BMC Medicine

Mutations in normal tissues—some diagnostic and clinical implications

Clare Fiala[1] and Eleftherios P. Diamandis[1,2,3*]

Abstract

Background: It has long been known that mutations are at the core of many diseases, most notably cancer. Mutational analysis of tissues and fluids is useful for cancer and other disease diagnosis and management.

Main body: The prevailing cancer development hypothesis posits that cancer originates from mutations in cancer-driving genes that accumulate in tissues over time. These mutations then confer special characteristics to cancer cells, known as the hallmarks of cancer. Mutations in specific driver genes can lead to the formation of cancerous subclones and mutation risk increases with age. New research has revealed an unexpectedly large number of mutations in normal tissues; these findings could have significant implications to the understanding of the pathobiology of cancer and for disease diagnosis and therapy. Here, we discuss how the prevalence of mutations in normal tissues provides novel and relevant insights about clonal development in cancer and other diseases. Specifically, this review will focus on discussing mutations in normal tissues in the context of developing specific, circulating tumor DNA (ctDNA) tests for cancer, and evaluating clonal hematopoiesis as a predictor of blood cancers and cardiovascular pathology, as well as their implications to the phenomena of neural mosaicism in the context of Alzheimer's disease.

Conclusions: In view of these new findings, the fundamental differences between the accumulation of genetic alterations in healthy, aging tissues compared to cancer and cardiovascular or neural diseases will need to be better delineated in the future.

Keywords: Normal tissues, Mutations, Circulating tumor DNA, Driver mutations, Passenger mutations

Background

In 1990, Fearon and Vogelstein presented a hypothesis, whereby cancer development is the result of accumulating mutations in some critical genes, occurring over years [1]. Hanahan and Weinberg provided more granularity to this model by identifying eight hallmarks of cancer that are a consequence of genetic instability [2]. However, very recent and continually expanding data suggest that normal tissues, and particularly aging

tissues, accumulate a surprisingly large number of mutations that are frequently inconsequential, prompting some to speculate that this is the "new normal." Several studies have found an approximate range of human germline mutation rates: $1.0–1.2 \times 10^{-8}$ per nucleotide per generation and up to 10% of these mutations are deleterious to some degree [3, 4].

The search to delineate the pathway of tissue mutations from harmless to pathological marks a new frontier for cellular and molecular research—beyond simply detecting and classifying genetic variations but determining whether these mutations will lead to disease. However, this challenge is complicated by various exogenous and endogenous factors. The transition from a benign clone to tumorigenesis or the accumulation of genetic

* Correspondence: Eleftherios.diamandis@sinaihealthsystem.ca
[1]Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, ON, Canada
[2]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada
Full list of author information is available at the end of the article

mutations that predispose an individual to Alzheimer's or cardiovascular disease is intimately influenced by the patient's immune system, the environment, other co-morbidities, and related epigenetic changes. For example, smoking is well known to augment mutagenesis in lung tissues [5], and cirrhosis (often caused by alcoholism or viral hepatitis) can also lead to increased driver mutations and clonal expansion in liver tissue [6]. Even water and oxygen, fundamentally necessary for life, can significantly mutate DNA. However, the percentage of oncogenic mutations in cells of normal tissues may be quite variable and may need highly sensitive assays in order to be detected or excluded. Clearly, more experimentation is necessary in order to make more robust predictions.
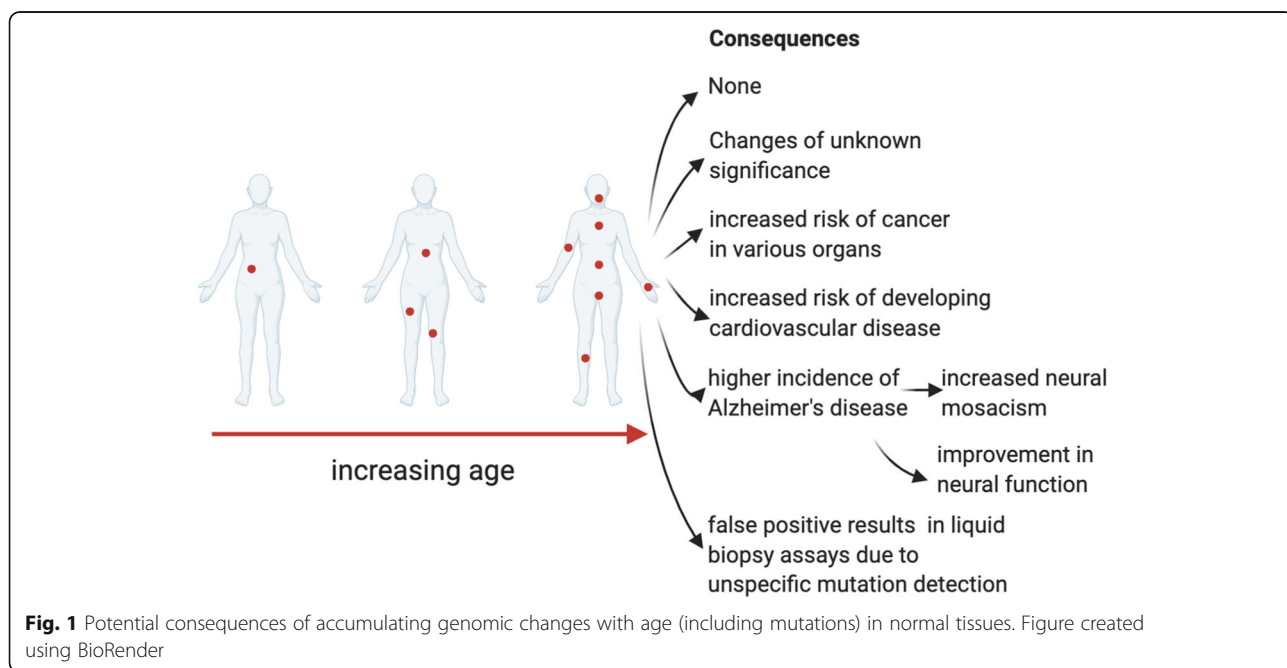
Further investigation into the mechanisms underlying mutations in normal tissues will likely uncover new opportunities for early detection as well as an increased understanding and even development of reference intervals to delineate which genetic variations are pathological and which are not. A visual summary of the potential consequences of the age-related accumulation of genomic changes in normal tissues can be found in Fig. 1.

Clonal development, circulating tumor DNA, and cancer detection

Biopsies are a cornerstone for cancer treatment, and in the last decade, tumor genetic sequencing has been extensively used to guide treatment selection and prognosis determination.

The inherent invasiveness of surgical resection and biopsy means these procedures are performed relatively infrequently. Previously, genome sequencing technologies have been focused on determining cancer-related mutations from patients diagnosed clinically or through screening. However, now, there is great interest in turning these technologies and knowledge towards early detection which necessitates knowing exactly which mutations signify cancer and which mutations constitute benign passenger changes.

Cells are thought to release their DNA into the circulation through apoptosis and necrosis; this genetic material is known as circulating free DNA (cfDNA). For example, during pregnancy, some fetal cells and fetal DNA escape into the maternal circulation which can be used to test the fetus for genetic abnormalities. Cancer cells also release DNA, denoted circulating tumor DNA (ctDNA). New tests (known widely as liquid biopsies) involve extracting and sequencing this ctDNA from a blood sample, a highly desirable alternative to tissue biopsies. Not only are liquid biopsies minimally invasive, ctDNA tests can be repeated frequently to analyze cancer progression or regression throughout treatment. Many investigators have used this technique to monitor cancer burden, treatment response, and tumor evolution in patients already diagnosed with the disease. The US Food and Drug Administration (FDA) has already approved a ctDNA test involving monitoring EGFR (epidermal growth factor receptor) mutations in the serum of non-small cell lung cancer patients to determine eligibility for the chemotherapy drug erlotinib [7],

Fig. 1 Potential consequences of accumulating genomic changes with age (including mutations) in normal tissues. Figure created using BioRender

and another test to determine ctDNA residue after surgery has been granted breakthrough device status [8].

However, now, investigators are endeavoring to develop a blood test for early detection of cancer in healthy individuals without symptoms or tumors invisible on imaging. The tests would involve detecting and sequencing the tiny amounts of ctDNA from the tumors that could be less than 2 mm in diameter (4 mm is the approximate detection limit of imaging for tumor detection) [9]. There are several challenges involved in developing such a test, including the high possibility of sampling error due to low mutant allele fraction caused by the miniscule tumor volume [10]. Moreover, there are also methodological difficulties involved with detecting somatic mutations in normal tissues that may only occur in small cell populations. Accuracy is crucial because false detection of cancer can lead to unnecessary procedures and surgery as well as significant patient anxiety and cost to the health care system. The concerns about specificity are most relevant to this review. A ctDNA test's success for early diagnosis is predicated on determining which mutations indicate cancer and which are associated with benign diseases, a task that is becoming more complicated as new research suggests that significant mutagenicity in normal tissues is the norm, not the exception.

Pre-cancerous clonal development
A recent study by Yizhak et al. examining more than 6700 histologically healthy samples spanning 29 types of tissues from almost 500 individuals underscores this phenomenon [11]. Employing analyses of RNA sequence databases to identify somatic mutations, the team found many mutations generally associated with cancer as well as macroscopic clones (population of mutated cells descended from a common ancestor, visible without a microscope) in tissues without any cancer pathology. Specifically, 95% of the study participants had at least one tissue with macroscopic clonal cell populations, 40% had one large mutational clone, and 5% of tissues had 5 or more large mutational clones. Moreover, 33% of the healthy participants carried a gene mutation connected to cancer. Unsurprisingly, the number of these mutations increased with age and with cell proliferation rates; high cell turnover tissues such as the skin and esophagus had increased levels of mutations compared to the less proliferative brain and muscle cells [11]. As expected, the tissues most affected by endogenous factors such as sun exposure and smoking—skin and lung—had a higher mutational burden, reinforcing the significant connection between environmental factors and mutagenicity.

The specificity of ctDNA analysis for early cancer diagnosis is complicated by the phenomenon of mosaicism: age-related genetic mutations occurring in a subset of an individual's cells. Consequently, even if a sensitive assay is developed to reliably detect mutant cfDNA at low concentrations in the circulation, investigators must be able to consistently determine whether the mutations are cause for concern. Sun-exposed human skin is the most highly mutated tissue—a 2015 study found driver mutations in about 25% of skin cells from healthy participants [12]. Another study employed mathematical modeling of tumor development and determined that more than half of the mutations found in some cancers were present in the normal cell where the cancer began—the founder clone. Thus, these somatic mutations found would have been present even if the tumor had not formed [13].

A recent study of more than 2000 colonic crypts further illustrates this challenge [14]. The authors found that approximately 1% of normal colorectal crypts in middle age individuals contain a driver mutation, yet incidence of colorectal cancer is much lower, suggesting that these tumors are rare incidences of normal mutational processes in aging colorectal epithelium. These discoveries dovetail with clinical finding from other investigations. Only 5% of people develop colorectal cancer across their lifespan [15] even though adenomas are found via routine colonoscopy in approximately 40% of individuals over 70 years old [16]. In fact, there is a less than 1 in 375,000 chance that the crypt micro neoplasms detected in the previous study would become macroscopically detectable adenomas and less than 1 in 3 million chance that they become carcinomas [14].

Mutations in esophageal carcinoma
Other research with esophageal squamous cell carcinoma has further elucidated the underlying process influencing whether a mutation will lead to the formation of cancerous clones. Martincorena et al. analyzed 844 upper esophageal tissue samples from 9 healthy donors (aged 20 to 75 years) [17]. Most importantly, the team discovered that cancer-associated genes were mutated far more frequently in healthy esophagus samples than in healthy skin. In fact, many of these mutated genes were under strong positive selection, leading to increased cell proliferation and the consequent formation of cell clones in the esophagus. Unsurprisingly, there was an increased number of mutations in both overall and in driver genes as well as a larger clone size on average in the tissue samples from older participants compared to younger participants. Another important finding was though mutations connected to benign clonal expansion may appear frequently in cancer cell genomes, these particular mutations may not necessarily contribute to carcinogenesis because these tissues already have a naturally high mutation frequency.

Furthermore, an analysis by Higa and DeGregori demonstrated that even classic "oncogenic drivers" like *NOTCH1* might not often contribute to carcinogenesis [18].

Even though the esophagus does not receive the same mutagenic exposure to ultraviolet radiation compared to skin and has a tenth of the mutations compared to skin, it is still exposed to a significant number of mutagens through a human's varied diet [19]. Consequently, Martinconera et al. found that some genes were frequently mutated in both skin and esophageal tissue, particularly *NOTCH1, TP53, NOTCH2, FAT1, NOTCH3*, and *ARID1A* [17]. Most of the abnormalities were missense and protein-truncating mutations though CpG islands were also heavily mutated [17].

Previous investigations reported mutations in *TP53* in more than 90% of esophageal squamous cell carcinoma (OSCC) cases [20]. Conversely, Martincorena et al. found *NOTCH1* to be the most commonly mutated gene in healthy esophageal tissue while *TP53* mutations were far less prevalent in their (healthy) samples [17]. Even though Marincorena's team detected a significant number of mutations in cancer-associated genes, their participants' tissues still appeared healthy. Notably, the rate of incidence of esophageal cancer is less than 0.01% of the general population [21], far less than the incidence suggested by the number of *NOTCH1* and *TP53* mutations observed by Marincorena et al. in healthy tissues [21].

Significantly, these results were echoed in a similar study by Yokoyama et al. which used both normal and cancerous esophageal tissue samples [22]. Age-related mutational signatures were most prevalent in the healthy samples whereas mutational signatures associated with the cytidine deaminase *APOBEC* (an mRNA modifying enzyme) or alcohol were most commonly found in cancer samples [22]. The studies with esophageal tissue provide compelling new information about the age-related accumulation of mutations and their transition to cancer. They also show the value of determining which specific gene mutations or mutational signatures most often lead to particular malignancies, crucial information for developing accurate ctDNA liquid biopsies for diagnosis as well as other related genetic tests.

Mutations in the female reproductive system

A study with paired peritoneal fluid and blood samples from women without cancer detected mutations in *TP53* at low frequencies (< 0.01%) in 16 out the 17 participants [23]. Moreover, the non-synonymous mutations were enriched, undergoing positive selection. Another investigation focused on endometriosis, a painful though benign pathology of the female reproductive system which rarely leads to cancer [24]. Yet, researchers found mutations in some important proto-oncogene driver genes

with a high mutant allele fraction including *KRAS, PIK3CA, PTEN, ARID1A*, and *TP53* in 19 out 24 patients' endometriotic lesions [24]. Notably, these women had no sign of neoplasia; however, some samples had mutant allele fraction (MAF) as high as 40%, showing the extensive prevalence of these mutations [24]. A similar study with uterine lavage fluid samples from healthy women echoed these results with more than half of 95 healthy women having cancerous driver mutations, positively correlating mutations in cancer-causing genes with age and postmenopausal status [25]. Finally, a recent study by Moore et al. on putative oncogenic driver mutations in normal endometrium found that its mutational burden increased at approximately 29 base substitutions each year [26]. Significantly, the particular mutational composition of the epithelium varied widely among the 28 women studied.

Mutations in other solid tissues

TP53 mutations have been reported in histologically normal oral, bronchial, bladder, and esophageal epithelial tissues [27]. *KRAS* mutations have also been detected in normal tissues adjacent to colorectal and lung cancers [28]. As expected, copy number variations are especially abundant in rapidly dividing tissues, particularly with genes involved in growth regulation [29]. Comparable mutations were also discovered in healthy eyelid skin, removed during blepharectomy [12]. Another study examined kidney, fat, and muscle cells from donors of different ages [30]. The investigations identified a mutation-prone cell type in the kidney while also delineating an age-related decline in DNA repair capacity. Similar work on mutational processes has also been conducted in the lung. One important finding was that following smoking cessation cells with a lower mutation burden are replenished since the lung is no longer exposed to mutagenic carcinogens [5].

A fascinating study showed a similar number of mutations in adult stem cells from small intestine, colon, and liver tissues even though there are very different incidences of these cancer in the general population [31]. Liver cancer and colon cancer are approximately 4 and 10 times, respectively, more prevalent in the USA than small intestine cancer [32–34], highlighting that the number of mutations is not necessarily a faultless indicator of cancer and there are likely tissue-dependent and external factors that influence tumor development.

The extensive range of studies of mutations in normal tissues highlights the specificity challenges relevant to ctDNA test development and mutation classification, spanning a range of cancers and tissue types [35]. Moreover, many of the mutations are present in multiple types of cancer, which would make it even more difficult to determine the tumor's tissue of origin after a positive

ctDNA test. A complex net of biochemical, epigenetic, and environmental factors likely influences whether mutations remain harmless or a person develops cancer. Consequently, there is a need for additional studies to delineate these mechanisms, determining which particular genes are more likely to be mutated for specific cancers, and which mutations are simply hallmarks of old age. In the meantime, the presence of mutations in normal tissues remains a major caveat to the development and specificity of a ctDNA test for the diagnosis of cancer.

Clonal hematopoiesis: hematological malignancies and cardiovascular disease

Clonal hematopoiesis, the formation of a genetically distinct population of blood cells, is another example of mutations in normal tissues. Clonal hematopoiesis is thought to precede many hematologic cancers, including acute myeloid leukemia (AML) and chronic lymphocytic leukemia [36–38]. For example, a recent study of peripheral blood cells obtained from 95 patients a mean of 6.3 years before AML diagnosis found that they had higher variant allele frequencies and greater clonal expansion compared to controls [39]. However, putative driver mutations connected to age-related clonal hematopoiesis (ARCH-PD) were not found exclusively in the pre-AML patients. 36.7% of controls had observable ARCH-PD compared to 73.4% of patients. Thirty-nine percent of patients above age 50 in the pre-AML cohort had a driver mutation with a variant allele frequency greater than 10%, in contrast to only 4% of controls. A landmark study of 12,380 patients elucidated some of the clinical implications of clonal hematopoiesis [40]. The researchers found clonal hematopoiesis increases with age, observing related somatic mutations in 10% of participants older than age 65. However, rates of clonal hematopoiesis dropped to less than 1% for participants younger than 50. *DNMT3A*, *ASXL1*, *TET2*, *PPMD1*, and *JAK2* were determined to be the most commonly mutated genes in the cohort's blood cells, with the number of mutations increasing with age [40].

Even though clonal hematopoiesis was a strong risk factor for hematologic cancers, with 42% of the hematologic cancers arising in people with clonal hematopoiesis, a significant portion of the cohort identified to have clonal hematopoiesis did not develop cancer over a 2- to 7-year follow-up period [40]. These individuals had clonal hematopoiesis of indeterminate potential, referring to expanded blood cell clones in the absence of other abnormalities. As a result, testing for clonal hematopoiesis is not necessarily a sensitive predictor of cancer, and it is not clear what factors precipitate the transition from clonal hematopoiesis to cancer.

A study of neonate cord blood samples found that around 1% had *TEL-AML1* and *AML1-ETO* gene fusions, which have been correlated with increased chance of developing leukemia [41]. However, longitudinal studies with children carrying such mutations would be useful to determine if these mutations actually predict leukemia incidence and what exogenous factors predict disease onset. Another large study involving nested case-control analyses of 4726 participants with heart disease and 3529 controls linked clonal hematopoiesis with a two times greater incidence of atherosclerosis, a condition characterized by arterial buildups of fat and cholesterol [42]. *TET2* is one of the most commonly mutated genes related to clonal hematopoiesis, and mice that received bone marrow grafts from homozygous or heterozygous *TET2* knockout mice developed larger atherosclerotic lesions compared to the control mice [42]. Yet again, not all the study participants with clonal hematopoiesis developed this disease. Further investigation in mice and in vitro models revealed that TET2 deficiency mediated by clonal hematopoiesis accelerated atherosclerosis, a disease mediated by chronic inflammation, by increasing macrophage proliferation [43, 44]. While factors such as diet and exercise likely play an important role, there is still a need to elucidate the biochemical processes underlying why some people with clonal hematopoiesis developed atherosclerosis and others did not. Clonal hematopoiesis has also been linked to greater risk of coronary heart disease, ischemic stroke, and overall mortality; however, more longitudinal studies are required to confirm these preliminary associations [36].

Even though clonal hematopoiesis testing for cancer and cardiac disease is attractive because it is essentially non-invasive, it still has a very long way to go before it reaches the clinic. Clonal hematopoiesis can also complicate chemotherapy as different subsets of cells have slightly different genomes, making them resistant to particular cytotoxic agents. In fact, recent research using blood samples from approximately 50,000 cancer-free individuals found that positive selection for advantageous clonal mutations, not neutral genetic drift, constituted the major influencer on clonal hematopoiesis and clone fitness, providing further insight into the complex role of clonal hematopoiesis and disease development [45, 46]. While the aforementioned studies represent a good start, more work and sequencing methods with greater sensitivity are needed to precisely determine which of the associated mutations or groups of mutations in blood cells are carcinogenic and pathological, or simply hallmarks of aging normal tissues.

Mosaicism in the brain and Alzheimer's disease

Increased mosaicism—genetic changes/mutations in a subset of an organism's cells—has also been found in neurons, both from healthy people and those with

Alzheimer's disease [47]. The exact mechanism is unknown; however, recent research suggests that amyloid-β precursor protein (*APP*) mRNA variants became permanently incorporated into the genome of neurons due to breaks in neuronal DNA and the action of reverse transcriptase. The investigators termed these novel genome sequences genomic complementary DNAs (gencDNAs) and identified them in the neurons of individuals with Alzheimer's disease as well as in those who were cognitively healthy [47]. Neurons are thought to be more susceptible to the incorporation of gencDNA due to their long lifespans and the fact that they rarely divide. Moreover, this mutational mechanism introduces another risk factor for individuals who suffered traumatic brain injury—the ensuing breaks in DNA could provide opportunities for Alzheimer-related gencDNA formation [47].

As expected, the number of gencDNAs increased with age, and the researchers found 10 times more gencDNAs in neurons from people with Alzheimer's disease than in cognitively healthy participants [47]. Some of these *APP* mRNA variants are thought to translate into cytotoxic proteins, further underscoring the role of mosaicism in Alzheimer-related neuron death. It is important to note that this is still very preliminary work and the authors were not able to conclude whether the accumulation of gencDNAs is a consequence or a cause of Alzheimer's disease [47]. Regardless, this phenomenon of DNA shuffling in somatic cells has only ever been reported in the immune system to generate antibody diversity, and its presence in the central nervous system is a fascinating discovery relating mutations found in normal tissues.

In fact, some mutations occurring in normal brain tissue may even be beneficial. The authors speculate that APP gencDNAs may contribute to synaptic proteome diversity, producing different genes that may improve neuron function [47]. Different clusters of neurons may be modified for selective activities, reducing the need for alternative splicing or RNA modification. gencDNAs may also be important for normal brain function, playing a role in plasticity, learning and memory, or giving neurons the ability to remember and employ genetic variants beyond the wild-type genes [47].

The discovery of the increase of gencDNAs in sporadic Alzheimer's disease and its reliance on reverse transcriptase may even offer a promising therapeutic approach. It is known that people who have human immunodeficiency virus and have been on long-term anti-retroviral therapy including reverse transcriptase inhibitors have a reduced incidence of Alzheimer's disease compared to the rest of the population [48, 49]. Thus, reverse transcriptase inhibitors may be a novel, promising avenue for Alzheimer's disease drug development.

The Lee et al. paper represents one of the earliest large-scale studies comparing mutations in normal and diseased tissues, and it already generated promising insights about mechanism, diagnosis, and potential therapy [47]. Another paper published around the same time performed comprehensive single nucleotide variant (SNV) identification using single-cell whole-genome sequencing [50]. The investigators took 161 neurons from the prefrontal cortex and hippocampus of fifteen healthy individuals of all ages as well as nine individuals who had early-onset neurodegeneration due to the genetic disorders of Cockayne syndrome and Xeroderma pigmentosum. While SNVs increased linearly in both cohorts, they were more common in the individuals with neurodegenerative disease. A higher rate of mutation was also observed in the hippocampus, and the authors posit that this may be relevant for Alzheimer's disease.

Altogether, the investigators delineated three mutational signatures [50]. The first signature, characterized by C>T and T>C mutations, increased with age, in a clock-like fashion. Perhaps related to early development, the second signature featured primarily C>T mutations and did not change with age. Finally, the third signature had C>A variants and was connected with oxidative DNA damage.

Some research has also been conducted surrounding aneuploidy (chromosome copy number variations) in the brain which occurs as a result of errors in mitosis [51]. While aneuploidy is highly common in cancer, surprisingly high levels of aneuploidy have also been observed in the normal brain [52, 53]. As expected, its frequency increases with age though the brain is still able to function normally despite elevated amounts of mosaic aneuploidy [54, 55]. At abnormally high levels, aneuploidy can lead to significant declines in neural function, including Alzheimer's and Parkinson's disease [56, 57].

While intriguing, this complex field is still in its infancy and the connections between neural mosaicism and normal function or neurodegeneration are unclear. More research is needed in order to conclusively determine what genetic changes are not cause for concern or those that are associated with devastating conditions including Alzheimer's disease and Cockayne syndrome.

Conclusions

Studying the transition of tissues from normal to diseased may enable early diagnosis as well as mechanisms for targeted pharmacology. Even though this branch of science is in its infancy, the scientific dogma is changing with the awareness that heavily mutated tissues are not necessarily hallmarks of disease but are actually the new normal [58]. There is greater awareness that an exclusive focus on studying tumor or diseased tissue limits our ability to learn about precancerous or pre-disease states.

Moreover, advances in sequencing techniques and bio-informatics may uncover more insights into this rapidly developing field, such as the prevalence of positive versus negative selection in cancer development [59]. Consequently, more studies and projects are addressing this dearth of information, studying genetic events and clonal properties from a wide variety of tissue samples and participants [21].

One such initiative is the National Cancer Institute's Pre-Cancer Genome Atlas [60], a large-scale collaboration to create a comprehensive database of histological and genomic/proteomic data from premalignant lesions which will produce new opportunities for early detection and risk stratification. Advances in nucleic acid sequencing including massively parallel as well as methylation-based techniques will also be helpful to accurately detect mutated or altered DNA in both normal and diseased tissues. Conversely, a new approach employing patterns in fragment length to distinguish cfDNA from ctDNA has also yielded some promising results for early cancer diagnosis though this method may have important limitations [61–63].

It is clear that it is unfeasible to diagnose cancer simply by examining mutations. Some recent studies go beyond mutations to consider chromatin organization. One recent study by Liu et al. studied DNA methylation as a further hallmark of a cancer cell, beyond simply mutation [64]. Another major investigation combined the detection of oncogenic ctDNA mutations, protein markers, and PET-CT (positron emission tomography—computerized tomography) in order to increase the accuracy of cancer detection while specifying the tissue of origin [65].

Other advances may come from creatively utilizing existing databases. For example, Garcia-Nieto et al. employed RNA-sequencing data to study mutagenesis in 7500 healthy tissue samples, coming from 36 distinct tissues [66]. Through examining the transcriptome, they were able to delineate somatic tissue evolution through examining the positive and negative selection for various mutations. Apart from some common cancer mutations that were highly enriched (like missense mutations in IDH2 or excess synonymous mutations in MP2K1), the team found strong negative selection acting on the majority missense and nonsense mutations.

A collection of articles published in Nature in May 2020 by the Genome Aggregation Database (gnomAD) Consortium constitutes another major step forward. Karczewski et al. examined predicted loss of function (pLoF) mutations using whole-genome sequencing (WGS) and exome sequencing of a diverse population of 141,456 individuals [67]. By comparing the expected and actual frequency of pLoF variants, the authors were able to classify the genes based on pLoF tolerance (from no effect to lethal). Another paper in the collection examined differences in tolerance of genes to pLoF, demonstrating that many of these variants appear in exons which minimize their effect [68]. The other papers discussed the potential utility of the gnomAD database in drug discovery [69] and conducted a large-scale analysis of structural variants in protein coding and non-coding regions [70].

Investigating mutations in normal tissues requires enormous amounts of data, particularly as these changes can be rare and subtle. It is particularly important to include genetic data from individuals from underrepresented racial and ethnic backgrounds as well as people of all ages. Continued innovations stemming from big data and precision medicine will be highly valuable; they will enable very high-resolution studies of normal tissues and diseased tissues as well as new biomarker discovery to facilitate better outcomes for individuals with cancer, Alzheimer's disease, cardiac pathologies and beyond.

Author details
[1]Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, ON, Canada. [2]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada. [3]Department of Clinical Biochemistry, Mount Sinai Hospital and University Health Network, 60 Murray St. Box 32, Floor 6, Rm L6-201, Toronto, ON M5T 3L9, Canada.

References
1. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell. 1990;61:759–67.
2. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144:646–74.

3. Kondrashov A. Genetics: the rate of human mutation. Nature. 2012;488: 467–8.
4. Zhang L, Vijg J. Somatic mutagenesis in mammals and its implications for human disease and aging. Annu Rev Genet. 2018;52:397–419.
5. Yoshida K, Gowers KHC, Lee-Six H, Chandrasekharan DP, Coorens T, Maughan EF, et al. Tobacco smoking and somatic mutations in human bronchial epithelium. Nature. 2020;578:266–72.
6. Brunner SF, Roberts ND, Wylie LA, Moore L, Aitken SJ, Davies SE, et al. Somatic mutations and clonal dynamics in healthy and cirrhotic human liver. Nature. 2019;574:538–42.
7. Kwapisz D. The first liquid biopsy test approved. Is it a new era of mutation testing for non-small cell lung cancer? AnnTransl Med. 2017;4:46.
8. Natera. FDA grants breakthrough device designation to Natera's Signatera test. Bloomberg. 2019; https://www.natera.com/press-releases/fda-grants-b.
9. Diamandis EP, Fiala C. Can circulating tumor DNA be used for direct and early stage cancer detection? F1000Research. 2017;6:2129.
10. Fiala C, Diamandis EP. Utility of circulating tumor DNA in cancer diagnostics with emphasis on early detection. BMC Med. 2018;16:166.
11. Yizhak K, Aguet F, Kim J, Hess JM, Kübler K, Grimsby J, et al. RNA sequence analysis reveals macroscopic somatic clonal expansion across normal tissues. Science. 2019;364:eaaw0726.
12. Martincorena I, Roshan A, Gerstung M, Ellis P, Van Loo P, McLaren S, et al. High burden and pervasive positive selection of somatic mutations in normal human skin. Science. 2015;348:880–6.
13. Tomasetti C, Vogelstein B, Parmigiani G. Half or more of the somatic mutations in cancers of self-renewing tissues originate prior to tumor initiation. Proc Natl Acad Sci. 2013;110:1999–2004.
14. Lee-Six H, Olafsson S, Ellis P, Osborne RJ, Sanders MA, Moore L, et al. The landscape of somatic mutation in normal colorectal epithelial cells. Nature. 2019;574:532–7 Springer US.
15. Cancer Research UK. Bowel cancer incidence statistics. Cancer Res. UK. 2014; https://www.cancerresearchuk.org/health-profession.
16. Corley DA, Jensen CD, Marks AR, Zhao WK, de Boer J, Levin TR, et al. Variation of adenoma prevalence by age, sex, race, and colon location in a large population: implications for screening and quality programs. Clin Gastroenterol Hepatol. 2013;11:172–80 Available from: https://linkinghub.elsevier.com/retrieve/pii/S1542356512010518.
17. Martincorena I, Fowler JC, Wabik A, Lawson ARJ, Cagan A, Stratton MR, et al. Somatic mutant clones colonize the human esophagus with age. Science. 2018;362:911–7.
18. Higa KC, DeGregori J. Decoy fitness peaks, tumor suppression, and aging. Aging Cell. 2019;18:e12938.
19. Ciccarelli F. Mutations differ in normal and cancer cells. Nature. 2019;565:301.
20. Kim J, Bowlby R, Mungall AJ, Robertson AG, Odze RD, Cherniack AD, et al. Integrated genomic characterization of oesophageal carcinoma. Nature. 2017;541:169–75.
21. Chanock SJ. The paradox of mutations and cancer. Science. 2018;362:893–4.
22. Yokoyama A, Kakiuchi N, Yoshizato T, Nannya Y, Suzuki H, Takeuchi Y, et al. Age-related remodelling of oesophageal epithelia by mutated cancer drivers. Nature. 2019;565:312–7.
23. Krimmel JD, Schmitt MW, Harrell MI, Agnew KJ, Kennedy SR, Emond MJ, et al. Ultra-deep sequencing detects ovarian cancer cells in peritoneal fluid and reveals somatic TP53 mutations in noncancerous tissues. Proc Natl Acad Sci. 2016;113:6005–10.
24. Anglesio MS, Papadopoulos N, Ayhan A, Nazeran TM, Noë M, Horlings HM, et al. Cancer-associated mutations in endometriosis without cancer. N Engl J Med. 2017;376:1835–48.
25. Nair N, Camacho-Vanegas O, Rykunov D, Dashkoff M, Camacho SC, Schumacher CA, et al. Genomic analysis of uterine lavage fluid detects early endometrial cancers and reveals a prevalent landscape of driver mutations in women without histopathologic evidence of cancer: a prospective cross-sectional study. PLoS Med. 2016;13:e1002206.
26. Schwaederle M, Husain H, Fanta PT, Piccioni DE, Kesari S, Schwab RB, et al. Detection rate of actionable mutations in diverse cancers using a biopsy-free (blood) circulating tumor cell DNA assay. Oncotarget. 2016;7:9707–17.
27. Waridel F, Estreicher A, Bron L, Flaman JM, Fontolliet C, Monnier P, et al. Field cancerisation and polyclonal p53 mutation in the upper aerodigestive tract. Oncogene. 1997;14:163–9.
28. Risques RA, Kennedy SR. Aging and the rise of somatic cancer-associated mutations in normal tissues. PLoS Genet. 2018;14:1–12 Available from: https://doi.org/10.1371/journal.pgen.1007108.
29. O'Huallachain M, Karczewski KJ, Weissman SM, Urban AE, Snyder MP. Extensive genetic variation in somatic human tissues. Proc Natl Acad Sci. 2012;109:18018–23.
30. Franco I, Helgadottir HT, Moggio A, Larsson M, Vrtačnik P, Johansson A, et al. Whole genome DNA sequencing provides an atlas of somatic mutagenesis in healthy human cells and identifies a tumor-prone cell type. Genome Biol. 2019;20:285.
31. Blokzijl F, De Ligt J, Jager M, Sasselli V, Roerink S, Sasaki N, et al. Tissue-specific mutation accumulation in human adult stem cells during life. Nature. 2016;538:260–4.
32. American Cancer Society. Key statistics about liver cancer. 2019. https://www.cancer.org/cancer/liver-cancer/about/w.
33. American Cancer Society. Key statistics for small intestine cancer. 2019. https://www.cancer.org/cancer/small-intestine-canc.
34. American Cancer Society. Key statistics for colorectal cancer. 2019. https://www.cancer.org/cancer/colon-rectal-cancer/.
35. Fiala C, Diamandis EP. Circulating tumor DNA (ctDNA) is not a good proxy for liquid biopsies of tumor tissues for early detection. Clin Chem Lab Med. 2020; EPUB. Available from: https://www.degruyter.com/view/journals/cclm/ahead-of-print/article-10.1515-cclm-2020-0083/article-10.1515-cclm-2020-0083.xml.
36. Shlush LI. Age-related clonal hematopoiesis. Blood. 2018;131:496–504.
37. Steensma DP, Bejar R, Jaiswal S, Lindsley RC, Sekeres MA, Hasserjian RP, et al. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. Blood. 2015;126:9–16.
38. Jan M, Ebert BL, Jaiswal S. Clonal hematopoiesis. Semin Hematol. 2017;54:43–50.
39. Abelson S, Collord G, Ng SWK, Weissbrod O, Mendelson Cohen N, Niemeyer E, et al. Prediction of acute myeloid leukaemia risk in healthy individuals. Nature. 2018;559:400–4 Springer US.
40. Genovese G, Kähler AK, Handsaker RE, Lindberg J, Rose SA, Bakhoum SF, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. N Engl J Med. 2014;371:2477–87.
41. Mori H, Colman SM, Xiao Z, Ford AM, Healy LE, Donaldson C, et al. Chromosome translocations and covert leukemic clones are generated during normal fetal development. Proc Natl Acad Sci. 2002;99:8242–7.
42. Jaiswal S, Natarajan P, Silver AJ, Gibson CJ, Bick AG, Shvartz E, et al. Clonal hematopoiesis and risk of atherosclerotic cardiovascular disease. N Engl J Med. 2017;377:111–21.
43. Luis TC, Wilkinson AC, Beerman I, Jaiswal S, Shlush LI. Biological implications of clonal hematopoiesis. Exp Hematol. 2019;77:1–5.
44. Fuster JJ, MacLauchlan S, Zuriaga MA, Polackal MN, Ostriker AC, Chakraborty R, et al. Clonal hematopoiesis associated with TET2 deficiency accelerates atherosclerosis development in mice. Science. 2017;355:842–7.
45. Watson CJ, Papula AL, Poon GYP, Wong WH, Young AL, Druley TE, et al. The evolutionary dynamics and fitness landscape of clonal hematopoiesis. Science. 2020;367:1449–54.
46. Curtis C. Quantifying mutations in healthy blood. Science. 2020;367:1426–7.
47. Lee MH, Siddoway B, Kaeser GE, Segota I, Rivera R, Romanow WJ, et al. Somatic APP gene recombination in Alzheimer's disease and normal neurons. Nature. 2018;563:639–45.
48. Turner RS, Chadwick M, Horton WA, Simon GL, Jiang X, Esposito G. An individual with human immunodeficiency virus, dementia, and central nervous system amyloid deposition. Alzheimers Dement (Amst). 2016;4:1–5.
49. Centers for Disease Control and Prevention. HIV Surveillance Report. 2016.
50. Lodato MA, Rodin RE, Bohrson CL, Coulter ME, Barton AR, Kwon M, et al. Aging and neurodegeneration are associated with increased mutations in single human neurons. Science. 2018;359:555–9.
51. Andriani GA, Vijg J, Montagna C. Mechanisms and consequences of aneuploidy and chromosome instability in the aging brain. Mech Ageing Dev. 2017;161:19–36.
52. Faggioli F, Vijg J, Montagna C. Chromosomal aneuploidy in the aging brain. Mech Ageing Dev. 2011;132:429–36.
53. McConnell MJ, Lindberg MR, Brennand KJ, Piper JC, Voet T, Cowing-Zitron C, et al. Mosaic copy number variation in human neurons. Science. 2013; 342:632–7.
54. Bushman DM, Chun J. The genomically mosaic brain: aneuploidy and more in neural diversity and disease. Semin Cell Dev Biol. 2013;24:357–69.
55. Westra JW, Peterson SE, Yung YC, Mutoh T, Barral S, Chun J. Aneuploid mosaicism in the developing and adult cerebellar cortex. J Comp Neurol. 2008;507:1944–51.

56. Yurov YB, Vorsanova SG, Liehr T, Kolotii AD, Iourov IY. X chromosome aneuploidy in the Alzheimer's disease brain. Mol Cytogenet. 2014;7:20.

57. Toft M, Ross OA. Copy number variation in Parkinson's disease. Genome Med. 2010;2:62 Available from: http://www.ncbi.nlm.nih.gov/pubmed/20828427.

58. Tomasetti C. Mutated clones are the new normal. Science. 2019;364:938–9.

59. Martincorena I, Raine KM, Gerstung M, Dawson KJ, Haase K, Van Loo P, et al. Universal patterns of selection in cancer and somatic tissues. Cell. 2017;171: 1029 1041.e21.

60. Campbell JD, Mazzilli SA, Reid ME, Dhillon SS, Platero S, Beane J, et al. The case for a pre-cancer genome atlas (PCGA). Cancer Prev Res. 2016;9:199–24.

61. Mouliere F, Chandrananda D, Piskorz AM, Moore EK, Morris J, Ahlborn LB, et al. Enhanced detection of circulating tumor DNA by fragment size analysis. Sci Transl Med. 2018;10:eaat4921.

62. Heitzer E, Speicher MR. One size does not fit all: Size-based plasma DNA diagnostics. Sci Transl Med. 2018;10:eaav3873.

63. Cristiano S, Leal A, Phallen J, Fiksel J, Adleff V, Bruhm DC, et al. Genome-wide cell-free DNA fragmentation in patients with cancer. Nature. 2019;570:385–9.

64. Liu MC, Oxnard GR, Klein EA, Swanton C, Seiden MV, Cummings SR, et al. Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. Ann Oncol. 2020;31:745–59.

65. Lennon AM, Buchanan AH, Kinde I, Warren A, Honushefsky A, Cohain AT, et al. Feasibility of blood testing combined with PET-CT to screen for cancer and guide intervention. Science. 2020;369:eabb9601.

66. García-Nieto PE, Morrison AJ, Fraser HB. The somatic mutation landscape of the human body. Genome Biol. 2019;20:298.

67. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020;581:434–43.

68. Cummings BB, Karczewski KJ, Kosmicki JA, Seaby EG, Watts NA, Singer-Berk M, et al. Transcript expression-aware annotation improves rare variant interpretation. Nature. 2020;581:452–8.

69. Minikel EV, Karczewski KJ, Martin HC, Cummings BB, Whiffin N, Rhodes D, et al. Evaluating drug targets through human loss-of-function genetic variation. Nature. 2020;581:459–64.

70. Collins RL, Brand H, Karczewski KJ, Zhao X, Alföldi J, Francioli LC, et al. A structural variation reference for medical and population genetics. Nature. 2020;581:444–51.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.