Molecular features of early onset adult myelodysplastic syndrome

Cassandra M. Hirsch,[1,2] Bartlomiej P. Przychodzen,[1] Tomas Radivoyevitch,[3] Bhumika Patel,[4] Swapna Thota,[4] Michael J. Clemente,[1] Yasunobu Nagata,[1] Thomas LaFramboise,[2] Hetty E. Carraway,[4] Aziz Nazha,[4] Mikkael A. Sekeres,[4] Hideki Makishima[1,5] and Jaroslaw P. Maciejewski[1,4]

[1]Department of Translational Hematology and Oncology Research, Taussig Cancer Institute, Cleveland Clinic, OH, USA; [2]Department of Genetics and Genome Science, Case Western Reserve University, Cleveland, OH, USA; [3]Department of Quantitative Health Sciences, Cleveland Clinic Lerner Research Institute, OH, USA; [4]Leukemia Program, Taussig Cancer Institute, Cleveland Clinic, OH, USA and [5]Department of Pathology and Tumor Biology, Graduate School of Medicine, Kyoto University, Japan

Haematologica 2017
Volume 102(6):1028-1034

Correspondence:
maciejj@ccf.org

Received: November 11, 2016.
Accepted: February 28, 2017.
Pre-published: March 2, 2017.

doi:10.3324/haematol.2016.159772

Check the online version for the most updated information on this article, online supplements, and information on authorship & disclosures: www.haematologica.org/content/102/6/1028

ABSTRACT

Myelodysplastic syndromes are typically diseases of older adults. Patients in whom the onset is early may have distinct molecular and clinical features or reflect a demographic continuum. The identification of differences between "early onset" patients and those diagnosed at a traditional age has the potential to advance understanding of the pathogenesis of myelodysplasia and may lead to formation of distinct morphological subcategories. We studied a cohort of 634 patients with various subcategories of myelodysplastic syndrome and secondary acute myeloid leukemia, stratifying them based on age at presentation and clinical parameters. We then characterized molecular abnormalities detected by next-generation deep sequencing of 60 genes that are commonly mutated in myeloid malignancies. The number of mutations increased linearly with age and on average, patients >50 years of age had more mutations. *TET2*, *SRSF2*, and *DNMT3A* were more commonly mutated in patients >50 years old compared to patients ≤50 years old. In general, patients >50 years of age also had more mutations in spliceosomal, epigenetic modifier, and RAS gene families. Although there are age-related differences in molecular features among patients with myelodysplasia, most notably in the incidence of *SRSF2* mutations, our results suggest that patients ≤50 years old belong to a disease continuum with a distinct pattern of early onset ancestral events.

Introduction

With the exception of treatment-related myelodysplastic syndromes (MDS), *de novo* MDS and the closely related secondary acute myeloid leukemia evolving from an antecedent MDS are diseases of the elderly. The median age at presentation is 71 years[1,2] and the yearly incidence rate increases from 2.5/10^5 in the sixth decade to 30/10^5 in the eighth decade of life to as high as 50/10^5 (females) to 100/10^5 (males) in patients >80 years of age.[1] These numbers may be underestimated, as MDS is likely underdiagnosed in the elderly.[3] In children, refractory cytopenia of childhood and juvenile myelomonocytic leukemia are considered distinct entities more related to congenital bone marrow failure and familial leukemia syndromes than to adult MDS.[4] Adult patients less than 50 years of age are sporadically affected by MDS, reflecting either an extreme polarity of age distribution of the disease, or, perhaps, similar to pediatric forms, constituting a separate pathological process. To determine whether MDS in younger adults should be considered a distinct disease subentity, we compared a cohort of what could be considered "early onset" MDS to a cohort of those diagnosed with MDS at a traditional age by contrasting specific clinical characteristics and molecular features.

Methods

Patients' samples

After obtaining informed consent according to protocols approved by the Cleveland Clinic Institution Review Board, marrow and blood samples were collected from patients (2003-2016) classified according to World Health Organization 2008 criteria,[5] and whole exome sequencing (n=95), and/or multi-amplicon deep sequencing (n=539) was performed. Tumor DNA was extracted from the patients' marrow and wherever possible germline DNA was obtained from selected $CD3^+$ T cells. Cases of clonal hematopoiesis of indeterminate potential (CHIP) were excluded from this study; all patients had to have dysplasia, increased blasts, or abnormal cytogenetics in order to be diagnosed with MDS.[6]

Next-generation sequencing

Whole exome capture was performed according to the manufacturer's protocol [SureSelect®, ver. 4 (Agilent Technology)].[7,8] The captured targets were subjected to massive parallel sequencing using the HiSeq 2000. Multi-amplicon-based, targeted, deep sequencing was performed for a panel of 64 genes, most commonly somatically mutated in MDS (*Online Supplementary Table S1*)[8,9] and those known to be affected by germline mutations (*Online Supplementary Table S2*).[10] Customized probe sets amplified exons of target genes. Sequencing libraries were generated according to an Illumina paired-end library protocol and subjected to deep sequencing on MiSeq (Illumina) sequencers according to the standard protocol. A bio-analytic pipeline, devised in-house, as previously described,[11] was applied to identify somatic mutations and (where appropriate) germline variants by comparison with germline controls, mutational databases (Entrez Gene[12] the Ensembl Genome Browser,[13] COSMIC[14]), and sequenced controls, Exome Aggregation Consortium (ExAC).[15] Variant allelic frequencies of mutations were adjusted according to the zygosity and copy number based on single nucleotide polymorphism results. Serial samples and variant allelic frequencies were compared to determine clonal hierarchy, i.e. dominant and secondary mutations.

Statistical analysis

Wilcoxon tests were performed for pairwise comparisons between continuous variables and the Fisher exact test was applied for categorical variables. Poisson regression was used to find the linear slope of mutations rates *versus* age. All *P* values were two-sided and values less than 0.05 were considered to be statistically significant. Analyses were performed using the R statistical program.

Results

Demographic features of adults with "early onset myelodysplastic syndrome"

We analyzed 634 patients with primary MDS (excluding treatment-related or secondary MDS) with a median presentation age of 68 years (range 20-94) (*Online Supplementary Table S3*). When patients were age-ranked in 5-year increments (Figure 1A), a unimodal age distribution was obtained, allowing the cohort to be empirically split into two groups: ≤50 years (n=65, with "early onset adult MDS"), and >50 years (n=569, with "traditional age of diagnosis MDS"), corresponding to 10% and 90% of patients, respectively. The split at the age of 50 was further justified by the separate clusters that were present when survival by age group was analyzed (*Online Supplementary Figure S1A*). Additionally, 15% of patients were over 79 years old and analyzed in contrast to the other two groups as "late onset MDS." Accordingly, the median age of the early onset MDS group was 44 years (range, 20-50), that of the group between 50 and 80 years was 70 years (range, 51-79), and the median age of the late onset group was 83 years (range, 80-94). Women were over-represented among early onset patients, whereas the converse was true among MDS patients diagnosed at a traditional age (60% *versus* 32%, $P<0.0001$) (Figure 1A). There was no significant difference between the numbers of females with 5q deletion in these groups. The Surveillance, Epidemiology and End Results (SEER) age-at-diagnosis distributions were comparable to those of our cohort[18] (Figure 1B).

Clinical features of "early onset myelodysplastic syndrome"

Higher-risk MDS (refractory anemia with excess blasts-1/2) was more common among early onset MDS patients (35% *versus* 24%, $P=0.048$) while lower-risk MDS (refractory cytopenia with multilineage dysplasia, refractory anemia with ringed sideroblasts, refractory cytopenia with unilineage dysplasia, refractory anemia) predominated in MDS patients >50 years of age (28% *versus* 41%, $P=0.042$). Additionally, 17% of patients ≤50 years old had MDS/myeloproliferative neoplasms (including chronic myelomonocytic leukemia), and 20% had secondary acute myeloid leukemia compared to 21% and 14%, respectively, in patients >50 years of age (*Online Supplementary Table S4*). When grouped according to cytogenetic risk categories, survival differences were seen between patients ≤50 and >50 years old (*Online Supplementary Figure 1B*). There was no difference between patients ≤50 and >50 years old with regards to a family history of cancer, including leukemia, or bone marrow failure (51% *versus* 54%, $P=0.51$; *Online Supplementary Table S4*). Family history of cancer was defined as a self-reported presence of one or more first-degree relative (parent, sibling, or child) with cancer, hematologic neoplasia, or bone marrow failure. When family history was restricted to only hematologic neoplasia or bone marrow failure, still no difference was seen between patients ≤50 and >50 years old.

Patterns of molecular lesions

When patients were screened for somatic mutations, more mutations were detected in patients >50 years old by whole exome and targeted sequencing overall ($P=0.05$ and $P=0.02$, respectively) (Figure 1C), as well as within molecular subtypes (Figure 1D). The average number of mutations per case increased in a linear fashion with age ($R^2=0.947$, Figure 1E). Considering the subset of patients analyzed by whole exome sequencing, no difference in the pattern of transversions or transitions was identified in the mutational signatures of patients ≤50 and >50 years old (*Online Supplementary Figure S2A*).

Focusing on specific lesions, *ASXL1* (9%), *TET2* (9%), *TP53* (9%), and *RUNX1* (8%) were the most frequently mutated genes in patients with early onset MDS (Figure 2A). However, *SRSF2* and *TET2* mutations were less common in patients ≤50 years old (2% *versus* 10%, $P=0.013$ and 9% *versus* 19%, $P=0.060$; respectively (Figure 2A). When categorized based on functional properties of

affected genes, patients >50 years old were found to have more spliceosomal gene mutations (*P*=0.025), epigenetic modifier mutations (*P*=0.007), and genes in the RAS family (*P*=0.08) (Figure 2A) when compared to patients ≤50 years old. Normal karyotype was present in about one-half of all patients but no major differences were found in distribution of individual lesions, including complex karyotype (14% *versus* 8%, *P*=0.13) (Figure 2B). In the linear fits of the average number of mutations *versus* age using Poisson regression, *TET2* and *SRSF2* mutation rates increased with age (*P*=0.001 and *P*=0.035, respectively (Figure 2C). This trend was recapitulated by applying age-adjusted frequencies of *TET2* mutants to SEER MDS demographics (Figure 2D), and parallels the trend seen in healthy controls (Figure 2E). Previously, it was suggested that hydroxymethylation may prevent C->T transitions by decreasing MeCpG levels. Thus, more C->T transitions via MeCpG deamination would be expected in cases with founder *TET2* mutations.[17,21] However, when molecular signatures of cases with and without *TET2* mutants were analyzed, age-related C->T transitions were similar (*Online Supplementary Figure S2B*). When compared to all other patients, *TET2* mutant cases tended to display a higher number of additional mutations (*P*<0.001; targeted, *P*=0.074 whole exome) (*Online Supplementary Figure S2C*). However, among patients without a *TET2* mutation, there was no difference in the number of mutations between patients ≤50 and >50 years old (*Online Supplementary Figure S2D*). Analysis of clonal architecture revealed that *RUNX1*, *SF3B1*, and *TP53* were the most common dominant mutations in patients with early onset MDS. In contrast, *TET2*, *SF3B1* and *STAG2* were the most common

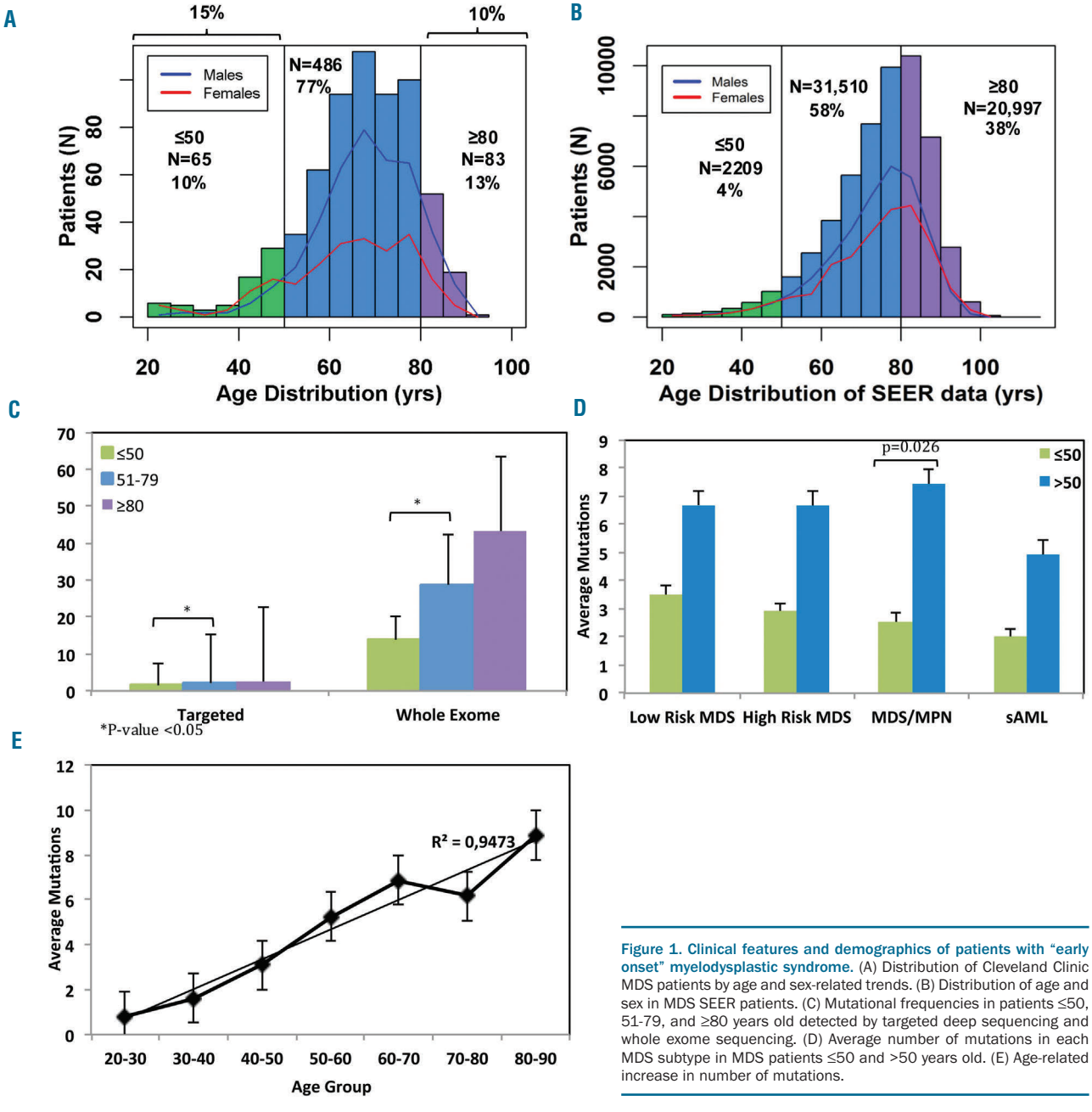

Figure 1. Clinical features and demographics of patients with "early onset" myelodysplastic syndrome. (A) Distribution of Cleveland Clinic MDS patients by age and sex-related trends. (B) Distribution of age and sex in MDS SEER patients. (C) Mutational frequencies in patients ≤50, 51-79, and ≥80 years old detected by targeted deep sequencing and whole exome sequencing. (D) Average number of mutations in each MDS subtype in MDS patients ≤50 and >50 years old. (E) Age-related increase in number of mutations.

dominant mutations in patients >50 years of age. Overall, there was no difference in the distribution of variant allele frequencies between the groups of patients (Figure 3A). Mutations were similarly distributed across functional gene families in MDS patients ≤50 and >50 years old (Figure 3B, C).

Known familial mutations

In an attempt to explain the early occurrence of MDS, the cohort of patients with early onset disease was analyzed for the presence of congenital mutations known to be associated with MDS, with a focus on mutations in genes frequently associated with a familial leukemia or

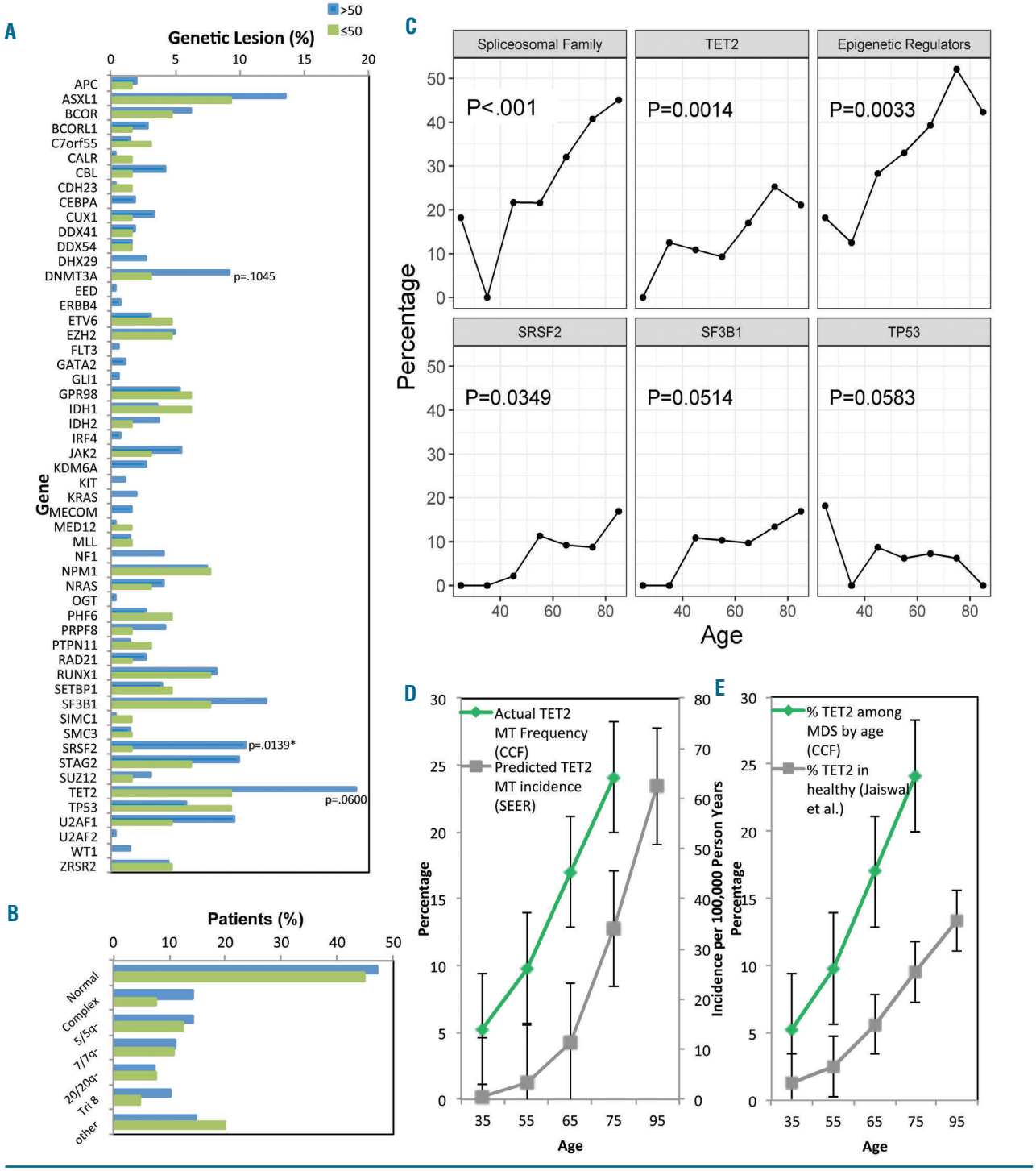

Figure 2. Typical somatic defects in myelodysplastic syndrome patients ≤50 and >50 years old. (A) Frequencies of somatic mutations in the most recurrently mutated genes in patients ≤50 and >50 years of age. (B) Frequencies of cytogenetic abnormalities in patients ≤ 50 and >50 years of age. (C) Percentages of somatic mutations according to age in selected genes. *P*-values correspond to the linear slope of mutation rates *vs.* age found using Poisson regression. Only the genes found to be significant are shown here. Epigenetic modifier genes include *DNMT3A*, *EZH2*, *KDM6A*, *IDH1/2*, and *TET2*. Spliceosomal gene mutations include *SRSF2*, *SF3B1*, *LUC7L2*, *U2AF1*, *ZRSRS*, *PRPF8*, and *DDX41*. (D) *TET2* mutational frequency per age group in MDS as predicted by SEER data and actual *TET2* mutation frequency in the Cleveland Clinic cohort. (E) Frequency of *TET2* mutations in healthy individuals and MDS patients by age.

marrow failure as previously described[10] (the panel of genes tested is shown in *Online Supplementary Table S2*). Familial mutations were found in 12% of patients with early onset MDS, but a higher incidence of telomerase complex or Fanconi anemia gene variants was not found (*versus* 6% P=0.04) (*Online Supplementary Table S2*).

Discussion

Investigations of disease demographics may reveal clues to pathogenic mechanisms, provide insight to correct diagnosis, and help identify disease variants or even new nosological entities. For instance, aplastic anemia shows a

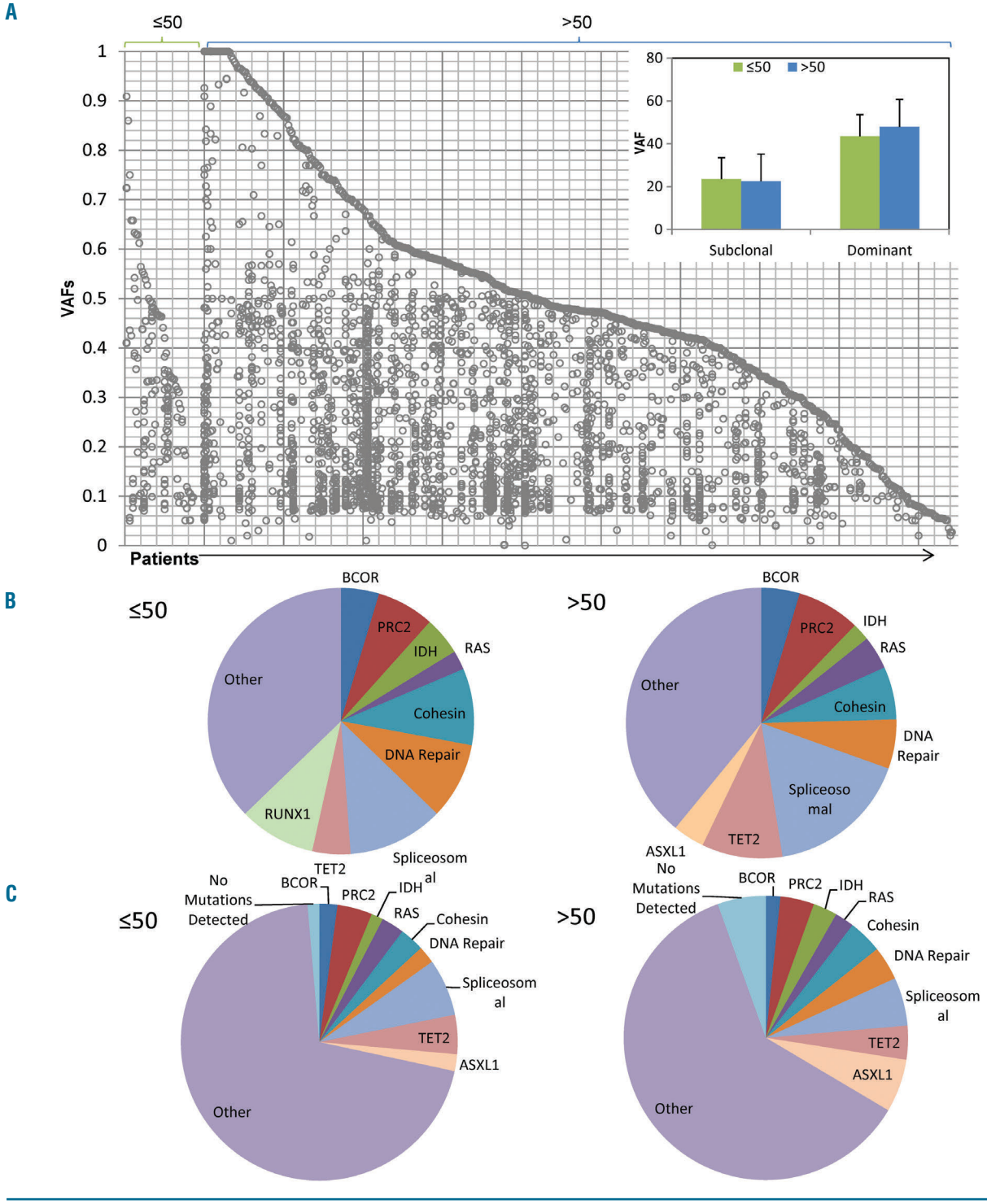

Figure 3. Molecular characteristics of patients with myelodysplastic syndrome. (A) Comparison of variant allele frequencies across patients ≤50 *vs.* >50 years old. (B) Distribution of dominant mutations in patients ≤50 *vs.* >50 years old. (C) Distribution of all mutations in patients ≤50 *vs.* >50 years old.

bimodal age distribution, with its first peak in children and young adults representing typical idiopathic disease, and the peak later in adults corresponding to a possible admix of patients with MDS, in particular hypocellular MDS.[4] In contrast, *de novo* MDS tends to be a disease of older adults. Our cohort demonstrated a unimodal age distribution similar to that of SEER data.[18] Patients with early onset MDS identified in this cohort could constitute an extreme outlier group, roughly 10% of the age continuum, or they could represent a separate disease sub-entity that pathogenically belongs to a different form of MDS such as childhood MDS or possibly juvenile myelomonocytic leukemia. With these hypotheses in mind, we analyzed the molecular profiles of MDS patients ≤50 and >50 years old, incorporating chromosomal abnormalities and somatic mutations identified through either exome or targeted sequencing approaches with a rationally selected panel of the most common mutations associations with myeloid neoplasms.[9] With age, the number of somatic events increased (as detected by both exome and targeted sequencing) while the percentage of patients with normal cytogenetics remained constant.[17,19,20] This observation suggests age-related accumulation of mutational events and a higher molecular complexity the later a patient is diagnosed with MDS. Analysis of mutations identified by exome sequencing in patients ≤50 and >50 years old did not reveal a higher rate of C to T transitions in older patients as previously described.[21] Furthermore, the clonal burden did not differ between patients ≤50 and >50 years old, suggesting that MDS is fully clonal at presentation.

Comparison of chromosomal and mutational patterns revealed several discrete differences suggesting that atypical patients with early onset MDS likely constitute an extreme group of a continuum rather than a separate entity. *TET2* and *SRSF2* were found to be more commonly mutated in the group of MDS patients >50 years old. Supporting this finding, control cohorts document an increase in the incidence of asymptomatic *TET2* mutations with age, raising the possibility that such mutations represent pre-leukemic founder lesions with a long latency period before disease manifestation since the presence of sub-clonal events was associated with a subsequent risk of malignancies.[6] Indeed, *TET2*, *RUNX1*, and *TP53* mutations were as frequent as in the Cancer Genomic Atlas data for a younger cohort of patients with acute myeloid leukemia.[22] In a recent study of elderly patients with acute myeloid leukemia, genes in the spliceosomal complex as well as *TET2* were found to be altered more frequently than in younger cases, similar to what is seen in MDS.[23] While there are marked differences in the mutational spectrum between early onset MDS and MDS patients diagnosed at a traditional age, overall our results suggest that patients with MDS ≤50 years old constitute part of a continuum rather than a specific group with a distinct molecular pathogenesis. Increased frequency of *TET2* mutations parallels the trend seen in healthy controls[17,20] with increasing frequency of *TET2* mutations with aging suggesting a disease-initiating role of these mutations in MDS that is consistent with increased MDS risk in asymptomatic mutant carriers.

Previously, somatic mutations of *SRSF2* and other spliceosomal genes were found exclusively in patients >70 years old and, therefore, associated with age-related clonal hematopoiesis.[17,19,20] In pediatric disease, including refractory cytopenia of childhood and juvenile myelomonocytic leukemia, we and others have also found a low rate of *TET2* and spliceosomal mutations.[4] Similarly, the incidence of observed *TET2* mutations in this population followed the trend observed among healthy controls as previously reported.[17,20] However, in contrast to juvenile myelomonocytic leukemia, *RAS* gene family mutations were not predominant in patients with early onset adult MDS.[24] We expect that early manifestation of MDS will be associated with familial disease, disease with a strong family history.[10,25,26] Furthermore, *RUNX1* was the most common dominant mutation in patients ≤50 years old, and two out of five cases were confirmed to be germline. As expected, the most common mutation in patients >50 years of age was *TET2*, well documented as a mutation associated with aging.[17,19,20] Overall, our study suggests that the molecular underpinnings of early onset adult MDS, while distinct from juvenile forms of the disease, do not differ enough from MDS diagnosed at a traditional age to warrant a separate categorization.

Acknowledgments

This work was supported by grants NIH/NHLBI, R01 HL118281, NIH/NHLBI, R01 HL123904, NIH/NHLBI, R01 HL128425, Edward P. Evans Foundation, 2015 Basic Research Grant, and American Cancer Society Research Scholar Grant 123436-RSG-12-159-01-DMC.

References

1. Ma X, Does M, Raza A, Mayne ST. Myelodysplastic syndromes: incidence and survival in the United States. Cancer. 2007;109(8):1536-1542.
2. Sekeres MA, Schoonen WM, Kantarjian H, et al. Characteristics of US patients with myelodysplastic syndromes: results of six cross-sectional physician surveys. J Natl Cancer Inst. 2008;100(21):1542-1551.
3. Burgstaller S, Wiesinger P, Stauder R. Myelodysplastic syndromes in the elderly: treatment options and personalized management. Drugs Aging. 2015;32(11): 891-905.
4. Niemeyer CM, Baumann I. Classification of childhood aplastic anemia and myelodysplastic syndrome. Hematol Am Soc Hematol Educ Program. 2011;2011:84-89.
5. Swerdlow SH. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. International Agency for Research on Cancer, 2008.
6. Steensma DP, Bejar R, Jaiswal S, et al. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. Blood. 2015; 126(1):9-16.
7. Thota S, Viny AD, Makishima H, et al. Genetic alterations of the cohesin complex genes in myeloid malignancies. Blood. 2014;124(11):1790-1798.
8. Yoshida K, Sanada M, Shiraishi Y, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. Nature. 2011;478(7367):64-69.
9. Haferlach T, Nagata Y, Grossmann V, et al. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. Leukemia. 2014;28(2):241-247.
10. Churpek JE, Pyrtel K, Kanchi KL, et al. Genomic analysis of germ line and somatic variants in familial myelodysplasia/acute myeloid leukemia. Blood. 2015;126(22): 2484-2490.
11. Makishima H, Yoshida K, Nguyen N, et al. Somatic SETBP1 mutations in myeloid malignancies. Nat Genet. 2013;45(8):942-946.
12. Maglott D, Ostell J, Pruitt KD, Tatusova T. Entrez Gene: gene-centered information at NCBI. Nucleic Acids Res. 2005;33 (Database issue):D54-58.
13. Yates A, Akanni W, Amode MR, et al. Ensembl 2016. Nucleic Acids Res. 2016;44 (D1):D710-716.

14. Forbes SA, Beare D, Gunasekaran P, et al. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic Acids Res. 2015;43(Database issue):D805-811.
15. Lek M, Karczewski K, Minikel E, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016;536(7616):285-291.
16. Makishima H, Visconte V, Sakaguchi H, et al. Mutations in the spliceosome machinery, a novel and ubiquitous pathway in leukemogenesis. Blood. 2012;119(14):3203-3210.
17. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. N Engl J Med. 2014;371(26):2488-2498.
18. Surveillance, Epidemiology, and End Results (SEER) Program Research Data. In: National Cancer Institute D, Surveillance Resarch Program, Surveillance Systems Branch ed, 1973-2012.
19. Mason CC, Khorashad JS, Tantravahi SK, et al. Age-related mutations and chronic myelomonocytic leukemia. Leukemia. 2016;30(4):906-913.
20. McKerrell T, Park N, Moreno T, et al. Leukemia-associated somatic mutations drive distinct patterns of age-related clonal hemopoiesis. Cell Rep. 2015;10(8):1239-1245.
21. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. Nature. 2013;500(7463):415-421.
22. Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med. 2013;368(22):2059-2074.
23. Silva P, Neumann M, Vosberg S, et al. Acute myeloid leukemia in the elderly is characterized by a distinct genetic landscape. Blood. 2015;126(23):804.
24. Chang TY, Dvorak CC, Loh ML. Bedside to bench in juvenile myelomonocytic leukemia: insights into leukemogenesis from a rare pediatric leukemia. Blood. 2014;124(16):2487-2497.
25. Hamadou WS, Bourdon V, Gaildrat P, et al. Mutational analysis of JAK2, CBL, RUNX1, and NPM1 genes in familial aggregation of hematological malignancies. Ann Hematol. 2016;95(7):1043-1050.
26. Astuto LM, Bork JM, Weston MD, et al. CDH23 mutation and phenotype heterogeneity: a profile of 107 diverse families with Usher syndrome and nonsyndromic deafness. Am J Hum Genet. 2002;71(2):262-275.