NIH Public Access
Author Manuscript
N Engl J Med. Author manuscript; available in PMC 2015 June 25.

Published in final edited form as:
N Engl J Med. 2014 December 25; 371(26): 2488–2498. doi:10.1056/NEJMoa1408617.

Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes

Siddhartha Jaiswal, M.D., Ph.D., **Pierre Fontanillas, Ph.D.**, **Jason Flannick, Ph.D.**, **Alisa Manning, Ph.D.**, **Peter V. Grauman, B.A.**, **Brenton G. Mar, M.D., Ph.D.**, **R. Coleman Lindsley, M.D., Ph.D.**, **Craig H. Mermel, M.D., Ph.D.**, **Noel Burtt, B.S.**, **Alejandro Chavez, M.D., Ph.D.**, **John M. Higgins, M.D.**, **Vladislav Moltchanov, Ph.D.**, **Frank C. Kuo, M.D., Ph.D.**, **Michael J. Kluk, M.D., Ph.D.**, **Brian Henderson, M.D.**, **Leena Kinnunen, M.Sc.**, **Heikki A. Koistinen, M.D., Ph.D.**, **Claes Ladenvall, Ph.D.**, **Gad Getz, Ph.D.**, **Adolfo Correa, M.D., Ph.D.**, **Benjamin F. Banahan, Ph.D.**, **Stacey Gabriel, Ph.D.**, **Sekar Kathiresan, M.D.**, **Heather M. Stringham, Ph.D.**, **Mark I. McCarthy, M.D.**, **Michael Boehnke, Ph.D.**[*], **Jaakko Tuomilehto, M.D., Ph.D.**[*], **Christopher Haiman, Sc.D.**, **Leif Groop, M.D., Ph.D.**, **Gil Atzmon, Ph.D.**, **James G. Wilson, M.D.**, **Donna Neuberg, Sc.D.**, **David Altshuler, M.D., Ph.D.**[*], and **Benjamin L. Ebert, M.D., Ph.D.**[†]

Department of Medicine, Division of Hematology (S.J., P.V.G., B.G.M., B.L.E.), the Department of Medical Oncology, Division of Hematological Malignancies (R.C.L.), and Center for Advanced Molecular Diagnostics and Department of Pathology (F.C.K., M.J.K.), Brigham and Women's Hospital, the Departments of Pathology (S.J., C.H.M., J.M.H., A. Chavez, G.G.) and Molecular Biology (J.F., D.A.), Center for Systems Biology (J.M.H.), Department of Medicine, Division of Cardiology and Cardiovascular Research Center (S.K.), Center for Human Genetic Research (S.K., D.A.), and Department of Medicine, Division of Endocrinology (D.A.), Massachusetts General Hospital, the Joint Program in Transfusion Medicine, Boston Children's Hospital, Brigham and Women's Hospital, and Dana–Farber Cancer Institute (S.J.), the Departments of Pediatric Oncology (B.G.M.) and Biostatistics and Computational Biology (D.N.), Dana–Farber Cancer Institute, Massachusetts General Hospital (J.M.H.), the Departments of Systems Biology (J.M.H.), Genetics (D.A.), and Medicine (D.A.), Harvard Medical School, and Harvard Stem Cell Institute (B.L.E.) — all in Boston; Program in Medical and Population Genetics, Broad Institute of the Massachusetts Institute of Technology and Harvard (P.F., J.F., A.M., N.B., S.G., S.K., D.A.), and the Department of Biology, Massachusetts Institute of Technology (D.A.), Cambridge, MA; the Department of Public Health, University of Helsinki, and Department of Chronic Disease Prevention, National Institute for Health and Welfare (V.M.), Minerva Foundation Institute for Medical Research, and the Department of Medicine, Division of Endocrinology, Helsinki University Central Hospital (H.A.K.), and the Diabetes Prevention Unit, National Institute for

Copyright © 2014 Massachusetts Medical Society.

Address reprint requests to Dr. Ebert at the Department of Medicine, Division of Hematology, Brigham and Women's Hospital, Harvard Medical School, 1 Blackfan Cir., Karp 5.210, Boston, MA 02115, or at bebert@partners.org..

[*]Dr. McCarthy is listed on behalf of the Type 2 Diabetes (T2D) Genetic Exploration by Next-Generation Sequencing in Multi-Ethnic Samples (T2D-GENES) study investigators; Dr. Boehnke, on behalf of the Genetics of Type 2 Diabetes (GoT2D) study investigators; and Dr. Altshuler, on behalf of the SIGMA T2D study investigators.

[†]A complete list of investigators in the T2D-GENES, GoT2D, and SIGMA T2D studies is provided in the Supplementary Appendix, available at NEJM.org.

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

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Health and Welfare (L.K., H.A.K, J.T.) — all in Helsinki; the Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles (B.H., C.H.); Lund University Diabetes Center, Department of Clinical Sciences, Clinical Research Center at Skåne University Hospital, Lund University, Lund, Sweden (C.L., L.G.); the Departments of Medicine (A. Correa) and Physiology and Biophysics (J.G.W.), University of Mississippi Medical Center, and the Center for Pharmaceutical Marketing and Management, University of Mississippi (B.F.B.), Oxford; Center for Statistical Genetics, University of Michigan, Ann Arbor (H.M.S., M.B.); Oxford Centre for Diabetes, Endocrinology, and Metabolism and the Wellcome Trust Centre for Human Genetics, University of Oxford, and Oxford National Institute for Health Research Biomedical Research Centre, Oxford, United Kingdom (M.I.M.); Center for Vascular Prevention, Danube University Krems, Krems, Austria, and Diabetes Research Group, King Abdulaziz University, Jeddah, Saudi Arabia (J.T.); and the Department of Medicine and Genetics, Albert Einstein College of Medicine, New York (G.A.).

Abstract

BACKGROUND—The incidence of hematologic cancers increases with age. These cancers are associated with recurrent somatic mutations in specific genes. We hypothesized that such mutations would be detectable in the blood of some persons who are not known to have hematologic disorders.

METHODS—We analyzed whole-exome sequencing data from DNA in the peripheral-blood cells of 17,182 persons who were unselected for hematologic phenotypes. We looked for somatic mutations by identifying previously characterized single-nucleotide variants and small insertions or deletions in 160 genes that are recurrently mutated in hematologic cancers. The presence of mutations was analyzed for an association with hematologic phenotypes, survival, and cardiovascular events.

RESULTS—Detectable somatic mutations were rare in persons younger than 40 years of age but rose appreciably in frequency with age. Among persons 70 to 79 years of age, 80 to 89 years of age, and 90 to 108 years of age, these clonal mutations were observed in 9.5% (219 of 2300 persons), 11.7% (37 of 317), and 18.4% (19 of 103), respectively. The majority of the variants occurred in three genes: *DNMT3A*, *TET2*, and *ASXL1*. The presence of a somatic mutation was associated with an increase in the risk of hematologic cancer (hazard ratio, 11.1; 95% confidence interval [CI], 3.9 to 32.6), an increase in all-cause mortality (hazard ratio, 1.4; 95% CI, 1.1 to 1.8), and increases in the risks of incident coronary heart disease (hazard ratio, 2.0; 95% CI, 1.2 to 3.4) and ischemic stroke (hazard ratio, 2.6; 95% CI, 1.4 to 4.8).

CONCLUSIONS—Age-related clonal hematopoiesis is a common condition that is associated with increases in the risk of hematologic cancer and in all-cause mortality, with the latter possibly due to an increased risk of cardiovascular disease. (Funded by the National Institutes of Health and others.)

Cancer is thought to arise through the stepwise acquisition of genetic or epi-genetic changes that transform a normal cell.[1] Hence, the existence of a premalignant state bearing only the initiating lesions may be detectable in some persons who have no other signs of disease. For example, multiple myeloma is frequently preceded by monoclonal gammopathy of unknown

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

significance,[2] and chronic lymphocytic leukemia is commonly preceded by monoclonal B-cell lymphocytosis.[3]

Several lines of evidence have suggested that clonal hematopoiesis resulting from an expansion of cells that harbor an initiating driver mutation might be an aspect of the aging hematopoietic system. Clonal hematopoiesis in the elderly was first demonstrated in studies that showed that approximately 25% of healthy women over the age of 65 years have a skewed pattern of X-chromosome inactivation in peripheral-blood cells,[4,5] which in some cases is associated with mutations in *TET2*.[6] Large-scale somatic events such as chromosomal insertions and deletions (indels) and loss of heterozygosity also occur in the blood of approximately 2% of persons older than 75 years of age.[7,8] Preleukemic hematopoietic stem cells harboring only the initiating driver mutation have been found in the bone marrow of patients with acute myeloid leukemia (AML) that is in remission.[9-11]

Sequencing studies have identified a set of recurrent mutations in several types of hematologic cancers.[12-24] However, the frequency of these somatic mutations in the general population is unknown. We tested the hypothesis that somatically acquired single-nucleotide variants (SNVs) and small indels might be detectable in the blood of older persons who are not known to have any hematologic abnormalities.

METHODS

SAMPLE ASCERTAINMENT

The study sample was selected from 22 population-based cohorts in three consortia (see Table S1 in the Supplementary Appendix, available with the full text of this article at NEJM.org). The protocols for these studies were approved by the ethics committees at all involved institutions; written informed consent was obtained from all participants. Persons with missing data on age (116 persons) or with cell lines as the source of DNA (492 persons) were excluded.

WHOLE-EXOME SEQUENCING AND TARGETED AMPLICON SEQUENCING

DNA was obtained from individual cohorts, and further processing was performed at the Broad Institute of Harvard and the Massachusetts Institute of Technology. In brief, genomic DNA was subject to hybrid capture, sequencing, and alignment with the use of the Broad genomics platform and Picard pipeline. We analyzed BAM files for SNVs using MuTect with OxoG filtering and for indels using Indelocator.[25,26] A clinically validated, targeted amplicon assay was used for sequencing 95 genes in select samples.

VARIANT CALLING

We searched the literature and the Catalog of Somatic Mutations in Cancer (COSMIC; http://cancer.sanger.ac.uk/cancergenome/projects/cosmic) (see Table S2 in the Supplementary Appendix) and compiled a list of pathogenic variants associated with human hematologic cancers in 160 genes. As a negative control, we also searched for variants that were recurrently seen in nonhemato-logic cancers (see Table S4 in the Supplementary Appendix).[27]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

STATISTICAL ANALYSIS

All the statistical analyses were performed with the use of the R statistical package (www.r-project.org). Full details of the statistical analysis are provided in the Methods section in the Supplementary Appendix.

RESULTS

IDENTIFICATION OF CANDIDATE SOMATIC MUTATIONS

To determine the extent of clonal hematopoiesis with somatic mutations, we analyzed whole-exome sequencing data from DNA in the peripheral-blood cells of 17,182 persons who were selected without regard to hematologic characteristics. Of these, 15,801 were case patients and controls ascertained from 22 cohorts in type 2 diabetes association studies, and the remaining 1381 were previously unsequenced participants in the Jackson Heart Study, a population-based cohort study (Table S1 in the Supplementary Appendix). The median age of the persons included in our study at the time DNA was obtained was 58 years (range, 19 to 108); 8741 were women, and 7860 had type 2 diabetes.

The identification of somatic driver mutations in cancer has come largely from studies that have compared differences in DNA sequence between tumor and normal tissue obtained from the same person. Once mutations are identified, investigators may genotype samples for these somatic variants without relying on matched normal tissue. Because we had DNA from only one source (blood), we limited our examination to variants that had been described previously in the literature in 160 recurrently mutated candidate genes in myeloid and lymphoid cancers (Table S2 in the Supplementary Appendix). We removed potential false positive variants by using variant-calling algorithms that had filters for known artifacts such as strand-bias and clustered reads, as well as by performing additional filtering for rare error modes using a “panel of normals” (sequence data from a panel of normal persons).[25] The lower limit of detection for variants depended on the depth of coverage. The median average sequencing depth for exons from the 160 genes was 84 reads (range, 13 to 144). At a sequencing depth of 84 reads, the limit of detection for SNVs was at a variant allele fraction of 0.035; the limit of detection for indels was 0.070.

With this approach, we identified a total of 805 candidate somatic variants (hereafter referred to as mutations) in 73 genes from 746 persons (Table S3 in the Supplementary Appendix). As a negative control, we searched for previously described, nonhematologic cancer-associated variants in 40 genes (Table S4 in the Supplementary Appendix)[27] and found only 10 such variants in these genes (Table S5 in the Supplementary Appendix), indicating that the rate of false discovery due to technical artifacts was low. We also verified a subset of the variants using amplicon-based, targeted sequencing; 18 of 18 variants were confirmed, with a correlation coefficient of 0.97 for the variant allele fraction between the two methods (Fig. S1 in the Supplementary Appendix).

INCREASE IN THE FREQUENCY OF CLONAL SOMATIC MUTATIONS WITH AGE

Hematologic cancers, as well as other cancers and premalignant states, increase in frequency with age. Mutations were very rare in samples obtained from patients younger than 40 years

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

of age but rose in frequency with each decade of life thereafter (Fig. 1). Mutations in genes implicated in hematologic cancers were found in 5.6% (95% confidence interval [CI], 5.0 to 6.3) of persons 60 to 69 years of age, 9.5% (95% CI, 8.4 to 10.8) of persons 70 to 79 years of age (219 of 2300 persons), 11.7% (95% CI, 8.6 to 15.7) of persons 80 to 89 years of age (37 of 317), and 18.4% (95% CI, 12.1 to 27.0) of persons 90 years of age or older (19 of 103). These rates greatly exceed the incidence of clinically diagnosed hematologic cancer in the general population.[28]

Though we searched for mutations in genes implicated in many different hematologic cancers, we primarily identified genes that were most frequently mutated in AML and the myelodys-plastic syndrome. The most commonly mutated gene was *DNMT3A* (403 variants) (Fig. 2A, and Fig. S2 in the Supplementary Appendix), followed by *TET2* (72 variants) and *ASXL1* (62 variants). In *TET2*, only exon 3 was obtained by exon capture (corresponding to approximately 50% of the coding region), and the portion of exon 12 of *ASXL1* that accounts for approximately 50% of the mutations in this gene had poor coverage depth. Thus, mutations in *TET2* and *ASXL1* are probably underrepresented in this study. Other frequently mutated genes included *TP53* (33 variants), *JAK2* (31 variants), and *SF3B1* (27 variants).

In sequencing studies of the myelodysplastic syndrome and AML, most patients have mutations in two or more driver genes (the median number of recurrently mutated genes in patients with de novo AML is five[17]). In this study, we found that 693 of 746 persons with a detectable mutation had only one mutation in the set of genes we examined (Fig. 2B, and Fig. S2 in the Supplementary Appendix), a finding that was consistent with the hypothesis that these persons had clones harboring only an initiating lesion.

The most common base-pair change in the somatic variants was a cytosine-to-thymine (C→T) transition (Fig. 2C), which is considered to be a somatic mutational signature of aging.[16,29] The median variant allele fraction for the identified mutations was 0.09 (Fig. 2D), suggesting that the variants are present in only a subset of blood cells and supporting their somatic rather than germ-line origin.

PERSISTENCE OF SOMATIC MUTATIONS OVER TIME

Blood-cell DNA obtained 4 to 8 years after the initial DNA collection was available for targeted sequencing in 13 persons with 17 somatic mutations (4 persons had 2 mutations). In all cases, the mutations detected at the earlier time point were still present at the later time point. For 10 mutations, the variant allele fraction stayed the same or decreased slightly, and for 7 mutations, the variant allele fraction increased; new mutations were detected in 2 persons. However, none of the 13 persons had a hematologic cancer (Fig. S4 in the Supplementary Appendix).

RISK FACTORS ASSOCIATED WITH SOMATIC MUTATION

To understand risk factors that contributed to having a detectable mutation, we performed a multivariable logistic-regression analysis that included age, sex, status with respect to type 2 diabetes, and ancestry as covariates (Tables S6 and S7 and Fig. S3B in the Supplementary

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Appendix). As expected, age was the largest contributor to the risk of a mutation. The incidence of the myelodysplastic syndrome is slightly higher among men than among women. In our study, among persons 60 years of age or older, men had an increased likelihood of having a detectable mutation as compared with women (odds ratio, 1.3; 95% CI, 1.1 to 1.5; P = 0.005 by logistic regression). Hispanics are reported to have a lower incidence of the myelodysplastic syndrome and myeloproliferative neoplasms than other groups in the United States.[30] In our study, we found that Hispanics had a lower risk of having a mutation than did those of European ancestry, whereas the risk in other groups did not differ significantly from the risk in persons of European ancestry (Table S6 and Fig. S5 in the Supplementary Appendix). Among the genes we queried, the spectrum of mutations did not differ significantly among ancestry groups (Fig. S6 in the Supplementary Appendix).

ASSOCIATION OF SOMATIC MUTATIONS WITH THE RISK OF HEMATOLOGIC CANCER

Clonal excess states such as monoclonal gammopathy of unknown significance are associated with an increased risk of cancer. Of the cohorts that contributed data to the study, two (the Jackson Heart Study cohort and the Multiethnic Cohort) had longitudinal follow-up information on cancer that was diagnosed after DNA collection. Together, these comprised 3342 persons, including 134 (4.0%) in whom we detected somatic mutations in the blood. In a median follow-up period of 95 months, 16 hematologic cancers were reported, of which 5 (31%) were in the group that had detectable mutations (Table S8 in the Supplementary Appendix).

In a fixed-effects meta-analysis of the two cohorts adjusted for age, sex, and status with respect to type 2 diabetes, hematologic cancers were more common by a factor of 11.1 (95% CI, 3.9 to 32.6) in persons with a detectable mutation (P<0.001). Among persons with a variant allele fraction of 0.10 or greater (indicating a higher proportion of cells in the blood carrying the mutation), the risk of a hematologic cancer was increased by a factor of nearly 50 (hazard ratio, 49; 95% CI, 21 to 120; P<0.001) (Fig. 3A). Consistent with this finding, the mean variant allele fraction at the time the blood sample was obtained was significantly higher among persons with a mutation who subsequently had a hematologic cancer than among those who did not subsequently have a hematologic cancer (25.2% vs. 12.0%, P = 0.002 by Wilcoxon rank-sum test) (Fig. 3C).

Although persons with detectable mutations had a markedly increased risk of hematologic cancer, the absolute risk remained small; overall, a hematologic cancer developed during the study period in approximately 4% of persons with a mutation (Fig. 3B). This translates to a risk of hematologic cancer of approximately 0.5% per year overall and approximately 1% per year among persons with a variant allele fraction greater than 0.10. Unfortunately, we were unable to evaluate the hematologic cancers to assess the relationship of the tumor to the mutant clone that preceded it.

BLOOD-CELL INDEXES OF PERSONS WITH SOMATIC MUTATIONS

Somatic mutations found in persons with the myelodysplastic syndrome and in those with AML lead to abnormal differentiation, ineffective hematopoiesis, and cytopenias. Data on blood counts were available for 3107 persons from five cohorts (the Jackson Heart Study

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

cohort, controls without diabetes from the Longevity Genes Project, the Botnia Study cohort, the Siblings in Malmö cohort, and the Helsinki Siblings with Diabetes cohort), including 139 persons with a detectable mutation. When we evaluated persons who had single mutations (*ASXL1*, *DNMT3A*, *JAK2*, *SF3B1*, or *TET2*) or mutations in more than one gene as compared with those who had no mutations, we found no significant differences in mean white-cell counts, hemoglobin levels, platelet counts, or white-cell differential counts after accounting for age and sex (Fig. S7 in the Supplementary Appendix). The only significant difference in blood-cell indexes was an increase in red-cell distribution width (13.8% vs. 13.4% in normocytic subjects, P = 0.002 by Wilcoxon rank-sum test), and this difference was driven by a large increase in this variable in a subgroup of persons with mutations (Fig. S8 in the Supplementary Appendix).

We also sought to determine whether the presence of a mutation was associated with an increased likelihood of abnormally low blood counts (Table S10 in the Supplementary Appendix). Most of the persons with a mutation had no cytopenias; among those who had a cytopenia, the frequency of any single cytopenia was not higher among those with mutations than among those without mutations. A small fraction of participants had multiple cytopenias, and these persons were more likely than those without multiple cytopenias to have mutations (odds ratio, 3.0; P = 0.04 by Fisher's exact test). Furthermore, among persons with anemia, those with mutations had a higher percentage of unexplained anemia than did those without mutations (Table S11 in the Supplementary Appendix).

ASSOCIATION OF SOMATIC MUTATIONS WITH OVERALL SURVIVAL

We next assessed whether the presence of a somatic mutation had an effect on overall survival, on the basis of available data from 5132 persons in seven cohorts (Fig. 4) with a median follow-up period of 96 months. In a model adjusted for age, sex, and status with respect to type 2 diabetes, carrying a mutation was associated with increased all-cause mortality (hazard ratio, 1.4; P = 0.02 by fixed-effects meta-analysis with beta coefficients derived from Cox proportional-hazards models for individual cohorts) (Fig. 4A). A Kaplan–Meier survival analysis of data from participants who were 70 years of age or older showed an increased risk of death among those with a mutation (P<0.001 by rank-sum test) (Fig. 4B; for results according to cohort, see Fig. S9 in the Supplementary Appendix). Death from hematologic neoplasms alone could not account for the observed increase in mortality, since only 1 person with a mutation died from a hematologic cancer. When we performed a cause-specific mortality analysis, we found that persons with mutations had a higher risk of death from cardiovascular causes but not from cancer (Fig. S10 in the Supplementary Appendix).

Because we found that the presence of a somatic mutation was significantly associated with a higher red-cell distribution width, we also examined whether harboring mutations was synergistic with an elevated red-cell distribution width with respect to the risk of death. High red-cell distribution width has been associated with increased all-cause mortality in the aging and critically ill population,[31-33] but the mechanism behind this association is uncertain. Information on red-cell distribution width was available for 2409 participants in two cohorts. In an analysis adjusted for age, sex, and status with respect to type 2 diabetes, we found that persons who had a mutation in conjunction with a red-cell distribution width

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

of 14.5% (the upper limit of the normal range) or higher had a markedly increased risk of death as compared with those who had a normal red-cell distribution width and did not have mutations (hazard ratio, 3.7; P<0.001 by fixed-effects meta-analysis with beta-coefficients estimated from Cox models for the two cohorts). In contrast, persons who had a high red-cell distribution width and no mutation had a more modest increase in mortality (Fig. 4C, and Fig. S11 in the Supplementary Appendix).

ASSOCIATION OF SOMATIC MUTATION WITH CARDIOMETABOLIC DISEASE

A recent study showed that large, somatic chromosomal alterations in peripheral-blood cells were associated with type 2 diabetes.[34] We also found that somatic mutations in genes known to cause hematologic cancers were modestly but significantly associated with an increased risk of type 2 diabetes, even after adjustment for potential confounding variables (odds ratio, 1.3; P<0.001) (Tables S6 and S7 in the Supplementary Appendix). Participants with type 2 diabetes were slightly more likely to have mutations than were those without type 2 diabetes in each age group (Fig. S5 in the Supplementary Appendix).

Cardiovascular disease is the leading cause of death worldwide. Given the association of somatic mutations with increased all-cause and cardiovascular-related mortality, we performed association analyses of data from two cohorts comprising 3353 persons for whom data on coronary heart disease and ischemic stroke were available. After excluding persons with prior events, we found an increased cumulative incidence of both coronary heart disease and ischemic stroke among those carrying a mutation (Fig. S12A and S12B in the Supplementary Appendix). In multivariable analyses that included age, sex, status with respect to type 2 diabetes, systolic blood pressure, and body-mass index as covariates, the hazard ratios for incident coronary heart disease and ischemic stroke among persons carrying a somatic mutation as compared with those without a mutation were 2.0 (95% CI, 1.2 to 3.5; P = 0.02) and 2.6 (95% CI, 1.3 to 4.8; P = 0.003), respectively (Fig. S12C through S12F in the Supplementary Appendix).

Data on the traditional risk factors of smoking, total cholesterol level, and high-density lipoprotein cholesterol level were also available for a subgroup of participants. The presence of a somatic mutation remained significantly associated with incident coronary heart disease and ischemic stroke even in the presence of these risk factors, and the risk was even greater among persons who had a variant allele fraction of 0.10 or greater (Table S12 in the Supplementary Appendix).

DISCUSSION

We found that somatic mutations leading to clonal outgrowth of hematopoietic cells were frequent in the general population we studied, since 10% of persons older than 70 years of age carried these lesions. The exact prevalence of clonal hematopoiesis is dependent on how cancer-causing mutations are defined and on the sensitivity of the technique used to detect mutations and thus may substantially exceed this estimate. Clonal hematopoiesis appeared to involve a substantial proportion of the affected tissue in most persons; on the basis of the proportion of alleles with the somatic mutation, we found that a median of 18% of peripheral-blood leukocytes were part of the abnormal clone. In the small number of cases

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

we were able to study longitudinally, clonal hematopoiesis also persisted over time; in all tested cases, the mutations were still present after 4 to 8 years.

We found that the genes that were most commonly mutated in clonal hematopoiesis were *DNMT3A*, *TET2*, and *ASXL1*. This finding is consistent with the results of previous studies that showed that *DNMT3A* and *TET2* mutations were frequent and early events in AML and the myelodysplastic syndrome.[9,10,14,16] Murine models of *DNMT3A* or *TET2* loss of function have shown that mutant hematopoietic stem cells have altered methylation patterns in pluripotency genes and a competitive advantage over wild-type stem cells, but cancer rarely develops in mice, and when it does, it develops only after a long latency period.[35-38] Similarly, our data show that humans with clonal hematopoiesis can live for many years without hematologic cancers developing, though they are at increased risk as compared with those without mutations.

Certain genes that are commonly mutated in AML and the myelodysplastic syndrome were absent or very rare in this study. Their rarity probably indicates that they are cooperating rather than initiating mutations. Although mutations in genes specific for lymphoid cancers were rarely detected, *TET2* and *DNMT3A* are frequently mutated in some lymphoid cancers, and the initiating event for such tumors may occur in a hematopoietic stem cell.[24,39-42] This may explain why some of the cancers that developed in persons with these mutations were lymphoid.

Our data showed that the majority of persons with clonal mutations in peripheral blood did not have the myelodysplastic syndrome or some other hematologic cancer; in addition, in a majority of the persons we evaluated, no clinically diagnosed cancer developed in the near term. At this time, it would be premature to genetically screen healthy persons for the presence of a somatic clone, since the positive predictive value for the presence of cancer or for the development of cancer is low. Further studies will be needed to definitively assess the natural history of clonal hematopoiesis.

Perhaps the most surprising finding in our study was the increased mortality among persons with clones as compared with those without clones. This effect is much larger than can be explained by hematologic cancers alone, is synergistic with high red-cell distribution width (which could be a marker of perturbation of hematopoiesis due to the clone), and may be related to the increased risk of incident coronary heart disease and ischemic stroke in persons with clones. The association of somatic mutations with nonhematologic disease may be due to confounding by variables that are currently unknown or may simply represent a shared consequence of the underlying process of aging. Alternatively, it may represent an underlying shared pathophysiology of seemingly unrelated disorders. For example, cells of the monocyte– macrophage lineage are considered to be important mediators of atherosclerosis and type 2 diabetes,[43,44] but it is unknown whether their function may be altered by somatic mutations that occur in stem cells.

In summary, we found that somatic mutations that drive clonal expansion of blood cells were a common finding in the elderly and most frequently involved *DNMT3A*, *TET2*, or *ASXL1*. We propose that age-related clonal hematopoiesis is a common premalignant

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

condition that is also associated with increased overall mortality and increased risk of cardiometabolic disease.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported by grants from the National Institutes of Health (NIH) (R01HL082945) and the Gabrielle's Angel Foundation, by Leukemia and Lymphoma Society Scholar and Specialized Center of Research (SCOR) Awards (to Dr. Ebert), and by an NIH Training Grant (5T32HL066987, to Dr. Jaiswal). The T2D-GENES (Type 2 Diabetes Genetic Exploration by Next-Generation Sequencing in Multi-Ethnic Samples) consortium was supported by grants from the NIH (U01s DK085526, DK085501, DK085524, DK085545, and DK085584). The GoT2D (Genetics of Type 2 Diabetes) consortium is supported by grants from the Medical Research Council (G0601261 and High Throughput Genomics Hub grant G0900747 91070), the Wellcome Trust (090367, 090532, 098381, 083948, 085475), and the NIH (DK088389). The Longevity Genes Project is supported by grants from the NIH (P01AG021654, 1R01AG042188, P30AG038072), and by a Paul Glenn Foundation Grant. The Jackson Heart Study is supported by contracts (HSN268201300046C, HHSN268201300047C, HHSN268201300048C, HHSN268201300049C, HHSN268201300-050C) from the National Heart, Lung, and Blood Institute and the National Institute on Minority Health and Health Disparities. Dr. McCarthy is a Wellcome Trust Senior Investigator and an NIH Research Senior Investigator. Exome sequencing in the Jackson Heart Study was funded by an award from the National Human Genome Research Institute (NHGRI) of the NIH (U54 HG003067, to Dr. Gabriel).

We thank the members of the Firehose Cancer Genome Analysis team for technical support and members of the Ebert Laboratory for helpful comments on the manuscript; Dr. Sifya Mill-man for providing blood count information for participants in the Longevity Genes Project; and all the members of the T2D-GENES, GoT2D, and SIGMA T2D consortia who were involved in ascertainment of the study sample and the sequencing effort.

References

1. Nowell PC. The clonal evolution of tumor cell populations. Science. 1976; 194:23–8. [PubMed: 959840]
2. Kyle RA, Therneau TM, Rajkumar SV, et al. A long-term study of prognosis in monoclonal gammopathy of undetermined significance. N Engl J Med. 2002; 346:564–9. [PubMed: 11856795]
3. Rawstron AC, Bennett FL, O'Connor SJ, et al. Monoclonal B-cell lymphocytosis and chronic lymphocytic leukemia. N Engl J Med. 2008; 359:575–83. [PubMed: 18687638]
4. Busque L, Mio R, Mattioli J, et al. Nonrandom X-inactivation patterns in normal females: lyonization ratios vary with age. Blood. 1996; 88:59–65. [PubMed: 8704202]
5. Champion KM, Gilbert JG, Asimakopoulos FA, Hinshelwood S, Green AR. Clonal haemopoiesis in normal elderly women: implications for the myeloproliferative disorders and myelodysplastic syndromes. Br J Haematol. 1997; 97:920–6. [PubMed: 9217198]
6. 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. [PubMed: 23001125]
7. Jacobs KB, Yeager M, Zhou W, et al. Detectable clonal mosaicism and its relationship to aging and cancer. Nat Genet. 2012; 44:651–8. [PubMed: 22561519]
8. 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. [PubMed: 22561516]
9. Jan M, Snyder TM, Corces-Zimmerman MR, et al. Clonal evolution of preleu kemic hematopoietic stem cells precedes human acute myeloid leukemia. Sci Transl Med. 2012; 4:149ra18.
10. Shlush LI, Zandi S, Mitchell A, et al. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. Nature. 2014; 506:328–33. [PubMed: 24522528]
11. Corces-Zimmerman MR, Hong WJ, Weissman IL, Medeiros BC, Majeti R. Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. Proc Natl Acad Sci U S A. 2014; 111:2548–53. [PubMed: 24550281]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

12. Mardis ER, Ding L, Dooling DJ, et al. Recurring mutations found by sequencing an acute myeloid leukemia genome. N Engl J Med. 2009; 361:1058–66. [PubMed: 19657110]
13. Bejar R, Stevenson K, Abdel-Wahab O, et al. Clinical effect of point mutations in myelodysplastic syndromes. N Engl J Med. 2011; 364:2496–506. [PubMed: 21714648]
14. Papaemmanuil E, Cazzola M, Boult-wood J, et al. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. N Engl J Med. 2011; 365:1384–95. [PubMed: 21995386]
15. Walter MJ, Ding L, Shen D, et al. Recurrent DNMT3A mutations in patients with myelodysplastic syndromes. Leukemia. 2011; 25:1153–8. [PubMed: 21415852]
16. Welch JS, Ley TJ, Link DC, et al. The origin and evolution of mutations in acute myeloid leukemia. Cell. 2012; 150:264–78. [PubMed: 22817890]
17. The Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med. 2013; 368:2059–74. [PubMed: 23634996] [Erratum, N Engl J Med 2013;369:98.]
18. Papaemmanuil E, Gerstung M, Malcovati L, et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood. 2013; 122:3616–27. [PubMed: 24030381]
19. Walter MJ, Shen D, Shao J, et al. Clonal diversity of recurrently mutated genes in myelodysplastic syndromes. Leukemia. 2013; 27:1275–82. [PubMed: 23443460]
20. Zhang J, Grubor V, Love CL, et al. Genetic heterogeneity of diffuse large B-cell lymphoma. Proc Natl Acad Sci U S A. 2013; 110:1398–403. [PubMed: 23292937]
21. Morin RD, Mendez-Lago M, Mungall AJ, et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. Nature. 2011; 476:298–303. [PubMed: 21796119]
22. Lenz G, Davis RE, Ngo VN, et al. Oncogenic CARD11 mutations in human diffuse large B cell lymphoma. Science. 2008; 319:1676–9. [PubMed: 18323416]
23. Lohr JG, Stojanov P, Lawrence MS, et al. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. Proc Natl Acad Sci U S A. 2012; 109:3879–84. [PubMed: 22343534]
24. Neumann M, Heesch S, Schlee C, et al. Whole-exome sequencing in adult ETPALL reveals a high rate of DNMT3A mutations. Blood. 2013; 121:4749–52. [PubMed: 23603912]
25. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol. 2013; 31:213–9. [PubMed: 23396013]
26. Costello M, Pugh TJ, Fennell TJ, et al. Discovery and characterization of artifactual mutations in deep coverage targeted capture sequencing data due to oxidative DNA damage during sample preparation. Nucleic Acids Res. 2013; 41(6):e67. [PubMed: 23303777]
27. Lawrence MS, Stojanov P, Mermel CH, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature. 2014; 505:495–501. [PubMed: 24390350]
28. US population data — 1969–2012. National Cancer Institute; Bethesda, MD: 2014. Surveillance, Epidemiology, and End Results Program.. (http://www.seer.cancer.gov/popdata)
29. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. Nature. 2013; 500:415–21. [PubMed: 23945592]
30. Rollison DE, Howlader N, Smith MT, et al. Epidemiology of myelodysplastic syndromes and chronic myeloproliferative disorders in the United States, 2001-2004, using data from the NAACCR and SEER programs. Blood. 2008; 112:45–52. [PubMed: 18443215]
31. Patel KV, Ferrucci L, Ershler WB, Longo DL, Guralnik JM. Red blood cell distribution width and the risk of death in middle-aged and older adults. Arch Intern Med. 2009; 169:515–23. [PubMed: 19273783]
32. Perlstein TS, Weuve J, Pfeffer MA, Beckman JA. Red blood cell distribution width and mortality risk in a community-based prospective cohort. Arch Intern Med. 2009; 169:588–94. [PubMed: 19307522]
33. Bazick HS, Chang D, Mahadevappa K, Gibbons FK, Christopher KB. Red cell distribution width and all-cause mortality in critically ill patients. Crit Care Med. 2011; 39:1913–21. [PubMed: 21532476]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

34. Bonnefond A, Skrobek B, Lobbens S, et al. Association between large detectable clonal mosaicism and type 2 diabetes with vascular complications. Nat Genet. 2013; 45:1040–3. [PubMed: 23852171]
35. Jeong M, Sun D, Luo M, et al. Large conserved domains of low DNA methylation maintained by Dnmt3a. Nat Genet. 2014; 46:17–23. [PubMed: 24270360]
36. Koh KP, Yabuuchi A, Rao S, et al. Tet1 and Tet2 regulate 5-hydroxymethylcyto-sine production and cell lineage specification in mouse embryonic stem cells. Cell Stem Cell. 2011; 8:200–13. [PubMed: 21295276]
37. Challen GA, Sun D, Jeong M, et al. Dnmt3a is essential for hematopoietic stem cell differentiation. Nat Genet. 2012; 44:23–31. [PubMed: 22138693]
38. Moran-Crusio K, Reavie L, Shih A, et al. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. Cancer Cell. 2011; 20:11–24. [PubMed: 21723200]
39. Quivoron C, Couronné L, Della Valle V, et al. TET2 inactivation results in pleio-tropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. Cancer Cell. 2011; 20:25–38. [PubMed: 21723201]
40. Odejide O, Weigert O, Lane AA, et al. A targeted mutational landscape of angioimmunoblastic T-cell lymphoma. Blood. 2014; 123:1293–6. [PubMed: 24345752]
41. Asmar F, Punj V, Christensen J, et al. Genome-wide profiling identifies a DNA methylation signature that associates with TET2 mutations in diffuse large B-cell lymphoma. Haematologica. 2013; 98:1912–20. [PubMed: 23831920]
42. Couronné L, Bastard C, Bernard OA. TET2 and DNMT3A mutations in human T-cell lymphoma. N Engl J Med. 2012; 366:95–6. [PubMed: 22216861]
43. Libby P. Inflammation in atherosclerosis. Nature. 2002; 420:868–74. [PubMed: 12490960]
44. Olefsky JM, Glass CK. Macrophages, inflammation, and insulin resistance. Annu Rev Physiol. 2010; 72:219–46. [PubMed: 20148674]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

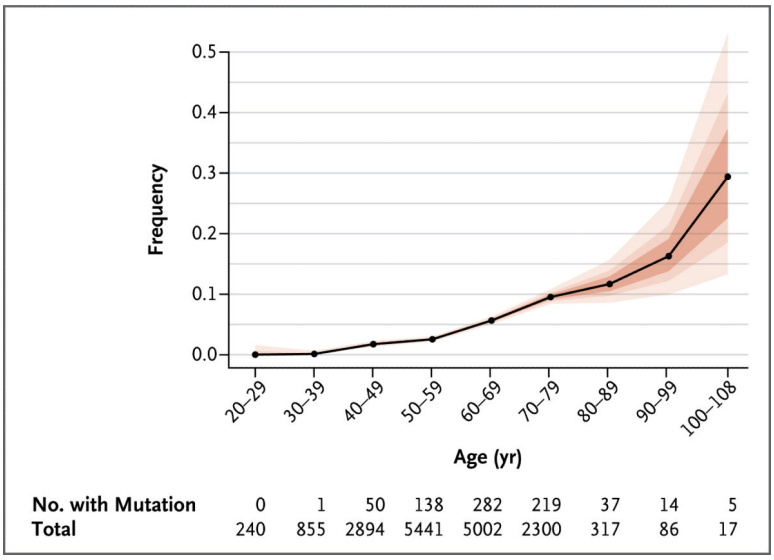

Figure 1. Prevalence of Somatic Mutations, According to Age

Colored bands, in increasingly lighter shades, represent the 50th, 75th, and 95th percentiles.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

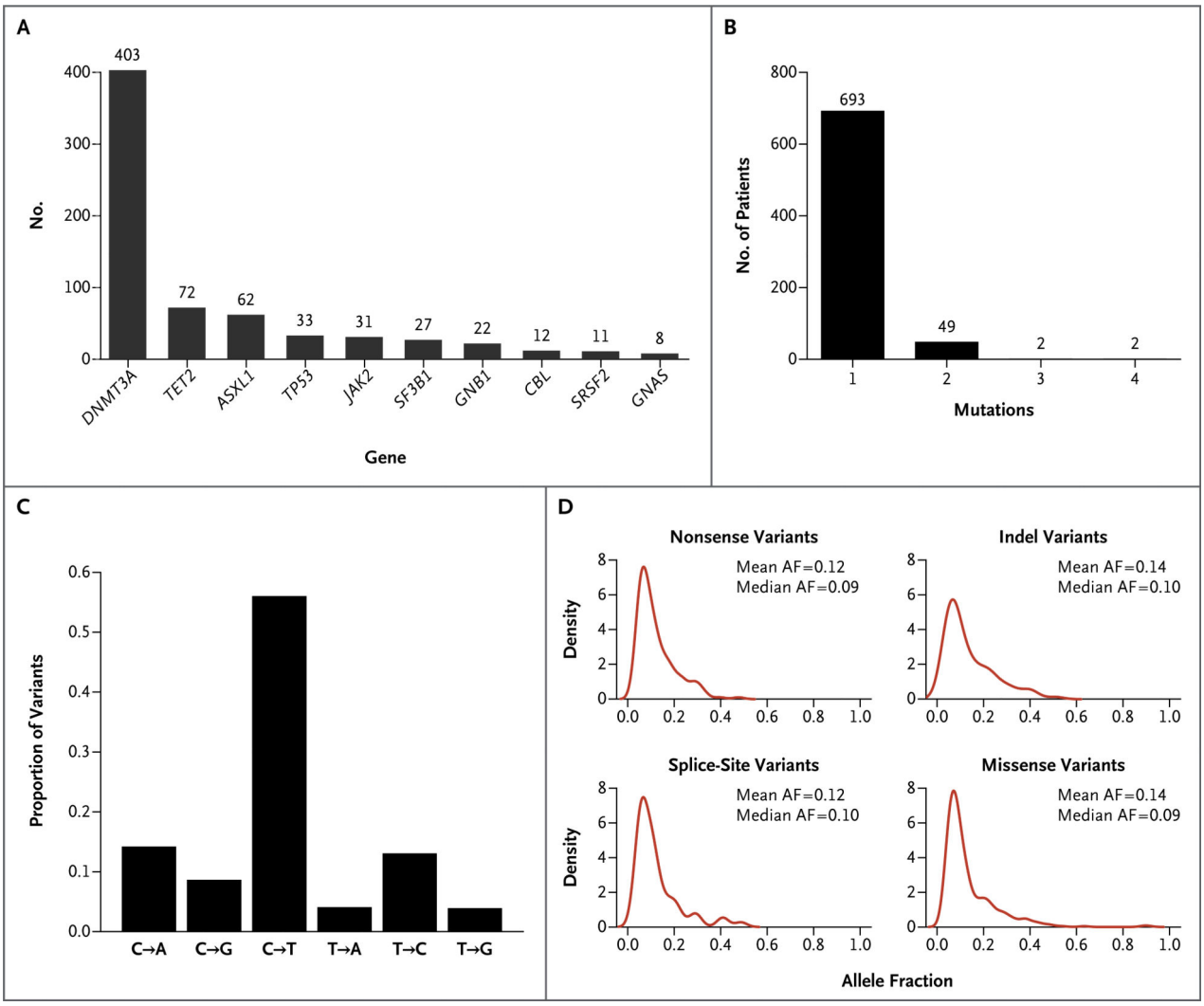

Figure 2. Characteristics of Candidate Somatic Variants

Panel A shows the 10 most frequently mutated genes implicated in hematologic cancers. Panel B shows the number of persons with 1, 2, 3, or 4 candidate variants. Panel C shows the distribution of the types of single-nucleotide base-pair changes seen in the candidate variants. Panel D shows the allele fractions (AFs) of candidate somatic variants. The allele fraction was calculated as the number of variant reads divided by the number of variant-plus-reference reads. For variants on the X chromosome in men, this number was divided by 2. Indel denotes insertions and deletions.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

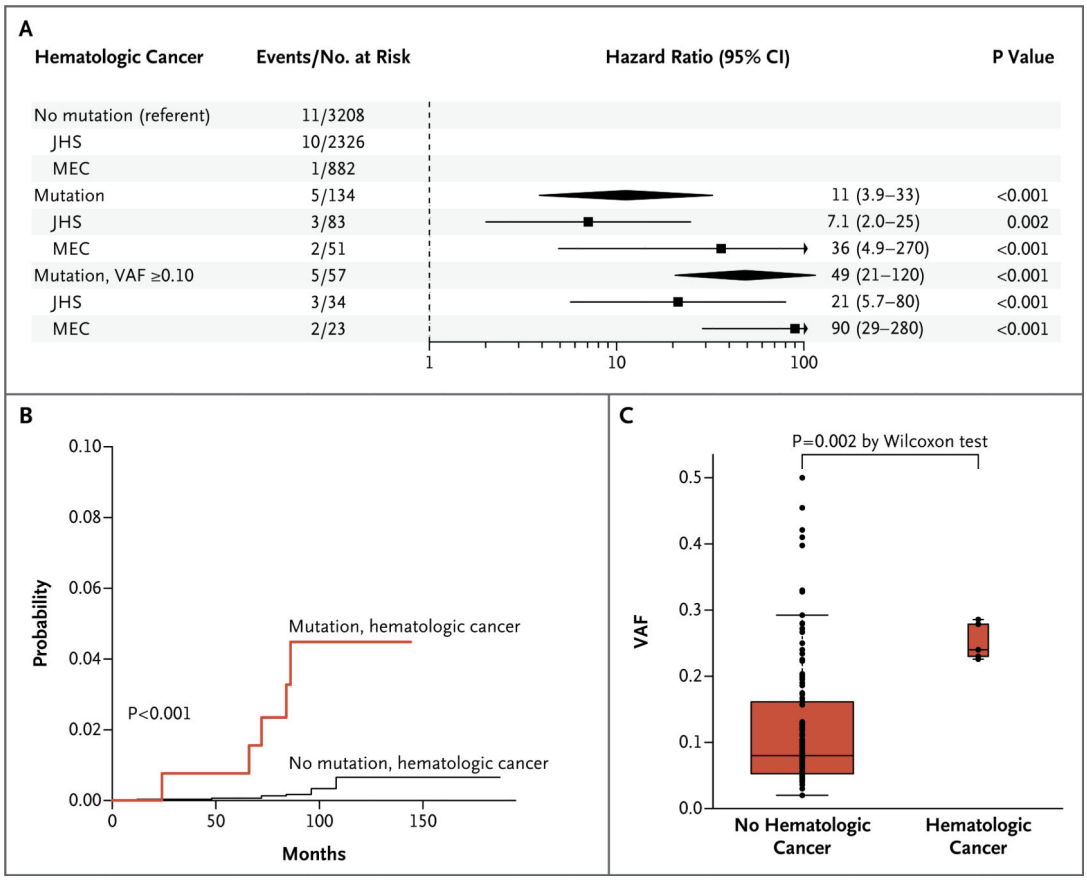

Figure 3. Development of Hematologic Cancers

Panel A is a forest plot of the risk of a hematologic cancer among persons with somatic mutations overall and among those with a variant allele fraction (VAF) of 0.10 or higher, as compared with those without mutations, in two cohorts: the Jackson Heart Study (JHS) cohort and the Multiethnic Cohort (MEC). Boxes indicate the hazard ratio for an individual cohort, with horizontal lines indicating 95% confidence intervals, and diamonds represent the results of a fixed-effects meta-analysis of the two cohorts. We estimated hazard ratios by means of competing risks regression, with death as the competing risk. The analysis includes adjudicated cancer information from the MEC and unadjudicated information ascertained through annual interviews with participants in the JHS. For interview data, leukemia, lymphoma, multiple myeloma, blood cancer, and spleen cancer were considered to be hematologic cancers. All models included age groups (<50 years, 50 to 59 years, 60 to 69 years, and ≥70 years), status with respect to type 2 diabetes, and sex as covariates. Panel B shows the cumulative incidence of hematologic cancer in the JHS cohort and the MEC. Curves were generated from competing-risks data, with death as the competing risk. Panel C shows the VAF in persons in whom a hematologic cancer developed and in those in whom a hemato-logic cancer did not develop. The top and bottom of each box represent the first and third quartiles, the horizontal line within the box represents the median, and the I bars represent 1.5 times the interquartile range.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

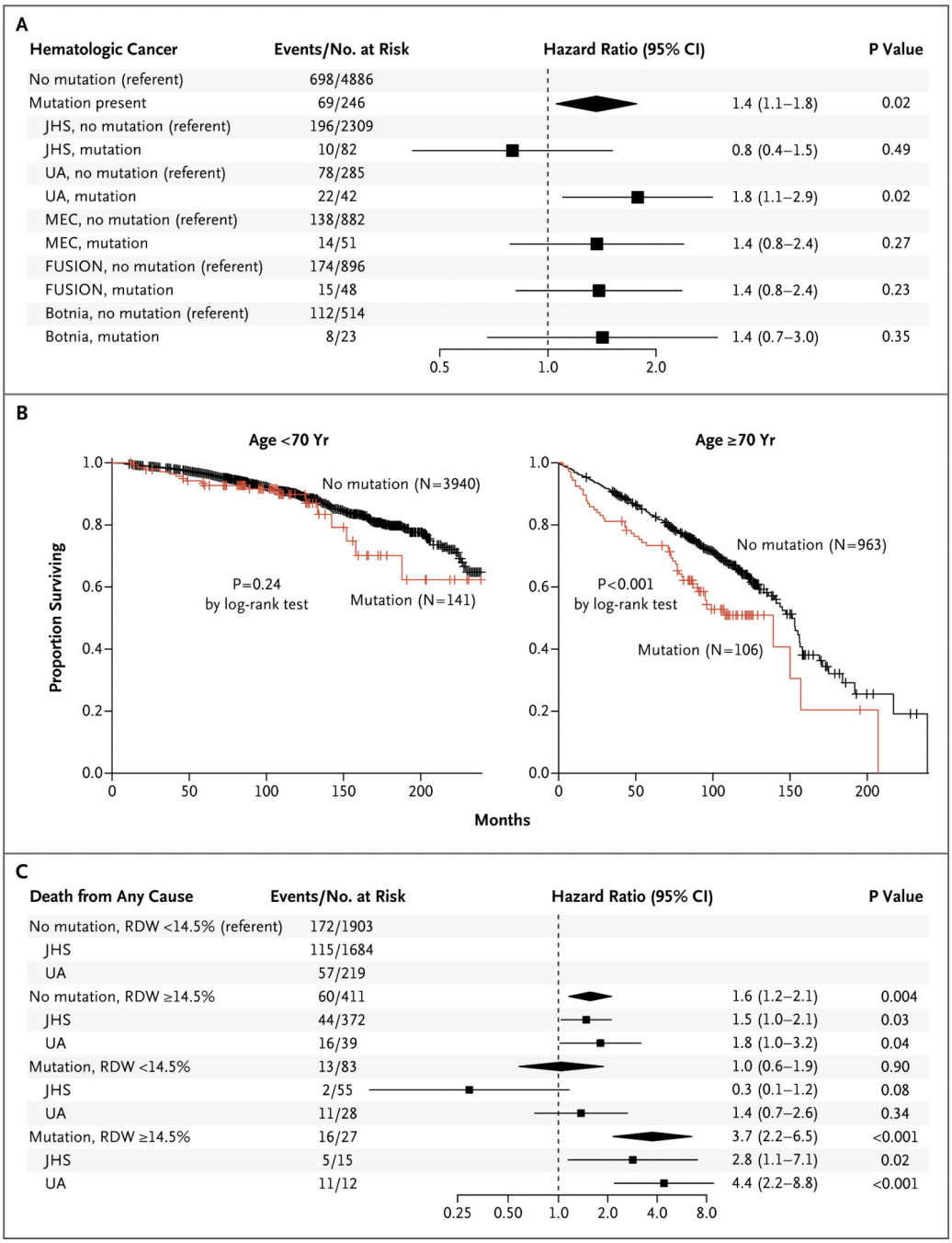

Figure 4. Effect of Somatic Mutations on All-Cause Mortality

Panel A is a forest plot of the risk of death from any cause associated with having a somatic clone, among participants from the JHS cohort, the Ashkenazi cohort of the Longevity Genes Project (UA), the MEC, the Finland–United States Investigation of NIDDM Genetics Study (FUSION) cohort, and the Botnia Study cohort. Boxes indicate the hazard ratio for an individual cohort, with horizontal lines indicating 95% confidence intervals, and diamonds represent the results of a fixed-effects meta-analysis of all cohorts. All models included age groups (<60 years, 60 to 69 years, 70 to 79 years, 80 to 89 years, and 90 years or older),

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

status with respect to type 2 diabetes, and sex as covariates in a Cox proportional-hazards analysis. The Botnia Study includes the Helsinki Siblings with Diabetes cohort and data from the Scania Diabetes Registry. Panel B shows Kaplan–Meier survival curves generated from data from the same cohorts as those included in Panel A. The left panel includes data from participants who were younger than 70 years of age at the time of DNA ascertainment, and the right panel data from participants who were 70 years of age or older. Panel C shows the results of a Cox proportional-hazards analysis of all-cause mortality among persons with and those without mutations, stratified according to normal or high red-cell distribution width (RDW). Boxes indicate the hazard ratio for an individual cohort, with horizontal lines indicating 95% confidence intervals, and diamonds represent the results of a fixed-effects meta-analysis of all cohorts. All models included age groups (<60 years, 60 to 69 years, 70 to 79 years, 80 to 89 years, and 90 years or older), status with respect to type 2 diabetes, and sex as covariates.