis. We note that the available coverage of the examined RNA-sequencing datasets permitted the study of a limited amount of splice junctions with high accuracy. More recent sequencing methods like nanopore sequencing are likely to capture additional, clinically relevant, aberrant splicing events along with their functional consequences (e.g., novel isoform expression). The potential of splice junction usage for risk prediction in AML has recently been demonstrated by collaborators [51].

Surprisingly, we observed a limited overlap between genes with differentially expressed isoforms and differentially spliced genes. In addition, a recent study by Liang et al. reported that the majority of differential binding events in *SRSF2*(P95H) mutants do not translate to alternative splicing [52]. Taken together, these findings indicate a "selection" or possibly a compensation of deregulatory events from differential binding through differential splicing to finally differential isoform expression. Furthermore, the enrichment of aberrant splicing in splicing-related genes opens the possibility of a cascading effect on transcription via the differential alternative splicing of transcriptional components. A congruent hypothesis was stated by Liang et al., where an enrichment of *SRSF2*(P95H) targets in RNA processing and splicing was shown, further supporting the notion of an indirect effect of mutant SRSF2 facilitated through additional splicing components. Future investigations may provide a mechanistic link between the differential splicing of selected genes and the impairment of transcription and specifically transcriptional pausing observed in SF mutant cells, which contributes to the MDS phenotype [32].

To the best of our knowledge our study represents the most comprehensive analysis of SF mutations in AML to date, both in terms of clinical characterization and differential splice junction usage. This enabled us to study *SRSF2*(P95H) and *SRSF2*(P95L) separately, thereby not only outlining their differences but also identifying common and likely core targets of differential splicing in *SRSF2* mutants, two of which presented with clinical significance. We conclude that SF mutated patients represent a distinct subgroup of AML patients with poor prognosis that is not attributable solely to the presence of SF mutations. SF mutations induce aberrant splicing throughout the genome including the dysregulation of several genes associated with AML pathogenesis, as well as a number of genes with immediate, functional implications on splicing and transcription. Further studies are required to identify which splicing events are critical in leukaemogenesis and whether they are accessible to new treatments options, such as splicing inhibitors [53] and immunotherapeutic approaches.

Data availability

Read counts and sample characteristics are available in the GEO database (GSE146173). Law restrictions prohibit us from publicly sharing raw sequencing data, which however can be made available upon reasonable request and permission of the local ethics committee.

Acknowledgements The authors thank all participants and recruiting centers of the AMLCG, BEAT and AMLSG trials.

Funding This work is supported by a grant of the Wilhelm-Sander-Stiftung (no. 2013.086.2) and the Physician Scientists Grant (G-509200–004) from the Helmholtz Zentrum München to T.H. and the German Cancer Consortium (Deutsches Konsortium für Translationale Krebsforschung, Heidelberg, Germany). K.H.M., K.S., and T.H. are supported by a grant from Deutsche Forschungsgemeinschaft (DFG SFB 1243, TP A06 and TP A07). S.K.B. is supported by Leukaemia & Blood Cancer New Zealand and the family of Marijanna Kumerich. A.M.N.B. is supported by the BMBF grant 01ZZ1804B (DIFUTURE).

Author contributions S.A.B., A.M.N.B., and T.H. conceived and designed the analysis. S.A.B., A.M.N.B., V.J., M.R.-T., H.J., A.G., S.C., N.K., K.S., K.H.M., and T.H. provided and analyzed data. A.M.N.B., V.J., and U.M. provided bioinformatics support. J.P.-M., S.K., and H.B. managed the HiSeq 1500 instrument and the RNA-sequencing of the AMLCG samples. M.R.-T., H.J., B.K., S.S., N.K., S.K.B., K.H.M., and K.S. characterized patient samples; M.C.S., D.G., W.B., B.W., J.B., and W.H. coordinated the AMLCG clinical trials. S.A.B. and T.H. wrote the manuscript. All authors approved the final manuscript.

Compliance with ethical standards

Conflict of interest H.J. is a current employee of Roche Pharma AG, Grenzach-Wyhlen, Germany. The authors declare that they have no conflict of interest.

References

1. Papaemmanuil E, Gerstung M, Malcovati L, Tauro S, Gundem G, Van Loo P, et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood 2013;122:3616–27.
2. Metzeler KH, Herold T, Rothenberg-Thurley M, Amler S, Sauerland MC, Görlich D, et al. Spectrum and prognostic relevance of driver gene mutations in acute myeloid leukemia. Blood 2016;128:686–98.
3. Makishima H, Visconte V, Sakaguchi H, Jankowska AM, Kar SA, Jerez A, et al. Mutations in the spliceosome machinery, a novel

and ubiquitous pathway in leukemogenesis. Blood 2012;119:3203–10.
4. Larsson CA, Cote G, Quintás-Cardama A. The changing mutational landscape of acute myeloid leukemia and myelodysplastic syndrome. Mol Cancer Res. 2013;11:815–27.
5. Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, Yamamoto R, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. Nature 2011;478:64–9.
6. Thol F, Kade S, Schlarmann C, Löffeld P, Morgan M, Krauter J, et al. Frequency and prognostic impact of mutations in SRSF2, U2AF1, and ZRSR2 in patients with myelodysplastic syndromes. Blood 2012;119:3578–84.
7. Dolatshad H, Pellagatti A, Fernandez-Mercado M, Yip BH, Malcovati L, Attwood M, et al. Disruption of SF3B1 results in deregulated expression and splicing of key genes and pathways in myelodysplastic syndrome hematopoietic stem and progenitor cells. Leukemia 2015;29:1092–103.
8. Wu S, Kuo Y, Hou H, Li L, Tseng M, Huang C, et al. The clinical implication of SRSF2 mutation in patients with myelodysplastic syndrome and its stability during disease evolution. Blood 2014;120:3106–12.
9. Graubert TA, Shen D, Ding L, Okeyo-owuor T, Cara L, Shao J, et al. Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. Nat Genet 2012;44:53–7.
10. Hou H-A, Liu C-Y, Kuo Y-Y, Chou W-C, Tsai C-H, Lin C-C, et al. Splicing factor mutations predict poor prognosis in patients with de novo acute myeloid leukemia. Oncotarget 2016;7:9084–101.
11. Cho Y-U, Jang S, Seo E-J, Park C-J, Chi H-S, Kim D-Y, et al. Preferential occurrence of spliceosome mutations in acute myeloid leukemia with a preceding myelodysplastic syndrome and/or myelodysplasia morphology. Leuk Lymphoma 2014;8194:1–25.
12. Moon H, Cho S, Loh TJ, Jang HN, Liu Y, Choi N, et al. SRSF2 directly inhibits intron splicing to suppresses cassette exon inclusion. BMB Rep. 2017;50:423–8.
13. Kim E, Ilagan JO, Liang Y, Daubner GM, Lee SCW, Ramakrishnan A, et al. SRSF2 mutations contribute to myelodysplasia by mutant-specific effects on exon recognition. Cancer Cell 2015;27:617–30.
14. Alsafadi S, Houy A, Battistella A, Popova T, Wassef M, Henry E, et al. Cancer-associated SF3B1 mutations affect alternative splicing by promoting alternative branchpoint usage. Nat Commun 2016;7:10615.
15. Okeyo-Owuor T, White BS, Chatrikhi R, Mohan DR, Kim S, Griffith M, et al. U2AF1 mutations alter sequence specificity of pre-mRNA binding and splicing. Leukemia 2015;29:909–17.
16. Przychodzen B, Jerez A, Guinta K, Sekeres MA, Padgett R, Maciejewski JP, et al. Patterns of missplicing due to somatic U2AF1 mutations in myeloid neoplasms. Blood 2013;122:999–1006.
17. Shirai CL, Ley JN, White BS, Kim S, Tibbitts J, Shao J, et al. Mutant U2AF1 expression alters hematopoiesis and pre-mRNA splicing in vivo. Cancer Cell 2015;27:631–43.
18. Yang J, Yao D, Ma J, Yang L, Guo H, Wen X, et al. The prognostic implication of SRSF2 mutations in Chinese patients with acute myeloid leukemia. Tumor Biol 2016;37:10107–14.
19. Papaemmanuil E, Gerstung M, Bullinger L, Gaidzik VI, Paschka P, Roberts ND, et al. Genomic classification and prognosis in acute myeloid leukemia. N. Engl J Med. 2016;374:2209–21.
20. Tyner JW, Tognon CE, Bottomly D, Wilmot B, Kurtz SE, Savage SL, et al. Functional genomic landscape of acute myeloid leukaemia. Nature 2018;562:526–31.
21. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21.
22. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 2017;14:417–9.
23. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43:e47–e47.
24. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol 2010;11:R25.
25. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2009;26:139–40.
26. Liu R, Holik AZ, Su S, Jansz N, Chen K, Leong HSan, et al. Why weight? Modelling sample and observational level variability improves power in RNA-seq analyses. Nucleic Acids Res. 2015;43:e97.
27. Law CW, Chen Y, Shi W, Smyth GK. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol 2014;15:R29.
28. Anders S, Reyes A, Huber W. Detecting differential usage of exons from RNA-seq data. Genome Res 2012;22:2008–17.
29. Tang AD, Soulette CM, van Baren MJ, Hart K, Hrabeta-Robinson E, Wu CJ. et al. Full-length transcript characterization of SF3B1 mutation in chronic lymphocytic leukemia reveals downregulation of retained introns. Nat Commun. 2020;11:1–12.
30. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. 2017.
31. Herold T, Metzeler KH, Vosberg S, Hartmann L, Ollig C, Olzel FS. et al. Isolated trisomy 13 defines a homogeneous AML subgroup with high frequency of mutations in spliceosome genes and poor prognosis. Blood J Am Soc Hematol 2014;124:1304–11.
32. Chen L, Chen J-Y, Huang Y-J, Gu Y, Qiu J, Qian H, et al. The augmented R-loop is a unifying mechanism for myelodysplastic syndromes induced by high-risk splicing factor mutations. Mol Cell 2018;69:412–25.
33. Yoshimi A, Lin K-T, Wiseman DH, Rahman MA, Pastore A, Wang B, et al. Coordinated alterations in RNA splicing and epigenetic regulation drive leukaemogenesis. Nature 2019;574:273–7.
34. Roe J-S, Vakoc CR. The essential transcriptional function of BRD4 in acute myeloid leukemia. Cold Spring Harb Symp Quant Biol. 2016;81:61–6.
35. Endo A, Tomizawa D, Aoki Y, Morio T, Mizutani S, Takagi M. EWSR1/ELF5 induces acute myeloid leukemia by inhibiting p53/p21 pathway. Cancer Sci 2016;107:1745–54.
36. Perner F, Jayavelu AK, Schnoeder TM, Mashamba N, Mohr J, Hartmann M, et al. The cold-shock protein Ybx1 is required for development and maintenance of acute myeloid leukemia (AML) in vitro and in vivo. Blood. 2017;130:792.
37. McNerney ME, Brown CD, Wang X, Bartom ET, Karmakar S, Bandlamudi C, et al. CUX1 is a haploinsufficient tumor suppressor gene on chromosome 7 frequently inactivated in acute myeloid leukemia. Blood 2013;121:975–83.
38. McGarvey T, Rosonina E, McCracken S, Li Q, Arnaout R, Mientjes E, et al. The acute myeloid leukemia-associated protein, DEK, forms a splicing-dependent interaction with exon-product complexes. J Cell Biol. 2000;150:309–20.
39. Fujita S, Honma D, Adachi N, Araki K, Takamatsu E, Katsumoto T, et al. Dual inhibition of EZH1/2 breaks the quiescence of leukemia stem cells in acute myeloid leukemia. Leukemia 2018;32:855–64.
40. Pallarès V, Hoyos M, Chillón MC, Barragán E, Prieto Conde MI, Llop M, et al. Focal adhesion genes refine the intermediate-risk

cytogenetic classification of acute myeloid leukemia. Cancers. 2018;10:E436.
41. Shiozawa Y, Malcovati L, Gallì A, Sato-Otsubo A, Kataoka K, Sato Y, et al. Aberrant splicing and defective mRNA production induced by somatic spliceosome mutations in myelodysplasia. Nat Commun 2018;9:3649.
42. Pellagatti A, Armstrong RN, Steeples V, Sharma E, Repapi E, Singh S, et al. Impact of spliceosome mutations on RNA splicing in myelodysplasia: dysregulated genes/pathways and clinical associations. Blood 2018;132:1225–40.
43. Skrdlant L, Stark JM, Lin R-J. Myelodysplasia-associated mutations in serine/arginine-rich splicing factor SRSF2 lead to alternative splicing of CDC25C. BMC Mol Biol. 2016;17:18.
44. Lemieux B, Blanchette M, Monette A, Mouland AJ, Wellinger RJ, Chabot B. A function for the hnRNP A1/A2 proteins in transcription elongation. PLoS ONE. 2015;10:e0126654.
45. Abelson S, Collord G, Ng SWK, Weissbrod O, Mendelson Cohen N, Niemeyer E, et al. Prediction of acute myeloid leukaemia risk in healthy individuals. Nature 2018;559:400–4.
46. Malcovati L, Papaemmanuil E, Bowen DT, Boultwood J, Della Porta MG, Pascutto C, et al. Clinical significance of SF3B1 mutations in myelodysplastic syndromes and myelodysplastic/myeloproliferative neoplasms. Blood 2011;118:6239–46.
47. Papaemmanuil E, Cazzola M, Boultwood J, Malcovati L, Vyas P, Bowen D, et al. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. N. Engl J Med. 2011;365:1384–95.
48. Wu L, Song L, Xu L, Chang C, Xu F, Wu D, et al. Genetic landscape of recurrent ASXL1, U2AF1, SF3B1, SRSF2, and EZH2 mutations in 304 Chinese patients with myelodysplastic syndromes. Tumor Biol 2016;37:4633–40.
49. Zhang S-J, Rampal R, Manshouri T, Patel J, Mensah N, Kayserian A, et al. Genetic analysis of patients with leukemic transformation of myeloproliferative neoplasms shows recurrent SRSF2 mutations that are associated with adverse outcome. Blood 2012;119:4480–5.
50. Döhner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Büchner T, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood. 2017;129:424–47.
51. Anande G, Deshpande NP, Mareschal S, Batcha AMN, Hampton HR, Herold T, et al. RNA splicing alterations induce a cellular stress response associated with poor prognosis in AML. bioRxiv. 2020;2020.01.10.895714.
52. Liang Y, Tebaldi T, Rejeski K, Joshi P, Stefani G, Taylor A, et al. SRSF2 mutations drive oncogenesis by activating a global program of aberrant alternative splicing in hematopoietic cells. Leukemia 2018;32:2659–71.
53. Lee SCW, Abdel-Wahab O. Therapeutic targeting of splicing in cancer [Internet]. Nat Med NIH Public Access. 2016;22:976–86.

Affiliations

Stefanos A. Bamopoulos[1,2] · Aarif M. N. Batcha[3,4] · Vindi Jurinovic[1,3] · Maja Rothenberg-Thurley[1] · Hanna Janke[1] · Bianka Ksienzyk[1] · Julia Philippou-Massier[5] · Alexander Graf[5] · Stefan Krebs[5] · Helmut Blum[5] · Stephanie Schneider[1,6] · Nikola Konstandin[1] · Maria Cristina Sauerland[7] · Dennis Görlich[7] · Wolfgang E. Berdel[8] · Bernhard J. Woermann[9] · Stefan K. Bohlander[10] · Stefan Canzar[11] · Ulrich Mansmann[3,4,12,13] · Wolfgang Hiddemann[1,12,13] · Jan Braess[14] · Karsten Spiekermann[1,12,13] · Klaus H. Metzeler[1,12,13] · Tobias Herold[1,12,15]

[1] Laboratory for Leukemia Diagnostics, Department of Medicine III, University Hospital, LMU Munich, Munich, Germany

[2] Department of Hematology, Oncology and Tumor Immunology (Campus Benjamin Franklin), Charité University Medicine Berlin, Berlin, Germany

[3] Institute for Medical Information Processing Biometry and Epidemiology, LMU Munich, Munich, Germany

[4] DIFUTURE (Data Integration for Future Medicine (DiFuture, www.difuture.de), LMU Munich, Munich, Germany

[5] Laboratory for Functional Genome Analysis (LAFUGA), Gene Center, LMU Munich, Munich, Germany

[6] Institute of Human Genetics, University Hospital, LMU Munich, Munich, Germany

[7] Institute of Biostatistics and Clinical Research, University of Münster, Münster, Germany

[8] Department of Medicine, Hematology and Oncology, University of Münster, Münster, Germany

[9] German Society of Hematology and Oncology, Berlin, Germany

[10] Leukemia and Blood Cancer Research Unit, Department of Molecular Medicine and Pathology, University of Auckland, Auckland, New Zealand

[11] Gene Center, LMU Munich, Munich, Germany

[12] German Cancer Consortium (DKTK), Partner Site Munich, Munich, Germany

[13] German Cancer Research Center (DKFZ), Heidelberg, Germany

[14] Department of Oncology and Hematology, Hospital Barmherzige Brüder, Regensburg, Germany

[15] Research Unit Apoptosis in Hematopoietic Stem Cells, Helmholtz Zentrum München, German Center for Environmental Health (HMGU), Munich, Germany