Published in final edited form as:

Semin Hematol. 2013 October ; 50(4): . doi:10.1053/j.seminhematol.2013.09.001.

Detectable Clonal Mosaicism in the Human Genome

Mitchell J. Machiela and Stephen J. Chanock*

Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland, USA. 20892-4605

Abstract

Human genetic mosaicism is the presence of two or more cellular populations with distinct genotypes in an individual who developed from a single fertilized ovum. While initially observed across a spectrum of rare genetic disorders, detailed assessment of data from genome-wide association studies now reveal that detectable clonal mosaicism involving large structural alterations (> 2 Mb) can also be seen in populations of apparently healthy individuals. The first generation of descriptive studies have generated new interest in understanding the molecular basis of the affected genomic regions, percent of the cellular subpopulation involved, and developmental timing of the underlying mutational event, which could reveal new insights into the initiation, clonal expansion and phenotypic manifestations of mosaic events. Early evidence indicates detectable clonal mosaicism increases in frequency with age and could preferentially occur in males. The observed pattern of recurrent events affecting specific chromosomal regions indicates some regions are more susceptible to these events, which could reflect inter-individual differences in genomic stability. Moreover, it is also plausible that the presence of large structural events could be associated with cancer risk. The characterization of detectable genetic mosaicism reveals that there could be important dynamic changes in the human genome associated with the aging process, which could be associated with risk for common disorders, such as cancer, cardiovascular disease, diabetes, and neurological disorders.

Keywords

aneuploidy; clonal expansion; copy-number; genome-wide association study; mutation

Introduction

Since the early 4th century BC, Greek and Roman cultures popularized the creation of large mosaic compositions from collections of small objects of varied size, color, and texture. While from a distance these mosaic depictions appear as one coherent and unified whole, a closer inspection reveals the many individual pieces of stone, shell, tile, or glass that outline the fine details of the illustration. The comparison of individual genetic heterogeneity to this artistic media was coined in humans as early as 1945¹. 'Genetic mosaicism' is now defined as the presence of two or more cellular populations with distinct genomic differences in an individual who developed from a single fertilized ovum² (Figure 1). It is postulated that mitotic errors and/or mutational events during development, adolescence, and adulthood result in genomically divergent cells, which, in turn, clonally expand to achieve a stable

*CORRESPONDING AUTHOR: Stephen J. Chanock, Laboratory of Translational Genomics, 8717 Grovemont Circle, ATC Room 134D, Office: (301) 435-7559, Fax: (301) 402-3134, chanocks@mail.nih.gov.

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

subset. Interestingly, genetic mosaicism has been noted in mitochondrial DNA^{3–5} as well as in the germline DNA of gametes^{6–12} and somatic cells^{13–15}. Furthermore, each of these types of mosaicism can give rise to phenotypic heterogeneity, resulting in a range of outcomes, including disease severity depending on the developmental timing of the mutation, cell lineages affected, genetic locations included, and relative percentage of the cellular sub-populations¹⁶. More recently, detectable genetic mosaicism, involving large structural abnormalities, has been observed in apparent healthy individuals based on new algorithms for examining intensity signals in single nucleotide polymorphism (SNP) microarrays^{13–15, 17–21}.

Just as in a mosaic piece of art, genetic mosaicism in humans reveals a high degree of complexity. Here we discuss the value of new genomic approaches in both characterizing and mapping genetic mosaicism in populations, a recently appreciated observation^{13–15}. The emerging picture suggests that a genome could erode slightly with aging and at the same time, provides new insights into the stability of chromosomes, which, in turn, could have important implications for human diseases and traits

Genetic Basis of Mosaicism

Genetic mosaicism arises from a post-zygotic mutational event in a cell or group of cells that undergoes positive selection and eventually achieves a stable, detectable, and distinct subpopulation characterized by an altered form of the genome. The range of events is wide, from single base pair changes to large structural abnormalities that may result in complete duplication or deletion of a chromosome. Manifestations of genetic mosaicism can be tissue specific, reflecting unique, cell-specific events. The timing of when events occur and their subsequent selection can vary greatly. The pediatric literature indicates that somatic events early in development can result in significant phenotypic effects, as observed in a small fraction of genetic disorders, such as neurofibromatosis type 1²² and trisomy 21²³. Early somatic alterations can arise in embryonic progenitors that remain below the threshold of detection until an event triggers a survival bottleneck. In turn, this leads to positive selection with rapid expansion of the second, ‘alternative’ clonal population. Alternatively, the timing of a somatic event can be later and due to either an increase in somatic alterations as a result of aging (e.g., breakdown in key DNA repair or stability pathways) alone or in combination with decreased genomic stability due to telomere attrition. Proliferation of the somatically altered clonal population then occurs, reaching a detectable fraction of one or more cell types. Since aging is associated with a decrease in cellular diversity and immunosenescence^{24, 25}, it is not surprising that the human genome can show signs of deterioration, namely the co-existence of two or more subpopulations of cells with distinct genomic features.

Perhaps the most well-known and extensively studied form of genetic mosaicism is cancer. Cancer is classically defined as a disease of genetic alterations that are “mosaic” with respect to the host germline. The landscape of genetic alterations that characterize a malignant tumor differs from the genotype of the surrounding tissue in many ways²⁶; the somatic alterations can drive oncogenesis, leading to de-differentiation, unrestrained growth and proliferation. It has been hypothesized that mutational events often occur during development when cells undergo numerous rounds of cellular division necessary to expand from a zygote to a multicellular organism²⁷. Mutations that arise during the exponential growth of early development may seed tissues with mosaic cellular populations harboring deleterious genetic alterations that may predispose to later onset diseases such as cancer- or in the case of childhood cancer, rapidly leads to detectable disease. Shared cellular mechanisms involved in the generation of somatic alterations and the clonal expansion of

cellular populations containing these alterations underscores the possible relationship between genetic mosaicism and future cancer risk.

It is unknown exactly what types of environmental or genetic factors are responsible for the development of detectable human clonal mosaicism. Studies in families, characterized by aneuploidy and cancer risk have uncovered key genes and pathways that, when inefficient can lead to detectable genetic mosaicism. For instance, the rare disorder known as mosaic variegated aneuploidy syndrome (MVAS) highlights the propensity to develop mosaic chromosomal aneuploidies and the subsequent development of one or more cancers. MVAS is an autosomal recessive disorder that can be characterized by a range of phenotypic manifestations such as growth retardation, heart defects, and facial deformities resulting from mosaic aneuploidies²⁸. Informative mutations have been identified in families with affected individuals in two or more generations. These include mutations in the gene, centrosomal protein 57kDa (*CEP57*), which encodes a centrosomal protein involved in nucleating and stabilizing microtubules, and in BUB1 homolog mitotic checkpoint serine/threonine kinase B (*BUB1B*), which encodes BUBR1 a critical checkpoint regulator of the spindle-assembly checkpoint^{28, 29}. Further work has shown that these mutations have an effect on cellular division, which, in turn, leads to mosaic karyotypes with multiple examples of aneuploidy chromosomes^{28, 29}. It is notable that the homozygous *BUB1B* mutation, c.2386-11A→G, increases risk for gastrointestinal neoplasia²⁹.

There are sparse data on how environmental factors can interact with genetic factors (e.g., less efficient variants due to polymorphisms or highly penetrant mutations) to result in detectable genetic mosaicism. Nonetheless, it is plausible that exposure to DNA mutagens such as ultraviolet radiation³⁰, ionizing radiation³¹, air pollution such as tobacco smoke and diesel exhaust^{32–34}, and industrial chemicals^{35, 36} can initiate the formation of DNA adducts and strand breaks that when not sufficiently repaired by cellular machinery may result in genomic changes that favor selection of subclones. Studying the effects of such environmental exposures on the genome and identifying exposure dosages and key developmental windows that are most susceptible to the effects of these environmental exposures will be critical in uncovering the basic mechanisms of genome instability as well as factors promoting clonal expansion.

Evolving Methods for Detecting Genetic Mosaicism

While many human diseases and traits have been linked to detectable genetic mosaicism, until recently, it has been daunting to detect the presence of at least two subpopulations of the same cell type with distinct genotypes. Moreover, the precision of detecting proportions of cells has been challenging, especially when the fractions approach zero or one. Initially, detection methods used light microscopy to look at karyotypes, but with the advent of molecular probes and the polymerase chain reaction (PCR) technology it is possible to cleverly design allele-specific molecular probes, and more recently, high-throughput array platforms that survey the genome, either with single nucleotide polymorphism (SNP) microarrays or next generation sequencing platforms (Figure 2). In the near future, next generation sequencing, which can generate deep coverage per base, should accelerate characterization of the scope and size of genetic mosaicism, across a range of tissue sources. Still, the larger challenge resides in developing stable and accurate analytical algorithms that can detect mosaic events, even small single base alterations with both high sensitivity and specificity.

A critical biological insight in the field of ‘mosaicism’ occurred when Mary Lyon first proposed that X-chromosomes can be randomly rendered transcriptionally inactive³⁷. Lyon’s hypothesis was based on phenotypic observation of mouse genetic models as well as

light microscopy photographs of Feulgen stained tissues³⁸. Although true genetic mosaicism should not be mistaken with the resulting phenotypic mosaicism of Lyon's X inactivation, the light microscopy used to detect inactivated X chromosomes first demonstrated the value of cytogenetic techniques to detect the coexistence of distinct chromosomal karyotypes. Initially, using karyotype analysis of cells in metaphase, investigators could determine the overall number of chromosomes and whether large portions were missing or exchanged across cell types, especially in metastatic cancer cells where chromosome counts and size may be altered from germline DNA. The advent of chromosomal banding techniques improved the capacity to distinguish between chromosomes of similar size by producing a series of consistent landmarks along the length of metaphase chromosomes^{39–42}. These staining patterns were particularly useful for recognition of segmental alterations of individual chromosomes and the investigation of translocations across chromosomes, particularly in cancer.

The efficient detection of sub-chromosomal events improved with the introduction of molecular probes designed to hybridize specific genetic sequences, which could be either radioisotope labeled or fluorescent labeled assays. Fluorescent in situ hybridization (FISH) uses fluorescently labeled probes to detect and localize the presence or absence of specific DNA sequences by fluorescence microscopy⁴³. Multicolor spectral karyotyping (SKY), an extension of FISH, enables mixes of multiple, spectrally overlapping probes to essentially "paint" each chromosome in a different color⁴⁴. While these technologies can determine the presence or absence of chromosomal regions and can detect translocations and other rearrangements, the methods are not optimized for high-throughput application. Since they examine one cell in metaphase at a time, clonal mosaic events could be detected by these methods, but the events had to be large (>5 Mb). Moreover, counting single cells in an unbiased manner created a formidable barrier for accurate assessment of the mosaic proportions.

Alternatives for copy number variation detection were developed using the polymerase chain reaction (PCR) technologies⁴⁵, such as multiplex ligation-dependent probe amplification (MLPA)⁴⁶, quantitative real-time PCR (qPCR)⁴⁷, and Sanger sequencing⁴⁸ to investigate mosaic copy loss and gain. These have been used to confirm the findings with SNP microarrays¹⁵ and accordingly, estimate mosaic proportions.

Array-based technologies have expanded upon PCR based approaches by constructing arrays with probes that span the genome. Array comparative genomic hybridization (array-CGH) uses test and reference DNA samples to compare relative copy number using thousands of probes across the genome⁴⁹. Since the development of array-CGH, arrays using complementary DNA (cDNA) and oligonucleotides have been used extensively to detect copy number variation⁵⁰. Array-CGH has been an effective tool for assessing copy number changes, particularly in cancer, but also in diagnostic tests for mosaicism, especially in the evaluation of pediatric developmental disorders^{51, 52}.

A major development that accelerated the field of detecting genetic mosaicism emerged from the rigorous quality metrics of single nucleotide polymorphism (SNP) genotyping arrays, primarily used for genome-wide association studies (GWAS)¹⁷. The careful evaluation of GWAS SNP chip data can now provide a high resolution virtual karyotype for cellular DNA¹⁷. Remarkably, the application of rigorous GWAS quality control procedures identified chromosomes in healthy controls with aberrant copy number states, which would have been filtered from GWAS analysis. Several groups have developed approaches to examine existing array data from GWAS and specifically examine the SNP log₂ R ratio (LRR) and B allele frequency (BAF)^{13–15, 19–21}. The LRR and BAF are combined to detect genomic locations of chromosomal loss and gain as well as copy neutral events such as

uniparental disomy (Figure 3). Using SNP genotyping arrays, the means for a linear string of BAF measurements can be used to estimate the proportion of cells in the population carrying the clonal mosaic event. The method is relatively high-throughput, allowing investigators to estimate the overall frequency of large structural mosaicism (e.g., events greater than 2 Mb in size) in both apparently healthy individuals as well as in a wide range of diseases. Due to the background noise of SNP microarrays, the present segmentation algorithms are undergoing further development to detect events less than 2 Mb in size as well as proportions near the extremes, namely above or below 5 and 95%^{13, 14}.

The Spectrum of Genetic Mosaicism

As studies of detectable genetic mosaicism accrue, a number of important features have begun to emerge, some of which begin to point towards underlying mechanisms as well as possible patterns, in both healthy and disease populations. The observation that detectable mosaic events cluster in the same genomic region in unrelated individuals and in rare cases in families suggests the disruption of key pathways, particularly related to chromosome integrity or DNA repair. While not all instances of genetic mosaicism have been demonstrated to result in phenotypically detectable changes, a wide spectrum of phenotypic manifestations has been associated in clinical series as well as in large population-based studies (Table 1). The potential molecular consequences associated with mosaic alterations at these sites suggest there could be physical properties (e.g., sequence motifs or tertiary structure) of the surrounding DNA that contribute to events that can lead to stable subclones. For genetic mosaicism to be detected, positive selection of the subpopulation has to occur—and in some instances, could be related to the copy number change of a region and its resident gene(s) that influence chromosomal integrity or DNA repair.

Mosaic events can affect an entire chromosome as well as single bases, but the latter are substantially harder to detect with current algorithms applied to available technologies. Several examples exist in the literature in which entire chromosomes are duplicated or deleted early in development and the resulting aberrant cell then clonally expands and its progeny are maintained as a stable subset of the adult cellular population. Turner's syndrome is an example in which loss or partial loss of an X chromosome results in variable phenotypes, which predominantly include short stature, lymphedema, broad chest, low hairline, and reproductive sterility⁵³. Mosaic Turner's syndrome, present in approximately 15 percent of cases, is the result of a mixture of X loss. Patients with a lower fraction of aberrant X chromosomes manifest a less severe phenotype; the region of X involved may also influence the phenotypic expression⁵⁴. Mosaic Down's syndrome is another instance in which an individual's cellular population is mosaic for an entire chromosome, namely, trisomy of chromosome 21²³. Approximately 2–4 percent of individuals with Down's syndrome exhibit mosaicism of trisomy 21; interestingly, these cases tend to have less intellectual and developmental impairment than individuals with complete trisomy 21²³. Mosaic chromosomal aneuploidies have also been reported in other autosomes in both phenotypically normal and abnormal individuals. Examples include partial tetrasomy on chromosome 5 and 9 and mosaic trisomies of chromosome 13^{55–60}.

The spectrum of mosaic events is also important in cancer and pre-cancerous states. For instance, it is possible to see evidence of a large structural alteration that tracks with a malignant disease, such as a deletion of region on chromosome 13q14, in a subset of the B cell population well before the diagnosis of chronic lymphocytic leukemia (CLL)⁶¹. Approximately 50% of CLL cases harbor a deletion in this region, which is suspected to function as a tumor suppressor⁶². Using SNP microarrays in large population-based studies, it is possible to detect 13q14 deletion mosaicism perhaps as much as 14 years prior to diagnosis, using prospective cohort studies¹³.

As early as the 1980's, investigators hypothesized that constitutionally present mutations, which would otherwise be lethal, could persist in an embryo in a mosaic state if the percent of the cellular population involved was sufficiently small enough as to not affect the overall developmental fate of the embryo^{63, 64}. Such mosaic individuals could survive to adulthood, but with a variety of phenotypic effects related to the type of mutation and percentage of cells harboring the mutation. McCune-Albright syndrome, a disease involving endocrine dysfunction, abnormal bone growth, and unilateral café-au-lait skin pigmentation^{65, 66}, is an example of a potentially embryonic lethal genetic mutation that can be seen in mosaic state. Activating mosaic post-zygotic mutations in *GNAS1* occurring early in development are observed in distinct tissue types in cases^{67, 68}. The *GNAS1* gene is involved in G-protein signaling and similar constitutive activating mutations have been identified in pituitary and thyroid tumors^{69, 70}. Proteus syndrome is a disfiguring disease characterized by irregular overgrowths, susceptibility to tumor development, and hyperplasia of multiple tissue types⁷¹. Recent evidence indicates a mosaic activating mutation in the *AKT1* oncogene, c. 49G→A, can account for a majority of cases⁷². It has been hypothesized that a constitutional mutation may be lethal, but in the mosaic state, the *AKT1* mutation results in observed clinical overgrowth. *AKT1* expression has a wide tissue distribution, is implicated in cellular growth and survival, and AKT activity has been associated with many other disease states such as cancer, cardiovascular disease, diabetes, and neurological disorders⁷³. This finding is remarkable because mutations in the *AKT1* oncogene have been implicated as drivers of lung cancer and neuroblastoma, yet in the tissue specific mosaic state, do not result in aggressive cancers^{74–78}.

Recently, two related skeletal disorders not traditionally considered to be classical Mendelian diseases (e.g. due to highly penetrant germline mutations), Ollier disease and Maffucci syndrome, have been explained by the presence of mosaic, single-base mutations in oncogenes of great interest. Ollier disease is a rare, sporadic disorder in which benign cartilaginous tumors develop close to the cartilage growth plate, and when accompanied by soft tissue hemangiomas, the disorder is classified as Maffucci syndrome. Molecular similarities between the associated cartilaginous tumors and growth plate tissue suggest that the tumors arise as a consequence of a growth or differentiation disorder. Recent evidence indicates that clinical manifestation can be linked to early post-zygotic mutations in the *IDH1*, c.394C→T and c.395G→A, and *IDH2* genes, c.516G→C^{79, 80}. *IDH* mutations are characterized by neomorphic enzymatic activity, which catalyzes the reduction of α-ketoglutarate to 2-hydroxyglutarate and results in a hypermethylation phenotype⁸¹. *IDH1* and *IDH2* mutations are found in up to 85% of high-grade gliomas, approximately 15% of AMLs, and in other cancers recently characterized using next generation sequencing approaches⁸².

The RAS proteins are a group of GTPases that play key roles in signal transduction growth pathways. Activating mutations in *RAS* genes have been observed in approximately 30% or more of cancers⁸³. Unexpectedly, mosaic point mutations that affect *HRAS*, *KRAS*, and *NRAS* have been associated with a wide range of phenotypic conditions including nevus sebaceous⁸⁴, Schimmelpenning syndrome⁸⁴, keratinocytic epidermal nevus⁸⁵, and keratinocytic epidermal nevus syndrome^{86, 87}. Some have suggested that the mosaic *RAS* states could be early precursors of neoplastic changes⁸⁸. Mosaic *RAS* mutations can be observed in the skin and occasionally in internal organs. Further characterization of the prevalence and pathologic relevance of mosaic *RAS* mutations is needed to better understand how these mosaic events may contribute to increased cancer risk as well as other chronic conditions. Depending on the type of *RAS* mutation and tissue involved, monitoring individuals could prove to be an important preventative strategy for detecting early precursor stages for *RAS*-associated cancers.

Genetic Mosaicism in Genome-wide Association Studies

Population-based studies of human clonal mosaicism are beginning to emerge from the evaluation of BAF and LRR in GWAS studies. So far, large scale surveys of GWAS inadvertently uncovered a higher than expected distribution of large structural events detected as mosaic events when scanning the genomes of tens of thousands of individuals in search of detectable clonal mosaic events^{13–15}. Serendipitously, unexpected deviations in GWAS quality control analyses led to the detection of two or more distinct subpopulations of cells with distinct, large structural alterations (> 2 Mb) within an individual. Since high quality SNP data can be derived from either blood or buccal cell DNA, it appears that detectable mosaicism is present in distinct compartments of the hematopoietic system as well as epithelial cells from buccal swabs. Validation of detected events as well as high concordance of detection algorithms across studies have demonstrated that using data from GWAS is a robust approach for detecting clonal mosaicism; the early study of Rodriguez et al. showed that classical techniques, namely, short tandem repeat (STR) assays, multiplication-dependent probe amplification (MLPA), and fluorescent in situ hybridization (FISH), could effectively confirm the mosaic state in either blood or buccal DNA¹⁵. The current algorithms for survey analyses of GWAS SNP microarray data are stable for mosaic events greater than 2 Mb in size and the same algorithms can be applied to regions of 300 kb in some instances¹⁴.

A surprising observation from these studies is that detectable clonal mosaicism of large events (e.g., greater than 2 Mb) has a much higher frequency than previously expected. Based on initial surveys drawing GWAS data from many studies, albeit of different epidemiological design, a few patterns have begun to emerge. These require further exploration in additional studies, preferably prospective studies in which blood/buccal retrieval occurs well before the diagnoses of interest. A remarkable finding is that detectable, large structural mosaicism increases with age. Frequency estimates are less than 0.5% for individuals under 50 years of age and then rapidly increase with age to around 2–3% for individuals 75 or older^{13, 14}. Interestingly, the age-related increase was also observed in a study of monozygotic twins²¹. Studies on telomere attrition also note reduced telomere length with age, which together with mosaicism data suggests the human genome may experience increased rates of somatic mutation or reduced genome maintenance capacity throughout the ageing process that may ultimately lead to immunosenescence and thus could be risk factors for chronic diseases, such as cancer and diabetes⁸⁹.

While not conclusively established, evidence from GWAS data indicates detectable genetic mosaicism may be enriched in one sex compared to the other. Males were found to have an overall significantly higher frequency of mosaic abnormalities when compared to females¹³. This trend was observed both among cancer cases and cancer-free individuals as well as consistently across non-sex specific cancer types. A higher frequency of mosaic abnormalities in males could be due to sex-specific differences in mutation and recombination rates⁹⁰ or differences in exposure to environmental mutagens according to gender. Additional studies are needed to further substantiate these sex-specific differences in frequency of mosaicism.

Data from GWAS also suggest individuals harboring detected mosaic events may be at elevated risk for developing hematologic malignancies^{13, 14, 20}. Here, it is plausible that the 'apparent' mosaicism could actually represent early detection of one or more structural alterations seen in one or more distinct malignant subclones^{91, 92}. Strong associations have been observed between mosaic status and incident hematological cancer with hazard ratio estimates that could exceed 10, but further studies are required to refine the estimates^{13, 14}. Additionally, many of the mosaic abnormalities detected in GWAS samples drawn from

either cancer-free controls or individuals with a solid tumor (e.g., prostate, lung or bladder cancer) are characteristic of those seen in hematologic cancers with one third of all duplications and deletions found in hematologic cancers having greater than 80% overlap with detected GWAS mosaic regions that harbor multiple overlapping anomalies¹⁴.

Preliminary analyses across a large number of GWAS studies raises the specter that mosaicism could be associated with the risk for solid tumors, though the precise mechanism is elusive¹³. It is not clear whether these findings reflect more global changes in individual genome stability or the actual regions are critical for specific tissue-specific carcinogenesis. So far, significant associations were observed in stratified analyses of lung and kidney cancer. The effect estimates for these solid tumor associations are substantially attenuated compared to those suggested for hematological malignancies, but nonetheless suggest clonal mosaicism and solid tumors may share common genetic features that could be instrumental in risk prediction and cancer prevention.

Recently, an association between detectable genetic mosaicism and type 2 diabetes (T2D) has been reported⁹³. Individuals with T2D have a higher frequency of detectable genetic mosaic events with an elevated estimated odds ratio, perhaps as high as 5, but further studies are required to verify the association. Interestingly, the observed association was stronger among non-obese individuals, and individuals with clonal mosaicism were observed to have an increased risk for micro- and macro-vascular complications. The proposed age-accelerating effects of T2D may be responsible for the development of these genetic mosaic events, although further replication is needed to substantiate this hypothesis. Such observations suggest genetic mosaic events may affect the severity of T2D phenotype and could ultimately modify a diabetic's risk of developing certain cancers associated with clonal mosaic events.

As catalogs of large structural events from over 100,000 subjects were analyzed with GWAS SNP microarrays, it has emerged that mosaic events cluster across particular chromosomal regions (Table 2). A substantial proportion of mosaic gains can include an entire chromosome, while mosaic losses usually occur in interstitial regions of the chromosomes that are not connected to telomeres or centromeres. Copy neutral events, on the other hand, are almost always seen involving a telomere with an approximate 2:1 preference for the q chromosomal arm over the p arm. These observations suggest distinct mechanisms underlie the generation and possible selection of different types of large structural mosaic events. It is plausible that events localized to telomeric chromosomal ends could be a consequence of errors in recombination and may play a role in the formation of copy neutral mosaic events. Chromosomal replications or deletions on the other hand could be derived from aberrant DNA repair mechanisms, which may be involved in the formation of mosaic gains of entire chromosomes and interstitial losses.

Detected mosaic events also appear to cluster in specific genomic regions. For example, mosaic deletions are common at 13q and 20q, and mosaic copy-neutral loss of heterozygosity is common at 9p and 14q. Many of the mosaic anomalies detected are characteristic of those found in hematologic cancers, such as 13q14 deletion in CLL, and represent common deleted genes previously associated with cancers⁹⁴. It is notable that in the GWAS surveys, there are more individuals with 13q14 deletion mosaicism than can be explained by the expected cumulative lifetime risk for hematologic malignancies in the study sets. Moreover, the number exceeds the estimates for MBL and CLL combined, suggesting that a small fraction of the population tolerates this event without evidence for development of cancer. It is also notable that there is no difference between the distribution of 13q14 deletion mosaicism between cases with solid (non-hematological) cancers and cancer free controls, many of whom have been observed for years in prospective cohorts.

The genomic distribution of mosaic event breakpoints also hints at underlying mechanisms that may contribute to the formation of somatic alterations at these sites. For example, breakpoints for 13q14 deletions appear to be enriched for regions that have been shown to contain DNase hypersensitivity peaks, transcription factor clusters, and hydroxyl radical cleavage sites in the Encyclopedia of DNA Elements (ENCODE) project^{95, 96}. This suggests boundaries of these deletions are often located in transcriptionally active areas of open chromatin that are accessible to potential mutagens that could cleave DNA.

Future Directions

Recent advances in analysis of SNP microarray technology coupled with more efficient bioinformatics algorithms of copy number state in germline DNA have uncovered an unanticipated frequency of large structural mosaic events in the ageing population. These findings perhaps represent part of a continuum, initially described in the investigation of unusual pediatric genetic diseases. Moreover, the unexpected findings in healthy ageing populations now raise new questions concerning the frequency and effect of smaller events predicted to be more common in the population. It is notable that these large structural events involve many genes, including oncogenes, tumor suppressor genes, and other drivers of cancer, yet in the mosaic state they are tolerated in otherwise healthy, cancer free controls. Many questions remain regarding how these events initially occur and why they are clonally selected.

The recent detection of large structural genetic mosaicism raises important questions concerning the distribution of the sizes, frequencies, and proportions of events that are tolerated. The advent of next-generation sequencing technologies promises to improve detection and provide more accurate surveys of the extent of genetic mosaicism, particularly as single cell sequencing is applied to multiple tissue types. By the use of massively parallel sequencing methods, high read depth can be obtained at a fraction of the cost of first generation chain-termination methods⁹⁷. Very high read depth affords high-resolution mapping down to the base pair for copy-number alterations, copy-neutral events, and chromosomal translocations. Fine-scale resolution will be helpful in determining which genes are involved and, at the same time, map the actual breakpoints, affording an opportunity to investigate the stability of distinct regions across the genome. The high cost and developing methodology has limited the use of next-generation sequencing in detecting clonal mosaic events, but as prices continue to drop greater numbers of samples will be sequenced and made available for analysis. Additional refinements in computational algorithms and advances in computational power are needed to distinguish specific events from the background noise of NGS as well as to partition out mutational heterogeneity that may lead to spurious associations of mosaic abnormalities with disease subtypes⁹⁸. Classical laboratory techniques, such as short tandem repeat (STR) assays, multiplication-dependent probe amplification (MLPA), and fluorescent in situ hybridization (FISH) may need to be used to confirm the accuracy of the new algorithms. Sequencing-based copy-number detection algorithms already exist and can be applied to large datasets^{97, 99–102}. Repurposing and modifying these algorithms for the detection of clonal mosaic events will be instrumental in gaining a more detailed understanding of where these mosaic events occur and elucidate unique characteristics that define their respective breakpoints.

A new generation of genomic studies in longitudinal studies is required to understand the evolution of genetic mosaicism over a lifetime. Prior consensus among geneticists was that the human genome remained static over time, but current findings suggest that human genomes may undergo alterations, particularly with ageing. In particular, it will be critical to understand, not only the determinants of generating somatic events, but also the factors that permit clonal expansion. Cell-sorting technologies will enable determination of rates and

types of cells that can tolerate mosaicism. Prospective studies should provide chronological insights into mosaic events that when coupled with biomarker and clinical data could be instrumental in developing effective screening tools capable of detecting early precursor stages of cancer as well as other complex diseases, such as diabetes, cardiovascular, and neurodegenerative disorders.

Large structural alterations resulting in detectable genetic mosaic events could represent the tip of the iceberg in that they illustrate the toleration of errors in DNA replication. Moreover, they could provide clues to better understand tissue specific genomic stability and the role it could contribute to complex diseases, such as cancer, diabetes, and perhaps neurodegenerative diseases. Surveys of genetic mosaicism are still preliminary but have uncovered an opportunity to understand early events in carcinogenesis and cancer susceptibility. An advanced understanding of genetic mosaicism as it relates to the ageing human genome and human disease has the potential to greatly improve classification of disease subtypes and ultimately result in more effective prevention strategies and treatments.

References

1. Zlotnikoff M. A Human Mosaic. *J Heredity*. 1945; 36:163–8.
2. Strachan, T.; Read, AP.; Strachan, T. Human molecular genetics. 4. New York: Garland Science; 2011. p. xxvp. 781
3. Larsson NG. Somatic mitochondrial DNA mutations in mammalian aging. *Annu Rev Biochem*. 2010; 79:683–706. [PubMed: 20350166]
4. Bianchi NO, Bianchi MS, Richard SM. Mitochondrial genome instability in human cancers. *Mutat Res*. 2001; 488:9–23. [PubMed: 11223402]
5. Taylor RW, Turnbull DM. Mitochondrial DNA mutations in human disease. *Nat Rev Genet*. 2005; 6:389–402. [PubMed: 15861210]
6. Schwab AL, Tuohy TM, Condie M, et al. Gonadal mosaicism and familial adenomatous polyposis. *Fam Cancer*. 2008; 7:173–7. [PubMed: 18026870]
7. Davidson S, Leshanski L, Rennert G, et al. Maternal mosaicism for a second mutational event--a novel deletion--in a familial adenomatous polyposis family harboring a new germ-line mutation in the alternatively spliced-exon 9 region of APC. *Hum Mutat*. 2002; 19:83–4. [PubMed: 11754114]
8. Niu DM, Huang JY, Li HY, et al. Paternal gonadal mosaicism of NIPBL mutation in a father of siblings with Cornelia de Lange syndrome. *Prenat Diagn*. 2006; 26:1054–7. [PubMed: 16958143]
9. Depienne C, Arzimanoglou A, Trouillard O, et al. Parental mosaicism can cause recurrent transmission of SCN1A mutations associated with severe myoclonic epilepsy of infancy. *Hum Mutat*. 2006; 27:389. [PubMed: 16541393]
10. Gennaro E, Santorelli FM, Bertini E, et al. Somatic and germline mosaicisms in severe myoclonic epilepsy of infancy. *Biochem Biophys Res Commun*. 2006; 341:489–93. [PubMed: 16430863]
11. Lund AM, Nicholls AC, Schwartz M, et al. Parental mosaicism and autosomal dominant mutations causing structural abnormalities of collagen I are frequent in families with osteogenesis imperfecta type III/IV. *Acta Paediatr*. 1997; 86:711–8. [PubMed: 9240878]
12. Somprasit C, Aguinaga M, Cisneros PL, et al. Paternal gonadal mosaicism detected in a couple with recurrent abortions undergoing PGD: FISH analysis of sperm nuclei proves valuable. *Reprod Biomed Online*. 2004; 9:225–30. [PubMed: 15333256]
13. Jacobs KB, Yeager M, Zhou W, et al. Detectable clonal mosaicism and its relationship to aging and cancer. *Nat Genet*. 2012 Jun;44:651–8. [PubMed: 22561519]
14. Laurie CC, Laurie CA, Rice K, et al. Detectable clonal mosaicism from birth to old age and its relationship to cancer. *Nat Genet*. 2012; 44:642–50. [PubMed: 22561516]
15. Rodriguez-Santiago B, Malats N, Rothman N, 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]

16. Youssoufian H, Pyeritz RE. Mechanisms and consequences of somatic mosaicism in humans. *Nat Rev Genet.* 2002; 3:748–58. [PubMed: 12360233]
17. Wang K, Bucan M. Copy Number Variation Detection via High-Density SNP Genotyping. *CSH Protoc.* 2008 pdb top46.
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. Gonzalez JR, Rodriguez-Santiago B, Caceres A, 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]
20. Schick UM, McDavid A, Crane PK, et al. Confirmation of the reported association of clonal chromosomal mosaicism with an increased risk of incident hematologic cancer. *PLoS One.* 2013; 8:e59823. [PubMed: 23533652]
21. Forsberg LA, Rasi C, Razzaghian HR, 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]
22. Kehrer-Sawatzki H, Cooper DN. Mosaicism in sporadic neurofibromatosis type 1: variations on a theme common to other hereditary cancer syndromes? *J Med Genet.* 2008; 45:622–31. [PubMed: 18511569]
23. Papavassiliou P, York TP, Gursoy N, et al. The phenotype of persons having mosaicism for trisomy 21/Down syndrome reflects the percentage of trisomic cells present in different tissues. *Am J Med Genet A.* 2009; 149A:573–83. [PubMed: 19291777]
24. Grubeck-Loebenstein B. Changes in the aging immune system. *Biologicals.* 1997 Jun; 25(2):205–8. [PubMed: 9236053]
25. Geiger H, de Haan G, Florian MC. The ageing haematopoietic stem cell compartment. *Nat Rev Immunol.* 2013; 13:376–89. [PubMed: 23584423]
26. Garraway LA, Lander ES. Lessons from the cancer genome. *Cell.* 2013 Mar 28; 153(1):17–37. [PubMed: 23540688]
27. Frank SA, Nowak MA. Cell biology: Developmental predisposition to cancer. *Nature.* 2003; 422:494. [PubMed: 12673242]
28. Snape K, Hanks S, Ruark E, et al. Mutations in CEP57 cause mosaic variegated aneuploidy syndrome. *Nat Genet.* 2011; 43:527–9. [PubMed: 21552266]
29. Rio Frio T, Lavoie J, Hamel N, et al. Homozygous BUB1B mutation and susceptibility to gastrointestinal neoplasia. *N Engl J Med.* 2010; 363:2628–37. [PubMed: 21190457]
30. Mouret S, Baudouin C, Charveron M, Favier A, Cadet J, Douki T. Cyclobutane pyrimidine dimers are predominant DNA lesions in whole human skin exposed to UVA radiation. *Proc Natl Acad Sci U S A.* 2006; 103:13765–70. [PubMed: 16954188]
31. Scientific Committee on the Effects of Atomic Radiation U. Effects of Ionising Radiation. 2006
32. Rossner P Jr, Svecova V, Schmuczerova J, et al. Analysis of biomarkers in a Czech population exposed to heavy air pollution. Part I: bulky DNA adducts. *Mutagenesis.* 2013; 28:89–95. [PubMed: 23047913]
33. Jantzen K, Roursgaard M, Desler C, Loft S, Rasmussen LJ, Moller P. Oxidative damage to DNA by diesel exhaust particle exposure in co-cultures of human lung epithelial cells and macrophages. *Mutagenesis.* 2012; 27:693–701. [PubMed: 22869610]
34. Phillips DH, Venitt S. DNA and protein adducts in human tissues resulting from exposure to tobacco smoke. *Int J Cancer.* 2012; 131:2733–53. [PubMed: 22961407]
35. Tung EW, Philbrook NA, Macdonald KD, Winn LM. DNA double-strand breaks and DNA recombination in benzene metabolite-induced genotoxicity. *Toxicol Sci.* 2012 Apr; 126(2):569–77. [PubMed: 22247006]
36. Costa C, Costa S, Silva S, et al. DNA damage and susceptibility assessment in industrial workers exposed to styrene. *J Toxicol Environ Health A.* 2012; 75:735–46. [PubMed: 22788361]
37. Lyon MF. Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature.* 1961; 190:372–3. [PubMed: 13764598]
38. Ohno S, Hauschka TS. Allocycl of the X-chromosome in tumors and normal tissues. *Cancer Res.* 1960; 20:541–5. [PubMed: 14428472]

39. Caspersson T, Farber S, Foley GE, et al. Chemical differentiation along metaphase chromosomes. *Exp Cell Res.* 1968; 49:219–22. [PubMed: 5640698]
40. Caspersson T, Zech L, Johansson C. Analysis of human metaphase chromosome set by aid of DNA-binding fluorescent agents. *Exp Cell Res.* 1970; 62:490–2. [PubMed: 5495462]
41. Latt SA. Optical studies of metaphase chromosome organization. *Annu Rev Biophys Bioeng.* 1976; 5:1–37. [PubMed: 60079]
42. Schreck RR, Distèche CM. Chromosome banding techniques. *Curr Protoc Hum Genet.* 2001 May. Chapter 4(Unit4):2. [PubMed: 18428280]
43. Langer-Safer PR, Levine M, Ward DC. Immunological method for mapping genes on *Drosophila* polytene chromosomes. *Proc Natl Acad Sci U S A.* 1982; 79:4381–5. [PubMed: 6812046]
44. Schrock E, du Manoir S, Veldman T, et al. Multicolor spectral karyotyping of human chromosomes. *Science.* 1996; 273:494–7. [PubMed: 8662537]
45. Saiki RK, Gelfand DH, Stoffel S, et al. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science.* 1988; 239:487–91. [PubMed: 2448875]
46. Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res.* 2002; 30:e57. [PubMed: 12060695]
47. Porcher C, Malinge MC, Picat C, Grandchamp B. A simplified method for determination of specific DNA or RNA copy number using quantitative PCR and an automatic DNA sequencer. *Biotechniques.* 1992; 13:106–14. [PubMed: 1503761]
48. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A.* 1977; 74:5463–7. [PubMed: 271968]
49. Pinkel D, Segraves R, Sudar D, et al. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet.* 1998; 20:207–11. [PubMed: 9771718]
50. Carter NP. Methods and strategies for analyzing copy number variation using DNA microarrays. *Nat Genet.* 2007; 39:S16–21. [PubMed: 17597776]
51. DeScipio C, Spinner NB, Kaur M, et al. Fine-mapping subtelomeric deletions and duplications by comparative genomic hybridization in 42 individuals. *Am J Med Genet A.* 2008; 146:730–9. [PubMed: 18257100]
52. Rabin KR, Man TK, Yu A, et al. Clinical utility of array comparative genomic hybridization for detection of chromosomal abnormalities in pediatric acute lymphoblastic leukemia. *Pediatr Blood Cancer.* 2008; 51:171–7. [PubMed: 18253961]
53. HHT. A syndrome of infantilism, congenital webbed neck, and cubitus valgus. *Endocrinology.* 1938; 23:566–74.
54. Sybert VP, McCauley E. Turner's syndrome. *N Engl J Med.* 2004; 351:1227–38. [PubMed: 15371580]
55. Papoulidis I, Kontodiu M, Tzimina M, et al. Tetrasomy 9p mosaicism associated with a normal phenotype in two cases. *Cytogenet Genome Res.* 2012; 136:237–41. [PubMed: 22487875]
56. McAuliffe F, Winsor EJ, Chitayat D. Tetrasomy 9p mosaicism associated with a normal phenotype. *Fetal Diagn Ther.* 2005; 20:219–22. [PubMed: 15824502]
57. Di Giacomo MC, Susca FC, Resta N, Bukvic N, Vimercati A, Guanti G. Trisomy 13 mosaicism in a phenotypically normal child: description of cytogenetic and clinical findings from early pregnancy beyond 2 years of age. *Am J Med Genet A.* 2007; 143:518–20. [PubMed: 17163529]
58. Aypar E, Yildirim MS, Sert A, Ciftci I, Odabas D. A girl with metopic synostosis and trisomy 13 mosaicism: case report and review of the literature. *Am J Med Genet A.* 2011; 155:638–41. [PubMed: 21344634]
59. Lorda-Sanchez I, Villa A, Urioste M, et al. Tetrasomy 5p mosaicism due to an extra i(5p) in a severely affected girl. *Am J Med Genet.* 1997; 68:481–4. [PubMed: 9021026]
60. Venci A, Bettio D. Tetrasomy 5p mosaicism due to an additional isochromosome 5p in a man with normal phenotype. *Am J Med Genet A.* 2009; 149:2889–91. [PubMed: 19938082]
61. Dohner H, Stilgenbauer S, Benner A, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med.* 2000; 343:1910–6. [PubMed: 11136261]

62. Stilgenbauer S, Nickolenko J, Wilhelm J, 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; 16:1891–7. [PubMed: 9583687]
63. Happle R. Cutaneous manifestation of lethal genes. *Hum Genet*. 1986; 72:280. [PubMed: 3957353]
64. Happle R. The McCune-Albright syndrome: a lethal gene surviving by mosaicism. *Clin Genet*. 1986; 29:321–4. [PubMed: 3720010]
65. DJM. Osteitis fibrosa cystica: the case of a nine-year-old girl who also exhibits precocious puberty, multiple pigmentation of the skin and hyperthyroidism. *Am J Dis Child*. 1936; 52:743–4.
66. Albright FBA, Hampton AO, Smith P. Syndrome characterized by osteitis fibrosa disseminata, areas, of pigmentation, and endocrine dysfunction, with precocious puberty in females: report of 5 cases. *N Engl J Med*. 1937; 216:727–46.
67. Weinstein LS, Shenker A, Gejman PV, Merino MJ, Friedman E, Spiegel AM. Activating mutations of the stimulatory G protein in the McCune-Albright syndrome. *N Engl J Med*. 1991; 325:1688–95. [PubMed: 1944469]
68. Schwindinger WF, Francomano CA, Levine MA. Identification of a mutation in the gene encoding the alpha subunit of the stimulatory G protein of adenylyl cyclase in McCune-Albright syndrome. *Proc Natl Acad Sci U S A*. 1992; 89:5152–6. [PubMed: 1594625]
69. Collins MT, Sarlis NJ, Merino MJ, et al. Thyroid carcinoma in the McCune-Albright syndrome: contributory role of activating Gs alpha mutations. *J Clin Endocrinol Metab*. 2003; 88:4413–7. [PubMed: 12970318]
70. Landis CA, Masters SB, Spada A, Pace AM, Bourne HR, Vallar L. GTPase inhibiting mutations activate the alpha chain of Gs and stimulate adenylyl cyclase in human pituitary tumours. *Nature*. 1989; 340:692–6. [PubMed: 2549426]
71. Biesecker LG. The multifaceted challenges of Proteus syndrome. *JAMA*. 2001; 285:2240–3. [PubMed: 11325326]
72. Lindhurst MJ, Sapp JC, Teer JK, et al. A mosaic activating mutation in AKT1 associated with the Proteus syndrome. *N Engl J Med*. 2011; 365:611–9. [PubMed: 21793738]
73. Hers I, Vincent EE, Tavare JM. Akt signalling in health and disease. *Cell Signal*. 2011; 23:1515–27. [PubMed: 21620960]
74. Janoueix-Lerosey I, Lequin D, Brugieres L, et al. Somatic and germline activating mutations of the ALK kinase receptor in neuroblastoma. *Nature*. 2008; 455:967–70. [PubMed: 18923523]
75. George RE, Sanda T, Hanna M, et al. Activating mutations in ALK provide a therapeutic target in neuroblastoma. *Nature*. 2008; 455:975–8. [PubMed: 18923525]
76. Chen Y, Takita J, Choi YL, et al. Oncogenic mutations of ALK kinase in neuroblastoma. *Nature*. 2008; 455:971–4. [PubMed: 18923524]
77. Choi YL, Takeuchi K, Soda M, et al. Identification of novel isoforms of the EML4-ALK transforming gene in non-small cell lung cancer. *Cancer Res*. 2008; 68:4971–6. [PubMed: 18593892]
78. Soda M, Choi YL, Enomoto M, et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature*. 2007; 448:561–6. [PubMed: 17625570]
79. Amary MF, Damato S, Halai D, et al. Ollier disease and Maffucci syndrome are caused by somatic mosaic mutations of IDH1 and IDH2. *Nat Genet*. 2011; 43:1262–5. [PubMed: 22057236]
80. Pansuriya TC, van Eijk R, d'Adamo P, 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; 43:1256–61. [PubMed: 22057234]
81. Figueroa ME, Abdel-Wahab O, Lu C, et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell*. 2010; 18:553–67. [PubMed: 21130701]
82. Dang L, Jin S, Su SM. IDH mutations in glioma and acute myeloid leukemia. *Trends Mol Med*. 2010; 16:387–97. [PubMed: 20692206]
83. Schubbert S, Shannon K, Bollag G. Hyperactive Ras in developmental disorders and cancer. *Nat Rev Cancer*. 2007; 7:295–308. [PubMed: 17384584]

84. Groesser L, Herschberger E, Ruetten A, et al. Postzygotic HRAS and KRAS mutations cause nevus sebaceous and Schimmelpenning syndrome. *Nat Genet.* 2012; 44:783–7. [PubMed: 22683711]
85. Hafner C, Toll A, Gantner S, et al. Keratinocytic epidermal nevi are associated with mosaic RAS mutations. *J Med Genet.* 2012 Apr;49:249–53. [PubMed: 22499344]
86. Bourdeaut F, Herault A, Gentien D, et al. Mosaicism for oncogenic G12D KRAS mutation associated with epidermal nevus, polycystic kidneys and rhabdomyosarcoma. *J Med Genet.* 2010; 47:859–62. [PubMed: 20805368]
87. Hafner C, Toll A, Real FX. HRAS mutation mosaicism causing urothelial cancer and epidermal nevus. *N Engl J Med.* 2011; 365:1940–2. [PubMed: 22087699]
88. Hafner C, Groesser L. Mosaic RASopathies. *Cell Cycle.* 2013; 12:43–50. [PubMed: 23255105]
89. Sahin E, Depinho RA. Linking functional decline of telomeres, mitochondria and stem cells during ageing. *Nature.* 2010; 464:520–8. [PubMed: 20336134]
90. Hedrick PW. Sex: differences in mutation, recombination, selection, gene flow, and genetic drift. *Evolution.* 2007; 61:2750–71. [PubMed: 17976181]
91. Ding L, Ley TJ, Larson DE, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature.* 2012; 481:506–10. [PubMed: 22237025]
92. Walter MJ, Shen D, Ding L, et al. Clonal architecture of secondary acute myeloid leukemia. *N Engl J Med.* 2012; 366:1090–8. [PubMed: 22417201]
93. Bonnefond A, Skrobek B, Lobbens S, et al. Association between large detectable clonal mosaicism and type 2 diabetes with vascular complications. *Nat Genet.* 2013; 45:1040–3. [PubMed: 23852171]
94. Forbes SA, Bindal N, Bamford S, et al. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* 2011; 39:D945–50. [PubMed: 20952405]
95. Consortium EP, Bernstein BE, Birney E, et al. An integrated encyclopedia of DNA elements in the human genome. *Nature.* 2012; 489:57–74. [PubMed: 22955616]
96. Chanock S. Toward mapping the biology of the genome. *Genome Res.* 2012 Sep; 22(9):1612–5. [PubMed: 22955973]
97. Chiang DY, Getz G, Jaffe DB, et al. High-resolution mapping of copy-number alterations with massively parallel sequencing. *Nat Methods.* 2009; 6:99–103. [PubMed: 19043412]
98. Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature.* 2013; 499:214–8. [PubMed: 23770567]
99. Sathirapongsasuti JF, Lee H, Horst BA, et al. Exome sequencing-based copy-number variation and loss of heterozygosity detection: ExomeCNV. *Bioinformatics.* 2011; 27:2648–54. [PubMed: 21828086]
100. Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* 2012; 22:568–76. [PubMed: 22300766]
101. Yoon S, Xuan Z, Makarov V, Ye K, Sebat J. Sensitive and accurate detection of copy number variants using read depth of coverage. *Genome Res.* 2009; 19:1586–92. [PubMed: 19657104]
102. Abyzov A, Urban AE, Snyder M, Gerstein M. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res.* 2011; 21:974–84. [PubMed: 21324876]
103. Morris SW, Kirstein MN, Valentine MB, et al. Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. *Science.* 1994; 263:1281–4. [PubMed: 8122112]
104. Rikova K, Guo A, Zeng Q, et al. Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell.* 2007; 131:1190–203. [PubMed: 18083107]

Figure 1.

Two example homologous chromosomes are depicted in the nucleus of a diploid individual for a cellular population. The underlying process of mosaicism is depicted where an individual with a normal karyotype carrying a complete paternal (white) and maternal (black) chromosome in all cells undergoes a mutational event that clonally expands to become a detectable subset of the total cellular population. This can result in a mosaic loss on the maternal chromosome in only a portion of the cellular population (orange), a mosaic gain on the maternal chromosome (yellow), or a mosaic copy neutral loss of heterozygosity (black) such as an acquired uniparental disomy.

Figure 2.

Illustrations of technologies used to detect mosaicism. (A) Karyotype as visualized using light microscopy and staining to visualize banding patterns. (B) “Painted” karyotype using SKY tagged probes. (C) Sanger sequencing reads from radioisotope labeled gel (left) and fluorescent labeled chromatogram (right). (D) Array CGH with comparative copy loss and gain on three example chromosomes. (E) BAF and LRR values extracted from SNP array plotting an example copy gain on a chromosome. (F) Next-generation sequencing reads generated for a chromosomal region of interest.

Figure 3.

Simulated plots of GWAS SNP array data depicting (A) a normal chromosome, (B) a chromosome exhibiting an acquired mosaic uniparental disomy, (C) chromosome with a mosaic gain, and (D) a chromosome with a mosaic loss. Log 2 R ratios (LRR) are points plotted in black on the left y axis and give an indication of copy number (gain>0, loss<0). B allele frequencies (BAF) are points plotted in red on the right y axis and are indicators of allelic imbalance. Each SNP on the genotyping array has a LRR (black) and BAF (red) point plotted at the SNP's chromosomal position along the x axis.

Table 1

Mosaic Disorders	Tissue Affected	Chromosome	References
Anaplastic Large-Cell Lymphoma	blood	t(2;5)	52
Chronic Lymphocytic Leukemia	blood	13q14	49
Down's Syndrome	multiple	21	42
Keratinocytic Epidermal Nevus	skin	11p	70
Maffucci Syndrome	multiple	2q,15q	65,66
McCune-Albright Syndrome	multiple	20q	58,59
Nevus Sebaceous	skin	11p,12p	69
Non-Small Cell Lung Cancer	lung	2p	53
Ollier Disease	connective	2q,15q	65,66
Proteus Syndrome	multiple	14q	63
Schimmelpenning Syndrome	multiple	11p,12p	69
Turner's Syndrome	multiple	X	41

Count and frequency of mosaic chromosomal events by event type and location combined from Laurie et al. and Jacobs et al.

Table 2

Event location	Mosaic chromosome count				Mosaic chromosome frequency (%)			
	Gain	Loss	CN LOH	Mixed	Total	Gain	Loss	CN LOH
Chromosome	70	11	45	5	131	6.7	1.0	4.3
Telomeric p	16	24	144	1	185	1.5	2.3	13.7
Telomeric q	22	26	232	0	280	2.1	2.5	22.1
Interstitial	46	379	2	1	428	4.4	36.1	0.2
Span centromere	2	1	2	0	5	0.2	0.1	0.2
Complex	1	5	9	7	22	0.1	0.5	0.9
Total	157	446	434	14	1051	14.9	42.4	41.3

CN LOH, copy-neutral loss of heterozygosity