Blood First Edition Paper, prepublished online January 29, 2016; DOI 10.1182/blood-2015-12-685925

Uncompromised ten-year survival of oldest old carrying somatic mutations in *DNMT3A* and *TET2*

Erik B. van den Akker[1,2,*], Steven J. Pitts[3], Joris Deelen[1,4], Matthijs H. Moed[1], Shobha Potluri[3], Jeroen van Rooij[5], H. Eka D. Suchiman[1], Nico Lakenberg[1], Wesley J. de Dijcker[1], André G. Uitterlinden[5,6], Robert Kraaij[5,6], Albert Hofman[6], Anton J.M. de Craen[7], Jeanine J. Houwing-Duistermaat[8], Gert-Jan B. van Ommen[9] on behalf of the Genome of the Netherlands Consortium[§], David R. Cox[3†], Joyce B.J. van Meurs[5], Marian Beekman[1,4], Marcel J.T. Reinders[2], P. Eline Slagboom[1,4]

1) Molecular Epidemiology, Leiden University Medical Center, Einthovenweg 20, 2333 ZC, Leiden, The Netherlands.
2) The Delft Bioinformatics Lab, Delft University of Technology, Mekelweg 4, 2628 CD, Delft, The Netherlands.
3) Rinat-Pfizer Inc, 230 E. Grand Ave., South San Francisco, CA 94080, United States of America.
4) Netherlands Consortium for Healthy Ageing, The Netherlands.
5) Internal Medicine, Erasmus Medical Center, Wytemaweg 80, 3015 CN, Rotterdam, The Netherlands.
6) Epidemiology, Erasmus Medical Center, Wytemaweg 80, 3015 CN, Rotterdam, The Netherlands.
7) Gerontology and Geriatrics, Leiden University Medical Center, Albinusdreef 2, 2333 ZA, Leiden, The Netherlands.
8) Medical Statistics, Leiden University Medical Center, Einthovenweg 20, 2333 ZC, Leiden, The Netherlands.
9) Human Genetics, Leiden University Medical Center, Einthovenweg 20, 2333 ZC Leiden, The Netherlands.

* To whom correspondence should be addressed:
Erik van den Akker (e.b.van_den_akker@lumc.nl)
Leiden University Medical Centre
Einthovenweg 20, 2333 ZC, Leiden, The Netherlands
Building 2 [Research], Room T5-46
Tel: +31 (0)71 526 97 44
Fax: +31 (0)71 526 82 80

§ Genome of the Netherlands Consortium members are listed in Supplementary Appendix S1
† In memoriam

Copyright © 2016 American Society of Hematology

To the editor:

Recent large-scale sequencing studies report recurrent somatic mutations in the blood of elderly individuals in genes previously linked to clonal expansion of hematopoietic stem cells[1–4]. Particularly for *DNMT3A* and *TET2*, a steep age-associated increase in the prevalence of somatic mutations is observed from middle age onward[2–4]. In addition, prospective analyses performed in predominantly middle-aged individuals, show an increased risk for all-cause mortality for carriers of such mutations as compared to non-carriers[3,4]. Jointly, these data suggest a rapidly increasing vulnerability amongst the elderly for adverse health effects associated with clonal expansion of hematopoietic stem cells. However, prospective data on elderly somatic mutations carriers are scarce. We therefore investigated the association between all-cause mortality and carriership of somatic mutations in genes linked to clonal expansion of hematopoietic stem cells in a large elderly subsample (N=864, 80 years and older) derived from two large-scale community-dwelling Dutch cohort studies[5,6].

For the present study we investigated whole blood-derived genomes of 646 individuals of 80 years and older from the Rotterdam Study[5] (RS; mean age at inclusion: 84.6 years; range: 80.0-105.8 years, Supplementary Appendix S2) and 218 individuals of 89 years and older from the Leiden Longevity Study[6] (LLS; mean age at inclusion: 94.0 years; range: 88.9-103.4 years, Supplementary Appendix S2). Jointly, this elderly subsample consists of 597 participants aged 80 to 89 years and 267 participants aged over 90 years, which is twice the number of participants for the respective age-categories as compared to any other study previously conducted on this topic[2–4]. Selected elderly participants of the RS and LLS were followed for all-cause mortality for a median 8.7 and 9.2 years respectively, which was sufficiently long to identify the age at death of 81.3% and 93.6% of the respective study subsamples. Methods of DNA sequencing and analysis are described in Supplementary

Appendix S3. The ethical committees of the involved institutes approved both studies and written informed consent was obtained from all study participants.

Using this unique cohort of sequenced oldest old, we first set out to confirm the recurrent acquisition of somatic mutations in genes linked to clonal hematopoiesis in the blood of highly aged individuals. For this we curated a list of 15 genes (Supplementary Appendix S4) reported to harbour recurrent somatic mutations in any of the large-scale sequencing studies in the blood of normal individuals conducted to date[2–4]. Thus identified genes were analysed for putative somatic mutations according to gene-specific inclusion criteria set by Jaiswal *et al.* (Supplementary Appendix S4).

The mutational analysis identified 39 (6.0%) and 40 (18.3%) unique carriers of respectively 42 and 46 mutations for the RS and LLS elderly subsamples respectively, predominantly in *DNMT3A* and *TET2* (Figure 1A, Supplementary Appendix Table S1, S2). The observed prevalence of somatic mutations in genes linked to clonal hematopoiesis in the RS and LLS elderly subsamples is consistent with the age-associated increase observed by Xie *et al.* (Figure 1B). Our observation thus confirms the age-associated increase of detectable somatic mutations in genes previously linked to hematopoietic malignancies reported by Xie *et al.* and extends this observation up to the highest ages.

The fraction of reads annotated to the alternative alleles (VAF) is generally much lower for the identified mutations than the 50% expected for germline heterozygous variants (RS: median: 21.6%, IQR: 14.1-29.6%; LLS: median: 23.2%, IQR=16.5-31.7%, Figure 1C). This finding indicates that the identified mutations were only present in a part of the sequenced blood cells, and thus corroborates the hypothesized clonal outgrowth of hematopoietic stem cells.

Compared to the three previous studies predominantly including middle-aged

participants, a high prevalence of mutation carriers with two mutations was observed in the RS and LLS elderly subsample (Figure 1D; 9 in 864 (1.04%) vs. 6 in 2,636 (0.28%)[2], 18 in 12,380 (0.15%)[3] and 49 in 17,182 (0.28%)[4]).

The spatial correlation between the identified variants within *DNMT3A* and *TET2* with respect to the primary protein sequence and previous reports in COSMIC[7] further corroborates our findings (Figure 1E, Supplementary Figure S1). Additional Sanger sequencing experiments in the LLS elderly subsample (Supplementary Appendix S5) confirmed the presence for 18 out of 19 tested mutations in *DNMT3A* (Figure 1E; diamonds and squares) and *TET2* (Supplementary Figure S1). Moreover, Sanger sequencing in siblings of 6 mutant carriers, who inherited the identical genetic alleles from their parents at these loci as the mutant carriers (Supplementary Appendix S6), did not show the identified somatic mutations (6 out of 6, Figure 1E; diamonds), thus confirming that these variants in *DNMT3A* and *TET2* were indeed acquired during life.

Having identified carriers of somatic mutations in genes linked to clonal hematopoiesis in the RS and LLS elderly subsample, we next assessed the impact on survival of carrying such mutations. When analysing the impact of carriership of the identified somatic mutations in the 15 genes previously linked to clonal expansion of hematopoietic stem cells, no difference in survival was observed between carriers and non-carriers for neither the RS (HR=0.83 (0.58-1.17), p=0.29, Figure 2A) nor the LLS (HR=0.94 (0.65-1.35), p=0.61, Figure 2B, Supplementary Appendix S6) elderly subsample. Also, a fixed-effect meta-analysis showed no indications of a compromised survival, (HR=0.88 (0.68-1.13), *p*=0.32, Figure 2C, Supplementary Appendix S6).

Using DNA sequencing data in an elderly subsample derived from two large-scale community-dwelling Dutch cohort studies[5,6], we confirm that somatic mutations in genes previously linked to hematopoietic malignancies are common in the oldest old, especially in *DNMT3A* and *TET2*. Yet, in two independent studies, jointly comprising the largest sample in this age range to date, we found no indications that the potentially premalignant mutations compromise the 8-10 year survival of highly aged carriers.

In contrast to our findings, two recent large-scale sequencing studies in peripheral blood performed in 12,300[3] and 17,182[4] normal mostly middle-aged individuals found a significant increased risk for all-cause mortality amongst carriers of premalignant somatic mutations, predominantly in *DNMT3A* and *TET2*. The difference in mortality risk between middle-aged and highly aged people may lie in the fact that the oldest old suffer from many other co-morbidities affecting their mortality rate. Causes of death or co-incident morbidities at the time of death may support this hypothesis, however, neither of such data are available for our studies. Lastly, there is intrinsic selection bias when investigating the oldest old – the success of these highly aged individuals in coping with at least some of the adverse effects during aging could limit the ability to detect adverse health effects associated with age-related clonal expansion.

A possible limitation of our study might relate to the lower sequencing depth, allowing for a less sensitive detection of variants characterized by a low allele fraction. However, when Jaiswal *et al.* stratified their mortality analyses on the median VAF, they observed that the observed increased risk on mortality could largely be attributed to carriers of variants with the largest VAF. Also the relatively modest size could limit our study. However, a power analysis (Supplementary Appendix S6) indicated that we should be able to detect a significant increased risk on all-cause mortality amongst mutation carriers compared to non-carriers.

We conclude that, unlike previous reports in predominantly middle-aged individuals[3,4], somatic mutations in genes linked to clonal expansion of hematopoietic stem cells do not compromise the 8-10 year survival in the oldest old.

Acknowledgements

This study has received funding from the Medical Delta (COMO), Pfizer Inc (USA) and European Union's Seventh Framework Programme (FP7/2007–2011) under grant agreement number 259679. The Leiden Longevity Study was financially supported by the Innovation-Oriented Research Program on Genomics (SenterNovem IGE05007), and the next gen sequencing by the Netherlands Consortium for Healthy Ageing (grant 050-060-810), all in the framework of the Netherlands Genomics Initiative, Netherlands Organization for Scientific Research (NWO) and by BBMRI-NL, a Research Infrastructure financed by the Dutch government (NWO 184.021.007).

The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The Exome Sequencing data set was funded by the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) sponsored Netherlands Consortium for Healthy Aging (NCHA; project nr. 050-060-810), by the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, and by the and by a Complementation Project of the Biobanking and Biomolecular Research Infrastructure Netherlands (BBMRI-NL; www.bbmri.nl; project number CP2010-41).

The authors thank the study participants of the Leiden Longevity Study and the Rotterdam Study, the staff and participating general practitioners and pharmacists of both studies. The generation and management of the exome sequencing data for the Rotterdam Study was executed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Netherlands. We thank Mr. Pascal Arp, Ms. Mila Jhamai, Mr. Marijn Verkerk for their help in creating the Rotterdam Study Exome

Sequencing database. In addition, we thank Prof. Veelken and the reviewers for their useful comments on the manuscript.

Author contributions

EBA, SJP, MJTR, MB, PES conceived the analyses and wrote the paper. EBA, SJP, JD, MHM, SP, JR, RK and JBM performed data processing and the statistical analyses. HEDS, NL, and WJD performed validation experiments. AGU, AH, AJMC, DRC, GJBO and PES conceived the employed studies and contributed reagents and materials for the experiments. JJHD provided statistical counselling.

Disclosure of Conflicts of Interest

The authors would like to disclose the following interests: SJP was previously employed by the for-profit health care company Pfizer Inc., for which he also has a stock ownership and research funding to disclose. SJP is currently employed by the for-profit health care company 23andME. SP is currently employed by the for-profit health care company Pfizer Inc., for which she also has a stock ownership and research funding to disclose. In addition, SP has an intellectual property interest to disclose.

References

1. Busque L, Patel JP, Figueroa ME, et al. Recurrent somatic TET2 mutations in normal elderly individuals with clonal hematopoiesis. *Nat. Genet.* 2012;44(11):1179–81.

2. Xie M, Lu C, Wang J, et al. Age-related cancer mutations associated with clonal hematopoietic expansion. *Nat. Med.* 2014;

3. Genovese G, Kähler AK, Handsaker RE, et al. Clonal Hematopoiesis and Blood-Cancer Risk Inferred from Blood DNA Sequence. *N. Engl. J. Med.* 2014;141126135947008.

4. Jaiswal S, Fontanillas P, Flannick J, et al. Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes. *N. Engl. J. Med.* 2014;1–11.

5. Hofman A, van Duijn CM, Franco OH, et al. The Rotterdam Study: 2012 objectives and design update. *Eur. J. Epidemiol.* 2011;26(8):657–86.

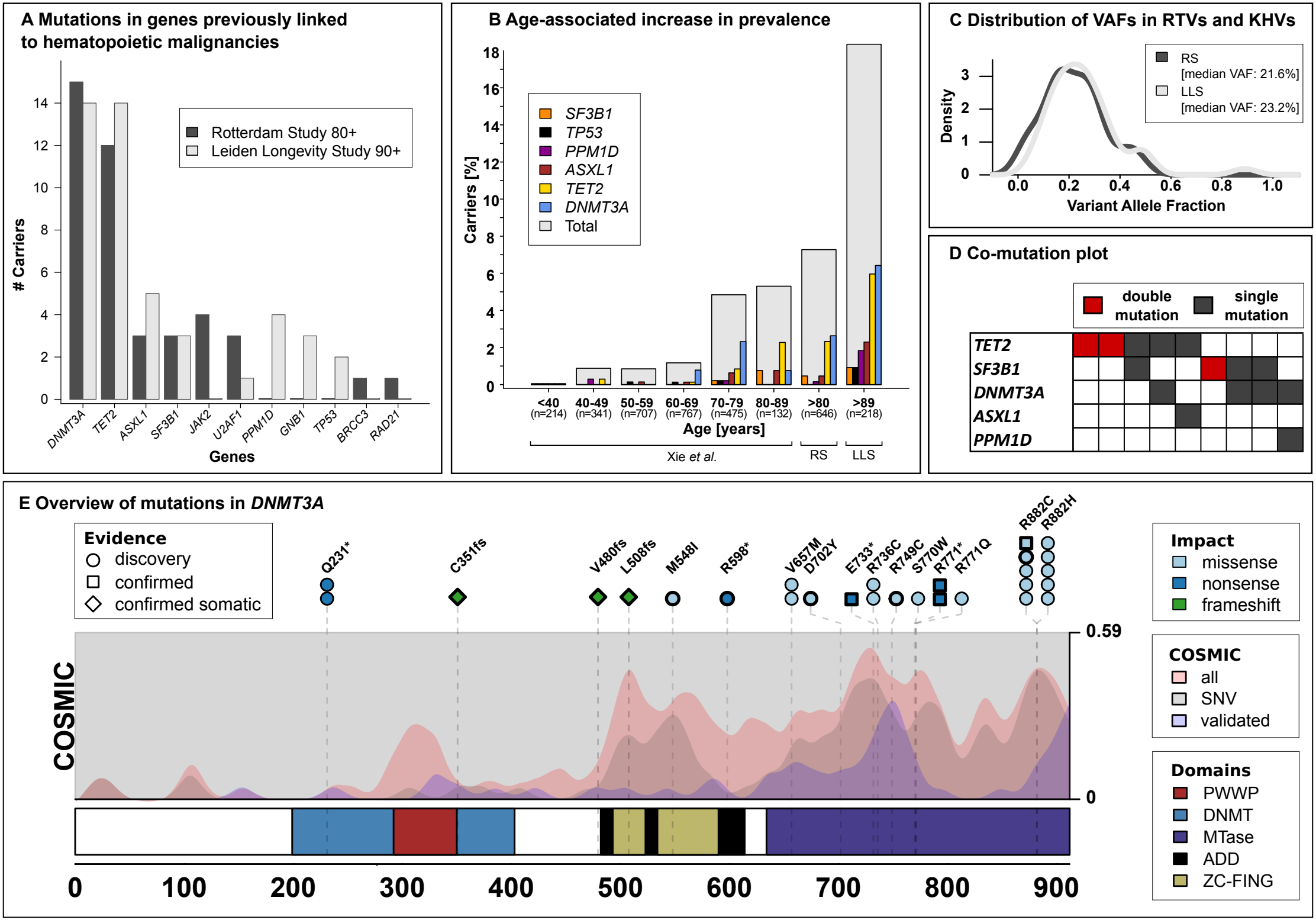
6. Schoenmaker M, de Craen AJM, de Meijer PHEM, et al. Evidence of genetic enrichment for exceptional survival using a family approach: the Leiden Longevity Study. *Eur. J. Hum. Genet.* 2006;14(1):79–84.

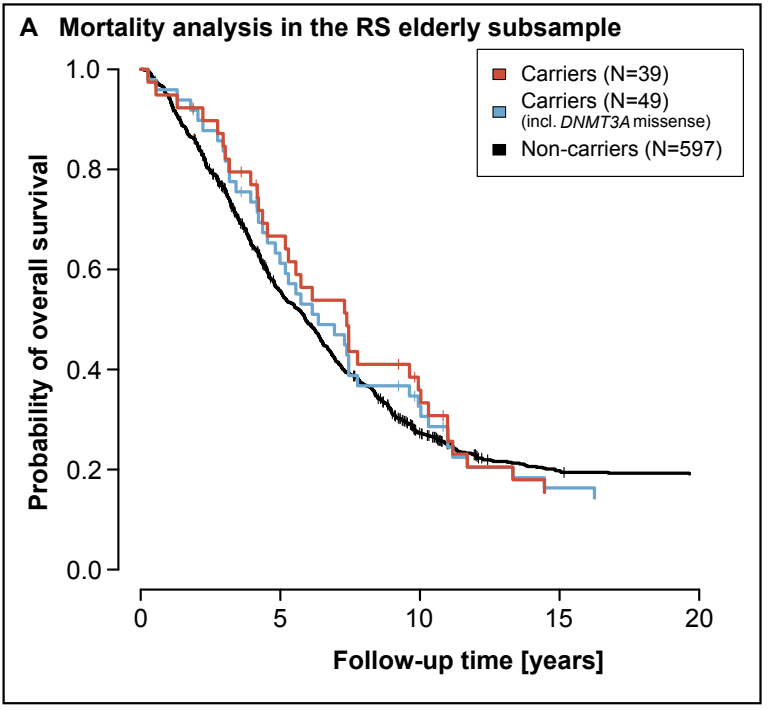
7. Forbes SA, Beare D, Gunasekaran P, et al. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res.* 2014;43(Database issue):D805–11.

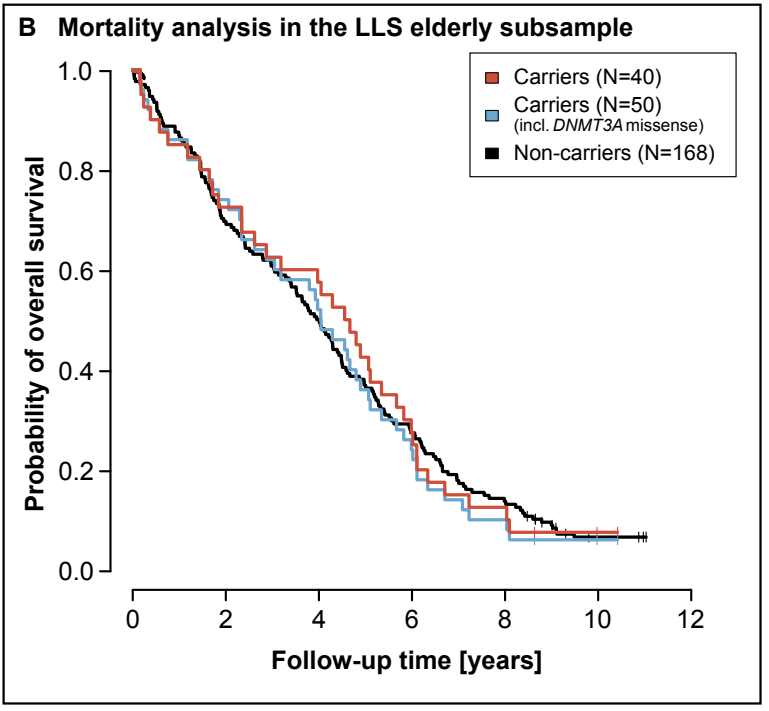
Figure Legends

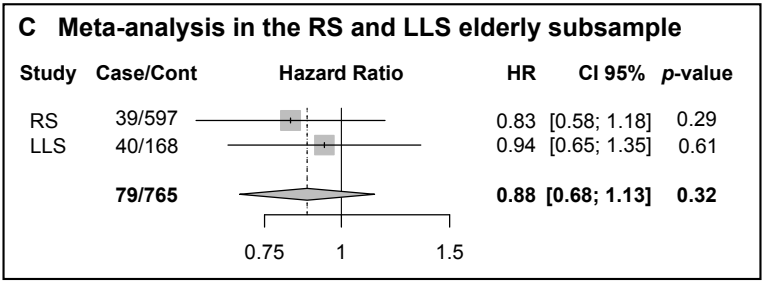
Figure 1: Characterization of identified variants in blood of the RS and LLS elderly subsample. Panel A: Barplot of the number of individuals carrying a mutation, split by genes and study. Note that only 11 out of the 15 investigated genes had a mutation (See for a complete list the S4 in the Supplementary Appendix). Panel B: Prevalence of carriers of somatic mutations stratified by age category using data of Xie *et al.* and the observations in the RS and LLS elderly subsample. Panel C: Distribution of Variant Allele Fractions (VAFs) of the identified mutations. Panel D: Co-mutation plot of carriers with two independent mutations. Panel E: Overview of mutations in *DNMT3A* identified in the RS and LLS elderly subsample. Variants are annotated at the top with a color-coding to indicate the impact and a shape to indicate the types of follow-up experiments. Circles indicate mutations detected in our sequencing data; squares indicate mutations also validated by Sanger sequencing; diamonds indicate mutations also validated by Sanger sequencing and absent in an IBD2 matched sib, i.e. confirming somatic variations. Mutations identified in multiple carriers are indicated with stacked annotations and those having bold borders were identified in the LLS. Missense variants are only included whenever they are present on a curated list of recurrently reported variants in Catalogue Of Somatic Mutations In Cancer (COSMIC)[7] assembled by Jaiswal *et al.*. Domains: DNMT: DNA methyltransferase interaction domain; PWWP: conserved DNA binding domain, MTase: methyltransferase domain; ADD: Histon binding domains; ZC-FING: Zinc Finger domains. COSMIC: Densities of somatic variants identified in hematopoietic or lymphoid tissue collected by the Catalogue Of Somatic Mutations In Cancer (COSMIC)[7] database: all small variants (red), missense SNVs (grey), small variants confirmed to be of somatic origin (blue).

Figure 2: Kaplan-Meier survival curves of the RS and LLS elderly subsample. Panel A: Kaplan-Meier curves for the 39 mutation carriers and 596 non-carriers in the RS elderly subsample. Panel B: Kaplan-Meier curves for the 40 mutation carriers and 168 non-carriers in the LLS elderly subsample. Since Jaiswal *et al.* and Genovese *et al.* do not agree on the status of *DNMT3A* missense mutations, we excluded *DNMT3A* missense mutation carriers from non-carriers in both the RS and LLS elderly subsample (Table S6 and S7 in the Supplementary Appendix). Panel C: Forest plot combining the Cox proportional hazards analyses in the RS and LLS elderly subsample.


A Mutations in genes previously linked to hematopoietic malignancies
Carriers
Rotterdam Study 80+
Leiden Longevity Study 90+
DNMT3A
TET2
ASXL1
SF3B1
JAK2
U2AF1
PPM1D
GNB1
TP53
BRCC3
RAD21
Genes
B Age-associated increase in prevalence
Carriers [%]
SF3B1
TP53
PPM1D
ASXL1
TET2
DNMT3A
Total
<40 (n=214)
40-49 (n=341)
50-59 (n=707)
60-69 (n=767)
70-79 (n=475)
80-89 (n=132)
>80 (n=646)
>89 (n=218)
Age [years]
Xie et al.
RS
LLS
C Distribution of VAFs in RTVs and KHVs
Density
RS [median VAF: 21.6%]
LLS [median VAF: 23.2%]
Variant Allele Fraction
D Co-mutation plot
double mutation
single mutation
TET2
SF3B1
DNMT3A
ASXL1
PPM1D
E Overview of mutations in DNMT3A
Evidence
discovery
confirmed
confirmed somatic
Q231*
C351fs
V480fs
L508fs
M548I
R598*
V657M
D702Y
E733*
R736C
R749C
S770W
R771*
R771Q
R882C
R882H
COSMIC
0.59
0
Impact
missense
nonsense
frameshift
COSMIC
all
SNV
validated
Domains
PWWP
DNMT
MTase
ADD
ZC-FING
0
100
200
300
400
500
600
700
800
900


A Mortality analysis in the RS elderly subsample
Carriers (N=39)
Carriers (N=49)
(incl. DNMT3A missense)
Non-carriers (N=597)
Probability of overall survival
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
Follow-up time [years]


B Mortality analysis in the LLS elderly subsample
Carriers (N=40)
Carriers (N=50)
(incl. DNMT3A missense)
Non-carriers (N=168)
Probability of overall survival
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
12
Follow-up time [years]


C Meta-analysis in the RS and LLS elderly subsample
Study
Case/Cont
Hazard Ratio
HR
CI 95%
p-value
RS
39/597
0.83
[0.58; 1.18]
0.29
LLS
40/168
0.94
[0.65; 1.35]
0.61
79/765
0.88
[0.68; 1.13]
0.32
0.75
1
1.5