Detectable clonal mosaicism and its relationship to aging and cancer

A full list of authors and affiliations appears at the end of the article.

Abstract

In an analysis of 31,717 cancer cases and 26,136 cancer-free controls drawn from 13 genome-wide association studies (GWAS), we observed large chromosomal abnormalities in a subset of clones from DNA obtained from blood or buccal samples. Mosaic chromosomal abnormalities, either aneuploidy or copy-neutral loss of heterozygosity, of size >2 Mb were observed in autosomes of 517 individuals (0.89%) with abnormal cell proportions between 7% and 95%. In cancer-free individuals, the frequency increased with age; 0.23% under 50 and 1.91% between 75 and 79 ($p=4.8\times10^{-8}$). Mosaic abnormalities were more frequent in individuals with solid-tumors (0.97% versus 0.74% in cancer-free individuals, OR=1.25, p=0.016), with a stronger association for cases who had DNA collected prior to diagnosis or treatment (OR=1.45, p=0.0005). Detectable clonal mosaicism was common in individuals for whom DNA was collected at least one year prior to diagnosis of leukemia compared to cancer-free individuals (OR=35.4, $p=3.8\times10^{-11}$). These findings underscore the importance of the role and time-dependent nature of somatic events in the etiology of cancer and other late-onset diseases.

*Correspondence should be addressed to (chanocks@mail.nih.gov).
‡L.A.P.J. and S.J.C. are co-last authors.

Classically, genetic mosaicism is defined as the co-existence of cells with two or more distinct karyotypes within an individual that results from a post-zygotic event during development and can occur in both somatic and germline cells[1,2]. Errors in chromosomal duplication and subsequent transmission to daughter cells may lead to aneuploidy, the gain or loss of chromosomes or segments of chromosomes, and reciprocal gain and loss events manifesting in copy-neutral loss of heterozygosity (cnloh) or acquired uniparental disomy. Somatic mosaicism has been established as a cause of miscarriage, birth defects, developmental delay, and cancer[3-9]. Because mosaicism can be benign or may manifest with diverse clinical phenotypes, there are no accurate estimates of its frequency in the general population[3,6]. On rare occasions the propensity to develop chromosomal abnormalities is inherited and leads to multiple phenotypic abnormalities including cancer predisposition as reported in families with mutations in *BUB1B* and *CEP57* [10,11]. Recently, two groups have identified somatic mosaic mutations in *IDH1* and *IDH2* in tumors of individuals with Ollier disease and Maffucci syndrome[12,13] while another group has characterized somatic mosaicism of a *HRAS* mutation in an individual with urothelial cancer and epidermal nevus[14]. Recent work in a population of twins has suggested that the detection of somatic structural variants in blood increases with aging and may be related to reduction in blood cell clonality[15]. In this report, we define mosaic chromosomal abnormalities broadly: the presence of both normal karyotypes and those with large structural genomic events resulting in alteration of copy number or loss of heterozygosity in distinct and detectable subpopulations of cells regardless of the clonal or developmental origin of the subpopulations.

Recently, we reported on 1,991 individuals from the Spanish Bladder Cancer/EPICURO population-based case-control study in which we had performed a GWAS of adult-onset bladder cancer using DNA obtained from blood or buccal samples[16]. The SNP array data generated for the GWAS was subsequently used to detect clonal mosaic abnormalities in the autosomes of 1.7% of study subjects, suggesting a higher frequency in adults than previously suspected. Even though somatic mosaicism has been implicated in several cancers, this study did not reveal a significant difference in frequency between cases and controls. A computational algorithm was used to detect 42 large mosaic events involving two or more distinct clones in DNA extracted from blood or buccal samples and we experimentally validated the findings using multiplex ligation-dependent probe amplification (MLPA) and microsatellite analysis (as well as fluorescent *in situ* hybridization in a subset), establishing the robustness of the software detection method. A similar proportion of cells carrying each event was found in 5 of 6 events (in four individuals with bladder cancer in whom three had one event and one individual with three separate events) in which it was possible to examine more than one tissue (whole blood and bladder mucosa), suggesting an early embryonic origin of the somatic mutation leading to the observed mosaic chromosomal abnormalities[16].

Results

In this report, we extend our analysis of clonal mosaic abnormalities in the autosomes to 57,853 individuals (including those previously published[16]). We tested 31,717 cancer cases and 26,136 cancer free controls for evidence of mosaic abnormalities using genome-wide

SNP array data generated as part of 13 distinct cancer GWAS drawn from 48 epidemiological case-control and case-cohort studies (Supplementary Table 1). DNA samples were extracted from blood or buccal samples using a variety of collection and extraction techniques and genotyped using one or more Infinium Human SNP arrays from Illumina Inc. (including versions of Hap300, Hap240, Hap550, Hap610, Hap660, Hap1, Omni Express, and Omni1). Genotype clusters were empirically estimated in 45 batches to optimize accuracy while minimizing potential batch effects (Online Methods).

Detection of clonal mosaic events was based on assessment of allelic imbalance and copy number changes. We used the B-allele frequency (BAF) measurement, derived from the ratio of probe values relative to the locations of the estimated genotype-specific clusters, for initial segmentation using the Mosaic Alteration Detection (MAD) algorithm implemented in GADA-R with modifications[17,18]. The BAF and $\log_2$ relative probe intensity ratio (LRR), which provides data on copy number, were used to classify each event as copy-altering (gain or loss) or neutral (reciprocal gain and loss resulting in loss of heterozygosity, LOH) and to assign the proportions of abnormal (p) and normal (1-p) cells. Mosaic proportions were required to deviate from levels expected from constitutional (non-mosaic) changes in order to exclude homozygous chromosomal segments inherited identical by descent and non-mosaic instances of trisomy, monosomy and uniparental disomy. A minimum event size threshold was set to detect only clonal mosaic events greater than 2 Mbps to minimize the false discovery of constitutional copy number variants. Copy-neutral LOH and copy-loss events could be detected for mosaic proportions between 7% and 95% (Figure 1) with sensitivity that was affected by the signal-to-noise ratio characteristic of each microarray assay and sample quality. There was reduced sensitivity to distinguish between copy-neutral LOH and copy-loss events for mosaic proportions less than 15% across the autosomes. The magnitude of BAF differences for single-copy gain events was 1/3 of the magnitude of copy-neutral LOH or copy-loss events, reducing the sensitivity for calling copy-gain events. As a result, single copy gain events could only be reliably detected for mosaic proportions between 22% and 88%, with ambiguity in distinguishing copy-gain from copy-neutral LOH for mosaic proportions of less than 20%. Since DNA was obtained for the purpose of performing a GWAS, it was not possible to further explore the developmental and clonal characteristics of mosaic events detected in these individuals (e.g. by studying DNA from fractionated blood and other tissue types, determining cell composition of buccal samples, or effect of DNA collection and extraction methods on detection and accuracy of the estimation of mosaic proportions). We report only autosomal chromosomal abnormalities, as the analysis of the sex chromosomes presents distinct technical and interpretative challenges.

We observed 681 mosaic segments of size greater than 2 Mb on 641 autosomal chromosomes in 517 individuals for an overall frequency of individuals with mosaicism of 0.87% (Tables 1 and 2). The most frequent type of event observed was copy-neutral LOH (48.2%), while copy-gains and copy-losses were observed for 15.1% and 34.8% of mosaic events, respectively (Table 1). A small proportion (1.9%) of mosaic chromosomes were complex, harboring more than one type of event. 18.7% of mosaic chromosomal events spanned the entire chromosome, including 62 complete trisomies, predominantly in chromosomes 8, 12 and 15. 47.9% of mosaic chromosomal events began at a telomere and extended across some portion of the chromosomal arm (Table 1 and Figure 2). The majority

of telomeric events were mosaic copy-neutral LOH (85.7%), most frequently on 9p (Table 3). The remaining mosaic chromosomal events were interstitial (31.5%) spanning neither telomere nor centromere, while an additional small proportion (1.8%) spanned the centromere or had more complex structure (e.g. distinct events involving both telomeres, but not the whole chromosome). The majority of interstitial events were mosaic copy-loss (91.6%), which was most frequently observed within specific regions of chromosomes 13q and 20q (Figure 2). We observed 69 individuals (46 cancer cases and 23 cancer-free individuals) with clonal mosaic events on multiple chromosomes. The distribution of the number of clonal mosaic chromosomal events per individual is shown in Supplementary Table 3. Among cancer-free individuals, the greatest number observed was 5 mosaic chromosomal events, whereas six individuals with cancer had greater than 5 events, including two individuals with gastric cancer who each had 20. A list of mosaic events with phenotype data is available as Supplementary Data.

The strongest predictor of mosaic autosomal abnormalities was age at DNA collection. We examined the effect of aging on the frequency of mosaicism across all studies, which were predominantly individuals over the age of 50. The frequency of cancer-free individuals with detectable clonal mosaic events increased with age, from 0.23% for those under 50 to 1.91% ($p=4.8\times10^{-8}$) for those between the ages of 75 and 79, and with slightly higher frequencies for individuals with cancer (Figure 3). In the early onset cancers (under age 40), which constituted less than 5% of analyzed cases (e.g., testicular cancer and osteogenic sarcoma), we did not observe an increase in mosaic abnormalities. Further studies are needed to investigate the relationship between mosaic abnormalities and cancer in children and young adults, particularly because of the strong association between mosaicism and many developmental disorders. There was no apparent relationship between age at DNA collection and the number, size of mosaic events, or the proportion of abnormal cells (Supplementary Figures 1 and 2).

We regressed the presence of detectable clonal mosaicism in 26,136 cancer-free individuals on age at DNA collection (in 5 year intervals), sex (male versus female), DNA source (buccal cells versus blood), smoking (ever versus never) and admixture coefficients for African and East Asian ancestry in a logistic model to determine the additional factors that influenced frequency of detectable clonal mosaicism. The source of DNA was known for 87% of individuals, of whom 19% were derived from buccal cells and the remainder from blood. DNA source was not significantly associated with mosaicism (OR=0.83, 0.55-1.26 95% confidence interval (CI), p=0.39). By admixture analysis, 75% of subjects were determined to be of European ancestry, 9% of African ancestry and 16% of East Asian ancestry. Although power was limited, we observed that cancer-free individuals with African admixture were at a lower risk of being mosaic (OR=0.43, 0.20-0.92 95% CI, p=0.03), but not in those with East Asian admixture (OR=0.60, 0.32-1.15 95% CI, p=0.12). We did not observe an association between smoking (ever/never) and frequency of mosaic abnormalities (OR=1.04, 0.75-1.44 95% CI, p=0.81).

In 26,136 cancer-free controls and 23,093 cancer cases drawn from non-sex specific and non-hematological cancer sites (i.e. excluding 8,470 individuals with leukemia, lymphoma, multiple myeloma and cancers of the breast, endometrium, ovary, testis, and prostate), we

observed a higher frequency of males with mosaic abnormalities than females. In cancer-free individuals, we observed mosaic events in 0.56% of females and 0.87% males (OR=1.35, 0.98-1.88 95% CI, p=0.07); for individuals with cancer we observed mosaic events in 0.79% of females and 1.21% of males (OR=1.48, 1.08-2.03 95% CI, p=0.015); and overall, 0.65% of females and 1.04% of males (OR=1.42, 1.14-1.80 95% CI, p=0.002) in logistic models adjusted for cancer diagnosis (if applicable), age at DNA collection, ancestry, DNA source and smoking. These differences could be due to a true sex-specific effect akin to sex-differential mutation and recombination rates[19]; however the complex and heterogeneous nature of the inclusion of individual studies and the differences in their entry and selection criteria could result in spurious associations. Although this observation was consistent across cancer types, it should be confirmed in additional studies better designed to address this question.

To determine the relationship between detectable mosaic autosomal abnormalities and non-hematological cancers, we regressed the presence of detectable clonal mosaicism on cancer diagnosis, age, sex, DNA source, smoking and ancestry in a logistic model. We observed a modest increase in cancer risk for mosaic individuals (OR=1.27, 1.05-1.52 95% CI, p=0.012) (Tables 2 & Supplementary Table 2). Notable associations were observed in stratified analyses of lung (OR=1.56, 1.18-2.08 95% CI, p=0.002) and kidney (OR=1.98, 1.27-3.06 95% CI, p=0.002) cancers, both tobacco-associated malignancies. However no cancer site-specific associations were observed for bladder, esophagus, stomach and pancreas cancers, which are also typically associated with tobacco use. There was no significant association in non-hematological cancer cases overall between smoking (ever/never) and frequency of mosaicism (OR=1.19, 0.92-1.54 95% CI, p=0.19) or when stratified by cancer site (results not shown).

In an analysis of the subset of 14,050 individuals with cancer for whom it was possible to determine that DNA was likely obtained before or at the time of diagnosis and prior to treatment with radiation or chemotherapy for a primary tumor (designated as "likely untreated"), we observed a stronger association between mosaic abnormalities and non-hematological cancer diagnosis (OR=1.45, 1.18-1.80 95% CI, p=0.0005). The associations for lung and kidney also increased in significance (Table 3). It is notable that the evidence for association with non-hematological cancer diminished in individuals who were potentially treated (OR=1.03, 0.81-1.30 95% CI, p=0.80). We had approached this analysis with the hypothesis that there could be an increased frequency in detectable clonal mosaicism in non-hematological cancers induced by chemotherapy or radiotherapy but were surprised to observe the frequency was reduced to virtually the same as in the cancer-free population. Although this attenuated effect could have many explanations (e.g., related to the diagnosis and treatment of a solid tumor leading to a decrease in populations of cells with mosaic alteration), we had a limited capacity to model and control for treatment-effects since many of the studies did not provide any treatment information or only provided incomplete, retrospective ascertainment of the specifics. Although many of the participating studies were prospectively ascertained cohorts, DNA collection often occurred after cancer diagnosis. Additional studies are needed in prospectively ascertained cohorts and longitudinal studies in which multiple DNA samples were collected prior to and after diagnosis in order to explore treatment and disease effects.

For the 43 individuals with hematological cancers for whom DNA was obtained at least a year prior to diagnosis, the frequency of detectable clonal mosaicism was 20% for myeloid leukemia and 22% for lymphocytic leukemia (predominantly chronic lymphocytic leukemia, Table 2) compared to 0.74% in 26,136 cancer free controls (overall OR=35.4, 14.7-76.6 95% CI, Fisher exact p=3.8×10^{-11}). Of the 8 mosaic individuals with leukemia for whom DNA samples were collected at least a year prior to diagnosis, 4 were diagnosed with chronic lymphocytic leukemia (CLL) of which 2 had a mosaic deletion in a region of chromosome 13q14 previously described to be deleted in CLL[20]. DNA was obtained more than 5 years prior to diagnosis for 6 mosaic individuals, with the longest interval being 14 years, suggesting that detectable clonal mosaicism could be a marker of hematological cancer or its precursors, i.e., monoclonal B cell lymphocytosis (MBL) for CLL and myelodysplastic syndrome for acute myelogenous leukemia. Recent work shows that the majority of MBL have mono- or biallelic 13q14 abnormalities[21]. However, further studies will be needed, preferably with serial pre- and post- diagnosis sampling to investigate the predictive nature of detectable clonal mosaicism, especially involving regions of chromosome 13 and 20 with respect to leukemia risk[20].

We further explored the 4 most recurrent altered regions (>20), which also harbor well known cancer genes (as noted in the COSMIC[22] and Mitelman databases: http:// cgap.nci.nih.gov/Chromosomes/Mitelman); these were on chromosomes 9p (cnloh), 13q (del), 14 (cnloh) and 20q (del) (Table 4). Notably, the most recurrent mosaic events were observed in cancer-free individuals as well as across multiple solid tumors. We observed a comparable frequency in non-hematologic cancer cases and cancer-free controls for three of the regions, whereas the chromosome 14 cnloh abnormalities were more frequent in non-hematological cancer cases (OR=3.32, 1.42-9.00 95% CI, Fisher's exact p=0.003), particularly in individuals with bladder or kidney cancer. Copy-neutral loh in this region of chromosome 14 has been associated with increased susceptibility to sporadic cancers and harbors imprinted genes, such as the tumor suppressing non-coding RNA, Maternally expressed gene 3 (*MEG3*)[8,23]. The recurrent segmental deletion of 13q14 was observed in 5 leukemia cases, but also in 18 individuals with solid tumors (9 with lung cancer and 4 with prostate cancer), and in 10 cancer-free individuals. This region includes the tumor suppressor gene *DLEU7* (Deleted In Leukemia 7) and related genes, *DLEU1* and *DLEU2*, the latter harboring two microRNAs within one of its introns (miR-15a and miR-16-1)[24-26]. The retinoblastoma gene, *RB1* was also included within a subset harboring a mosaic deletion of 13q14. It cannot be ruled out that these individuals have either undiagnosed CLL or MBL. The 20q- was seen in two individuals with myeloid leukemia as has been described previously[27] but also in cancer-free and individuals with solid tumors.

The accuracy of our software methods to detect clonal mosaic abnormalities was previously addressed and we were able to validate 100% of 42 events in 34 individuals from the Spanish Bladder Cancer Study using confirmatory cytogenetic assays[16] (Supplementary Figure 3). We have also performed a comparison of mosaic events in samples from the EAGLE and PLCO lung cancer studies which were independently analyzed as part of the Gene, Environment Association studies consortium (GENEVA) report on mosaic events[28]. A total of 83 mosaic events in individuals from the EAGLE and PLCO lung cancer studies

were detected in common, 20 additional events of size less than 2 Mb and 8 events greater than 2 Mb were detected by GENEVA and not by our study, while we detected 20 additional events (size > 2 Mb) that were not detected by GENEVA. Although additional cytogenetic or molecular validation was not performed, neither method detected notable false-positive events based on manual review of the data. The concordance rate is 75% if considering events > 2 Mb (the cut-off for this analysis) or 63% if considering all events, both of which are considerably better than the 25-50% concordance rates observed across CNV detection methods[29-31]. Our method is more conservative in the size of events detected, while the GENEVA method is more conservative with respect to sample quality, but provides calls for smaller events when assay quality is sufficient. Better approaches are needed to characterize smaller size events accurately as either mosaic or constitutional and to estimate their frequency. Further improvements to data normalization, segmentation and event classification methods will also likely reduce false-negative rates.

Discussion

Our study has important implications for the design and analysis of molecular epidemiology studies in cancer as well as the somatic characterization of cancer genomes, like The Cancer Genome Atlas[32] and International Cancer Genome Consortium[33]. Investigators will need to carefully analyze samples used as exemplars of germline DNA for somatic alterations, such as detectable clonal mosaicism. Otherwise, comparisons between "grmline" and tumor DNA may result in implausible somatic changes (e.g. large gains of heterozygosity) and it may be impossible to determine whether somatic events pre-date changes secondary to driver mutations. Since how to detect mosaic events with next generation sequencing technologies is neither routine nor well understood, for the near future it may be prudent to continue to utilize SNP microarrays for such analyses. Due to the increased frequency of detectable clonal mosaicism with age, this will be particularly important for the analysis of epithelial cancers, which characteristically occur in the older population. For future large-scale GWAS in prospective studies, it may be wise to consider analyzing the earliest, pre-diagnosis DNA samples and to consider time from collection to diagnosis in the analysis of longitudinally collected biospecimens.

We have extended our initial observation that detectable clonal mosaicism of the autosomes is present in the population with surprising frequency and particularly in the aging genome. A recent study of detectable clonal mosaicism in twins reported an increase in frequency with age and suggested that this reduction could lead to a less diverse blood cell population and immune system[15]. These emerging data raise a number of critical issues in mechanisms underlying the possible shift in the repertoire of clones with large structural abnormalities. Thus cells with abnormal karyotypes could have an early developmental origin in which a somatic event in a single stem cell progenitor during embryogenesis could become apparent when cellular diversity decreases with age and cell populations become increasingly oligoclonal. Higher rates of detectable clonal mosaicism in older cancer-free individuals could also be due to increased rates of somatic mutation or diminished capacity for genomic maintenance, such as with telomere attrition[34] leading to proliferation of somatically altered cell populations. A survival bottleneck of cellular progenitors could also lead to observable mosaic alterations that were previously below the threshold of detection but subsequently

expanded due to positive selection. Further work is required to begin to unravel the underlying mechanisms that result in mosaic abnormalities, particularly as it relates to how and when altered clones are created, tissue-specificity, and the timing and expansion of distinct populations of cells with age. Finally, these findings underscore the importance of considering the role and time-dependent nature of somatic events in the etiology of cancer as well as other late-onset diseases.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Kevin B Jacobs[1,2], Meredith Yeager[1,2], Weiyin Zhou[1,2], Sholom Wacholder[1], Zhaoming Wang[1,2], Benjamin Rodriguez-Santiago[3,5], Amy Hutchinson[1,2], Xiang Deng[1,2], Chenwei Liu[1,2], Marie-Josephe Horner[1], Michael Cullen[1,2], Caroline G Epstein[1], Laurie Burdett[1,2], Michael C Dean[6], Nilanjan Chatterjee[1], Joshua Sampson[1], Charles C Chung[1], Joseph Kovaks[1], Susan M Gapstur[7], Victoria L Stevens[7], Lauren T Teras[7], Mia M Gaudet[7], Demetrius Albanes[1], Stephanie J Weinstein[1], Jarmo Virtamo[8], Philip R Taylor[1], Neal D Freedman[1], Christian C Abnet[1], Alisa M Goldstein[1], Nan Hu[1], Kai Yu[1], Jian-Min Yuan[9,10], Linda Liao[1], Ti Ding[11], You-Lin Qiao[12], Yu-Tang Gao[13], Woon-Puay Koh[14], Yong-Bing Xiang[13], Ze-Zhong Tang[11], Jin-Hu Fan[12], Melinda C Aldrich[15,16], Christopher Amos[17], William J Blot[16,18], Cathryn H Bock[19,20], Elizabeth M Gillanders[21], Curtis C Harris[22], Christopher A Haiman[23], Brian E Henderson[23], Laurence N Kolonel[24], Loic Le Marchand[24], Lorna H McNeill[25,26], Benjamin A Rybicki[27], Ann G Schwartz[19,20], Lisa B Signorello[16,18,28], Margaret R Spitz[29], John K Wiencke[30], Margaret Wrensch[30], Xifeng Wu[17], Krista A Zanetti[21,22], Regina G Ziegler[1], Jonine D Figueroa[1], Montserrat Garcia-Closas[1,31], Nuria Malats[32], Gaelle Marenne[32], Ludmila Prokunina-Olsson[1], Dalsu Baris[1], Molly Schwenn[33], Alison Johnson[34], Maria Teresa Landi[1], Lynn Goldin[1], Dario Consonni[35,36], Pier Alberto Bertazzi[35,36], Melissa Rotunno[1], Preetha Rajaraman[1], Ulrika Andersson[37], Laura E Beane Freeman[1], Christine D Berg[38], Julie E Buring[39,40], Mary A Butler[41], Tania Carreon[41], Maria Feychting[42], Anders Ahlbom[42], J Michael Gaziano[40,43,44], Graham G Giles[45,46], Goran Hallmans[47], Susan E Hankinson[48], Patricia Hartge[1], Roger Henriksson[37,49], Peter D Inskip[1], Christoffer Johansen[50], Annelie Landgren[1], Roberta McKean-Cowdin[23], Dominique S Michaud[51,52], Beatrice S Melin[37], Ulrike Peters[53,54], Avima M Ruder[41], Howard D Sesso[40], Gianluca Severi[45,46], Xiao-Ou Shu[28], Kala Visvanathan[55], Emily White[53,54], Alicja Wolk[56], Anne Zeleniuch-Jacquotte[57], Wei Zheng[28], Debra T Silverman[1], Manolis Kogevinas[58,61], Juan R Gonzalez[59,61], Olaya Villa[3,5], Donghui Li[62], Eric J Duell[63], Harvey A Risch[64], Sara H Olson[65], Charles Kooperberg[53], Brian M Wolpin[48,66], Li Jiao[67], Manal Hassan[62], William Wheeler[68], Alan A Arslan[69,71], H Bas Bueno-de-Mesquita[72,73], Charles S Fuchs[48,66], Steven Gallinger[74], Myron D Gross[75], Elizabeth A Holly[76], Alison P Klein[77], Andrea LaCroix[53], Margaret T Mandelson[53,78], Gloria Petersen[79], Marie-Christine Boutron-Ruault[80], Paige M Bracci[76], Federico Canzian[81], Kenneth

Chang[82], Michelle Cotterchio[83], Edward L Giovannucci[48,84,85], Michael Goggins[76,86,87], Judith A Hoffman Bolton[55], Mazda Jenab[88], Kay-Tee Khaw[89], Vittorio Krogh[90], Robert C Kurtz[91], Robert R McWilliams[92], Julie B Mendelsohn[1], Kari G Rabe[79], Elio Riboli[51], Anne Tjønneland[50], Geoffrey S Tobias[1], Dimitrios Trichopoulos[84,93], Joanne W Elena[38], Herbert Yu[24], Laufey Amundadottir[1], Rachael Z Stolzenberg-Solomon[1], Peter Kraft[48,84], Fredrick Schumacher[23], Daniel Stram[23], Sharon A Savage[1], Lisa Mirabello[1], Irene L Andrulis[74,94], Jay S Wunder[74,94], Ana Patiño García[95], Luis Sierrasesúmaga[95], Donald A Barkauskas[23], Richard G Gorlick[96,97], Mark Purdue[1], Wong-Ho Chow[1], Lee E Moore[1], Kendra L Schwartz[98], Faith G Davis[99], Ann W Hsing[1], Sonja I Berndt[1], Amanda Black[1], Nicolas Wentzensen[1], Louise A Brinton[1], Jolanta Lissowska[100], Beata Peplonska[101], Katherine A McGlynn[1], Michael B Cook[1], Barry I Graubard[1], Christian P Kratz[1,102], Mark H Greene[1], Ralph L Erickson[103], David J Hunter[48,84], Gilles Thomas[103], Robert N Hoover[1], Francisco X Real[3,105], Joseph F Fraumeni Jr.[1], Neil E Caporaso[1], Margaret Tucker[1], Nathaniel Rothman[1], Luis A Pérez-Jurado[3,4,‡], and Stephen J Chanock[1,‡,*]

Affiliations

[1] Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD, USA [2] Core Genotyping Facility, SAIC-Frederick, Inc., NCI-Frederick, Frederick, Maryland, USA [3] Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, E-08003 Barcelona, Spain [4] Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), E-08003 Barcelona, Spain [5] Quantitative Genomic Medicine Laboratory, qGenomics, E-08003 Barcelona, Spain [6] Laboratory of Experimental Immunology, Center for Cancer Research, NCI- Frederick, Frederick, MD, USA [7] Epidemiology Research Program, American Cancer Society, Atlanta, Georgia, USA [8] Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland [9] Department of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN, USA [10] University of Pittsburgh Cancer Institute, Pittsburgh, PA, USA [11] Shanxi Cancer Hospital, Taiyuan, Shanxi, People's Republic of China [12] Department of Epidemiology, Cancer Institute (Hospital), Chinese Academy of Medical Sciences, Beijing, People's Republic of China [13] Shanghai Cancer Institute, Shanghai, People's Republic of China [14] Saw Swee Hock School of Public Health, National University of Singapore, Singapore [15] Department of Thoracic Surgery, Vanderbilt University School of Medicine, Nashville, Tennessee, USA [16] Division of Epidemiology in the Department of Medicine, Vanderbilt Epidemiology Center,Nashville, Tennessee,USA [17] Department of Epidemiology, Division of Cancer Prevention and Population Sciences, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA [18] International Epidemiology Institute, Rockville, Maryland 20850, USA [19] Karmanos Cancer Institute, Wayne State University, Detroit, Michigan, USA [20] Department of Oncology, Wayne State University School of Medicine, Detroit, Michigan, USA [21] Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland 20892, USA [22] Laboratory of Human

Carcinogenesis, Center for Cancer Research, National Cancer Institute, Bethesda, Maryland 20892, USA [23] Department of Preventive Medicine, Keck School of Medicine, University of Southern California/Norris Comprehensive Cancer Center, Los Angeles, California 90033, USA [24] Epidemiology Program, University of Hawaii Cancer Center, Honolulu, Hawaii 96813, USA [25] Department of Health Disparities Research, Division of Office of the Vice-President, Cancer Prevention and Population Sciences, The University of Texas MD Anderson Cancer Center, Houston, Texas , USA [26] Center for Community Engaged Translational Research , Duncan Family Institute, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA [27] Department of Public Health Sciences, Henry Ford Hospital, Detroit, Michigan 48202, USA [28] Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee 37203,USA [29] Dan L Duncan Cancer Center, Baylor College of Medicine, Houston, TX, USA [30] Department of Neurological Surgery, University of California San Francisco , San Francisco, California 94158, USA [31] Division of Genetics and Epidemiology, Institute of Cancer Research, London, UK [32] Genetics and Molecular Epidemiology Group, Centro Nacional de Investigaciones Oncológicas (CNIO), Madrid, Spain [33] Maine Cancer Registry, Augusta, Maine, USA [34] Vermont Cancer Registry, Burlington, Vermont, USA [35] Department of Occupational and Environmental Health , University of Milan, Milan, 20122, Italy [36] Unit of Epidemiology, Fondazione Istituto di Ricevero e Cura a Carattere Scientifico (IRCCS), Ospedale Maggiore Policlinico, Mangiagalli e Regina Elena, Milan, 20122, Italy [37] Department of Radiation Sciences, Oncology, Umeå University, Umeå, Sweden [38] Clinical and Translational Epidemiology Branch, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD, USA [39] Department of Ambulatory Care and Prevention, Harvard Medical School, Boston, MA, USA [40] Division of Preventive Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA [41] National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, Cincinnati, OH, USA [42] Division of Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden [43] Division of Aging, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA [44] Massachusetts Veteran's Epidemiology, Research and Information Center, Geriatric Research Education and Clinical Center, VA Boston Healthcare System, Boston, MA, USA [45] Cancer Epidemiology Centre, The Cancer Council of Victoria, Melbourne, Australia [46] Centre for Molecular, Environmental, Genetic, and Analytic Epidemiology, University of Melbourne, Melbourne, Australia [47] Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden [48] Channing Laboratory, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA [49] Department of Oncology, Karolinska University Hospital, SE-171 76 Stockholm, Sweden [50] Institute of Cancer Epidemiology, Danish Cancer Society, Copenhagen, Denmark [51] Division of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK [52] Department of Epidemiology, Division of Biology and Medicine, Brown University, Providence, RI, USA [53] Division of Public Health Sciences, Fred

Hutchinson Cancer Research Center, Seattle, WA, USA [54] Department of Epidemiology, University of Washington, Seattle, WA, USA [55] Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA [56] Division of Nutritional Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden [57] Division of Epidemiology, Department of Environmental Medicine, NYU School of Medicine, New York, NY [58] National School of Public Health, Athens, Greece [59] Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Spain [60] Institut Municipal d'Investigació Mèdica (IMIM), Barcelona, Spain [61] Centro de Investigacion Biomedica en Red Epidemiologia y Slaud Publica (CIBERESP), Barcelona [62] Department of Gastrointestinal Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA [63] Catalan Institute of Oncology (ICO), Institut d'Investigació Biomèdica de Bellvitge (IDIBELL), Barcelona, Spain [64] Yale University School of Public Health, New Haven, CT, USA [65] Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY, USA [66] Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA [67] Department of Medicine, Baylor College of Medicine, Houston, TX, USA [68] Information Management Services, Silver Spring, MD, USA [69] Department of Obstetrics and Gynecology, New York University School of Medicine, New York, NY, USA [70] Department of Environmental Medicine, New York University School of Medicine; New York, NY, USA [71] New York University Cancer Institute, New York, NY, USA [72] Department of Gastroenterology and Hepatology, University Medical Center Utrecht, Utrecht, The Netherlands [73] National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands [74] Fred. A. Litwin Centre for Cancer Genetics, Samuel Lunenfeld Research Institute, Toronto, Ontario, Canada [75] Department of Laboratory Medicine and Pathology, School of Medicine, University of Minnesota, Minneapolis, MN, USA [76] Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA, USA [77] Department of Oncology, the Johns Hopkins University School of Medicine, Baltimore, MD, USA [78] Group Health Center for Health Studies, Seattle, WA, USA [79] Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA [80] Institut National de la Santé et de la Recherche Médicale (INSERM), Paris-Sud University, Institut Gustave-Roussy, Villejuif, France [81] Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany [82] Comprehensive Digestive Disease Center, University of California, Irvine Medical Center, Orange, CA, USA [83] Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada [84] Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA [85] Department of Nutrition, Harvard School of Public Health, Boston, MA, USA [86] Department of Medicine, The Johns Hopkins University School of Medicine, Baltimore, MD, USA [87] Department of Pathology, The Johns Hopkins University School of Medicine, Baltimore, MD, USA [88] International Agency for Research on Cancer (IARC/WHO), Lyon, France [89] Department of Public Health and Primary Care, Clinical Gerontology, Addenbrooke's Hospital, University of Cambridge, Cambridge, UK [90]

Nutritional Epidemiology Unit, Fondazione Istituto Di Ricovero e Cura a Carattere Scientifico (IRCCS), Istituto Nazionale dei Tumori, Milan, Italy [91] Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY, USA [92] Division of Medical Oncology, Mayo Clinic, Rochester, MN, USA [93] Bureau of Epidemiologic Research, Academy of Athens, Athens, Greece [94] Mount Sinai Hospital Christopher Sharp Centre for Surgery and Oncology, Toronto, Ontario, Canada [95] Department of Pediatrics, Clínica Universidad de Navarra, E31080 Pamplona, Spain [96] Department of Molecular Pharmacology , Albert Einstein College of Medicine of Yeshiva University, Bronx, NY, USA [97] Department of Pediatrics, Albert Einstein College of Medicine of Yeshiva University, Bronx, NY, USA [98] Department of Family Medicine and Public Health Sciences, Wayne State University, MI, USA [99] Division of Epidemiology and Biostatistics, School of Public Health, University of Illinois at Chicago, Chicago, Illinois, USA [100] Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland [101] Nofer Institute of Occupational Medicine, Lodz, Poland [102] Zentrum für Kinderheilkunde und Jugendmedizin, Klinik für Pädiatrische Hämatologie und Onkologie, Medizinische Hochschule Hannover, Hannover, Germany [103] Walter Reed Army Institute of Research, Silver Spring, MD, USA [104] Synergie-Lyon-Cancer, Universite Lyon 1, Centre Leon Berard, Lyon, France [105] Epithelial Carcinogenesis Group, Centro Nacional de Investigaciones Oncológicas (CNIO), Madrid, Spain

Acknowledgments

References

1. Youssoufian H, Pyeritz RE. Mechanisms and consequences of somatic mosaicism in humans. Nat Rev Genet. 2002; 3:748–58. [PubMed: 12360233]

2. Notini AJ, Craig JM, White SJ. Copy number variation and mosaicism. Cytogenet Genome Res. 2008; 123:270–7. [PubMed: 19287164]

3. Hsu LY, et al. Proposed guidelines for diagnosis of chromosome mosaicism in amniocytes based on data derived from chromosome mosaicism and pseudomosaicism studies. Prenat Diagn. 1992; 12:555–73. [PubMed: 1508847]

4. Menten B, et al. Emerging patterns of cryptic chromosomal imbalance in patients with idiopathic mental retardation and multiple congenital anomalies: a new series of 140 patients and review of published reports. J Med Genet. 2006; 43:625–33. [PubMed: 16490798]

5. Lu XY, et al. Genomic imbalances in neonates with birth defects: high detection rates by using chromosomal microarray analysis. Pediatrics. 2008; 122:1310–8. [PubMed: 19047251]

6. Conlin LK, et al. Mechanisms of mosaicism, chimerism and uniparental disomy identified by single nucleotide polymorphism array analysis. Hum Mol Genet. 2010; 19:1263–75. [PubMed: 20053666]

7. Heim, S.; Mitelman, F. Nonrandom chromosome abnormalities in cancer - an overview. In: Mitelman, F.; Heim, S., editors. Cancer Cytogenetics. John Wiley & Sons, Inc.; Hoboken, NJ: 2009.

8. Tuna M, Knuutila S, Mills GB. Uniparental disomy in cancer. Trends Mol Med. 2009; 15:120–8. [PubMed: 19246245]

9. Solomon DA, et al. Mutational inactivation of STAG2 causes aneuploidy in human cancer. Science. 2011; 333:1039–43. [PubMed: 21852505]

10. Rio Frio T, et al. Homozygous BUB1B mutation and susceptibility to gastrointestinal neoplasia. N Engl J Med. 2010; 363:2628–37. [PubMed: 21190457]

11. Snape K, et al. Mutations in CEP57 cause mosaic variegated aneuploidy syndrome. Nat Genet. 2011; 43:527–9. [PubMed: 21552266]

12. Amary MF, et al. Ollier disease and Maffucci syndrome are caused by somatic mosaic mutations of IDH1 and IDH2. Nat Genet. 2011

13. Pansuriya TC, et al. Somatic mosaic IDH1 and IDH2 mutations are associated with enchondroma and spindle cell hemangioma in Ollier disease and Maffucci syndrome. Nat Genet. 2011

14. Hafner C, T.A. Real FX. HRAS mutation mosaicism causing urothelial cancer and epidermal nevus. N Engl J Med. 2011; 365:1940–2. [PubMed: 22087699]

15. Forsberg LA, et al. Age-related somatic structural changes in the nuclear genome of human blood cells. Am J Hum Genet. 2012; 90:217–28. [PubMed: 22305530]

16. Rodriguez-Santiago B, et al. Mosaic uniparental disomies and aneuploidies as large structural variants of the human genome. Am J Hum Genet. 2010; 87:129–38. [PubMed: 20598279]

17. Gonzalez JR, et al. A fast and accurate method to detect allelic genomic imbalances underlying mosaic rearrangements using SNP array data. BMC Bioinformatics. 2011; 12:166. [PubMed: 21586113]

18. Pique-Regi R, Caceres A, Gonzalez JR. R-Gada: a fast and flexible pipeline for copy number analysis in association studies. BMC Bioinformatics. 2010; 11:380. [PubMed: 20637081]

19. Hedrick PW. Sex: differences in mutation, recombination, selection, gene flow, and genetic drift. Evolution. 2007; 61:2750–71. [PubMed: 17976181]

20. Dohner H, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med. 2000; 343:1910–6. [PubMed: 11136261]

21. Lanasa MC, et al. Immunophenotypic and gene expression analysis of monoclonal B-cell lymphocytosis shows biologic characteristics associated with good prognosis CLL. Leukemia. 2011; 25:1459–66. [PubMed: 21617698]

22. Forbes SA, et al. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. Nucleic Acids Res. 2011; 39:D945–50. [PubMed: 20952405]

23. Benetatos L, Vartholomatos G, Hatzimichael E. MEG3 imprinted gene contribution in tumorigenesis. Int J Cancer. 2011; 129:773–9. [PubMed: 21400503]

24. Lee S, et al. Forerunner genes contiguous to RB1 contribute to the development of in situ neoplasia. Proc Natl Acad Sci U S A. 2007; 104:13732–7. [PubMed: 17702869]

25. Migliazza A, et al. Nucleotide sequence, transcription map, and mutation analysis of the 13q14 chromosomal region deleted in B-cell chronic lymphocytic leukemia. Blood. 2001; 97:2098–104. [PubMed: 11264177]

26. Pekarsky Y, Zanesi N, Croce CM. Molecular basis of CLL. Semin Cancer Biol. 2010; 20:370–6. [PubMed: 20863894]

27. Gurvich N, et al. L3MBTL1 polycomb protein, a candidate tumor suppressor in del(20q12) myeloid disorders, is essential for genome stability. Proc Natl Acad Sci U S A. 2010; 107:22552–7. [PubMed: 21149733]

28. Laurie CC, Laurie CA, Kenneth R, Doheny KF. Somatic mosaicism for large chromosomal anomalies from birth to old age and its relationship to cancer. submitted to Nature Genetics. 2011

29. Pinto D, et al. Comprehensive assessment of array-based platforms and calling algorithms for detection of copy number variants. Nat Biotechnol. 2011; 29:512–20. [PubMed: 21552272]

30. Dellinger AE, et al. Comparative analyses of seven algorithms for copy number variant identification from single nucleotide polymorphism arrays. Nucleic Acids Res. 2010; 38:e105. [PubMed: 20142258]

31. Marenne G, et al. Assessment of copy number variation using the Illumina Infinium 1M SNP-array: a comparison of methodological approaches in the Spanish Bladder Cancer/EPICURO study. Hum Mutat. 2011; 32:240–8. [PubMed: 21089066]

32. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature. 2008; 455:1061–8. [PubMed: 18772890]

33. Hudson TJ, et al. International network of cancer genome projects. Nature. 2010; 464:993–8. [PubMed: 20393554]

34. Sahin E, Depinho RA. Linking functional decline of telomeres, mitochondria and stem cells during ageing. Nature. 2010; 464:520–8. [PubMed: 20336134]

35. Petersen GM, et al. A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. Nat Genet. 2010; 42:224–228. [PubMed: 20101243]

36. Staaf J, et al. Normalization of Illumina Infinium whole-genome SNP data improves copy number estimates and allelic intensity ratios. BMC Bioinformatics. 2008; 9:409. [PubMed: 18831757]

37. Diskin SJ, et al. Adjustment of genomic waves in signal intensities from whole-genome SNP genotyping platforms. Nucleic acids research. 2008; 36:e126. [PubMed: 18784189]

38. Peiffer DA, et al. High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. Genome Res. 2006; 16:1136–1148. [PubMed: 16899659]

39. Itsara A, et al. Population analysis of large copy number variants and hotspots, of human genetic disease. Am J Hum Genet. 2009; 84:148–161. [PubMed: 19166990]

40. Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 1974; 19:716–723.

41. Krzywinski MI, et al. Circos: An information aesthetic for comparative genomics. Genome Research. 2009

42. Agresti A, Coull BA. Approximate is better than "exact" for interval estimation of binomial proportions. Am Stat. 1998; 52:119–126.

43. Frazer KA, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature. 2007; 449:851–861. [PubMed: 17943122]
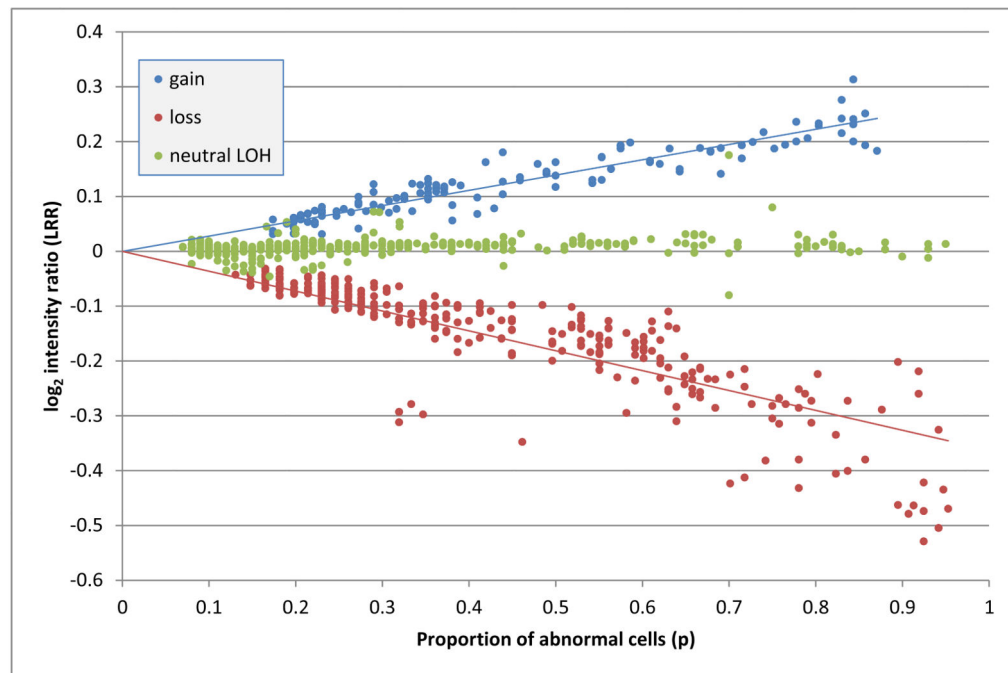
Figure 1. Characteristics of detectable clonal mosaic events
Detectable clonal mosaic events plotted by proportion of abnormal cells (p) and Log R Ratio
(LRR) for 681 events in 517 individuals.
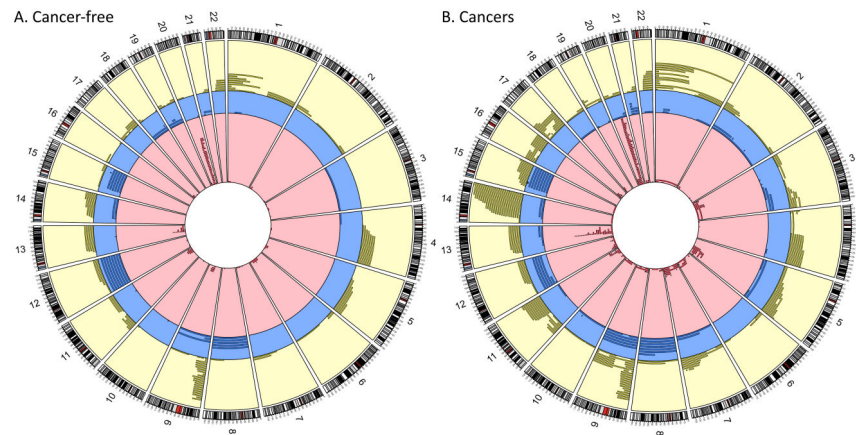
Figure 2. Circular genomic plot of detectable clonal mosaic events
Genomic location of detectable clonal mosaic events. Outer rings are the autosomes 1 to 22.
Yellow region denotes events of copy neutral loss of heterozygosity. Blue region denotes
copy gain events. Red region denotes copy loss events. Panel A includes events in cancer
free controls. Panel B includes events in cancer cases.
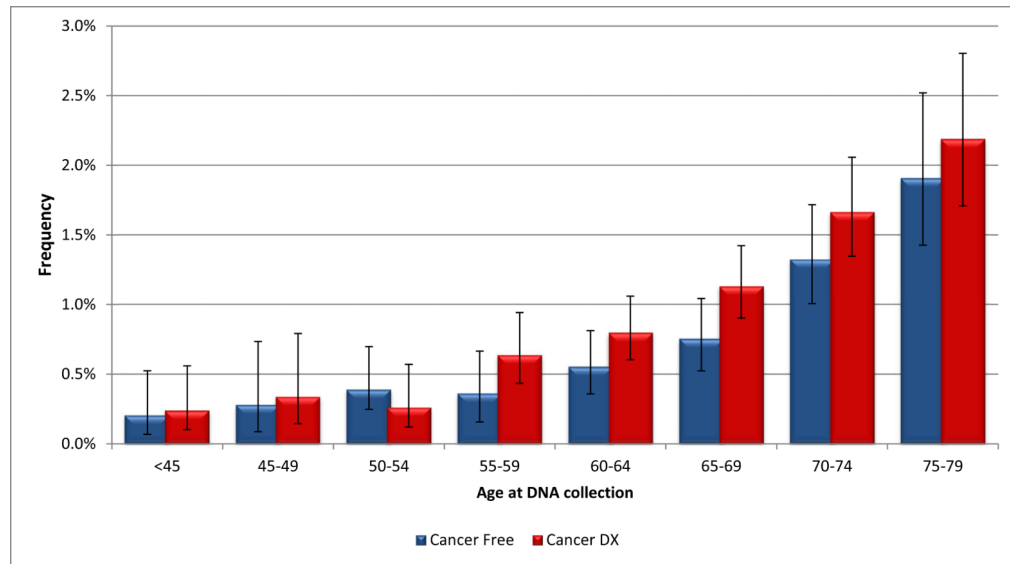
Figure 3. Frequency of detectable clonal mosaic events by age and cancer status
Analysis excluded 1,000 individuals with unknown age at DNA collection. 95% confidence intervals are shown.

Table 1

Count and frequency of mosaic chromosomal events by event type and location

Event Location	Mosaic Chromosome Count				Mosaic Chromosome Frequency (%)				
	gain	loss	cnloh	mixed	Total	gain	loss	cnloh	mixed	Total
chromosome	62	11	42	5	**120**	9.7	1.7	6.6	0.8	**18.7**
telomeric P	11	13	114	1	**139**	1.7	2.0	17.8	0.2	**21.7**
telomeric Q	9	10	149	0	**168**	1.4	1.6	23.2	0.0	**26.2**
interstitial	14	185	2	1	**202**	2.2	28.9	0.3	0.2	**31.5**
span centromere	1	1	2	0	**4**	0.2	0.2	0.3	0.0	**0.6**
complex	0	3	0	5	**8**	0.0	0.5	0.0	0.8	**1.2**
Total	**97**	**223**	**309**	**12**	**641**	**15.1**	**34.8**	**48.2**	**1.9**	

Table 2

Count and frequency of individuals with detectable clonal mosaic events for cancer-free individuals and by first diagnosed cancer site

Non-hematological cancers are listed by first cancer site and exclude anyone diagnosed with a hematological cancer, shown separately.

Site of first cancer	Mosaic Counts			Non-Mosaic Counts			Mosaic Frequency (%)		
	Likely Untreated	Possibly Treated	Total	Likely Untreated	Possibly Treated	Total	Likely Untreated	Possibly Treated	Overall
overall *			**498**			**57,201**			**0.86**
cancer-free			**194**			**25,942**			**0.74**
First non-hematologic cancer	**185**	**119**	**304**	**13,865**	**17,394**	**31,259**	**1.32**	**0.68**	**0.96**
bladder	37	6	43	2,240	973	3,213	1.62	0.61	1.32
breast	4	8	12	1,060	1,753	2,813	0.38	0.45	0.42
endometrium	3	6	9	247	624	871	1.20	0.95	1.02
esophagus	1	6	7	53	1,855	1,908	1.85	0.32	0.37
glioma	7	2	9	1,279	441	1,720	0.54	0.45	0.52
kidney	21	3	24	1,241	325	1,566	1.66	0.91	1.51
lung	73	26	99	4,647	2,605	7,252	1.55	0.99	1.35
osteosarcoma	0	3	3	0	760	760		0.39	0.39
ovary	1	3	4	260	283	543	0.38	1.05	0.73
pancreas	2	29	31	379	3,513	3,892	0.52	0.82	0.79
prostate	32	11	43	2,116	1,410	3,526	1.49	0.77	1.20
stomach	2	13	15	99	2,194	2,293	1.98	0.59	0.65
testis	2	0	2	144	503	647	1.37	0.00	0.31
other sites	0	3	3	100	155	255	0.00	1.90	1.16
Any hematologic cancer	**8**	**11**	**19**	**34**	**62**	**96**	**19.05**	**15.07**	**16.52**
leukemia	**8**	**9**	**17**	**34**	**11**	**45**	**19.05**	**45.00**	**27.42**
lymphoid	4	5	9	14	5	19	22.22	50.00	32.14
myeloid	4	3	7	16	5	21	20.00	37.50	25.00
other/nos	0	1	1	4	1	5	0.00	50.00	16.67
lymphoma	0	2	2	0	42	42		4.55	4.55
multiple myeloma	0	0	0	0	9	9	0.00	0.00	0.00

*
overall total of cancer-free individuals and those with non-hematologic cancers

Table 3

Distribution and frequency of recurrent detectable clonal mosaic events

	Mosaic Counts						Mosaic Frequency (%)				
	13q (del)	20q (del)	9p (cnloh)	14q (cnloh)	other	Total	13q (del)	20q (del)	9p (cnloh)	14q (cnloh)	other
Overall	**33**	**77**	**56**	**35**	**480**	**681**	**5**	**11**	**8**	**5**	**70**
Cancer free	**10**	**30**	**28**	**7**	**150**	**225**	**4**	**13**	**12**	**3**	**67**
Cancer DX	**23**	**47**	**28**	**28**	**330**	**456**	**5**	**10**	**6**	**6**	**72**
First non-hematologic cancer											
bladder	2	4	2	7	35	50	4	8	4	14	70
breast	1	1	0	1	13	16	6	6		6	81
endometrium	0	1	3	1	6	11		9	27	9	55
esophagus	0	1	1	1	6	9		11	11	11	67
glioma	0	2	2	0	6	10		20	20		60
kidney	1	3	0	4	20	28	4	11		14	71
lung	9	14	10	7	90	130	7	11	8	5	69
osteosarcoma	0	0	1	0	6	7			14		86
ovary	0	2	0	0	2	4		50			50
pancreas	1	4	1	1	29	36	3	11	3	3	81
prostate	4	10	4	1	33	52	8	19	8	2	63
stomach	0	1	4	3	55	63		2	6	5	87
testis	0	0	0	0	4	4					100
other sites	0	0	0	1	3	4				25	75
Any hematologic cancer											
leukemia	5	3	0	1	20	29	17	10		3	69
lymphoma	0	1	0	0	2	3		33			67