Published in final edited form as:

Exp Hematol. 2016 September ; 44(9): 857–865.e5. doi:10.1016/j.exphem.2016.05.009.

Clonal hematopoiesis as determined by the *HUMARA* assay is a marker for acquired mutations in epigenetic regulators in older women

J Erin Wiedmeier^{a,†}, Catherine Kato^{a,†}, Zhenzhen Zhang^d, Hyunjung Lee^a, Jennifer Dunlap^{b,c}, Eric Nutt^c, Rogan Rattray^{b,c}, Sarah McKay^a, Christopher Eide^{a,e}, Richard Press^{b,c}, Motomi Mori^{b,d}, Brian Druker^{a,e}, and Kim-Hien Dao^a

^aKnight Cancer Institute, Hematology and Medical Oncology, Oregon Health & Science University, Portland, OR, USA., Oregon Health and Sciences University, 3181 SW Sam Jackson Park Rd, Portland, OR 97239

^bKnight Cancer Institute, Oregon Health & Science University, Portland, OR, USA., Oregon Health and Sciences University, 3181 SW Sam Jackson Park Rd, Portland, OR 97239

^cDepartment of Pathology, Oregon Health & Science University, Portland, OR, USA., Oregon Health and Sciences University, 3181 SW Sam Jackson Park Rd, Portland, OR 97239

^dSchool of Public Health, Oregon Health & Science University, Portland, OR, USA., Oregon Health and Sciences University, 3181 SW Sam Jackson Park Rd, Portland, OR 97239

^eHoward Hughes Medical Institute, Portland, OR, USA., Oregon Health and Sciences University, 3181 SW Sam Jackson Park Rd, Portland, OR 97239

Abstract

Recent large cohort studies revealed that healthy older individuals harbor somatic mutations that increase their risk for hematologic malignancy and all-cause cardiovascular deaths. The majority of these mutations are in chromatin and epigenetic regulatory genes (CERGs). CERGs play a key role in regulating DNA methylation (*DNMT3A* and *TET2*) or histone function (*ASXL1*) and in clonal proliferation of hematopoietic stem cells. We hypothesize that older women demonstrating

Corresponding Author: Kim-Hien Dao, DO, PhD, Oregon Health and Sciences University, 3181 SW Sam Jackson Park Rd, Mail Code: UHN 73C, Portland, OR 97239, 1-503-494-7894, daok@ohsu.edu.

[†]Equally contributed as co-first authors.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Author contributions:

Study concept and design: Wiedmeier, Dao, Kato, Dunlap.

Acquisition of data: Wiedmeier, Dao, Kato, Lee, Nutt.

Technical advice and instructions: Rattray, Eide, Press.

Statistical analysis: Zhang and Mori.

Other analysis and interpretation of data: Wiedmeier, Dao, Kato, Zhang, Dunlap, McKay.

Critical revision of the manuscript: All authors

Administrative, technical, or material support: All authors

Conflict of interest disclosure

No potential conflicts of interest were disclosed.

clonal hematopoiesis, defined here as a functional phenomenon in which a hematopoietic stem cell has acquired a survival and proliferative advantage, harbor a higher frequency of somatic mutations in CERGs. The human androgen receptor gene (*HUMARA*) assay was utilized in our study to detect the presence of nonrandom X-inactivation in women, a marker for clonal hematopoiesis. In our pilot study, we tested 127 blood samples from women 65 years old without a history of invasive cancer or hematologic malignancies. Applying stringent qualitative criteria, we found 26% displayed clonal hematopoiesis, 52.8% displayed polyclonal hematopoiesis, and 21.3% had indeterminate patterns (too close to call by qualitative assessment). Using Illumina MiSeq next generation sequencing, we identified somatic mutations in CERGs in 15.2% of subjects displaying clonal hematopoiesis (3 *ASXL1* and 2 *DNMT3A* mutations with an average variant allele frequency of 15.7%, range: 6.3%–23.3%). In a more limited sequencing analysis, we evaluated the frequency of *ASXL1* mutations by Sanger sequencing and found mutations in 9.7% of the clonal samples and 0% in the polyclonal samples. By comparing several recent studies (with some caveats as described), we determined the fold-enrichment of detecting CERG mutations by using the *HUMARA* assay as functional screen for clonal hematopoiesis. We conclude that a functional assay of clonal hematopoiesis enriches for older women with somatic mutations in CERGs, particularly for *ASXL1* and *TET2* mutations, and less so for *DNMT3A* mutations.

Graphical Abstract

Keywords

Clonal hematopoiesis; *HUMARA* assay; methylation; chromatin; hematologic malignancy

Introduction

Hematological malignancies, like most other cancers, occur after the accrual of multiple somatic mutations over the course of years.(1) Pre-malignant somatic mutations in the older population were long suspected as the driving force behind clonal hematopoiesis (CH) – a phenomenon that has been studied for decades by taking advantage of the fact that 1) random X-inactivation by methylation occurs during embryonic development and remains a permanent feature for the lifetime of the cell, and 2) certain genes on the X chromosome are highly polymorphic among individuals and the alleles can be resolved by molecular analysis

(e.g., *G6PD* and *HUMARA* genes).(2) A hematopoietic stem cell that acquires a survival and/or proliferative advantage is expected to dominate hematopoiesis. This leads to overrepresentation of active genes on one X-chromosome (non-random X-inactivation) in hematopoietic cells. While somatic mutations may underlie many cases of clonal hematopoiesis, another potential cause of clonal hematopoiesis is selective attrition or fitness of a hematopoietic stem cell conferred by the presence of polymorphic genes on the X chromosome(3) or other factors that have not been well studied.

Busque *et al.* first determined that clonal hematopoiesis evaluated by the *HUMARA* assay selectively enriches for women with mutations in *TET2*.(4) Of 182 older women displaying clonal hematopoiesis, *TET2* mutations occurred at a 5.5% frequency. The authors did not identify any mutations in older women without clonal hematopoiesis or in younger subjects with clonal hematopoiesis. Remarkably, one in seven subjects with clonal hematopoiesis and a *TET2* mutation developed a hematologic malignancy.(4) Several recent large cohort studies revealed that mutations in chromatin and epigenetic regulatory genes (CERGs) are not uncommon in the older population without any evidence of hematologic malignancy.(5–8) In two studies, the incidence of somatic mutations was determined by whole exome sequencing of blood samples from subjects of all ages from various disease cohorts without a cancer focus.(5, 6) The most common mutations occurred in *ASXL1*, *DNMT3A*, and *TET2*. Without considering the age of the subjects or the presence or absence of functional clonal hematopoiesis, the authors reported that mutations in the CERGs *ASXL1*, *DNMT3A*, *IDH1/2*, and *TET2* occurred at a frequency of 2.1% (257 out of 12,380)(5) and 3.1% (540 out of 17,182)(6) of all subjects tested, respectively. The authors mention that some mutation hotspots in *ASXL1* and *TET2* were not well covered. Using no mutation as the referent, the hazard ratio (95% confidence interval) for developing a hematologic malignancy with any somatic mutation in the Jaiswal *et al.* study and in the Genovese *et al.* study was 11.1 (3.9 to 32.6) and 12.9 (5.8 to 28.7), respectively. Additionally, all-cause cardiovascular deaths were significantly increased in the cohort of patients harboring mutations compared to age-matched controls adjusting for diabetes.(6) When considering a variant allele frequency of 10%, there was a much greater impact on the risk for hematologic malignancy and cardiovascular deaths.(6)

In animal models, a single gene defect in a CERG is sufficient to cause clonal expansion of the hematopoietic stem/progenitor compartment and development of myeloid malignancies. (9–13) However, in humans, these same mutations do not translate to a highly penetrant incidence of myeloid malignancies. In fact, the majority of older subjects who carry somatic mutations in CERGs do not develop myeloid malignancies in the time frame that these subjects were followed. This led to the introduction of the term clonal hematopoiesis of indeterminate potential (CHIP).(14) The next logical step in this field is to define the genetic and laboratory features that might identify those at the highest risk for hematologic malignancy. Because the incidence of hematologic malignancy in the general population over the age of 65 is less than 0.5%, the combined data from recent studies suggest that screening for clonal hematopoiesis by the *HUMARA* assay followed by next generation sequencing may help identify a population at the highest risk for hematologic malignancy. (4–8) Since the clone size (variant allele frequency) impacts risk for hematologic malignancy and cardiovascular disease, excessive allelic skewing (non-random X-

inactivation) may be an excellent functional marker for mutations driving this phenotype. Here we conclude that a functional screen for clonal hematopoiesis helped identify older women with somatic mutations in CERGs. Our initial goal was to confirm the frequency of clonal hematopoiesis and *TET2* mutations in older women without history of invasive cancer or hematologic malignancy.(4) We expanded on the work by Busque *et al.* by performing a more robust interrogation of a hematologic malignancy gene panel in terms of the number of genes evaluated and the depth of coverage for each gene. We sought to determine whether certain somatic mutations are more prominent in driving functional clonal hematopoiesis. While our work was underway, the two large cohort studies were published,(5, 6) allowing us to further compare the role of clonal hematopoiesis as defined by the *HUMARA* assay in enriching for older women with somatic mutations in CERGs.

Material and methods

Study population

Between March 2015 and April 2015, we obtained discarded blood samples from women aged 65 or older without a history of invasive cancer or known hematologic malignancy from Oregon Health and Science University's hospital and affiliated clinics. This study was approved by our Institutional Review Board. We reviewed the subjects' medical records to confirm their lack of a previous or current cancer diagnosis. Of the samples we collected, 127 samples were informative (met assay quality control and had heterozygous *HUMARA* alleles). Approximately 10% of samples displayed homozygous alleles (consistent with previous reports as reviewed in Busque and Gilliland(2)) and were excluded from further analysis.

Genomic DNA and the *HUMARA* assay

Genomic DNA (gDNA) was prepared from mononuclear cells using Qiagen's DNeasy Blood and Tissue Kit. Typically 0.5 µg to 1.0 µg of gDNA was digested with 20 units *HpaII* or a no enzyme control in separate tubes. Reactions were incubated at 37°C for 12 hours to overnight then terminated at 85°C for 20 minutes. Two microliters of digested DNA were added to a PCR reaction mixture containing a final concentration of 1X GoTaq Flexi buffer, 2mM MgCl₂, 10 mM of dNTPs, 0.2 µM of each primer, and 1.25 units of GoTaq DNA polymerase (Promega) in a 50 µL reaction. The primer sequences were: *HUMARA*-P1 5'-GCTGTGAAGGTTGCTGTTCCCTCAT and *HUMARA*-P2 5'-TCC AGAATCTGTTCCAGAGCGTGC (the primers were 5' labeled with DY782 and HPLC-purified). PCR reactions were subjected to 95°C for 5 minutes, then 32 cycles of 95°C for 45 seconds, 60°C for 30 seconds, 72°C for 30 seconds, and then held at 10°C indefinitely. PCR products were analyzed using the QIAxcel (Qiagen Corporation) instrument. The QX alignment marker of 15bp/1kb and the QX DNA size marker of 50bp–800bp were used for analyses. Results of each run were evaluated using the ScreenGel software (Qiagen Corporation).

Definition of clonal hematopoiesis

We applied qualitative criteria for identifying samples with clonal hematopoiesis. We required that one allele became the major peak while the other allele became a minor peak

when comparing the peaks in the *HpaII*-digested reaction relative to the control-digested reaction. Figure 1a illustrates a sample that was definitively clonal. A sample was polyclonal if the peaks of the PCR products for both *HUMARA* alleles were essentially identical in the *HpaII*-digested reaction compared to the control-digested reaction. Samples that did not clearly meet qualitative standards for either clonal or polyclonal were categorized as indeterminate. Because we noted that 21.3% of samples were categorized as indeterminate, we also evaluated these samples quantitatively (area under the curve (AUC) analysis) and compared the concordance rate with our qualitative interpretation. We used a previously published equation (Supplemental Table 1).(15) We also applied a different equation (Figure 2a) that evaluates each *HUMARA* allele separately and quantifies absolute change between the two alleles. This equation produced values that were normally distributed, which facilitated the analysis performed in Figure 2b. The allele frequency for A and B for any given sample occasionally exceeded 100% mathematically due to allele specific variations of the PCR amplification (technical or inherent). We defined a sum of greater than 135% as a sample that required a quality control check. An absolute AF difference between allele A versus allele B of 50% was considered to be a sample displaying clonal hematopoiesis (e.g. AF of 75% for allele A and 25% for allele B), whereas a sample with less than 50% absolute difference between alleles was considered to display polyclonal hematopoiesis, e.g., AF of 60% for allele A and 40% for allele B.

Next-generation sequencing

We sequenced samples displaying clonal hematopoiesis at the University of Pennsylvania Health System, Center for Personalized Diagnostics. A targeted hematologic malignancy panel interrogating 67 genes (Supplementary Table 2, *CEBPA* was not included) was selected for this study (average coverage 2,500X for entire panel). Oligonucleotide probe pairs were designed using Illumina design studio to capture hotspots and exons, where each oligonucleotide probe pair defines the content of one amplicon. Each oligonucleotide probe consists of complementary sequences to the target of interest and a tail of sequences that enables incorporation of sequencing adaptors. These oligonucleotide probes were manufactured and pooled together. Sequencing library preparations were performed following the Illumina TruSeq Custom Amplicon (TSCA) library preparation guide. Sequencing data was analyzed by an internal bioinformatics pipeline, which was created using both open source tools and custom algorithms (proprietary information was not provided). This pipeline was designed to extract essential performance statistics and to identify unique recurring mutations that eluded open source tools. During assay validation, consistently poor performing regions (18/1041) were removed from bioinformatics analysis (these 18 regions do not contain recurrent hotspots), and resulted in 1023 amplicons in 67 genes defining the regions ultimately analyzed in this study. The % gene coverage of *ASXL1*, *DNMT3A*, and *TET2* was 99%, 98%, and 99%, respectively, with greater than 250X coverage. All variants listed are with reference to the hg19 Genome build.

Targeted Sanger sequencing of specific mutations in *ASXL1*

Target DNA was amplified by PCR (primers are listed in Supplemental Table 3). A 50 μ L reaction mixture contained a final concentration of 5 μ L of 10X PCR buffer II, 1 μ L of each primer stock at 10 μ M, one unit of AccuPrime *Taq*, and 30 ng of gDNA. PCR reactions were

performed in a thermocycler according to the following cycling protocol: 94°C for 20 seconds, then 35–42 cycles of 94°C for 20 seconds, 57°C for 20 seconds, 68°C for 25 seconds; and then held at 10°C indefinitely. The PCR product was purified then combined with a sequencing primer for sequencing.

Statistical analysis

Categorical variables were compared using Fischer's exact test due to the relatively small sample size in certain categories. For continuous variables, we conducted Shapiro Wilks tests to check the normality. For variables with normal distribution, we used either a t-test or Analysis of Variance (ANOVA) test. All statistical analyses were performed using SAS 9.4 (SAS Institute Inc., Cary, NC, USA).

Results

Clonal hematopoiesis is prevalent in women over 65 years old

We evaluated blood samples from women without a history of invasive cancer or known hematologic malignancy. A total of 127 samples met assay quality control and displayed heterozygous *HUMARA* alleles. The median age of the subjects was 74 years old (range: 66–100 years old). Based on our stringent qualitative criteria for defining clonality (see methods section), 26% were clonal, 52.8% were polyclonal, and 21.3% were indeterminate. Indeterminate samples did not fit perfectly in the definitively clonal or polyclonal categories or the changes were too-close-to-call by qualitative criteria. If we excluded the indeterminate samples (no call was made qualitatively), 33% were clonal and 67% were polyclonal. The subjects' ages according to the various categories are shown in Figure 1b. There was no significant difference in subjects' ages between the three sample categories ($p = 0.44$ by ANOVA test). Age was not a discriminating factor for increased frequency of clonal hematopoiesis in women over 65 years old.

Next generation sequencing reveals frequent mutations in *ASXL1* and *DNMT3A* in samples displaying clonal hematopoiesis

Next-generation sequencing was performed on 33 samples displaying clonal hematopoiesis (median age 72, range 66–92 years) using the Illumina MiSeq technology. A 67-gene panel (Supplemental Table 2) was evaluated with an average of 2,500X coverage to identify mutations that occur at a frequency $\geq 5\%$. Somatic mutations occurred in 24.2% of samples displaying clonal hematopoiesis (average AF: 16.5%, range: 6.3%–27.2%). Most of these mutations occurred in CERGs (Table 1). If we only consider mutations in CERGs, the rate of mutation was 15.2% in samples displaying clonal hematopoiesis (average AF: 15.7%, range: 6.3%–23.3%). Specifically, we found three *ASXL1* and two *DNMT3A* mutations. We confirmed the *ASXL1* mutations but did not confirm the *DNMT3A* mutations because the variant allele frequencies of these mutations were less than 10% (the lower limit of detection of Sanger sequencing). Additional mutated genes included *TP53*, *C-MYC*, and *ZRSR2*. The mean ages of subjects displaying clonal hematopoiesis without mutations versus with mutations were 73.7 years and 77.1 years, respectively. We did not identify any *TET2* mutations and had expected 1 or 2 mutations based on the study by Busque *et al.*(4)

Some samples had two peaks (doublets) representing one *HUMARA* allele. We attribute the double peaks to *Taq* polymerase slippage, an experimental artifact not uncommon in areas of short repeats (the *HUMARA* gene contains CAG repeats).(16) There could also be multiple alignments of the primer in a CAG-rich region. We note that the extra peaks for one allele occur predominately with the larger DNA fragment, suggesting that the longer the length the more likely an allele is represented by multiple peaks. When multiple peaks represent one *HUMARA* allele, this made both qualitative and quantitative analyses more complicated. Five of the samples we initially defined as clonal qualitatively were categorized as polyclonal based on quantitative analysis using the equation previously published (Supplemental Table 1). These samples were not found to have mutations. Therefore, the frequency of 1somatic mutations in samples displaying clonal hematopoiesis by a *HUMARA* screen may be even higher. If the five polyclonal samples are removed, the total number of clonal samples sent for sequencing is 28 with 28.6% of samples displaying somatic mutations in any of the genes in the panel tested and 17.9% of samples displaying somatic mutations in CERGs. The concordance rate between qualitative and quantitative methodologies for determining sample category was overall very good: it was 85% (28/33) for clonal samples and 99% (66/67) for polyclonal samples, using the qualitative determination as the denominator. When we re-evaluated the indeterminate samples (those that by our estimation had <50% absolute allele frequency difference but greater than 10% absolute allele frequency difference), the vast majority of these samples (85% or 23/27) did not meet the criteria defined for clonal hematopoiesis. The four samples that met criteria for clonal hematopoiesis had Psup values very close to 0.25 or 0.75 (Supplemental Table 1).

Degree of *HUMARA* allelic skewing is more pronounced in samples with somatic mutations than in samples without somatic mutations

We then investigated whether the degree of allelic skewing or clonal hematopoiesis evaluated by the *HUMARA* assay is different in samples that harbored mutations versus samples that do not harbor mutations in the gene panel tested. This was evaluated by a different quantitative AUC analyses but we used the same threshold for calling clonal hematopoiesis, i.e., representation of one *HUMARA* allele by more than 75%, please refer to methods section. The equation and schematic for performing this analysis is illustrated in Figure 2a. When we compared the absolute AF difference, samples with clonal hematopoiesis but without mutations (n=25) display a lower degree of allelic skewing compared to samples with clonal hematopoiesis and somatic mutations (n=8). The mean absolute AF difference was 64.7% versus 97.4%, respectively (Figure 2b; $p = 0.035$). Across sample categories (clonal, polyclonal, indeterminate), there was a significant difference in the absolute allele frequency difference of samples displaying clonal hematopoiesis (mean 72.6%) versus samples categorized as indeterminate (mean 28.7%) and polyclonal (mean 14.1%, $p < 0.0001$). The difference between the groups and the mean % absolute AF difference was expected based on the definition of clonal hematopoiesis as defined.

Clonal hematopoiesis determination by the *HUMARA* assay enriches for samples with somatic mutations in *ASXL1*

To verify the mutations identified in the next-generation sequencing and to determine the degree of enrichment using the *HUMARA* assay as an initial screen for mutations in

ASXL1, we performed Sanger sequencing of select mutations in *ASXL1* (Supplemental Table 3 and Table in Figure 3). The mutations we evaluated *ASXL1*-S547, E635, G645/6 represent approximately 80% of all *ASXL1* mutations in acute myeloid leukemia.¹⁸ Among samples with sufficient gDNA material available for Sanger re-sequencing analysis, we evaluated 31 clonal hematopoiesis samples and randomly chose 34 polyclonal hematopoiesis samples. We confirmed the *ASXL1* mutations in sample #12, 37, and 132 (Figure 3a) and found no additional mutations (n=31). The sequencing chromatograms are shown in Figure 3b. We found no *ASXL1* mutations in the polyclonal samples (n=34). The difference between *ASXL1* mutations in the clonal versus polyclonal samples was not significantly different due to the relatively small sample size ($p = 0.103$). Nonetheless, the rate of mutation in *ASXL1* was 9.7% in clonal samples and 0% in polyclonal samples, and there was a trend towards a difference. These findings are similar to the Busque *et al.* study evaluating *TET2* mutations in clonal and polyclonal samples as determined by the *HUMARA* assay.

Discussion

Our goal was to identify the predominant somatic mutations leading to functional clonal hematopoiesis as represented by non-random X-inactivation. We modeled our study in a manner similar to the Busque *et al.* study.⁽⁴⁾ In this study, the authors used the *HUMARA* assay to identify older women with mutations in *TET2* or in four other genes tested. We expanded on this work by interrogating a 67-gene panel using Illumina MiSeq next-generation sequencing on those samples that displayed clonal hematopoiesis. The Busque *et al.* study, and our follow-up study to expand the number of genes interrogated, for the first time identify the spectrum of mutations associated with clonal hematopoiesis as evident by non-random X-inactivation. Although we did not find any *TET2* mutations in our study, we expected only 1 or 2 cases based on the sample size and previous estimates.⁽⁴⁾ Our study was performed on mononuclear cells from whole blood. We considered the possibility that the use of mononuclear cells may have preferentially evaluated the somatic mutations occurring predominately in lymphocytes and monocytes. In a pilot study, we performed the *HUMARA* assay on mononuclear cells and whole blood cells after RBC lysis from the same subjects and found no difference in the clonal or polyclonal hematopoiesis patterns (Supplemental Figure 2). Therefore, it appears that clonal hematopoiesis determination can be performed in any of these cell fractions, which would be consistent with the premise that these early somatic mutations occur in a multipotent, hematopoietic stem cell.

We found a clonal hematopoiesis frequency of 33% in older women using the *HUMARA* assay and applying relatively stringent qualitative criteria. This frequency is within range of previous reports.^(17–19) However, using quantitative methods (Supplemental Table 1), the frequency of clonal hematopoiesis was slightly lower. We believe that quantitative analysis may occasionally mask quality control issues, incorporate high background tracings, or miscall or misalign AUC peak values with alleles represented by multiple peaks (especially for the suspected clonal samples), thus, a review of each tracing is still an essential part of data analysis. We note very high concordance rates for the qualitative and both quantitative methods for samples that were determined indeterminate or polyclonal (Supplemental Table 1). The clonal hematopoiesis rates could also be slightly lower due to the use of

mononuclear cells as well. Nevertheless, clonal hematopoiesis defined in our study identifies older women with a high frequency of harboring somatic mutations in CERGs (15.2%). Additionally, we report a higher degree of allelic skewing in subjects with clonal hematopoiesis and somatic mutations versus subjects with clonal hematopoiesis but without somatic mutations in the gene panel we tested (Figure 2b). When we supplemented our data with targeted Sanger sequencing analysis to confirm the mutations found by next generation sequencing and to evaluate the polyclonal samples without undue costs, we found that screening for clonal hematopoiesis in older women by the *HUMARA* assay enriches for those with mutations in *ASXL1* compared to those without clonal hematopoiesis (9.7% versus 0%, Table in Figure 3). We limited our sequencing to specific mutation hotspots as described in Figure 3 for this analysis. Nonetheless, using one method to compare the role of screening by the *HUMARA* assay (Sanger sequencing studies performed here), we show that clonal hematopoiesis defined by the *HUMARA* assay enriches for the identification of somatic mutations in CERGs. The authors of the Busque *et al.* study arrived to the same conclusion for their evaluation of the *TET2* gene.(4)

Mutations in *ASXL1*, *DNMT3A*, and *TET2* are not uncommon in older people with evidence of hematologic malignancy.(5–8) *ASXL1*, a histone function modulator and transcriptional regulator, is frequently mutated in myeloid malignancies. Loss of *ASXL1* in mice leads to the development of myelodysplastic syndrome (MDS)-like disease.(9, 20) The mechanism of its function in hematopoiesis is still unknown. Abnormal DNA methylation occurs in many types of cancers. *DNMT3A*, a DNA methyltransferase involved in *de novo* DNA methylation, is frequently mutated in hematologic malignancies.(21) Hematopoietic stem cells harboring *DNMT3A* mutations gain a self-renewal advantage and clonally expand over time.(22, 23) *TET2* is a methylcytosine dioxygenase and converts methylated cytosines to hydroxymethylcytosines, a step necessary for DNA demethylation. *TET2* loss-of-function mutations are present in 25–40% of all blood cancers. *TET2* deficiency in murine models leads to an accelerated rate of clonal evolution and myelodysplasia development.(20, 24) The fact that mutations in CERGs are present in pre-malignant states(5–8) and occur in more than 50% of all myeloid malignancies suggests an important role for these mutations early in disease pathogenesis. Acknowledging that different methods were employed and is a limitation of comparing various recent studies, we list the mutation rates of *ASXL1*, *DNMT3A*, and *TET2* in Table 2 in a highly-selected population (women 65 years old with clonal hematopoiesis by the *HUMARA* assay) versus an unselected population (all age groups in preexisting cohorts, men and women, and no functional clonal hematopoiesis screen). The fold-enrichment using a highly-selected population versus an unselected population is particularly high for *ASXL1* and *TET2* mutations (29.3-fold, and 16.6-fold, respectively), indicating that clonal expansion may be a more prominent feature with these mutations and less prominent with *DNMT3A* mutations (3.6-fold). Under-evaluation of *ASXL1* and *TET2* mutations in the two largest studies as reported is unlikely to change this overall conclusion.(5, 6)

Our study supports the conclusion that clonal hematopoiesis defined by a functional assay (non-random X-inactivation) is a marker for subjects carrying CERG mutations. There may be additional mutation(s) that modify or drive clonal hematopoiesis and specific somatic mutations (gene and/or type of mutation) may confer varying degrees of clonal

hematopoiesis. The median age of patients diagnosed with myelodysplasia and acute myeloid leukemia is reported as 71 years old and 69 years old, respectively. We therefore speculate that sequential somatic mutations typically occur over years before overt cancer develops, if it ever develops. This is based on the frequency of clonal hematopoiesis and somatic mutations, using age as a risk factor for these events. It is possible that by the time clonal hematopoiesis is detectable by a functional assay, there are already multiple somatic mutations in genes that drive this phenomenon in humans, but this remains to be studied on a whole genome/exome scale with deep coverage. Identifying the somatic mutation(s) that manifest as clonal hematopoiesis will provide biological insights into the risks for cancer progression. The predictive value for a future hematologic malignancy using a functional clonal hematopoiesis test (e.g. *HUMARA*) combined with next-generation sequencing remains to be determined. Screening by this approach will also advance our understanding of the natural history of clonal hematopoiesis and somatic mutations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported by grants sponsored by Collins Medical Trust (K.T.D.), OHSU Knight Cancer Institute (K.T.D.), and the NIH/NHLBI (1K08HL111280) (K.T.D.). Biostatistics support is provided by the Biostatistics Shared Resource of the Knight Cancer Institute (Cancer Center Support Grant, NCI 5P30CA069533) (M.M.). J.E.W was supported by a Howard Hughes Medical Institute fellowship, and B.J.D is supported by the Howard Hughes Medical Institute.

References

1. Nowell PC. The clonal evolution of tumor cell populations. *Science*. 1976 Oct 1; 194(4260):23–28. [PubMed: 959840]
2. Busque L, Gilliland DG. X-inactivation analysis in the 1990s: promise and potential problems. *Leukemia*. 1998 Feb; 12(2):128–135. [PubMed: 9519772]
3. Abkowitz JL, Taboada M, Shelton GH, Catlin SN, Gutter P, Kiklevich JV. An X chromosome gene regulates hematopoietic stem cell kinetics. *Proc Natl Acad Sci U S A*. 1998 Mar 31; 95(7):3862–3866. [PubMed: 9520458]
4. Busque L, Patel JP, Figueroa ME, Vasanthakumar A, Provost S, Hamilou Z, et al. Recurrent somatic TET2 mutations in normal elderly individuals with clonal hematopoiesis. *Nat Genet*. 2012 Nov; 44(11):1179–1181. [PubMed: 23001125]
5. Genovese G, Kahler AK, Handsaker RE, Lindberg J, Rose SA, Bakhoum SF, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med*. 2014 Dec 25; 371(26):2477–2487. [PubMed: 25426838]
6. Jaiswal S, Fontanillas P, Flannick J, Manning A, Grauman PV, Mar BG, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med*. 2014 Dec 25; 371(26):2488–2498. [PubMed: 25426837]
7. Xie M, Lu C, Wang J, McLellan MD, Johnson KJ, Wendl MC, et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat Med*. 2014 Dec; 20(12):1472–1478. [PubMed: 25326804]
8. McKerrill T, Park N, Moreno T, Grove CS, Ponstingl H, Stephens J, et al. Leukemia-associated somatic mutations drive distinct patterns of age-related clonal hemopoiesis. *Cell reports*. 2015 Mar 3; 10(8):1239–1245. [PubMed: 25732814]
9. Wang J, Li Z, He Y, Pan F, Chen S, Rhodes S, et al. Loss of *Asx1* leads to myelodysplastic syndrome-like disease in mice. *Blood*. 2014 Jan 23; 123(4):541–553. [PubMed: 24255920]

10. Chen E, Schneider RK, Breyfogle LJ, Rosen EA, Poveromo L, Elf S, et al. Distinct effects of concomitant Jak2V617F expression and Tet2 loss in mice promote disease progression in myeloproliferative neoplasms. 2015 Jan 8; 125(2):327–335.
11. De Vita S, Schneider RK, Garcia M, Wood J, Gavillet M, Ebert BL, et al. Loss of function of TET2 cooperates with constitutively active KIT in murine and human models of mastocytosis. *PLoS One*. 2014; 9(5):e96209. [PubMed: 24788138]
12. Li Z, Cai X, Cai CL, Wang J, Zhang W, Petersen BE, et al. Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood*. 2011 Oct 27; 118(17):4509–4518. [PubMed: 21803851]
13. Moran-Crusio K, Reavie L, Shih A, Abdel-Wahab O, Ndiaye-Lobry D, Lobry C, et al. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell*. 2011 Jul 12; 20(1):11–24. [PubMed: 21723200]
14. Steensma DP, Bejar R, Jaiswal S, Lindsley RC, Sekeres MA, Hasserjian RP, et al. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood*. 2015 Jul 2; 126(1):9–16. [PubMed: 25931582]
15. Bolduc V, Chagnon P, Provost S, Dube MP, Belisle C, Gingras M, et al. No evidence that skewing of X chromosome inactivation patterns is transmitted to offspring in humans. *J Clin Invest*. 2008 Jan; 118(1):333–341. [PubMed: 18097474]
16. Shinde D, Lai Y, Sun F, Arnheim N. Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT)_n and (A/T)_n microsatellites. *Nucleic Acids Res*. 2003 Feb 1; 31(3):974–980. [PubMed: 12560493]
17. Tonon L, Bergamaschi G, Dellavecchia C, Rosti V, Lucotti C, Malabarba L, et al. Unbalanced X-chromosome inactivation in haemopoietic cells from normal women. *Br J Haematol*. 1998 Sep; 102(4):996–1003. [PubMed: 9734650]
18. Busque L, Paquette Y, Provost S, Roy DC, Levine RL, Mollica L, et al. Skewing of X-inactivation ratios in blood cells of aging women is confirmed by independent methodologies. *Blood*. 2009 Apr 9; 113(15):3472–3474. [PubMed: 19202126]
19. Swierczek SI, Agarwal N, Nussenzweig RH, Rothstein G, Wilson A, Artz A, et al. Hematopoiesis is not clonal in healthy elderly women. *Blood*. 2008 Oct 15; 112(8):3186–3193. [PubMed: 18641369]
20. Abdel-Wahab O, Gao J, Adli M, Dey A, Trimarchi T, Chung YR, et al. Deletion of Asxl1 results in myelodysplasia and severe developmental defects in vivo. *J Exp Med*. 2013 Nov 18; 210(12):2641–2659. [PubMed: 24218140]
21. Yang L, Rau R, Goodell MA. DNMT3A in haematological malignancies. *Nat Rev Cancer*. 2015 Mar; 15(3):152–165. [PubMed: 25693834]
22. Challen GA, Sun D, Jeong M, Luo M, Jelinek J, Berg JS, et al. Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat Genet*. 2012 Jan; 44(1):23–31. [PubMed: 22138693]
23. Mayle A, Yang L, Rodriguez B, Zhou T, Chang E, Curry CV, et al. Dnmt3a loss predisposes murine hematopoietic stem cells to malignant transformation. *Blood*. 2015 Jan 22; 125(4):629–638. [PubMed: 25416277]
24. Gelsi-Boyer V, Brecqueville M, Devillier R, Murati A, Mozziconacci MJ, Birnbaum D. Mutations in ASXL1 are associated with poor prognosis across the spectrum of malignant myeloid diseases. *Journal of hematology & oncology*. 2012; 5:12. [PubMed: 22436456]
25. Pratorcorona M, Abbas S, Sanders MA, Koenders JE, Kavelaars FG, Erpelinck-Verschueren CA, et al. Acquired mutations in ASXL1 in acute myeloid leukemia: prevalence and prognostic value. *Haematologica*. 2012 Mar; 97(3):388–392. [PubMed: 22058207]

Highlights

- Older women with non-random X-inactivation are more likely to harbor mutations in blood cells
- The majority of mutations occur in chromatin and epigenetic regulatory genes
- Non-random X-inactivation significantly enriches for *ASXL1* and *TET2* mutations

Figure 1. A clonal hematopoiesis sample by qualitative assessment and subjects evaluated according to age and sample category

A, A sample displaying a clonal pattern evaluated qualitatively. Blue curves: PCR amplification after *HpaII* digestion; red curves: PCR amplification after no enzyme control digestion. A sample was labeled definitively clonal when one allele (a blue curve) has a massive expansion (more than 200% of AUC of the reference red curve), while the other allele (the other blue curve) diminished substantially (less than 50% of AUC of the reference red curve). The red curves provide a “fingerprint” of each *HUMARA* allele regardless of methylation status for each subject and is specific for any given subject. B, Scatter plot showing the age distribution and median age of subjects (horizontal bar) displaying clonal, indeterminate, and polyclonal hematopoiesis as determined by the *HUMARA* assay. We included a category for the ages of all the subjects as well.

Equation:

$$AF_A = (HUMARA-A_H \cdot HUMARA-A_C) \div 0.5 \div 100$$

$$AF_B = (HUMARA-B_H \cdot HUMARA-B_C) \div 0.5 \div 100$$

$$\text{abs AF difference} = |AF_A| - |AF_B|$$

$$\Sigma AF = |AF_A| + |AF_B|$$

Sample calculation:

$$AF_A = (120 \cdot 60) \div 0.5 \div 100 = 100\%$$

$$AF_B = (15 \cdot 90) \div 0.5 \div 100 = 8.3\%$$

$$\text{abs AF difference} = |100\%| - |8.3\%| = 91.7\%$$

$$\Sigma AF = |100\%| + |8.3\%| = 108.3\%$$

Figure 2. Area under the curve (AUC) or quantitative analysis of allele frequency (AF) and evaluating degree of allelic skewing of CH samples with and without mutations

A, Schematic of AUC and allele frequency determination. The equation shows how each allele is evaluated independently based on the control (no enzyme control digestion) PCR products. There are two *HUMARA* alleles in women, A and B. Red, control digest (C); Blue, *HpaII* digest (H); AUC, area under curve; AF, allele frequency, and Σ , sum of. Parallel bars indicate absolute value in the equation. See methods section for additional details. B, Scatter plot of absolute AF difference for CH samples with and without somatic mutations, p value = 0.035.

A.

<i>ASXL1</i> mutations	Clonal hematopoiesis, n = 31	Polyclonal hematopoiesis, n = 34
Frequency	3 (9.7%)	0
Mutation	S547fs, S547fs, G646W/fs	-
Total	3 (9.7%)	0

B.

Figure 3. Targeted Sanger sequencing of mutation hotspots in *ASXL1*
A, *ASXL1* mutations identified by Sanger sequencing. This was done in part to confirm the *ASXL1* mutations found in the next generation sequencing. Mutations with a variant allele frequency below 10% as reported in the next generation sequencing were not confirmed by Sanger sequencing (below the lower limit of detection).**ASXL1* codons covered include S547, E635, G645/6, which accounts for approximately 80% of *ASXL1* mutations in AML. (25) †Nonsense mutaXon; fs, frameshiY mutaXon. B, Chromatograms of mutations identified by Sanger sequencing. The bar indicates the codon specifically evaluated.

Table 1

Mutations found by next generation sequencing of 33 samples displaying clonal hematopoiesis

Sample Number	Age of Subject	Gene	Protein Change	Mutation Type	Variant Allele Frequency	Expected Outcome
132	92	<i>ASXL1</i>	G646W/fs	Frame Shift	23.33	Pathogenic
12	68	<i>ASXL1</i>	S547fs	Frame Shift	22.78	Pathogenic
37	68	<i>ASXL1</i>	S547fs	Frame Shift	16.57	Pathogenic
92	75	<i>C-Myc</i>	S388R	Missense	27.21	VOUS
151	81	<i>DNMT3A</i>	W313 [†]	Missense	9.5	Pathogenic
177	87	<i>DNMT3A</i>	W753C	Missense	6.28	VOUS
37	68	<i>TP53</i>	P153H	Missense	11.45	VOUS
156	69	<i>ZRSR2</i>	H265Y	Missense	14.93	VOUS

VOUS, Variants of Unknown Significance;

[†] Stop gained.

Table 2

Comparing mutation rates across recent studies in a highly-selected versus unselected population as defined

	ASXL1			DNMT3A			TET2		
	highly-selected population ^a	unselected population ^b	highly-selected population ^a	unselected population ^b	highly-selected population ^a	unselected population ^b	highly-selected population ^a	unselected population ^b	highly-selected population ^a
This study	9.09%	ND	6.06%	ND	0%	ND	0%	ND	ND
Busque <i>et al.</i>	ND	ND	0%	ND	5.49%	ND	5.49%	ND	ND
Genovese <i>et al.</i> ^c	ND	0.35%	ND	1.57%	ND	0.25%	ND	0.25%	0.42%
Jaiswal <i>et al.</i> ^c	ND	0.36%	ND	2.35%	ND	0.42%	ND	0.42%	ND
McKerrell <i>et al.</i> ^d	ND	ND	ND	1.16%	ND	ND	ND	ND	0.33%
Xie <i>et al.</i> ^e	ND	0.22%	ND	0.66%	ND	0.33%	ND	0.33%	ND
Average	NA	0.31%	NA	1.70 %	NA	0.33%	NA	0.33%	16.6X
Fold-enrichment in a highly-selected versus unselected population	NA	29.3X	NA	3.6X	NA	16.6X	NA	16.6X	16.6X

^a highly selected population: women 65 years old displaying clonal hematopoiesis as determined by the *HUMARA* assay.

^b unselected population: subjects of all ages, men and women, and no functional screen for clonal hematopoiesis.

^c whole exome sequencing with incomplete coverage of *ASXL1* and *TET2* genes as described in each study.

^d genomic DNA sequencing using Hi-Plex PCR –MPS (massively parallel sequencing). For the *DNMT3A* gene, only *DNMT3A*-R882 mutations were evaluated as described in study.

^e whole exome sequencing, The Cancer Genome Atlas.

ND=not determined; NA=not applicable.

Please refer to the discussion section for citation of the full references and the original manuscripts for additional details on the respective methods.