Introduction

Hematological malignancies are a group of malignant diseases originating from blood, bone marrow cells or lymph, including leukemia [myeloid originated: chronic myeloid leukemia (CML), acute myeloid leukemia (AML); lymphoid originated: chronic lymphoblastic leukemia (CLL) and acute lymphoblastic leukemia (ALL)], plasma multiple myeloma (MM) (plasma cell originated) and lymphoma [non-Hodgkin lymphoma (NHL) and Hodgkin lymphoma (HL)] (1). Depending on the type of leukemia and the age of the patient at diagnosis, the prognosis of leukemia patients differs significantly. However, in general, leukemia is the 5th and 6th most common cancer death in men and women, respectively (Facts and Statistics 2015, Leukemia and Lymphoma Society). AML remains one of the worst clinically devastating diseases. The 5-year
overall survival rate is about 20% in adult AML patients (2,3). These poor outcomes highlight the unmet need for a better understanding of leukemogenesis and novel, targeted therapies to replace chemotherapy, which has not been changed for more than four decades (4-7).

Leukemia stem cell (LSC) was described as source of origin and progression of AML in more than 10 years ago (8). Accumulating evidence supports the fact that these LSC populations acquire self-renewal function and sustain the disease (9,10). They are rare, but functionally and phenotypically different from bulk of blast cells (11,12). AML LSCs are generally insensitive to the conventional chemotherapy, instead, they are more enriched after chemotherapy (13,14). Physically, LSC populations reside in the bone marrow microenvironment and are poised to propagate, leading to the therapy failure and disease relapse (15-18). Taken together, these discoveries indicate that the LSC is the culprit for the dismal prognosis of AML and selectively targeting LSC could be a promising strategy for AML treatment (19-23).

The advent of next generation sequencing (NGS) technologies has revolutionized the field of genomics, enabling us to gain a deep and broad understanding of human diseases on a whole genome level, including cancers (24-26). Recent studies using NGS platform have revealed substantially frequency of recurrent somatic points and copy number changes in genes associated with spliceosome in leukemias and myelodysplastic syndromes (MDS), which are chronic myeloid neoplasms, often progressing to AML (27-30). This review will first describe the complex and general function of RNA spliceosome. We will outline various mis-spliced mRNA in AML and their clinical significance, followed by summary of mutations in spliceosome and related factors, particularly its role in AML LSC. Finally, we will discuss the promises and challenges in an attempt to exploit spliceosomal machinery therapeutically.

The spliceosomal machinery and RNA splicing regulation

In mammalian cells, genes are transcribed as messenger RNA precursors (pre-mRNAs), containing introns, which are intervening sequences, noncoding regions (31). RNA splicing, the process from nuclear pre-mRNA into mature mRNA where introns are excised and the exons (coding regions) are joined together is mediated by a large complex, called spliceosome or spliceosomal machinery (32,33). In human genome, 60% of the mature mRNAs are spliced by alternative splicing in which pre-mRNAs can be spliced in more than one way (34,35). Alternative splicing tremendously increases the diversity of human transcriptome, resulting in translation of more complex proteome in human. Alternative splicing also acts as a form of gene regulation because dominant negative protein translated from alternative spliced mRNA can inhibit its wild type protein. Thus, RNA splicing ultimately influences assorted cellular functions, tissue specificity, and developmental states of human.

The spliceosome is a large molecular machinery, consisting of five small nuclear ribonucleoproteins (snRNPs) (RNA-protein complex), U1, U2, U4, U5 and U6, as well as about 200 associated protein factors, such as serine and arginine-rich proteins (SR proteins) and heterogeneous nuclear ribonucleoprotein (hnRNP) family (36,37). The boundaries between introns and exons of the pre-mRNAs are distinguished by specific nucleotide sequences. The 5’ splice site at the 5’ (left) end of the intron includes almost an invariant sequence GU, while the splice acceptor site at the 3’ splice site (right) end of the intron includes an consensus sequence AG (GU-AG rule) (38). In general, the RNA splicing process involves three steps, recognizing the appropriate splice sites, bringing those sites together, and catalyzing the splicing reactions. U1 snRNP initiates splicing by recognizing the 5’ splice site and binding it through an RNA-RNA base pairing reaction (canonical Watson-Crick or wobble base pairs), resulting in the assembly of the commitment (E) complex. U1 snRNP plays an important role in the binding of U2 snRNP to the branch point region of the intron. Subunit of U2 snRNP, U2 auxiliary factors (U2AFs) bind to polypyrimidine tract region and the other to the intron’s 3’ splice site-AG dinucleotide. The association of both of the U1 snRNP and U2 snRNP assembles the pre-spliceosome (complex A). U4, U5 and U6 form the tri-snRNP complex. When pre-spliceosome recruits this tri-snRNP complex, it converts the A complex into B complex, which comprises all the necessary splicing components and poises to catalyze. The B complex passes through a series of RNA-RNA rearrangements and transform into C complex, in which U2 and U6 are brought together to create the catalytic active site. In human, 99% of the introns is cleaved by this way (major spliceosome). An alternative splicing pathway uses the minor spliceosome which contains U12 and another set of snRNPs in a similar fashion. For the purpose of this review, here we only briefly outline the spliceosome
formation and splicing process. We recommend a number of comprehensive and elegant reviews for a detailed coverage of this topic (39-42).

**Abnormally spliced (AS) mRNAs in AML**

Before the advent of NGS technology, abnormal RNA splicing of certain oncogenes, tumor suppressor gene and epigenetic regulators had been noticed in AML, although in an unsystematic fashion. Here we have concisely summary some of these main findings.

Myc gene belongs to Myc family of transcription factors, which comprises C-Myc, N-Myc and L-Myc genes. Myc gene is a prominent oncogene and has been implicated in transformation of many types of cancers, including leukemia (43). A unique pattern of L-Myc mRNA processing with 40% of them lacking exon III and intron I has been revealed in AML (44). Importantly, L-Myc expression is very low in adult bone marrow and in fetal spleen and thymus. However, whether this splicing variant functionally contributes to leukemogenicity has not been defined. In contrast, two isoforms of Kit transcripts created through the alternate use of 5’ splice donor sites (alternative splicing) are detected in AML (45), but further analysis showed no apparent association with pathology of AML, indicating naturally occurring changes in splicing mechanisms as stem cells differentiate (46). An alternatively spliced IL-6R mRNA, encoding soluble IL-6R (sIL-6R) expressed in 64% of the primary blast cells of AML patients and all AML cell lines tested, supporting the notion of alternative splicing as a mechanism of sIL-6R production in AML (47).

PTPN6 is a 68 kDa SH2 domain-containing tyrosine phosphatase. PTPN6 regulates hematopoietic cell development, proliferation and receptor-mediated mitogenic signaling pathways (48). A novel PTPN6 mRNA species, derived from aberrant splicing within the N-SH2 domain leading to retention of intron 3 has been discovered in CD34(+)/CD117(+) blasts from AML patients (49). The level of the aberrant intron-retaining splice variant, is lower in CD117(+)–AML bone marrow mononuclear cells at remission than at diagnosis, suggesting the involvement of post-transcriptional PTPN6 processing in leukemogenesis (49).

Survivin is a member of the inhibitors of apoptosis protein family, playing important roles in cell proliferation and survival (50). Survivin is highly expressed in CD34+/ CD38– leukemia progenitor cells and associated poor prognosis and drug resistance in AML patients (51-54). In addition, surviving has been shown to selectively modulate genes the epidermal growth factor receptor signaling pathway in AML LSCs (55). It has been long recognized that alternative splicing of its pre-mRNA generates four different mRNAs: survivin, survivin-2B, survivin-DeltaEx3 and survivin-3B (56,57). In AML cells, survivin is the predominant transcript variant, whereas significantly survivin-2B and survivin-DeltaEx3 express at lower level (58). Expression patterns of survivin variants are associated with clinical outcome. For example, expression of survivin-3B is detected in AML cell lines and may associate with G2/M phase of cell cycle (59). Low expression of survivin-2B correlated with a better overall survival and event-free-survival, whereas high survivin-DeltaEx3 expression was associated with a shorter overall survival (58). However, these splice variants don’t correlate with FAB subtypes, immunophenotype or cytogenetic risk groups (58). These studies support the conclusion that certain survivin splice variants have prognostic values and could be implicated in the leukemogenesis of AML.

Hoxa9 belongs to a family of homeodomain containing transcription factors (60). Hox family regulates genes which control the anterior-posterior body plan and assign tissue fate in human (61). Dysregulation of Hoxa9 has been found in more than 50% of AML patients and highly predicts worse survival (62). Research data indicate splicing play a central role in Hox gene mediated leukemogenesis as a full-length Hoxa9 engineered to prevent natural splicing significantly reduced in vivo leukemogenicity (62).

Oncogene Wilms’ tumor gene 1 (WT1) is a zinc-finger motif containing transcription factor with a proline/glutamine-rich DNA-binding domain. WT1 has been a target for immunotherapy and biomarker commonly used in monitoring of minimal residual disease (MRD) in AML patients. A large assortment of isoforms of WT1 transcripts as many as 36 has been identified. Among them, +5/+KTS are the predominant variants at diagnoses, but their ratio vary between diagnoses (63-66). Increased ratio of the +5/–KTS is associated with aggressive and/or resistant characteristics in FAB subtype M3 and secondary AML (sAML) (63). Together, these data suggest the ratio of certain WT1 isoforms might be crucial for transformation of AML or relapse (66).

The FMS-like tyrosine kinase 3 (FLT3) is a class III RTK family, sharing structural similarity with platelet-derived growth factors (PDGFs), the colony-stimulating factor 1 receptor (CSF1-R) and steel factor receptor (KIT)
FL T3 mutations are identified in about one-third of adult (AML) (69,70). FL T3 mutations induce constitutive activation of phosphoinositide 3-kinase (PI3K)-AKT, RAS-MEK-mitogen-activated protein kinase (MAPK), and signal transducers and activators of transcription (STAT) five pathways, leading uncontrolled cell proliferation, blockage of differentiation and cell survival (52). Thus, FL T3 mutations play a central role in leukemogenesis (71-73). Surprisingly, FL T3, together with NOTCH2 has been identified as the most commonly mis-spliced genes in more than 70% of AML patients (74). The splice variants of NOTCH2 and FL T3 are produced through complete or partial exon skipping and utilization of cryptic splice sites (75). NOTCH-2Va and FL T3-Va transcripts are detected in a significant number of AML patients and high level of NOTCH-2Va predicts worse outcome independent of other known clinical indicators (75).

Taken together, accumulating evidence clearly demonstrate that mis-splicing of certain genes is a common characteristic of AML and some of these mis-spliced mRNAs could translate into proteins with altered function which contribute to leukemogenesis. Notably, there is great interest in identification of splice variants of these genes as disease markers for both diagnosis and stratification and targets for novel therapeutics of leukemia patients.

**Aberrant RNA splicing and drug resistance in AML**

Although initial response to chemotherapy, most of AML patients will develop resistance (76). The appearance of resistance poises a major therapeutic challenge in the treatment of AML and is the primer cause of mortality (77). The phenomenon when an untreated patient does not respond to chemotherapy is termed as primary resistance or intrinsic resistance. Secondary resistance (or acquired resistance) is almost unavoidable when chemotherapy is used for a long period. Alternative RNA splicing also contributes to drug resistance in AML. Cytarabine (Ara-C) is one of the core drugs in the combination chemotherapy against AML (78). Low expression or activity of deoxycytidine kinase (dCK) is responsible for the in vitro cellular resistance to Ara-C in AML cells. Only wild-type mRNA of dCK is amplified from healthy control samples, while splicing variants translating inactive dCK protein resulting from exon skipping are detected in 7 out of 12 purified AML specimen from resistant patients (79). Further work indicates that the alternatively spliced dCK forms render the AML cells to evade to Ara-C attack when there is no wide type dCK (80). On the other side of the coin, exon-array analysis has been performed in isogenic sensitive and (secondary) resistant AML cell lines to Ara-C, doxorubicin (Dox) and hypomethylating agent, azacitidine (Aza) and produced novel insight of alternative exon usages (AEUs) globally (81). Significant alternations were identified in near 1,000 AEU events in a few thousand genes on average between these sensitive and resistant cell lines. GO analysis uncovers five common functional pathways that are shared with resistance to Ara-C, Dox or Aza, including T cell receptor signaling pathway, focal adhesion, axon guidance, regulation of actin cytoskeleton, and ECM-receptor interaction. Taken together, aberrant RNA splicing

<table>
<thead>
<tr>
<th>Gene</th>
<th>Clinical implications of AS forms</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myc</td>
<td>40% of L-myc mRNA lacks exon III and intron I in AML</td>
<td>(44)</td>
</tr>
<tr>
<td>Kit</td>
<td>Two isoforms Kit and KitA are associated with obvious biological and clinical features of AML</td>
<td>(45,46)</td>
</tr>
<tr>
<td>IL-6R</td>
<td>Alternative splicing produces soluble IL-6R, which stimulates growth of AML cells</td>
<td>(47)</td>
</tr>
<tr>
<td>PTPN6</td>
<td>Intron-retaining PTPN6 mRNA is high at diagnosis and low at remission samples of AML</td>
<td>(49)</td>
</tr>
<tr>
<td>Survivin</td>
<td>Low survivin-2B correlates with a better survival, whereas high survivin-DeltaEx3 predicts poor survival</td>
<td>(51,54,58)</td>
</tr>
<tr>
<td>Hoxa9</td>
<td>A full-length Hoxa9 engineered to prevent natural splicing significantly reduces leukemogenicity</td>
<td>(62)</td>
</tr>
<tr>
<td>WT1</td>
<td>+5/+KTS are the predominant variants. Increased ratio of the +5/−KTS is associated with aggressiveness, drug resistant and relapse</td>
<td>(63-66)</td>
</tr>
<tr>
<td>FLT3</td>
<td>FLT3-Va expresses 50% of AML cases. Its expression decreases at remission and increases at relapse</td>
<td>(74,75)</td>
</tr>
<tr>
<td>NOTCH2</td>
<td>High NOTCH2-Va expression is an independent prognostic factor for worse survival</td>
<td>(74,75)</td>
</tr>
</tbody>
</table>

AML, acute myeloid leukemia.
could act as both cause and consequence of drug resistance in AML, suggesting that targeting RNA splicing process might be a novel approach to overcome resistance in anti-AML therapy (82).

### Aberrant RNA splicing and genetic alternations of spliceosome or splicing factors in AML

NGS analysis of MDS in 2011 unexpectedly discovered recurrent somatic mutation of splicing factor 3b subunit 1 (SF3B1) in 20% of cases (27,83). SF3B1 is a component of the U2 snRNP complex and binds pre-mRNA upstream of the intron’s branch site (39). Following this initial finding, a number of studies aiming to categorize genetic alternations in spliceosome and related factors in AML have been performed. Although the frequencies of mutations in genes involving splicing process is relatively low in AML as compared to MDS, these abnormalities influence the overall survival rate and contribute to pathogenesis of AML. One recent study collecting more than 1,500 samples provided an unrivalled understanding of how different driver mutations cooperate and lead to AML (84). In this elegant study, a new molecular subgroup of AML with mutations in genes encoding chromatin, RNA-splicing regulators, or both is defined and occurs in 18% of patients. The common mutated splicing modulators encompass SRSF2, SF3B1, U2AF1 and ZRSR2. SRSF2 is a member of the serine (S) and arginine (R) rich family of pre-mRNA splicing factors, consisting of an RNA recognition motif (RRM) and an RS domain (84). SRSF2 protein regulates diverse RNA-related processes, enhancing the U1 snRNP complex binding to the 5’ splice site of pre-mRNA, U2 snRNP binding at the branch point, and mRNA stabilization. Clinically, AML patients carrying spliceosome-chromatin mutations has poor prognosis. Furthermore, co-occurrence of TP53 mutations in this group of patients makes their prognosis even worse. In practice, these patients should be the candidates for clinical trials of investigational agents or bone marrow transplant (BMT) (84). Recurrent somatic mutations in U2AF1 (also known as U2AF35) have been identified in 3–4% of AML cases. There were 369 splicing alterations significantly associated with U2AF1 mutation (84).

DEAD-box polypeptide 41 (DDX41) (also known as ABS) is an evolutionarily conserved germ cell marker in a wide range of animals. DDX41 belongs to an RNA helicase family, characterized by the conserved motif Asp-Glu-Ala-Asp (DEAD) (85). DDX4 plays an important role in RNA spliceosome assembly, and many other processes including translation initiation, pre-ribosomal RNA processing, small RNA biogenesis and chromosome condensation (85). DDX41 mutations have been identified in about 3% in inherited hematologic malignancies (HM) (86). MDS and AML are the most common malignancies with lower-age onset (87). Defection of DDX41 sizably impair pre-mRNA splicing on the evidence of more avid exon skipping and more exon retention in 61 and 95 genes, respectively, observed in mutant cases than controls (87). Recently, resurgence of AML in donor cells in a patient after allogeneic BMT in a family with a germline DDX41 mutation (88). Taken together, these studies strongly support DDX41 as a tumor suppressor gene and defective DDX41 contributes to AML development.

PRE-mRNA processing factor 8 (PRPF8) shares high identity (61%) in amino acid sequences from yeast to human and plays a central position in the catalytic core of the spliceosome (89). Keightley and colleagues first reported that mutant PRPF8 carrying an early premature STOP codon instigate aberrantly spliced transcripts retaining both U2- and U12-type introns in zebrafish model (90). These authors also observed that myeloid differentiation was impaired within early haematopoiesis (90). It is worthy of pointing out that blocked differentiation of myeloid cells is one of the characteristic feature of AML. Combining these evidences, we could imply that PRPF8 is required for haematopoietic development and defective PRPF8 could be conducive to myeloid malignancies. Indeed, Kurtovic-Kozaric and coworkers identified either recurrent somatic PRPF8 mutations or hemizygous deletions in 3.3% (15 out of 447 cases) and 5.3% (24 out of 450 cases) of myeloid neoplasms, respectively (91). Notably, 50% of PRPF8 mutant and deletion of chromosome 17p cases were found in AML. Survival analysis confirmed that PRPF8 was associated with poor prognosis (91).

Splicing factor proline and glutamine rich (SFPQ) is one of the three member Drosophila behavior/human splicing (DBHS) family in human (92). The DBHS family shares highly conserved tandem N-terminal RRM, a NonA/paraspeckle domain (NOPS) and a C-terminal coiled-coil (93). SFPQ binds wide range of nucleic acids and can be found in the pre-assembled spliceosome complex. However, SFPQ is not an essential member of spliceosome, but rather act as a molecular scaffold and involve co-translational and alternative splicing (93). Mutations in SFPQ and in the nonclassic regulators of mRNA processing CTCF and RAD21 have been discovered in 10% of AML patients in a mutually exclusive manner (94). However, the
clinical importance of this mutation has not been confirmed. In addition to mutations occurring the spliceosome and related factors, the expression levels of some splicing factor SR protein family are also downregulated in AML samples relative to healthy controls (95). Moreover, abnormal caspase-8 pre-mRNA splicing is observed only in AML patients, has significant correlation with several splicing factors (95). These data imply that aberrant expression of splicing factors may potentially rewire apoptosis pathway in AML.

**Distinct deregulation of RNA splicing in sAML and LSC**

sAML or therapy-related AML refers to this disease arises from patients who previously received chemotherapy and/or radiation therapy (96). These patients often have featured genetic changes, including complex karyotype, chromosome 11q abnormalities (MLL gene arrangements), and monosomy of chromosome 5 or 7, p53 mutations (97). Clinically, patients with sAML often are resistant to chemotherapy and have extremely dismal outcome with average survival of 7 months in 40% of patients (97). Thus, a better apprehension of the molecular pathology of sAML and finding novel treatment option are the areas of unmet need.

A recent study by Crews et al. sparks substantial excitement surrounding the pivotal role of spliceosome in sAML and the therapeutic potential of targeting LSCs in this subtype of AML often unresponsive to current therapy (98). In this report, whole-transcriptome sequencing was performed from purified progenitors isolated from sAML, de novo AML and MDS patients. Although mutation in U2 splicing factor SF3B1 was only identified in one sAML sample, increased expression of wild-type SF3B1 was confirmed in a subset of sAML LSCs (98). A splice isoform signature of sAML LSCs was created, which was represented by several alternatively spliced signal transduction and cell adhesion gene products. Pro-survival long isoforms of BCL2 family was elevated in sAML compared to young and aged normal hematopoietic progenitor cells (HPCs), indicating existence of pro-survival splice isoform switching in LSC transformation (98). Over the past 10 years, a list of natural products, including FD-895, pladienolide B, herboxidiene, and spliceostatin A, has been identified as spliceosome modulators (99). They have been shown to have anti-cancer effect in vitro and in vivo models. However, these compounds demonstrate poor metabolic stability and short half-lives in vivo, excluding them from entering clinical evaluation (99,100). 17S-FD-895, an analog of FD-895, was synthesized through the combination of total synthesis and synthetic methods, demonstrating improved stability and on-target effect (101). This new spliceosome targeting compound was evaluated in different sAML models and showed potent efficacy in inhibition of AML LSC and disruption of AML maintenance in vitro and in mouse xenograft models (98).

Importantly, 17S-FD-895 minimal impacted normal BM HSPCs. To validate its on-target effect, comparative RNA-seq analysis revealed that the splice isoform signature of sAML LSCs has been reversed to a normal BM transcriptome profile in treated mice (11). In sum, these evidence underscore that eradication of LSC by targeting spliceosome could represent a novel therapeutic strategy for sAML and further support the development of 17S-FD-895 in clinic.

**Outlook**

With the aid of rapid progress in NGS technology and bioinformatics, novel recurrent mutations and expression level changes in the spliceosome and splicing regulating factors have been uncovered. A flurry of studies characterizing spliceosomal abnormalities in various hematological malignancies and solid tumors have been published since 2011. Also, the promise of the spliceosome as novel anticancer target has drawn substantial funding and focus from both academic and pharma-industrial. We would anticipate a line of such novel therapeutic agents will be tested in clinical trials in near future. However, caution must be highlighted as we still don’t understand well how these aberrant RNA splicing exactly contribute the development of cancers. Importantly, as proper splicing is required for normal hematopoiesis, the potential side effect of manipulating spliceosome has not been addressed in human in a long term.

In addition to the improvement of current compounds in terms of better pharmacokinetics (PK), distribution and pharmacodynamics (PD) in human, substantial efforts have been devoted into additional small molecular inhibitor library screening to discover new class of spliceosome inhibitors with much favorable therapeutic index. Furthermore, another therapeutic approach by gene-silencing molecules with antisense oligonucleotide derivatives or small interfering RNA (siRNA) for targeting aberrant splicing activity has attracted considerable attention. As compared to small molecular inhibitors,
oligonucleotide-based therapeutics offer the promise of precise targeting any mutant splicing factors without impeding any gene sharing sequencing similarity. Although currently no case of such gene-silencing molecules has been approved by government regulators, the difficulty of, with the aid of rapid progress in in vivo delivery field, we would anticipate a line of such oligonucleotide-based novel therapeutic agents will be tested in anti-AML trials in near future.

**Acknowledgements**

Due to space limit, the authors apologize to those whose important works in this field could not be discussed. **Funding:** This research is supported by the National Research Foundation Singapore and the Singapore Ministry of Education under its Research Centres of Excellence initiative and NMRC Clinician-Scientist IRG Grant CNIG11nov38 (J Zhou). WJ Chng is also supported by NMRC Clinician Scientist Investigator award.

**Footnote**

*Conflicts of Interest:* The authors have no conflicts of interest to declare.

**References**


52. Zhou J, Bi C, Janakakumara JV, et al. Enhanced activation of STAT pathways and overexpression of survivin confer resistance to FLT3 inhibitors and could be therapeutic...


