

## ORIGINAL ARTICLE

# Somatic Mutations and Clonal Hematopoiesis in Aplastic Anemia

T. Yoshizato, B. Dumitriu, K. Hosokawa, H. Makishima, K. Yoshida, D. Townsley, A. Sato-Otsubo, Y. Sato, D. Liu, H. Suzuki, C.O. Wu, Y. Shiraishi, M.J. Clemente, K. Kataoka, Y. Shiozawa, Y. Okuno, K. Chiba, H. Tanaka, Y. Nagata, T. Katagiri, A. Kon, M. Sanada, P. Scheinberg, S. Miyano, J.P. Maciejewski, S. Nakao, N.S. Young, and S. Ogawa

## ABSTRACT

**BACKGROUND**

In patients with acquired aplastic anemia, destruction of hematopoietic cells by the immune system leads to pancytopenia. Patients have a response to immunosuppressive therapy, but myelodysplastic syndromes and acute myeloid leukemia develop in about 15% of the patients, usually many months to years after the diagnosis of aplastic anemia.

**METHODS**

We performed next-generation sequencing and array-based karyotyping using 668 blood samples obtained from 439 patients with aplastic anemia. We analyzed serial samples obtained from 82 patients.

**RESULTS**

Somatic mutations in myeloid cancer candidate genes were present in one third of the patients, in a limited number of genes and at low initial variant allele frequency. Clonal hematopoiesis was detected in 47% of the patients, most frequently as acquired mutations. The prevalence of the mutations increased with age, and mutations had an age-related signature. *DNMT3A*-mutated and *ASXL1*-mutated clones tended to increase in size over time; the size of *BCOR*- and *BCORL1*-mutated and *PIGA*-mutated clones decreased or remained stable. Mutations in *PIGA* and *BCOR* and *BCORL1* correlated with a better response to immunosuppressive therapy and longer and a higher rate of overall and progression-free survival; mutations in a subgroup of genes that included *DNMT3A* and *ASXL1* were associated with worse outcomes. However, clonal dynamics were highly variable and might not necessarily have predicted the response to therapy and long-term survival among individual patients.

**CONCLUSIONS**

Clonal hematopoiesis was prevalent in aplastic anemia. Some mutations were related to clinical outcomes. A highly biased set of mutations is evidence of Darwinian selection in the failed bone marrow environment. The pattern of somatic clones in individual patients over time was variable and frequently unpredictable. (Funded by Grant-in-Aid for Scientific Research and others.)

The authors' full names, academic degrees, and affiliations are listed in the Appendix. Address reprint requests to Dr. Ogawa at the Dept. of Pathology and Tumor Biology, Graduate School of Medicine, Kyoto University, Yoshida-Konoe-cho, Kyoto-shi Sakyo-ku, Kyoto 606-8501, Japan, or at [sogawa-tky@umin.ac.jp](mailto:sogawa-tky@umin.ac.jp); or to Dr. Young at the Hematology Branch, National Heart, Lung, and Blood Institute, 10 Center Dr., Bldg. 10/CRC, Rm. 3E-5140, Bethesda, MD 20892-1202, or at [youngns@mail.nih.gov](mailto:youngns@mail.nih.gov).

Drs. Yoshizato and Dumitriu and Drs. Young and Ogawa contributed equally to this article.

N Engl J Med 2015;373:35-47.

DOI: 10.1056/NEJMoa1414799

Copyright © 2015 Massachusetts Medical Society.

**A**QUIRED APLASTIC ANEMIA IS CAUSED by immune-mediated destruction of hematopoietic stem and progenitor cells.<sup>1</sup> CD34+ cells and early progenitors are uniformly reduced in aplastic anemia.<sup>2</sup> Bone marrow transplantation is curative, and patients may also have a response to immunosuppressive therapy.<sup>3,4</sup> With improved survival, the late development of myelodysplastic syndromes, acute myeloid leukemia (AML), or both has been noted in about 15% of patients and termed “clonal evolution.”<sup>5</sup>

Although “clonal evolution” historically has been used to describe the development of cancer in patients with an immune disease, this term is a misnomer, since there is evidence of clonal hematopoiesis associated with aplastic anemia in patients in whom myelodysplastic syndromes or AML do not develop. Small numbers of stem-cell clones were inferred from X-chromosome skewing.<sup>6</sup> In about half the patients with aplastic anemia, there is evidence of a second disease, paroxysmal nocturnal hemoglobinuria (PNH),<sup>1</sup> which results from expansion of somatically mutated hematopoietic stem cells. Cytogenetic abnormalities,<sup>7,8</sup> as well as uniparental disomy of the 6p arm (6pUPD),<sup>9</sup> have been described in aplastic anemia without clinical evidence of myelodysplastic syndromes. However, the origin, importance, and dynamics over time of clonal hematopoiesis in aplastic anemia, and its relationship to the development of myelodysplastic syndromes, AML, or both, have not been defined. Clonality before the development of cancer and thus early events in at least one form of leukemogenesis can be examined in patients with aplastic anemia.

We used samples obtained from three institutions that specialize in the treatment of patients with bone marrow failure to perform targeted deep sequencing of genes implicated in myelodysplastic syndromes, AML, or both, with correlation of clonal populations of mutated cells to clinical outcomes. In selected cases, we then used whole-exome sequencing to characterize hematopoietic clonal architecture in patients over time.

## METHODS

### STUDY DESIGN

We enrolled 439 patients with aplastic anemia from three centers that specialize in the treatment of patients with bone marrow failure: 256

patients from the National Institutes of Health (NIH), 24 patients from the Cleveland Clinic, and 159 patients from Kanazawa University (Fig. S1 and Tables S1 and S2 in the Supplementary Appendix, available with the full text of this article at NEJM.org). Blood, bone marrow, and buccal samples were obtained from the patients after written informed consent was received according to protocols approved by the institutional review board at each of the three institutions. Germline DNA was available from CD3+ T cells obtained from patients from the NIH and Cleveland Clinic and from buccal smears obtained from 22 patients from Kanazawa University. Serially obtained specimens were available from 82 patients. In total, 668 blood samples were analyzed. The criteria for diagnosis, disease severity, response to immunosuppressive therapy, and relapse have been described previously.<sup>3</sup> For the NIH cohort, per protocol, clonal evolution to myelodysplastic syndromes or AML was defined as identification of a chromosomal abnormality by means of metaphase karyotyping, the presence of frank morphologic dysplasia, or an increase in the level of CD34+ cells detected on examination of the bone marrow.

The last two authors designed the study, and five of the authors gathered the data. The last two authors vouch for the integrity and completeness of the data and analyses and the fidelity of the study to the protocol. Sequencing and genotype data from the study have been deposited in the European Genome-Phenome Archive under accession number EGAS00001001153.

### DNA SEQUENCING AND SINGLE-NUCLEOTIDE POLYMORPHISM ARRAY

Methods of whole-exome and targeted sequencing and single-nucleotide polymorphism (SNP)-array karyotyping were reported previously<sup>10-12</sup> (see the Supplementary Appendix). The panel of 106 genes for targeted sequencing included most of the genes known to be mutated in myeloid cancers, as well as *PIGA* (Table S3 in the Supplementary Appendix). The mean depth for targeted sequencing was 1248×, and the mean depth for whole-exome sequencing was 112× (Table S4 in the Supplementary Appendix). Somatic mutations were detected with the use of a frequency threshold of 0.07 (whole-exome sequencing) or 0.02 (targeted sequencing) for variant allele frequency and were individually validated with the

use of deep sequencing of polymerase-chain-reaction–amplified targets (with assessment of specific variant reads in healthy controls whenever possible).

#### STATISTICAL ANALYSIS

Random survival forests<sup>13</sup> and the penalized likelihood approach to variable selection<sup>14</sup> were used to identify subsets of mutated genes that correlated with clinical outcomes. Multiple logistic-regression and Cox proportional-hazards models were used to evaluate mutations in gene sets and competing risk factors (see the Supplementary Appendix). Full details of the statistical analysis are provided in the Methods section in the Supplementary Appendix.

## RESULTS

### TARGETED SEQUENCING AND SNP ARRAY KARYOTYPING

Targeted sequencing of a panel of genes that are recurrently mutated in myeloid cancers was performed with the use of blood obtained from all 439 patients who had aplastic anemia. Overall, 249 somatic mutations were detected among 156 patients (36%); 56 of these 156 patients (36%) had multiple mutations (range, 1 to 7 mutations) (Fig. S2A in the Supplementary Appendix). The most frequently mutated genes were *BCOR* and *BCORL1* (in 9.3% of the patients), *PIGA* (in 7.5%), *DNMT3A* (in 8.4%), and *ASXL1* (in 6.2%), which together accounted for 77% of all mutation-positive patients. The overall mutation rate and the predominant mutations were similar in the U.S. and Japanese cohorts (Fig. 1, and Fig. S2A and Table S5 in the Supplementary Appendix).

Except in the case of the *PIGA* and *BCOR* and *BCORL1* mutations, both the presence and the number of mutations per patient positively correlated with age ( $P < 0.001$  for both comparisons) (Fig. 1B). In 6 patients, multilineage involvement of mutations was confirmed by means of amplicon sequencing of genomic DNA from flow-sorted fractions of hematopoietic stem cells, common myeloid progenitors, and myeloid erythroid progenitors, although, as expected, the numbers of cells in these fractions were reduced (Fig. S3 and Table S6 in the Supplementary Appendix).

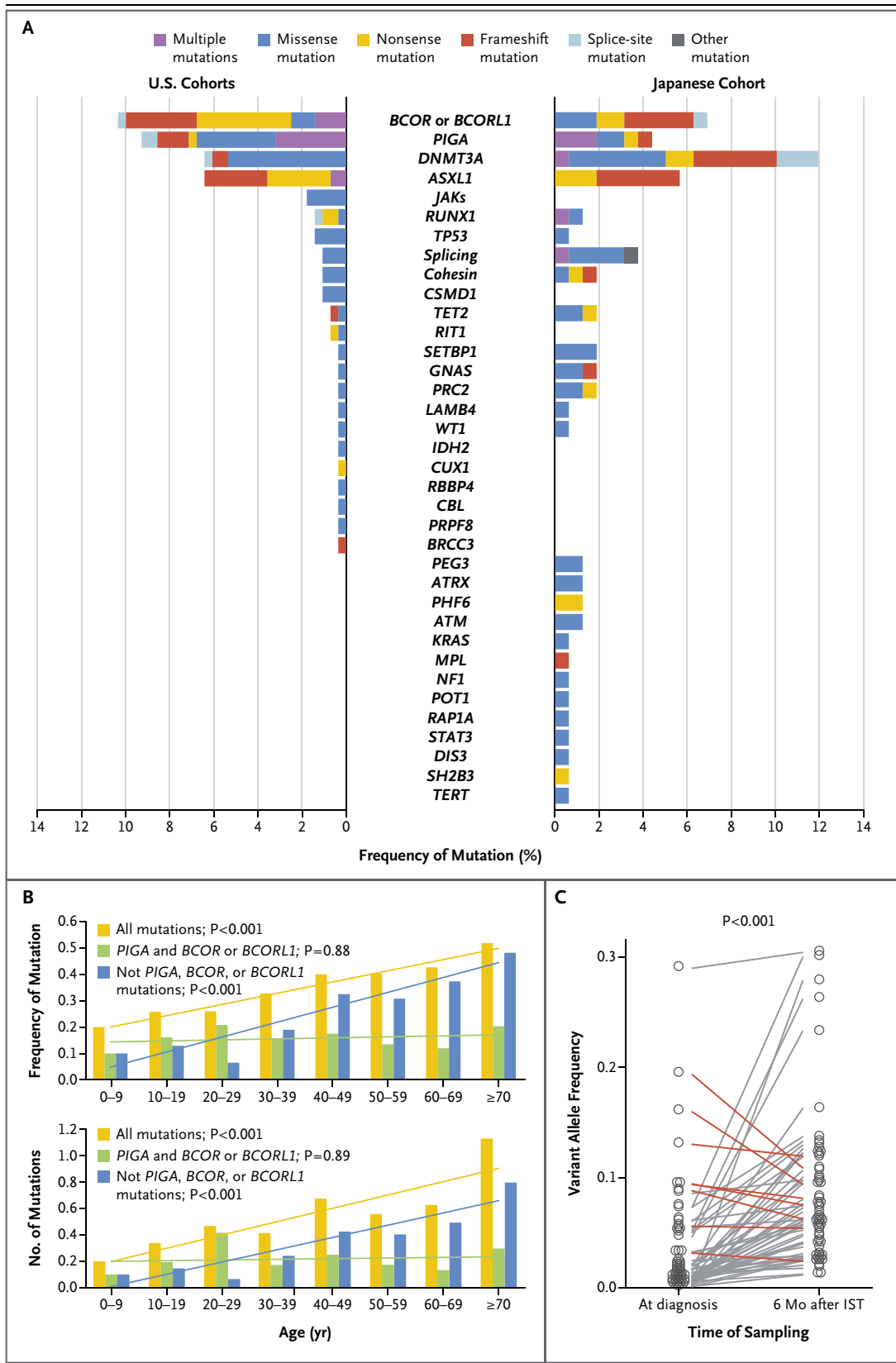
Samples in the NIH cohort were obtained 6 months after the initiation of immunosuppressive therapy. In that cohort, diagnostic samples

were available for 45 patients; 79 mutations were found in these patients. Amplicon sequencing confirmed that the majority of these mutations (58 of 79) were present at diagnosis at a variant allele frequency that was much lower than that at 6 months after treatment ( $P < 0.001$ ) (Fig. 1C). There was no significant difference in the relative number of mutations in each gene at diagnosis and at 6 months after treatment (Fig. S2B in the Supplementary Appendix); this indicates that these mutations were present before treatment.

As has been reported previously,<sup>9</sup> the most prevalent abnormality in SNP array karyotyping was 6pUPD, which was present in 13% of the patients in each cohort (Fig. S4 in the Supplementary Appendix). Copy-number variations were detected in 18 of 417 samples analyzed and included  $-7$  (in 7 patients),  $\text{del}(13q)$  (in 2 patients), and others (in 10 patients) (Fig. S4 in the Supplementary Appendix). With the use of both platforms, clonal hematopoiesis was identified in 47% of patients.

Candidate-gene mutations in aplastic anemia recapitulated those in myelodysplastic syndromes, AML, or both with respect to positional distribution and a strong bias toward nonsense, frameshift, and splice-site changes (Fig. S5A in the Supplementary Appendix).<sup>15-19</sup> However, the mean allelic burden of mutations in aplastic anemia was substantially lower than that in myelodysplastic syndromes (9.3% vs. 30.4%) (Fig. S5B and S5C in the Supplementary Appendix).<sup>20</sup> Although *DNMT3A* and *ASXL1* mutations were frequent both in aplastic anemia and in myelodysplastic syndromes and AML, we observed overrepresentation of *PIGA* and *BCOR* and *BCORL1* mutations and underrepresentation of mutations in *TET2*, splicing-factor genes, *JAK2*, *RUNX1*, and *TP53* in aplastic anemia, as compared with myelodysplastic syndromes; this suggested discrete mechanisms of clonal selections in aplastic anemia.

Some patients carried multiple mutations in the same genes, including *PIGA*, *BCOR*, *DNMT3A*, *ASXL1*, *RUNX1*, and *ZRSR2*. In 11 male patients, there were 22 mutations in X-chromosome genes; therefore, they were probably within independent subclones (Fig. S5D in the Supplementary Appendix). This was explicitly observed with the use of single-colony sequencing in 2 patients carrying 2 *BCOR* mutations (Patient NIH076) and 3 *PIGA* mutations (Patient NIH207) (Table S7 in the Supplementary Appendix).



**Figure 1 (facing page). Somatic Mutations Identified by Means of Targeted Sequencing.**

Panel A shows the frequency of mutated genes and type of mutations in each gene identified in the U.S. cohorts (from the National Institutes of Health [NIH] and the Cleveland Clinic) and the Japanese cohort (from Kanazawa University). Panel B shows the frequency (top) and mean number (bottom) of mutations in various age groups for indicated sets of mutations with corresponding regression lines. Panel C shows the variant allele frequency of mutations detected at the time of diagnosis and 6 months after immunosuppressive therapy (IST) in the NIH cohort. Red lines indicate mutations with a higher allele frequency (or the same allele frequency) at diagnosis than at 6 months after IST; gray lines indicate the rest of the mutations.

**CLINICAL CORRELATIONS**

The effect of gene mutations on clinical outcomes was assessed first in the NIH cohort, in which samples were uniformly obtained 6 months after treatment and all patients had severe aplastic anemia. There was no apparent relationship between the presence of mutations and the response to immunosuppressive therapy ( $P=0.28$ ) (Fig. S6A in the Supplementary Appendix), but when the mutated genes were assessed individually, *BCOR* and *BCORL1* mutations favored a good response to immunosuppressive therapy ( $P=0.013$ ) (Fig. S6B in the Supplementary Appendix). Logistic regression with the use of multiple variables that were previously reported to correlate with a response to treatment at 6 months<sup>21</sup> showed that *BCOR* and *BCORL1* remained independently associated with a response (Table S8 in the Supplementary Appendix).

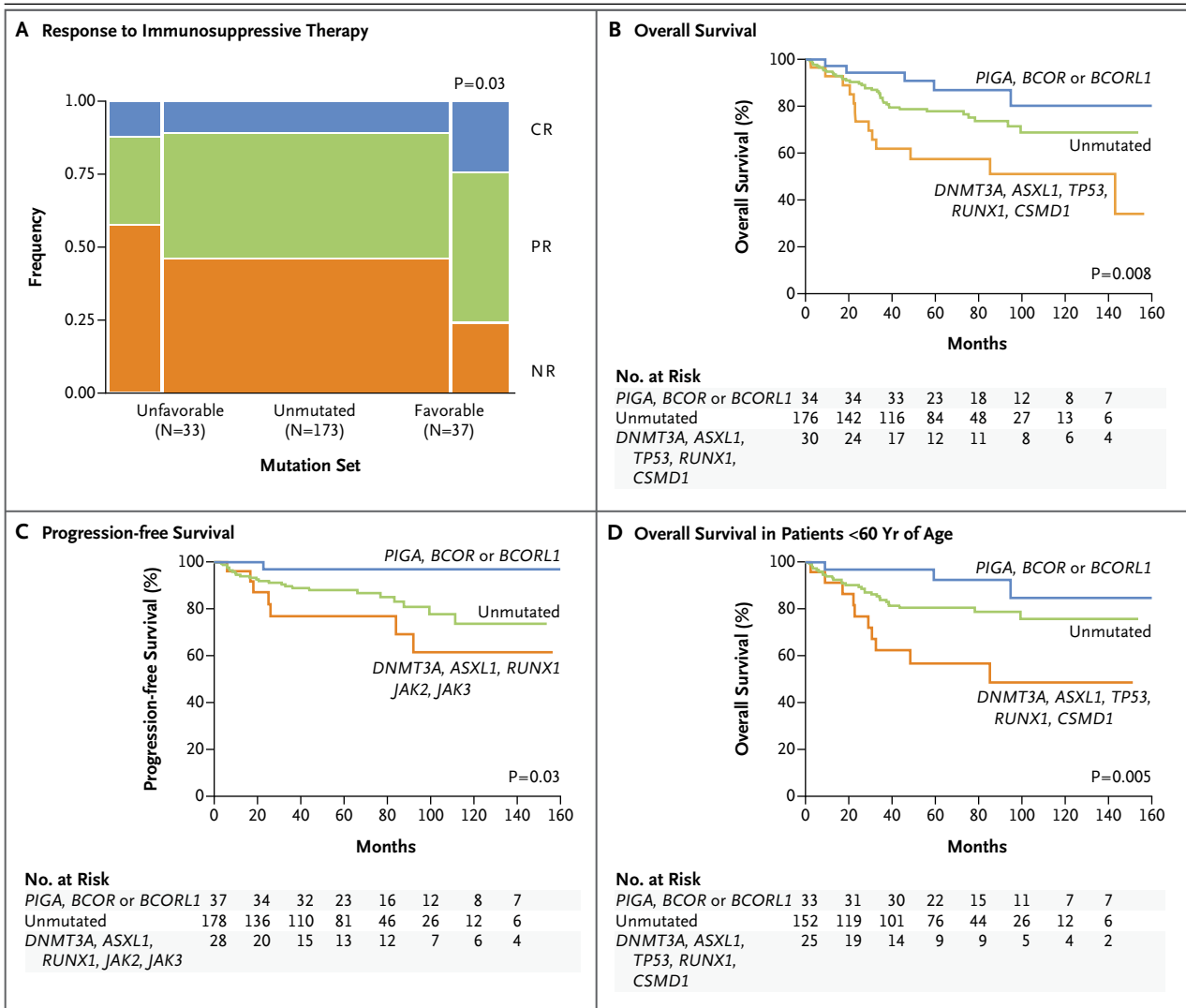
Similarly, there was no significant difference in overall survival and progression-free survival (evolution to myelodysplastic syndromes, AML, or both) between all patients with somatic mutations and those without mutations (Fig. S7A in the Supplementary Appendix). We next used machine-learning methods to assess associations between gene mutations and clinical outcomes. Both random survival forests and algorithms for the penalized likelihood approach to variable selection identified patients with better responses to immunosuppressive therapy (those with *PIGA* and *BCOR* and *BCORL1*) and patients with worse responses (those with *ASXL1*, *DNMT3A*, *TP53*, *RUNX1*, *JAK2*, *JAK3*, and *CSMD1*) than the responses in patients without these mutations (“unmutated” group) (Fig. 2A). We also identified

patients with better overall survival (those with *PIGA* and *BCOR* and *BCORL1*) and patients with worse overall survival (those with *ASXL1*, *DNMT3A*, *TP53*, *RUNX1*, and *CSMD1*) than overall survival in the unmutated group (Fig. 2B), and patients with better progression-free survival (those with *PIGA* and *BCOR* and *BCORL1*) and patients with worse progression-free survival (those with *ASXL1*, *DNMT3A*, *RUNX1*, *JAK2*, and *JAK3*) than progression-free survival in the unmutated group (Fig. 2C). Assessment of the Japanese cohort with the use of the same gene set confirmed the relationship of *ASXL1*, *DNMT3A*, *TP53*, *RUNX1*, and *CSMD1* to lower overall survival among patients with severe aplastic anemia (Fig. S7B in the Supplementary Appendix). Patients with “mixed” mutations — both “favorable” (*PIGA* or *BCOR* and *BCORL1*) and any other mutated gene — were excluded from these analyses, but stratification in models was enhanced if these patients were grouped with the favorable gene set for response to therapy and overall survival (see below) and with unfavorable genes for progression-free survival (data not shown).

Since age correlated with the presence of mutations and is a known risk factor for poor overall survival among patients with aplastic anemia,<sup>22</sup> a Cox proportional-hazards model was used to determine whether mutations were independent risk factors (Table 1). “Favorable” somatic mutations were associated with good overall survival (model 1), and “unfavorable” mutations, older age, male sex, and a lower initial reticulocyte count were associated with poor overall survival (model 2). Overall survival among younger patients was particularly influenced by the presence of unfavorable mutations (Fig. 2D, and Fig. S8 in the Supplementary Appendix).

**CHRONOLOGY OF CLONAL ARCHITECTURE IN APLASTIC ANEMIA**

We used whole-exome sequencing to define clonal hematopoiesis over time. Comprehensive detection of coding-sequence mutations was achieved in 135 samples obtained from 52 patients, including serially obtained samples from 28 patients (Table S9 in the Supplementary Appendix). Sample selection was biased toward patients with known mutations in targeted genes, since the analysis was intended to depict clonal architecture preceding clinical evolution



**Figure 2. Clinical Correlations with Somatic Mutations.**

Gene-set enrichment analysis with the use of algorithms for a penalized likelihood approach to variable selection was used to identify sets of genes that were associated with a good or poor response to immunosuppressive therapy (Panel A), overall survival (Panel B), and progression-free survival (Panel C) in the NIH cohort. Panel A shows an inferior response to immunosuppressive therapy in a group of patients with “unfavorable” mutations (*DNMT3A*, *ASXL1*, *TP53*, *RUNX1*, *JAK2*, *JAK3*, or *CSMD1*) and a superior response in patients with “favorable” mutations (*PIGA* or *BCOR* and *BCORL1*) as compared with patients in an “unmutated” group ( $P=0.03$  by the chi-square test). The width of each column represents the number of patients in each group. CR denotes complete response, NR nonresponse, and PR partial response. In Panel B, Kaplan–Meier curves for overall survival are shown for three groups: patients with favorable mutations in *PIGA* or *BCOR* and *BCORL1*, patients in the unmutated group, and patients with unfavorable mutations in *DNMT3A*, *ASXL1*, *TP53*, *RUNX1*, or *CSMD1*. In Panel C, Kaplan–Meier curves for progression-free survival among patients with favorable mutations in *PIGA*, *BCOR* or *BCORL1*, patients in the unmutated group, and patients with unfavorable *DNMT3A*, *ASXL1*, *RUNX1*, *JAK2*, or *JAK3* mutations are shown. In Panel D, Kaplan–Meier curves for overall survival among patients younger than 60 years of age are shown for three groups: patients with favorable mutations in *PIGA* or *BCOR* and *BCORL1*, patients in the unmutated group, and patients with unfavorable mutations in *DNMT3A*, *ASXL1*, *TP53*, *RUNX1*, or *CSMD1*. Log-rank tests were used for statistical comparisons among the groups. The unmutated group included patients with other candidate-gene mutations that did not cluster in gene-set enrichment analysis with either favorable or unfavorable groups. Results were similar when patients in the unmutated group (no candidate-gene mutations detected) were used as the reference group. Thirteen patients with “mixed” mutations were excluded from the gene-set enrichment analysis.

**Table 1. Risk Factors Associated with Poor Overall Survival.\***

Variable	Model 1		Model 2	
	Hazard Ratio for Death (95% CI)	P Value	Hazard Ratio for Death (95% CI)	P Value
Age >60 yr	2.72 (1.51–4.90)	<0.001	2.50 (1.41–4.43)	0.002
Male sex	2.30 (1.27–4.17)	0.006	2.27 (1.25–4.10)	0.007
Somatic-mutation category				
Favorable	0.27 (0.09–0.78)	0.016	—	—
Unmutated	0.56 (0.29–1.09)	0.09	—	—
Mixed	0.23 (0.05–1.05)	0.06	—	—
Combined	—	—	0.48 (0.25–0.91)	0.03
Blood counts at diagnosis				
Reticulocytes	0.43 (0.22–0.84)	0.014	0.47 (0.25–0.88)	0.02
Neutrophils	0.67 (0.26–1.72)	0.42	0.67 (0.26–1.75)	0.41
Lymphocytes	1.63 (0.89–2.97)	0.11	1.46 (0.83–2.60)	0.19
Platelets	1.18 (0.57–2.46)	0.66	1.34 (0.65–2.79)	0.43

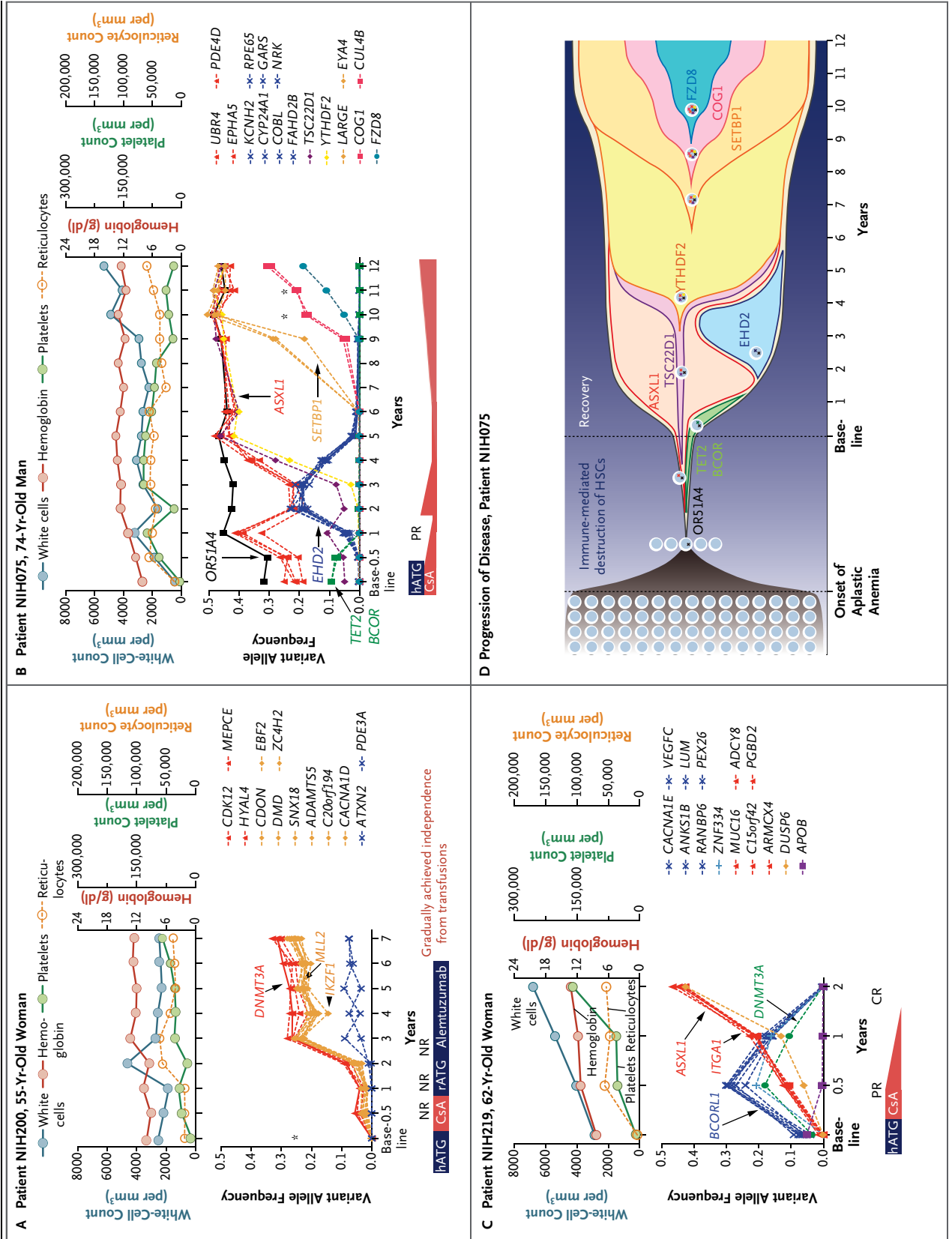
\* A multivariate Cox proportional-hazards model was used to assess risk factors. Reference (baseline) groups for categorical variables were as follows: an age of 60 years or younger, female sex, and “unfavorable” mutations. Blood counts at the time of diagnosis were treated as continuous variables after  $\log_{10}$  transformation. The P value for each variable was obtained with the use of a likelihood-ratio test. In model 1, each individual gene set (“favorable,” “unmutated,” and “mixed” mutations) was compared separately with the unfavorable gene set as a reference. In model 2, all other gene sets were combined for comparison with the unfavorable gene set. Categories of somatic mutations were derived with the use of algorithms for a penalized likelihood approach to variable selection for favorable mutations (*PIGA* and *BCOR* and *BCORL1*), unfavorable mutations (*ASXL1*, *DNMT3A*, *TP53*, *RUNX1*, and *CSMD1*), and unmutated cases (none of these mutations or no mutations); 13 patients had mixed mutations.

to myelodysplastic syndromes, AML, or both. In total, 95 somatic-mutation events in 25 of 52 patients (48%) were detected and validated by means of amplicon sequencing. Acquired mutations in exomes of genes of unknown significance defined hematopoietic-cell clones in another 5 of the 52 patients who were evaluated (Fig. S9A in the Supplementary Appendix).

When all platforms were combined, clonal hematopoiesis was detected in 85% of the 52 samples (Table S10 in the Supplementary Appendix). The mean number of nonsynonymous mutations per sample was 1.21 (when synonymous mutations were included, the total mean number of mutations was 1.83) (Fig. S9A in the Supplementary Appendix), and the mean number of nonsynonymous mutations increased significantly with age (Fig. S9B in the Supplementary Appendix). Mutations were prominently biased toward an age-related cytosine-to-thymine (C→T) transition at CpG dinucleotides; this was true as well of mutations detected by means

of targeted sequencing (Fig. S9C in the Supplementary Appendix). Transitions also accumulated in healthy older adults (Fig. S10 in the Supplementary Appendix).<sup>23–25</sup> With the exception of genes included in targeted sequencing, only four very large genes (*ITGA1*, *HDLBP*, *AKAP9*, and *MUC16*) harbored recurrent mutations. The presence of multiple subclones, inferred from significantly different variant allele frequencies of distinct mutations in targeted sequencing at 6 months after treatment, was confirmed by means of whole-exome sequencing that revealed additional new mutations.

A more detailed analysis of clonal evolution was obtained by the sequencing of serial samples from 35 patients that spanned years. In most patients, clonal hematopoiesis originated from a minor clone that was already present at the time of diagnosis. However, the subsequent temporal course of these clones was highly variable. In some cases, clones containing mutations were stable over a period of years (median, 5 years;





**Figure 3 (facing page). Temporal Profile of Mutations Detected by Means of Whole-Exome Sequencing.**

Panels A through C show the results of a representative secular analysis of somatic mutations and their relationship to peripheral-blood counts and response to immunosuppressive therapy. White-cell counts (top) and the variant allele frequency of mutations identified by means of whole-exome sequencing (bottom) are shown. The treatment received is shown at the bottom of each panel. CR denotes complete response, CsA cyclosporine, hATG horse antithymocyte globulin, NR nonresponse, PR partial response, and rATG rabbit antithymocyte globulin. Asterisks indicate the samples that were not used in whole-exome sequencing. Panel D shows the chronologic history of clonal evolution from the onset of disease to the last follow-up in Patient NIH075. Each mutated gene is depicted in the representative cells. The vertical axis indicates the absolute volume of the clones. Representative mutations in each clone are also shown. HSC denotes hematopoietic stem cell.

range, 2 to 10), generally with persistent cytopenias (Fig. S11 in the Supplementary Appendix). In others (Figs. 3A, 3B, and 3C, and Fig. S12 in the Supplementary Appendix), a large hematopoietic stem-cell clone gradually dominated the blood compartment, and new subclones were observed preceding the development of myelodysplastic syndromes, AML, or both (Fig. S13 in the Supplementary Appendix).

In Patient NIH075 from the NIH cohort, in whom mutations were fully assessed over the course of 12 years, clonal hematopoiesis was shaped by multiple rounds of acquisition of new mutations and clonal selection with progression of thrombocytopenia after an initial response to immunosuppressive therapy (Fig. 3B). The original clone, which was characterized by an *OR51A4* mutation, gave rise to several independent clones carrying *ASXL1*, *EHD2*, or coexisting *TET2* and *BCOR* mutations (Table S7 in the Supplementary Appendix), some of which appeared and disappeared, with the ultimate emergence of new subclones within the compartment carrying the *ASXL1* mutation (Fig. 3D). Patient NIH075 had persistent thrombocytopenia, which was responsive to cyclosporine. Serial bone marrow samples showed decreased megakaryocytes, but there was no apparent dysplasia, nor were there increased blasts or cytogenetic abnormalities. In other patients (Patients NIH044 and NIH169), a dominant clone was compatible with normal or near-normal blood counts. Nevertheless, clones

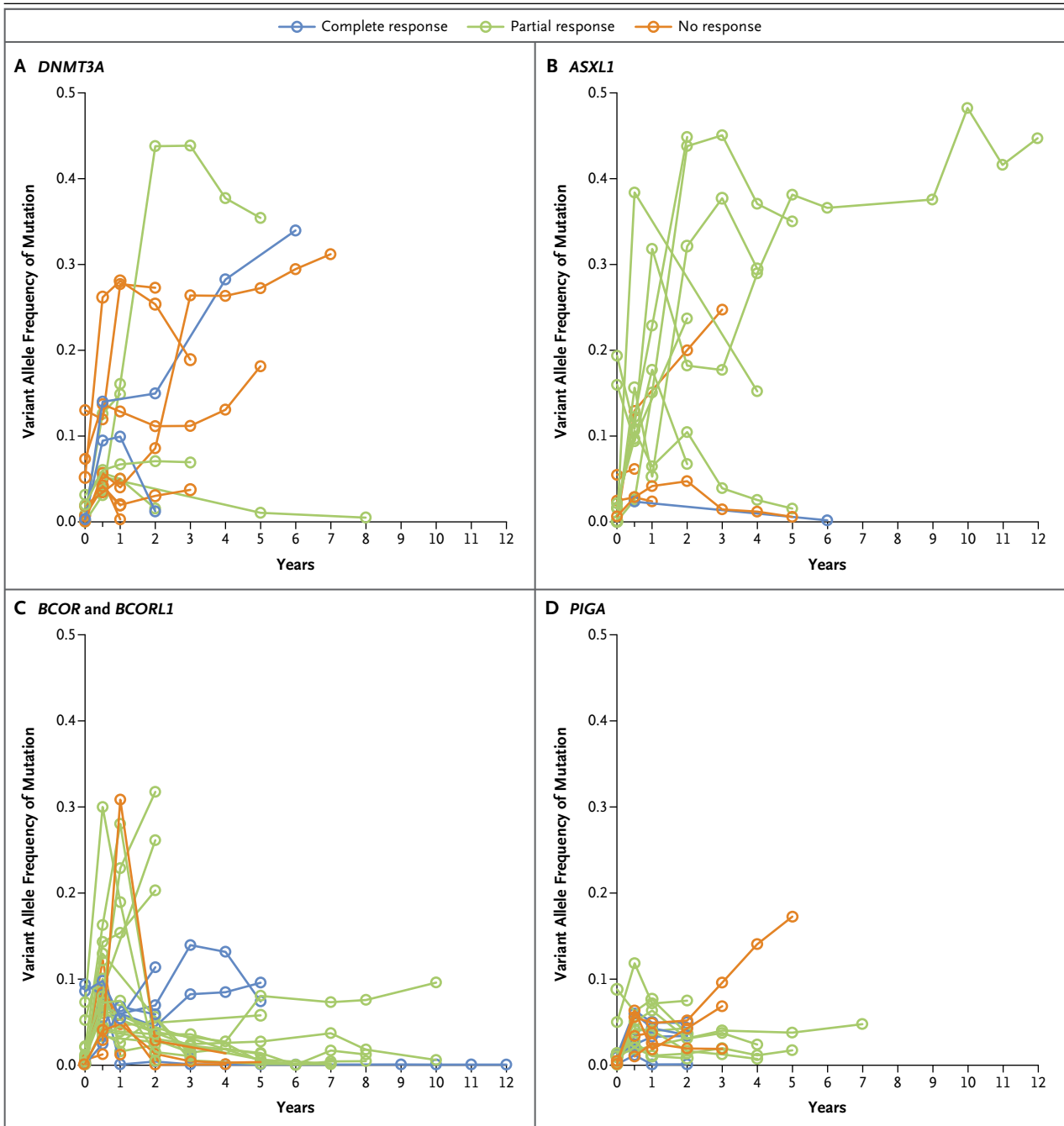
with *DNMT3A*, *ASXL1*, *RUNX1*, or *U2AF1* mutations tended to continue to become larger over time, whereas the size of clones with *PIGA* and *BCOR* and *BCORL1* mutations was stable or decreased (Fig. 4, and Fig. S14 in the Supplementary Appendix).

## DISCUSSION

Through targeted deep-sequencing, SNP array karyotyping, and whole-exome sequencing, we delineated a broad registry of genetic alterations in aplastic anemia and described their dynamics over long clinical courses. The combined results of all the assays that we used showed that about half the patients with aplastic anemia in the study had evidence of expanded hematopoietic cell clones, and about one third had acquired mutations in candidate genes for myelodysplastic syndromes, AML, or both. These results have implications for bone marrow failure, for early events in leukemogenesis, and for normal aging.

Clonal hematopoiesis in aplastic anemia was manifested most frequently as somatic mutations in a very few common driver genes (*DNMT3A*, *ASXL1*, and *BCOR* and *BCORL1*) for myelodysplastic syndromes, AML, or both. Our findings are consistent with those of recent studies<sup>26-28</sup> in which candidate-gene targeted sequencing also showed recurrent mutations in a similar spectrum of genes. In the first study, mutations were described as rare, but 2 of 39 patients had mutations with low variant allele frequency in *ASXL1* and 1 had a *DNMT3A* mutation.<sup>26</sup> In two studies<sup>27,28</sup> in which samples were obtained at various times after diagnosis, mutations in *ASXL1*, *DNMT3A*, and *BCOR* were prevalent in 19% of patients seen at King's College Hospital, London (samples were obtained relatively late after diagnosis) and were rare in a German study in which myelodysplastic syndrome mutations were detected in 2 of 38 patients (mutations with low variant allele frequency were excluded).

As compared with these recent studies, our data set was larger and involved a more consistent collection of samples at specified time points over the course of disease. Our analysis by means of whole-exome and deep sequencing, the confirmation of putative mutations with the use of multiple methods, and the consistency of findings among three cohorts define a spectrum of genomic abnormalities in aplastic anemia. Re-



**Figure 4. Temporal Profile of Common Mutations in the NIH Cohort.**

The variant allele frequencies of mutations in commonly affected genes such as *DNMT3A* (Panel A), *ASXL1* (Panel B), *BCOR* and *BCORL1* (Panel C), and *PIGA* (Panel D) are plotted for all relevant mutations. Each line indicates the time course of variant allele frequencies for an individual mutation and the response to immunosuppressive therapy at 6 months.

sults from recent studies as well as our study concerning the specific genes mutated and their low variant allele frequency in aplastic anemia raise important questions about disease classification, pathophysiology, and clinical practice.

First, despite a substantial overlap in recurrently mutated genes between aplastic anemia and myelodysplastic syndromes, these mutations were generally detected in small subpopulations, mostly at less than 10% variant allele frequen-

cies, and at least 6 months after treatment; usually years are required before clones achieve dominance. Various sets of mutations show distinct behavior and clinical effect; *BCOR*-mutant and *BCORL1*-mutant clones and *PIGA*-mutant clones tended to disappear or remain small, were associated with a better response to immunosuppressive therapy, and predicted favorable outcomes. In contrast, clones carrying mutations in *DNMT3A*, *ASXL1*, and a few other genes were more likely to increase in size over time, and these mutations (dominated by *DNMT3A* and *ASXL1*) as a group were associated with a poorer response to immunosuppressive therapy, inferior overall survival, and progression to myelodysplastic syndromes, AML, or both. Thus, in these cases, clonal hematopoiesis may represent the earliest stages of leukemogenesis visualized by deep sequencing before evident neoplastic expansion. Preleukemic clones in primary AML also frequently harbor *DNMT3A* mutations.<sup>29,30</sup> The presence of monosomy 7 at 6 months after treatment also predicted poor overall survival and progression to myelodysplastic syndromes (Fig. S15 in the Supplementary Appendix).

Second, our results show parallels between bone marrow failure and normal aging of the hematopoietic compartment. The characteristic mutation signature and correlation of mutations with patient age suggested age-related, spontaneous conversion of methylated cytosine (<sup>m</sup>C) to thymidine at <sup>m</sup>CpG sites<sup>31</sup> as a major source of nucleotide alterations in aplastic anemia.<sup>32</sup> Similar C-to-T conversion mutations accumulate in hematopoietic progenitors in healthy persons (Fig. S11 in the Supplementary Appendix).<sup>23-25</sup> Age-related clonal hematopoiesis, as detected by means of SNP array karyotyping<sup>33,34</sup> and exome sequencing,<sup>24,25,35</sup> was identified in 1 to 3% of patients with nonhematologic cancers and in healthy persons, showed a strong correlation with age, and was associated with a higher risk of the development of subsequent hematologic cancers.<sup>24,25,33-36</sup> Mutations generally appeared at a low variant allele frequency, were predominantly C-to-T conversions, and involved common mutational targets in myeloid cancers,<sup>24,25,33-36</sup> which suggests that the origin and clonal selection of these mutations are similar to those in aplastic anemia. The incidence of clonal hematopoiesis was much higher among patients with aplastic anemia, and *BCOR* and *BCORL1* and *PIGA* mutations were overrepresented, whereas *TET2* and

*JAK2* were infrequently mutated, which suggests that distinct mechanisms of clonal selection operate in the aplastic bone marrow environment. Nevertheless, acquisition of specific mutations (detected by means of candidate-gene sequencing), an increased mutational burden in dominant clones (detected by means of whole-exome sequencing), and accelerated telomere attrition preceding development of myelodysplastic syndromes, AML, or both<sup>37</sup> suggest a common mechanism that links genome aging in bone marrow failure with that in the healthy elderly.

Third, specific mutations probably have functional relevance,<sup>38,39</sup> but the exact mechanism of selection of mutated cells in aplastic anemia is unclear. Cells containing *DNMT3A* or *ASXL1* mutations may preferentially self-renew rather than differentiate in response to extrinsic signals that are especially abundant in bone marrow failure. We speculate that clones containing mutated genes that are present in aging bone marrow are selected in bone marrow failure; the lower prevalence of such clones in healthy elderly persons as compared with persons with aplastic anemia may be due to the insensitivity of deep sequencing for detecting very small populations of abnormal cells (since flow cytometry is far more sensitive than *PIGA* sequencing for the detection of small PNH clones). That clones containing *DNMT3A* or *ASXL1* mutations might initiate a pathogenic immune response would seem unlikely given the inverse correlation of their presence with response to immunosuppressive therapy. In contrast, the striking overrepresentation of *BCOR* and *BCORL1* and *PIGA* mutations as well as frequent 6pUPD involving specific HLA classes<sup>9</sup> suggest a mechanism of protection of mutated cells from immune-mediated destruction by pathogenic T cells.<sup>40,41</sup> Stratification in our models was enhanced if patients who had a mixed pattern of both favorable and unfavorable mutations were grouped in the favorable gene set for early outcomes (response to therapy and overall survival) and with unfavorable genes for later progression-free survival — consistent with the dynamics of their clonal behavior over time and presumed functionality.

Finally, despite the association of particular gene mutations observed early in the course of disease with the response to therapy and survival, it should be underscored that the complex dynamics of clonal hematopoiesis are highly vari-

able and not necessarily determinative. Monosomy 7 can develop in the absence of candidate-gene mutations in patients with aplastic anemia,<sup>37</sup> and virtually monoclonal hematopoiesis, including multiple unfavorable mutations, can be present in patients who have a response to immunosuppressive therapy, maintain good blood counts, and have prolonged survival. Close monitoring of clonal hematopoiesis by means of both deep sequencing and SNP array karyotyping will need to be combined with clinical evaluation to estimate prognosis and to guide treatment of patients with aplastic anemia.

Supported by Grant-in-Aid for Scientific Research (KAKENHI 22134006, 23249052, 26253060, and 26221308) from the Ministry of Health, Labor, and Welfare of Japan and the Japan Society for the Promotion of Science through the Funding Program for World-Leading Innovative Research and Development on Science and Technology, the Intramural Research Program of the National Heart, Lung, and Blood Institute, a grant from the Aplastic Anemia and MDS International Foundation, and a research grant from the Scott Hamilton Cancer Alliance for Research Education and Survivorship Foundation.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank Drs. X. Feng, A. LaRochelle, and C. Hourigan for their careful reading of an earlier version of the manuscript, and S. Wong, Y. Mori, M. Nakamura, and H. Higashi for technical assistance.

#### APPENDIX

The authors' full names and academic degrees are as follows: Tetsuichi Yoshizato, M.D., Bogdan Dumitriu, M.D., Kohei Hosokawa, M.D., Ph.D., Hideki Makishima, M.D., Ph.D., Kenichi Yoshida, M.D., Ph.D., Danielle Townsley, M.D., Aiko Sato-Otsubo, Ph.D., Yusuke Sato, M.D., Delong Liu, Ph.D., Hiromichi Suzuki, M.D., Colin O. Wu, Ph.D., Yuichi Shiraishi, Ph.D., Michael J. Clemente, M.S., Keisuke Kataoka, M.D., Ph.D., Yusuke Shiozawa, M.D., Yusuke Okuno, M.D., Ph.D., Kenichi Chiba, Ph.D., Hiroko Tanaka, B.A., Yasunobu Nagata, M.D., Ph.D., Takamasa Katagiri, Ph.D., Ayana Kon, M.D., Masashi Sanada, M.D., Ph.D., Phillip Scheinberg, M.D., Satoru Miyano, Ph.D., Jaroslaw P. Maciejewski, M.D., Ph.D., Shinji Nakao, M.D., Ph.D., Neal S. Young, M.D., and Seishi Ogawa, M.D., Ph.D.

The authors' affiliations are as follows: the Department of Pathology and Tumor Biology, Graduate School of Medicine, Kyoto University, Kyoto (T.Y., K.Y., A.S.-O., Y. Sato, H.S., K.K., Y. Shiozawa, Y.N., A.K., M.S., S.O.), Department of Cellular Transplantation Biology, Division of Cancer Medicine, Graduate School of Medical Sciences, Kanazawa University, Kanazawa (K.H., T.K., S.N.), Laboratory of DNA Information Analysis, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo (Y. Shiraishi, K.C., H.T., S.M.), and the Department of Pediatrics, Nagoya University Graduate School of Medicine, Nagoya (Y.O.) — all in Japan; the Hematology Branch (B.D., K.H., D.T., D.L., P.S., N.S.Y.) and Office of Biostatistics Research (C.O.W.), National Heart, Lung, and Blood Institute, Bethesda, MD; and the Department of Translational Hematology and Oncology Research, Taussig Cancer Institute, Cleveland Clinic, Cleveland (H.M., M.J.C., J.P.M.).

#### REFERENCES

- Young NS, Calado RT, Scheinberg P. Current concepts in the pathophysiology and treatment of aplastic anemia. *Blood* 2006;108:2509-19.
- Maciejewski JP, Salleri C, Sato T, Anderson S, Young NS. A severe and consistent deficit in marrow and circulating primitive hematopoietic cells (long-term culture-initiating cells) in acquired aplastic anemia. *Blood* 1996;88:1983-91.
- Scheinberg P, Young NS. How I treat acquired aplastic anemia. *Blood* 2012;120:1185-96.
- Scheinberg P, Nunez O, Weinstein B, et al. Horse versus rabbit antithymocyte globulin in acquired aplastic anemia. *N Engl J Med* 2011;365:430-8.
- Socié G, Rosenfeld S, Frickhofen N, Gluckman E, Tichelli A. Late clonal diseases of treated aplastic anemia. *Semin Hematol* 2000;37:91-101.
- Mortazavi Y, Chopra R, Gordon-Smith EC, Rutherford TR. Clonal patterns of X-chromosome inactivation in female patients with aplastic anaemia studies using a novel reverse transcription polymerase chain reaction method. *Eur J Haematol* 2000;64:385-95.
- Afable MG II, Wlodarski M, Makishima H, et al. SNP array-based karyotyping: differences and similarities between aplastic anemia and hypocellular myelodysplastic syndromes. *Blood* 2011;117:6876-84.
- Mikhailova N, Sessarego M, Fugazza G, et al. Cytogenetic abnormalities in patients with severe aplastic anemia. *Haematologica* 1996;81:418-22.
- Katagiri T, Sato-Otsubo A, Kashiwase K, et al. Frequent loss of HLA alleles associated with copy number-neutral 6pLOH in acquired aplastic anemia. *Blood* 2011;118:6601-9.
- Yoshida K, Sanada M, Shiraishi Y, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* 2011;478:64-9.
- Nannya Y, Sanada M, Nakazaki K, et al. A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. *Cancer Res* 2005;65:6071-9.
- Yamamoto G, Nannya Y, Kato M, et al. Highly sensitive method for genome-wide detection of allelic composition in non-paired, primary tumor specimens by use of affymetrix single-nucleotide-polymorphism genotyping microarrays. *Am J Hum Genet* 2007;81:114-26.
- Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. *Ann Appl Stat* 2008;2:841-60.
- Simon N, Friedman J, Hastie T, Tibshirani R. Regularization paths for Cox's proportional hazards model via coordinate descent. *J Stat Softw* 2011;39:1-13.
- Ley TJ, Ding L, Walter MJ, et al. DNMT3A mutations in acute myeloid leukemia. *N Engl J Med* 2010;363:2424-33.
- Gelsi-Boyer V, Trouplin V, Adélaïde J, et al. Mutations of polycomb-associated gene ASXL1 in myelodysplastic syndromes and chronic myelomonocytic leukaemia. *Br J Haematol* 2009;145:788-800.
- Damm F, Chesnais V, Nagata Y, et al. BCOR and BCORL1 mutations in myelodysplastic syndromes and related disorders. *Blood* 2013;122:3169-77.
- Mortazavi Y, Merk B, McIntosh J, Marsh JC, Schrezenmeier H, Rutherford TR. The spectrum of PIG-A gene mutations in aplastic anemia/paroxysmal nocturnal hemoglobinuria (AA/PNH): a high incidence of multiple mutations and evidence of a mutational hot spot. *Blood* 2003;101:2833-41.
- Nafa K, Bessler M, Castro-Malaspina H, Jhanwar S, Luzzatto L. The spectrum of somatic mutations in the PIG-A gene in paroxysmal nocturnal hemoglobinuria includes large deletions and small duplications. *Blood Cells Mol Dis* 1998;24:370-84.

20. Haferlach T, Nagata Y, Grossmann V, et al. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. *Leukemia* 2014;28:241-7.
21. Scheinberg P, Wu CO, Nunez O, Young NS. Predicting response to immunosuppressive therapy and survival in severe aplastic anaemia. *Br J Haematol* 2009;144:206-16.
22. Scheinberg P, Cooper JN, Sloand EM, Wu CO, Calado RT, Young NS. Association of telomere length of peripheral blood leukocytes with hematopoietic relapse, malignant transformation, and survival in severe aplastic anemia. *JAMA* 2010;304:1358-64.
23. Welch JS, Ley TJ, Link DC, et al. The origin and evolution of mutations in acute myeloid leukemia. *Cell* 2012;150:264-78.
24. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med* 2014;371:2488-98.
25. Genovese G, Kähler AK, Handsaker RE, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med* 2014;371:2477-87.
26. Lane AA, Odejide O, Kopp N, et al. Low frequency clonal mutations recoverable by deep sequencing in patients with aplastic anemia. *Leukemia* 2013;27:968-71.
27. Kulasekararaj AG, Jiang J, Smith AE, et al. Somatic mutations identify a subgroup of aplastic anemia patients who progress to myelodysplastic syndrome. *Blood* 2014;124:2698-704.
28. Heuser M, Schlarman C, Dobbernack V, et al. Genetic characterization of acquired aplastic anemia by targeted sequencing. *Haematologica* 2014;99(9):e165-e167.
29. Shlush LI, Zandi S, Mitchell A, et al. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature* 2014;506:328-33.
30. Yasuda T, Ueno T, Fukumura K, et al. Leukemic evolution of donor-derived cells harboring IDH2 and DNMT3A mutations after allogeneic stem cell transplantation. *Leukemia* 2014;28:426-8.
31. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature* 2001;409:860-921.
32. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature* 2013;500:415-21.
33. Jacobs KB, Yeager M, Zhou W, et al. Detectable clonal mosaicism and its relationship to aging and cancer. *Nat Genet* 2012;44:651-8.
34. Laurie CC, Laurie CA, Rice K, et al. Detectable clonal mosaicism from birth to old age and its relationship to cancer. *Nat Genet* 2012;44:642-50.
35. Xie M, Lu C, Wang J, et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat Med* 2014;20:1472-8.
36. Busque L, Patel JP, Figueroa ME, et al. Recurrent somatic TET2 mutations in normal elderly individuals with clonal hematopoiesis. *Nat Genet* 2012;44:1179-81.
37. Dumitriu B, Feng X, Townsley DM, et al. Telomere attrition and candidate gene mutations preceding monosomy 7 in aplastic anemia. *Blood* 2015;125:706-9.
38. Challen GA, Sun D, Jeong M, et al. Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat Genet* 2012;44:23-31.
39. Abdel-Wahab O, Gao J, Adli M, et al. Deletion of *Asxl1* results in myelodysplasia and severe developmental defects in vivo. *J Exp Med* 2013;210:2641-59.
40. Murakami Y, Kosaka H, Maeda Y, et al. Inefficient response of T lymphocytes to glycosylphosphatidylinositol anchor-negative cells: implications for paroxysmal nocturnal hemoglobinuria. *Blood* 2002;100:4116-22.
41. Gargiulo L, Papaioannou M, Sica M, et al. Glycosylphosphatidylinositol-specific, CD1d-restricted T cells in paroxysmal nocturnal hemoglobinuria. *Blood* 2013;121:2753-61.

Copyright © 2015 Massachusetts Medical Society.

#### CLINICAL TRIAL REGISTRATION

The *Journal* requires investigators to register their clinical trials in a public trials registry. The members of the International Committee of Medical Journal Editors (ICMJE) will consider most reports of clinical trials for publication only if the trials have been registered. Current information on requirements and appropriate registries is available at [www.icmje.org/faq\\_clinical.html](http://www.icmje.org/faq_clinical.html).