



Published in final edited form as:

*J Hum Genet.* 2016 May ; 61(5): 411–418. doi:10.1038/jhg.2015.166.

## Mosaic 13q14 deletions in peripheral leukocytes of non-hematologic cancer cases and healthy controls

Mitchell J. Machiela<sup>1</sup>, Weiyin Zhou<sup>1,2</sup>, Neil Caporaso<sup>1</sup>, Michael Dean<sup>1,3</sup>, Susan M. Gapstur<sup>4</sup>, Lynn Goldin<sup>1</sup>, Victoria L. Stevens<sup>4</sup>, Meredith Yeager<sup>1,2</sup>, and Stephen J. Chanock<sup>1</sup>

<sup>1</sup>Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, Maryland, USA

<sup>2</sup>Cancer Genomics Research Laboratory, Leidos Biomedical Research, Inc, Frederick, Maryland, USA

<sup>3</sup>Laboratory of Experimental Immunology, Center for Cancer Research, Frederick National Laboratory, Frederick, Maryland, USA

<sup>4</sup>Epidemiology Research Program, American Cancer Society, Atlanta, GA 30303, USA

### Abstract

Loss of 13q14.3 is a chromosomal event found in approximately 50 percent of B-cell chronic lymphocytic leukemia (CLL) and monoclonal B-cell lymphocytosis (MBL) cases. Surveys of somatic alterations in solid tumors have shown sporadic 13q14.3 loss in many different tumor types, but not at high frequency in any specific tumor type. In our recent survey of the single nucleotide polymorphism (SNP) microarray data from 127,000 cancer free or solid tumor cases, we observed mosaic 13q14.3 loss as a common autosomal somatic large structural events (>2 Mb in size) in blood and buccal-derived DNA. Herein, we examined this region more closely investigating structural mosaic events <2 Mb using SNP microarray data in 46,254 non-hematologic cancer cases and 36,229 controls. We detected 60 individuals with 13q14.3 mosaic loss, one mosaic copy neutral uniparental disomy, and 13 individuals with homozygosity. While 13q14.3 loss size was variable, the minimally deleted region (MDR) (chr13:49,590,000-49,983,100; GRCh36) was comparable to what is classically reported in MBL and CLL. Breakpoint analysis of the estimated boundaries reveals enrichment for genes and open chromatin. The frequency of 13q14.3 loss significantly increases with increasing age (P-value=0.028), but was not significantly different between non-hematological cancer cases and controls (0.084% versus 0.058%; P-value=0.19). These findings suggest mosaic 13q14.3 losses accumulate with age. Individuals with detected mosaic 13q14.3 deletions may be early, undetected cases of MBL or CLL, but not necessarily all will develop MBL and CLL.

---

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: [http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

**CORRESPONDING AUTHOR:** Stephen J. Chanock, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 9609 Medical Center Drive, Gaithersburg, MD 20852, Office: (301) 435-7559, Fax: (301) 402-3134, chanocks@mail.nih.gov.

### CONFLICT OF INTEREST DISCLOSURE

The authors report no relevant conflicts of interest.

## INTRODUCTION

Hemizygous and homozygous deletions on the long arm of chromosome 13 are the most common genetic aberrations in B-cell chronic lymphocytic leukemia (CLL) and monoclonal B-cell lymphocytosis (MBL), occurring in approximately 50 percent of cases<sup>1-4</sup>. The reported deletions are usually large and heterogeneous in size and typically include a minimally deleted region (MDR) on 13q14.3 (~130 kilobases)<sup>3,5-7</sup>. The MDR is telomeric to the retinoblastoma gene (*RB1*) and includes *DLEU1* and *DLEU2* as well as two notable micro-RNAs, *MIR15A* and *MIR16-1*. Loss at 13q14 as a single large abnormality can indicate a more favorable CLL clinical course<sup>2</sup>; however, a higher percentage of cells with 13q14 loss and increased size of 13q14 deletions have been associated with shorter time-to-treatment<sup>6-11</sup>. Deletions at 13q14 are postulated to result in loss of tumor-suppressor activity, but the variable size of the 13q14 deletion suggests more than one tumor-suppressor gene is likely present. Further work has shown that down-regulation of *MIR15A* and *MIR16-1* in CLL coupled with a CLL like phenotype in mouse models deficient for the *DLEU2/MIR-15A/MIR16-1* cluster indicate *MIR15A* and *MIR16-1* may be important regulators in CLL pathogenesis<sup>12-14</sup>.

Somatic alterations of 13q14 have been reported in solid tumors, suggesting a possible role in the carcinogenesis of select non-hematological malignancies. Approximately 6% of retinoblastoma cases have a 13q14.3 deletion of the *RB1* gene<sup>15,16</sup>. Sporadic observation of 13q14 events has been reported in other solid tumors, but not with the consistency observed in CLL or retinoblastoma. For example, pan-cancer analyses of The Cancer Genome Atlas (TCGA) data show evidence for 13q14 loss in bladder, breast, colon, glioblastoma, head/neck, kidney, lung, ovarian, and endometrial tumors<sup>17</sup>. Allelic loss at 13q14 has been reported in one third of prostate tumors<sup>18,19</sup>. Other studies have reported 13q14 loss with high prostate tumor grade and stage<sup>20</sup> as well as increased proliferation and invasiveness of untransformed prostate cells after down regulation of *MIR-15a* and *MIR16-1*<sup>21</sup>. Loss of heterozygosity (LOH) data from breast cancer studies described a region including 13q12.2-14.3 can be commonly deleted in breast cancer<sup>22</sup>. Moreover, it has been suggested that 13q14 allelic imbalance may be associated with breast cancer mortality more than 5 years after diagnosis<sup>23</sup>, but further studies are needed to confirm this finding. Additional association evidence indicates that ADP-ribosylation factor-like tumor suppressor gene 1 (*ARLTS1*, also called ADP-ribosylation factor-like 11 *ARL11*) located on 13q14.3 approximately 400 Kb centromeric to *MIR15a* and *MIR16-1* may function as a tumor suppressor gene for lung cancer<sup>24</sup>, familial breast cancer<sup>25,26</sup>, melanoma<sup>27</sup>, ovarian cancer<sup>28</sup>, prostate cancer<sup>29</sup>, and colorectal cancer<sup>30</sup>. Overall, the many reports of 13q14 deletions in solid cancers point towards a possible role as a contributing 'driver' event, but further studies are needed to confirm the reported frequencies and more importantly, establish the functional implications of 13q14 deletion beyond CLL and MBL.

Genetic mosaicism is the coexistence of clonal cellular populations harboring two or more distinct genotypes in an individual<sup>31,32</sup>. In previous studies of detectable clonal mosaicism, an increased frequency of large-structural events detected in a fraction of circulating cells was noted in pre-diagnostic samples of individuals who later developed a lymphoid malignancy<sup>33,34</sup>. One of the most common sites for large structural deletions of more than 2

Mb in size included at least 300 Kb of the 13q14.3 MDR region<sup>33,35</sup>, the region commonly deleted in CLL. In this context, it is notable that mosaic 13q14.3 loss was observed in a fraction of DNA samples collected from individuals diagnosed with solid tumors or who were cancer free<sup>33,35</sup>. Since the screening detection algorithm for SNP microarrays conducted as part of cancer GWAS is stable for events greater than 2 Mb in size, the aim of this analysis was to re-examine our large GWAS data set of more than 80,000 individuals to identify additional, smaller 13q14.3 events as well. In addition, we wanted to determine if there is a possible relationship between 13q14.3 events in blood DNA and risk for solid tumor.

## MATERIALS AND METHODS

### Study Population

The Division of Cancer Epidemiology and Genetics and the Cancer Genome Research (CGR) Laboratory in the NCI have conducted a series of genome-wide association studies with commercial SNP microarrays (Illumina Hap300, Hap240, Hap550, Hap610, Hap660, Hap 1, Omni Express, Omni 1, Omni 2.5, and Omni 5). The study set included 82,483 participants (46,254 non-hematological cancer cases and 36,229 cancer-free controls) with blood or buccal DNA, previously known as Total GWAS Set (TGS) I and II, actually scanned in the CGR<sup>35</sup>. No cases of retinoblastoma were included in either TGS I or II. The institutional review board of the participating study centers and the NIH approved the study protocols and informed consent was received for each study participant.

### 13q14.3 Copy Number Alteration Detection

B-allele frequency (BAF) and log R ratio (LRR) are two metrics used to assess 13q14.3 mosaic loss in our study population. BAF is a measure of allelic imbalance used to determine stretches of heterozygous SNPs that deviate from the expected value of 0.5. LRR provides information on copy number in which LRR values greater than 0 are evidence for allelic gains, LRR values less than 0 indicate allelic loss, and LRR values not deviating from 0 indicate copy neutral events. BAF and LRR values were estimated using methods previously described<sup>36</sup> and normalized and corrected using a framework from our previous work<sup>33</sup>. Based on copy number status, mean heterozygous BAF bands were used to calculate the percentage of cells in the cellular population that contained copy number alterations at 13q14.3. This copy number alteration detection method has been previously validated with additional laboratory techniques for autosomes (e.g., single tandem repeat, multiplex ligation-dependent probe amplification, and fluorescent in situ hybridization)<sup>37</sup> and is described in greater detail in the methods and supplementary material of the original analysis<sup>33</sup>. Stretches of homozygosity were detected as genomic stretches across 13q14.3 with no BAF band for heterozygous SNPs.

Two segmentation algorithms were used to scan genotypes of the non-hematological cancer population for mosaic alterations at the 13q14.3 locus: a LRR-based approach and a BAF-based approach. Both approaches scanned the sample population and surveyed a window around the CLL MDR region on chromosome 13q14.3 between bases 49,139,793 and 50,269,706 (GRCh36). The LRR-based approach isolated individuals with evidence of

13q14.3 loss by finding individuals with MDR windows that had mean LRR values significantly less than mean LRR values on chromosome 1 (used as the reference) by applying t-tests with unequal variance and a P-value  $<1.0 \times 10^{-10}$  as the significance threshold. A mean difference of  $-0.05$  was used as a filter to ensure statistically significant 13q14.3 LRR values displayed meaningful differences from the chromosome 1 reference LRR. The BAF-based approach detected individuals with allelic imbalances at 13q14.3 by selecting individuals with MDR windows that had heterozygous  $B_{dev}$  means ( $B_{dev} = |BAF - 0.5|$  for heterozygous BAF bands) that significantly differed from  $B_{dev}$  means on chromosome 1 (P-value  $<1.0 \times 10^{-10}$ ). A mean  $B_{dev}$  difference of 0.005 was required in statistically significant regions to ensure meaningful differences in  $B_{dev}$  were detected. Results from the two approaches were merged and manually reviewed to confirm a copy number alteration at the 13q14.3 MDR. Copy number status was assigned and the percentage of cells with altered 13q14.3 nuclei was estimated from heterozygous BAF bands. For individuals with 13q14.3 events that extended beyond the investigated chromosome 13 window, boundaries were adjusted to include the entire 13q event.

### Breakpoint Analysis

Mosaic event breakpoint regions were mapped to within 100 Kb from the best estimate of 13q14.3 deletion start and end points using the SNP intensity data. The 100 Kb window was chosen as a conservative estimate of the true breakpoint since start and stop coordinates for events were defined based on the nearest array tagging SNPs; an estimate which is affected by SNP density and the quality of LRR and BAF data around these SNPs. The 60 detected 13q14.3 mosaic deletions resulted in a total of 120 breakpoint regions, roughly 200 Kb in size.

1,000 permutations of 120 randomly selected 200 Kb regions on chromosome 13 were selected to estimate the underlying distribution of chromosome 13 structural elements flanking the 13q14.3 region. To generate these breakpoints, 60 random events were selected from a Gaussian distribution that reflected observed event length ( $\mu = 10.80$  Mb,  $\sigma = 16.90$  Mb) and were constrained to overlap the 393 Kb minimally deleted region. The first and last Illumina microarray SNPs within simulated events were used to define the estimated breakpoints, and 100 Kb windows were calculated to model the underlying distribution of observed event breakpoints.

University of California Santa Cruz (UCSC) and Encyclopedia of DNA Elements (ENCODE) data tracks were downloaded from the UCSC FTP data portal (<ftp://hgdownload.cse.ucsc.edu/>) to investigate local DNA characteristics in breakpoint regions as well as randomly selected regions. Features of interest included gene-rich regions (RefSeq genes, CpG Islands), indicators of open chromatin (ORChID, DNaseI HS Peaks, FAIRE-Seq Peaks), recombination rate (deCODE sex-averaged), and repetitive elements (SINEs, LINEs, LTRs, and Segmental Duplications). Comparisons between observed breakpoints and random breakpoints were made using mean differences and mean counts of feature elements. Statistical significance was assessed by permutation P-value.

## Statistical Analyses

Univariate analyses were performed characterizing the relationship of age group (<50, 50–54, 55–59, 60–64, 65–69, 70–74, and 75 years), gender (male, female), estimated ancestry (%European, %African, and %Asian ancestry estimated by SNP genotypes), and non-hematologic cancer (disease-free controls, solid tumor cases overall and substrata of cases by solid tumor type) with 13q14.3 loss. An additional multivariable analysis that included age group, gender, estimated ancestry, and non-hematologic cancer as well as adjustment for contributing study (indicator variables) were carried out using logistic regression. Strata specific analyses were performed within strata of non-hematological cancer type to investigate whether 13q14.3 deletion was associated with specific non-hematological cancers.

To investigate differences in frequency between 13q14.3 alterations in our population and CLL prevalence in the US population, age and sex-specific limited duration prevalence measures from the November 2011 submission of Surveillance Epidemiology and End Results<sup>38</sup> data were used based on the US Census data. All statistical analyses were performed in R version 2.15.1 (R Foundation for Statistical Computing).

## RESULTS

In a combined analysis of 82,483 non-hematologic cancer cases and disease free controls DNA drawn from peripheral leukocytes (81%) and buccal samples (19%), we detected 60 individuals (0.073%, 95% CI=0.054–0.091%) with 13q14.3 mosaic loss between 19.7 and 90.1%. The distribution of events across strata of gender, age group, ancestry, and non-hematologic cancer status is displayed in Table 1. Non-hematological cancer cases and controls in our analysis had a frequency of 13q14.3 mosaic loss of 0.084 (95% CI=0.066–0.115) and 0.069 (95% CI= 0.038–0.089) percent, respectively. The majority of the detected 13q14.3 mosaic losses were in DNA derived from blood, however, 3 mosaic 13q14.3 deletions were observed in buccal DNA. In addition to the 60 individuals detected with mosaic loss at 13q14.3, our analysis detected one lung cancer case with a copy neutral uniparental disomy that spanned the 13q14.3 region. Stretches of homozygosity within the 13q14.3 region were also observed in 13 individuals, which included 6 controls, 2 endometrial cancer cases, 3 lung cancer cases, 1 ovary cancer case, and 1 pancreatic cancer case. Individuals with mosaic 13q14.3 deletions showed no preference for developing additional mosaic autosomal events in other chromosomal locations as compared to others with non-13q14.3 mosaic events ( $p=0.17$ ).

Detected 13q14.3 mosaic losses are plotted in Figure 1 and vary in size and estimated breakpoint locations. The median mosaic loss was 1.9 Mb (interquartile range: 1.1–17.1 Mb) in size with the smallest mosaic loss 246 Kb and the largest greater than 75 Mb in size. Overall, smaller mosaic 13q14 events are more common than larger events (Figure 2A), which parallels what has been seen for large structural mosaic events across the autosomes<sup>35</sup>. Common breakpoint regions on the centromeric side included clusters at 29.1–32.0 Mb, 40.1–43.7 Mb, and 49.3–49.6 Mb; while breakpoints on the telomeric side primarily clustered around 50.3–52.9 Mb (GRCh36). At a minimum, mosaic 13q14.3 losses detected in peripheral blood or buccal DNA cover a region on chromosome 13 spanning an

approximately 393 Kb region from base pairs 49,590,000 to 49,983,100 (GRCh36). This region includes *DLEU1* and *DLEU2*, the *ST13P4* pseudogene, and is approximately 70 Kb telomeric to the location of *MIR15A* and *MIR16-1*. The variable size of the detected 13q14.3 mosaic losses resulted in deletions of other known tumor suppressor genes. When we examined the mapping of the deletions, we observed inclusion of *BRCA2* in 10% of individuals, *RBI* in 40%, *MIR15A* in 87%, *MIR16-1* in 87%, *DLEU1* in 95%, *DLEU2* in 98%, and *DLEU7* in 97%.

In an analysis of average LRR values, an indicator of allelic copy number, the majority of individuals with 13q14.3 mosaic losses affect only one allele (Figure 2A). There was, however, one control individual with evidence for mosaic loss of both alleles at the 13q14.3 locus (LRR=-1.58). Overall, the observed mosaic proportion, a measurement of the percentage of cells containing 13q14 loss, was between 19.7 to 90.1 percent, with a mean of 41.4 percent (Figure 2A). As expected, we observed an association between LRR and mosaic proportion in our data, indicating that despite the small size of many 13q14.3 mosaic losses, we were able to calculate highly correlated measures of mosaic percentage ( $p=8.16\times 10^{-9}$ ,  $\text{adj } R^2=0.52$ ) (Figure 2B). No significant associations were observed between the size of 13q14.3 deletion and either LRR or mosaic percentage ( $p=0.097$  and  $0.211$ , respectively).

Analysis of the 13q14.3 deletion breakpoints using select bioinformatic data tracks from the UCSC browser and ENCODE indicates substantial clustering of active elements in the region of 13q14.3 loss breakpoints. In comparison to a random sampling of similar-sized regions on other autosomes, we observed a statistically significant enrichment for regions rich in genes and regulatory elements (Figure 3). For instance, RefSeq genes and CpG islands were significantly enriched at 13q14.3 breakpoints ( $p<0.002$ , Figure 3A), indicating 13q14.3 breakpoint regions often occur near genes or promoters of genes. Indicators of open chromatin, OH Radical Cleavage Intensity Database (ORChID), DNaseI hypersensitivity peaks (DNase HS), and Formaldehyde-Assisted Isolation of Regulatory Elements peaks (FAIRE-Seq), were also significantly enriched ( $p<0.002$ , Figure 3B). There was no evidence indicating an association between recombination rate and location of 13q14.3 mosaic breakpoints ( $p=0.39$ , Figure 3C). The distribution of repetitive elements such as short interspersed nuclear elements (SINE), long interspersed nuclear elements (LINE), long terminal repeats (LTR), and segmental duplications suggests clustering in 13q14.3 breakpoint regions (Figure 3D). In particular, SINEs are significantly enriched for ( $p<0.002$ ), while LTRs are less common in these regions ( $p=0.002$ ). A marginally significant enrichment for segmental duplications was also observed in 13q14.3 breakpoint regions ( $p=0.062$ ).

Unadjusted associations with 13q14.3 mosaic deletions were investigated (Table 1). Mosaic loss of 13q14 was slightly more common in males compared to females (0.09 versus 0.05, respectively,  $p=0.051$ ). An unadjusted positive association was also observed between mosaic 13q14.3 loss and 5-year age group (<50, 50-54, 55-59, 60-64, 65-69, 70-74, 75+,  $p=1.8\times 10^{-3}$ ), ranging from 0.02% for individuals less than 50 and increases up to a frequency of 0.18% for individuals 75 and older. An unadjusted association with ancestry was also detected (P-value=0.0004) with mosaic loss of 13q14.3 most commonly observed

in individuals of European ancestry, but less commonly in individuals of African ancestry. Table 1 further illustrates the breakdown of 13q14.3 mosaic events by solid tumor subtype. Overall, no significant difference in 13q14.3 loss was observed between non-hematological cancers (0.08%) and controls (0.06%) (P-value=0.19).

In exploratory multivariable models including study, gender, 5-year age group, ancestry, and non-hematologic cancer status, a positive association was only observed for age (P-value=0.028). The increased odds for a 13q14.3 mosaic deletion for 75 years or older compared to those less than age 50 years was 8.70 (95% CI=1.65–45.78). In the multivariable analysis, there were no associations of overall solid tumor (P-value=0.89) or specific tumor subtype with 13q14.3 mosaic loss.

To investigate whether detection of 13q14 deletion could be detecting early, undiagnosed cases of CLL, age and gender specific frequencies of 13q14.3 mosaic loss in our dataset were compared to SEER population-based CLL prevalence data (Table 2). Assuming 50 percent of all CLL cases have 13q14.3 loss, expected age and gender stratified counts of 13q14.3 loss were estimated for our sample set based on SEER limited-duration prevalence data (0 to <34 years) and US census data. Results indicate that although on average slightly more instances of 13q14.3 mosaic loss are observed among non-hematologic cancer cases and controls, this does not significantly depart from the 13q14.3 losses attributable to undiagnosed CLL that would be expected in a population of this size, age, and gender distribution ( $p=0.46$ ).

## DISCUSSION

Our analysis of 13q14.3 mosaic losses in DNA isolated from blood or buccal cell samples revealed an overall frequency of 0.073 percent (95% CI=0.056–0.094) based on detection of 13q14.3 mosaicism in 60 individuals out of 82,483 who were included in non-hematologic cancer GWAS. No significant difference in frequency of 13q14.3 mosaic loss was detected between non-hematologic cancer cases and controls. We show that the prevalence of mosaic 13q14.3 loss increases with age, a phenomenon observed for all large mosaic events<sup>35,39</sup>. The size of 13q14.3 mosaic losses can vary, but almost always included a minimally deleted region on chromosome 13, from 49,590,000 to 49,983,100 (GRCh36). In more than 85% of instances, mosaic losses included the *MIR15A*, *MIR16-1*, *DLEU1*, *DLEU2*, and *DLEU7* loci. We did not identify a clearly defined set of breakpoints, suggesting that events may not be mapped to specific basepairs but instead to regions of open chromatin.

An enrichment of genes, promoter sites, and enhancers around 13q14.3 breakpoints indicates cellular mechanisms related to transcription and gene expression may be important in breakpoint formation. Regions of open chromatin around enhancers and gene rich regions may also expose more of the DNA backbone to environmental mutagens and could result in higher probabilities for mutational events. The enrichment of SINEs in 13q14.3 breakpoint regions, as well as a marginally significant enrichment of segmental duplications, suggest genomic repeat regions could play a role in initiation of events leading to 13q14.3 mosaic losses. It is less clear what functional role these repetitive elements could contribute to the initiation of somatic 13q14.3 loss, but transposon activity and mismatch repair may be two

important mechanisms to investigate further. While these findings do not conclusively demonstrate a specific cellular mechanism responsible for the detected 13q14.3 mosaic losses, it does suggest transcription coupled repair, exposure to DNA mutagens, and transposon activity may be important mechanisms capable of initiating DNA breaks leading to mosaic 13q14.3 deletions.

In an unadjusted analysis, we observed 13q14.3 mosaic loss were associated with age (5-year age group), continental ancestry, and endometrial cancer, but with small numbers these latter observations are most likely a false positives. In fact, only the 5-year age group association was observed in the adjusted multivariable analysis. The association between mosaicism and age has been observed in previous studies reporting on overall autosomal mosaicism<sup>33–35,39</sup>. Notably, 13q14 deletion events are among the most common large structural somatic events observed in detectable mosaicism. The increasing frequency of MBL and CLL with age<sup>38,40,41</sup>, makes it challenging to delineate whether detected 13q14.3 mosaic losses are early biomarkers for potential MBL and CLL risk or if such deletions are sentinel events that frequently appear when genomic maintenance capacities begin to deteriorate with age. Additionally, many of the contributing cancer genome-wide association studies did not screen for MBL/CLL, thus, there is a substantive possibility that some undiagnosed cases may have existed in our dataset.

Our study does not provide sufficient evidence that mosaic 13q14.3 deletions in blood or buccal DNA are significantly associated with solid tumors, despite the literature reports of sporadic observations of 13q14.3 deletion events in a spectrum of cancers. Publically available TCGA data on the cBioPortal suggests 13q14.3 deletions are present in up to 11% of prostate tumors, 7% of bladder tumors, 4% of endometrial tumors, and 3% of ovary and colorectal tumors<sup>42</sup>. We surmise that many of these may be passengers and not necessarily drivers of the solid tumors. However, it is important to note that our study on DNA obtained from blood and buccal cells does not provide information about the frequency of 13q14.3 loss in other tissues, which may be a more relevant prognostic feature for tumors that develop in those tissues. In addition, even in our large survey of cancer GWAS, we had limited power to investigate the association between 13q14.3 mosaic loss and solid tumors.

Our study raises an important question regarding the implications of harboring a mosaic 13q14.3 deletion. The most likely possibility is that individuals with mosaic 13q14.3 deletions are instances of early undetected MBL/CLL. The commonly deleted region in our analysis is nearly identical to that seen in MBL/CLL and the frequency estimate we observe for 13q14.3 deletions (0.07%) is statistically indistinguishable from the expected age and sex prevalence of CLL, assuming 50 percent of CLL cases have a 13q14.3 deletion. Additionally, deletions of 13q14.3 as the sole abnormality are generally associated with indolent clinical disease<sup>2</sup> and often appear early in early-stage CLL<sup>43</sup>. Alternatively, it is plausible that these 13q14.3 deletions are somatic changes in B or T cell populations that serve as sentinel events that frequently appear when genomic maintenance capacities begin to deteriorate with age. The last, and least likely possibility, is that these are inherited events passed from parents to offspring through the germline. The range of allelic proportions for heterozygous SNPs spanning these events (20 to 90%) substantially deviates from the 50% expectation, making this interpretation very unlikely. Further large studies are required to



capture a sufficiently large enough set of mosaic 13q14.3 deletions in a prospective manner so as to accurately assess cancer risk.

## Acknowledgments

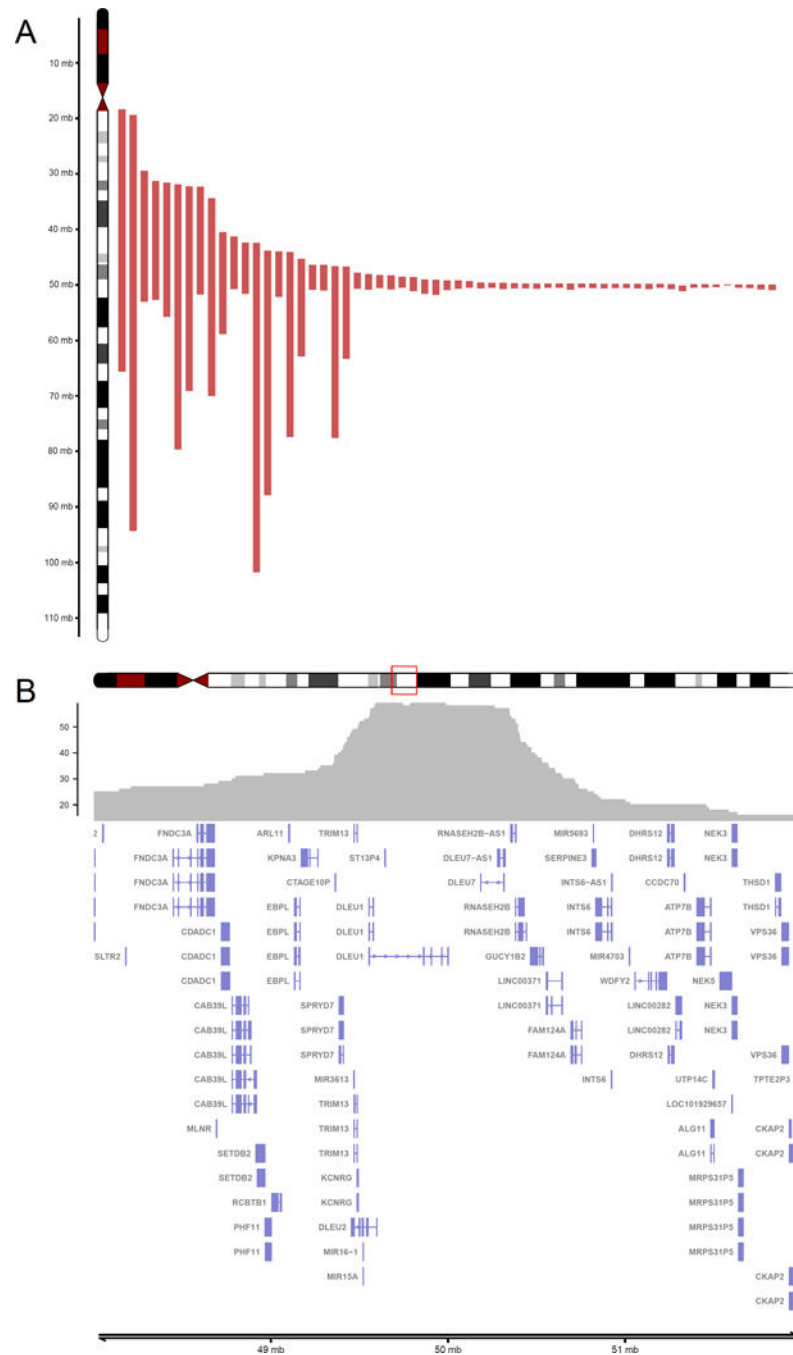
This study was funded by the Intramural Research Program of the National Cancer Institute's Division of Cancer Epidemiology and the American Cancer Society.

## References

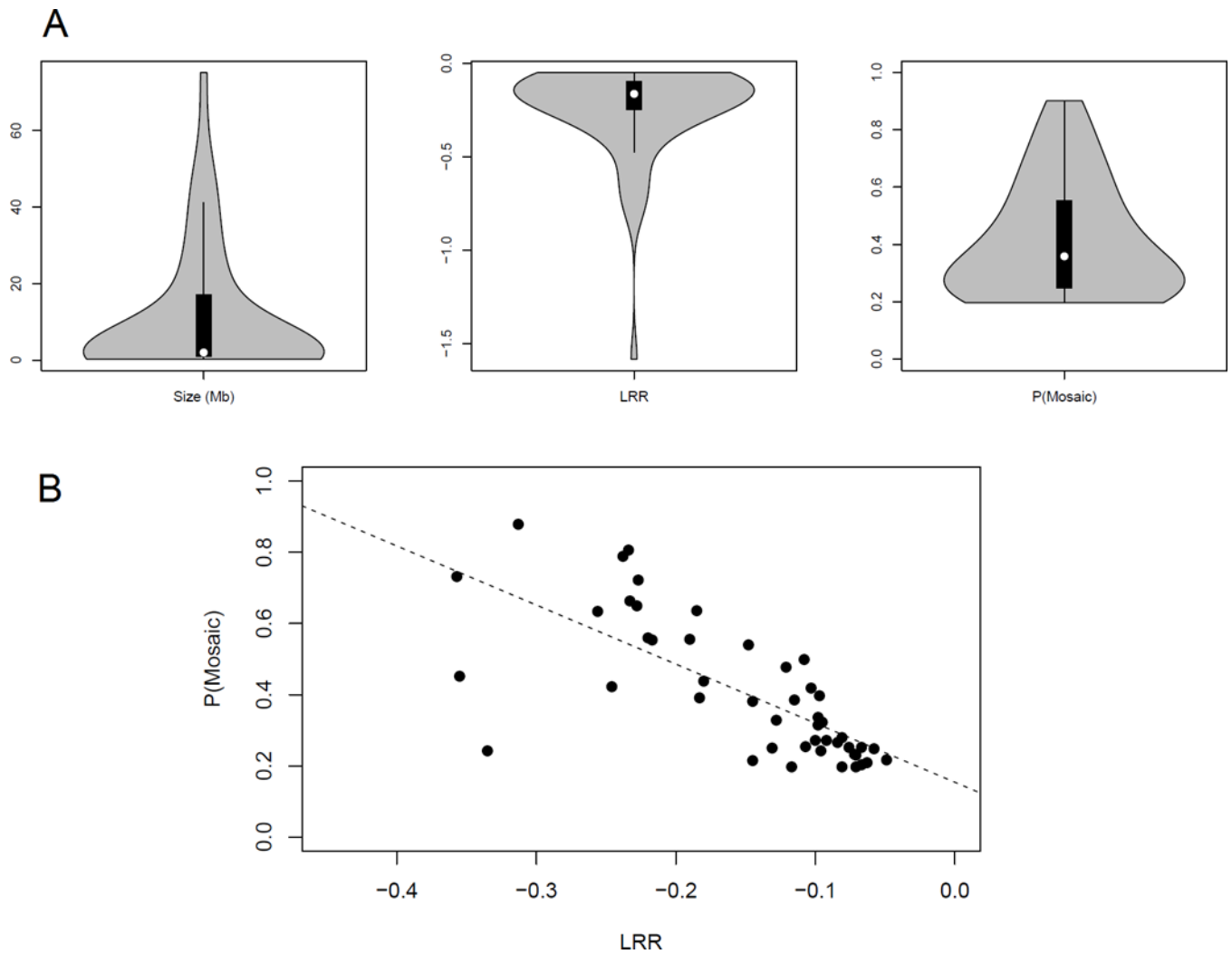
1. Stilgenbauer S, Nickolenko J, Wilhelm J, Wolf S, Weitz S, Döhner K, et al. Expressed sequences as candidates for a novel tumor suppressor gene at band 13q14 in B-cell chronic lymphocytic leukemia and mantle cell lymphoma. *Oncogene*. 1998; 16:1891–1897. [PubMed: 9583687]
2. Döhner H, Stilgenbauer S, Benner A, Leupolt E, Kröber A, Bullinger L, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med*. 2000; 343:1910–1916. [PubMed: 11136261]
3. Edelmann J, Holzmann K, Miller F, Winkler D, Bühler A, Zenz T, et al. High-resolution genomic profiling of chronic lymphocytic leukemia reveals new recurrent genomic alterations. *Blood*. 2012;4783–4794. [PubMed: 23047824]
4. Rawstron A, Bennett F, O'Connor S, Kwok M, Fenton J. Monoclonal B-cell lymphocytosis and chronic lymphocytic leukemia. *N Engl J Med*. 2008;575–583. [PubMed: 18687638]
5. Mosca L, Fabris S, Lionetti M, Todoerti K. Integrative genomics analyses reveal molecularly distinct subgroups of B-cell chronic lymphocytic leukemia patients with 13q14 deletion. *Clin Cancer Res*. 2010; 16:5641–5653. [PubMed: 20947517]
6. Ouillette P, Erba H, Kujawski L, Kaminski M, Shedden K, Malek SN. Integrated genomic profiling of chronic lymphocytic leukemia identifies subtypes of deletion 13q14. *Cancer Res*. 2008; 68:1012–1021. [PubMed: 18281475]
7. Ouillette P, Fossum S, Parkin B, Ding L, Bockenstedt P, Al-Zoubi A, et al. Aggressive chronic lymphocytic leukemia with elevated genomic complexity is associated with multiple gene defects in the response to DNA double-strand breaks. *Clin Cancer Res*. 2010; 16:835–847. [PubMed: 20086003]
8. Hernández JA, Rodríguez AE, González M, Benito R, Fontanillo C, Sandoval V, et al. A high number of losses in 13q14 chromosome band is associated with a worse outcome and biological differences in patients with B-cell chronic lymphoid leukemia. *Haematologica*. 2009; 94:364–371. [PubMed: 19252174]
9. Van Dyke DL, Shanafelt TD, Call TG, Zent CS, Smoley Sa, Rabe KG, et al. A comprehensive evaluation of the prognostic significance of 13q deletions in patients with B-chronic lymphocytic leukaemia. *Br J Haematol*. 2010; 148:544–550. [PubMed: 19895615]
10. Bo MD, Rossi FM, Rossi D, Deambrogi C, Bertoni F, Del Giudice I, et al. 13q14 Deletion Size and Number of Deleted Cells Both Influence Prognosis in Chronic Lymphocytic Leukemia. *Genes Chromosomes Cancer*. 2011; 643:633–643.
11. Mian M, Rinaldi A, Mensah A, Rossi D, Ladetto M, Forconi F, et al. Del(13q14.3) length matters: an integrated analysis of genomic, fluorescence in situ hybridization and clinical data in 169 chronic lymphocytic leukaemia patients with 13q deletion alone or a normal karyotype. *Hematol Oncol*. 2012; 53:46–49. [PubMed: 21671253]
12. Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, Noch E, et al. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A*. 2002; 99:15524–15529. [PubMed: 12434020]
13. Klein U, Lia M, Crespo M, Siegel R, Shen Q, Mo T, et al. The DLEU2/miR-15a/16-1 cluster controls B cell proliferation and its deletion leads to chronic lymphocytic leukemia. *Cancer Cell*. 2010; 17:28–40. [PubMed: 20060366]
14. Lia M, Carette A, Tang H, Shen Q, Mo T, Bhagat G, et al. Functional dissection of the chromosome 13q14 tumor-suppressor locus using transgenic mouse lines. *Blood*. 2012; 119:2981–2990. [PubMed: 22174151]

15. Houdayer C, Gauthier-Villars M, Lauge A, Pages-Berhouet S, Dehainault C, Caux-Moncoutier V, et al. Comprehensive Screening for Constitutional RB1 Mutations by DHPLC and QMPSP. *Hum Mutat.* 2004; 23:193–202. [PubMed: 14722923]
16. Albrecht P, Ansperger-Rescher B, Schuler A, Zeschnigk M, Gallie B, Lohmann DR. Spectrum of Gross Deletions and Insertions in the RB1 Gene in Patients With Retinoblastoma and Association With Phenotypic Expression. *Hum Mutat.* 2005; 26:437–445. [PubMed: 16127685]
17. Zack TI, Schumacher SE, Carter SL, Cherniack AD, Saksena G, Tabak B, et al. Pan-cancer patterns of somatic copy number alteration. *Nat Genet.* 2013; 45:1134–1140. [PubMed: 24071852]
18. Yin Z, Spitz MR, Babaian RJ, Strom SS, Troncoso P, Kagan J. Limiting the location of a putative human prostate cancer tumor suppressor gene at chromosome 13q14.3. *Oncogene.* 1999; 18:7576–7583. [PubMed: 10602517]
19. Cooney KA, Wetzel JC, Merajver SD, Cooney A, Macoska JA, Singleton P. Distinct Regions of Allelic Loss on 13q in Prostate Cancer. *Cancer Res.* 1996:1142–1145. [PubMed: 8640774]
20. Dong JT, Boyd JC, Frierson HF. Loss of heterozygosity at 13q14 and 13q21 in high grade, high stage prostate cancer. *Prostate.* 2001; 49:166–171. [PubMed: 11746261]
21. Bonci D, Coppola V, Musumeci M, Addario A, Giuffrida R, Memeo L, et al. The miR-15a-miR-16-1 cluster controls prostate cancer by targeting multiple oncogenic activities. *Nat Med.* 2008; 14:1271–1277. [PubMed: 18931683]
22. Osborne R, Hamshire M. A genome-wide map showing common regions of loss of heterozygosity/allelic imbalance in breast cancer. *Cancer Res.* 2000:3706–3712. [PubMed: 10919637]
23. Voegtly LM, Mamula K, Campbell JL, Shriver CD, Ellsworth RE. Molecular alterations associated with breast cancer mortality. *PLoS One.* 2012; 7:e46814. [PubMed: 23056464]
24. Yendamuri S, Trapasso F, Ferracin M, Cesari R, Seignani C, Shimizu M, et al. Tumor suppressor functions of ARLTS1 in lung cancers. *Cancer Res.* 2007; 67:7738–7745. [PubMed: 17699778]
25. Akisik E, Yazici H, Dalay N. ARLTS1, MDM2 and RAD51 gene variations are associated with familial breast cancer. *Mol Biol Rep.* 2011; 38:343–348. [PubMed: 20358297]
26. Frank B, Hemminki K, Meindl A, Wappenschmidt B, Klaes R, Schmutzler RK, et al. Association of the ARLTS1 Cys148Arg variant with familial breast cancer risk. *Int J Cancer.* 2006; 118:2505–2508. [PubMed: 16353159]
27. Frank B, Meyer P, Boettger MB, Hemminki K, Stapelmann H, Gast A, et al. ARLTS1 variants and melanoma risk. *Int J Cancer.* 2006; 119:1736–1737. [PubMed: 16646072]
28. Petrocca F, Iliopoulos D, Qin HR, Nicoloso MS, Yendamuri S, Wojcik SE, et al. Alterations of the tumor suppressor gene ARLTS1 in ovarian cancer. *Cancer Res.* 2006; 66:10287–10291. [PubMed: 17079447]
29. Siltanen S, Wahlfors T, Schindler M, Saramäki OR, Mpindi JP, Latonen L, et al. Contribution of ARLTS1 Cys148Arg (T442C) variant with prostate cancer risk and ARLTS1 function in prostate cancer cells. *PLoS One.* 2011; 6:e26595. [PubMed: 22028916]
30. Castellví-Bel S, Castells A, de Cid R, Muñoz J, Balaguer F, Gonzalo V, et al. Association of the ARLTS1 Cys148Arg variant with sporadic and familial colorectal cancer. *Carcinogenesis.* 2007; 28:1687–1691. [PubMed: 17449901]
31. Strachan, T.; Read, A. *Human Molecular Genetics.* Wiley-Liss; 1999.
32. Machiela MJ, Chanock SJ. Detectable clonal mosaicism in the human genome. *Semin Hematol.* 2013; 50:348–359. [PubMed: 24246702]
33. Jacobs KB, Yeager M, Zhou W, Wacholder S, Wang Z, Rodriguez-Santiago B, et al. Detectable clonal mosaicism and its relationship to aging and cancer. *Nat Genet.* 2012; 44:651–658. [PubMed: 22561519]
34. Laurie CC, Laurie CA, Rice K, Doheny KF, Zelnick LR, McHugh CP, et al. Detectable clonal mosaicism from birth to old age and its relationship to cancer. *Nat Genet.* 2012; 44:642–650. [PubMed: 22561516]
35. Machiela MJ, Zhou W, Sampson JN, Dean MC, Jacobs KB, Black A, et al. Characterization of Large Structural Genetic Mosaicism in Human Autosomes. *Am J Hum Genet.* 2015; 96:487–497. [PubMed: 25748358]
36. Wang K, Bucan M. Copy Number Variation Detection via High-Density SNP Genotyping. *CSH Protoc.* 2008 pdb.top46 (2008).

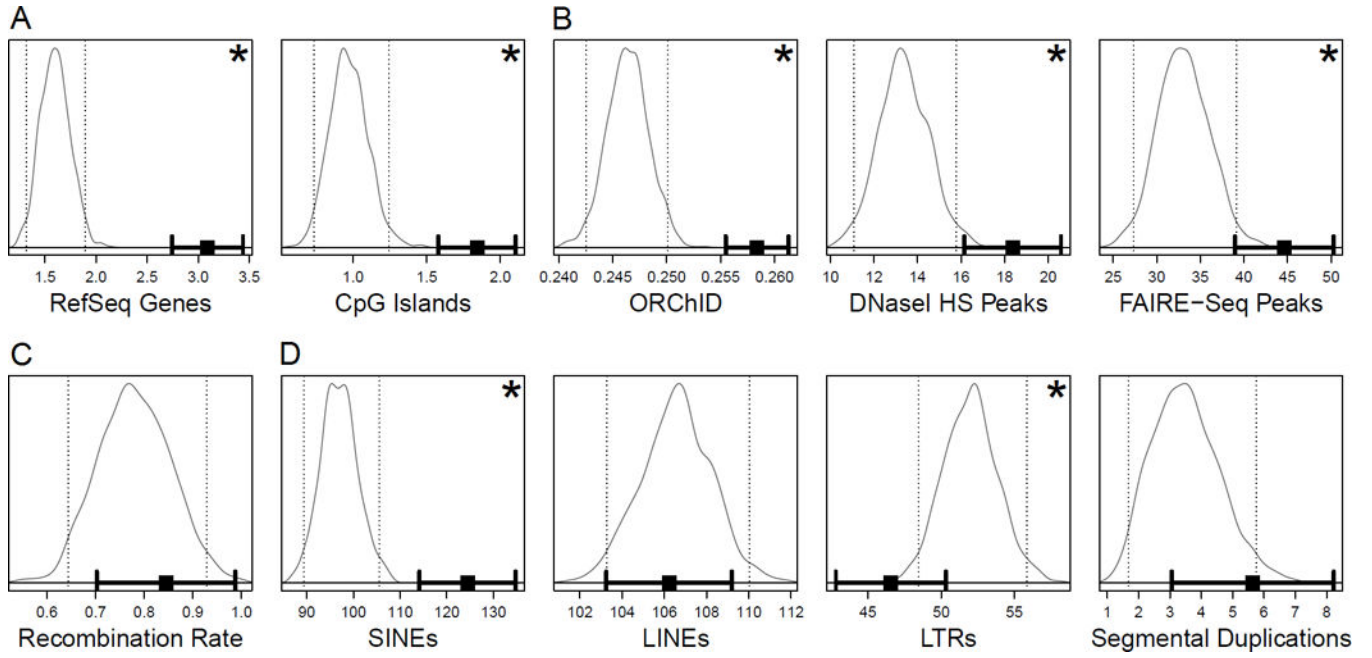
37. Rodríguez-Santiago B, Malats N, Rothman N, Armengol L, Garcia-Closas M, Kogevinas M, et al. Mosaic uniparental disomies and aneuploidies as large structural variants of the human genome. *Am J Hum Genet.* 2010; 87:129–138. [PubMed: 20598279]
38. Surveillance Research Program. Fast Stats: An interactive tool for access to SEER cancer statistics. National Cancer Institute at. <<http://seer.cancer.gov/faststats>>
39. Dumanski J, Rasi C, Lönn M, Davies H. Smoking is associated with mosaic loss of chromosome Y. *Science (80-)*. 2015; 347:81–83.
40. Goldin LR, Lanasa MC, Slager SL, Cerhan JR, Vachon CM, Strom SS, et al. Common occurrence of monoclonal B-cell lymphocytosis among members of high-risk CLL families. *Br J Haematol.* 2010; 151:152–158. [PubMed: 20738309]
41. Scarfò L, Dagklis A, Scielzo C, Fazi C, Ghia P. CLL-like monoclonal B-cell lymphocytosis: are we all bound to have it? *Semin Cancer Biol.* 2010; 20:384–390. [PubMed: 20816789]
42. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2012; 2:401–404. [PubMed: 22588877]
43. Shanafelt TD, Witzig TE, Fink SR, Jenkins RB, Paternoster SF, Smoley Sa, et al. Prospective evaluation of clonal evolution during long-term follow-up of patients with untreated early-stage chronic lymphocytic leukemia. *J Clin Oncol.* 2006; 24:4634–4641. [PubMed: 17008705]



**FIGURE 1.** Distribution of 13q14.3 losses. (A) Graphical illustration of the 60 detected mosaic losses that span the minimally deleted region at 13q14.3. (B) Zoom in of the 13q14.3 minimally deleted region showing overlap of detected events (top panel) and genes in the region (bottom panel).



**FIGURE 2.** Event size, log R ratio (LRR), and proportion of mosaic cells (P(Mosaic)) for detected 13q14.3 losses. (A) Violin plots of event size, LRR, and P(Mosaic) distributions. (B) Relationship between LRR and P(Mosaic) for monoallelic 13q14.3 losses.



**FIGURE 3.** Breakpoint analysis of 13q14 breakpoint regions in comparison to breakpoints from random permutations of similarly-sized regions spanning the 13q14.3 minimally deleted region. Features investigated include (A) gene-rich regions, (B) indicators of open chromatin, (C) recombination rate, and (D) repetitive elements. Gray distributions are means across 1,000 permutations. Black boxes and error bars represent the mean across detected 13q14.3 breakpoints and 95% confidence interval around the mean. Asterisks indicate  $P < 0.05$ .

TABLE 1

Counts and percentages for detected 13q14.3 mosaic loss across population characteristics.

	13q14.3 Deletion	No 13q14.3 Deletion	Total	P-value
<b>Gender</b>				
Male	40 (0.09%)	44,012 (99.91%)	44,052	(53.41%)
Female	20 (0.05%)	38,411 (99.95%)	38,431	(46.59%)
	60 (0.07%)	82,423 (99.93%)	82,483	(100.00%)
<b>Age Group</b>				0.0513
<50	2 (0.02%)	9,882 (99.98%)	9,884	(13.12%)
50–54	2 (0.03%)	6,437 (99.97%)	6,439	(8.55%)
55–59	4 (0.04%)	9,963 (99.96%)	9,967	(13.23%)
60–64	8 (0.06%)	14,048 (99.94%)	14,056	(18.65%)
65–69	10 (0.07%)	15,174 (99.93%)	15,184	(20.15%)
70–74	13 (0.11%)	11,639 (99.89%)	11,652	(15.46%)
75	15 (0.18%)	8,153 (99.82%)	8,168	(10.84%)
	54 (0.07%)	75,296 (99.93%)	75,350	(100.00%)
<b>Ancestry</b>				0.0018 *
>85% European	53 (0.10%)	54,765 (99.90%)	54,818	(72.82%)
>85% African	0 (0.00%)	2,647 (100.00%)	2,647	(03.52%)
>85% Asian	3 (0.02%)	17,810 (99.98%)	17,813	(23.66%)
	56 (0.07%)	75,222 (99.93%)	75,278	(100.00%)
<b>Non-Hematologic Cancer</b>				0.0004 *
Cancer-free	21 (0.06%)	36,208 (99.94%)	36,229	(43.92%)
Cancer	39 (0.08%)	46,215 (99.92%)	46,254	(56.08%)
Bladder	4 (0.08%)	5,036 (99.92%)	5,040	(10.90%)
Breast	2 (0.07%)	2,827 (99.93%)	2,829	(06.12%)
Colorectal	0 (0.00%)	95 (100.00%)	95	(00.21%)
Endometrial	6 (0.21%)	2,908 (99.79%)	2,914	(06.30%)
Esophagus	0 (0.00%)	1,915 (100.00%)	1,915	(04.14%)
Glioma	0 (0.00%)	1,729 (100.00%)	1,729	(03.74%)
Head/Neck	0 (0.00%)	25 (100.00%)	25	(00.05%)
Kidney	2 (0.13%)	1,589 (99.87%)	1,591	(03.44%)

	<b>13q14.3 Deletion</b>	<b>No 13q14.3 Deletion</b>	<b>Total</b>	<b>P-value</b>
Liver	0 (0.00%)	13 (100.00%)	13 (00.03%)	1.0000
Lung	12 (0.09%)	13,251 (99.91%)	13,263 (28.67%)	0.2378
Osteosarcoma	0 (0.00%)	982 (100.00%)	982 (02.12%)	1.0000
Other	1 (1.20%)	82 (98.80%)	83 (00.18%)	0.0491 *
Ovary	0 (0.00%)	548 (100.00%)	548 (01.18%)	1.0000
Pancreas	4 (0.10%)	3,919 (99.90%)	3,923 (08.48%)	0.3002
Prostate	7 (0.10%)	7,238 (99.90%)	7,245 (15.66%)	0.3051
Skin	0 (0.00%)	49 (100.00%)	49 (00.11%)	1.0000
Stomach	1 (0.03%)	3,360 (99.97%)	3,361 (07.27%)	1.0000
Testis	0 (0.00%)	649 (100.00%)	649 (01.40%)	1.0000
	60 (0.07%)	82,423 (99.93%)	82,483 (100.00%)	0.1932
<b>Total Sample</b>	<b>60 (0.07%)</b>	<b>82,423 (99.93%)</b>	<b>82,483 (100.00%)</b>	

P-values are from Fisher exact tests with an asterisk (\*) indicating  $P < 0.05$ . For Ancestry,  $>85\%$  European, for example, indicates the individual has greater than 85% inferred European ancestry.



Comparison of observed 13q14 mosaic deletions to that expected in a sample of this size, age, and sex composition attributable to undiagnosed CLL cases.

**TABLE 2**

Sex	Age Group	N	Observed		Expected		P-value
			Count	Percent	Count	Percent	
Males	Ages <50	4,068	1	0.025%	0.04	0.001%	0.6858
	Ages 50–64	15,959	6	0.038%	4.62	0.029%	
	Ages 65–74	16,973	16	0.094%	17.50	0.103%	
	Ages 75+	5,225	11	0.211%	9.08	0.174%	
All Ages		42,225	34	0.081%	31.24	0.074%	
Females	Ages <50	5,816	1	0.017%	0.03	0.001%	0.5403
	Ages 50–64	14,489	6	0.041%	2.30	0.016%	
	Ages 65–74	9,831	5	0.051%	5.60	0.057%	
	Ages 75+	2,935	2	0.068%	3.50	0.119%	
All Ages		33,071	14	0.042%	11.43	0.035%	
Overall		75,296	48	0.064%	42.67	0.057%	0.4595