DISSERTATION

INTERROGATING AND PREDICTING IONIZING RADIATION EFFECTS ON TELOMERES AND CHROMOSOMES, AND IMPLICATIONS FOR LONG-TERM RISKS FOR HUMAN HEALTH

Submitted by

Jared James Luxton

Graduate Degree Program in Cell and Molecular Biology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2020

Doctoral Committee:

Advisor: Susan M. Bailey

Jennifer G. DeLuca
Juan L. Argueso
Takamitsu A. Kato

ABSTRACT

INTERROGATING AND PREDICTING IONIZING RADIATION EFFECTS ON TELOMERES AND CHROMOSOMES, AND IMPLICATIONS FOR LONG-TERM RISKS FOR HUMAN HEALTH

Space remains the final frontier of exploration. NASA plans to send humans back to the lunar surface by 2024, establishing a permanent lunar base thereafter and embarking humanity on a new era of endeavors. As we approach these grand ventures, the long-term impacts of spaceflight on human health, particularly from chronic exposure to space radiation, remain poorly understood in terms of risks for degenerative disease and cancer. Even less is understood about the effects of deep space flight. These collective unknowns present significant challenges to ensuring and safeguarding astronaut performance during and after spaceflight missions.

Telomeres are the ends of linear chromosomes; they shorten with each cell division and also in response to stress, but can be elongated via lifestyle choices and environmental factors, thus telomeres provide an integrated view of an individual's health status during, and after, given exposures. We hypothesized that longitudinally monitoring telomeres as well as chromosome rearrangements (genomic instability) in astronauts aboard the International Space Station (ISS) throughout spaceflight missions would provide an informative view of astronaut health and long-term risks incurred from spaceflight.

For astronauts during spaceflight aboard the ISS, we observed telomeric and cytogenetic evidence for transient activation of Alternative Lengthening of Telomeres (ALT), concurrent with significant increases in chromosomal rearrangements (inversions) during spaceflight – RNA

sequencing of individuals in extreme environments also yielded indications of transient ALT activation in response to high levels of chronic stress. Telomere elongation concurrent with significant levels of DNA damage has implications for cancer risk. Methods for predicting telomeric responses to given exposures such as ionizing radiation would provide material improvements in projecting disease risk. A high performance machine learning framework for accurate predictions of telomere length is provided here. Our work provides novel indications of transient ALT activation in humans during chronic exposure to extreme environments, as well as a framework for accurately predicting how an individual's telomere length will change throughout.

ACKNOWLEDGEMENTS

To my family: Peggy, you have always supported me with your love and guided me through the difficult times. All the best things that I have ever done have been through you – I will always love you and cherish our lives together. *Mom*, thank you for always loving and believing in me – there have been so many years, but it will be all right – all of it. *Dad*, we are indeed in the future; I like to think that you do come here every now and again. We have all done good things, shown kindness to others, and explored rich, creative, and joyful avenues. *India*, thank you for finding Colorado State University and venturing out here in the first place! You were the anchor and line that brought Peggy and I here. You and Pascal are great together. *Audrey*, I will always love and support you. These past few years have been difficult, but I believe that you and Mom are present-day heroes. *Lee and Candy,* we did not find enough time together but I suspect there will be plenty at other times and places. *Jim*, I will always be so thankful that we have reconnected these past few years. *Jimma and R.P.,* I would like to think that somewhere along the way I picked up the interest in radiation – through diffusion! *Jen, Mankin, and Sally,* you know – the intersection between you all was the impetus for me learning programming, the most interesting parts of this research – and enabled me to start my career. All started with the 2012 Macbook, which this dissertation may have killed.

To my colleagues: Thank you for helping me along this journey. My deepest gratitude to Dr. Susan Bailey for *everything* along the way. Thank you Lynn Taylor for making all the lab work happen and keeping things moving. Special thanks to David Maranon for teaching me microscopy, and to Paula Genik for helping me to advocate for myself.

There are too many others to list here. I have enjoyed my time with everyone – thank you.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

Telomere biology, structure, and maintenance

Telomere biology and structure

Telomeres are repetitive nucleic elements that form the natural ends of linear chromosomes (Figure 1.1). Telomeres address the end replication problem of DNA, serving as a buffer against loss of DNA due to cellular replication, and also protect against inappropriate fusion between DNA termini and inappropriate recognition of the termini as DNA double strand breaks (DSBs) (Arnoult and Karlseder, 2015). Mammalian telomeres are mainly comprised of the repetitive 5'-TTAGGG$_n$-3' DNA motif which is encapsulated in a myriad number of proteins collectively referred to as shelterin (Figure 1.1) (Lange, 2005; Moyzis et al., 1988). While estimates of telomere length vary by species, age, tissue type, and method of measurement, average telomere length ranges from 5-15 kb in humans, 5-25 kb in cats, up to 30 kb in bats, and, reportedly beyond 90 kb in dolphin and 140 kb in lab mice (Calado and Dumitriu, 2013; Heidinger et al., 2012; McKevitt et al., 2003; Okuda et al., 2002; Whittemore et al., 2019). While a consensus on the reasons behind the wide variations in telomere length between species has not been reached, the core functionality of telomeres as protective features of linear chromosome termini was correctly posited before DNA was discovered or the composition of telomeres was known.

Muller and McClintock were the first to suggest that telomeres serve as protective elements of linear chromosome termini (McClintock, 1941; Muller, 1938). Muller worked with Drosophila and observed that despite the repeated exposure of chromosomes to X-rays, he was unable to generate a chromosome without a telomere. With *Zea mays*, McClintock deduced that

one function of telomere was to prevent fusions between chromosome termini and with DNA DSBs. Much later experiments by Griffith resulted in the discovery of the T-loop, a structure where the telomere's 3' single-stranded DNA (ssDNA) overhang loops back into (i.e T-loop) and displaces tracts of double-stranded telomere DNA (i.e D-loop) (Calado and Dumitriu, 2013; Griffith et al., 1999; McClintock, 1941; Muller, 1938; Whittemore et al., 2019). The structure of T-loops, "hiding" the 3' telomere overhang, hinted at later confirmations that they serve to prevent recognition of chromosome termini as dsDNA breaks (Longhese, 2008). Understanding the structural features of telomeres chromosome termini remains an active area of research. Indeed, new sequencing technologies in recent years have revealed that the 5'-$TTAGGG_n$-3' motif may not be the only, but rather the primary motif in mammalian telomeres; alternative motifs may well be playing previously unappreciated roles in telomere structure and therefore function (Lee et al., 2014; Somanathan and Baysdorfer, 2018). Discoveries about the core structural elements of telomeres, i.e TElomeric Repeat-containing RNA (TERRA), continue to be made, enhancing the understanding of the telomere's structure and also function (Azzalin et al., 2007). However, the structure of telomeres remains in many ways more obscure than the many purposes they serve.

Telomeres physically shorten with each cellular division due to the end-replication problem of DNA, which results from the semi-conservative synthesis of DNA and the issue of DNA polymerases synthesizing DNA only in the 5' to 3' direction, resulting in a natural loss of DNA with each round of DNA replication (Figure 1.1) (Olovnikov, 1973; Watson, 1972). Telomeres provide a buffer against this natural loss of DNA, enabling cells to lose telomere DNA, rather than coding DNA, with each cellular division. While telomere loss to replication varies greatly between species and tissue type, the estimates of human telomere loss per cell

division tend to land around 30-100 bp (Allsopp et al., 1992; Levy et al., 1992). If cells are dividing in the absence of a telomere maintenance mechanism, their telomere lengths will erode with each cell division. As cells approach critically short lengths or complete ablation of the telomere, apoptosis or replicative senescence will be triggered; ceasing continued replication for that cell (Shay et al., 1991). This "limit" of cellular division is known as the Hayflick limit, termed after an observation that cell lines derived from human fetal tissue ceased division after 40-50 cell cycles (Hayflick and Moorhead, 1961). Hence, telomeres serve the critical purpose of essentially acting as a mitotic clock, preventing runaway cellular division and forming an important barrier against tumorigenesis. However, the exact timing on this clock and the conceptual utility varies greatly depending on the absence or presence of telomere maintenance mechanisms.

Telomere maintenance mechanisms

In the search for proteins with telomere elongation activity, in 1984 and under the guidance of Dr. Elizabeth Blackburn, Dr. Carol Greider isolated the enzyme telomerase in Tetrahymena, and noted its ability to elongate synthetic telomere primers (Blackburn and Gall, 1978). Soon after, telomerase activity was observed in a human cancer cell line (HeLa), plants, and many other Eukaryotes, indicating that for Eukaryotic cells telomerase is a major player in telomere length maintenance, capable of replenishing telomere content lost to cell division (Morin, 1989; Moyzis et al., 1988; Shippen and McKnight, 1998). The telomerase holoenzyme is a dynamic, intricate ribonucleoprotein complex composed of two core subunits: the catalytic telomerase reverse transcriptase (TERT) protein, and telomerase RNA component (TERC), along with many other proteins and co-regulatory factors (Greider and Blackburn, 1987;

Shippen-Lentz and Blackburn, 1990). Telomerase (TERT) adds telomere nucleotide repeats to the 3' end of the G-rich strand of telomeres by priming reverse transcription via an RNA/DNA complex between TERC and the 3' overhang (Figure 1.2) (Podlevsky and Chen, 2012). In humans, the telomere elongation activity of telomerase is typically relegated to germline, stem, and cancer cells, along with a choice few cell types, i.e leukocytes (Shay and Wright, 2010). High expression of telomerase ensures that telomeres will not reach critically short lengths due to division, and so these choice cell types are able to "bypass" the Hayflick limit. For this reason, telomerase is typically considered essential for immortalization of cell lines by enabling unlimited cellular proliferation; this point mirrors the near ubiquitous presence (85%) of telomerase expression in human cancers (Figure 1.2) (Kim et al., 1994; Morales et al., 1999; Shay and Bacchetti, 1997). Perhaps in line with its importance in tumorigenesis, numerous and varied reports have been made about the extratelomeric, and even extranuclear, moonlighting functions of telomerase in both cancerous and normal somatic mammalian cells.

One of the first reports of the extra-telomeric functions of telomerase came in 2009, when Maida et al. observed the catalytic domain of telomerase, TERT, forming a complex with the RNA component of the mitochondrial RNA processing endoribonuclease (RMRP) *in vitro*, reminiscent of TERT's association with TERC (Maida et al., 2009). The TERT-RMPR complex was observed to have RNA-dependent RNA polymerase activity, in sharp contrast to TERT's canonical reverse transcriptase activity at the telomere. TERT was demonstrated able synthesizing from the RMPR RNA template double stranded RNAs, which were of correct length and sequence content to be processed by DICER into small-interfering RNAs – indicating a potential mechanism of gene expression regulation by TERT (Maida et al., 2009). Maida et al. later demonstrated that TERT has the capacity to complex with a variety of RNAs and perform

de novo synthesis of small RNAs *in vitro*, which again were of correct length and sequence content to be processed by DICER and exert gene regulatory functions (Maida et al., 2016). In line with the mitochondrial relationship evidenced by Maida et al., TERT has also been observed to associate with and reverse transcribe mitochondrial tRNAs (Sharma et al., 2012). The impacts on cellular biology, from the reported associations of TERT with various RNAs for RNA-dependent RNA polymerase activity or with mitochondrial tRNAs for reverse transcription, are unclear, though these studies strongly suggest a mechanistic basis for TERT to exert regulation over gene expression and mitochondrial function. Reports have also been made of TERT's clear exertion over gene expression in another context – the cytoplasm of neuronal cells.

Indeed, TERT has been demonstrated to associate a wide variety of cytoplasmic mRNAs within the RNA granules of neuronal hippocampal cells (Iannilli et al., 2013). Through an undefined mode of regulation, TERT was demonstrated to associate with the mRNA transcript of the pro-survival gene p15INK4B under normal conditions. Upon induction of cellular stress, TERT dissociated from p15INK4B, allowing expression and engagement of its pro-survival functions. A comprehensive view of TERT's associations with RNAs of various types and in various cellular locations and physiology contexts, and the specific consequences for each, have not yet been established. However, the many and also varied relationships between TERT and the mitochondria has sustained high interest during recent decades.

The first evidence for extranuclear localization of TERT was provided in 2000 by Seimiya et al., who identified a nuclear export motif at the C-terminus of TERT (Seimiya et al., 2000). Yan et al. (2004) later performed immunohistochemistry on a variety of human tissues and demonstrated that TERT localizes not only to the nucleus but also the nucleolus and cytoplasm (Yan et al., 2004). A complementary report by Haendeler et al. identified a bona fide

mitochondrial localization sequence at the N-terminus of TERT, and also demonstrated that TERT is exported from the nucleus and imported into the mitochondria during oxidative stress (Haendeler Judith et al., 2009).

Curiously, and particularly during periods of chronic oxidative stress, TERT will localize within the mitochondrial matrix and bind directly to mtDNA with a preference for the ND1 and ND2 gene sequences, which encode subunits 1 and 2 of the NADH:unibquinone oxidoreductase (complex I); the first protein in the respiratory chain (Haendeler Judith et al., 2009). The regulatory functions of TERT on these genes, if any, have not been elucidated, though it's tempting to speculate that TERT's binding of these mtDNA genetic elements implies direct regulation over transcription. A potential role for TERT in mtDNA metabolism seems plausible, given direct binding with mtDNA and the reported binding and reverse transcription of many mt-tRNAs (Gordon and Santos, 2010; Sharma et al., 2012) TERT's relation to genes encoding core respiratory chain elements hints at its general involvement with redox reactions and reactive oxygen species (ROS), though it's not clear whether TERT is entirely protective of mitochondrial function.

Conflicting reports have been made as to whether the localization of TERT to mitochondria during chronic oxidative stress provides a protective or deleterious influence on mitochondrial function. During chronic oxidative stress and upon import of TERT, Santos et al. (2004) reported increased mtDNA damage, while Ahmed et al. (2008) observed decreased peroxide levels and increased superoxide levels, and Haendeler et al. (2009) observed increased respiratory chain activity (Ahmed et al., 2008; Haendeler Judith et al., 2009; Santos et al., 2004). The seemingly contradictory nature of these reports can be reconciled by taking the view that mitochondrial import of TERT during oxidative stress results in dampened mtROS production

and thus altered redox states. If TERT were to engage an overall decrease of mtROS in mitochondria, this would provide protection against endogenous oxidative damage. However, reduction of mtROS would also modify the mitochondria redox environment and alter mtROS signaling pathways, thus modifying signaling to various pathways (i.e. autophagy), which could provide seemingly contradictory inferences about TERT and mtROS given different experimental contexts. Green et al. (2019) recently provided strong evidence that during chronic oxidative stress, TERT localization to the mitochondria and by an unknown mechanism alters the redox environment, causing reduced mtROS and thus reduced endogenous oxidative damage (Green et al., 2019). Hence, as part of the cellular response to chronic oxidative stress, TERT is exported from the nucleus and localizes to the mitochondria, where it enacts a process that results in reduced mtROS and hence reduced oxidative damage, at the expense of pathways reliant upon mtROS for molecular signaling (Green et al., 2019). By definition, chronic oxidative stress is experienced over an extended period of time, implying long-term cellular export of TERT and absence of a telomere maintenance mechanism (TMM). How cells with otherwise active and high telomerase activity (stem, germ, cancer) maintain telomere length during periods of chronic oxidative stress, with TERT relegated to the mitochondria, is not well understood. However, there is a second TMM that accounts for the telomere elongation in 5-15% cancers and also remains elusive in its definitive contribution to cellular biology in normal somatic cells (Heaphy et al., 2011).

The first suggestions of a mammalian TMM alternative to telomerase in were made during a study on the relationship between telomerase activation cellular immortalization. Bryan et al. (1994) were investigating cellular immortalization by telomerase in 35 cell lines transfected with Simian Virus 40 (SV40) (Bryan and Reddel, 1994). Of the 35 immortalized cell lines, 15

were observed to be telomerase-negative and yet possessed very long and heterogenous telomere lengths, suggesting the existence of a TMM alternative to telomerase – a pathway dubbed Alternative Lengthening of Telomeres (ALT) (Figure 1.2). Subsequent studies by Bryan et al. (1997) reported ALT in human tumors and tumor-derived cell lines, suggesting that like telomerase, ALT could immortalize cells and likely played key roles in tumorigenesis when telomerase is absent or non-functional (Bryan et al., 1997). Soon after, Dunham et al. (2000) discovered ALT to be a homologous recombination (HR)-based process by observing amplification and movement of a fluorescent telomere tag between telomeres in human cancer ALT cell lines (Dunham et al., 2000). In sum, these works identified ALT as an important TMM that is commonly engaged during tumorigenesis.

The revelation that ALT is a recombination-based process re-cast light on prior studies by Lundblad and Blackburn (1993), and McEachern and Blackburn (1996), both of which observed telomere-recombination phenomena in telomerase-defective yeast, where HR proteins were engaged for telomere elongation, similar to ALT (Lundblad and Blackburn, 1993; McEachern and Blackburn, 1996). The first evidence for ALT's role in normal mammalian biology came from a critical study by Liu et al. (2007), demonstrating that during early embryogenesis telomeres become highly elongated and heterogeneous with high rates of sister chromatid exchange at the telomere (T-SE). These observations were in line with the hallmarks of ALT (Liu et al., 2007), and provided strong supportive evidence that ALT serves a key role in "reprogramming" telomere length during embryogenesis. ALT however is almost entirely characterized as a mechanism of key importance to human cancers, with ALT cells typically containing a number of the following features: highly elongated telomeres, extremely heterogenous telomere lengths, elevated frequencies of T-SCE, extrachromosomal telomeric

DNA, and ALT-associated promyelocytic leukemia nuclear bodies (APBs), which contain telomere and HR-related proteins (Cesare and Reddel, 2010). Though many models for the mechanisms underpinning ALT have been propose [e.g. landmark review by Cesare and Reddel (2010)], the regulation of ALT activation in tumorigenesis and embryogenesis remain unclear (Cesare and Reddel, 2010). The contributions of ALT to normal somatic cellular biology, if any, remains uncertain, though strides have been made in identifying the somatic contexts of "normal" ALT, herein referred to as normal ALT.

The first potential observations of normal ALT came from normal human fibroblasts, when Martens et al. (2000) observed rapid deletion events of long telomeres with the simultaneous maintenance of shorter telomeres (Martens et al., 2000). In addition, the mean telomere length, rather than the shortest telomeres, was observed to be strongly correlated with extended replicative capacity and hence, increasing the number of cellular divisions before senescence (Martens et al., 2000). These results strongly suggested regularly occurring, non-reciprocal recombination events between the longest and shortest telomeres in normal human fibroblasts, which served to extend replicative capacity – i.e., normal ALT.

Numerous observations have also implicated normal ALT as a means of adaptation to chronic oxidative stress. Repeated treatments of hydrogen peroxide on human primary fibroblasts has been shown to significantly elongate telomeres, with concurrent increases in telomere length heterogeneity and T-SCE (Coluzzi et al., 2017). Furthermore, treatment of human foreskin fetal fibroblasts with X-rays has shown similar results, with concurrent and significant elevations in telomere length, length heterogeneity, and T-SCE (De Vitis et al., 2019). Both studies suggest that normal ALT may play important functions in adapting to chronic oxidative stress, though the studies are limited by their *in vitro* nature. There is but one report

suggesting telomere recombination, hence normal ALT, as part of routine mammalian biology. Neumann et al. (2013) inserted a fluorescent tag into the telomeres of lab mice and, much like earlier works by Dunham et al. (2000), observed movement and amplification of the tag in normal somatic murine cells (Dunham et al., 2000; Neumann et al., 2013). No follow up studies for this tantalizing report have been made. Altogether, normal ALT remains elusive in terms of its contextual activation and also overall contributions to normal mammalian biology. It is tempting to speculate that ALT may play an important adaptive role in the cellular response to periods chronic oxidative stress, given that during these periods TERT is known to be excluded from the nucleus and localize to the mitochondria, while ALT has been shown to transiently activate. Whether this interplay exists is not known; and certainly is unknown for human cells *in vivo*, as no previous reports have demonstrated activation of normal ALT in normal human cells.

An extremely recent report by Vaz et. al (2020) suggested that human T-cells may receive vesicles containing telomere content and also recombination factors (RAD50, BRCA2) from antigen presenting cells (APC) (Vaz et al., 2020). The authors observed that exposure of telomerase-negative T-cells to APC vesicles resulted in telomere elongation for the T-cells with a concurrent decrease in telomere length for the APC cells which released the vesicles (Vaz et al., 2020). These results suggest non-reciprocal exchange of telomere content from APCs to T-cells via vesicles, followed by recombination-factors mediating recombination (attachment; normal ALT) of the APC telomere content onto the T-cell's telomeres. These results are reminiscent of the non-reciprocal telomere length events reported decades earlier by Martens et al. (2000) (Martens et al., 2000). The recency of the report by Vaz et. al (2020) tempers the interpretation, but this work could potentially represent a third TMM, in addition to telomerase

and ALT, by which cells can maintain telomere length via intercellular vesicular transfer of telomere content – time will tell.

Altogether, maintenance of telomere length is of critical importance to cellular health and function – this importance is underscored by the complexity and nuances of the TMMs. Not only does telomerase maintain telomere length in normal cells and also as part of tumorigenesis in ~85% of cancers, telomerase also has important, nuanced, and poorly understood roles in the mitochondria; thus telomerase bridges telomeric and mitochondrial function for reasons that are still unfolding (Maida et al., 2009, 2016; Sharma et al., 2012; Shay and Bacchetti, 1997). ALT is known to elongate telomeres in telomerase negative-cancers and potentially during mammalian embryogenesis, and reports have indicated transient ALT activation in normal cells during periods of chronic oxidative stress, potentially suggesting an adaptation function (Bryan and Reddel, 1994; Bryan et al., 1997; Coluzzi et al., 2017; De Vitis et al., 2019; Liu et al., 2007). The extremely recent demonstration of intercellular telomere exchange highlights the complexity of TMMs and more importantly, the cellular emphasis on maintaining telomere length (Vaz et al., 2020). The importance of appropriate maintenance and regulation of telomere length for cellular stability and lifespan, translates into into overall tissue health and longevity.

Telomere length: individual variation and roles in disease

Telomere length: aging, individual variation, and lifestyle factors

With each cellular division, telomeres erode until reaching a critical length and at which point, if a TMM is not present, cells will cease dividing and activate apoptosis or enter replicative senescence (Figure 1.3) (Aubert and Lansdorp, 2008). As senescent cells accumulate,

tissue function declines, and as tissues decline so do organs, underpinning the physiological phenomena of organismal aging until mortality is reached (van Deursen, 2014). Hence, telomere length provides a mitotic clock that, when exhausted for any given cell, incurs senescence, pushing the overall organism down the path of aging. Thus, the rate at which an organism's overall telomere length declines provides an important measure of aging and also risk for aging-related degenerative pathologies, including cardiovascular disease and cancer (Figure 1.3) (Aubert and Lansdorp, 2008). Indeed, aging – cellular division – is the central contributing factor to telomere erosion. The rate of telomere erosion in humans is a synthesis of age, genetics, and lifestyle factors; as such, telomere length is a highly heritable trait that can be modulated by lifestyle choices, in particular nutrition and physical activity (Shammas, 2011).

For humans, the baseline telomere length of an individual is determined at the moment of conception, based on the telomere length of the child's parents (Turner et al., 2010). It is generally agreed that telomere length of sperm increases with the father's age while telomere length of oocytes declines with the mother's age (Broer et al., 2013; Eisenberg and Kuzawa, 2018). At conception, the telomeres from each parent provide the baseline telomere length in the embryo. During the early stages of embryogenesis telomere length undergoes a "reprogramming", likely to be highly plastic and receiving cues from the maternal environment, which involves extensive lengthening of telomeres via a recombination based process (essentially, ALT) (Entringer et al., 2018; Liu et al., 2007). The oocyte has been suggested to be the primary regulator of the telomere length "reprogramming", in line with indications that maternal inheritance of telomere length predominates (Broer et al., 2013). The overall heritability of telomere length has been estimated in various studies; early reports on twins and siblings indicated 34-82% of telomere length is inherited, while a recent, large meta-analysis on

six independent cohorts of subjects (nearly 20,000 subjects) estimated the heritability of telomere length to be approximately 70% (Broer et al., 2013). While telomere length is a highly polygenic and malleable trait, many genomic loci and single nucleotide polymorphisms important for determining telomere length in humans at birth and throughout life have been identified through Genome Wide Association Studies. Some of the most important genomic loci for telomere length include TERT and TERC (Telomerase RNA Component) but also MRE11 and MEN1 (ALT-/HR-related) among many others (Mirabello et al., 2010). Altogether, telomere length is a highly heritable trait with genetic background playing an important but not complete role in determining telomere length throughout life. Indeed, it has been demonstrated that an individual can modulate their telomere length, and thus the rate of their aging and disposition to aging-related diseases, through lifestyle choices – both positive and negative.

In parallel with conventional wisdom about "eating properly", modulation of telomere length via eating habits, food choices, and nutritional intake has been demonstrated. Studies on the relationship between vitamin intake and telomere length have included positive associations between folate, vitamin B12, nicotinamide, and vitamins A/D/C/E; the antioxidant capacities of the above mentioned vitamins is generally considered protective against telomere shortening due to oxidative damage (Vidaček et al., 2017). Positive associations between mineral intake and telomere length have also been reported, these include magnesium and zinc – zinc is notably important for facilitating reverse transcriptase activity (i.e telomerase) (Poiesz et al., 1974). In contrast, iron intake has been correlated with shorter telomeres, likely because iron is a prooxidant, capable of generating ROS (Vidaček et al., 2017). Consumption of omega-3 fatty acids from seafood has also been positively associated with telomere length (Vidaček et al., 2017). More broadly, positive associations have been reported between telomere length and

consumption of seafood, seaweed, legumes, nuts, coffee; negative associations include alcohol, red/processed meats, sugary beverages, and white bread; proposed mechanisms of action by foods on telomere length tend to focus on the antioxidant or prooxidant content of the specific food (Vidaček et al., 2017). In addition to the effects of nutrition on telomere length, links between physical exercise, obesity, and smoking have also been demonstrated.

Physical activity is well understood to be positively associated with healthy aging and reduced risk of aging-related diseases; many reports support telomeres as an important link between these phenomena. Telomeres are known to shorten in response to inflammation and oxidative stress, and if the inflammation and/or oxidative stress is chronic, telomeres will continue to shorten at an accelerated rate irrespective of age, contributing to increased risks for aging-related diseases (Aubert and Lansdorp, 2008). Many reports have demonstrated that the net effect of physical activity on inflammation is a reduction in markers of inflammation, including TNF-α, CRP, IL6, and INF- γ, with simultaneous increases in levels of anti-inflammatory factors such as IL-4, IL-10, IL-12, and TGFβ1 (Sallam and Laher, 2016). Physical activity has also been demonstrated to reduce age-related oxidative stress in the heart, arteries, plasma, liver, and skeletal muscles in older individuals (Bejma et al., 2000; Kim et al., 1996; Navarro et al., 2004; Radák et al., 2004). Physical activity has also been reported to engage mitochondrial biogenesis and turnover of defective mitochondria, reducing oxidative stress (Jornayvaz and Shulman, 2010). Reduced levels of inflammation and oxidative stress via physical activity will reduce levels of ROS, which in turn enhances maintenance of telomere length; many reports have evidenced this line of reasoning (Arsenis et al., 2017). Altogether, physical activity enables individuals to improve their health via reduced inflammation and oxidative stress, resulting in improved telomere maintenance and healthspan.

In sharp contrast to physical activity, obesity is understood to be negatively associated with healthy aging. Obese individuals have markedly increased levels inflammation and oxidative stress which consequently decreases telomere length and increases the incidence and burden of aging-related diseases, including cardiovascular disease and cancer. Importantly, telomere length has been evidenced to decrease linearly with increasing body mass index (BMI) in obese individuals (Kim et al., 2009), in line with observations that morbid obesity coincides with even higher incidence of disease burden. Considering the effects of obesity on inflammation, oxidative stress, telomere length, and disease burden, obesity has been proposed to accelerate the aging process (Salvestrini et al., 2019). Similar to obesity, cigarette smoking is also understood to be negatively associated with healthy aging, causing increased levels of inflammation and oxidative stress, along with increased incidence and burden of aging-related diseases, including cancer and cardiovascular disease (Astuti et al., 2017). Expectedly, smoking also shortens telomere length, likely due to ROS generated via smoking, which result in inflammation and oxidative stress. Similar to obesity, the amount of cigarette smoking is linearly correlated with decreases in telomere length (Astuti et al., 2017). Indeed, obesity and cigarette smoking are two lifestyle factors that bear negative consequences on telomeres and disease risk, providing the important demonstration that individuals can modulate their telomere length – and therefore aging trajectory – for better or worse.

Telomere length is determined at birth via parental telomere length and embryonic reprogramming (Turner et al., 2010); thereafter, genetics and lifestyle choices such as nutrition and physical exercise play determinant roles in positively or negatively affecting telomere length throughout life (Mirabello et al., 2010; Ornish et al., 2013). However, aging – cellular division – is the prime, inevitable factor controlling telomere length attrition. As we age, telomeres shorten

and cells enter replicative senescence, limiting the replicative capacity of cells (Aubert and Lansdorp, 2008). This hard-stop on cellular division is thought to provide an important tumor suppressive role, as unlimited proliferation and thus clonal dominance is prevented (Aubert and Lansdorp, 2008). If the majority of telomeres in a hypothetical organism are too short, unlimited clonal expansion will indeed be avoided (Aubert and Lansdorp, 2008); however an unfortunate secondary problem is created where this hypothetical organism would be in a state of complete degeneration (Aubert and Lansdorp, 2008). For this hypothetical organism, the "ideal" telomere length would be sufficiently long to promote healthy aging and avoidance of aging-related diseases/cancer, while also sufficiently short to cease cell division before oncogenic mutations can be acquired and tumorigenesis initiated (Aubert and Lansdorp, 2008). These opposing forces on telomere length are colloquially referred to as the cancer-aging paradox, and are of paramount importance to understanding the links between telomere length and aging-related diseases such as cancer and cardiovascular disease (Shay, 2016).

Roles of telomeres and telomerase in aging-related diseases

Cardiovascular diseases (CVD) are degenerative diseases that affect all aspects of the circulatory system, e.g. heart, blood vessels arteries, and as a general category are the number one global cause of death. Approximately 17 million individuals died from CVD in 2008, representing 30% of all global deaths; 23 million individuals are projected to die from CVD in 2030, mainly from heart disease and stroke (Balakumar et al., 2016). Modifiable risk factors for CVD include non-physical activity, hypertension, cigarette smoking, alcohol abuse, unhealthy dietary choices (poor nutrition), and other lifestyle choices (Balakumar et al., 2016). Non-modifiable risk factors for CVD include sex, family history, and age (Balakumar et al., 2016).

Telomere length is of high interest to CVD research, given that CVD is a degenerative disease, the strong mechanistic links between telomere length and CVD, and the high global mortality of CVD (Balakumar et al., 2016; Zhan and Hägg, 2019).

As cells divide, telomeres shorten and cells enter senescence, senescent cells accumulate, tissue function deteriorates which manifests as aging and degenerative age-related diseases (Aubert and Lansdorp, 2008). Indeed, senescent cells have been tightly linked to age-related CVD. Tissue degeneration increases with age, but also due to chronic consumption of poor dietary choices (malnutrition), alcohol abuse, non-physical activity (Aubert and Lansdorp, 2008). The important risk factors for CVD, such age, diet, and lifestyle choices, which are all capable of driving tissue degeneration and CVD progression, are also capable of driving telomere shortening, engaging accumulation of senescent cells and CVD progression (Aubert and Lansdorp, 2008; Balakumar et al., 2016; Vidaček et al., 2017; Zhan and Hägg, 2019).

One of the major physiological drivers of CVD is atherosclerosis, which is the accumulation of fats, proteins, and other substances at sites, termed plaques, of damaged endothelial cells lining the vascular system (Balakumar et al., 2016; Zhan and Hägg, 2019). The integrity of the endothelial cells lining the vascular system is essential for vascular health and prevention of atherosclerosis and CVD (Balakumar et al., 2016; Zhan and Hägg, 2019). Curiously, significantly shortened telomeres and high numbers of senescent cells in vascular smooth muscle cells (VSMCs) isolated from human atherosclerotic endothelium have been observed (Matthews et al., 2006). These cells had rampant expression of oxidative stress markers and very low levels of telomerase. Expression of exogenous telomerase (TERT) rescued the senescent cells, suggesting that dysregulation of telomerase activity in VSMCs via chronic oxidative stress leads to telomere shortening and eventual senescence, driving atherosclerosis

(Matthews et al., 2006). Within atherosclerotic plaques, senescent VSMCs with highly shortened telomeres have been observed to excrete matrix-degrading enzymes, promoting dissolution of the plaque, which can lead to life-threatening conditions like thrombosis or stroke; suggesting that telomere length dysregulation can strongly drive pathogenesis and mortality risks during CVD (Matthews et al., 2006).

A condition related to atherosclerosis and important in the context of CVD is cardiomyopathy, where heart muscle becomes diseased and less able to pump blood throughout the body (Balakumar et al., 2016). An important study by Leri et al. (2003) demonstrated that telomerase double knock-out mice have shortened telomeres and exhibit myocyte cell death, ventricular dilation, thin vascular walls, and general cardiac dysfunction; all mirroring cardiomyopathy in humans (Leri et al., 2003). In contrast, high exogenous expression of telomerase in mice has been shown to prevent the above described cardiac symptoms (Bär et al., 2014). Indeed, in endomyocardial biopsies of humans with heart failure/CVD shortened telomeres, accumulation of senescent cells, and high levels of apoptosis signaling have been observed, further implicating dysregulation of telomeres/telomerase as core processes in CVD (Chimenti et al., 2003).

Physical activity is well appreciated as a preventative and therapeutic measure for CVD as it improves vascular function, and has also been demonstrated to increase telomere length; the links between physical activity, telomere length and CVD may be more direct than previously appreciated. One important signaling molecule and parameter of general vascular health is nitric oxide, which is produced by endothelial nitric oxide synthase (eNOS) (Förstermann Ulrich and Münzel Thomas, 2006). Extended periods of physical activity in mice has been observed to cause increased expression of telomerase activity and decreased expression of vascular apoptosis

markers (Werner et al., 2008). Curiously, these observations were ablated in both telomerase negative and eNOS negative mice, suggesting that the positive effects of exercise on vascular health are mediated by a telomerase-eNOS axis. Thus, telomeres and telomerase play fundamental roles in facilitating heart function (Werner et al., 2008).

Telomeres and telomerase also play key roles in the progression and prevention of CVD. Non-modifiable factors like age and lifestyle lead to shortened telomeres, promoting accumulation of senescent cells which lead to vascular tissue degeneration and pathologies at the organismal level (atherosclerosis, cardiomyopathy) which contribute to CVD progression mortality (Aubert and Lansdorp, 2008; Zhan and Hägg, 2019). The results of mechanistic and observational studies linking telomere length to CVD, in various contexts and pathologies, are supported by a body of literature that leverages an advanced statistical approach (Mendelian Randomization) to test the hypothesis of causality between shortened telomeres and CVD – strong causality has indeed been observed (Codd et al., 2013; Protsenko et al., 2020; Scheller Madrid et al., 2016; Xu et al., 2020; Zhan et al., 2017). Conversely, positive lifestyle choices like adequate diet and regular exercise promote telomere elongation and reduction of vascular pathologies (i.e., via TERT-eNOS axis), potentially preventing or attenuating CVD progression. With shorter telomeres comes increased burden of degenerative diseases, like CVD; the benefit being reduced rates of tumorigenesis, with the reverse being true for longer telomere lengths (Protsenko et al., 2020).

Normal somatic cells with longer telomeres are inherently able to persist longer than cells with shorter telomeres (Aubert and Lansdorp, 2008). Therefore, longer lived cells are able to accumulate relatively more oncogenic mutations; e.g., inappropriate activation of telomerase or ALT combined with p53 loss of function, which could enable bypass of senescence and

apoptosis, unlimited cellular proliferation, and potentially carcinogenesis (Hanahan and Weinberg, 2011). Individuals with overall longer telomeres are thus at higher risk for developing cancer; though longer telomeres are certainly not a requirement for tumorigenesis. The increased incidence of cancer from longer telomeres is, interestingly, nearly equivalent to the increased incidence of degenerative diseases (i.e CVD) due to shorter telomeres(Protsenko et al., 2020).

If an actively dividing cell has lost the ability to sense DNA damage and halt cell cycle progression, e.g. via p53 or retinoblastoma loss of function mutations, that cell is likely to ignore activation of senescence or apoptosis pathways upon reaching critically short telomere lengths (Hanahan and Weinberg, 2011). If the cell does escape and continue to divide, it will enter a state of "telomere crisis", characterized by genomic instability and dysfunctional telomeres that are liable to fuse, initiating break-fusion-bridge cycles (Ishikawa, 1997; McClintock, 1941). A recent report in yeast by Heasley et al. (2020) has proposed that even a single aberrant round of mitosis can lead to highly reconfigured genomes with many aneuploidies (Heasley et al., 2020). This observation is highly reminiscent of telomere crisis, which is known to severely disrupt mitosis has been observed to yield all manner of genomic reconfigurations, ranging from aneuploidy, chromosome rearrangements, and loss of heterozygosity, to the more extreme events such as chromothripsis, kataegis, whole genome duplication, and even tetraploidization (Ishikawa, 1997; Maciejowski et al., 2015). Human cells that undergo and successfully emerge from telomere crisis have very high rates of death - but for some, telomerase or ALT-mediated pathways (~85% and ~15%, respectively) have been activated and/or upregulated, granting unlimited cellular proliferation and initiating tumorigenesis (Ishikawa, 1997; Maciejowski et al., 2015).

While telomerase is primarily activated in cancers for telomere maintenance, it is possible that TERT's non-canonical functions such as localizing to the mitochondria for ROS

reduction or the potential capacity for gene regulation are important for tumorigenesis (Green et al., 2019; Iannilli et al., 2013; Maida et al., 2009, 2016). As well, it is important to note that telomerase and ALT activity have been observed in the "same" tumor, though they are typically considered mutually exclusive processes in the same cell (Hu et al., 2016). Cells with highly reconfigured genomes and capacity for unlimited proliferation either via telomerase or ALT will continue down the path of tumorigenesis and may become life-threatening to the organism if they escape immune surveillance. Telomere length has been analyzed in most types and stages of cancers, with reports of long telomeres in some cancers and short telomeres in others, depending on the tissue type and cancer staging (Okamoto and Seimiya, 2019; Wentzensen et al., 2011). Indeed, using telomere length to assess risks for cancer or monitor cancer progression is of high interest and an area of active research for personalized medicine approaches, particularly for ionizing radiation (IR) therapies in the treatment of cancer (Ferrandon et al., 2013; Mirjolet et al., 2015; Wu et al., 2003).

Ionizing radiation induced genomic instability

Ionizing radiation and telomere length dynamics

IR is a potent inducer of DNA damage and for that reason, can both cause cancer and be used to effectively treat it. In a therapeutic context, many varieties of IR delivery and strategies have been developed to optimize treatment success, but again is a double-edged sword, as secondary malignancies are a concern (Baskar et al., 2012, 2014; Chew et al., 2020). The magnitude of risk is contextual and varies greatly depending on the size and location of the tumor, the extent of normal tissue exposed, as well as the type and amount dose received, and the patient's demographics (Chew et al., 2020). IR is essentially energy that travels as

electromagnetic waves (e.g. X-rays, γ-rays) or particles (e.g. β-/α-particles, neutrons, protons, highly energetic heavy ions (HZE)), and that have the capacity to strip electrons from atoms or molecules, thereby ionizing them (Keith et al., 2012; US EPA, 2014). This transfer of energy from IR to encountered matter via ionization is known as linear energy transfer (LET) and is formally described as the amount of energy deposited via IR per unit distance within the encountered matter (Keith et al., 2012; US EPA, 2014). With respect to human cells, there are two general modes of action by which IR damages DNA.

Exposure of water to IR causes radiolysis, where the IR lyses the water molecules into H and OH radicals, which are chemically reactive species, that initiate chain reactions that yields various types of ROS, e.g. hydroxyls, peroxides, superoxides, etc. Human cells are approximately 80% water; as electromagnetic IR passes through cells it yields ROS; if the ROS are in proximately to DNA they will produce single-stranded (ss) or double-stranded breaks (DSBs) (Feldberg and Carew, 1981). The interaction between electromagnetic IR and cells are considered low LET events as the DNA damage is indirectly caused (Feldberg and Carew, 1981). Interactions with heavy charged particles are high LET events as the IR particulate will directly ionize DNA molecules, causing highly complex clusters of DNA damage along with secondary ionization events yielding ROS which incurs further damage (Hada and Georgakilas, 2008). In addition to heavy charged particles, neutrons are another type of high LET and deposit their energy primarily via nuclear interactions (Hada and Georgakilas, 2008). Between low and high LET events, low LET accounts for approximately 70% of DNA damage for any general IR exposure, given the high water content of cells (Desouky et al., 2015). Exposure of cells to IR is well appreciated to induce oxidative stress; chronic IR exposure thus induces chronic oxidative stress (Azzam et al., 2012). While IR exposure is destructive for DNA in general, telomeres are

particularly sensitive to IR by virtue of their sequence content (Coluzzi et al., 2019; Maeda et al., 2013; Zglinicki, 2000).

Telomeres are rich in guanine (5'-TTAGGG$_n$-3') and one of the most common DNA lesions caused by ROS from IR exposure is 8-oxoguanine (8-oxoG), due to its relatively low redox potential (Figure 1.4) (Aguiar et al., 2013; Singh et al., 2011). As such, telomeres are particularly sensitive and liable to shorten from IR exposure due to the resulting ROS causing 8-oxoG bases (Figure 1.4) (Lee et al., 2017). Oxidization of guanine residues within telomeres inhibits the binding of TRF1 and TRF2, deprotecting the oxidized bases. DNA replication is also perturbed by the presence of 8-oxoG bases, as eukaryotic polymerases will insert an adenine opposite the site (Aguiar et al., 2013). While telomerase is able to insert 8-oxoG bases from the free-nucleotide pool, upon insertion the elongation process will terminate immediately, impairing the ability of telomerase to maintain telomeres in the presence of free 8-oxoG bases (Fouquerel et al., 2016). However, if the 8-oxoG bases have been incorporated close to the terminus of the telomere, G-quadruplex formation is impaired, enabling telomerase easier access to the 3' end of the telomere, facilitating elongation (Fouquerel et al., 2016). The dynamics and net effects of 8-oxoG on telomere length *in vivo*, cell biology, and tumorigenesis are not clear, though the near opposite impacts of 8-oxoG bases on telomerase processivity (free 8-oxoG bases terminate, incorporated 8-oxoG bases enhance) may explain a portion of contradictory reports on telomere shortening and lengthening after IR exposure (Sishc et al., 2015).

Structural variation – chromosome aberrations (CAs) – and recombination events

The impacts of IR exposure on genomic stability is of particular interest in the context of radiotherapy and environmental exposures. IR in general is a potent DNA damaging agent, and while the types and amount of DNA damage incurred will vary based on type, dose, and

frequency of IR exposure, it is clear that IR exposure is capable of driving tumorigenesis long after exposure (Barcellos-Hoff and Nguyen, 2009; Little, 2000; Mozdarani, 2012). For this reason, understanding the incidence and types of genomic damage incurred after different types and doses of IR exposure, particularly for normal tissues, is critical for improving IR-based therapies and also understanding and predicting the long-term effects of specific environmental exposures.

There are many types of direct and indirect damage, and mutagenesis events inflicted by IR upon DNA that can be described; these range in size and complexity from simple lesions, ssDNA breaks, DSBs, stalled or collapsed replication forks, and also larger structural rearrangements such as inversions, translocations, and dicentrics, or more complex damage such as chromothripsis and kataegis (Cornforth, 2001; Koltsova et al., 2019; Morishita et al., 2016). Though not directly induced by IR exposure, sister chromatid exchange (SCE) represent a recombination event that can occur as a result of IR-induced oxidative stress, consequent replication stress, fork stalling, and template switching (Conrad et al., 2011; Nagasawa and Little, 1992). Other events that can trigger SCE include collapsed replication forks, replication fork reversal, template switching, gap repair, and break-induced replication (Wojcik et al., 2004). Importantly, SCEs are frequently observed to be elevated after IR exposure and are acknowledged to indicate genomic instability via replication stress (Conrad et al., 2011; Nagasawa and Little, 1992; Rudd et al., 2007). T-SCE are particularly important as these are a hallmark of ALT (Blagoev et al., 2010; Cesare and Reddel, 2010; Londoño-Vallejo et al., 2004).

IR is also a potent inducer of prompt DSBs. The resulting intra-chromosomal structural variants (SVs) include deletions, duplications, and inversions (Morin et al., 2017; Pai et al., 1980; Weckselblatt and Rudd, 2015). Deletions and duplications have historically been easier to

detect relative to inversions, as array-based methods can detect the difference in copy number variation, while sequencing can identify deletions and duplications using a reference genome for comparison (Weckselblatt and Rudd, 2015). Inversions, particularly small inversions, have historically been difficult to detect, given that they're copy number neutral and not captured via sequencing without high read depths (Weckselblatt and Rudd, 2015). Recent advances in cytogenetics, namely directional genomic hybridization have enabled straightforward detection of even small inversions(Ray et al., 2013, 2014). The importance of inversion detection is underscored by the high rates at which they occur, relative to other structural variants such as translocations or dicentrics (McKenna et al., 2019).

Inter-chromosomal SVs include translocations and inversions, which minimally require two DSBs, close in time and space, and insertions that require at least three (Figure 1.5) (Weckselblatt and Rudd, 2015). Translocations are typically reciprocal, where two chromosomes exchange genetic material; the manner in which these derivate chromosomes segregate determines whether the event is copy-neutral or results in a loss of genetic information (Weckselblatt and Rudd, 2015). Per the name, insertions are the interstitial insertion of chromosomal material into a non-homologous chromosome (Weckselblatt and Rudd, 2015). Dicentrics are another result of inter-chromosomal rearrangement that result from the fusion of two broken centric chromosomes (Weckselblatt and Rudd, 2015). Dicentrics have long been associated with IR exposure; indeed, dicentrics are regarded as the "gold standard" for biodosimetry (Lloyd et al., 1980; Wilkins et al., 2008). Dicentrics can also occur when telomeres become dysfunctional, either by becoming critically short or lost entirely, or by loss of end-capping function, yielding "sticky" chromosome ends that are susceptible to fusion (Stimpson et al., 2012). It is important to note that if one of the centromeres in the dicentric chromosome is

successfully deactivated, and thus does not influence segregation during mitosis, dicentrics can persist indefinitely as a stable structure, assuming its formation did not involve loss of critical genomic information (Stimpson et al., 2012).

Intra- and inter-chromosomal SVs represent informative signatures of IR exposure that help contextualize potential long-term health outcomes. SCEs are also important given that their incidence indicates the relative degree of replication stress and instability. Thus, monitoring SVs can provide valuable, personalized information on individual susceptibility and long-term risks for tumorigenesis (Schuster et al., 2018; Vodenkova et al., 2020). This is especially evident in the context of personalized IR therapy regimes, as well as for populations of occupationally exposed individuals, such as NASA astronauts on missions aboard the International Space Station (ISS).

Interrogating and predicting telomeric and chromosomal responses to specific exposures

Safeguarding human health on the ISS and beyond

Space remains the final frontier of human exploration and knowledge. By 2024, NASA aims to return humans to the moon and establish a permanent lunar base, marking a new era of space exploration and human endeavors (Dunbar, 2019). To date, the short- and long-term impacts of long-duration spaceflight on human health remain poorly understood, owing to the complex interactions and myriad effects on cellular biology that result from microgravity and space radiation exposure, and other spaceflight stressors like isolation, endurance exercise, and diet (Chancellor et al., 2014, 2018; Cucinotta, 2015; Garrett-Bakelman et al., 2019). The potential impacts of long-term, deep-space travel are markedly less clear (Chancellor et al., 2014). To ensure continued high performance of astronauts during and after spaceflight missions,

understanding the effects of spaceflight on human health is of critical importance (Chancellor et al., 2014).

Space radiation is composed of types of radiation not typically encountered on Earth, galactic cosmic rays (GCRs) and solar particle events (SPEs) (Chancellor et al., 2014). GCRs are HZE particles that range in size from hydrogen to nickel and originate from outside our solar system. SPEs are emissions of high-energy particles, mainly protons, released by the Sun during solar eruptions or flares (Chancellor et al., 2014). Both GCRs and SPEs are capable of producing profound DNA and tissue damage, which could potentially initiate tumorigenesis or promote tissue degeneration during or after long-duration spaceflight (Chancellor et al., 2014, 2018; Cucinotta, 2015). The ISS is in Low Earth Orbit (LEO), and so has some protection from space radiation due to Earth's magnetosphere, however, space radiation exposure is still of major concern (Chancellor et al., 2014).

With regards to space radiation exposure in LEO aboard the ISS, astronauts are on average exposed to 70 μGy from GCRs per day, and up to 500 μGy when depending on orbit (Dachev et al., 2017). During and after periods of magnetic storms, daily dose rates have been observed to increase and range from 3000 μGy per day and even 5000 μGy per hour; the highest recorded exposure was a particularly strong SPE at 1400 mGy per hour (Dachev et al., 2017; Hu et al., 2009). These ranges in exposure highlight the complex and often unpredictable dynamics of space radiation exposure. Importantly, these observations are all within LEO and as astronauts venture beyond into deep-space, exposure will only increase. Indeed, it is anticipated that astronauts traveling through deep-space will have every cell nucleus exposed to at least one hydrogen ion every few days, and exposure to heavier GCRs every few months (Chancellor et al., 2018). In deep-space, GCR dose rates are estimated to range from 50 mGy to 300 mGy per

year depending on solar conditions, while SPEs can yield dose rates up to 2800 mGy per hour (Chancellor et al., 2018). For perspective, in contrast to the doses described above humans on Earth are on average exposed to less than 5 μGy per year, with half generally from natural sources. Altogether, spaceflight poses significant and continuing challenges with regards to understanding biological impacts of space radiation.

The recent NASA Twins Study interrogated the impacts of spaceflight on human health by investigating identical twin NASA astronauts (Garrett-Bakelman et al., 2019). One twin undertook a long-duration 340 day mission aboard the ISS, while the other remaining on Earth for that year (Garrett-Bakelman et al., 2019). Our part of the NASA Twins Study focused on telomere length and chromosomal rearrangements to better understand the impacts of long-term chronic exposure to space radiation on genomic stability, and thus consequences for human health (Garrett-Bakelman et al., 2019). Exposure to space radiation and elevated oxidative stress were primary concerns for the astronaut's health. Surprisingly and contrary to expectations, telomere elongation was observed during spaceflight, which rapidly shortened to baseline upon post-flight return (Garrett-Bakelman et al., 2019). Indications of DNA damage responses, specifically elevated chromosomal inversions and translocations, were also observed mid-flight and persisted at post-flight (Garrett-Bakelman et al., 2019). These results are concerning with respect to cancer risk. Elongated telomeres naturally enable increased proliferative capacity, and in combination with elevated chromosomal rearrangements leads to increased risks for cancer (Aubert and Lansdorp, 2008; Garrett-Bakelman et al., 2019; Protsenko et al., 2020). The overall effects of long-duration spaceflight missions on health remain to be seen and understood. Whether telomere elongation is a spaceflight-specific phenomena, and what underlying mechanisms are responsible for the observed telomere elongation, remain to be elucidated.

Clarifying these mechanisms will be critical to ensuring and safeguarding astronaut performance during and after spaceflight.

Utility of predicting telomeric and chromosomal responses to exposures

It is well understood that telomere length has mechanistic, causal relationships with many aspects of health, and that individuals with overall shorter telomeres have increased risks for degenerative diseases while those with overall longer telomeres have increased risks for cancer (Aubert and Lansdorp, 2008; Shammas, 2011; Shay, 2016; Xu et al., 2020; Zhan and Hägg, 2019; Zhan et al., 2017). As telomeres shorten with age, genetic background, lifestyle factors and environmental exposures intersect and contribute to telomere shortening, and lengthening, thus modifying risks for degenerative disease and cancer respectively (Aubert and Lansdorp, 2008; Shammas, 2011; Shay, 2016; Xu et al., 2020; Zhan and Hägg, 2019; Zhan et al., 2017). While it is well appreciated that telomere length changes in response to environmental exposures (e.g. IR exposure during therapy, space radiation/spaceflight, etc.) and that those changes differentially modify disease risk, it is not known whether an individual's telomeric response to a given exposure can be predicted (Cucinotta, 2015; Garrett-Bakelman et al., 2019; Lustig et al., 2016; Maeda et al., 2013; Mirjolet et al., 2015).

Having the ability to predict an individual's telomeric response to a given exposure would provide significant improvements to understanding changes in disease risk as a function of telomeric length and would be relevant to any situation where changes in telomere length are anticipated or unavoidable from specific exposures (Mirjolet et al., 2015; Protsenko et al., 2020; Vodenkova et al., 2020; Zhan and Hägg, 2019). In addition to being able to predict an individual's telomeric response to a given exposure, specifically IR, it would also be of material

value to predict an individual's incidence of CAs, given the well understood links between CAs and cancer risk and progression (Baskar et al., 2012; Chew et al., 2020; Vodenkova et al., 2020). One circumstance where predicting an individual's specific change in disease risk due to IR exposure would be radiotherapy; especially for pediatric patients where secondary cancers are of extreme concern (Choi et al., 2014; Lin et al., 2019; Ng et al., 2010). Indeed, there is high and increasing interest in leveraging new computational approaches and biomarkers for predicting, monitoring, and understanding an individual's outcomes with respective to IR therapies in the era of personalized precision medicine (Holzinger, 2014; Kang et al., 2015; Krittanawong et al., 2017).

Another circumstance where personalized monitoring is of extreme importance would be astronaut health during and after spaceflight missions aboard the ISS (Chancellor et al., 2014, 2018; Cucinotta, 2015). Spaceflight missions aboard the ISS are inherently physically and mentally taxing – to the extreme. A common, unattributable but apt phrase, which describes the responsibilities demanded by spaceflight, is as follows: "… *if something breaks aboard the ISS, you have the rest of your life to fix it".* Monitoring astronaut health to identify issues and intervene early if complications arise is critical for ensuring astronaut safety and mission success, given that a single misstep during a spaceflight mission has the potential to end in catastrophe. Additionally, monitoring astronaut health after spaceflight is of equivalent concern with regards to maintaining a robust spaceflight program. Given the uncertainties around spaceflight's impacts on human health, a framework for predicting an astronauts' telomeric and chromosomal responses during and after spaceflight would also be of interest for contextualizing long-term health risks (Garrett-Bakelman et al., 2019). One framework that could potentially enable

predictions of telomeric and chromosomal responses to IR in the context of radiotherapy and even spaceflight is machine learning.

In essence machine learning is the process of training a computational algorithm to "identify" patterns in data (Li et al., 2018). With a robustly trained algorithm, future data points, events, and patterns can be predicted (forecasted) when new data is presented to the algorithm (Li et al., 2018). In addition, the trained algorithm will generally be plastic and able to continually "learn" to improve its ability to make inferences from new data (Li et al., 2018). Today, the most sophisticated, cutting edge machine learning algorithms are widely available, free, and accessible for anyone with programming experience, a modest laptop, and an Internet connection (see python.org and scikit-learn.org).

In biological research it is common to engage with extremely large and complex sets of data, e.g. sequencing, cytogenetic imaging, individual telomere length measurements (Li et al., 2018). Thus, machine learning (under various names, e.g. deep learning, neural networks, etc.) has seen an explosion in interest within biological and namely cancer research in recent years (Boldrini et al., 2019; Holzinger, 2014; Kang et al., 2015; Krittanawong et al., 2017; Li et al., 2018). With regards to radiotherapy the utilization and application of machine learning has provided significant improvements to areas such as treatment planning, personalized monitoring, and treatment decisions (Boldrini et al., 2019). Machine learning approaches have extremely important roles to play in heralding new personalized medicine approaches. To date, no publications have reported attempts to predict an individual's telomeric and/or chromosomal responses to a given exposure, including IR, representing a major knowledge gap in both the fields of telomere biology and radiobiology. Possessing the ability to predict an individual's specific telomeric and chromosomal responses to IR, in any context, would provide significant

and material improvements in the utilization of telomeres and CAs in personalized monitoring after such exposures.

Figures

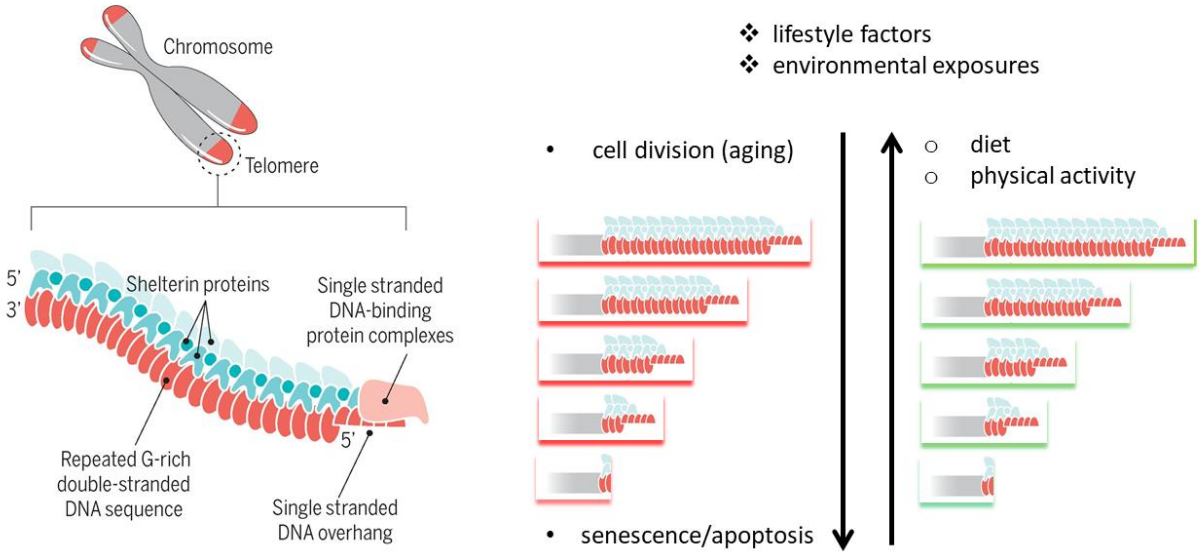

Figure 1.1. Telomere structure and relation with aging. Telomeres are repetitive nucleic acid elements encapsulated in a plethora of proteins, collectively termed shelterin. Telomeres shorten with each cell division and therefore with age, but also shorten and lengthen in response to lifestyle choices. Adapted from (Blackburn et al., 2015).

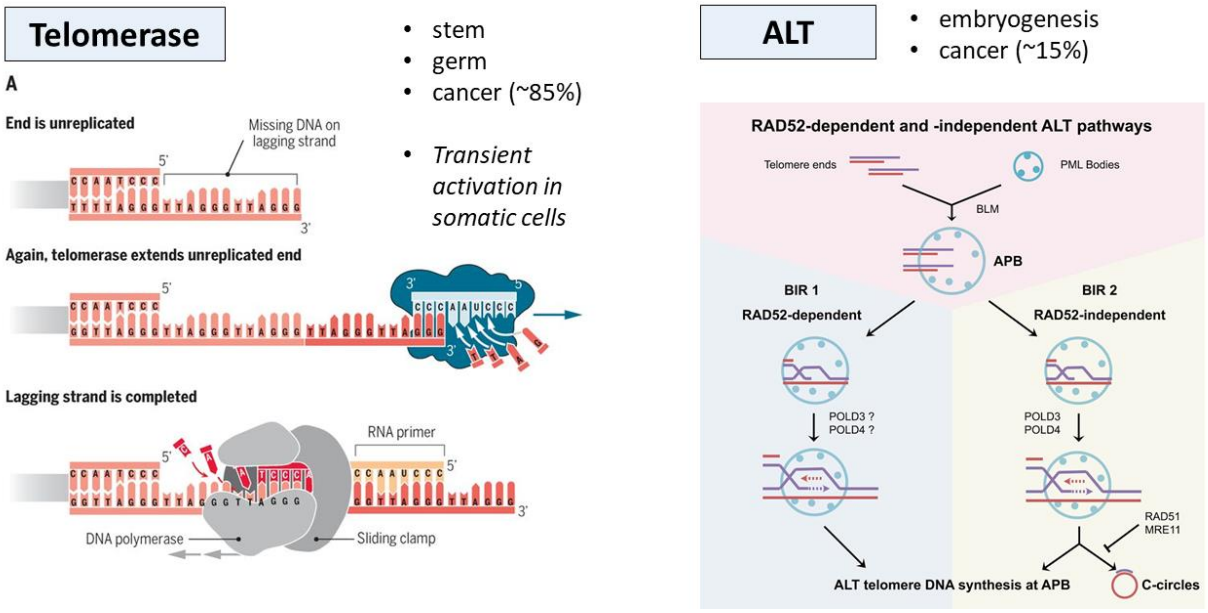

Figure 1.2. Telomere maintenance mechanisms. Telomerase is a ribonucleoprotein reverse transcriptase comprised of two core components, telomerase reverse transcriptase (TERT) and telomerase RNA component (TERC), which together enable elongation of telomeres. Alternative Lengthening of Telomeres (ALT) is a recombination-based mechanism that utilizes HR proteins for telomere elongation. Adapted from (Blackburn et al., 2015).

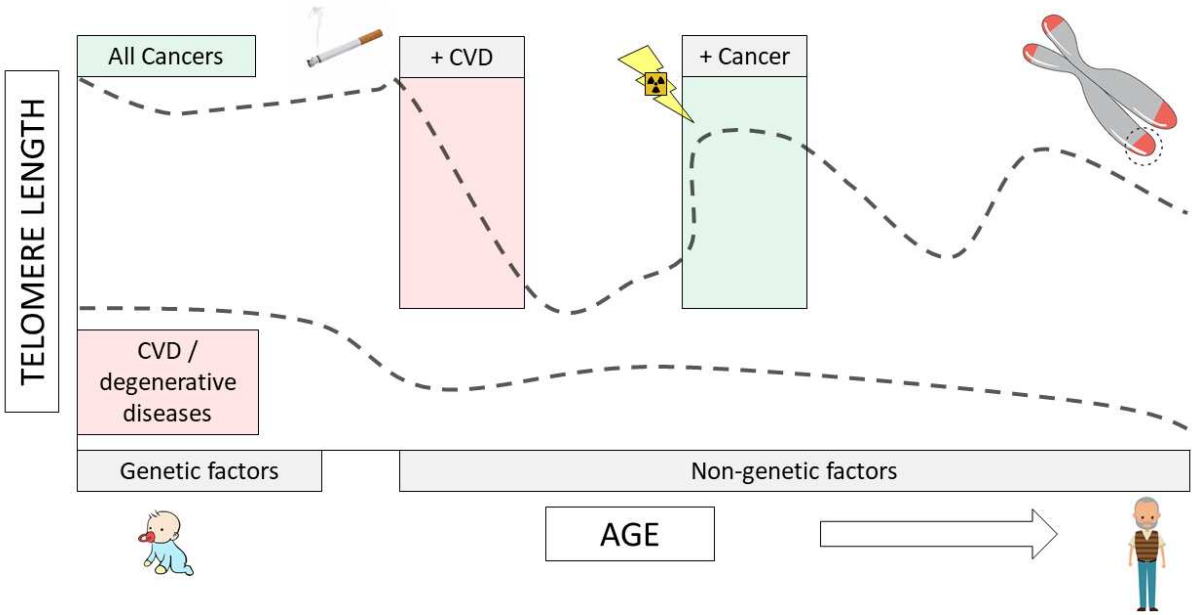

Figure 1.3. Telomere length and disease risk: genetic and non-genetic factors. At birth telomere length is most strongly influenced by genetic factors but with aging the influence of non-genetic factors on telomere length is greatest. Reductions in telomere length yield increase risks for cardiovascular disease (CVD) and other degenerative diseases i.e. Alzheimer's, whereas longer telomeres, especially in combination with DNA damage (indicated in yellow bolt), increase risks for cancer. Adapted from (Blackburn et al., 2015).

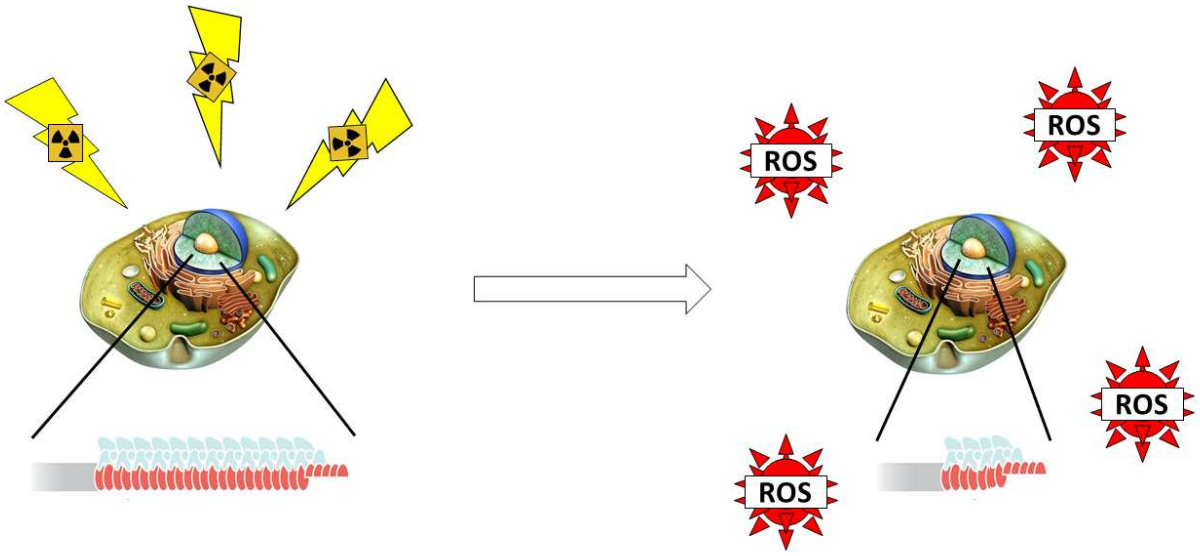

Figure 1.4. Telomeres, IR and oxidative stress. Telomeres are particularly sensitive to IR due to their high content of guanine bases which readily oxidize to 8-oxoG bases in the presence of reactive oxygen species (ROS) yielded from IR exposure. due to the induction of oxidative stress after exposure via reactive oxygen species (ROS). The 8-oxoG bases contribute to replication stress and have the net effect of shortening telomeres. Adapted from (Blackburn et al., 2015).

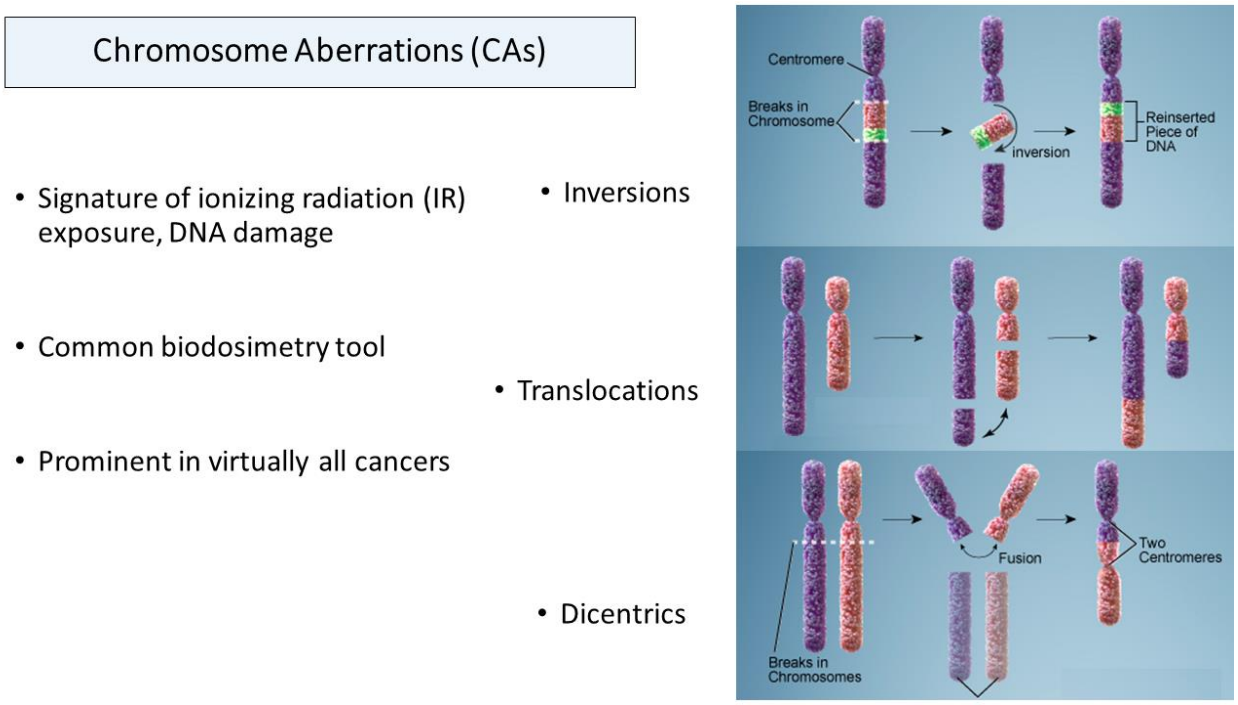

Figure 1.5. Chromosome aberrations. IR can induce prompt DSBs and with two or more DSBs close in space and time chromosome aberrations (CAs) can be formed via mis-repair of DNA damage. CAs thus can provide valuable signatures of IR exposure. Adapted from (Griffiths et al., 1999)

REFERENCES

Aguiar, P.H.N., Furtado, C., Repolês, B.M., Ribeiro, G.A., Mendes, I.C., Peloso, E.F., Gadelha, F.R., Macedo, A.M., Franco, G.R., Pena, S.D.J., et al. (2013). Oxidative Stress and DNA Lesions: The Role of 8-Oxoguanine Lesions in Trypanosoma cruzi Cell Viability. PLoS Negl. Trop. Dis. *7*, e2279.

Ahmed, S., Passos, J.F., Birket, M.J., Beckmann, T., Brings, S., Peters, H., Birch-Machin, M.A., von Zglinicki, T., and Saretzki, G. (2008). Telomerase does not counteract telomere shortening but protects mitochondrial function under oxidative stress. J. Cell Sci. *121*, 1046–1053.

Allsopp, R.C., Vaziri, H., Patterson, C., Goldstein, S., Younglai, E.V., Futcher, A.B., Greider, C.W., and Harley, C.B. (1992). Telomere length predicts replicative capacity of human fibroblasts. Proc. Natl. Acad. Sci. U. S. A. *89*, 10114–10118.

Arnoult, N., and Karlseder, J. (2015). Complex interactions between the DNA-damage response and mammalian telomeres. Nat. Struct. Mol. Biol. *22*, 859–866.

Arsenis, N.C., You, T., Ogawa, E.F., Tinsley, G.M., and Zuo, L. (2017). Physical activity and telomere length: Impact of aging and potential mechanisms of action. Oncotarget *8*, 45008–45019.

Astuti, Y., Wardhana, A., Watkins, J., and Wulaningsih, W. (2017). Cigarette smoking and telomere length: A systematic review of 84 studies and meta-analysis. Environ. Res. *158*, 480–489.

Aubert, G., and Lansdorp, P.M. (2008). Telomeres and aging. Physiol. Rev. *88*, 557–579.

Azzalin, C.M., Reichenbach, P., Khoriauli, L., Giulotto, E., and Lingner, J. (2007). Telomeric Repeat–Containing RNA and RNA Surveillance Factors at Mammalian Chromosome Ends. Science *318*, 798–801.

Azzam, E.I., Jay-Gerin, J.-P., and Pain, D. (2012). Ionizing radiation-induced metabolic oxidative stress and prolonged cell injury. Cancer Lett. *327*, 48–60.

Balakumar, P., Maung-U, K., and Jagadeesh, G. (2016). Prevalence and prevention of cardiovascular disease and diabetes mellitus. Pharmacol. Res. *113*, 600–609.

Bär, C., de Jesus, B.B., Serrano, R., Tejera, A., Ayuso, E., Jimenez, V., Formentini, I., Bobadilla, M., Mizrahi, J., de Martino, A., et al. (2014). Telomerase expression confers cardioprotection in the adult mouse heart after acute myocardial infarction. Nat. Commun. *5*, 5863.

Barcellos-Hoff, M.H., and Nguyen, D.H. (2009). RADIATION CARCINOGENESIS IN CONTEXT: HOW DO IRRADIATED TISSUES BECOME TUMORS? Health Phys. *97*, 446–457.

Baskar, R., Lee, K.A., Yeo, R., and Yeoh, K.-W. (2012). Cancer and Radiation Therapy: Current Advances and Future Directions. Int. J. Med. Sci. *9*, 193–199.

Baskar, R., Dai, J., Wenlong, N., Yeo, R., and Yeoh, K.-W. (2014). Biological response of cancer cells to radiation treatment. Front. Mol. Biosci. *1*.

Bejma, J., Ramires, P., and Ji, L.L. (2000). Free radical generation and oxidative stress with ageing and exercise: differential effects in the myocardium and liver. Acta Physiol. Scand. *169*, 343–351.

Blackburn, E.H., and Gall, J.G. (1978). A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. J. Mol. Biol. *120*, 33–53.

Blackburn, E.H., Epel, E.S., and Lin, J. (2015). Human telomere biology: A contributory and interactive factor in aging, disease risks, and protection. Science *350*, 1193–1198.

Blagoev, K.B., Goodwin, E.H., and Bailey, S.M. (2010). Telomere sister chromatid exchange and the process of aging. Aging *2*, 727–730.

Boldrini, L., Bibault, J.-E., Masciocchi, C., Shen, Y., and Bittner, M.-I. (2019). Deep Learning: A Review for the Radiation Oncologist. Front. Oncol. *9*.

Broer, L., Codd, V., Nyholt, D.R., Deelen, J., Mangino, M., Willemsen, G., Albrecht, E., Amin, N., Beekman, M., de Geus, E.J.C., et al. (2013). Meta-analysis of telomere length in 19 713 subjects reveals high heritability, stronger maternal inheritance and a paternal age effect. Eur. J. Hum. Genet. *21*, 1163–1168.

Bryan, T.M., and Reddel, R.R. (1994). SV40-induced immortalization of human cells. Crit. Rev. Oncog. *5*, 331–357.

Bryan, T.M., Englezou, A., Dalla-Pozza, L., Dunham, M.A., and Reddel, R.R. (1997). Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. Nat. Med. *3*, 1271–1274.

Calado, R.T., and Dumitriu, B. (2013). Telomere Dynamics in Mice and Humans. Semin. Hematol. *50*, 165–174.

Cesare, A., and Reddel, R. (2010). Alternative Lengthening of telomeres in Mammalian Cells. Madam Curie Biosci. DatabaseInternet.

Chancellor, J.C., Scott, G.B.I., and Sutton, J.P. (2014). Space Radiation: The Number One Risk to Astronaut Health beyond Low Earth Orbit. Life Open Access J. *4*, 491–510.

Chancellor, J.C., Blue, R.S., Cengel, K.A., Auñón-Chancellor, S.M., Rubins, K.H., Katzgraber, H.G., and Kennedy, A.R. (2018). Limitations in predicting the space radiation health risk for exploration astronauts. Npj Microgravity *4*, 1–11.

Chew, M.T., Jones, B., Hill, M., and Bradley, D.A. (2020). Radiation, a two-edged sword: From untoward effects to fractionated radiotherapy. Radiat. Phys. Chem. 108994.

Chimenti, C., Kajstura, J., Torella, D., Urbanek, K., Heleniak, H., Colussi, C., Di Meglio, F., Nadal-Ginard, B., Frustaci, A., Leri, A., et al. (2003). Senescence and death of primitive cells and myocytes lead to premature cardiac aging and heart failure. Circ. Res. *93*, 604–613.

Choi, D.K., Helenowski, I., and Hijiya, N. (2014). Secondary malignancies in pediatric cancer survivors: Perspectives and review of the literature. Int. J. Cancer *135*, 1764–1773.

Codd, V., Nelson, C.P., Albrecht, E., Mangino, M., Deelen, J., Buxton, J.L., Jan Hottenga, J., Fischer, K., Esko, T., Surakka, I., et al. (2013). Identification of seven loci affecting mean telomere length and their association with disease. Nat. Genet. *45*, 422-427e2.

Coluzzi, E., Buonsante, R., Leone, S., Asmar, A.J., Miller, K.L., Cimini, D., and Sgura, A. (2017). Transient ALT activation protects human primary cells from chromosome instability induced by low chronic oxidative stress. Sci. Rep. *7*, 43309.

Coluzzi, E., Leone, S., and Sgura, A. (2019). Oxidative Stress Induces Telomere Dysfunction and Senescence by Replication Fork Arrest. Cells *8*.

Conrad, S., Künzel, J., and Löbrich, M. (2011). Sister chromatid exchanges occur in G2-irradiated cells. Cell Cycle *10*, 222–228.

Cornforth, M.N. (2001). Analyzing Radiation-Induced Complex Chromosome Rearrangements by Combinatorial Painting. Radiat. Res. *155*, 643–659.

Cucinotta, F.A. (2015). A New Approach to Reduce Uncertainties in Space Radiation Cancer Risk Predictions. PLoS ONE *10*.

Dachev, T.P., Bankov, N.G., Tomov, B.T., Matviichuk, Y.N., Dimitrov, P.G., Häder, D.-P., and Horneck, G. (2017). Overview of the ISS Radiation Environment Observed during the ESA EXPOSE-R2 Mission in 2014–2016. Space Weather *15*, 1475–1489.

De Vitis, M., Berardinelli, F., Coluzzi, E., Marinaccio, J., O'Sullivan, R.J., and Sgura, A. (2019). X-rays Activate Telomeric Homologous Recombination Mediated Repair in Primary Cells. Cells *8*.

Desouky, O., Ding, N., and Zhou, G. (2015). Targeted and non-targeted effects of ionizing radiation. J. Radiat. Res. Appl. Sci. *8*, 247–254.

van Deursen, J.M. (2014). The role of senescent cells in ageing. Nature *509*, 439–446.

Dunbar, B. (2019). What is Artemis?

Dunham, M.A., Neumann, A.A., Fasching, C.L., and Reddel, R.R. (2000). Telomere maintenance by recombination in human cells. Nat. Genet. *26*, 447–450.

Eisenberg, D.T.A., and Kuzawa, C.W. (2018). The paternal age at conception effect on offspring telomere length: mechanistic, comparative and adaptive perspectives. Philos. Trans. R. Soc. B Biol. Sci. *373*, 20160442.

Entringer, S., de Punder, K., Buss, C., and Wadhwa, P.D. (2018). The fetal programming of telomere biology hypothesis: an update. Philos. Trans. R. Soc. B Biol. Sci. *373*, 20170151.

Feldberg, R.S., and Carew, J.A. (1981). Water Radiolysis Products and Nucleotide Damage in γ-irradiated DNA. Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. *40*, 11–17.

Ferrandon, S., Saultier, P., Carras, J., Battiston-Montagne, P., Alphonse, G., Beuve, M., Malleval, C., Honnorat, J., Slatter, T., Hung, N., et al. (2013). Telomere profiling: toward glioblastoma personalized medicine. Mol. Neurobiol. *47*, 64–76.

Förstermann Ulrich, and Münzel Thomas (2006). Endothelial Nitric Oxide Synthase in Vascular Disease. Circulation *113*, 1708–1714.

Fouquerel, E., Lormand, J., Bose, A., Lee, H.-T., Kim, G.S., Li, J., Sobol, R.W., Freudenthal, B.D., Myong, S., and Opresko, P.L. (2016). Oxidative guanine base damage regulates human telomerase activity. Nat. Struct. Mol. Biol. *23*, 1092–1100.

Garrett-Bakelman, F.E., Darshi, M., Green, S.J., Gur, R.C., Lin, L., Macias, B.R., McKenna, M.J., Meydan, C., Mishra, T., Nasrini, J., et al. (2019). The NASA Twins Study: A multidimensional analysis of a year-long human spaceflight. Science *364*.

Gordon, D.M., and Santos, J.H. (2010). The Emerging Role of Telomerase Reverse Transcriptase in Mitochondrial DNA Metabolism. J. Nucleic Acids *2010*.

Green, P.D., Sharma, N.K., and Santos, J.H. (2019). Telomerase Impinges on the Cellular Response to Oxidative Stress Through Mitochondrial ROS-Mediated Regulation of Autophagy. Int. J. Mol. Sci. *20*.

Greider, C.W., and Blackburn, E.H. (1987). The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. Cell *51*, 887–898.

Griffith, J.D., Comeau, L., Rosenfield, S., Stansel, R.M., Bianchi, A., Moss, H., and de Lange, T. (1999). Mammalian Telomeres End in a Large Duplex Loop. Cell *97*, 503–514.

Griffiths, A.J., Gelbart, W.M., Miller, J.H., and Lewontin, R.C. (1999). Chromosomal Rearrangements. Mod. Genet. Anal.

Hada, M., and Georgakilas, A.G. (2008). Formation of clustered DNA damage after high-LET irradiation: a review. J. Radiat. Res. (Tokyo) *49*, 203–210.

Haendeler Judith, Dröse Stefan, Büchner Nicole, Jakob Sascha, Altschmied Joachim, Goy Christine, Spyridopoulos Ioakim, Zeiher Andreas M., Brandt Ulrich, and Dimmeler Stefanie (2009). Mitochondrial Telomerase Reverse Transcriptase Binds to and Protects Mitochondrial DNA and Function From Damage. Arterioscler. Thromb. Vasc. Biol. *29*, 929–935.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646–674.

Hayflick, L., and Moorhead, P.S. (1961). The serial cultivation of human diploid cell strains. Exp. Cell Res. *25*, 585–621.

Heaphy, C.M., Subhawong, A.P., Hong, S.-M., Goggins, M.G., Montgomery, E.A., Gabrielson, E., Netto, G.J., Epstein, J.I., Lotan, T.L., Westra, W.H., et al. (2011). Prevalence of the Alternative Lengthening of Telomeres Telomere Maintenance Mechanism in Human Cancer Subtypes. Am. J. Pathol. *179*, 1608–1615.

Heasley, L.R., Watson, R.A., and Argueso, J.L. (2020). Punctuated Aneuploidization of the Budding Yeast Genome (Genomics).

Heidinger, B.J., Blount, J.D., Boner, W., Griffiths, K., Metcalfe, N.B., and Monaghan, P. (2012). Telomere length in early life predicts lifespan. Proc. Natl. Acad. Sci. U. S. A. *109*, 1743–1748.

Holzinger, A. (2014). Trends in Interactive Knowledge Discovery for Personalized Medicine: Cognitive Science meets Machine Learning. IEEE Intell. Inform. Bull. *15*, 6–14.

Hu, S., Kim, M.-H.Y., McClellan, G.E., and Cucinotta, F.A. (2009). Modeling the acute health effects of astronauts from exposure to large solar particle events. Health Phys. *96*, 465–476.

Hu, Y., Shi, G., Zhang, L., Li, F., Jiang, Y., Jiang, S., Ma, W., Zhao, Y., Songyang, Z., and Huang, J. (2016). Switch telomerase to ALT mechanism by inducing telomeric DNA damages and dysfunction of ATRX and DAXX. Sci. Rep. *6*, 32280.

Iannilli, F., Zalfa, F., Gartner, A., Bagni, C., and Dotti, C.G. (2013). Cytoplasmic TERT Associates to RNA Granules in Fully Mature Neurons: Role in the Translational Control of the Cell Cycle Inhibitor p15INK4B. PLOS ONE *8*, e66602.

Ishikawa, F. (1997). Telomere Crisis, the Driving Force in Cancer Cell Evolution. Biochem. Biophys. Res. Commun. *230*, 1–6.

Jornayvaz, F.R., and Shulman, G.I. (2010). Regulation of mitochondrial biogenesis. Essays Biochem. *47*.

Kang, J., Schwartz, R., Flickinger, J., and Beriwal, S. (2015). Machine Learning Approaches for Predicting Radiation Therapy Outcomes: A Clinician's Perspective. Int. J. Radiat. Oncol. *93*, 1127–1135.

Keith, S., Doyle, J.R., Harper, C., Mumtaz, M., Tarrago, O., Wohlers, D.W., Diamond, G.L., Citra, M., and Barber, L.E. (2012). OVERVIEW OF BASIC RADIATION PHYSICS, CHEMISTRY, AND BIOLOGY (Agency for Toxic Substances and Disease Registry (US)).

Kim, J.D., McCarter, R.J., and Yu, B.P. (1996). Influence of age, exercise, and dietary restriction on oxidative stress in rats. Aging Milan Italy *8*, 123–129.

Kim, N.W., Piatyszek, M.A., Prowse, K.R., Harley, C.B., West, M.D., Ho, P.L., Coviello, G.M., Wright, W.E., Weinrich, S.L., and Shay, J.W. (1994). Specific association of human telomerase activity with immortal cells and cancer. Science *266*, 2011–2015.

Kim, S., Parks, C.G., DeRoo, L.A., Chen, H., Taylor, J.A., Cawthon, R.M., and Sandler, D.P. (2009). Obesity and Weight Gain in Adulthood and Telomere Length. Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol. *18*, 816.

Koltsova, A.S., Pendina, A.A., Efimova, O.A., Chiryaeva, O.G., Kuznetzova, T.V., and Baranov, V.S. (2019). On the Complexity of Mechanisms and Consequences of Chromothripsis: An Update. Front. Genet. *10*.

Krittanawong, C., Zhang, H., Wang, Z., Aydar, M., and Kitai, T. (2017). Artificial Intelligence in Precision Cardiovascular Medicine. J. Am. Coll. Cardiol. *69*, 2657–2664.

Lange, T. de (2005). Shelterin: the protein complex that shapes and safeguards human telomeres. Genes Dev. *19*, 2100–2110.

Lee, H.-T., Bose, A., Lee, C.-Y., Opresko, P.L., and Myong, S. (2017). Molecular mechanisms by which oxidative DNA damage promotes telomerase activity. Nucleic Acids Res. *45*, 11752–11765.

Lee, M., Hills, M., Conomos, D., Stutz, M.D., Dagg, R.A., Lau, L.M.S., Reddel, R.R., and Pickett, H.A. (2014). Telomere extension by telomerase and ALT generates variant repeats by mechanistically distinct processes. Nucleic Acids Res. *42*, 1733–1746.

Leri, A., Franco, S., Zacheo, A., Barlucchi, L., Chimenti, S., Limana, F., Nadal-Ginard, B., Kajstura, J., Anversa, P., and Blasco, M.A. (2003). Ablation of telomerase and telomere loss leads to cardiac dilatation and heart failure associated with p53 upregulation. EMBO J. *22*, 131–139.

Levy, M.Z., Allsopp, R.C., Futcher, A.B., Greider, C.W., and Harley, C.B. (1992). Telomere end-replication problem and cell aging. J. Mol. Biol. *225*, 951–960.

Li, Y., Wu, F.-X., and Ngom, A. (2018). A review on machine learning principles for multi-view biological data integration. Brief. Bioinform. *19*, 325–340.

Lin, C.-C., Wu, L.S.-H., and Lee, K.-F. (2019). The Potential Effect of Different Doses of Ionizing Radiation on Genes and Disease. Dose-Response *17*.

Little, J.B. (2000). Radiation carcinogenesis. Carcinogenesis *21*, 397–404.

Liu, L., Bailey, S.M., Okuka, M., Muñoz, P., Li, C., Zhou, L., Wu, C., Czerwiec, E., Sandler, L., Seyfang, A., et al. (2007). Telomere lengthening early in development. Nat. Cell Biol. *9*, 1436–1441.

Lloyd, D.C., Purrott, R.J., and Reeder, E.J. (1980). The incidence of unstable chromosome aberrations in peripheral blood lymphocytes from unirradiated and occupationally exposed people. Mutat. Res. *72*, 523–532.

Londoño-Vallejo, J.A., Der-Sarkissian, H., Cazes, L., Bacchetti, S., and Reddel, R.R. (2004). Alternative Lengthening of Telomeres Is Characterized by High Rates of Telomeric Exchange. Cancer Res. *64*, 2324–2327.

Longhese, M.P. (2008). DNA damage response at functional and dysfunctional telomeres. Genes Dev. *22*, 125–140.

Lundblad, V., and Blackburn, E.H. (1993). An alternative pathway for yeast telomere maintenance rescues est1- senescence. Cell *73*, 347–360.

Lustig, A., Shterev, I., Geyer, S., Shi, A., Hu, Y., Morishita, Y., Nagamura, H., Sasaki, K., Maki, M., Hayashi, I., et al. (2016). Long term effects of radiation exposure on telomere lengths of leukocytes and its associated biomarkers among atomic-bomb survivors. Oncotarget.

Maciejowski, J., Li, Y., Bosco, N., Campbell, P.J., and de Lange, T. (2015). Chromothripsis and Kataegis Induced by Telomere Crisis. Cell *163*, 1641–1654.

Maeda, T., Nakamura, K., Atsumi, K., Hirakawa, M., Ueda, Y., and Makino, N. (2013). Radiation-associated changes in the length of telomeres in peripheral leukocytes from inpatients with cancer. Int. J. Radiat. Biol. *89*, 106–109.

Maida, Y., Yasukawa, M., Furuuchi, M., Lassmann, T., Possemato, R., Okamoto, N., Kasim, V., Hayashizaki, Y., Hahn, W.C., and Masutomi, K. (2009). An RNA-dependent RNA polymerase formed by TERT and the RMRP RNA. Nature *461*, 230–235.

Maida, Y., Yasukawa, M., and Masutomi, K. (2016). De Novo RNA Synthesis by RNA-Dependent RNA Polymerase Activity of Telomerase Reverse Transcriptase. Mol. Cell. Biol. *36*, 1248–1259.

Martens, U.M., Chavez, E.A., Poon, S.S., Schmoor, C., and Lansdorp, P.M. (2000). Accumulation of short telomeres in human fibroblasts prior to replicative senescence. Exp. Cell Res. *256*, 291–299.

Matthews, C., Gorenne, I., Scott, S., Figg, N., Kirkpatrick, P., Ritchie, A., Goddard, M., and Bennett, M. (2006). Vascular smooth muscle cells undergo telomere-based senescence in human atherosclerosis: effects of telomerase and oxidative stress. Circ. Res. *99*, 156–164.

McClintock, B. (1941). The Stability of Broken Ends of Chromosomes in Zea Mays. Genetics *26*, 234–282.

McEachern, M.J., and Blackburn, E.H. (1996). Cap-prevented recombination between terminal telomeric repeat arrays (telomere CPR) maintains telomeres in Kluyveromyces lactis lacking telomerase. Genes Dev. *10*, 1822–1834.

McKenna, M.J., Robinson, E., Taylor, L., Tompkins, C., Cornforth, M.N., Simon, S.L., and Bailey, S.M. (2019). Chromosome Translocations, Inversions and Telomere Length for Retrospective Biodosimetry on Exposed U.S. Atomic Veterans. Radiat. Res. *191*, 311–322.

McKevitt, T.P., Nasir, L., Wallis, C.V., and Argyle, D.J. (2003). A cohort study of telomere and telomerase biology in cats. Am. J. Vet. Res. *64*, 1496–1499.

Mirabello, L., Yu, K., Kraft, P., De Vivo, I., Hunter, D.J., Prescott, J., Wong, J.Y.Y., Chatterjee, N., Hayes, R.B., and Savage, S.A. (2010). The association of telomere length and genetic variation in telomere biology genes. Hum. Mutat. *31*, 1050–1058.

Mirjolet, C., Boidot, R., Saliques, S., Ghiringhelli, F., Maingon, P., and Créhange, G. (2015). The role of telomeres in predicting individual radiosensitivity of patients with cancer in the era of personalized radiotherapy. Cancer Treat. Rev. *41*, 354–360.

Morales, C.P., Holt, S.E., Ouellette, M., Kaur, K.J., Yan, Y., Wilson, K.S., White, M.A., Wright, W.E., and Shay, J.W. (1999). Absence of cancer–associated changes in human fibroblasts immortalized with telomerase. Nat. Genet. *21*, 115–118.

Morin, G.B. (1989). The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. Cell *59*, 521–529.

Morin, S.J., Eccles, J., Iturriaga, A., and Zimmerman, R.S. (2017). Translocations, inversions and other chromosome rearrangements. Fertil. Steril. *107*, 19–26.

Morishita, M., Muramatsu, T., Suto, Y., Hirai, M., Konishi, T., Hayashi, S., Shigemizu, D., Tsunoda, T., Moriyama, K., and Inazawa, J. (2016). Chromothripsis-like chromosomal rearrangements induced by ionizing radiation using proton microbeam irradiation system. Oncotarget *7*, 10182–10192.

Moyzis, R.K., Buckingham, J.M., Cram, L.S., Dani, M., Deaven, L.L., Jones, M.D., Meyne, J., Ratliff, R.L., and Wu, J.R. (1988). A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. Proc. Natl. Acad. Sci. U. S. A. *85*, 6622–6626.

Mozdarani, H. (2012). Biological Complexities in Radiation Carcinogenesis and Cancer Radiotherapy: Impact of New Biological Paradigms. Genes *3*, 90–114.

Muller, H.J. (1938). The remaking of chromosomes. Collect. Net *13*, 181–198.

Nagasawa, H., and Little, J.B. (1992). Induction of sister chromatid exchanges by extremely low doses of alpha-particles. Cancer Res. *52*, 6394–6396.

Navarro, A., Gomez, C., López-Cepero, J.M., and Boveris, A. (2004). Beneficial effects of moderate exercise on mice aging: survival, behavior, oxidative stress, and mitochondrial electron transfer. Am. J. Physiol. Regul. Integr. Comp. Physiol. *286*, R505-511.

Neumann, A.A., Watson, C.M., Noble, J.R., Pickett, H.A., Tam, P.P.L., and Reddel, R.R. (2013). Alternative lengthening of telomeres in normal mammalian somatic cells. Genes Dev. *27*, 18–23.

Ng, A.K., Kenney, L.B., Gilbert, E.S., and Travis, L.B. (2010). Secondary malignancies across the age spectrum. Semin. Radiat. Oncol. *20*.

Okamoto, K., and Seimiya, H. (2019). Revisiting Telomere Shortening in Cancer. Cells *8*.

Okuda, K., Bardeguez, A., Gardner, J.P., Rodriguez, P., Ganesh, V., Kimura, M., Skurnick, J., Awad, G., and Aviv, A. (2002). Telomere length in the newborn. Pediatr. Res. *52*, 377–381.

Olovnikov, A.M. (1973). A theory of marginotomy. The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. J. Theor. Biol. *41*, 181–190.

Ornish, D., Lin, J., Chan, J.M., Epel, E., Kemp, C., Weidner, G., Marlin, R., Frenda, S.J., Magbanua, M.J.M., Daubenmier, J., et al. (2013). Effect of comprehensive lifestyle changes on telomerase activity and telomere length in men with biopsy-proven low-risk prostate cancer: 5-year follow-up of a descriptive pilot study. Lancet Oncol. *14*, 1112–1120.

Pai, G.S., Thomas, G.H., Mahoney, W., and Migeon, B.R. (1980). Complex chromosome rearrangements:Report of a new case and literature review. Clin. Genet. *18*, 436–444.

Podlevsky, J.D., and Chen, J.J.-L. (2012). It all comes together at the ends: telomerase structure, function, and biogenesis. Mutat. Res. *730*, 3–11.

Poiesz, B.J., Seal, G., and Loeb, L.A. (1974). Reverse transcriptase: correlation of zinc content with activity. Proc. Natl. Acad. Sci. U. S. A. *71*, 4892–4896.

Protsenko, E., Rehkopf, D., Prather, A.A., Epel, E., and Lin, J. (2020). Are long telomeres better than short? Relative contributions of genetically predicted telomere length to neoplastic and non-neoplastic disease risk and population health burden. PLOS ONE *15*, e0240185.

Radák, Z., Chung, H.Y., Naito, H., Takahashi, R., Jung, K.J., Kim, H.-J., and Goto, S. (2004). Age-associated increase in oxidative stress and nuclear factor kappaB activation are attenuated in rat liver by regular exercise. FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol. *18*, 749–750.

Ray, F.A., Zimmerman, E., Robinson, B., Cornforth, M.N., Bedford, J.S., Goodwin, E.H., and Bailey, S.M. (2013). Directional genomic hybridization for chromosomal inversion discovery and detection. Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol. *21*, 165–174.

Ray, F.A., Robinson, E., McKenna, M., Hada, M., George, K., Cucinotta, F., Goodwin, E.H., Bedford, J.S., Bailey, S.M., and Cornforth, M.N. (2014). Directional genomic hybridization: inversions as a potential biodosimeter for retrospective radiation exposure. Radiat. Environ. Biophys. *53*, 255–263.

Rudd, M.K., Friedman, C., Parghi, S.S., Linardopoulou, E.V., Hsu, L., and Trask, B.J. (2007). Elevated Rates of Sister Chromatid Exchange at Chromosome Ends. PLoS Genet *3*, e32.

Sallam, N., and Laher, I. (2016). Exercise Modulates Oxidative Stress and Inflammation in Aging and Cardiovascular Diseases. Oxid. Med. Cell. Longev. *2016*.

Salvestrini, V., Sell, C., and Lorenzini, A. (2019). Obesity May Accelerate the Aging Process. Front. Endocrinol. *10*.

Santos, J.H., Meyer, J.N., Skorvaga, M., Annab, L.A., and Van Houten, B. (2004). Mitochondrial hTERT exacerbates free-radical-mediated mtDNA damage. Aging Cell *3*, 399–411.

Scheller Madrid, A., Rode, L., Nordestgaard, B.G., and Bojesen, S.E. (2016). Short Telomere Length and Ischemic Heart Disease: Observational and Genetic Studies in 290 022 Individuals. Clin. Chem. *62*, 1140–1149.

Schuster, B., Ellmann, A., Mayo, T., Auer, J., Haas, M., Hecht, M., Fietkau, R., and Distel, L.V. (2018). Rate of individuals with clearly increased radiosensitivity rise with age both in healthy individuals and in cancer patients. BMC Geriatr. *18*, 105.

Seimiya, H., Sawada, H., Muramatsu, Y., Shimizu, M., Ohko, K., Yamane, K., and Tsuruo, T. (2000). Involvement of 14-3-3 proteins in nuclear localization of telomerase. EMBO J. *19*, 2652–2661.

Shammas, M.A. (2011). Telomeres, lifestyle, cancer, and aging: Curr. Opin. Clin. Nutr. Metab. Care *14*, 28–34.

Sharma, N.K., Reyes, A., Green, P., Caron, M.J., Bonini, M.G., Gordon, D.M., Holt, I.J., and Santos, J.H. (2012). Human telomerase acts as a hTR-independent reverse transcriptase in mitochondria. Nucleic Acids Res. *40*, 712–725.

Shay, J.W. (2016). Role of Telomeres and Telomerase in Aging and Cancer. Cancer Discov. *6*, 584–593.

Shay, J.W., and Bacchetti, S. (1997). A survey of telomerase activity in human cancer. Eur. J. Cancer Oxf. Engl. 1990 *33*, 787–791.

Shay, J.W., and Wright, W.E. (2010). Telomeres and telomerase in normal and cancer stem cells. FEBS Lett. *584*, 3819–3825.

Shay, J.W., Pereira-Smith, O.M., and Wright, W.E. (1991). A role for both RB and p53 in the regulation of human cellular senescence. Exp. Cell Res. *196*, 33–39.

Shippen, D.E., and McKnight, T.D. (1998). Telomeres, telomerase and plant development. Trends Plant Sci. *3*, 126–130.

Shippen-Lentz, D., and Blackburn, E.H. (1990). Functional evidence for an RNA template in telomerase. Science *247*, 546–552.

Singh, S.K., Szulik, M.W., Ganguly, M., Khutsishvili, I., Stone, M.P., Marky, L.A., and Gold, B. (2011). Characterization of DNA with an 8-oxoguanine modification. Nucleic Acids Res. *39*, 6789–6801.

Sishc, B.J., Nelson, C.B., McKenna, M.J., Battaglia, C.L.R., Herndon, A., Idate, R., Liber, H.L., and Bailey, S.M. (2015). Telomeres and Telomerase in the Radiation Response: Implications for Instability, Reprograming, and Carcinogenesis. Front. Oncol. *5*, 257.

Somanathan, I., and Baysdorfer, C. (2018). A bioinformatics approach to identify telomere sequences. BioTechniques *65*, 20–25.

Stimpson, K.M., Matheny, J.E., and Sullivan, B.A. (2012). Dicentric chromosomes: unique models to study centromere function and inactivation. Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol. *20*, 595–605.

Turner, S., Wong, H.P., Rai, J., and Hartshorne, G.M. (2010). Telomere lengths in human oocytes, cleavage stage embryos and blastocysts. Mol. Hum. Reprod. *16*, 685–694.

US EPA, O. (2014). Radiation Basics.

Vaz, B., Vuotto, C., Valvo, S., D'Ambra, C., Esposito, F.M., Chiurchiù, V., Devine, O., Sanchez, M., Borsellino, G., Gilroy, D., et al. (2020). Intercellular telomere transfer extends T cell lifespan. BioRxiv 2020.10.09.331918.

Vidaček, N.Š., Nanić, L., Ravlić, S., Sopta, M., Gerić, M., Gajski, G., Garaj-Vrhovac, V., and Rubelj, I. (2017). Telomeres, Nutrition, and Longevity: Can We Really Navigate Our Aging? J. Gerontol. Ser. A.

Vodenkova, S., Kroupa, M., Polivkova, Z., Musak, L., Ambrus, M., Schneiderova, M., Kozevnikovova, R., Vodickova, L., Rachakonda, S., Hemminki, K., et al. (2020). Chromosomal damage and telomere length in peripheral blood lymphocytes of cancer patients. Oncol. Rep. *44*, 2219–2230.

Watson, J.D. (1972). Origin of Concatemeric T7DNA. Nature. New Biol. *239*, 197–201.

Weckselblatt, B., and Rudd, M.K. (2015). Human Structural Variation: Mechanisms of Chromosome Rearrangements. Trends Genet. TIG *31*, 587–599.

Wentzensen, I.M., Mirabello, L., Pfeiffer, R.M., and Savage, S.A. (2011). The Association of Telomere Length and Cancer: A Meta-Analysis. Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol. *20*, 1238–1250.

Werner, C., Hanhoun, M., Widmann, T., Kazakov, A., Semenov, A., Pöss, J., Bauersachs, J., Thum, T., Pfreundschuh, M., Müller, P., et al. (2008). Effects of physical exercise on myocardial telomere-regulating proteins, survival pathways, and apoptosis. J. Am. Coll. Cardiol. *52*, 470–482.

Whittemore, K., Vera, E., Martínez-Nevado, E., Sanpera, C., and Blasco, M.A. (2019). Telomere shortening rate predicts species life span. Proc. Natl. Acad. Sci. U. S. A. *116*, 15122–15127.

Wilkins, R.C., Romm, H., Kao, T.-C., Awa, A.A., Yoshida, M.A., Livingston, G.K., Jenkins, M.S., Oestreicher, U., Pellmar, T.C., and Prasanna, P.G.S. (2008). Interlaboratory Comparison of the Dicentric Chromosome Assay for Radiation Biodosimetry in Mass Casualty Events. Radiat. Res. *169*, 551–560.

Wojcik, A., Bruckmann, E., and Obe, G. (2004). Insights into the mechanisms of sister chromatid exchange formation. Cytogenet. Genome Res. *104*, 304–309.

Wu, X., Amos, C.I., Zhu, Y., Zhao, H., Grossman, B.H., Shay, J.W., Luo, S., Hong, W.K., and Spitz, M.R. (2003). Telomere Dysfunction: A Potential Cancer Predisposition Factor. JNCI J. Natl. Cancer Inst. *95*, 1211–1218.

Xu, C., Wang, Z., Su, X., Da, M., Yang, Z., Duan, W., and Mo, X. (2020). Association between leucocyte telomere length and cardiovascular disease in a large general population in the United States. Sci. Rep. *10*, 80.

Yan, P., Benhattar, J., Seelentag, W., Stehle, J.-C., and Bosman, F.T. (2004). Immunohistochemical localization of hTERT protein in human tissues. Histochem. Cell Biol. *121*, 391–397.

Zglinicki, T.V. (2000). Role of Oxidative Stress in Telomere Length Regulation and Replicative Senescence. Ann. N. Y. Acad. Sci. *908*, 99–110.

Zhan, Y., and Hägg, S. (2019). Telomere length and cardiovascular disease risk. Curr. Opin. Cardiol. *34*, 270–274.

Zhan, Y., Karlsson, I.K., Karlsson, R., Tillander, A., Reynolds, C.A., Pedersen, N.L., and Hägg, S. (2017). Exploring the Causal Pathway From Telomere Length to Coronary Heart Disease: A Network Mendelian Randomization Study. Circ. Res. *121*, 214–219.

CHAPTER 2: TEMPORAL TELOMERE AND DNA DAMAGE RESPONSES IN THE SPACE RADIATION ENVIRONMENT[1]

Chapter summary

The NASA Twins study investigated the impacts of spaceflight on health during a one-year mission aboard the ISS for the space twin Scott Kelly. During spaceflight, significant telomere elongation was observed which rapidly returned to baseline, ending shorter at post-flight than before. DNA damage responses were also observed in-flight and persisted at post-flight. Here, we aimed to determine if the spaceflight-specific telomere elongation observed in the space twin was experienced by other astronauts, as well as ascertain the mechanism of elongation and the incidence of DNA damage responses. Much like the results of NASA's one-year mission twin astronaut, significantly longer telomeres were observed during spaceflight for two astronauts undergoing six-month missions. Significant telomere elongation was observed regardless of cell or sample type (blood, urine), or analysis (qPCR, Telo-FISH, Nanopore-sequencing). As expected for chronic exposure to the space radiation environment, DNA damage responses were significantly elevated in-flight and persisted post-flight (inversions). Importantly, we find telomeric and cytogenetic evidence implicating ALT as the mechanism responsible for telomere elongation observed during spaceflight.

[1] This chapter is a draft of a first author manuscript accepted at Cell Reports. Anticipated publication date is Nov-Dec 2020.

Authors:

Jared J. Luxton[1,2], Miles J. McKenna[1,2], Lynn E. Taylor[1], Kerry A. George[3], Sara R. Zwart[4], Brian Crucian[5], Viktor R. Drel[6], Francine E. Garrett-Bakelman[7,8,9], Daniel Butler[10], Jonathan Foox[10], Kirill Grigorev[10], Daniela Bezdan[10], Cem Meydan[10] Scott M. Smith[5], Kumar Sharma[6], Christopher E. Mason[10,11,12*], and Susan M. Bailey[1,2*]

Affiliations:

[1]Department of Environmental and Radiological Health Sciences, and [2]Cell and Molecular Biology Program, Colorado State University, Fort Collins, CO, USA

[3]KBR, Houston, TX, USA

[4] UTMB, Galveston, TX, USA

[5]NASA, Houston, TX, USA

[6]Center for Renal Precision Medicine, UT Health San Antonio, TX, USA

[7]Department of Medicine, Weill Cornell Medicine, New York, NY, USA

[8]Department of Medicine, and [9]Department of Biochemistry & Molecular Genetics, University of Virginia, Charlottesville, VA, USA

[10]Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY, USA

[11]The WorldQuant Initiative for Quantitative Prediction, Weill Cornell Medicine, New York, NY, USA

[12]The HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine, Weill Cornell Medicine, New York, NY, USA

Overview

Telomeres, repetitive terminal features of chromosomes essential for maintaining genome integrity, shorten with cell division, lifestyle factors and stresses, and environmental exposures, and so provide a robust biomarker of health, aging, and age-related diseases. We assessed telomere length dynamics in three unrelated astronauts before, during, and after one-year or six-month missions aboard the International Space Station (ISS). Similar to our results for NASA's One-Year Mission twin astronaut (Garrett-Bakelman et al., 2019), significantly longer telomeres were observed during spaceflight for two six-month mission astronauts. Furthermore, telomere length shortened rapidly after return to Earth for all three crewmembers and, overall, telomere length tended to be shorter after spaceflight than before. Consistent with chronic exposure to the space radiation environment, signatures of persistent DNA damage responses were also detected, including mitochondrial and oxidative stress, inflammation, and telomeric and chromosomal aberrations, which together provide potential mechanistic insight into spaceflight-specific telomere elongation.

One Sentence Summary:

Consistent with findings first observed for NASA's One-Year Mission twin astronaut, spaceflight-specific telomere elongation was also observed in six-month ISS mission crewmembers, as were signatures of persistent DNA damage responses in the space radiation environment, which together provide potential mechanistic insight into telomere maintenance pathways.

Introduction

The ends of human chromosomes are capped by telomeres, tandem arrays of repetitive G-rich sequence bound by a plethora of associated proteins that protect chromosomal termini from degradation and loss (de Lange, 2009; Moyzis et al., 1988). Telomeres preserve genome stability by preventing the ends of chromosomes from being recognized as broken DNA (double-strand breaks [DSBs]) and triggering inappropriate DNA damage responses (DDRs). Due to the end-replication problem (inability to replicate to the very end of a linear DNA molecule), telomere length erodes with cell division, causing telomeres to shorten until they reach a critically short length, at which point a permanent cell cycle arrest known as replicative senescence is entered (Harley et al., 1990; Olovnikov, 1971; Watson, 1972). The specialized reverse transcriptase, telomerase, can counteract telomere attrition via *de novo* addition of telomeric repeats onto the ends of newly replicated chromosomes (Greider and Blackburn, 1985). However, telomerase activity is prominent only in highly proliferative stem, germline, and cancer cells; its levels are insufficient to maintain telomere length in normal somatic cells.

While it is well established that telomere length is an inherited trait that diminishes with normal aging (Aubert and Lansdorp, 2008; Honig et al., 2015), oxidative stress, and inflammation (von Zglinicki, 2000; Zhang et al., 2016), it is becoming increasingly appreciated that telomere length is influenced by a variety of other factors as well, including sex (Gardner et al., 2013) and lifestyle factors [e.g., diet (Gu et al., 2015), smoking and obesity (Valdes et al., 2005), physical activity (Cherkas et al., 2008), psychological stress (Epel et al., 2004)], chronic stress and disease (Blackburn and Epel, 2012). Environmental exposures, such as air pollution, UV and ionizing radiations, also affect telomere length (Miri et al., 2019; Rochette and Brash, 2010; Shim et al., 2014). Therefore, telomere length maintenance represents a key integrating

component of the cumulative effects of genetic, lifestyle, and environmental factors; i.e., the rate at which telomeres shorten provides an informative biomarker of general health and aging. Furthermore, telomere dysfunction and/or altered telomere length can be linked to age-related pathologies, ranging from reduced immune function and dementia, to cardiovascular disease (CVD) and cancer (Cohen et al., 2013; Haycock et al., 2014; Honig et al., 2012; Shay, 2013a, b).

Reports of both shortened and elongated telomeres after exposure to various types of radiation (Berardinelli et al., 2013; Maeda et al., 2013) have implications for a variety of degenerative late effects, particularly CVD, as an inverse relationship between telomere length in peripheral blood cells and coronary heart disease has been demonstrated (Haycock et al., 2014). Such late health effects relevant to spaceflight are largely unknown and controversial (Cucinotta et al., 2016; Delp et al., 2016), yet they have very real potential for influencing performance during long-duration missions. Considering the unique stresses and environmental exposures (galactic cosmic rays [GCR], solar particle events [SPE], and secondary particles that arise from interactions with spacecraft shielding) experienced during long-duration spaceflight, we proposed that telomere length dynamics (changes over time) is of particular relevance for astronauts because it provides insight into health status during a mission, as well as potential implications for aging and long-term health outcomes thereafter.

Here, we assessed telomere length dynamics and DDRs in three unrelated astronauts before, during, and after one-year or six-month missions aboard the International Space Station (ISS). Importantly, in all samples analyzed (blood, urine) and irrespective of mission duration or means of measurement (qPCR, Telo-FISH, Nanopore sequencing), significantly longer telomeres were observed during spaceflight. Additionally, and as for the One-Year Mission astronaut, telomere length shortened rapidly after return to Earth for the six-month mission crewmembers.

Although overall, average telomere length tended to be shorter after spaceflight than before, individual differences were detected. Consistent with chronic exposure to the space radiation environment and concomitant DDRs during spaceflight, persistent mitochondrial and oxidative stress, inflammation, and telomeric and chromosomal aberrations also were observed. In addition to providing further support for spaceflight-specific telomere elongation, the findings reported here provide potential mechanistic insight involving persistent DDRs and telomere length maintenance pathways, as well as reveal important differences in individual long-term responses.

Results

Telomere elongation during spaceflight

We previously observed spaceflight-specific telomere elongation for the One-Year Mission twin astronaut (TW), as compared to his ground-based twin (HR), during NASA's first One Year Mission (Garrett-Bakelman et al., 2019). Supportive of that somewhat surprising finding, we now report significantly longer average telomere length and increased numbers of long telomeres during spaceflight for two additional (unrelated) astronauts on shorter duration, six-month missions aboard the ISS.

For all three astronauts (and all ground control subjects), telomere length was measured in blood before, during, and after spaceflight using two independent assays. First, quantitative (q)PCR provided a measure of average telomere length in DNA isolated from peripheral blood mononuclear cells (PBMCs) (**Figure 2.1A)**. Cell-by-cell analysis was also performed using telomere fluorescence *in situ* hybridization (Telo-FISH) on metaphase chromosomes (stimulated T-cells), which provided telomere length measures for thousands of individual telomeres based on relative fluorescence intensity (**Figure 2.1B; Figure 2.2**). In addition to average telomere

length, Telo-FISH data was used to generate individual telomere length distributions, which facilitated monitoring shifts in the populations of short and long telomeres (**Figure 2.3)**. While telomere length in the healthy age and sex-matched ground control cohort (n=11) remained relatively stable over the course of the study (**Figure 2.3A**), dramatic shifts in the populations of short and long telomeres were observed in all three astronauts (**Figures 2.3B-D**). Compared to before and after spaceflight, and consistent with increased average telomere length during spaceflight, many more *long* telomeres were present during spaceflight. Inter-individual differences in the magnitude of telomere elongation during spaceflight were detected; however, differences in telomere lengthening as a function of spaceflight duration (one-year vs. six-months) were not observed.

For the One-Year Mission astronaut, and for the first time in astronauts, longer telomeres during spaceflight were also validated using Nanopore sequencing (DNA isolated from ambient return, sorted CD4+ T-cells) (**Figure 2.4A**). Although sequencing through extensive tracks of repetitive DNA has historically been challenging (Lee et al., 2017), recent technological advances have made obtaining significantly longer reads, and thus sequencing of telomeres possible (Grigorev/Mason, et al., *companion submission to Genome Res*). Importantly, here sequencing results were independently verified by both qPCR and Telo-FISH assessment of telomere length in the same blood-derived samples (Garrett-Bakelman et al., 2019).

Also unprecedented, and consistent with results first reported for the One-Year Mission astronaut (Garrett-Bakelman et al., 2019), telomere length shortened rapidly (within days) after return to Earth for the six-month mission astronauts (**Figures 2.1 and 2.2**). Although overall, average telomere length for the three crewmembers stabilized to near baseline values by the end of the study (**Figure 2.1**), average telomere length tended to be shorter after spaceflight than

before. Furthermore, shifts in individual telomere length distributions revealed the presence of many more *short* telomeres (and fewer long) after spaceflight than before (**Figure 2.3**). Inter-individual differences in individual telomere length distributions were also observed, in that one six-month mission crewmember displayed a decreased number of short telomeres after spaceflight, as well as a corresponding increase in the number of long telomeres.

Lastly, telomere length was measured in DNA extracted from urine samples collected during the NASA Twins Study (**Figure 2.4B**). Similar to the telomere elongation seen in all blood samples collected during spaceflight (Garrett-Bakelman et al., 2019), significantly longer average telomere length was observed in urine samples collected and frozen on the ISS, and returned to Earth by cargo vehicle **(Figure 2.5)**. Telomere length in urine also shortened rapidly after return to Earth, and was shorter after spaceflight than before. These findings represent the first report of telomere length in urine during spaceflight, provide additional support for spaceflight-specific telomere elongation, and may be a more practical means of assessing in-flight telomere length during future long-duration missions.

Telomerase activity

Although telomerase activity is the most obvious contender for *de novo* telomere elongation during spaceflight, it was not possible to directly access telomerase activity in samples collected during spaceflight. Similar to our results with samples collected during spaceflight for the One-Year Mission twin astronaut, telomerase activity was "*lost in space*" for both six-month mission crewmembers, again likely due to the unavoidable heat and/or time associated with return of ambient blood samples from the ISS to Earth on the Soyuz vehicle (Garrett-Bakelman et al., 2019). All three unrelated astronauts had lower levels of telomerase activity than the ground control cohort before and after spaceflight (**Figure 2.6A**), consistent

with TW having lower levels of telomerase activity than HR (**Figure 2.6B**) (Garrett-Bakelman et al., 2019), and similar to astronauts in general (*Luxton et al., companion paper Cell, 2020*). TW's post-flight levels of telomerase activity were similar to his pre-flight levels; however, when taken together, the three unrelated crewmembers had significantly elevated telomerase activity after spaceflight (**Figure 2.6A**). Although this finding could be potentially suggestive of a role for telomerase in the telomere elongation observed during spaceflight, it is more likely related to individual differences, as elevated telomerase activity would be expected after spaceflight for all crewmembers, which was not the case (**Figure 2.6B, Figure 2.7**, *Luxton et al., companion paper Cell, 2020*).

Inter-individual differences in telomerase activity post-spaceflight most likely reflect individual differences in responding to life stress (Deng et al., 2016). Indeed, *highly* elevated levels of telomerase activity were observed at the final collection time point of the Twin Study (270 days after Return to Earth; R+270) – for *both* TW and HR; therefore, the response was not related to spaceflight (**Figure 2.6C**). Consistent with acute psychological stress increasing telomerase activity, the twins had experienced a traumatic personal family loss just prior to sample collection, explaining such high levels of telomerase activity at this particular timepoint, and supportive of the intimate relationships between human real-life stresses and factors that regulate telomere length maintenance. It is worth noting, that although telomerase activity was elevated at R+270 for both twins, a corresponding increase in telomere length was not observed (Garrett-Bakelman et al., 2019), suggestive of non-canonical, extra-telomeric roles of telomerase as previously demonstrated (Low and Tergaonkar, 2013).

Persistent mitochondrial stress during spaceflight

The multi-omics approach employed by the NASA Twins Study provided ample evidence of mitochondrial stress during spaceflight for the One Year Mission astronaut [(Garrett-Bakelman et al., 2019); *companion papers Silveira/ Beheshti et al., Cell, 2020; Bezdan et al., iScience, 2020*]. Persistent mitochondrial stress during spaceflight likely occurs in response to chronic DNA damage; mtDNA has been shown to trigger innate immune responses and inflammation, which was significantly elevated during spaceflight in the three crewmembers reported here (**Figure 2.7**). Furthermore, mtDNA has recently been proposed as a stress sentinel, eliciting a protective signaling response that enhances nuclear DNA repair (Wu et al., 2019), a view also consistent with findings reported in the Twins Study (Garrett-Bakelman et al., 2019). The findings of increased mtDNA and mtRNA during spaceflight are supportive of increased mitochondrial biogenesis, and potentially related telomere elongation, as a positive correlation between mtDNA copy number and telomere length in healthy adults has been demonstrated (Tyrka et al., 2015).

Interestingly, AMP-activated protein kinase (AMPK) is activated in skeletal muscle during aerobic exercise and is involved in regulating diverse processes, including mitochondrial biogenesis (Winder et al., 2000). Astronauts routinely exercise during spaceflight as a countermeasure for the muscle wasting and bone loss associated with living in a microgravity environment (Loehr et al., 2015; Scott et al., 2019). Consistent with previous reports [e.g., (Puterman et al., 2018)], we have shown that 12 weeks of aerobic exercise training significantly increased telomere length in skeletal muscle ($p<0.001$), and further, that telomere length measured in blood is representative of telomere length in skeletal muscle (Konopka et al., 2019). Moreover, during the One Year Mission, telomere length increased in CD8+ T-cells (Garrett-Bakelman et al., 2019), a cell population shown to experience increased telomere length

immediately after acute bouts of aerobic exercise (Simpson et al., 2010). Taken together, such experimental and observational studies are suggestive of mechanistic relationships between chronic DNA damage, persistent mitochondrial stress, intense aerobic exercise, and the telomere elongation observed during spaceflight; however, they are far from conclusive.

Chronic oxidative stress during spaceflight

In addition to persistent mitochondrial stress during spaceflight, astronauts in general (Silveira/Beheshti et al., *companion paper, Cell, 2020*), including those reported here, experience increased oxidative stress and inflammation during spaceflight (**Figure 2.8**) (Garrett-Bakelman et al., 2019). Ionizing radiation exposures, particularly high Linear Energy Transfer (LET)/High atomic number (Z) and Energy (HZE) exposures, induce persistent reactive oxygen species (ROS) production by mitochondria, DNA damage signaling, and chronic oxidative stress (Azzam et al., 2012; Sridharan et al., 2015), conditions under which telomerase is excluded from the nucleus and relocated to mitochondria, thereby improving mitochondrial function (Haendeler et al., 2003; Saretzki, 2009). Additionally, oxidative base damage has been shown to play a dual role in regulating telomerase-mediated elongation of human telomeres (Fouquerel et al., 2016). Incorporation of free oxidized dNTPs (8-oxodG) by telomerase *inhibits* telomerase activity and terminates telomere elongation, presumably due to replication fork arrest at telomeres (Coluzzi et al., 2019). This finding helps to explain why acute *in vitro* exposures to oxidative stress shorten telomeres. In contrast, preexisting 8-oxodG within telomeric DNA, as would be expected from chronic exposure to oxidative stress during spaceflight, *enhances* telomerase activity by destabilizing G-quadruplex secondary structures, therefore unrepaired 8-oxodG lesions at telomeres may promote telomere lengthening.

Furthermore, due to their G-rich nature, telomeric regions are especially susceptible to oxidative damage, which tends to accumulate because telomeres are refractory to repair (Miller et al., 2011). Of potential relevance to spaceflight-specific telomere elongation, persistent DDRs have been implicated in activating Alternative Lengthening of Telomeres (ALT) – a telomerase independent, recombination-based pathway of telomere maintenance (Bryan et al., 1997; Dunham et al., 2000; Murnane et al., 1994). In particular, chronic oxidative stress has been shown to transiently activate ALT (Coluzzi et al., 2017; Hu et al., 2012). In the present study, cytogenetic analysis revealed evidence of oxidative damage in all three astronauts during spaceflight (**Figure 2.9A**). Frequencies of sister chromatid exchange (SCE), recombination events indicative of replication stress (oxidative damage) (Wilson and Thompson, 2007), increased during spaceflight, and declined after spaceflight. Frequencies of terminal SCE, specifically involving the heterochromatin junction of the subtelomere of chromosome 2p, also increased during spaceflight. Terminal SCE likely reflect persistent DNA damage at a fragile site, consistent with the individual variation observed. SCE within subtelomeres themselves was not significantly elevated during spaceflight. Importantly, and specifically involving telomeres, frequencies of heterogeneous sister telomere lengths (**Figure 2.9A, C**), a recombination event characteristic of ALT (Bailey et al., 2004), were significantly elevated specifically during spaceflight. Lastly, frequencies of "fragile" telomeres, those displaying stringy, diffuse, and/or multiple telomeric signals on a single chromatid (Sfeir et al., 2009), as might be expected from microgravity-mediated decompaction of chromatin, were not elevated in samples collected during spaceflight.

Chronic space radiation exposure

In addition to the oxidative damage and replication stress that occurs at telomeres during spaceflight, chronic low dose exposure to the space radiation environment would result in persistent breakage of telomeric strands, and at least occasionally, produce telomeric DSBs. It was recently proposed that persistent telomere DSBs are also capable of enhancing recombination-mediated telomere elongation and induction of ALT phenotypes (Liu et al., 2018), including, telomere length heterogeneity, as observed here during spaceflight (**Figure 2.9A**).

Chromosome aberrations are well-established signatures of ionizing radiation-induced DNA damage (specifically, DSBs) that have been previously evaluated in blood lymphocytes of astronauts (Cucinotta, 2014; Garrett-Bakelman et al., 2019; George et al., 2013). In the present study, strand-specific directional genomic hybridization (dGH) paints (Ray et al., 2013) for chromosomes 1, 2, and 3, together with chromosome-specific subtelomere probes to distinguish terminal rearrangements [as in (McKenna and Bailey, 2017)] were used to simultaneously detect rearrangements *between* (dicentrics, translocations), and *within* (inversions) chromosomes. Consistent with our previous works and proposition that inversions are particularly informative biomarkers of exposure to high LET and space radiations (Garrett-Bakelman et al., 2019; McKenna et al., 2019; Ray et al., 2014), inversions were significantly increased during spaceflight for all three astronauts (**Figure 2.9B**). Furthermore, increased frequencies of inversions persisted after spaceflight for all three crewmembers, potentially suggestive of stem cell damage, clonal hematopoiesis (growth advantage), and/or instability [(Garrett-Bakelman et al., 2019), Trinchant/Hassane et al., *companion paper, Cell Press 2020*].

Satellite associations of acrocentric chromosomes have been proposed as early biomarkers of chronic exposure to low levels of ionizing radiation in occupationally exposed

hospital workers (Caradonna, 2015). Here, significantly increased frequencies of satellite associations were observed in all three astronauts specifically during spaceflight (**Figure 2.9B).** Satellite associations are nucleoprotein bridges involving the short p arms of 2-5 human acrocentric chromosomes, which were recently confirmed as linkages between ribosomal (r)DNA (Potapova et al., 2019), suggesting that transcriptionally active DNA loci engage in physical inter-chromosomal connections, perhaps as an integral feature of genome organization. Consistent with this view, we find that satellite associations are DNA interactions that do not involve telomeres, subtelomeres, or the chromosome most proximal to them (**Figure 2.9D-H**). Occasionally, satellite associations displayed chromosome 1, 2, or 3 sequence-specific signal, likely representing rDNA. With regard to space radiation exposure, satellite associations may be reflective of enhanced transcription of DNA repair related genes – consistent with (Garrett-Bakelman et al., 2019) – and, as such, could provide valuable indicators of chronic exposure during spaceflight.

Discussion

Marking a new era of human endeavors in space, NASA is planning to send the first woman and the next man to the moon by 2024, and targeting human exploration of Mars by the mid-2030s. Historically, a total of only 563 individuals have participated in spaceflight, the vast majority of whom have been males aged 35-55 years (average age of first-time astronaut is 40 years), on missions shorter than 20 days duration (Smith, 2020). As the number and diversity of people experiencing significantly longer duration spaceflight increases and “space tourism” becomes a reality, a better understanding of how spaceflight affects short- and long-term health is critical to maintaining astronaut performance during, and improving health outcomes

following, future long-duration missions. This information will also serve to inform both preemptive interventions [e.g., (Marshall-Goebel et al., 2019)], and personalized monitoring of individual crewmembers after their mission. Importantly, longitudinal monitoring of telomere length (dynamics) provides an integrated and informative assessment of general health, as well as a robust biomarker of aging and age-related disease trajectories, notably CVD and cancer, during and after long-duration missions.

For the One Year Mission astronaut, significantly longer telomeres specifically during spaceflight were observed, and then telomere length decreased rapidly upon return to Earth (Garrett-Bakelman et al., 2019). Increased mitochondrial and oxidative stress were also observed during spaceflight, as were elevated frequencies of chromosomal inversions that persisted after spaceflight, suggestive of on-going genomic instability. In the present study, we assessed telomere length dynamics, oxidative stress and inflammation, and telomeric and chromosomal DNA damage in two unrelated astronauts who participated in six-month missions aboard the ISS for whom in-flight samples were available, and compared results to those of the One-Year Mission astronaut. Irrespective of mission duration and consistent with our previous findings, similar biological changes were observed during spaceflight, and in general most factors returned to their preflight baseline values after spaceflight; however, some did not, specifically, telomere length and inversion frequencies.

Interestingly, the telomerase-independent, recombination-mediated ALT mechanism for maintaining telomere length can be transiently activated by persistent DNA damage induced by exposures to high LET radiations (Berardinelli et al., 2010), and low levels of chronic oxidative stress, including those induced by ionizing radiation (Coluzzi et al., 2017; De Vitis et al., 2019), and/or telomeric DSBs (Liu et al., 2018). We speculate that persistent, unrepaired telomeric

DNA damage induced by exposure to the space radiation environment during long-duration spaceflight (persistent production of ROS by mitochondria, chronic oxidative stress, persistent telomeric DSBs) triggers ALT or ALT-like mechanisms, particularly in normal cells possessing very low levels of telomerase, thereby resulting in ALT phenotypes, herein evidenced as elongated and heterogenous sister telomere lengths. During periods of chronic oxidative stress, telomerase can be exported from the nucleus to the mitochondria, improving their function (Haendeler et al., 2003; Saretzki, 2009), consistent with observations of increased mtDNA and mtRNA during spaceflight (Garrett-Bakelman et al., 2019) (*Silveira/Beheshti, companion paper Cell, 2020*). When persistent space radiation-induced ROS and oxidative stress are reduced upon return to Earth, telomerase would be expected to return to the nucleus, possibly accounting for the rapid reduction in telomere length observed, as the enzymatic component of telomerase (hTERT) has been shown to mediate shortening of excessively long telomeres in ALT and human primary cells (Zheng et al., 2014). Even so, the dramatic telomere length changes associated with spaceflight occurred very quickly, and there tended to be many more short telomeres after spaceflight than before.

While the definitive mechanisms involved in these extraterrestrial processes remain elusive, previously proposed changes in radiation-induced cell population dynamics may also contribute to the overall telomere lengthening observed during spaceflight (Garrett-Bakelman et al., 2019), but again, the rapid reduction in telomere length after return to Earth must considered. Additionally, while both one-year and six-month mission astronauts presented with more long telomeres during spaceflight, inter-individual variations in the magnitude of the response were observed. Furthermore, significantly more short telomeres were observed after spaceflight for two astronauts, independent of mission duration (one-year and six-month mission crewmembers

at R+270), whereas significantly more long telomeres were observed after spaceflight for one astronaut (six-month mission crewmember, at R+180). These results highlight the importance of monitoring individual astronauts, as they demonstrate individual variation in response and, therefore, individual differences in susceptibilities that influence long-term health outcomes. Indeed, accumulating evidence supports an emerging view of telomere length not only as a biomarker, but also a determinant of CVD and cancer; i.e., shorter telomeres and increased risk of age-related (degenerative) CVD on the one hand, longer telomeres and increased cancer risk (proliferative pathology) on the other (Stone et al., 2016).

We report for the first time that telomere length in urine was also increased during spaceflight for the One-Year Mission astronaut, and decreased rapidly after spaceflight, mirroring the telomere length changes observed in blood. These findings could reflect a general increase in telomere length and/or a kidney-specific increase in telomere length during spaceflight. Limited information exists regarding telomere length and telomerase activity in the human kidney; however, in general, telomere length is thought to decrease with age, acute oxidative stress, inflammation, and dysregulation of the renin-angiotensin system (Vasan et al., 2008; Yang and Fogo, 2010). Consistent with studies demonstrating that exercise stimulates mitochondrial biogenesis in the kidneys of mice (Dugan et al., 2013; Liu et al., 2019), endurance exercise during spaceflight could potentially increase mtRNA, mtDNA, and telomere length in urine, as observed here.

Cytogenetic analyses revealed that all three unrelated astronauts had higher background frequencies of inversions than other types of chromosome aberrations, findings consistent with our previous studies (Garrett-Bakelman et al., 2019; McKenna et al., 2019; Ray et al., 2014). Also consistent, and as would be expected from chronic, low dose exposure to the space

radiation environment (and induction of DSBs), intra-chromosomal inversions significantly increased during spaceflight and did so at a higher rate than other types of aberrations. Elevated frequencies of inversions persisted after spaceflight, supportive of on-going genomic instability and potentially increased cancer risk as a function of spaceflight (Behjati et al., 2016; Bonassi et al., 2004). Given this additional evidence, inversions are proving to be particularly informative for monitoring exposures to densely ionizing high-LET radiations, such as those encountered during spaceflight (Garrett-Bakelman et al., 2019; Hada et al., 2007; Ray et al., 2014). Furthermore, satellite associations dramatically increased during spaceflight and returned to baseline after spaceflight, consistent with them serving as early biomarkers of chronic, low-dose ionizing radiation exposure, as previously proposed (Caradonna, 2015).

In summary, multiple assays (including sequencing), in all cell-types evaluated (including urine), confirmed spaceflight-specific telomere elongation in three unrelated astronauts during one-year and six-month missions aboard the ISS; rapid telomere shortening after return to Earth was also observed in all three crewmembers. We propose that in the context of chronic exposure to the space radiation environment, persistent DDRs to ROS production by mitochondria, elevated levels of oxidative damage and replication stress, and low dose high LET radiations, together at damaged telomeres, transiently enhance telomerase activity and/or activate ALT or ALT-like recombinational mechanisms that are at least partially responsible for the dramatic shifts in telomere length dynamics observed. Current and future studies are aimed at assessing larger cohorts of astronauts, which include women and various duration missions, and will help to further elucidate the mechanisms involved in telomere elongation during spaceflight, as well as the potential influence and sources of persistent DNA damage. Furthermore, identifying inter-individual differences in response to spaceflight represents a critical next step for ensuring

astronaut health and performance during, and improving health outcomes following, upcoming lunar missions and beyond.

Limitations: There are a myriad of limitations associated with long-duration spaceflight that unavoidably influence experimental design and execution. In addition to a relatively small and similar population of astronauts, and the challenges associated with collection and return of in-flight samples, contributing factors such as microgravity, space radiation exposure, psychological stress, nutrition, and exercise, cannot be isolated and tested, so definitive mechanisms underlying the dramatic shifts in telomere length dynamics observed during spaceflight are difficult, if not impossible to identify. Furthermore, these studies involve high-profile humans whose health and confidentiality must be maintained, living in an extreme environment that cannot be fully replicated on Earth.

Figures

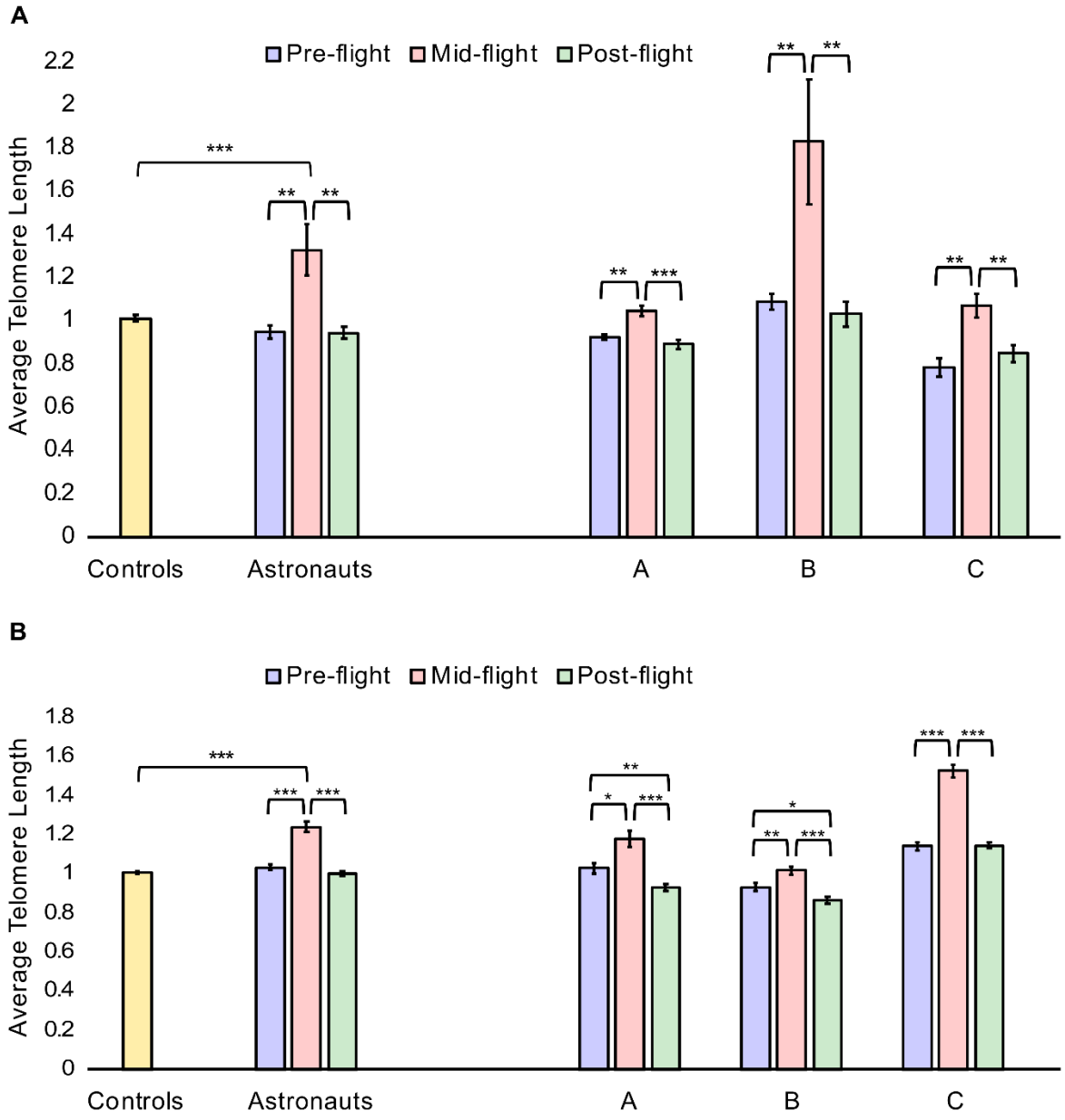

Figure 2.1. Telomere length dynamics before, during, and after long-duration spaceflight. Relative average telomere length in one-year and six-month mission astronauts (n=3). A. *Telomere length assessment by qPCR* (DNA isolated from PBMCs). Astronauts combined, and individual crewmember means (A, B, C), normalized to healthy age and sex matched ground control cohort (n=11). B. *Telomere length assessment by Telo-FISH* (metaphase chromosomes; T-cells). Astronauts combined, and individual crewmember means (A, B, C), normalized to healthy age and sex matched ground control cohort (n=11). For both assays, significance was tested by one-way ANOVA and post-hoc Student's t-tests for multiple comparisons. Error bars are SEM. $p<0.05$*, $p<0.01$**, $p<0.001$***.

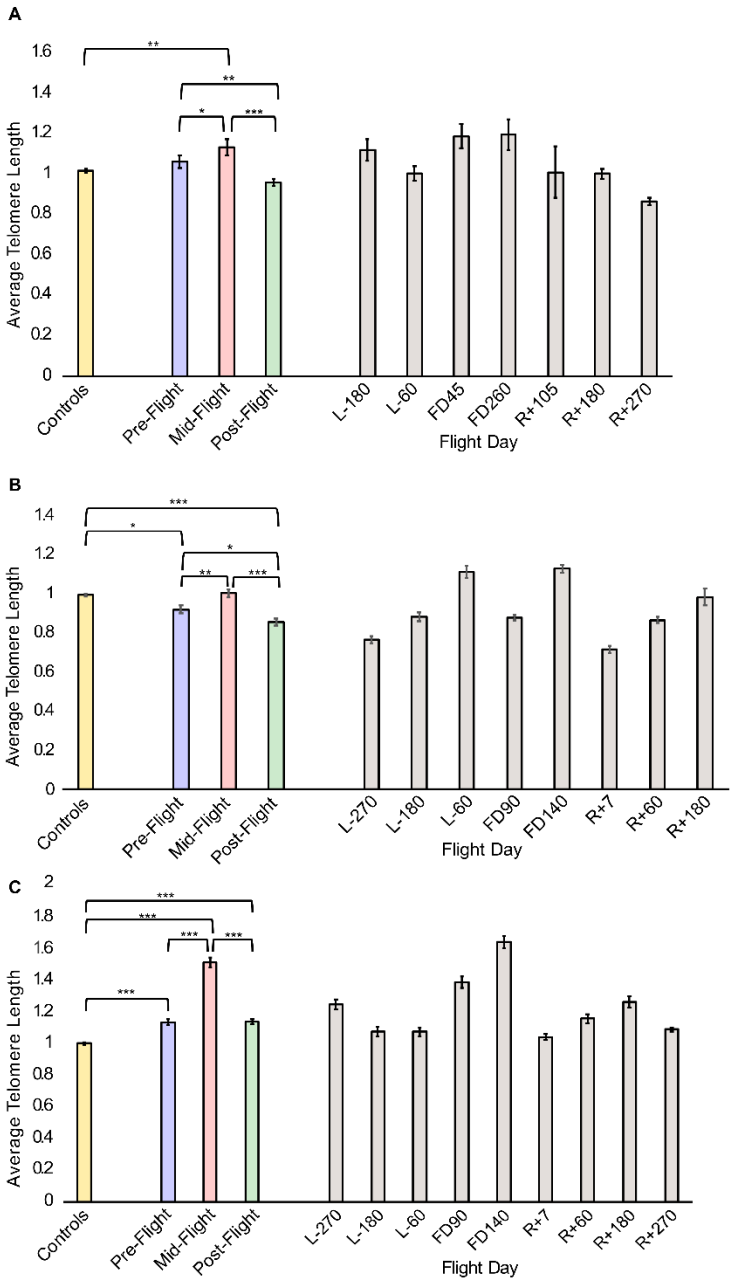

Figure 2.2. Telomere length time-course for one-year and six-month mission astronauts. Relative average telomere length (in blood) assessed by Telo-FISH before (Launch minus X days; L-X), during (Flight Day; FD), and after (Return plus X days; R+X) spaceflight for **A.** crewmember A, **B.** crewmember B, and **C.** crewmember C. For each crewmember, combined pre-, mid-, and post-flight means, and individual time-course (gray bars), normalized to age and sex matched ground control cohort (yellow bar; n=11) are shown. Significance tested by one-way ANOVA and post-hoc Student's t-tests for multiple comparisons. Error bars are SEM. $p<0.05$*, $p<0.01$**, $p<0.001$***.

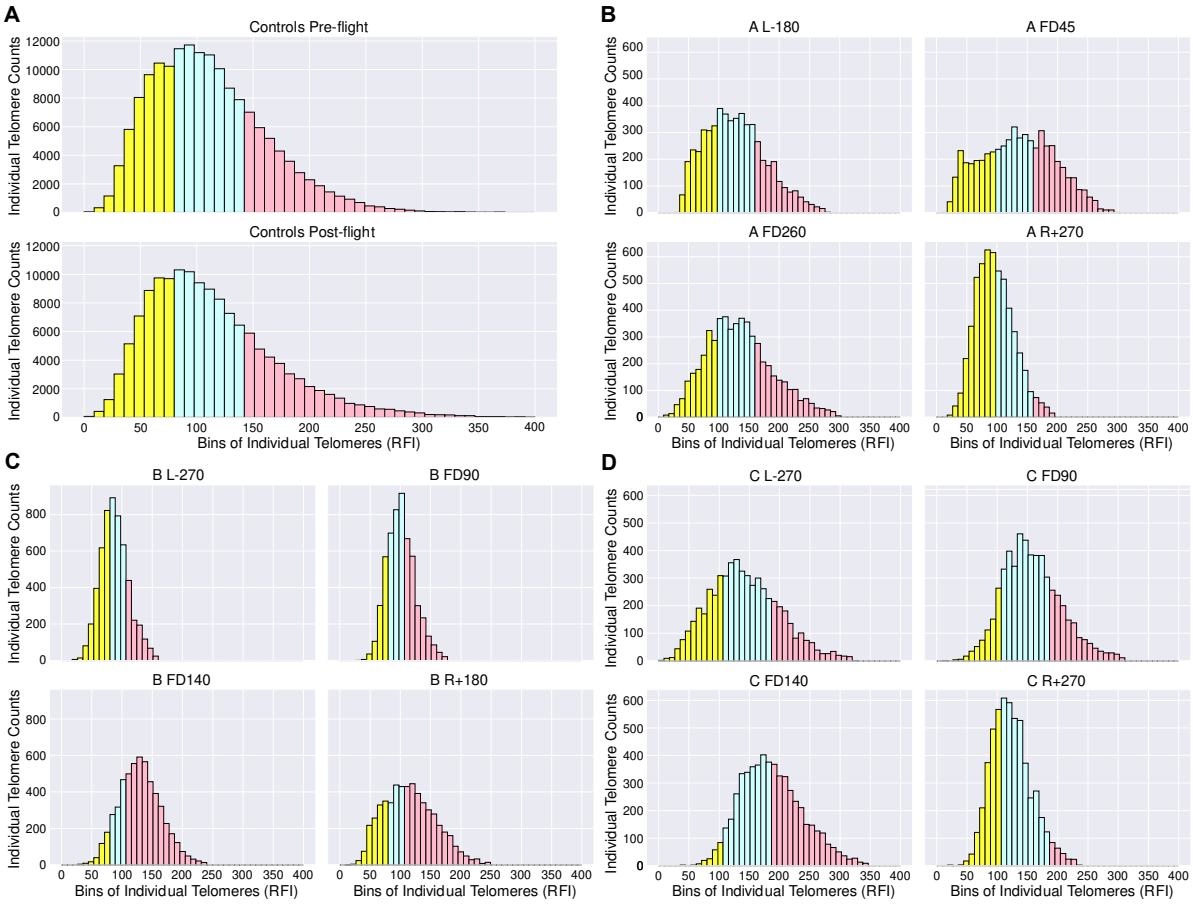

Figure 2.3. Telomere length distributions. Telo-FISH assessment of individual telomeres facilitated monitoring shifts in populations of short and long telomeres before (Launch minus X days; L-X), during (Flight Day; FD), and after (Return plus X days; R+X) spaceflight. Telomere length distributions from the earliest available pre-flight sample for each individual were split into quartiles, with telomeres designated as shortest 25% (yellow), middle 50% (blue), and longest 25% (red). Quartile cut-off values, unique for each individual, were then applied to subsequent timepoints. A. Age-sex matched ground control cohort distributions remain relatively stable over the course of the study. In contrast, astronauts' distributions shift dramatically: B. crewmember A; C. crewmember B; D. crewmember C. Thirty metaphases per timepoint/individual were scored. RFI: Relative fluorescence intensity.

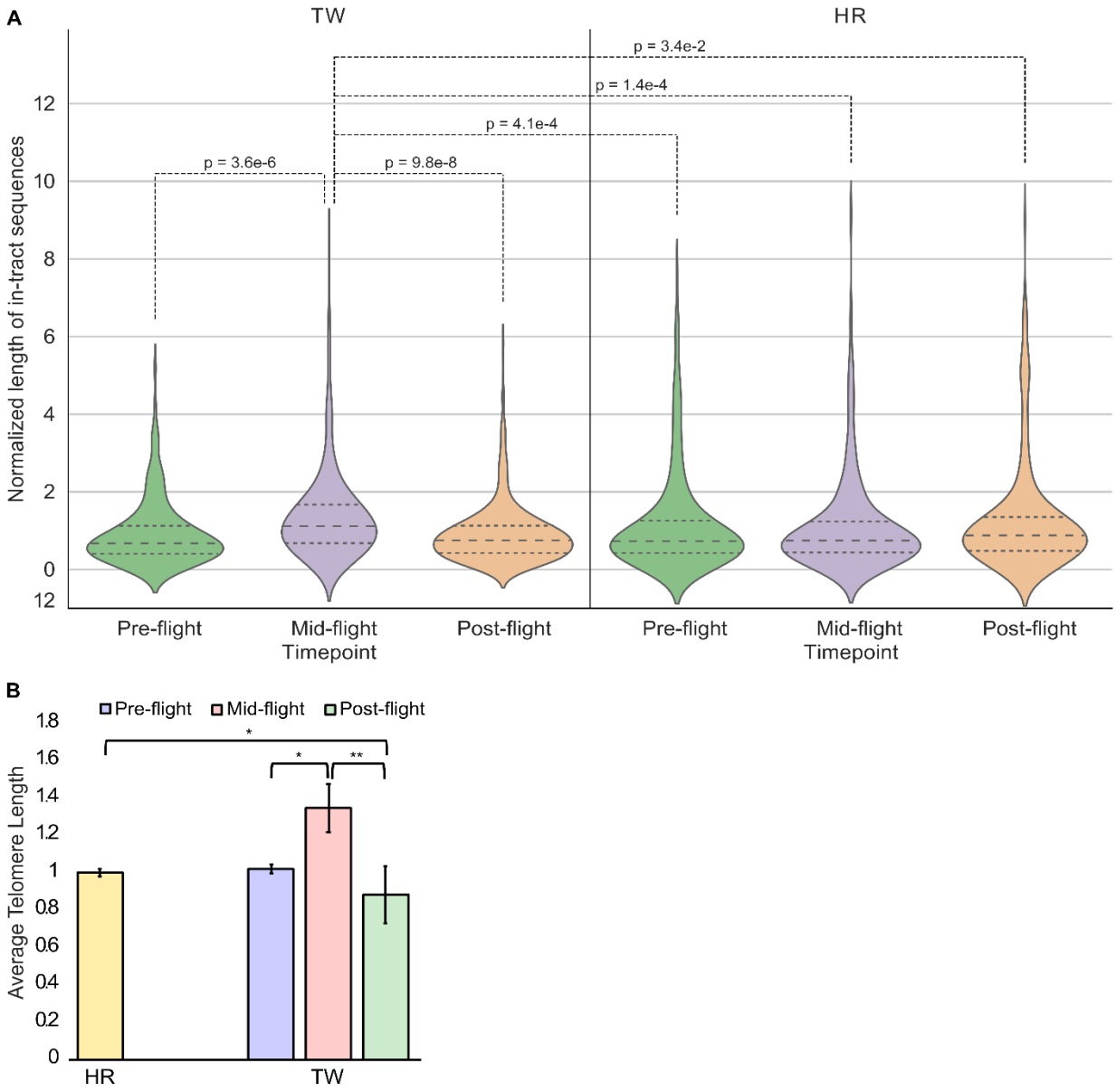

Figure 2.4. Telomere length assessment by Nanopore sequencing (blood) and qPCR in urine. Telomere length determinations for TW (space twin) and HR (ground twin) before, during, and after the One Year Mission. **A.** Telomere elongation during spaceflight was corroborated by Nanopore sequencing of DNA isolated from ambient return, sorted CD4+ T-cells; median (long dashes) and upper/lower quartiles (short dashes) are annotated. Significance was tested by Mann-Whitney U test and a Bonferroni correction for p-values. **B.** Similar telomere length dynamics for both twins were observed in DNA extracted from urine (qPCR); all values averaged for HR. Significance was tested by one-way ANOVA and post-hoc Student's t-tests for multiple comparisons. Error bars are SEM. p $p<0.05$*, $p<0.01$**, $p<0.001$***.

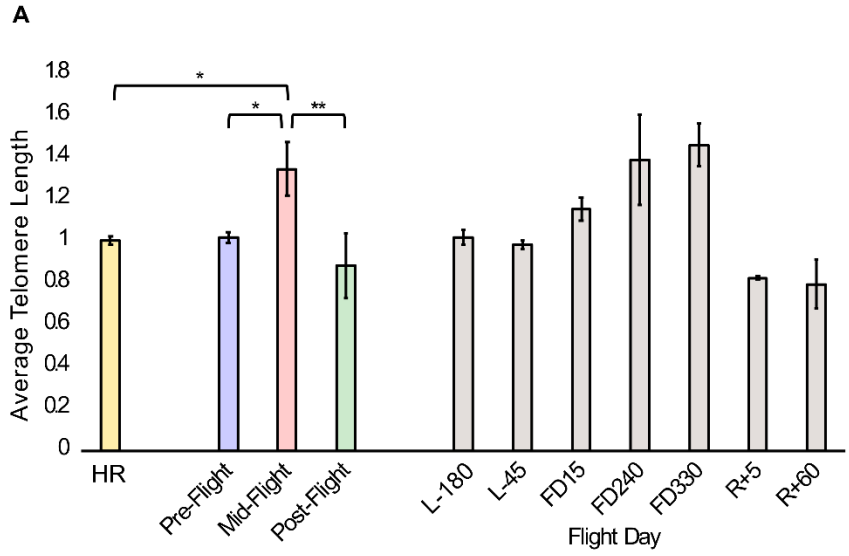

Figure 2.5. Telomere length in urine during long-duration spaceflight. Relative average telomere length assessed by qRT-PCR in DNA extracted from urine collected before, during and after the One Year Mission. **A.** Combined pre-, mid-, and post-flight means, and individual time-course for TW (space twin), normalized to the ground-based control (HR; all values averaged), are shown. L-, FD, and R+ refer to days pre-, mid-, and post-spaceflight, respectively. Significance tested by one-way ANOVA and post-hoc Student's t-tests for multiple comparisons. Error bars are SEM. $p<0.05$*, $p<0.01$**, $p<0.001$***.

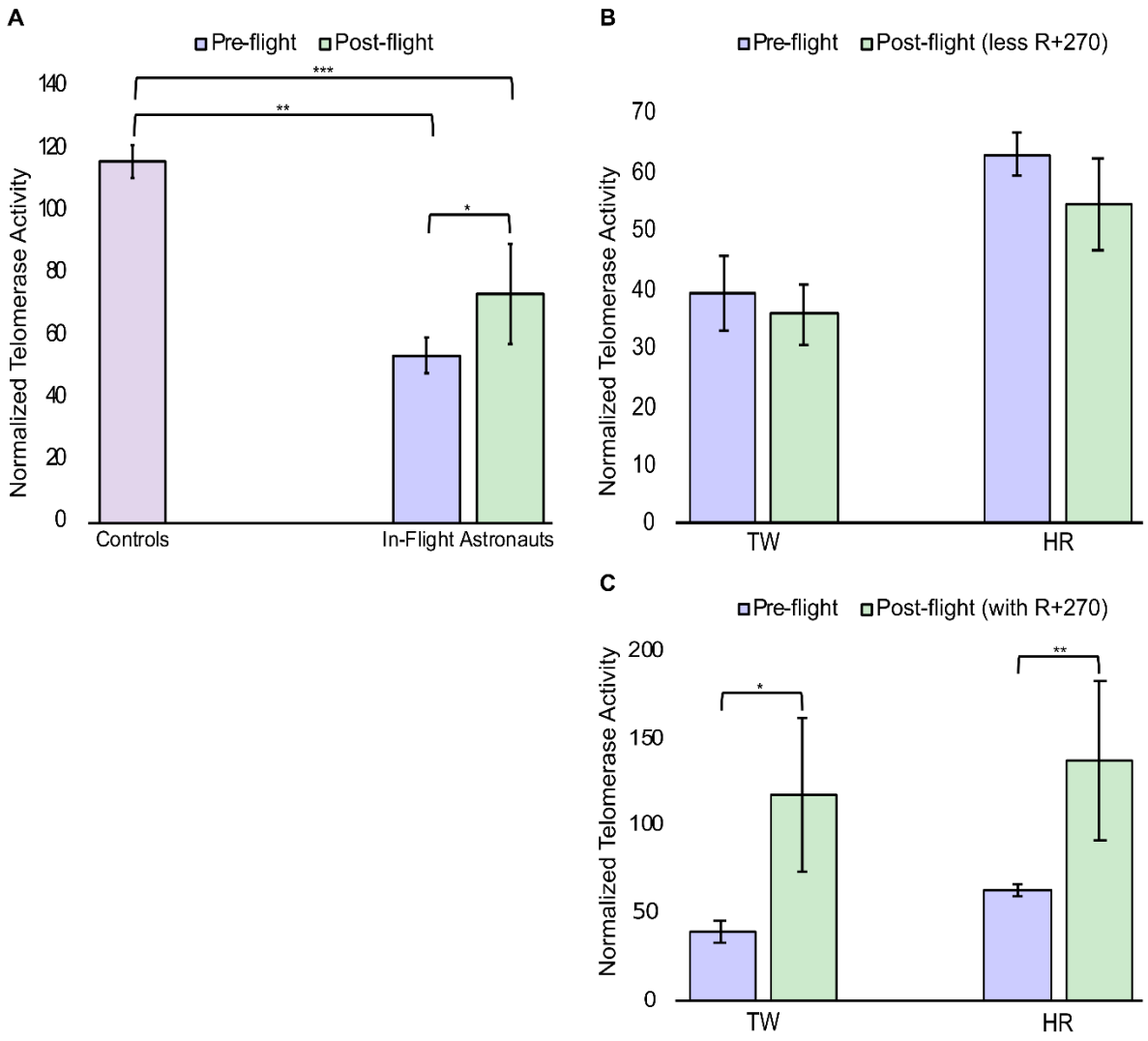

Figure 2.6. Telomerase activity before and after long-duration spaceflight. Telomerase activity was assessed by qPCR-TRAP in PBMC protein lysates. **A.** One-year and six-month mission astronauts (n=3) and age-sex matched ground control cohort (n=11). **B.** TW (space twin) and HR (ground twin) before and after the One Year Mission, minus the latest timepoint, Return +270 days (R+270), and **C.** including R+270 timepoint. Significance was tested by one-way ANOVA and post-hoc Student's t-tests adjusted for multiple comparisons. Error bars are SEM. p $p<0.05$*, $p<0.01$**, $p<0.001$***.

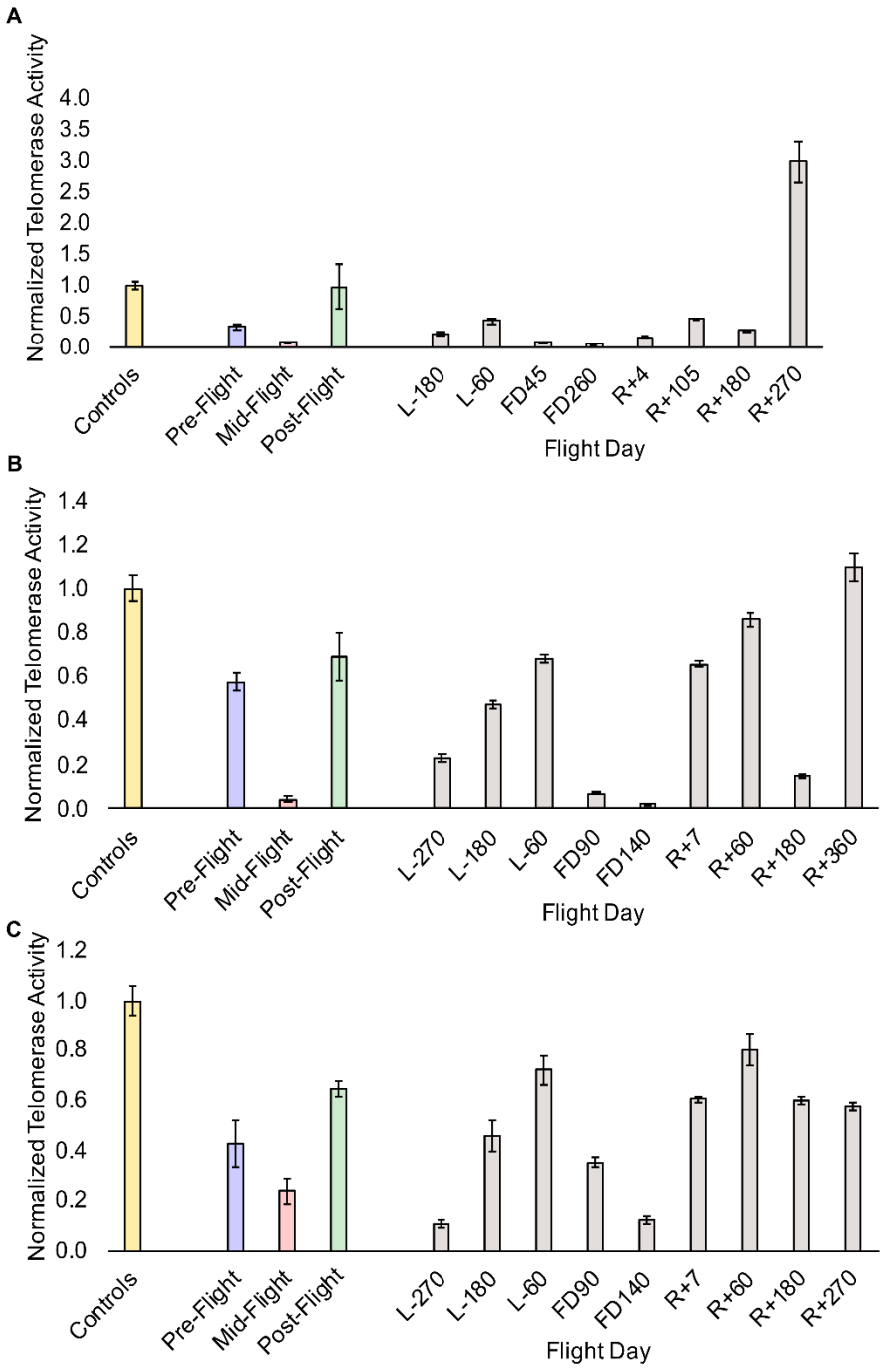

Figure 2.7. Telomerase activity time-course for one-year and six-month mission astronauts. Telomerase activity assessed by qRT-PCR TRAP (in PBMC protein lysates), normalized to internal HeLa controls for **A**. crewmember A, **B**. crewmember B, and **C**. crewmember C. For each crewmember, combined pre-, mid-, and post-flight telomerase activity, and their individual time-course (gray bars), normalized to age and sex matched ground control cohort (yellow bar; n=11), are shown. L-, FD, and R+ refer to days pre-, mid-, and post-spaceflight, respectively.

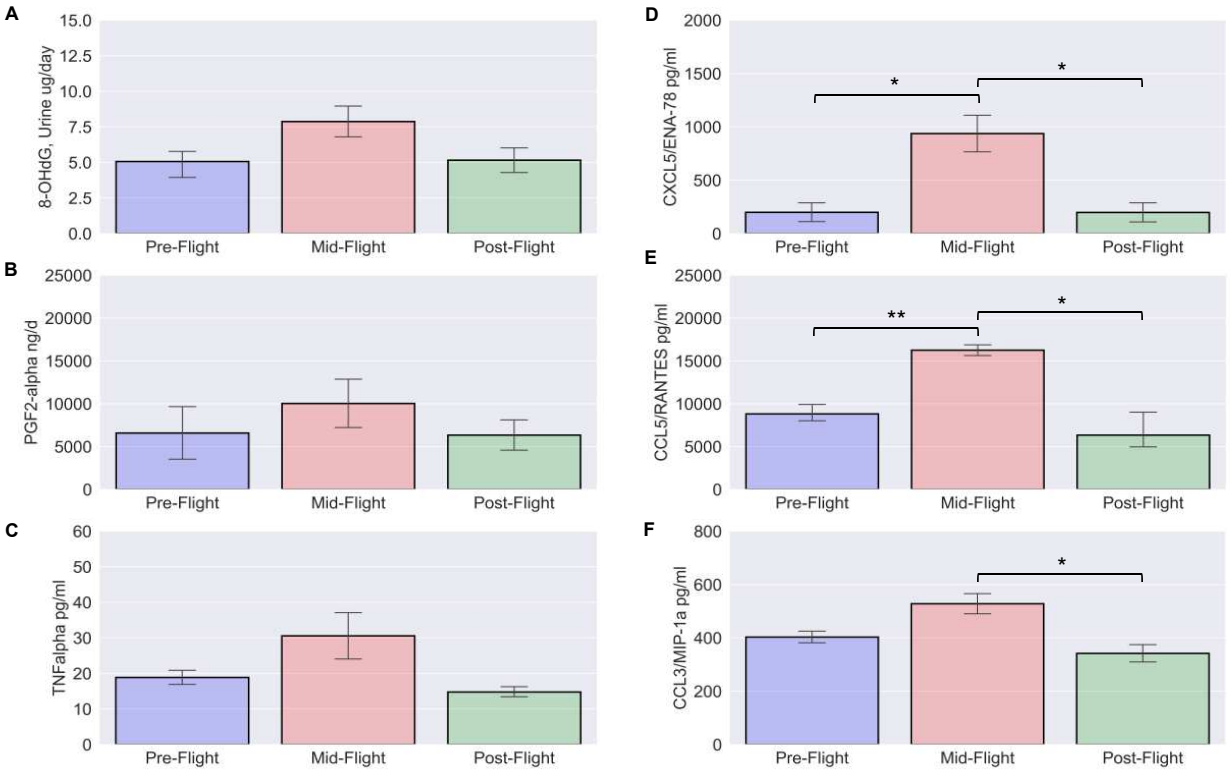

Figure 2.8. Oxidative stress and inflammation elevated during long-duration spaceflight. Urinary biomarkers of oxidative stress before, during, and after one-year and six-month missions (n=3); **A.** 8-OHdG (8-hydroxy-2'-deoxyguanosine), and **B.** PGF2-alpha (prostaglandin F-2 alpha). Plasma concentrations of inflammation biomarkers were significantly elevated during spaceflight; **C.** cytokine TNF-alpha (Tumor Necrosis Factor alpha), and **D-F.** chemokines (CXCLS, CCL5, and CCL3). Significance tested by one-way ANOVA and post-hoc t-tests adjusted with a Bonferroni correction. Error bars are SEM. $p<0.05$*, $p<0.01$**, $p<0.001$***.

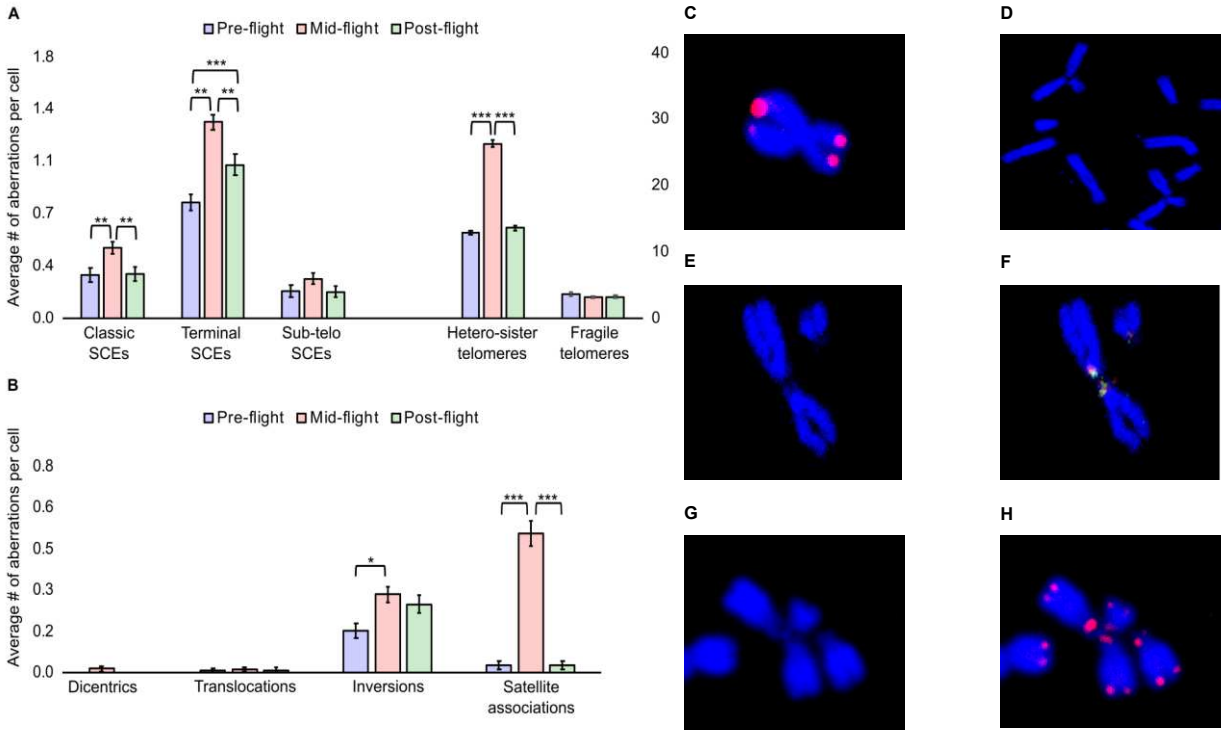

Figure 2.9. Chromosomal and telomeric recombination and rearrangement events before, during, and after long-duration spaceflight. Cytogenetic analyses (metaphase chromosomes; T-cells) for astronauts on one-year or six-month missions (n=3). A. Left. Sister Chromatid Exchange (SCE) recombination events detected by dGH and chromosome-specific subtelomere probes (subtelo-dGH): classic SCEs along the length of a chromosome; terminal or subtelo-SCEs, complete or partial exchange of the subtelomere between sister chromatids, respectively. Right. Telomere recombination events detected by Telo-FISH: hetero-sister telomeres, telomeres on sister chromatids of heterogeneous size and/or intensity; fragile telomeres, individual telomeres with stringy, diffuse, and/or multiple discrete signals. B. Ionizing radiation-induced chromosome rearrangements (dicentrics, translocations, inversions) and satellite associations (involving short p-arms). Thirty metaphases per individual per timepoint scored. Overall means shown. Significance tested by one-way ANOVA per aberration and post-hoc Student's t-tests for multiple comparisons. Error bars are SEM. $p<0.05$*, $p<0.01$**, $p<0.001$***. C-H. Illustrative examples of spaceflight-specific anomalies (arrows). C. Hetero-sister telomeres; DNA (DAPI, blue), telomeres (Cy3, red). D-F. Satellite associations with subtelo-dGH; DNA (DAPI, blue), dGH chromosome (Cy3, red) and subtelomere (Cy5, green) probes. D, E. Satellite associations without and F. with Cy3/Cy5 channels. G. Satellite associations with Telo-FISH, without and H. with Cy3 (red) channel.

STAR Methods

RESOURCE AVAILABILITY

Lead Contact

Further information on methods and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Dr. Susan M. Bailey (susan.bailey@colostate.edu).

Materials Availability

This study did not generate new unique reagents.

Data and Code Availability

The NASA Life Sciences Data Archive (LSDA) is the repository for all human and animal research data, including that associated with this study. LSDA has a public facing portal where data requests can be initiated (https://lsda.jsc.nasa.gov/Request/dataRequestFAQ). The LSDA team provides the appropriate processes, tools, and secure infrastructure for archival of experimental data and dissemination while complying with applicable rules, regulations, policies, and procedures governing the management and archival of sensitive data and information. The LSDA team enables data and information dissemination to the public or to authorized personnel either by providing public access to information or via an approved request process for information and data from the LSDA in accordance with NASA Human Research Program and Johnson Space Center (JSC) Institutional Review Board direction. See Perspective piece (Afshinneko E., et al *Cell, 2020*) for additional information.

The code necessary to recreate all data processing, statistical analyses and figures will be made available at: https://github.com/Jared-Luxton/NASA-astronauts-telomeres-chromosomes.

The code necessary to recreate the nanopore measurements of telomere length is available at: https://github.com/LankyCyril/edgecase.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human subjects

Three International Space Station (ISS) astronauts participated in this study; one one-year mission crewmember, and two six-month mission crewmembers for whom in-flight samples were available. The ground control cohort consisted of eleven healthy age and sex-matched NASA volunteers (Luxton et al, *Cell 2020*). Institutional review board approval was obtained from the Committee for the Protection of Human Subjects at JSC, Houston, TX, USA. Informed consent was obtained from all subjects who participated in the study. In order to protect personally identifiable information for this small and select group, sex and age are not reported.

Collection of biological samples on/off the International Space Station

Samples were collected and processed as pioneered and previously described for the NASA Twins Study (Garrett-Bakelman et al., 2019). Briefly, peripheral blood was drawn from crewmembers (and ground control subjects), before (~270, 180, and 60 days prior to Launch), during (Flight Day ~45 or 90, and 140 or 260), and after (~1-7, 60, 180, and 270 days after Return) spaceflight, either on Earth or on the ISS. Samples were kept under ambient (room temperature) conditions for ~24hr during shipment to research laboratory for processing, or for

in-flight samples, ~48hr during return from space via Soyuz vehicle to Russia, and then jet aircraft to JSC for processing on-site. Urine samples were collected before (180 and 45 days prior to Launch), during (Flight Day 15, 240, and 330), and after spaceflight, (~1-7, and 60 days after Return), either on Earth or on the ISS, and promptly frozen at -80ºC. Upon arrival at JSC, samples were thawed, pooled, and aliquoted, frozen and shipped to research laboratory for processing. For biochemical profiles, crewmember blood and urine samples were collected as previously described (Smith et al., 2012; Zwart et al., 2011). Briefly, samples were collected before (~80 and 45 days prior to Launch), during (Flight Day 15, 30, 60, 120, and 180), and after spaceflight, in the first 24 hr after landing (Return+0), and again 30 days later (Return+30).

Cell line culture

LY-S cells were cultured to confluency in T-75 tissue culture flasks at 1:9 in Gibco PB-Max Karyotyping Medium supplemented with phytohaemagglutinin A (PHA) (ThermoFisher #12557021) and incubated at 37ºC. HeLa cells (ATTC CCL-2) were cultured similarly, except in alpha-mem medium supplemented with 10% fetal bovine serum. For harvesting, cells were split after reaching confluency and KaryoMax Colcemid (ThermoFisher #15210040) added [0.1 µg/mL final concentration], then incubated for 4 hours. Metaphase chromosome spreads were prepared using standard cytogenetic techniques [e.g., (Robinson et al., 2019)]. Briefly, cells were treated with 75mM hypotonic KCl at 37ºC for 20 minutes, centrifuged (1000rpm; ~270xg) and washed with 5ml 3:1 methanol-acetic acid fixative four times, then a final wash and dropping on slides or storage at -20ºC.

METHOD DETAILS

Collection and Processing of Blood for qPCR assessment of Telomere Length and Telomerase Activity

Peripheral blood was drawn from crewmembers (and ground control subjects), before (~270, 180, and 60 days prior to Launch), during (Flight Day ~45 or 90, and 140 or 260), and after, (~1-7, 60, 180, and 270 days after Return) spaceflight, either on Earth or on the ISS, and collected into one 10 mL ethylenediaminetetraacetic acid (K2EDTA) tube (Becton, Dickinson, and Co #368589), which was kept under ambient (room temperature) conditions for ~24hr during shipment to research laboratory for processing, or for in-flight samples, ~48hr during return from space to JSC for on-site processing. Upon arrival, cells were transferred into CPT mononuclear cell preparation tubes (Becton, Dickinson, and Co # 362761) and centrifuged at 1700 x g for 20 min per manufacturer's recommendations. Separated peripheral blood mononuclear cells (PBMCs) were collected and washed in PBS and any remaining red blood cell contamination was removed using ACK lysis buffer. PBMCs (1 x 10^6 cells/mL) were divided for either qRT-PCR based-telomere length measurement or qRT-PCR TRAP assessment of telomerase activity. DNA extraction for qPCR telomere length measurements was performed using the DNeasy Blood and Tissue Kit (Qiagen #69504). PBMCs were incubated in proteinase K for 3 hr at 37°C. DNA was isolated from PBMC populations, and from separated CD4+ and CD8+ T-cells, CD19+ B-cells, and lymphocyte depleted (LD) populations. For telomerase activity, PBMC protein lysates were prepared in mammalian protein extraction buffer (M-PER) (ThermoFisher #78503) containing cOmplete Protease Inhibitor Cocktail (Millipore Sigma #11836170001) and RNasin Ribonuclease Inhibitors (Promega #N2515) at 1 x 10^6 cells per 100 µL. Lysates were cleared by incubating on ice for 10 min, followed by centrifugation at 18,000 rpm for 15 min at

40°C, aliquoted, and stored at -80°C. Protein concentrations were determined using the Bio-Rad Protein Assay Kit (Bradford Assay; Bio-Rad # 5000001).

Telomere Length assessment by Multiplexed Quantitative (q)PCR

Multiplexed qPCR measurements of telomere length were carried out as previously described (Cawthon, 2009). Here, a 22 μL master mix was prepared using SYBR green GoTaq qPCR master mix (Promega #A6001) combined with the telomere forward primer (TelG; 5'-ACACTAAGGTTTGGGTTTGGGTTTGGGTTTGGGTTAGTGT-3'), telomere reverse primer (TelC; 5'-TGTTAGGTATCCCTATCCCTATCCCTATCCCTATCCCT AACA-3'), albumin forward primer (AlbU; 5'-CGGCGGCGGGCGGCGCGGGCTGGGCGGA AATGCTGCACAGAATCCTTG-3'), albumin reverse primer (AlbD; 5'-GCCCGGCCCGCCG CGCCCGTCCCGCCGGAAAAGCATGGTCGCCTGTT-3') at 10 μM per primer (Integrated DNA Technologies), and RNase/DNase free water. To the master mix, 3 μL of DNA at 3.33 ng/uL was added for a final volume of 25 μL. The TelG/C primers were at a final concentration of 900 nM and the AlbU/D primers at 400 nM. Telomere length measurements were carried out using a Bio-Rad CFX-96 qPCR machine. The cycle design was as follows: 95°C for 3 min; 94°C for 15 sec, 49°C for 15 sec, for 2 cycles; 94°C for 15 sec, 62°C for 10 sec, 74°C for 15 sec, 84°C for 10 sec, and 88°C for 15 sec, for 32 cycles. The melting curve was established by a 72°C to 95°C ramp at 0.5°C/sec increase with a 30 sec hold. Multiplexing both telomere and albumin primers using a single fluorescent DNA-intercalating dye is possible because the telomere sequences are amplified at a lower quantification cycle (Cq) than the albumin

sequences. Standard curves were prepared using human genomic DNA (Promega cat # G3041) with 3-fold dilutions ranging from 50 ng to 0.617 ng in 3 µL per dilution. Negative controls included a no-template TelG/C only and AlbU/D only, and a combined TelG/C and AlbU/D control. Samples were normalized across plates using a human genomic DNA standard. Each sample was run in triplicate on a 96-well plate format and relative telomere length was established using a telomere (T) to albumin (A) ratio.

Telomerase Activity assessment by qPCR Telomere Repeat Amplification Protocol (TRAP)

Telomerase activity was monitored using qRT-PCR TRAP as previously described (Hou et al., 2001). Here, protein lysates were suspended at 0.20 µg in 25 µL of SYBR green GoTaq qPCR Master Mix (Promega #A6001) containing 0.20 µg T4 gene 32 protein (New England Biolabs #M0300S), 0.10 µg of primers TS (5'-AATCCGTCGAGCAG AGTT- 3') and ACX (5'-GCGCGG(CTTACC)3CTAACC-3') (Integrated DNA Technologies) and RNase/DNase free water. The qRT PCR reactions took place using a Bio-Rad CFX-96 qPCR machine using the following steps: 25ºC for 20 min; 95ºC for 3 min; 95ºC for 20 sec, 50ºC for 30 sec, 72ºC for 90 sec for 50 cycles; establish a melting curve and prevent primer dimerization for 10 sec per cycle for 80 cycles. Each sample was run in triplicate on a 96-well plate format allowing for an average Cq to be obtained per sample. Relative telomerase activity was then established using the delta delta Ct method (Livak and Schmittgen, 2001). In addition to the test samples, the following controls were included: no template controls with TS only, ACX only, and TS/ACX primers combined, as well as a HeLa cell line (ATTC CCL-2) heat-inactivated negative control, and a HeLa positive control (very high telomerase activity).

Collection and Processing of Urine for qPCR assessment of Telomere Length

Urine samples were collected before (180, and 45 days prior to Launch), during (Flight Day 15, 240, and 330), and after (~1-7, and 60 days after Return) spaceflight, either on Earth or on the ISS, and promptly frozen at -80C as previously described for the NASA Twins Study (Garrett-Bakelman et al., 2019). Upon arrival at JSC, samples were thawed, pooled, and aliquoted (300 uL), frozen and shipped to research laboratory for processing. To remove cell debris, urine samples thawed on ice were centrifuged at 2000 x g for 5 min at 4°C by using Eppendorf centrifuge 5417R. Total DNA concentration in the urine was measured by a NanoDrop 2000 Spectrophotometer (ThermoFisher). Telomere length was quantified using a described qPCR assay developed to measure relative telomere length (O'Callaghan and Fenech, 2011), which we adjusted to measure relative telomere length in the urine. Briefly, we measured relative telomere length by determining the ratio of telomere repeat copy number to single control gene copy number. For a single control gene, we used β-globin. Forward and reverse telomeric primers were (Tel1b; 5'-CGGTTTGTTTGGGTTTGGGTTTGGGTTTGGGTTTGGGTT-3'), and (Tel2b; 5'- GGCTTGCCTTACCCTTACCCTTACCCTTACCCTTACCCT-3'), respectively, and forward and reverse primers for the β-globin gene were (HBG1; 5'-GCTTCTGACACAACTGTGTTCACTAGC-3') and (HBG2; 5'-CACCAACTTCATCCACGTTCACC-3'), respectively (Integrated DNA Technologies). A 20 µL master mix was prepared using AzuraQuant Green Fast qPCR Mix LoRox (Azura Genomics, # AZ-2105) combined with the telomere forward and reverse primers or with β-globin forward and reverse primers at 10 µM per primer (Integrated DNA Technologies), and RNase/DNase free water. To the Master Mix, 5 uL of diluted urine DNA in DNase/RNase free water at 200 ng/uL

was added for a final volume of 20 µL. Approximately 20 times dilution of urine in the final volume of master mix reduced the effect of urea and other potential inhibitors on PCR. Each sample was run in duplicate on a 96-well plate format. PCRs were performed in a QuantStudio 3 machine (ThermoFisher). PCR conditions for both reactions were 50°C for 2 min; 95°C for 10 min followed by 40 cycles of 95°C for 15 sec, 60°C for 1 min. A standard curve is established by dilution of known quantities of a synthesized 84 mer oligonucleotide containing only TTAGGG repeats (Integrated DNA Technologies) with 10-fold dilutions starting from 60 pg to 1 pg.

Collection and Processing of Blood for Nanopore sequencing assessment of Telomere Length

Peripheral blood was drawn, collected, and processed as previously described for the NASA Twins Study (Garrett-Bakelman et al., 2019). Briefly, samples were collected before (~160 days prior to Launch), during (Flight Day ~260), and after (~180 days after Return) spaceflight, either on Earth or on the ISS, and kept under ambient (room temperature) conditions for ~24hr or 48hr prior to processing. Fresh processing of CPT (BD Biosciences Cat # 362760) tubes involved cell separation by centrifugation (1800 X g for 20 min), positive selection of CD4+ cells using magnetic beads per manufacturer's recommendations (Miltenyi Biotec Cat # 130-097-048), preparation of CD4 T-cell lysates in RLT+ solution, and DNA extraction using QIAGEN's Allprep kit (Cat # 80204). Quantitation of DNA was assessed using Qubit 2.0 (Invitrogen) per manufacturer's recommendations. DNA quality was assessed by visualization on 0.7% agarose gels using standard techniques. Extracted DNA was library prepped for Oxford Nanopore Promethion sequencing using an LSK-108 ligation prep and PRO-002 flowcell. Runs were performed for 48 hours and basecalled using guppy to generate fastq files.

Adapting the approach from (Grigorev, Foox et al., *Cell* 2020), nanopore reads from each experiment were aligned with minimap2 (Li, 2018) to the hg38 reference that was extended and annotated with subtelomeric sequences (Stong et al., 2014). Reads overhanging the end of the subtelomeric region (extending into the telomeric tract) were selected and the segments of these reads fully contained in the telomeric tract were measured.

Collection and Processing of Blood for Telomere Fluorescence *in situ* Hybridization (Telo-FISH) and directional Genomic Hybridization (dGH)

Peripheral blood was drawn from crewmembers (and ground control subjects), before (270, 180, and 60 days prior to Launch), during (Flight Day 45 or 90, and 140 or 260), and after (~1-7, 60, 180, and 270 days after Return) spaceflight, either on Earth or on the ISS, and collected into one 10 mL sodium heparin tube (Becton, Dickinson, and Co #367874), which was kept under ambient (room temperature) conditions for ~24hr during shipment to research laboratory, or for in-flight samples, ~48hr during return from space via Soyuz vehicle and jet aircraft to JSC for on-site processing. Upon arrival, heparinized blood was split into two T-25 tissue culture flasks at 1:9 in Gibco PB-Max Karyotyping Medium supplemented with phytohaemagglutinin A (PHA) (ThermoFisher #12557021) and cultured at 37ºC. For directional Genomic Hybridization (dGH), 5.0 mM 5-bromo-deoxyuridine (BrdU) and 1.0 mM 5- bromo-deoxycytidine (BrdC) were added to the medium. Following 44 hr incubation/stimulation, KaryoMax Colcemid (ThermoFisher #15210040) was added [0.1 μg/mL final concentration] for the final 4 hr prior to harvest (at 48 hr). Metaphase chromosome spreads were prepared using standard cytogenetic techniques [e.g., (Robinson et al., 2019)]. Briefly, cells were treated with 75mM hypotonic KCl

at 37ºC for 20 minutes, centrifuged (1000rpm; ~270xg) and washed with 5ml 3:1 methanol-acetic acid fixative four times, then a final wash and dropping on slides or storage at -20ºC.

Telomere Length assessment by Telo-FISH

Prepared slides with metaphase chromosome spreads (PHA stimulated T-cells) were hybridized with a fluorescently-labeled telomere probe as previously described (Sishc et al., 2015). Briefly, slides were washed in PBS for 5 min, dehydrated through a cold ethanol series (75%, 85% and 100%) for 2 min each, and air dried. Slides were then denatured in 70% formamide/2x sodium chloride and sodium citrate (2xSSC) at 75°C for 2 min, followed by a second cold ethanol series. A hybridization mixture containing a G-rich (TTAGGG3) peptide nucleic acid (PNA) telomere probe directly labeled with Cyanine-3 (Cy3; Biosynthesis) was prepared by diluting 0.5 µL probe [5nM final concentration] in 36 µL of formamide, 12 µL of 0.5 M pH 7.5 Tris-HCl, 2.5 µL of 0.1 M KCl, and 0.6 µl of 0.1 M MgCl2; mixture was incubated at 75°C for 5 min, then 50 µL applied to each slide. Coverslips were sealed with rubber cement and slides incubated at 37 °C for 24 hr. After hybridization, slides were washed in a series of 43.5°C washes for 3 min each: washes one and two: 50% formamide in 2x sodium chloride and sodium citrate (2xSSC); washes three and four: 2x SSC; and washes five and six: 2x SSC plus 0.1% Nonidet P-40. Slides were counterstained with 50 µL DAPI in Prolong Gold Antifade (ThermoFisher #P36931), coverslipped, and stored at 4°C for 24 hr before imaging. Metaphase spreads (30 per sample) were imaged on a Nikon Eclipse Ni-U epifluorescent microscope equipped with an Andor Zyla 5.5 sCMOS camera and SpectraX LED light source. Stacked images (5 images/stack, 0.3 µm step size) were taken of each metaphase spread. Fluorescence intensity was quantified using the ImageJ (http://rsb.info.nih.gov/ij/) plugin, Telometer

(http://demarzolab.pathology.jhmi.edu/telometer/index.html), and normalized using the mouse lymphoma control cell line LY-S (Wong and Slijepcevic, 2004) and age-sex matched controls.

Cytogenetic analyses by directional Genomic Hybridization (dGH)

High resolution detection of inversions and translocations on metaphase chromosome spreads (PHA stimulated T-cells) was performed utilizing Directional Genomic Hybridization (dGH) single color (Cy-3) whole chromosome and single color (Cy5) subtelomere 1-3 paints (KromaTiD Inc., Ft. Collins, CO) as previously described (Ray et al., 2014; Ray et al., 2013). Briefly, prepared slides with metaphase chromosomes substituted with bromonucleotides (BrdU/BrdC) were submersed in Hoechst 33258 (Millipore Sigma #B1155) for 15 min, selectively photolysed using a SpectroLinker UV Crosslinker equipped with 365 nm UV bulbs for 35 min, and then the nicked DNA was exonucleolytically degraded with Exonuclease III (New England Biolabs #M0206L) for 30 min. Hybridization cocktail was applied to the slides, which were coverslipped and sealed using rubber cement, then denatured at 68°C for 3 min. Slides were hybridized overnight at 37°C followed by 5 washes in 2x SSC at 43°C prior to imaging. Metaphase spreads were imaged on a Nikon Eclipse Ni-U epifluorescent microscope equipped with an Andor Zyla 5.5 sCMOS camera and SpectraX LED light source. 30 metaphase spreads were scored per timepoint per individual.

Collection and Processing of Blood and Urine for Biochemical analyses

Blood and urine samples were collected as previously described (Smith et al., 2012; Zwart et al., 2011). Briefly, samples were collected before (80 and 45 days prior to Launch), during (Flight Day 15, 30, 60, 120, and 180), and after spaceflight, in the first 24 hr after landing (Return+0)

and again 30 days later (Return+30). The R+0 samples were not necessarily fasting, given the time of day and nature of return from flight. Pre- and post-flight collections included two consecutive 24 hr urine collections, and inflight collections included one 24 hr urine collection. PGF2-α was measured using a commercially available ELISA kit (Oxford Biomedical Research, Rochester Hills, MI) and 8-OHdG was measured by LC-MS/MS (Wu and Ho, 2009). Plasma concentrations were evaluated simultaneously in duplicate using a multiplex bead immunoassay (R&D Systems). Samples were processed according to the manufacturer's instructions. Plasma cytokines were measured as previously described (Crucian et al., 2014). Briefly, 50 μL of plasma was incubated with 22 types of beads, each bound to unique cytokine capture antibodies. Post-incubation the plasma/bead mixture was washed and incubated with fluorescent secondary antibodies to each of the 22 primary antibodies. Assays were performed in a 96-well plate and analyzed by a Luminex 100 instrument (Luminex, Inc.).

QUANTIFICATION AND STATISTICAL ANALYSIS

Power calculations were performed to predetermine minimum sample size for our endpoints. Although 10 crewmembers were desired, practical limitations associated with collection of samples during spaceflight resulted in 3 crewmembers with in-flight samples. Even so, because of the large number of telomeres and chromosome aberrations scored, determination of statistical significance was possible. Data are generally displayed as mean with SEM except where noted (Nanopore sequencing, Figure 1.3A). Sample numbers (n) are provided with each figure.

Mean Telomere Length and Telomerase Activity by qPCR

All samples at each timepoint were run in technical triplicate. Statistical significance was evaluated using a one-way ANOVA with post-hoc Tukey's HSD in Python using statsmodels library.

Relative Telomere Length by Nanopore sequencing

Given the differences in the lengths of reads in the experiments, the lengths of telomeric segments were normalized by the median length of the sequencing pool in each experiment. Distributions of the normalized lengths were compared for each pair of experiments using the Mann-Whitney U one-tailed test, and the P-values were adjusted with the Bonferroni correction. The normalized lengths of telomeric segments were significantly higher in the in-flight TW sample than in all other samples; other pairwise comparisons were not significant (not shown). Annotated within each distribution: median (long dashed line), upper and lower quartile (short dashed line).

Telomere Length by Telo-FISH and generation of individual telomere length distributions

Fluorescence intensity quantifications of individual telomeres, output by ImageJ, were stored in Excel files, with one file per astronaut per timepoint. Custom scripts were written in Python with extensive use of the pandas library to extract and parse the individual telomere length measurements into one dataframe object for data visualizations and statistical analyses. For timepoint samples with individual telomere distributions less than the theoretical number of telomeres, given 30 metaphase spreads (5520), individual telomere counts were imputed up to 5520 using random sampling (numpy library) from the observed distribution of telomeres for

that subject's timepoint, facilitating visual and statistical comparisons for those distributions. For each astronaut, histograms of individual telomeres were colored by taking the distribution of individual telomeres at the earliest timepoint pre-spaceflight and splitting the data into quartiles, with the bottom 25% of telomeres (shortest) in yellow, the middle 50% in blue, and the top 25% in red (longest). The quartile cutoff values, unique to each astronaut, were then applied to that astronaut's subsequent timepoints, to visualize shifts in individual telomeres into and out of the colored quartiles.

Analysis of Telomere Length distributions

Statistically significant differences between the distributions of individual telomeres were assessed per timepoint for each astronaut using the Mann-Whitney U test in Python using the statsmodels library.

Mean Telomere Length by Telo-FISH

Thirty metaphases (184 telomeres/metaphase) were scored per timepoint per individual and statistics were performed on the mean telomere length for a given timepoint. Statistical significance was evaluated using a one-way ANOVA with post-hoc Tukey's HSD in Python using the statsmodels library.

Chromosome Aberration analyses by dGH

Thirty metaphases were scored per timepoint per individual and statistics were performed on the average number of aberrations per cell for a given timepoint. Statistical significance was

evaluated using a one-way ANOVA with post-hoc Tukey's HSD in Python using the statsmodels library.

Biochemical/analyte analyses

Statistical significance was evaluated using a one-way ANOVA with post-hoc Tukey's HSD in Python using the statsmodels library.

REFERENCES

Aubert, G., and Lansdorp, P.M. (2008). Telomeres and aging. Physiol Rev *88*, 557-579.
Azzam, E.I., Jay-Gerin, J.P., and Pain, D. (2012). Ionizing radiation-induced metabolic oxidative stress and prolonged cell injury. Cancer Lett *327*, 48-60.

Bailey, S.M., Brenneman, M.A., and Goodwin, E.H. (2004). Frequent recombination in telomeric DNA may extend the proliferative life of telomerase-negative cells. Nucleic Acids Res *32*, 3743-3751.

Behjati, S., Gundem, G., Wedge, D.C., Roberts, N.D., Tarpey, P.S., Cooke, S.L., Van Loo, P., Alexandrov, L.B., Ramakrishna, M., Davies, H.*, et al.* (2016). Mutational signatures of ionizing radiation in second malignancies. Nat Commun *7*, 12605.

Berardinelli, F., Antoccia, A., Buonsante, R., Gerardi, S., Cherubini, R., De Nadal, V., Tanzarella, C., and Sgura, A. (2013). The role of telomere length modulation in delayed chromosome instability induced by ionizing radiation in human primary fibroblasts. Environ Mol Mutagen *54*, 172-179.

Berardinelli, F., Antoccia, A., Cherubini, R., De Nadal, V., Gerardi, S., Cirrone, G.A., Tanzarella, C., and Sgura, A. (2010). Transient activation of the ALT pathway in human primary fibroblasts exposed to high-LET radiation. Radiat Res *174*, 539-549.

Blackburn, E.H., and Epel, E.S. (2012). Telomeres and adversity: Too toxic to ignore. Nature *490*, 169-171.

Bonassi, S., Znaor, A., Norppa, H., and Hagmar, L. (2004). Chromosomal aberrations and risk of cancer in humans: an epidemiologic perspective. Cytogenet Genome Res *104*, 376-382.

Bryan, T.M., Englezou, A., Dalla-Pozza, L., Dunham, M.A., and Reddel, R.R. (1997). Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. Nat Med *3*, 1271-1274.

Caradonna, F. (2015). Nucleoplasmic bridges and acrocentric chromosome associations as early markers of exposure to low levels of ionising radiation in occupationally exposed hospital workers. Mutagenesis *30*, 269-275.

Cawthon, R.M. (2009). Telomere length measurement by a novel monochrome multiplex quantitative PCR method. Nucleic Acids Res *37*, e21.

Cherkas, L.F., Hunkin, J.L., Kato, B.S., Richards, J.B., Gardner, J.P., Surdulescu, G.L., Kimura, M., Lu, X., Spector, T.D., and Aviv, A. (2008). The association between physical activity in leisure time and leukocyte telomere length. Arch Intern Med *168*, 154-158.

Cohen, S., Janicki-Deverts, D., Turner, R.B., Casselbrant, M.L., Li-Korotky, H.S., Epel, E.S., and Doyle, W.J. (2013). Association between telomere length and experimentally induced upper respiratory viral infection in healthy adults. Jama *309*, 699-705.

Coluzzi, E., Buonsante, R., Leone, S., Asmar, A.J., Miller, K.L., Cimini, D., and Sgura, A. (2017). Transient ALT activation protects human primary cells from chromosome instability induced by low chronic oxidative stress. Sci Rep *7*, 43309.

Coluzzi, E., Leone, S., and Sgura, A. (2019). Oxidative Stress Induces Telomere Dysfunction and Senescence by Replication Fork Arrest. Cells *8*.

Crucian, B.E., Zwart, S.R., Mehta, S., Uchakin, P., Quiriarte, H.D., Pierson, D., Sams, C.F., and Smith, S.M. (2014). Plasma cytokine concentrations indicate that in vivo hormonal regulation of immunity is altered during long-duration spaceflight. J Interferon Cytokine Res *34*, 778-786.

Cucinotta, F.A. (2014). Space radiation risks for astronauts on multiple International Space Station missions. PLoS One *9*, e96099.

Cucinotta, F.A., Hamada, N., and Little, M.P. (2016). No evidence for an increase in circulatory disease mortality in astronauts following space radiation exposure. Life Sciences in Space Research *10*, 53-56.

de Lange, T. (2009). How telomeres solve the end-protection problem. Science *326*, 948-952.
De Vitis, M., Berardinelli, F., Coluzzi, E., Marinaccio, J., O'Sullivan, R.J., and Sgura, A. (2019). X-rays Activate Telomeric Homologous Recombination Mediated Repair in Primary Cells. Cells *8*.

Delp, M.D., Charvat, J.M., Limoli, C.L., Globus, R.K., and Ghosh, P. (2016). Apollo Lunar Astronauts Show Higher Cardiovascular Disease Mortality: Possible Deep Space Radiation Effects on the Vascular Endothelium. Sci Rep *6*, 29901.

Deng, W., Cheung, S.T., Tsao, S.W., Wang, X.M., and Tiwari, A.F. (2016). Telomerase activity and its association with psychological stress, mental disorders, lifestyle factors and interventions: A systematic review. Psychoneuroendocrinology *64*, 150-163.

Dugan, L.L., You, Y.H., Ali, S.S., Diamond-Stanic, M., Miyamoto, S., DeCleves, A.E., Andreyev, A., Quach, T., Ly, S., Shekhtman, G.*, et al.* (2013). AMPK dysregulation promotes diabetes-related reduction of superoxide and mitochondrial function. J Clin Invest *123*, 4888-4899.

Dunham, M.A., Neumann, A.A., Fasching, C.L., and Reddel, R.R. (2000). Telomere maintenance by recombination in human cells. Nat Genet *26*, 447-450.

Epel, E.S., Blackburn, E.H., Lin, J., Dhabhar, F.S., Adler, N.E., Morrow, J.D., and Cawthon, R.M. (2004). Accelerated telomere shortening in response to life stress. Proceedings of the National Academy of Sciences of the United States of America *101*, 17312-17315.

Fouquerel, E., Lormand, J., Bose, A., Lee, H.T., Kim, G.S., Li, J., Sobol, R.W., Freudenthal, B.D., Myong, S., and Opresko, P.L. (2016). Oxidative guanine base damage regulates human telomerase activity. Nat Struct Mol Biol *23*, 1092-1100.

Gardner, M., Bann, D., Wiley, L., Cooper, R., Hardy, R., Nitsch, D., Martin-Ruiz, C., Shiels, P., Sayer, A.A., Barbieri, M.*, et al.* (2013). Gender and telomere length: Systematic review and meta-analysis. Exp Gerontol.

Garrett-Bakelman, F.E., Darshi, M., Green, S.J., Gur, R.C., Lin, L., Macias, B.R., McKenna, M.J., Meydan, C., Mishra, T., Nasrini, J.*, et al.* (2019). The NASA Twins Study: A multidimensional analysis of a year-long human spaceflight. Science *364*.

George, K., Rhone, J., Beitman, A., and Cucinotta, F.A. (2013). Cytogenetic damage in the blood lymphocytes of astronauts: effects of repeat long-duration space missions. Mutat Res *756*, 165-169.

Greider, C.W., and Blackburn, E.H. (1985). Identification of a Specific Telomere Terminal Transferase-Activity in Tetrahymena Extracts. Cell *43*, 405-413.

Gu, Y., Honig, L.S., Schupf, N., Lee, J.H., Luchsinger, J.A., Stern, Y., and Scarmeas, N. (2015). Mediterranean diet and leukocyte telomere length in a multi-ethnic elderly population. Age (Dordr) *37*, 24.

Hada, M., Cucinotta, F.A., Gonda, S.R., and Wu, H. (2007). mBAND analysis of chromosomal aberrations in human epithelial cells exposed to low- and high-LET radiation. Radiat Res *168*, 98-105.

Haendeler, J., Hoffmann, J., Brandes, R.P., Zeiher, A.M., and Dimmeler, S. (2003). Hydrogen peroxide triggers nuclear export of telomerase reverse transcriptase via Src kinase family-dependent phosphorylation of tyrosine 707. Mol Cell Biol *23*, 4598-4610.

Harley, C.B., Futcher, A.B., and Greider, C.W. (1990). Telomeres Shorten During Aging of Human Fibroblasts. Nature *345*, 458-460.

Haycock, P.C., Heydon, E.E., Kaptoge, S., Butterworth, A.S., Thompson, A., and Willeit, P. (2014). Leucocyte telomere length and risk of cardiovascular disease: systematic review and meta-analysis. BMJ *349*, g4227.

Honig, L.S., Kang, M.S., Cheng, R., Eckfeldt, J.H., Thyagarajan, B., Leiendecker-Foster, C., Province, M.A., Sanders, J.L., Perls, T., Christensen, K.*, et al.* (2015). Heritability of telomere length in a study of long-lived families. Neurobiol Aging *36*, 2785-2790.

Honig, L.S., Kang, M.S., Schupf, N., Lee, J.H., and Mayeux, R. (2012). Association of shorter leukocyte telomere repeat length with dementia and mortality. Arch Neurol *69*, 1332-1339.

Hou, M., Xu, D., Bjorkholm, M., and Gruber, A. (2001). Real-time quantitative telomeric repeat amplification protocol assay for the detection of telomerase activity. Clin Chem *47*, 519-524.

Hu, J., Hwang, S.S., Liesa, M., Gan, B., Sahin, E., Jaskelioff, M., Ding, Z., Ying, H., Boutin, A.T., Zhang, H.*, et al.* (2012). Antitelomerase therapy provokes ALT and mitochondrial adaptive mechanisms in cancer. Cell *148*, 651-663.

Konopka, A.R., Laurin, J.L., Schoenberg, H.M., Reid, J.J., Castor, W.M., Wolff, C.A., Musci, R.V., Safairad, O.D., Linden, M.A., Biela, L.M.*, et al.* (2019). Metformin inhibits mitochondrial adaptations to aerobic exercise training in older adults. Aging Cell *18*, e12880.

Lee, M., Napier, C.E., Yang, S.F., Arthur, J.W., Reddel, R.R., and Pickett, H.A. (2017). Comparative analysis of whole genome sequencing-based telomere length measurement techniques. Methods *114*, 4-15.

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics *34*, 3094-3100.

Liu, H., Xie, Y., Zhang, Z., Mao, P., Liu, J., Ma, W., and Zhao, Y. (2018). Telomeric Recombination Induced by DNA Damage Results in Telomere Extension and Length Heterogeneity. Neoplasia *20*, 905-916.

Liu, H.W., Kao, H.H., and Wu, C.H. (2019). Exercise training upregulates SIRT1 to attenuate inflammation and metabolic dysfunction in kidney and liver of diabetic db/db mice. Nutr Metab (Lond) *16*, 22.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*, 402-408.

Loehr, J.A., Guilliams, M.E., Petersen, N., Hirsch, N., Kawashima, S., and Ohshima, H. (2015).

Physical Training for Long-Duration Spaceflight. Aerosp Med Hum Perform *86*, A14-A23.
Low, K.C., and Tergaonkar, V. (2013). Telomerase: central regulator of all of the hallmarks of cancer. Trends Biochem Sci *38*, 426-434.

Maeda, T., Nakamura, K., Atsumi, K., Hirakawa, M., Ueda, Y., and Makino, N. (2013). Radiation-associated changes in the length of telomeres in peripheral leukocytes from inpatients with cancer. Int J Radiat Biol *89*, 106-109.

Marshall-Goebel, K., Laurie, S.S., Alferova, I.V., Arbeille, P., Aunon-Chancellor, S.M., Ebert, D.J., Lee, S.M.C., Macias, B.R., Martin, D.S., Pattarini, J.M.*, et al.* (2019). Assessment of Jugular Venous Blood Flow Stasis and Thrombosis During Spaceflight. JAMA Netw Open *2*, e1915011.

McKenna, M.J., and Bailey, S.M. (2017). Chromosomal and Telomeric biomarkers of normal tissue injury to evaluate risk of degenerative health effects (secondary malignancy, cardiovascular disease) post radiation therapy. Transl Cancer Res *6*, S789-S794.

McKenna, M.J., Robinson, E., Taylor, L., Tompkins, C., Cornforth, M.N., Simon, S.L., and Bailey, S.M. (2019). Chromosome Translocations, Inversions and Telomere Length for Retrospective Biodosimetry on Exposed U.S. Atomic Veterans. Radiat Res *191*, 311-322.

Miller, D., Reynolds, G.E., Mejia, R., Stark, J.M., and Murnane, J.P. (2011). Subtelomeric regions in mammalian cells are deficient in DNA double-strand break repair. DNA Repair (Amst) *10*, 536-544.

Miri, M., Nazarzadeh, M., Alahabadi, A., Ehrampoush, M.H., Rad, A., Lotfi, M.H., Sheikhha, M.H., Sakhvidi, M.J.Z., Nawrot, T.S., and Dadvand, P. (2019). Air pollution and telomere length in adults: A systematic review and meta-analysis of observational studies. Environ Pollut *244*, 636-647.

Moyzis, R.K., Buckingham, J.M., Cram, L.S., Dani, M., Deaven, L.L., Jones, M.D., Meyne, J., Ratliff, R.L., and Wu, J.R. (1988). A Highly Conserved Repetitive Dna-Sequence ; (Ttaggg)N ; Present At the Telomeres of Human-Chromosomes. Proceedings of the National Academy of Sciences of the United States of America *85*, 6622-6626.

Murnane, J.P., Sabatier, L., Marder, B.A., and Morgan, W.F. (1994). Telomere Dynamics in an Immortal Human Cell-Line. Embo Journal *13*, 4953-4962.

O'Callaghan, N.J., and Fenech, M. (2011). A quantitative PCR method for measuring absolute telomere length. Biol Proced Online *13*, 3.

Olovnikov, A.M. (1971). Principle of marginotomy in template synthesis of polynucleotides. Dokl Akad Nauk SSSR *201*, 1496-1499.

Potapova, T.A., Unruh, J.R., Yu, Z., Rancati, G., Li, H., Stampfer, M.R., and Gerton, J.L. (2019). Superresolution microscopy reveals linkages between ribosomal DNA on heterologous chromosomes. J Cell Biol *218*, 2492-2513.

Puterman, E., Weiss, J., Lin, J., Schilf, S., Slusher, A.L., Johansen, K.L., and Epel, E.S. (2018). Aerobic exercise lengthens telomeres and reduces stress in family caregivers: A randomized controlled trial - Curt Richter Award Paper 2018. Psychoneuroendocrinology *98*, 245-252.

Ray, F.A., Robinson, E., McKenna, M., Hada, M., George, K., Cucinotta, F., Goodwin, E.H., Bedford, J.S., Bailey, S.M., and Cornforth, M.N. (2014). Directional genomic hybridization: inversions as a potential biodosimeter for retrospective radiation exposure. Radiat Environ Biophys *53*, 255-263.

Ray, F.A., Zimmerman, E., Robinson, B., Cornforth, M.N., Bedford, J.S., Goodwin, E.H., and Bailey, S.M. (2013). Directional genomic hybridization for chromosomal inversion discovery and detection. Chromosome Res *21*, 165-174.

Robinson, E., McKenna, M.J., Bedford, J.S., Goodwin, E.H., Cornforth, M.N., Bailey, S.M., and Ray, F.A. (2019). Directional Genomic Hybridization (dGH) for Detection of Intrachromosomal Rearrangements. Methods Mol Biol *1984*, 107-116.

Rochette, P.J., and Brash, D.E. (2010). Human telomeres are hypersensitive to UV-induced DNA Damage and refractory to repair. PLoS Genet *6*, e1000926.

Saretzki, G. (2009). Telomerase, mitochondria and oxidative stress. Exp Gerontol *44*, 485-492.

Scott, J.P.R., Weber, T., and Green, D.A. (2019). Introduction to the Frontiers Research Topic: Optimization of Exercise Countermeasures for Human Space Flight - Lessons From Terrestrial Physiology and Operational Considerations. Front Physiol *10*, 173.

Sfeir, A., Kosiyatrakul, S.T., Hockemeyer, D., MacRae, S.L., Karlseder, J., Schildkraut, C.L., and de Lange, T. (2009). Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication. Cell *138*, 90-103.

Shay, J.W. (2013a). Are Short Telomeres Hallmarks of Cancer Recurrence? Clin Cancer Res.

Shay, J.W. (2013b). Are short telomeres predictive of advanced cancer? Cancer Discov *3*, 1096-1098.

Shim, G., Ricoul, M., Hempel, W.M., Azzam, E.I., and Sabatier, L. (2014). Crosstalk between telomere maintenance and radiation effects: A key player in the process of radiation-induced carcinogenesis. Mutat Res Rev Mutat Res.

Simpson, R.J., Cosgrove, C., Chee, M.M., McFarlin, B.K., Bartlett, D.B., Spielmann, G., O'Connor, D.P., Pircher, H., and Shiels, P.G. (2010). Senescent phenotypes and telomere lengths of peripheral blood T-cells mobilized by acute exercise in humans. Exerc Immunol Rev *16*, 40-55.

Sishc, B.J., Nelson, C.B., McKenna, M.J., Battaglia, C.L., Herndon, A., Idate, R., Liber, H.L., and Bailey, S.M. (2015). Telomeres and Telomerase in the Radiation Response: Implications for Instability, Reprograming, and Carcinogenesis. Front Oncol *5*, 257.

Smith, M.G., Kelley, M., Basner, M. (2020). A brief history of spaceflight from 1961 to 2020: an analysis of missions and astronaut demographics. Acta Astronautica *175*, 290-299.

Smith, S.M., Heer, M., Wang, Z., Huntoon, C.L., and Zwart, S.R. (2012). Long-duration space flight and bed rest effects on testosterone and other steroids. J Clin Endocrinol Metab *97*, 270-278.

Sridharan, D.M., Asaithamby, A., Bailey, S.M., Costes, S.V., Doetsch, P.W., Dynan, W.S., Kronenberg, A., Rithidech, K.N., Saha, J., Snijders, A.M.*, et al.* (2015). Understanding cancer development processes after HZE-particle exposure: roles of ROS, DNA damage repair and inflammation. Radiat Res *183*, 1-26.

Stone, R.C., Horvath, K., Kark, J.D., Susser, E., Tishkoff, S.A., and Aviv, A. (2016). Telomere Length and the Cancer-Atherosclerosis Trade-Off. PLoS Genet *12*, e1006144.

Stong, N., Deng, Z., Gupta, R., Hu, S., Paul, S., Weiner, A.K., Eichler, E.E., Graves, T., Fronick, C.C., Courtney, L.*, et al.* (2014). Subtelomeric CTCF and cohesin binding site organization using improved subtelomere assemblies and a novel annotation pipeline. Genome Res *24*, 1039-1050.

Tyrka, A.R., Carpenter, L.L., Kao, H.T., Porton, B., Philip, N.S., Ridout, S.J., Ridout, K.K., and Price, L.H. (2015). Association of telomere length and mitochondrial DNA copy number in a community sample of healthy adults. Exp Gerontol *66*, 17-20.

Valdes, A.M., Andrew, T., Gardner, J.P., Kimura, M., Oelsner, E., Cherkas, L.F., Aviv, A., and Spector, T.D. (2005). Obesity, cigarette smoking, and telomere length in women. Lancet *366*, 662-664.

Vasan, R.S., Demissie, S., Kimura, M., Cupples, L.A., Rifai, N., White, C., Wang, T.J., Gardner, J.P., Cao, X., Benjamin, E.J.*, et al.* (2008). Association of leukocyte telomere length with circulating biomarkers of the renin-angiotensin-aldosterone system: the Framingham Heart Study. Circulation *117*, 1138-1144.

von Zglinicki, T. (2000). Role of oxidative stress in telomere length regulation and replicative senescence. Ann N Y Acad Sci *908*, 99-110.

Watson, J.D. (1972). Origin of concatameric T7 DNA. Nature: New Biology *239*, 197-201.
Wilson, D.M., 3rd, and Thompson, L.H. (2007). Molecular mechanisms of sister-chromatid exchange. Mutat Res *616*, 11-23.

Winder, W.W., Holmes, B.F., Rubink, D.S., Jensen, E.B., Chen, M., and Holloszy, J.O. (2000). Activation of AMP-activated protein kinase increases mitochondrial enzymes in skeletal muscle. J Appl Physiol (1985) *88*, 2219-2226.

Wong, H.P., and Slijepcevic, P. (2004). Telomere length measurement in mouse chromosomes by a modified Q-FISH method. Cytogenet Genome Res *105*, 464-470.

Wu, J.Z., and Ho, P.C. (2009). Comparing the relative oxidative DNA damage caused by various arsenic species by quantifying urinary levels of 8-hydroxy-2'-deoxyguanosine with isotope-dilution liquid chromatography/mass spectrometry. Pharm Res *26*, 1525-1533.

Wu, Z., Oeck, S., West, A.P., Mangalhara, K.C., Sainz, A.G., Newman, L.E., Zhang, X.O., Wu, L.Z., Yan, Q., Bosenberg, M.*, et al.* (2019). Mitochondrial DNA stress signalling protects the nuclear genome. Nat Metab *1*, 1209-+.

Yang, H., and Fogo, A.B. (2010). Cell senescence in the aging kidney. J Am Soc Nephrol *21*, 1436-1439.

Zhang, J., Rane, G., Dai, X., Shanmugam, M.K., Arfuso, F., Samy, R.P., Lai, M.K., Kappei, D., Kumar, A.P., and Sethi, G. (2016). Ageing and the telomere connection: An intimate relationship with inflammation. Ageing Res Rev *25*, 55-69.

Zheng, Y.L., Zhang, F., Sun, B., Du, J., Sun, C., Yuan, J., Wang, Y., Tao, L., Kota, K., Liu, X.*, et al.* (2014). Telomerase enzymatic component hTERT shortens long telomeres in human cells. Cell Cycle *13*, 1765-1776.

Zwart, S.R., Booth, S.L., Peterson, J.W., Wang, Z., and Smith, S.M. (2011). Vitamin K status in spaceflight and ground-based models of spaceflight. J Bone Miner Res *26*, 948-954.

CHAPTER 3: TELOMERE LENGTH DYNAMICS AND DNA DAMAGE RESPONSES ASSOCIATED WITH LONG-DURATION SPACEFLIGHT[2]

Chapter summary

Here, we follow our previous investigation on three unrelated astronauts undergoing one-year and 6-month missions aboard the ISS. Previously we observed telomere elongation in-flight which returned to baseline at post-flight though trended shorter; and found telomeric and cytogenetic evidence implicating ALT as the responsible mechanism. We also observed significant elevations in DNA damage responses for all astronauts during and after spaceflight. Now we aimed to evaluate general trends in telomeric and cytogenetic responses to spaceflight in a larger cohort of 11 astronauts, as well as continue our inquiries into the possibility of ALT-activation during spaceflight; namely whether previous observations represent a general adaptive response to chronic oxidative stress. Consistent with our previous studies, we again observe a trend towards shorter telomere length after spaceflight compared to before, significant elevations in DNA damage responses from spaceflight, and indications of chronic oxidative stress during spaceflight. Consistent with our telomeric and cytogenetic evidence for ALT-activation during spaceflight, we find telomeric and sequencing indications of ALT-activation during exposure to another extreme environment, Mt. Everest, suggesting that ALT-activation may potentially be a general adaptive response to chronic oxidative stress.

[2] This chapter is a draft of a first author manuscript accepted at Cell Reports. Anticipated publication date is Nov-Dec 2020.

Authors:

Jared J. Luxton[1,2], Miles J. McKenna[1,2], Aidan Lewis[1], Lynn E. Taylor[1], Kerry A. George[3], Sameer M. Dixit[4], Matthew Moniz[5], Willie Benegas[6], Daniel Butler[7], Daniela Bezdan[7], Cem Meydan[7], Brian E. Crucian[8], Sara R. Zwart[9], Scott M. Smith[8], Christopher E. Mason[7,10,11], and Susan M. Bailey[1,2*]

Affiliations:

[1]Department of Environmental and Radiological Health Sciences, and [2]Cell and Molecular Biology Program, Colorado State University, Fort Collins, CO, USA

[3]KBR, Houston, TX, USA

[4] Center for Molecular Dynamics - Nepal (CMDN), Nepal

[5]Dartmouth College, Hanover, NH USA

[6]Benegas Brothers Expedition, Sandy, Utah USA

[7]Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY, USA

[8]Human Health and Performance Directorate, NASA Johnson Space Center, Houston, TX, USA

[9]University of Texas Medical Branch, Galveston, TX, USA

[10]The WorldQuant Initiative for Quantitative Prediction, Weill Cornell Medicine, New York, NY, USA

[11]The HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine, Weill Cornell Medicine, New York, NY, USA

Overview

Telomeres are repetitive terminal features of chromosomes essential for maintaining genome stability, that shorten with cell division, and thus with aging. Telomere length is also influenced by a variety of lifestyle factors, stresses, and environmental exposures, and so over time provides a robust biomarker of health, aging, and aging trajectories. Astronauts on long-duration missions work in an extreme environment fraught with unique physiological stressors. We assessed telomere length dynamics and DNA damage responses in a cohort of unrelated astronauts (n=11) before, during, and after one-year or shorter duration missions aboard the International Space Station. Astronauts are generally healthy individuals, yet they tended to have significantly shorter telomeres than age and sex matched ground control subjects before spaceflight. Astronauts also had significantly lower telomerase activity than the control cohort, and levels were similar before and after spaceflight. Although in-flight samples were not available for most of the astronauts, telomere elongation during spaceflight was observed for all crewmembers analyzed, irrespective of mission duration. All crewmembers experienced chronic oxidative stress during spaceflight, supportive of our supposition that the telomerase-independent Alternative Lengthening Telomere pathway may be transiently activated in extreme environments. Increased frequencies of telomeric and chromosomal aberrations further strengthened mechanistic connections to chronic exposure to the space radiation milieu. Upon return to Earth, telomere length shortened rapidly and, overall, astronauts had shorter telomeres after spaceflight than they did before. This larger cohort of astronauts facilitates establishing significance and importantly, also identifies inter-individual differences in response to long-duration spaceflight across these key biological pathways.

Introduction

The ends of human chromosomes are capped by telomeres, tandem arrays of repetitive G-rich sequence bound by a plethora of associated proteins that protect chromosomal termini from degradation and loss (de Lange, 2009; Moyzis et al., 1988). Telomeres preserve genome stability by preventing the natural ends of chromosomes from being recognized as broken DNA (double-strand breaks [DSBs]) and triggering inappropriate DNA damage responses (DDRs). Due to the end-replication problem, telomere length erodes with cell division (~50-100 bp/cell division), causing telomeres to shorten until they reach a critically short length, at which point a permanent cell cycle arrest known as replicative senescence is entered (Harley et al., 1990; Olovnikov, 1971; Watson, 1972). The specialized reverse transcriptase, telomerase, counteracts telomere attrition via *de novo* addition of telomeric repeats onto the ends of newly replicated chromosomes (Greider and Blackburn, 1985). However, telomerase activity is prominent only in highly proliferative stem, germline, and cancer cells; its levels are insufficient to maintain telomere length in normal somatic cells (Kim et al., 1994; Yuan et al., 2019).

While it is well established that telomere length is an inherited trait that erodes with normal aging (Aubert and Lansdorp, 2008; Honig et al., 2015), oxidative stress, and inflammation (von Zglinicki, 2000; Zhang et al., 2016), it is becoming increasingly appreciated that telomere length is influenced by a variety of other factors as well, including sex (Gardner et al., 2013) and lifestyle factors [e.g., diet (Gu et al., 2015), smoking and obesity (Valdes et al., 2005), physical activity (Cherkas et al., 2008), psychological stress (Epel et al., 2004)], chronic stress and disease (Blackburn and Epel, 2012). Telomeres are also regarded as sentinels of environmental exposures, as their length is affected by air pollution, ultraviolet (UV) and ionizing radiations (Miri et al., 2019; Rochette and Brash, 2010; Shim et al., 2014). Therefore,

maintenance of telomere length represents a central integrating component of the cumulative effects of genetic, environmental, and lifestyle factors; i.e., the rate at which telomeres shorten provides an informative biomarker of general health, as well as an indication how quickly and/or how well one might be aging. Furthermore, altered telomere length can be linked to age-related pathologies, including dementias, cardiovascular disease (CVD), and cancer (Cohen et al., 2013; Haycock et al., 2014; Honig et al., 2012; Shay, 2013a, b). Quantitative estimates recently demonstrated that indeed, both short and long telomere length are associated with increased disease risk – of approximately equal degree – supporting the concept of a cancer-aging/disease trade-off (Protsenko et al., 2020; Stone et al., 2016). Thus, reports of both shortened and elongated telomeres after exposure to various types of radiation (Berardinelli et al., 2013; Maeda et al., 2013) have implications for cancer risk and a variety of degenerative late effects, including CVD, as an inverse relationship between telomere length in peripheral blood cells and coronary heart disease has been demonstrated (Haycock et al., 2014). Such late health effects relevant to spaceflight are largely unknown and controversial (Cucinotta et al., 2016; Delp et al., 2016), yet they have very real potential for influencing astronaut performance during, and health following, long-duration missions.

Astronauts participating in long duration missions work in an extreme environment and experience a myriad of chronic stressors, including space radiation exposure (galactic cosmic rays [GCR], solar particle events [SPE], and secondary particles that arise from interactions with spacecraft shielding), microgravity, psychological isolation, altered nutrition and microbiome, in addition to intermittent bouts of acute stress (e.g., extravehicular activities [EVAs], endurance exercise) (*Afshinneko E., et al 2020, Cell perspective*). Considering the unique combination of stressors associated with spaceflight, even in low Earth orbit, and the adverse effects on multiple

physiological systems (e.g., dysregulated immunity, inflammation, infection), we proposed that telomere length dynamics (changes over time) are of particular relevance for astronauts because they provide insight into individual health status during a mission, as well as potential implications for aging and aging trajectories thereafter.

Results

Telomere length dynamics associated with spaceflight

For all NASA astronauts (n=11) and healthy age and sex matched ground control subjects (n=11), and as for our Twins Study investigation (Garrett-Bakelman et al., 2019), whole blood was collected at multiple timepoints before, during, and after spaceflight (one-year or shorter duration missions), and telomere length measured using two independent assays. DNA was isolated from separated peripheral blood mononuclear cells (PBMCs) and mean telomere length assessed by quantitative polymerase chain reaction (qPCR) (**Figure 3.1A**). Cell-by-cell analyses using telomere fluorescence *in situ* hybridization (Telo-FISH) on metaphase chromosomes (stimulated T-cells) was also performed, and provided telomere length measures for thousands of individual telomeres based on relative fluorescence intensity (**Figure 3.1B)** (Luxton et al., *Cell Reports 2020*). Overall, trends were similar for both assays, and the sizeable number of subjects and timepoints served to establish significance. Telomere length was similarly evaluated in the ground control cohort and remained relatively stable over the course of the study, just as it had for the Earth-bound twin astronaut (Garrett-Bakelman et al., 2019); therefore, all values from all control subjects were combined and averaged to establish a baseline for standardization. Average telomere length at baseline was, or trended toward, significantly shorter for the astronauts than for the ground control cohort. As previously reported for the One Year Mission crewmember,

and irrespective of mission duration or means of measurement, significantly longer telomeres were observed at all in-flight timepoints, in all samples analyzed, and for all crewmembers for whom inflight samples were available (n=3) (Garrett-Bakelman et al., 2019) (Luxton et al., *Cell Reports 2020*). Even so, overall, astronauts had, or trended toward, significantly shorter telomeres after spaceflight than before.

Longitudinal analyses revealed dynamic shifts in telomere length for the astronaut cohort in general, and clustering analyses of individual telomeric responses identified groups of crewmembers displaying similar telomeric responses before and after spaceflight (**Figure 3.2**). In general, dramatic decreases in telomere length (qPCR; PBMCs) were observed soon after return to Earth (1-7 days after return; R+7), with an overall trend of shorter mean telomere length after spaceflight than before (**Figure 3.2A**). Hierarchical clustering of all astronauts by mean telomere length at all pre- and post-flight timepoints, failed to identify any clustered groups (**Figure 3.2B**). However, when astronauts were clustered by changes in mean telomere length specifically at the earliest timepoint, 270 days before spaceflight (L-270), soon after return to Earth (R+7), and at the last timepoint, 270 days after return to Earth (R+270), three groups of similarly responding crewmembers clearly emerged, thereby identifying distinct telomeric responses to spaceflight during these decisive periods (**Figure 3.2C**). The distinguishing features of the groups included an overall downward trend in mean telomere length after spaceflight (R+270), with either a sharp decrease (group 2) or no change (group 3) soon after return to Earth (R+7), and for one crewmember (group 1), an *increase* in mean telomere length after spaceflight at both R+7 and R+270 (**Figure 3.2D**).

Analyses of mean telomere length using Telo-FISH (stimulated T-cells) before and after spaceflight, revealed cyclical-like changes over the course of the study (**Figure 3.3A**), an

observation potentially reflective of the on-going nature of T-cell exhaustion and replacement during hematopoiesis. Longitudinal hierarchical clustering of all astronauts by mean telomere length at all pre- and post-flight timepoints, did not identify any clustered groups **(Figure 3.3B**). However, when astronauts were clustered by changes in mean telomere length specifically at the earliest pre-flight (L-270), immediate post-flight (R+7), and last post-flight (R+270) timepoints, three groups of similar responders again emerged (**Figure 3.3C**), and displayed the same general trends in telomeric responses to spaceflight as the qPCR results (**Figure 3.3D**).

In addition to average telomere length, Telo-FISH data can be used to generate individual telomere length distributions and monitor shifts in populations of short and long telomeres over time (**Figure 3.4**). Individual telomere length distributions at the earliest pre-flight (L-270) and final post-flight (R+270) timepoints for the three groups of crewmembers identified by hierarchical clustering, confirmed more *long* telomeres after spaceflight for group 1 (n=3 astronauts), more *short* telomeres after spaceflight for group 2 (n=5 astronauts), and relatively unchanged telomere length distributions for group 3 (n=3 astronauts).

For the astronauts for whom in-flight samples were available (n=3) (Luxton et al., *Cell Reports, 2020*), longitudinal hierarchical clustering of mean telomere length (qPCR, Telo-FISH) before, during, and after spaceflight identified two groups of similar responders (**Figure 3.5A-D**). Trends between the assays were similar, and interestingly, clustered group 1 included one 6-month mission crewmember, whereas group 2 included one-year and 6-month mission crewmembers, both of whom displayed very similar telomeric responses to spaceflight. These findings further support our supposition that individual variation in response, rather than mission duration, is the stronger determinant of spaceflight-specific telomere elongation.

Telomerase activity

Although telomerase activity is the most obvious candidate for *de novo* telomere elongation during spaceflight, it was not possible to directly access telomerase activity in samples collected on the ISS. Similar to our experience and results from the Twins Study (Garrett-Bakelman et al., 2019), telomerase activity was "*lost in space*" for all blood samples collected during flight, likely due to the unavoidable heat and/or time associated with return of ambient samples from the ISS to Earth on the Soyuz vehicle (Luxton et al., *Cell Reports, 2020*). Overall, for this cohort of astronauts (n=11), levels of telomerase activity before spaceflight were significantly lower than for the healthy age and sex matched ground control cohort, and were similar before and after spaceflight for the crewmembers (**Figure 3.6A**). There was a great deal of individual variation in the levels of telomerase activity between timepoints (**Figure 3.6B**), and neither hierarchical clustering of telomerase activity at all timepoints, nor at the earliest pre-flight (L-270), immediate post-flight (R+7), and last post-flight (R+270) timepoints, resulted in clearly defined groups of responses (**Figure 3.6C, D**). A notable exception however, was the first timepoint after return to Earth (R+7), which showed a striking compression of variation in telomerase activity (**Figure 3.6B**). This observation is potentially suggestive of a shared mechanism of regulation/dysregulation during spaceflight, or in response to the abrupt absence of spaceflight specific factors (e.g., microgravity, space radiation, environmental conditions), and/or to the severe stresses associated with return to Earth (e.g., physical, psychological, immunological).

Biochemical Profiles and Telomere Length Dynamics

Potential mechanistic insight into spaceflight-specific telomere elongation was sought via interrogation of biochemical profiles, specifically for this astronaut cohort (**Tables 3.1, 3.2**). Biochemical analytes in blood and urine for individual crewmembers were correlated with mean telomere length (qPCR, Telo-FISH) before, during and after spaceflight (**Figure 3.7**). There was a strong positive correlation between mean telomere length (qPCR) and serum ferritin (**Figure 3.7A, B**); although high body iron status has been associated with shorter telomere length in adults on Earth (Liu et al., 2019a). However, iron stores during spaceflight have been correlated with increased oxidative stress (Zwart et al., 2013) and indeed, urinary excretion of two biomarkers of oxidative damage, 8-iso-prostaglandin F2-α (PGF-2α) and 8-hydroxydeoxyguanosine (8-OHdG), were highly correlated with mean telomere length (**Figure 3.7C, F**). Biologically related transferrin was negatively correlated with mean telomere length (**Figure 3.7B, E**), but the relationship to telomere length regulation during spaceflight is unclear. Serum levels of folate were also negatively correlated with mean telomere length, a finding consistent with a previously demonstrated inverse relationship between folate concentration and telomere length when folate levels are low (Paul et al., 2009).

Interestingly, plasma concentrations of inflammatory cytokines (interleukins; IL-4, IL-10, IL-5, IL-1a, IL-2) and chemokines (CCL5, CCL4, CXCL5), and vascular endothelial growth factor 1 (VEGF-1) were strongly correlated with mean telomere length before, during, and after spaceflight (**Figure 3.7A, B, D, E**). IL-4 is involved in Th2 mediated humoral and adaptive immune responses and regulation of hematopoiesis (induces differentiation of naïve helper T-cells), alterations in which have been associated with spaceflight (Garrett-Bakelman et al., 2019). Consistent with the elevated plasma concentrations of VEGF-1 during spaceflight observed in this cohort of astronauts, gene expression of VEGF-1, and the related hypoxia-inducible factor 1-

alpha (HIF1-α) were increased during spaceflight for the one-year mission astronaut (Garrett-Bakelman et al., 2019). Of particular relevance to telomere elongation during spaceflight, HIF-1α has been implicated in mediating upregulation and activation of hTERT, the catalytic reverse transcriptase subunit of human telomerase (Minamino et al., 2001; Nishi et al., 2004); this finding is potentially indicative of elevated expression of hTERT during spaceflight as previously suggested for the one-year mission astronaut (Garrett-Bakelman et al., 2019). Interestingly, hTERT has a number of non-canonical roles related to cell proliferation and differentiation, including in VEGF-mediated angiogenesis as indicated here, and supportive of intimate co-regulation of TERT and VEGF as previously suggested (Hartwig et al., 2012). Regulation of HIF-1 is also complex and not exclusively related to reduced levels of oxygen. Consistent with elevated oxidative stress during spaceflight, reactive oxygen species (ROS) have been shown to stabilize HIF-1α protein and stimulate transcription of downstream targets (Jung et al., 2008); e.g., hTERT/VEGF, and tumor necrosis factor-alpha (TNF-α) (Yu et al., 2012), which also positively correlated with telomere length. Together these findings suggest that while the ISS is generally a "normoxic" environment, transiently altered atmospheres (e.g., elevated CO_2; hypercapnia), and/or a myriad of other spaceflight stressors, including chronic oxidative stress, physiological dysregulation, and even endurance exercise (to counteract ever present deconditioning due to hypokinesia), may result in stimulation of the HIF-1α/hTERT/VEGF-1 axis (Lerman et al., 2010; Ohno, 2012) (Luxton et al., *Cell Reports, 2020*).

Chronic space radiation exposure and DNA damage responses

Oxidative damage. Consistent with reports of persistent mitochondrial and chronic oxidative stress during spaceflight (Garrett-Bakelman et al., 2019) (Luxton et al, *Cell Reports, 2020;*

Silveira et al., Cell 2020; Bezdan et al., *iScience 2020*), biomarkers of oxidative stress, 8-iso-prostaglandin F2-α (PGF-2α) and 8-hydroxydeoxyguanosine (8-OHdG), were highly correlated with mean telomere length in this astronaut cohort (**Figure 3.7C, F**). Interestingly, increased urinary excretion of 8-OHdG has been associated with occupational low-dose radiation exposure, including in the urine of intercontinental airline pilots (Silva et al., 2013). Ionizing radiation exposures, particularly high Linear Energy Transfer (LET)/High atomic number (Z) and Energy (HZE) exposures such as those encountered during spaceflight, induce persistent ROS production by mitochondria, which were increased in the one-year mission astronaut (Garrett-Bakelman et al., 2019), DNA damage signaling, and chronic oxidative stress (Azzam et al., 2012; Sridharan et al., 2015). Furthermore, due to their G-rich nature, telomeric regions are especially susceptible to oxidative damage, which tends to accumulate because telomeres are refractory to repair (Miller et al., 2011). Of relevance to spaceflight-specific telomere elongation, chronic oxidative stress has been shown to transiently activate the Alternative Lengthening of Telomeres (ALT) pathway (Coluzzi et al., 2017; Hu et al., 2012) – a telomerase independent, recombination-based mechanism of telomere maintenance (Bryan et al., 1997; Dunham et al., 2000; Murnane et al., 1994). Cytogenetic analysis revealed evidence of oxidative damage at telomeres specifically during spaceflight, in that heterogeneous telomere lengths, a hallmark of ALT, were significantly elevated (Luxton et al., *Cell Reports, 2020).*

Recently, we found significantly longer telomeres in humans experiencing a “high altitude” extreme environment on Earth, one also associated with increased low dose radiation exposure, altered atmospheres, and endurance exercise – climbing Mt. Everest. Similar to astronauts during spaceflight, the climbers (n=2) had longer average telomere length (in blood) immediately post-summit, which returned to pre-climb baseline values after decent (**Figure**

3.8A, B). Additionally, RNA-seq analysis revealed evidence of chronic oxidative stress and activation of the ALT pathway of telomere maintenance. Expression of the catalytic subunit of human telomerase, hTERT, essentially flatlined during the climb, while the non-coding RNA component, hTERC, was significantly upregulated (**Figure 3.8C**). Independent of telomerase, TERC has been shown to promote inflammatory responses, act as a downstream signal of mitochondrial function in response to stress, and protect from oxidative stress (Eitan et al., 2016; Liu et al., 2019b; Zheng et al., 2019). Expression of RAD50, a key homologous recombinational repair protein involved in ALT, was also significantly elevated during the climb in both individuals (**Figure 3.8C**). Pathway analysis provided additional support for telomere maintenance by recombination and ALT, but not for the telomerase-mediated pathway (**Figure 3.9A, B, C; Table 3.3**). Together, these results provide intriguing mechanistic insight into potential telomere adaptation in extreme environments involving chronic exposure to oxidative stress and transient activation of ALT.

In the present study, with the inclusion of additional astronauts and pre- and post-flight timepoints, other recombination events indicative of replication stress and oxidative damage (Wilson and Thompson, 2007), specifically frequencies of sister chromatid exchange (SCE) that occurred along the length of chromosomes (classic) or into subtelomeric regions, were not significantly increased during spaceflight (**Figure 3.10A**). Frequencies of terminal SCE, those specifically occurring at the subtelomeric heterochromatin junction of chromosome 2p, a likely fragile site, were significantly increased during spaceflight, and remained elevated post-flight. Satellite associations, nucleoprotein bridges involving the short p-arms of 2-5 human acrocentric chromosomes that represent linkages between ribosomal (r)DNA (Potapova et al., 2019), have been proposed as early biomarkers of chronic exposure to low levels of ionizing radiation in

occupationally exposed hospital workers (Caradonna, 2015). Here, significantly increased frequencies of satellite associations were observed during spaceflight.

IR-induced double-strand breaks (DSBs). Chromosome aberrations represent informative signatures of ionizing radiation-induced DNA damage (specifically DSBs) that have been previously evaluated in blood lymphocytes of astronauts (Cucinotta, 2014; Garrett-Bakelman et al., 2019; George et al., 2013). Here, strand-specific directional genomic hybridization (dGH) paints (Ray et al., 2013) for chromosomes 1, 2, and 3, together with chromosome-specific subtelomere probes to distinguish terminal inversions from SCEs (McKenna and Bailey, 2017), were used to simultaneously detect rearrangements *between* (dicentrics, translocations), and *within* (inversions) chromosomes. For the astronaut cohort, and consistent with our previous findings and proposition that inversions are particularly informative biomarkers of exposure to high-LET and space radiations (Garrett-Bakelman et al., 2019; McKenna et al., 2019; Ray et al., 2014), background frequencies of inversions were higher than for dicentrics and translocations, and frequencies of inversions significantly increased during spaceflight (**Figure 3.10B**) (Luxton et al., *Cell Reports, 2020*). Furthermore, significantly increased frequencies of inversions persisted after spaceflight for the astronaut cohort, suggestive of genome instability, stem cell damage, and/or clonal hematopoiesis (growth advantage), [(Garrett-Bakelman et al., 2019), Trinchant/Hassane et al*., Cell Stem Cell 2020*]. Although frequencies of dicentrics were much lower, they too were significantly increased during spaceflight, consistent with ionizing radiation exposure and their historical utility for biodosimetry (Lloyd et al., 1973).

For the astronaut cohort, frequencies of inversions approximately doubled between pre- and post-spaceflight measures (**Figure 3.10B**). Longitudinal hierarchical clustering of individual

astronaut inversion frequencies (**Figure 3.11**) identified two groups of similar DNA damage responses to spaceflight. Importantly, the astronaut with the most accumulated time in space had the highest yield of inversions (**Figure 3.11A, B**), consistent with inversions being informative, and persistent, biomarkers of space radiation exposure. Similarly, longitudinal hierarchical clustering analyses of inversion frequencies for the crewmembers for whom in-flight samples were available (**Figure 3.11C, D**), identified two groups of responses based on mission duration; the one-year mission astronaut (group 1), and the 6-month mission astronauts (group 2). Inversion frequencies were highest during spaceflight, then diminished and converged to relatively similar, elevated levels post-flight. Additionally, and as would be expected for radiation-induced rearrangements, a positive and reasonably strong relationship was observed between post-flight inversion frequencies and available total lifetime radiation dose estimates (**Figure 3.12A**). Evidence of dose-dependent cell killing was also apparent, as inversion frequencies, a proxy for radiation dose, negatively correlated with white blood cell (WBC) counts after spaceflight exposure, but not before (**Figure 3.12B, C**). Consistent with redistribution of leukocyte subsets and elevated WBC counts during long-duration spaceflight (Crucian et al., 2015), a dramatic spike in WBC counts was observed immediately upon return to Earth (R+0) in this astronaut cohort (**Figure 3.12D**), which was likely stress-induced and indicative of a heightened immune response (Garrett-Bakelman et al., 2019).

Telomeric and DNA damage signatures of spaceflight

Here, we demonstrate that clustering analysis is capable of identifying groups of astronauts who responded similarly to spaceflight across independent telomeric and cytogenetic endpoints. We further hypothesized that spaceflight-specific signatures could be identified

through a more holistic analysis across these fundamental endpoints; specifically, telomere length (qPCR, Telo-FISH) and DNA damage responses (inversions, translocations, dicentrics, terminal SCEs, satellite associations). For this, we employed Uniform Manifold Approximation and Projection (UMAP) for dimension reduction (McInnes, 2018), an approach that provides powerful and intuitive visualizations, where the distance between points can be interpreted literally in terms of similarity. To determine if the data could distinguish between the different spaceflight stages (pre-, mid-, post-spaceflight), astronaut IDs and timepoint information was stripped from the data, UMAP performed, and labels reapplied for graphing purposes. Remarkably, there was clear separation of the mid-flight data, and while there were outliers in the pre- and post-flight datapoints, the vast majority trended toward separation (**Figure 3.13A**). These results suggest that spaceflight associated signatures exist and can be distinguished.

To determine whether astronauts exhibited similar responses to spaceflight across all endpoints analyzed, all data at all timepoints were combined for each individual astronaut (all data in a single row); biochemical analyte data could not be included due to the inability of UMAP to handle missing values. The data was stripped of astronaut IDs and timepoint information, UMAP performed, and IDs reapplied for graphing purposes. When considering all telomeric and cytogenetic data together, no specific groups of similarly responding astronauts emerge. Rather strikingly, each astronaut displays their own individual, integrated response to spaceflight across these key biological pathways (**Figure 3.13B**).

Discussion

America recently witnessed the successful launch and safe return of the first commercial rocket, SpaceX Falcon 9, with a Crew Dragon capsule carrying NASA astronauts Bob Behnken

and Doug Hurley into orbit and to the ISS. This milestone achievement ushered in a new era of spaceflight, one that targets the first woman and the next man on the Moon by 2024, and human exploration of Mars by the mid-2030s. Historically, a total of only 563 individuals have participated in spaceflight, the vast majority of whom have been males aged 35-55 years, on missions shorter than 20 days duration (Smith, 2020). Thus, as the number and diversity of space travelers and tourists increase, a better understanding of how longer duration spaceflight affects human health is critical to maintaining astronaut performance during, and improving aging trajectories following, future exploration missions.

Here, and as for our NASA Twins Study investigation (Garrett-Bakelman et al., 2019), we assessed telomere length dynamics and DDRs in a cohort of unrelated astronauts (n=11) before, during, and after one-year or shorter duration missions aboard the ISS. Although astronauts are generally healthy individuals, overall, they had significantly shorter telomeres than healthy age and sex matched ground control subjects (n=11) before flight, a finding perhaps reflective of the many stresses that astronauts experience during intense training, selection and preparation for spaceflight. Astronauts also had significantly lower telomerase activity than the control cohort, and levels were similar before and after spaceflight.

Most unexpected was the observation of significantly longer telomeres during spaceflight in all crewmembers (n=3) and all in-flight samples analyzed, irrespective of mission duration or means of measurement (Luxton et al., *Cell Reports, 2020*). Interestingly, the only other aging study ever done on the ISS reported slightly elongated telomeres in worms (*C. elegans)* flown for 11 days (Zhao et al., 2006). The larger cohort of astronauts reported here (n=11) served to further validate the significance of increases in telomere length and frequencies of telomeric and chromosomal aberrations observed during spaceflight, and provided additional support for our

previously proposed mechanistic connections to chronic exposure to the space radiation environment (Luxton et al., *Cell Reports, 2020*). Interestingly, upon return to Earth, telomere length shortened rapidly and significantly and, overall, astronauts had many more short telomeres after spaceflight than they did before, although inter-individual differences were observed.

While the definitive mechanisms involved in these extraterrestrial processes remain elusive, we propose that transient activation of ALT/ALT-like phenotypes in response to chronic oxidative stress, and/or changes in cell population dynamics resulting from radiosensitivity and redistribution of leukocyte subsets, likely contribute to the telomere elongation observed during spaceflight (Garrett-Bakelman et al., 2019) (Luxton et al., *Cell Reports, 2020*). During spaceflight, astronauts experience dysregulation of multiple physiological systems, and they undergo persistent low-level inflammation and oxidative stress, consistent with chronic space radiation exposure, and/or any combination of spaceflight specific stressors, including endurance/aerobic exercise (Afshinneko E., et al *Cell perspective, 2020*). Consistent with elevated WBC counts during spaceflight, and as is common for ISS crewmembers (Crucian et al., 2015), a dramatic increase in WBC counts was observed immediately upon return to Earth (R+0) in this cohort of astronauts, which was followed by a rapid decrease at post-flight day 3 (R+3). This is typically due to stress associated demargination of neutrophils, while absolute counts of lymphocytes and monocytes remain generally unaffected. Interestingly, telomere length dynamics followed a somewhat similar timeline, and elevated plasma concentrations of inflammatory cytokines and chemokines were highly correlated with telomere length before, during, and after spaceflight. The chronic, low-level inflammatory response induced during long-duration spaceflight, and the potential for on-going recruitment and release of increased numbers

of WBCs into the circulation, may inappropriately result in subsets of more naïve, stem-like/progenitor cells with inherently longer telomeres. While generally not present in the peripheral circulation of healthy individuals, the physiological dysregulation associated with spaceflight may facilitate such a process. Altered leukocyte distributions during spaceflight may result from a variety of stressors, including legitimate immune responses to pathogens; e.g., external bacteria and viruses, and/or latent herpesvirus reactivation, which cannot be mitigated by pre-flight quarantine. The health effects of spaceflight associated dysregulation of the immune system and dramatic fluctuations in telomere length dynamics are currently unknown. Here, a radiation dose-dependent decrease in WBC counts post-spaceflight was observed, providing additional support for appropriate individualized countermeasures during upcoming deep space exploration missions (Crucian et al., 2018).

We also find strong relationships between post-flight inversion frequencies and lifetime radiation dose estimates, additional evidence of the persistence of these intra-chromosomal rearrangements after exposure (McKenna et al., 2019), and further support of inversions as informative biomarkers of radiation exposure during spaceflight (Garrett-Bakelman et al., 2019; Hada et al., 2007; Ray et al., 2014). Chronic radiation exposure is one of the primary hazards of long-duration space travel, particularly as astronauts venture deeper into space and outside of the protection the Earth. The mechanistic links between exposure to the space radiation environment and the telomeric and DNA damage responses reported here, and elsewhere (Garrett-Bakelman et al., 2019) (Luxton et al., *Cell Reports, 2020*), support continued efforts and development of effective radiation mitigators for future long-duration missions.

The relatively large cohort of astronauts reported here facilitated characterization of general trends, and clustering analyses identified groups of crewmembers presenting similar

telomeric *or* DNA damage responses to spaceflight. Furthermore, when considering combined telomeric *and* DNA damage responses, spaceflight-specific signatures were evident (clustering of pre-, mid- and post-flight responses), however, no groups of similarly responding astronauts emerged. Rather, each astronaut displayed their own individual, integrated response to spaceflight across these key biological pathways. These results indicate that inter-individual variation in response to the combined stressors and exposures associated with spaceflight predominates over general trends, and highlights the necessity of personalized monitoring of astronauts. Our future studies aim to address differing telomere lengths in, and redistribution of, leukocyte subsets during spaceflight for individual astronauts on various duration missions.

Together, our findings related to telomere length dynamics and DNA damage responses during spaceflight have important implications for the health and performance of astronauts participating in long-duration missions, as well as for long-term aging and disease risk outcomes. Accumulating evidence supports an emerging view of telomere length not only as a biomarker but also a determinant of CVD and cancer (Stone et al., 2016), and chromosome aberrations have a long and infamous history of being associated with genome instability and a vast majority of cancers (Mitelman, 2000). Therefore, as more and more women and men journey back to the Moon and beyond in the coming years, identifying inter-individual differences, including sex-dependent ones, in response to the extreme environment, experiences and exposures associated with long-duration spaceflight, represents a critical next step in ensuring astronaut health during, and after, such missions.

Limitations: There are a myriad of limitations associated with long-duration spaceflight that unavoidably influence experimental design and execution. In addition to a relatively small and

similar population of astronauts, and the challenges associated with collection and return of in-flight samples, contributing factors such as microgravity, space radiation exposure, psychological stress, nutrition, and exercise, cannot be isolated and tested, so definitive mechanisms underlying the dramatic shifts in telomere length dynamics observed during spaceflight are difficult, if not impossible to identify. Furthermore, these studies involve high-profile humans whose health and confidentiality must be maintained, living in an extreme environment that cannot be fully replicated on Earth.

Figures

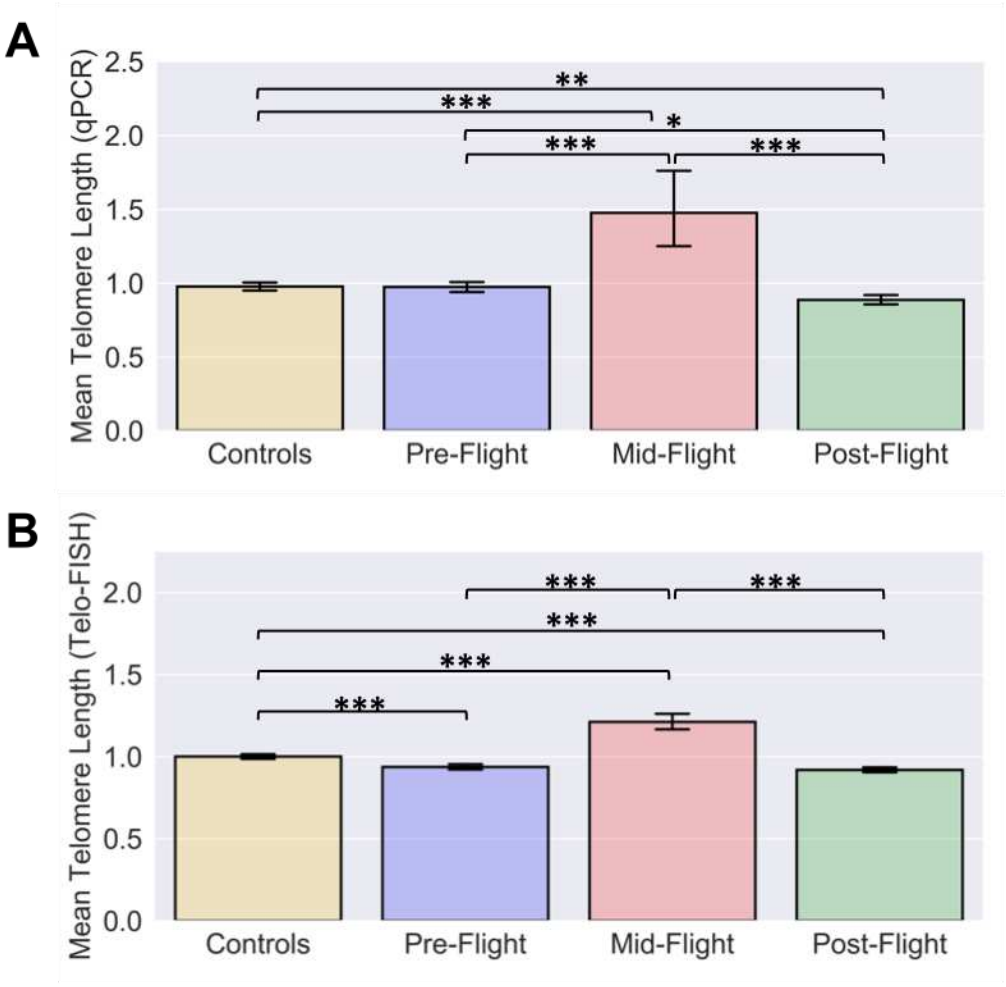

Figure 3.1. Telomere length dynamics in astronauts. Telomere length was assessed by two independent assays in a cohort of unrelated astronauts (n=11) before, during and after spaceflight (one-year and shorter duration missions), and standardized to healthy age and sex-matched ground control subjects (n=11). **A.** Relative mean telomere length determined by qPCR for peripheral blood mononuclear cells (PBMCs). Three technical replicates were performed per individual/time point. N = total number of measurements (including replicates); controls n=233; pre-flight n=95; mid-flight n=18; post-flight n=127. **B.** Relative mean telomere length determined by Telo-FISH on metaphase chromosomes (stimulated T-cells). N = total number of cells analyzed; controls n=2010; pre-flight n=957; mid-flight n=180; post-flight n=1206. A/B. Significance was assessed by one-way ANOVA with a post-hoc Tukey's HSD test. Error bars are 95% confidence intervals of the mean. $p<0.05$*, $p<0.01$**, $p<.001$***.

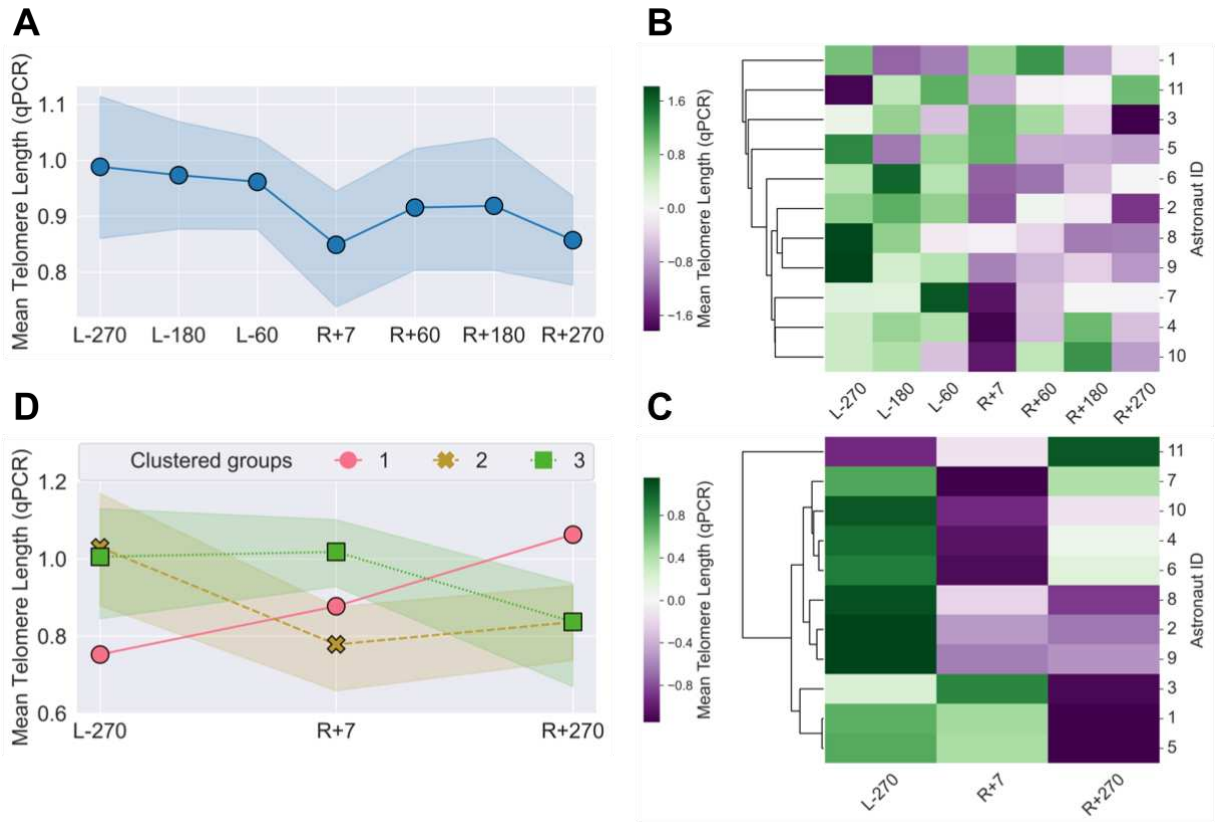

Figure 3.2. Longitudinal clustering of astronaut telomere length before and after spaceflight (qPCR). Longitudinal analyses of telomere length (PBMCs) for cohort of unrelated astronauts (n=11) at all timepoints pre- (Launch-) and post- (Return+) spaceflight (one-year and shorter duration missions), standardized to healthy age and sex-matched ground control cohort (n=11). **A.** Relative mean telomere length over time. Dark line represents the median, lighter bands the 95% confidence interval. Three technical replicates were performed per individual/time point. **B.** Longitudinal hierarchical clustering of relative mean telomere length at all pre- (L-) and post-(R+) spaceflight timepoints, and **C.** the earliest/latest (L-270/R+270) and immediately post spaceflight (R+7) timepoints. **D.** Relative mean telomere length for the three clustered groups identified in C.

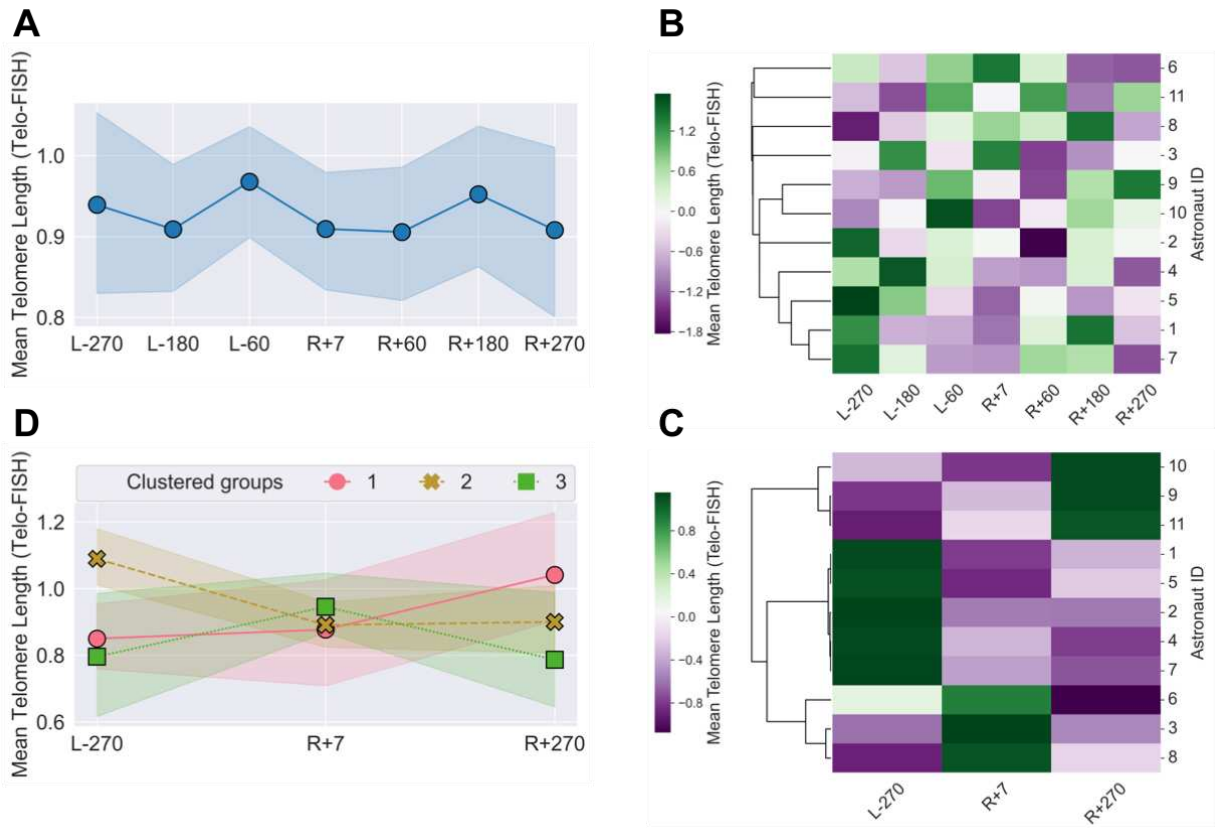

Figure 3.3. Longitudinal clustering of astronaut telomere length before and after spaceflight (Telo-FISH). Longitudinal analyses of telomere length on metaphase chromosomes (stimulated T-cells) for cohort of unrelated astronauts (n=11) at all timepoints pre- (L-) and post- (R+) spaceflight (one-year and shorter duration missions), standardized to healthy age and sex-matched ground control cohort (n=11). **A.** Relative mean telomere length over time. Dark line represents the median, lighter bands the 95% confidence interval. Thirty cells per astronaut per timepoint were analyzed. **B**. Longitudinal hierarchical clustering of relative mean telomere length at all pre- (L-) and post (R+) -spaceflight timepoints, and **C**. at the earliest/latest (L-270/R+270) and immediately post spaceflight (R+7) timepoints. **D.** Relative mean telomere length for the three clustered groups identified in C.

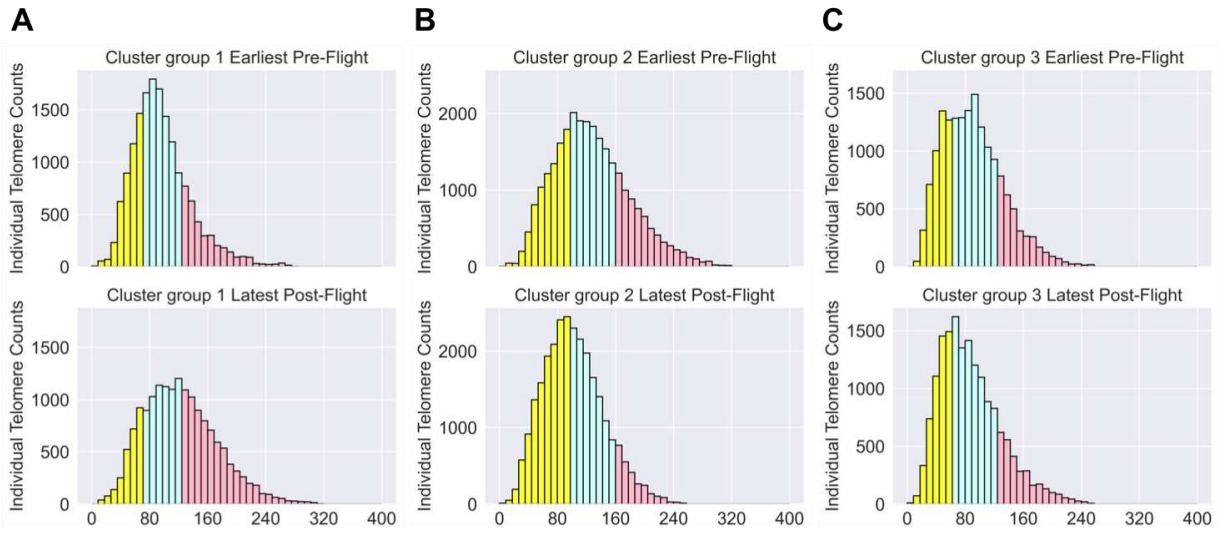

Figure 3.4. Telomere length distributions for clustered groups of astronauts before and after spaceflight. Individual telomere length distributions at earliest pre- (L-270) and final post- (R+270) spaceflight timepoints for the three groups of astronauts identified by hierarchical clustering of mean telomeric length (Telo-FISH). All measurements are standardized to the healthy age and sex-matched ground control cohort (n=11). RFI: relative fluorescent intensity. Short, medium, and long telomeres are yellow, blue, and red, respectively. N = total number of individual telomeres measured. **A.** Clustered group 1: astronauts #9, 10, 11; N=33120. **B.** Clustered group 2: astronauts #1, 2, 4, 5, 7; N=55200. **C.** Clustered group 3: astronauts #3, 6, 8; N=33120.

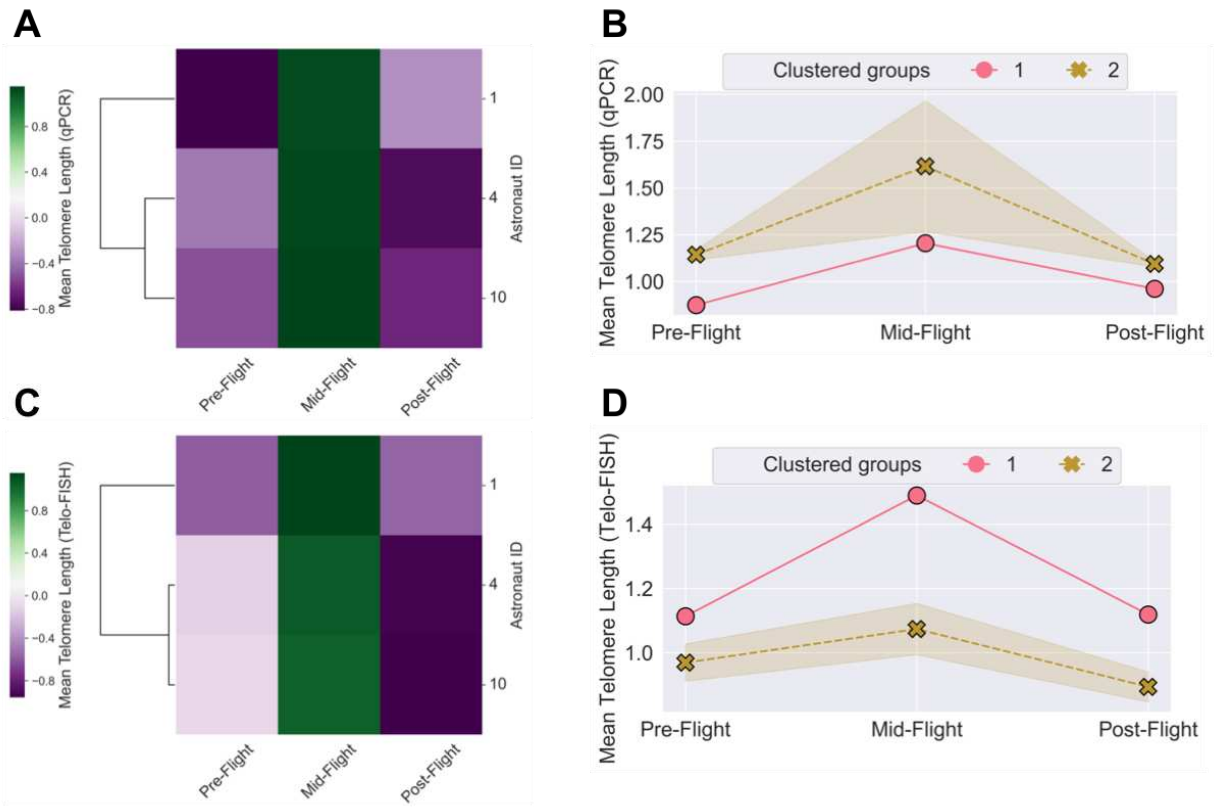

Figure 3.5. Longitudinal clustering of telomere length *during* spaceflight. Telomere length measurement by qPCR (PBMCs) and Telo-FISH (stimulated T-cells) for three unrelated astronauts on one-year or 6-month missions, pre-, mid-, and post-spaceflight, and standardized to healthy age and sex-matched ground control cohort (n=11). **A/C.** Longitudinal hierarchical clustering of mean telomere length before, during and after spaceflight for each assay identified two groups of similar responders. **B/D.** Mean telomere length of the two clustered groups. Clustered group 1: astronaut (n=1); 6-month mission. Clustered group 2: astronauts (n=2); one-year and 6-month missions.

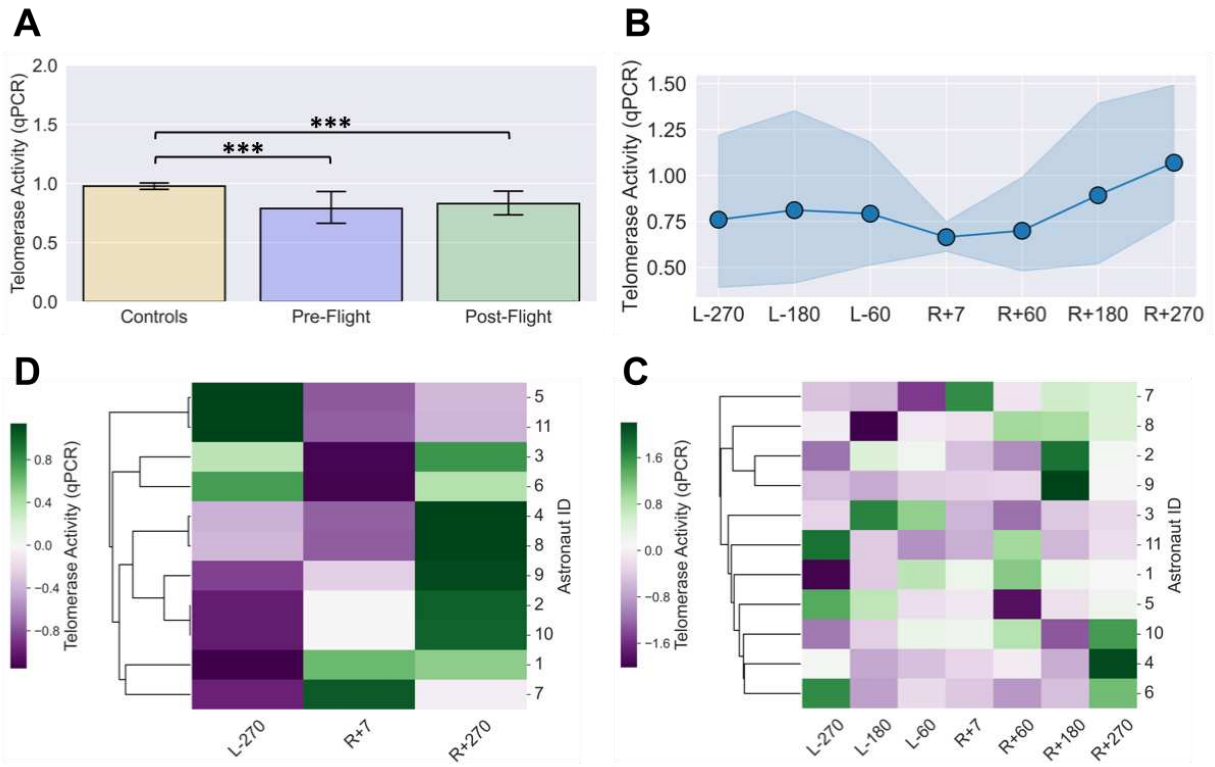

Figure 3.6. Longitudinal clustering of telomerase activity before and after spaceflight. Telomerase activity for the cohort of unrelated astronauts (n=11) at all timepoints pre- and post-spaceflight (one-year and shorter duration missions), standardized to healthy age and sex-matched ground control subjects (n=11). L- and R+ refer to days pre- and post-spaceflight, respectively. It was not possible to directly assess telomerase activity in samples collected onboard the ISS. **A**. Telomerase activity determined by qPCR-TRAP in PBMCs overall, and **B.** over time. Dark line represents the median, lighter bands the 95% confidence interval. Three technical replicates performed per individual/time point. **C**. Longitudinal hierarchical clustering of telomerase activity at all pre- and post-spaceflight timepoints, and **D**. the earliest/latest (L-270/R+270) and immediately post-spaceflight (R+7) timepoints. Significance was assessed by a one-way ANOVA with a post-hoc Tukey's HSD test. Error bars are 95% confidence intervals of the mean. $p<0.05$*, $p<0.01$**, $p<.001$***.

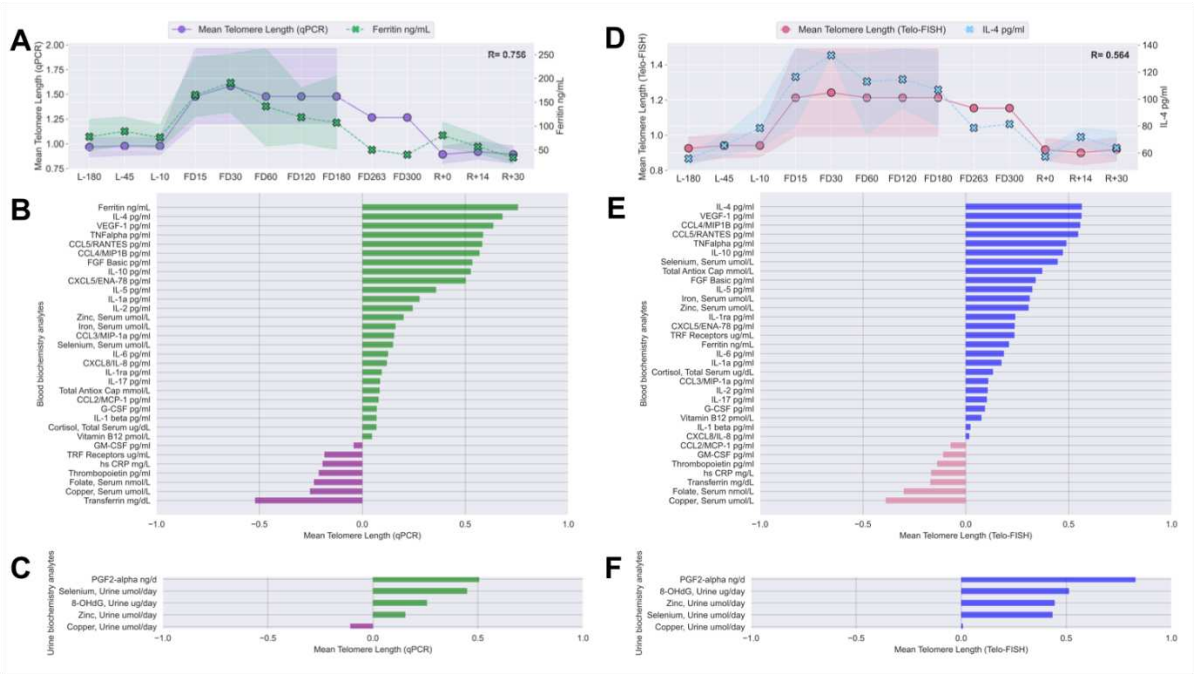

Figure 3.7. Telomere length dynamics and biochemical analytes in blood and urine. Mean telomere length (determined by qPCR, left; Telo-FISH, right) for entire astronaut cohort (n=11), correlated with biochemical analytes measured before, during, and after spaceflight (one-year and shorter duration missions). Correlations were performed on a per astronaut basis, between each astronaut's mean telomere length at pre-, mid-, and post-spaceflight timepoints, and that astronaut's respective biochemical measurements (when available). L-, FD, and R+ refer to days before, during, and after spaceflight, respectively. **A/D.** Full time-course with longitudinal correlations between mean telomere length and the analyte with the strongest positive correlation. **B/E**. Diverging bars plots for correlations between mean telomere length and all biochemical analytes in blood, or **C/F** urine**.**

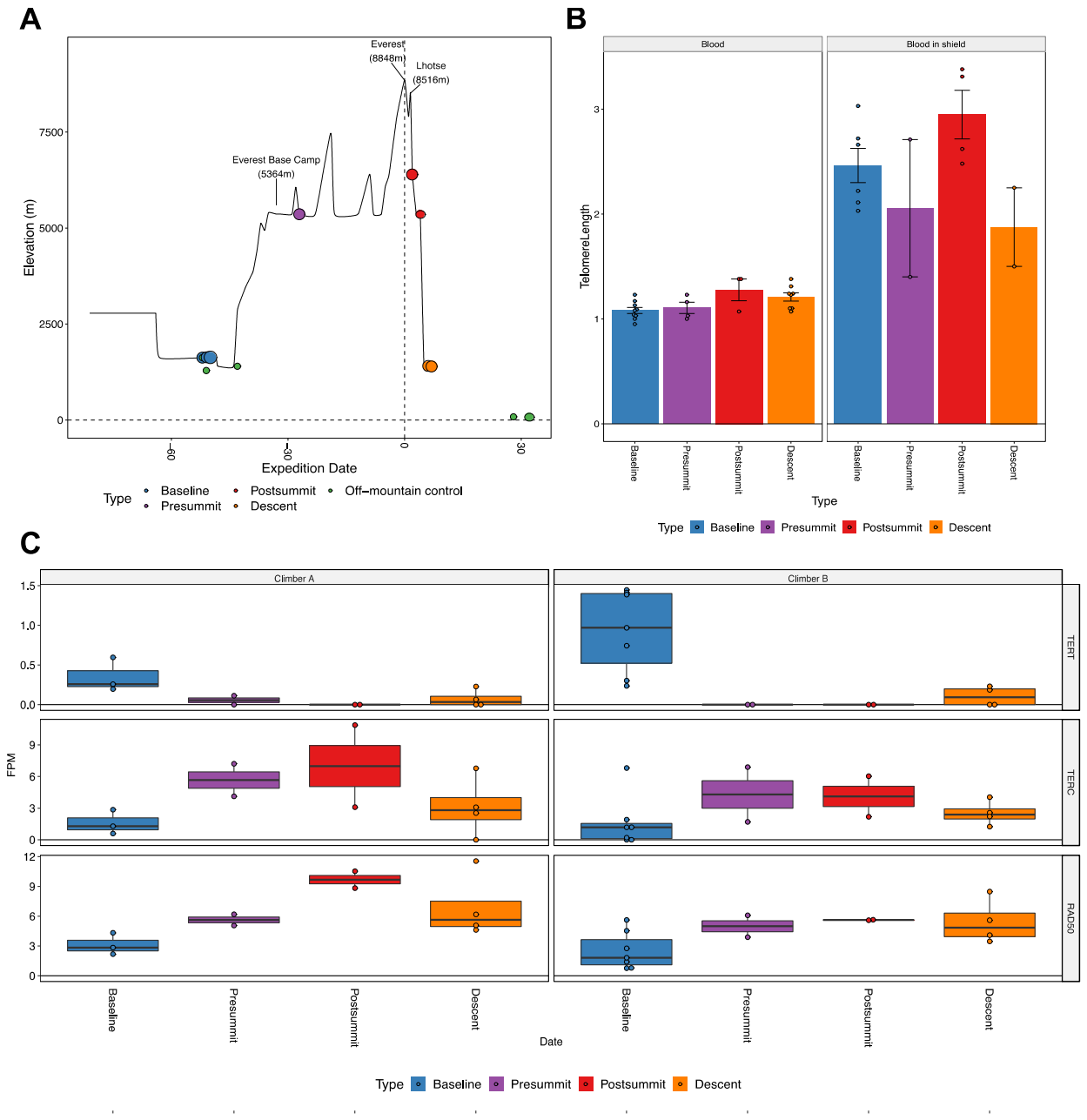

Figure 3.8. Mt. Everest telomere length and gene expression. **A**. Schematic of elevation and location of blood sample collection on/off Mt. Everest/Mt. Lhotse: for the climbers (n=2) at baseline, pre-summit (at base camp), post-summit, and descent (to Katmandu); and for the off-mountain twin controls (n=2) at baseline. **B**. Telomere length assessment (qPCR) for the climbers in blood, and in blood with shield (preservative) show similar trends, including significantly longer telomere length post-summit vs. baseline. **C**. RNA-seq/gene expression data for the climbers show similar patterns of decreased hTERT, increased hTERC, and increased RAD50 pre-summit and post-summit.

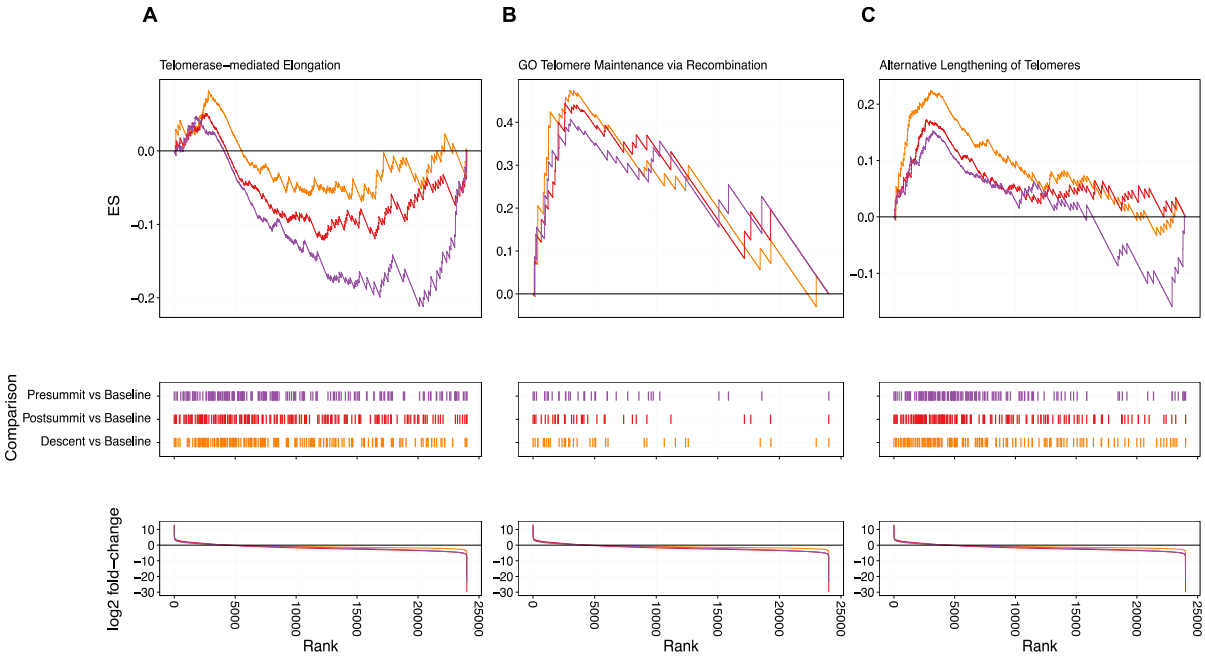

Figure 3.9. Mt. Everest pathway analyses. Pathway analyses provided evidence supportive of activation of the ALT pathway of telomere maintenance associated with this extreme environment. **A**. Telomerase-mediated telomere elongation pathway was not involved. In contrast, **B.** telomere maintenance via recombination, and **C.** Alternative Lengthening of Telomeres (ALT) pathways are significantly involved (Table 3.3).

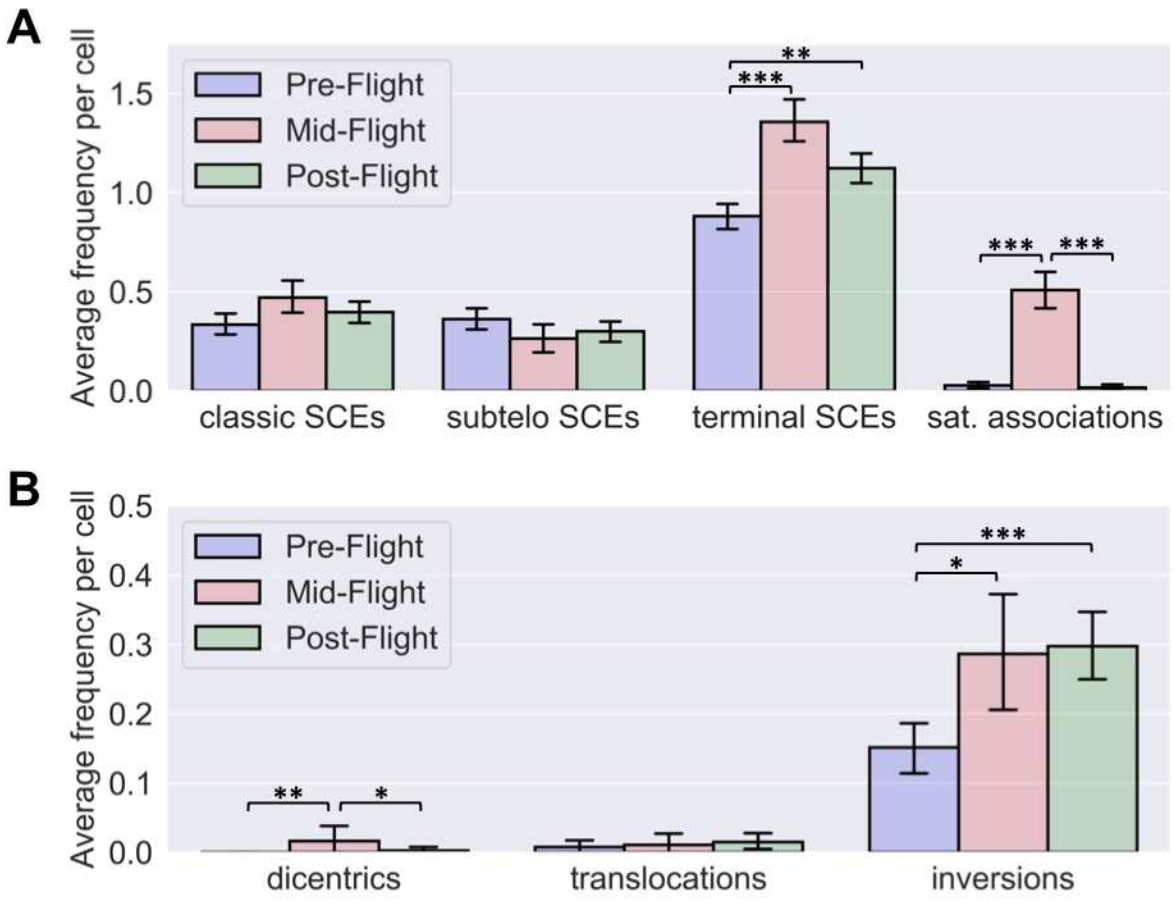

Figure 3.10. Cytogenetic signatures of DNA damage responses. Cytogenetic analyses for the astronaut cohort (n=11) before, during, and after spaceflight, utilizing directional Genomic Hybridization (dGH); n=30 cells per individual per timepoint. **A.** Oxidative damage (replication stress)-induced sister chromatid exchange (SCE): classic SCE along length of chromosome; subtelo and terminal SCE, partial or complete exchange of the subtelomere between sister chromatids, respectively. Satellite (sat) associations between short (p)-arms of acrocentric chromosomes. **B.** Ionizing radiation (double-strand break)-induced chromosome rearrangements. Significance was assessed by a one-way ANOVA per aberration with a post-hoc Tukey's HSD test. Error bars are SEM. $p<0.05$*, $p<0.01$**, $p<.001$***.

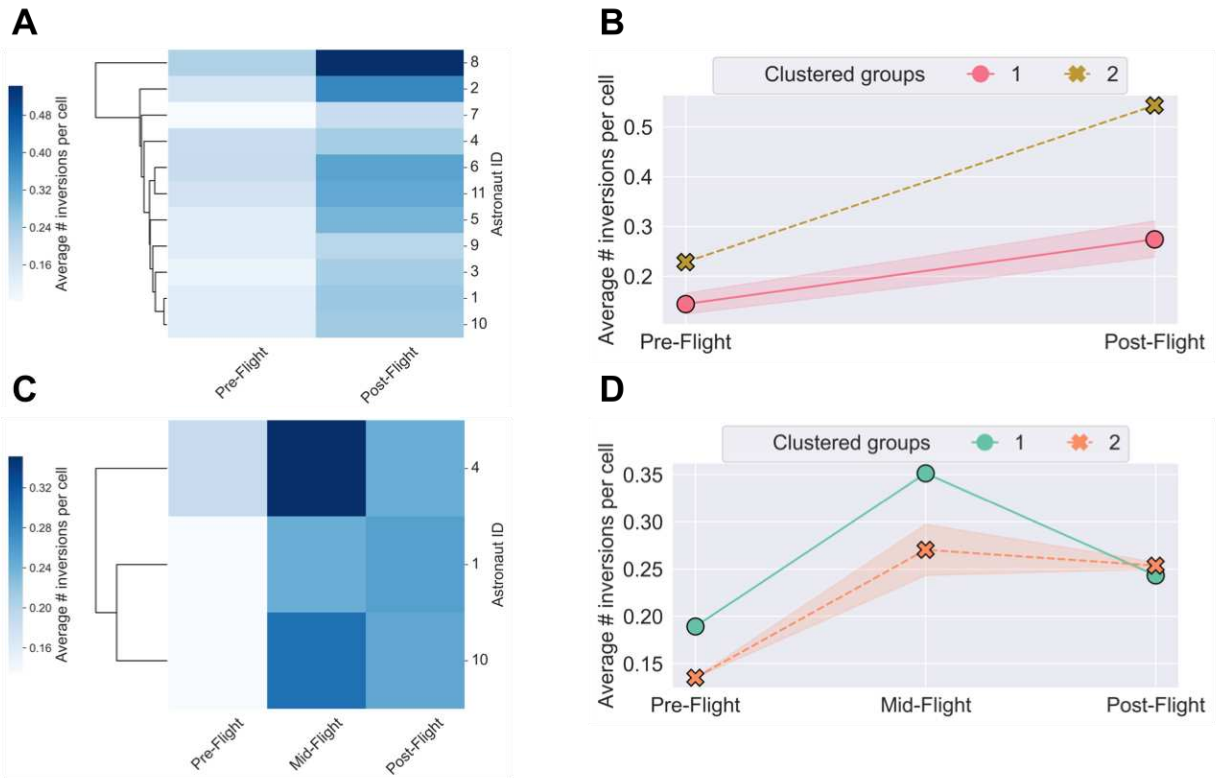

Figure 3.11. Longitudinal clustering of inversions for the astronauts. A. Hierarchical longitudinal clustering of inversions before and after spaceflight for the astronaut cohort (n=11), and **C.** before, during, and after spaceflight for the astronauts with mid-flight data (n=3). **B/D.** Inversion frequencies from the two clustered groups derived from A/C, respectively. For **C**. Clustered group 1: n=10; group 2: n=1. For **D**: Clustered group 1: one-year mission (n=1); group 2: 6-month missions (n=2).

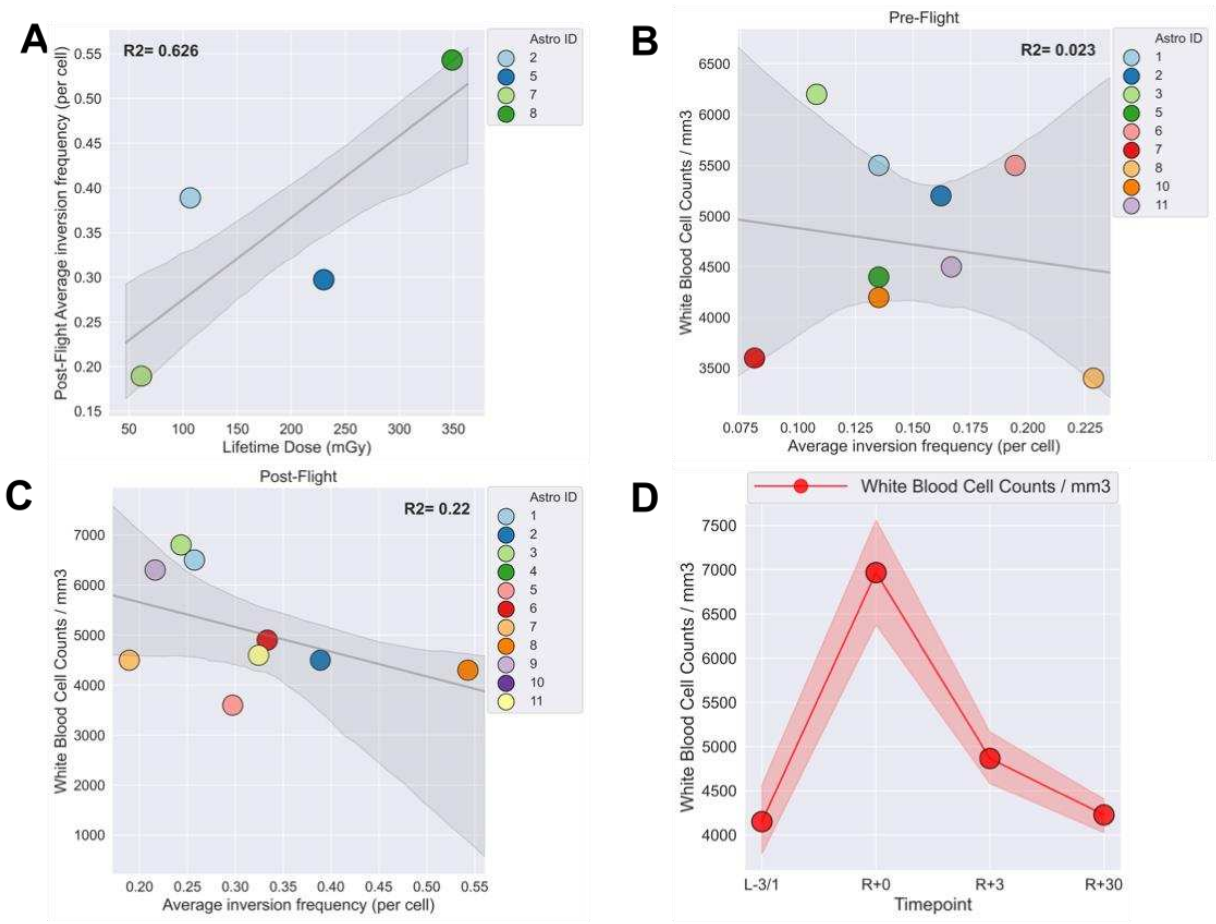

Figure 3.12. Radiation dose estimates, inversions, and white blood cell counts. A. Simple linear regression of post-flight (R+60 or R+180) inversion frequencies versus lifetime dose estimates (mGy), where available. **B.** Simple linear regression of inversion frequencies pre-flight (L-60 or L-180) versus white blood cell (WBC) counts pre-flight (L-3 or L-1). **C.** Simple linear regression of inversion frequencies post-flight (R+60 or R+180) versus WBC counts post-flight (R+30). R2 coefficients of determination are noted (bold); dark lines denote the linear regression, shaded areas the 95% confidence interval. **D.** White blood cell counts pre-flight (L-3 or L-1), immediately upon return to earth (R+0), return plus 3 days (R+3), and return plus 30 days (R+30); center lines are medians, shaded areas the 95% confidence interval.

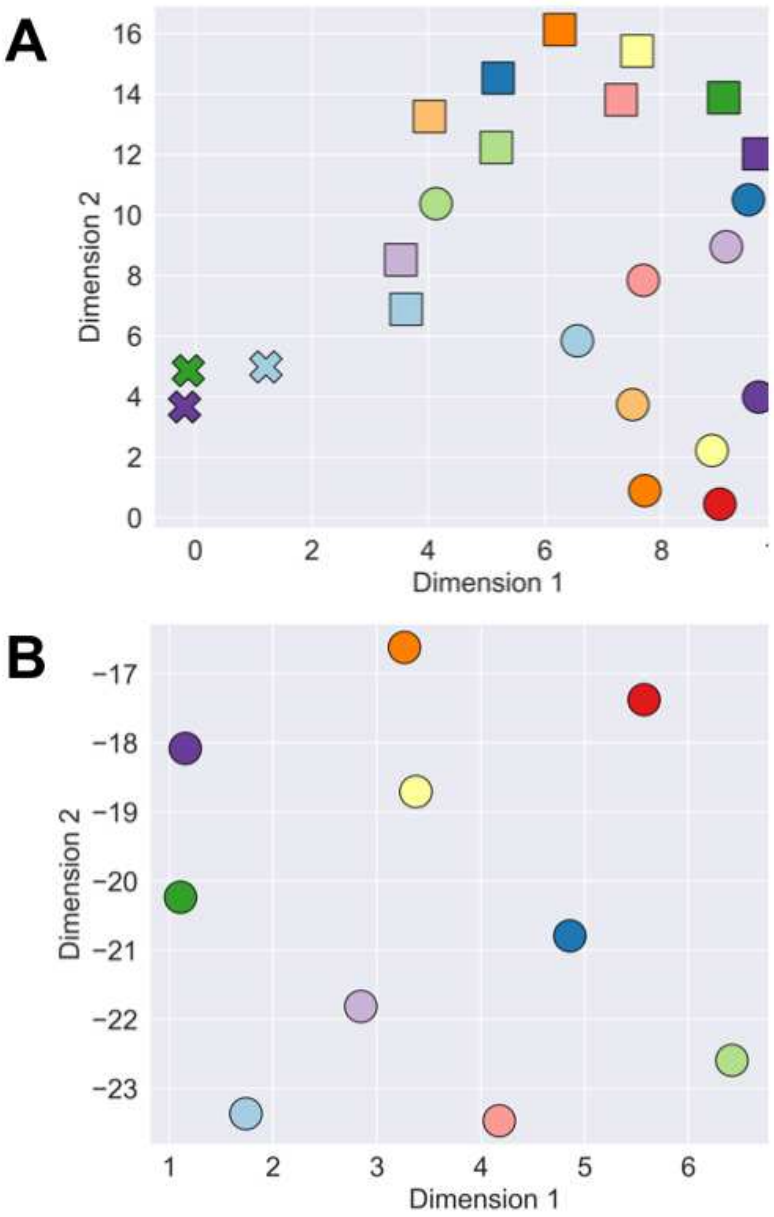

Figure 3.13. UMAP dimensionality reduction of telomeric and DNA damage responses to spaceflight. Uniform Manifold Approximation and Projection (UMAP) was applied to all telomere length (qPCR, Telo-FISH) and cytogenetic data (inversions, translocations, dicentrics, terminal SCEs, satellite associations). **A.** Each point represents the dimensionality reduction of all data for an individual astronaut at specific timepoints (pre-, mid-, or post-spaceflight), or **B.** for all timepoints for that astronaut. Astronaut IDs and timepoint information were stripped prior to UMAP dimensionality reduction and reapplied for graphing purposes. Dimensions (X/Y; 1, 2) are arbitrary units, though distance between points can be interpreted literally.

Tables

Table 3.1. Biochemical analytes for astronaut cohort (blood). Biochemical analyte concentrations before, during, and after spaceflight were evaluated using linear mixed models; each individual astronaut (n=11) was modeled with a random intercept. N represents the number of astronauts per flight status; p-values were generated with respect to pre-flight and rounded to four digits (where p<0.0000 is represented as 0.0).

Analyte	Flight Status	N	Mean	SD	P value	Analyte	Flight Status	N	Mean	SD	P value
CCL2/MCP-1 pg/ml	Pre-Flight	8	83.26	17.13		GM-CSF pg/ml	Pre-Flight	8	22.45	9.14	
	Mid-Flight	8	103.75	24.31	0.594		Mid-Flight	8	23.5	3.57	0
	Post-Flight	8	128.7	103.28	0.005		Post-Flight	8	18.47	6.51	0.094
CCL3/MIP-1a pg/ml	Pre-Flight	8	453.99	192.7		IL-1 beta pg/ml	Pre-Flight	8	5.34	2.14	
	Mid-Flight	8	522.83	70.86	0		Mid-Flight	8	6.26	1.19	0
	Post-Flight	8	363.58	105.31	0.075		Post-Flight	8	4.58	1.58	0.195
CCL4/MIP1B pg/ml	Pre-Flight	8	52.93	17.23		IL-10 pg/ml	Pre-Flight	8	3.9	1.43	
	Mid-Flight	8	90.86	34.24	0		Mid-Flight	8	5.7	1.65	0
	Post-Flight	8	49.54	15.11	0.301		Post-Flight	8	3.99	1.52	0.663
CCL5/RANTES pg/ml	Pre-Flight	8	7368.89	3814.92		IL-17 pg/ml	Pre-Flight	8	12.36	2.66	
	Mid-Flight	8	16144.16	2415.01	0		Mid-Flight	8	12.12	1.65	0
	Post-Flight	8	6000.11	4046.44	0.121		Post-Flight	8	11.38	2.88	0.013
CXCL5/ENA-78 pg/ml	Pre-Flight	8	294.25	569.77		IL-1a pg/ml	Pre-Flight	8	18.19	12.95	
	Mid-Flight	8	928.29	446.11	0		Mid-Flight	8	25.1	7.32	0
	Post-Flight	8	248.37	382.5	0.681		Post-Flight	8	12.1	5.89	0.098
CXCL8/IL-8 pg/ml	Pre-Flight	8	9.32	3.38		IL-1ra pg/ml	Pre-Flight	8	541.26	148.42	
	Mid-Flight	8	11.58	3.87	0		Mid-Flight	8	858.78	366.67	0
	Post-Flight	8	9.19	2.88	0.916		Post-Flight	8	657.2	356.65	0.056
Copper, Serum umol/L	Pre-Flight	11	15.86	5.22		IL-2 pg/ml	Pre-Flight	8	47.1	28.9	
	Mid-Flight	11	12.82	1.5	0.97		Mid-Flight	8	59.25	14.96	0
	Post-Flight	11	14.02	2.95	0.009		Post-Flight	8	35.28	10.27	0.117
Cortisol, Total Serum ug/dL	Pre-Flight	11	11.49	4.77		IL-4 pg/ml	Pre-Flight	8	67.61	18.31	
	Mid-Flight	11	13.52	4.59	0		Mid-Flight	8	111.16	27.62	0
	Post-Flight	11	10.29	4.46	0.241		Post-Flight	8	62.39	13.77	0.089
FGF Basic pg/ml	Pre-Flight	8	94.09	31.55		IL-5 pg/ml	Pre-Flight	8	2.37	0.98	
	Mid-Flight	8	135.52	39.53	0		Mid-Flight	8	2.97	0.6	0.247
	Post-Flight	8	80.02	19.35	0.04		Post-Flight	8	2.53	0.95	0.911
Ferritin ng/mL	Pre-Flight	11	81.02	49.74		IL-6 pg/ml	Pre-Flight	8	6.38	1.57	
	Mid-Flight	11	129.19	77.29	0		Mid-Flight	8	6.93	0.32	0.977
	Post-Flight	11	57.04	42.04	0.004		Post-Flight	8	6.93	2.35	0.274
Folate, Serum nmol/L	Pre-Flight	11	25.1	8.85		G-CSF pg/ml	Pre-Flight	8	13.94	3.85	
	Mid-Flight	11	19.84	6.24	0		Mid-Flight	8	16.17	2.59	0
	Post-Flight	11	22.48	12.85	0.044		Post-Flight	8	14.76	4.69	0.393

Table 3.2. Biochemical analytes for astronaut cohort (blood, urine). Biochemical analyte concentrations before, during, and after spaceflight in blood and urine (source), were evaluated using linear mixed models; each individual astronaut (n=11) was modeled with a random intercept. N represents the number of astronauts per flight status; pvalues were generated with respect to pre-flight and rounded to four digits (where p<0.0000 is represented as 0.0).

Source	Analyte	Flight Status	N	Mean	SD	P value	Source	Analyte	Flight Status	N	Mean	SD	P value
Blood	TRF Receptors ug/mL	Pre-Flight	11	4.12	1.24		Blood	Selenium, Serum umol/L	Pre-Flight	11	1.65	0.3	
Blood		Mid-Flight	11	4.37	0.96	0.896	Blood		Mid-Flight	11	1.78	0.2	0.987
Blood		Post-Flight	11	4.87	1.42	0	Blood		Post-Flight	11	1.66	0.22	0.925
Blood	Thrombopoietin pg/ml	Pre-Flight	8	603.94	313.74		Blood	TNFalpha pg/ml	Pre-Flight	8	16.85	5.98	
Blood		Mid-Flight	8	578.73	201.11	0	Blood		Mid-Flight	8	29.7	10.55	0
Blood		Post-Flight	8	496.4	178.9	0.254	Blood		Post-Flight	8	14.12	4.14	0.113
Blood	Total Antiox Cap mmol/L	Pre-Flight	11	1.49	0.1		Blood	Iron, Serum umol/L	Pre-Flight	11	22.64	7.56	
Blood		Mid-Flight	11	1.48	0.08	0.272	Blood		Mid-Flight	11	24.96	6.45	0.195
Blood		Post-Flight	11	1.43	0.1	0.003	Blood		Post-Flight	11	18.75	7.71	0.042
Blood	Transferrin mg/dL	Pre-Flight	11	247.77	41.06		Urine	8-OHdG, Urine ug/day	Pre-Flight	11	4.97	2.08	
Blood		Mid-Flight	11	221.19	35.05	0.275	Urine		Mid-Flight	11	7.39	3.57	0
Blood		Post-Flight	11	240.35	43.52	0.185	Urine		Post-Flight	11	4.62	2.11	0.37
Blood	VEGF-1 pg/ml	Pre-Flight	8	199.91	203.4		Urine	Copper, Urine umol/day	Pre-Flight	11	0.24	0.08	
Blood		Mid-Flight	8	660.63	383.69	0	Urine		Mid-Flight	11	0.25	0.06	0.862
Blood		Post-Flight	8	139.35	131.52	0.078	Urine		Post-Flight	11	0.31	0.22	0.049
Blood	Vitamin B12 pmol/L	Pre-Flight	11	278.3	113.03		Urine	PGF2-alpha ng/d	Pre-Flight	11	3559.09	2601.43	
Blood		Mid-Flight	11	304.25	44.8	0	Urine		Mid-Flight	11	9957.42	6053.34	0
Blood		Post-Flight	11	292.45	120.87	0.469	Urine		Post-Flight	11	4012.34	2894.56	0.535
Blood	Zinc, Serum umol/L	Pre-Flight	11	15.65	3.76		Urine	Selenium, Urine umol/day	Pre-Flight	11	1.09	0.39	
Blood		Mid-Flight	11	15.96	1.58	0.421	Urine		Mid-Flight	11	1.44	0.39	0.807
Blood		Post-Flight	11	13.38	2.51	0.001	Urine		Post-Flight	11	1.05	0.33	0.696
Blood	hs CRP mg/L	Pre-Flight	11	0.99	1.06		Urine	Zinc, Urine umol/day	Pre-Flight	11	7.55	2.37	
Blood		Mid-Flight	11	0.64	0.5	0.376	Urine		Mid-Flight	11	11.99	7.69	0.148
Blood		Post-Flight	11	1.39	2.08	0.131	Urine		Post-Flight	11	9.91	5.8	0.06

Table 3.3. Mt. Everest Twins Study statistical analyses. RNA-seq, pathway analysis, and telomere length measurements were performed on climbers (n=2) at multiple timepoints (baseline, pre-climb, post-climb, descent). Shown are the statistical significance results for the comparisons made (shown in Figures 3.8, 3.9): RNA seq/gene expression (RAD50, TERC, TERT); pathway enrichment (annotations shown); and telomere length (qPCR). NES: normalized enrichment score. FDR: false discovery rate.

Experiment	Comparison	Gene	log2FoldChange	p-value adj
RNA Seq	Preclimb_vs_Baseline	RAD50	0.9840	0.0459
RNA Seq	Postclimb_vs_Baseline	RAD50	1.5564	0.0012
RNA Seq	Descent_vs_Baseline	RAD50	1.2381	0.0028
RNA Seq	Preclimb_vs_Baseline	TERC	1.5903	0.0840
RNA Seq	Descent_vs_Baseline	TERC	0.8247	0.2822
RNA Seq	Postclimb_vs_Baseline	TERC	1.7970	0.0430
RNA Seq	Descent_vs_Baseline	TERT	-2.1765	0.0066
RNA Seq	Postclimb_vs_Baseline	TERT	-4.7582	0.0004
RNA Seq	Preclimb_vs_Baseline	TERT	-4.4508	0.0014
Experiment	**Comparison**	**Annotation**	**NES**	**FDR.q.val**
Pathway analysis	Descent_vs_Baseline	GO_TELOMERE_MAINTENANCE_VIA_RECOMBINATION	2.3737	0.0033
Pathway analysis	Postclimb_vs_Baseline	GO_TELOMERE_MAINTENANCE_VIA_RECOMBINATION	2.5891	0.0000
Pathway analysis	Preclimb_vs_Baseline	GO_TELOMERE_MAINTENANCE_VIA_RECOMBINATION	2.2714	0.0098
Pathway analysis	Descent_vs_Baseline	TELNET_ALT	1.6011	0.0135
Pathway analysis	Postclimb_vs_Baseline	TELNET_ALT	1.1073	0.0000
Pathway analysis	Preclimb_vs_Baseline	TELNET_ALT	-0.5458	0.9996
Pathway analysis	Descent_vs_Baseline	TELNET_TELOMERASEMEDIATED	0.6035	1.0000
Pathway analysis	Postclimb_vs_Baseline	TELNET_TELOMERASEMEDIATED	-0.5129	0.9993
Pathway analysis	Preclimb_vs_Baseline	TELNET_TELOMERASEMEDIATED	-0.7327	1.0000
Experiment	**Comparison**		**Coefficient**	**p-value**
Telomere length (qPCR)	Presummit_vs_Baseline		-0.0661	0.6650
Telomere length (qPCR)	Postsummit_vs_Baseline		0.4349	0.0045
Telomere length (qPCR)	Descent_vs_Baseline		0.0112	0.9320

STAR Methods

RESOURCE AVAILABILITY

Lead Contact

Further information on methods and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Dr. Susan M. Bailey (susan.bailey@colostate.edu).

Materials Availability

This study did not generate new unique reagents.

Data and Code Availability

The NASA Life Sciences Data Archive (LSDA) is the repository for all human and animal research data, including that associated with this study. LSDA has a public facing portal where data requests can be initiated (https://lsda.jsc.nasa.gov/Request/dataRequestFAQ). The LSDA team provides the appropriate processes, tools, and secure infrastructure for archival of experimental data and dissemination while complying with applicable rules, regulations, policies, and procedures governing the management and archival of sensitive data and information. The LSDA team enables data and information dissemination to the public or to authorized personnel either by providing public access to information or via an approved request process for information and data from the LSDA in accordance with NASA Human Research Program and Johnson Space Center (JSC) Institutional Review Board direction. See Perspective piece (Afshinneko E., et al *Cell, 2020*) for additional information.

The code necessary to recreate all data processing, statistical analyses and figures will be made available at: https://github.com/Jared-Luxton/NASA-astronauts-telomeres-chromosomes.

For the Mt. Everest Study, data will be made available in dbGAP, StudyID: phs002258.v1.p1.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human subjects

Eleven International Space Station (ISS) astronauts participated in this study; one one-year mission crewmember, and ten six-month mission crewmembers. The ground control cohort consisted of eleven healthy age and sex-matched NASA volunteers (Luxton et al, *Cell 2020*). Institutional review board approval was obtained from the Committee for the Protection of Human Subjects at JSC, Houston, TX, USA. Informed consent was obtained from all subjects who participated in the study. In order to protect personally identifiable information for this small and select group, sex and age are not reported.

Two male climbers (ages 20 and 49 yr) participating in a climbing expedition of Mt. Everest and Mt. Lhotse, and their non-climbing twins, also contributed to the study. Institutional review board approval was obtained from the Committee for the Protection of Human Subjects at Weill Cornell Medicine, New York, NY. Informed consent was obtained from all participants.

Collection of biological samples on/off the International Space Station

Samples were collected and processed as pioneered and previously described for the NASA Twins Study (Garrett-Bakelman et al., 2019). Briefly, peripheral blood was drawn from crewmembers (and ground control subjects), before (~270, 180, and 60 days prior to Launch),

during (Flight Day ~45 or 90, and 140 or 260), and after (~1-7, 60, 180, and 270 days after Return) spaceflight, either on Earth or on the ISS. Samples were kept under ambient (room temperature) conditions for ~24hr during shipment to research laboratory for processing, or for in-flight samples, ~48hr during return from space via Soyuz vehicle to Russia, and then jet aircraft to JSC for processing on-site. Urine samples were collected before (180 and 45 days prior to Launch), during (Flight Day 15, 240, and 330), and after spaceflight, (~1-7, and 60 days after Return), either on Earth or on the ISS, and promptly frozen at -80ºC. Upon arrival at JSC, samples were thawed, pooled, and aliquoted, frozen and shipped to research laboratory for processing. For biochemical profiles, crewmember blood and urine samples were collected as previously described (Smith et al., 2012; Zwart et al., 2011). Briefly, samples were collected before (~80 and 45 days prior to Launch), during (Flight Day 15, 30, 60, 120, and 180), and after spaceflight, in the first 24 hr after landing (Return+0), and again 30 days later (Return+30).

Collection of blood samples on/off Mt. Everest

Peripheral blood was drawn and collected from two male climbers, and their non-climbing twins, at baseline (before the expedition), pre-summit/at Mt. Everest Base Camp (after acclimatization rotation from Everest Camp 1), post-summit of Mt. Everest and Lhotse, and again after descent to Kathmandu. Tubes were stored in climbers' packs for the duration of the climb and immediately centrifuged and frozen at -80°C upon arrival at Kathmandu, and then shipped to USA research laboratories for processing and analyses.

Cell line culture

LY-S cells were cultured to confluency in T-75 tissue culture flasks at 1:9 in Gibco PB-Max Karyotyping Medium supplemented with phytohaemagglutinin A (PHA) (ThermoFisher #12557021) and incubated at 37ºC. HeLa cells (ATTC CCL-2) were cultured similarly, except in alpha-mem medium supplemented with 10% fetal bovine serum. For harvesting, cells were split after reaching confluency and KaryoMax Colcemid (ThermoFisher #15210040) added [0.1 µg/mL final concentration], then incubated for 4 hours. Metaphase chromosome spreads were prepared using standard cytogenetic techniques [e.g., (Robinson et al., 2019)]. Briefly, cells were treated with 75mM hypotonic KCl at 37ºC for 20 minutes, centrifuged (1000rpm; ~270xg) and washed with 5ml 3:1 methanol-acetic acid fixative four times, then a final wash and dropping on slides or storage at -20ºC.

METHOD DETAILS

Collection and Processing of Blood for qPCR assessment of Telomere Length and Telomerase Activity

Peripheral blood was drawn from NASA crewmembers and healthy age and sex-matched ground control subjects, before (~270, 180, and 60 days prior to Launch), during (Flight Day ~45 or 90, and 140 or 260), and after, (~1-7, 60, 180, and 270 days after Return) spaceflight, either on Earth or on the ISS, and collected into one 10 mL ethylenediaminetetraacetic acid (K_2EDTA) tube (Becton, Dickinson, and Co #368589), which was kept under ambient (room temperature) conditions for ~24hr during shipment to research laboratory for processing, or for in-flight samples, ~48hr during return from space to JSC for on-site processing. Upon arrival, cells were transferred into CPT mononuclear cell preparation tubes (Becton, Dickinson, and Co # 362761) and centrifuged at 1700 x g for 20 min per manufacturer's recommendations. Separated

peripheral blood mononuclear cells (PBMCs) were collected and washed in PBS and any remaining red blood cell contamination was removed using ACK lysis buffer. PBMCs (1 x 10^6 cells/mL) were divided for either qRT-PCR based-telomere length measurement or qRT-PCR TRAP assessment of telomerase activity. DNA extraction for qPCR telomere length measurements was performed using the DNeasy Blood and Tissue Kit (Qiagen #69504). PBMCs were incubated in proteinase K for 3 hr at 37°C. DNA was isolated from PBMC populations, and from separated CD4+ and CD8+ T-cells, CD19+ B-cells, and lymphocyte depleted (LD) populations. For telomerase activity, PBMC protein lysates were prepared in mammalian protein extraction buffer (M-PER) (ThermoFisher #78503) containing cOmplete Protease Inhibitor Cocktail (Millipore Sigma #11836170001) and RNasin Ribonuclease Inhibitors (Promega #N2515) at 1 x 10^6 cells per 100 µL. Lysates were cleared by incubating on ice for 10 min, followed by centrifugation at 18,000 rpm for 15 min at 40°C, aliquoted, and stored at -80°C. Protein concentrations were determined using the Bio-Rad Protein Assay Kit (Bradford Assay; Bio-Rad # 5000001).

For the Mt. Everest climbers, blood was self-collected by the subjects at baseline, pre-summit, post-summit, and after decent, into BD Vacutainer CPT Sodium Citrate tubes (BD 362761) and Zymo DNA/RNA Shield blood collection tubes (Zymo R1150). DNA extraction was performed using Macherey-Nagel NucleoSpin Blood, Mini kit for DNA from blood (740951.50).

Telomere Length assessment by Multiplexed Quantitative (q)PCR

Multiplexed qPCR measurements of telomere length were carried out as previously described (Cawthon, 2009). Here, a 22 µL master mix was prepared using SYBR green GoTaq qPCR master mix (Promega #A6001) combined with the telomere forward primer (TelG; 5'-

ACACTAAGGTTTGGGTTTGGGTTTGGGTTTGGGTTAGTGT-3'), telomere reverse primer (TelC; 5'-TGTTAGGTATCCCTATCCCTATCCCTATCCCTATCCCT AACA-3'), albumin forward primer (AlbU; 5'-CGGCGGCGGGCGGCGCGGGCTGGGCGGA AATGCTGCACAGAATCCTTG-3'), albumin reverse primer (AlbD; 5'-GCCCGGCCCGCCG CGCCCGTCCCGCCGGAAAAGCATGGTCGCCTGTT-3') at 10 μM per primer (Integrated DNA Technologies), and RNase/DNase free water. To the master mix, 3 μL of DNA at 3.33 ng/uL was added for a final volume of 25 μL. The TelG/C primers were at a final concentration of 900 nM and the AlbU/D primers at 400 nM. Telomere length measurements were carried out using a Bio-Rad CFX-96 qPCR machine. The cycle design was as follows: 95°C for 3 min; 94°C for 15 sec, 49°C for 15 sec, for 2 cycles; 94°C for 15 sec, 62°C for 10 sec, 74°C for 15 sec, 84°C for 10 sec, and 88°C for 15 sec, for 32 cycles. The melting curve was established by a 72°C to 95°C ramp at 0.5°C/sec increase with a 30 sec hold. Multiplexing both telomere and albumin primers using a single fluorescent DNA-intercalating dye is possible because the telomere sequences are amplified at a lower quantification cycle (Cq) than the albumin sequences. Standard curves were prepared using human genomic DNA (Promega cat # G3041) with 3-fold dilutions ranging from 50 ng to 0.617 ng in 3 μL per dilution. Negative controls included a no-template TelG/C only and AlbU/D only, and a combined TelG/C and AlbU/D control. Samples were normalized across plates using a human genomic DNA standard. Each sample was run in triplicate on a 96-well plate format and relative telomere length was established using a telomere (T) to albumin (A) ratio.

Telomerase Activity assessment by qPCR Telomere Repeat Amplification Protocol (TRAP)

Telomerase activity was monitored using qRT-PCR TRAP as previously described (Hou et al., 2001). Here, protein lysates were suspended at 0.20 μg in 25 μL of SYBR green GoTaq qPCR Master Mix (Promega #A6001) containing 0.20 μg T4 gene 32 protein (New England Biolabs #M0300S), 0.10 μg of primers TS (5'-AATCCGTCGAGCAG AGTT- 3') and ACX (5'-GCGCGG(CTTACC)3CTAACC-3') (Integrated DNA Technologies) and RNase/DNase free water. The qRT PCR reactions took place using a Bio-Rad CFX-96 qPCR machine using the following steps: 25ºC for 20 min; 95ºC for 3 min; 95ºC for 20 sec, 50ºC for 30 sec, 72ºC for 90 sec for 50 cycles; establish a melting curve and prevent primer dimerization for 10 sec per cycle for 80 cycles. Each sample was run in triplicate on a 96-well plate format allowing for an average Cq to be obtained per sample. Relative telomerase activity was then established using the delta delta Ct method (Livak and Schmittgen, 2001). In addition to the test samples, the following controls were included: no template controls with TS only, ACX only, and TS/ACX primers combined, as well as a HeLa cell line (ATTC CCL-2) heat-inactivated negative control, and a HeLa positive control (very high telomerase activity).

Collection and Processing of Blood for Telomere Fluorescence *in situ* Hybridization (Telo-FISH) and directional Genomic Hybridization (dGH)

Peripheral blood was drawn from NASA crewmembers and healthy age and sex-matched ground control subjects, before (270, 180, and 60 days prior to Launch), during (Flight Day 45 or 90, and 140 or 260), and after (~1-7, 60, 180, and 270 days after Return) spaceflight, either on Earth or on the ISS, and collected into one 10 mL sodium heparin tube (Becton, Dickinson, and Co

#367874), which was kept under ambient (room temperature) conditions for ~24hr during shipment to research laboratory, or for in-flight samples, ~48hr during return from space via Soyuz vehicle and jet aircraft to JSC for on-site processing. Upon arrival, heparinized blood was split into two T-25 tissue culture flasks at 1:9 in Gibco PB-Max Karyotyping Medium supplemented with phytohaemagglutinin A (PHA) (ThermoFisher #12557021) and cultured at 37ºC. For directional Genomic Hybridization (dGH), 5.0 mM 5-bromo-deoxyuridine (BrdU) and 1.0 mM 5- bromo-deoxycytidine (BrdC) were added to the medium. Following 44 hr incubation/stimulation, KaryoMax Colcemid (ThermoFisher #15210040) was added [0.1 µg/mL final concentration] for the final 4 hr prior to harvest (at 48 hr). Metaphase chromosome spreads were prepared using standard cytogenetic techniques [e.g., (Robinson et al., 2019)]. Briefly, cells were treated with 75mM hypotonic KCl at 37ºC for 20 minutes, centrifuged (1000rpm; ~270xg) and washed with 5ml 3:1 methanol-acetic acid fixative four times, then a final wash and dropping on slides or storage at -20ºC.

Telomere Length assessment by Telo-FISH

Prepared slides with metaphase chromosome spreads (PHA stimulated T-cells) were hybridized with a fluorescently-labeled telomere probe as previously described (Sishc et al., 2015). Briefly, slides were washed in PBS for 5 min, dehydrated through a cold ethanol series (75%, 85% and 100%) for 2 min each, and air dried. Slides were then denatured in 70% formamide/2x sodium chloride and sodium citrate (2xSSC) at 75°C for 2 min, followed by a second cold ethanol series. A hybridization mixture containing a G-rich (TTAGGG3) peptide nucleic acid (PNA) telomere probe directly labeled with Cyanine-3 (Cy3; Biosynthesis) was prepared by diluting 0.5 µL probe [5nM final concentration] in 36 µL of formamide, 12 µL of 0.5 M pH 7.5 Tris-HCl, 2.5 µL of 0.1

M KCl, and 0.6 µl of 0.1 M MgCl2; mixture was incubated at 75°C for 5 min, then 50 µL applied to each slide. Coverslips were sealed with rubber cement and slides incubated at 37 °C for 24 hr. After hybridization, slides were washed in a series of 43.5°C washes for 3 min each: washes one and two: 50% formamide in 2x sodium chloride and sodium citrate (2xSSC); washes three and four: 2x SSC; and washes five and six: 2x SSC plus 0.1% Nonidet P-40. Slides were counterstained with 50 µL DAPI in Prolong Gold Antifade (ThermoFisher #P36931), coverslipped, and stored at 4°C for 24 hr before imaging. Metaphase spreads (30 per sample) were imaged on a Nikon Eclipse Ni-U epifluorescent microscope equipped with an Andor Zyla 5.5 sCMOS camera and SpectraX LED light source. Stacked images (5 images/stack, 0.3 µm step size) were taken of each metaphase spread. Fluorescence intensity was quantified using the ImageJ (http://rsb.info.nih.gov/ij/) plugin, Telometer (http://demarzolab.pathology.jhmi.edu/telometer/index.html), and normalized using the mouse lymphoma control cell line LY-S (Wong and Slijepcevic, 2004) and age-sex matched controls.

Cytogenetic analyses by directional Genomic Hybridization (dGH)

High resolution detection of inversions and translocations on metaphase chromosome spreads (PHA stimulated T-cells) was performed utilizing Directional Genomic Hybridization (dGH) single color (Cy-3) whole chromosome and single color (Cy5) subtelomere 1-3 paints (KromaTiD Inc., Ft. Collins, CO) as previously described (Ray et al., 2014; Ray et al., 2013). Briefly, prepared slides with metaphase chromosomes substituted with bromonucleotides (BrdU/BrdC) were submersed in Hoechst 33258 (Millipore Sigma #B1155) for 15 min, selectively photolysed using a SpectroLinker UV Crosslinker equipped with 365 nm UV bulbs for 35 min, and then the nicked DNA was exonucleolytically degraded with Exonuclease III

(New England Biolabs #M0206L) for 30 min. Hybridization cocktail was applied to the slides, which were coverslipped and sealed using rubber cement, then denatured at 68°C for 3 min. Slides were hybridized overnight at 37°C followed by 5 washes in 2x SSC at 43°C prior to imaging. Metaphase spreads were imaged on a Nikon Eclipse Ni-U epifluorescent microscope equipped with an Andor Zyla 5.5 sCMOS camera and SpectraX LED light source. 30 metaphase spreads were scored per timepoint per individual.

Collection and Processing of Blood and Urine for Biochemical analyses

Urinary 8-hydroxy 2'-deoxyguanosine (8OHdG) and 8-iso-prostaglandin F2-α (PGF2α) and plasma cytokines IL-1β, IL-1a, and IL-1ra data from before, during, and after long-duration spaceflights have previously been reported (Silveira et al., *Cell, 2020*). Subsets of these data have also been previously published (Lee et al., 2020). The data presented here are specifically from the 11 crewmembers in the cohort, whose spaceflight missions ranged in length from approximately 4-12 months. Briefly, blood and urine samples were collected twice before flight: approximately 180 and 45 days before launch. Blood samples were collected in flight, at approximately flight days 15, 30, 60, 120, and 180. Post-flight samples were collected in the first 24 hr after landing (R+0) and again 30 days later (R+30). The R+0 samples were not necessarily fasting, given the time of day and nature of return from flight. Pre and post-flight collections included two consecutive 24 hr urine collections, and inflight collections included one 24 hr urine collection. These collection techniques have been previously described (Smith et al., 2015; Zwart et al., 2013). PGF2α was measured using a commercially available ELISA kit (Oxford Biomedical Research, Rochester Hills, MI) and 8OHdG was measured by LC-MS/MS

(Wu and Ho, 2009), as previously reported (Lee et al., 2020; Zwart et al., 2013). Plasma cytokines were measured as previously described (Crucian et al., 2014).

Collection and Processing of Blood for RNA sequencing

For the Mt. Everest climbers, blood was self-collected by subjects at baseline, pre-summit, post-summit, and after decent, into BD Vacutainer CPT Sodium Citrate tubes (BD 362761) and Zymo DNA/RNA Shield blood collection tubes (Zymo R1150). RNA extraction was performed using the Macherey-Nagel Nucleospin RNA Set for NucleoZOL (Machery-Nagel 740406.10). RNA was treated with Zymo DNAse I (Zymo E1010) after extraction. Library prep of the RNA was performed with the Illumina Truseq Total RNA library prep gold with rRNA reduction chemistry (Illumina 20020599). Libraries were sequenced PE50 on the NovaSEQ 6000 instrument using an S1 flowcell. Fastq files were generated using basespace for downstream analysis. Sequenced results were trimmed using Trim Galore and mapped to Ensembl transcripts using Salmon and aggregated into gene level quantifications (Patro et al., 2017).

QUANTIFICATION AND STATISTICAL ANALYSIS

Power calculations were performed to predetermine minimum sample size for our endpoints. Results indicated that 10 crewmembers were desired; practical limitations associated with collection of samples during spaceflight resulted in 3 crewmembers with in-flight samples. Even so, because of the large number of telomeres and chromosome aberrations scored, determination of statistical significance was possible. Data are generally displayed as mean with SEM except where noted. Sample numbers (n) are provided with each figure.

Mean Telomere Length and Telomerase Activity by qPCR

All samples at each timepoint were run in technical triplicate. Statistical significance was evaluated using a one-way ANOVA with post-hoc Tukey's HSD in Python using statsmodels library.

Telomere Length by Telo-FISH and generation of individual telomere length distributions

Fluorescence intensity quantifications of individual telomeres, output by ImageJ, were stored in Excel files, with one file per astronaut per timepoint. Custom scripts were written in Python with extensive use of the pandas library to extract and parse the individual telomere length measurements into one dataframe object for data visualizations and statistical analyses. For timepoint samples with individual telomere distributions less than the theoretical number of telomeres, given 30 metaphase spreads (5520), individual telomere counts were imputed up to 5520 using random sampling (numpy library) from the observed distribution of telomeres for that subject's timepoint, facilitating visual and statistical comparisons for those distributions. For each astronaut, histograms of individual telomeres were colored by taking the distribution of individual telomeres at the earliest timepoint pre-spaceflight and splitting the data into quartiles, with the bottom 25% of telomeres (shortest) in yellow, the middle 50% in blue, and the top 25% in red (longest). The quartile cutoff values, unique to each astronaut, were then applied to that astronaut's subsequent timepoints, to visualize shifts in individual telomeres into and out of the colored quartiles.

Analysis of Telomere Length distributions

Statistically significant differences between the distributions of individual telomeres were assessed per timepoint for each astronaut using the Mann-Whitney U test in Python using the statsmodels library.

Mean Telomere Length by Telo-FISH

Thirty metaphases (184 telomeres/metaphase) were scored per timepoint per individual and statistics were performed on the mean telomere length for a given timepoint. Statistical significance was evaluated using a one-way ANOVA with post-hoc Tukey's HSD in Python using the Statsmodels library.

Chromosome Aberration analyses by dGH

Thirty metaphases were scored per timepoint per individual and statistics were performed on the average number of aberrations per cell for a given timepoint. Statistical significance was evaluated using a one-way ANOVA with post-hoc Tukey's HSD in Python using the Statsmodels library.

Biochemical/analyte analyses

Statistical significance was evaluated using a one-way ANOVA with post-hoc Tukey's HSD in Python using the Statsmodels library.

Statistical and clustering analyses for telomere length, telomerase activity, and chromosome aberration data

Statistical and clustering analyses were conducted with Python in Jupyter Notebooks (see Code availability). With the Statsmodels library (Seabold, 2010), mean telomere length (qPCR, Telo-FISH), telomerase activity (qPCR-TRAP), and chromosome aberration frequencies (dGH) were analyzed by separate one-way ANOVAs and post-hoc Tukey's HSD test. Hierarchical clustering analyses were performed on z-score normalized data (all except for inversions, which had too few timepoints to be z-score normalized), using the Scipy library with a Pearson correlation metric and single linkage method.

Uniform manifold approximation and projection (UMAP) for telomere and chromosome aberration data

UMAP was utilized for dimensionality reduction due to its capacity to robustly handle sparse matrix data, as well as the intuitive interpretation offered by the graphed results. To analyze data patterns per spaceflight status (pre-, mid-, post-), mean telomere length data (qPCR, Telo-FISH) and cytogenetic data (inversions, translocations, dicentrics, terminal SCEs, sat. associations) were arranged in tidy data format per the flight status column for each astronaut. Astronaut IDs and flight status labels were stripped, UMAP performed, and astronaut IDs and flight status labels reapplied for graphing purposes. To analyze individual variation in astronaut responses for all data, data was pivoted on the flight status column per astronaut, in effect putting all data associated with each astronaut in a respective row. Astronaut IDs were stripped, UMAP performed, and astronaut IDs reapplied for graphing purposes.

Analyte correlations with mean telomere length (qPCR, Telo-FISH)

For each astronaut (n=11) and assay (qPCR, Telo-FISH), mean telomere length pre-flight (3 timepoints), mid-flight (2 timepoints, where available), and post-flight (4 timepoints) was calculated and merged onto corresponding timepoints for biochemical analyte data, where available, using astronaut ID and flight status as the join keys; e.g., pre-flight = L-180, L-45, L-10, etc. Pearson correlations were calculated between the astronaut's mean telomere length (for a given flight status) and biochemical analytes (for a given timepoint; 13 total) for all astronauts. This approach was used to facilitate merging of the telomere length and biochemical data, which were collected at different timepoints, and to reduce the potential for spurious correlations.

Linear mixed modeling of biochemical analytes

The impacts of flight status (pre-, mid-, post-flight) on biochemical analyte concentrations (in blood, urine) were evaluated using linear mixed models. The analyses were done in Python with the mixedlm object from the Statsmodels library. Modeling was performed for each analyte and written as follows: {analyte} ~ C(flight status), Treatment(reference=Pre-Flight), where flight status was treated as a categorical variable and p-values were calculated with respect to Pre-Flight using a Patsy Treatment formula from the Patsy library. Summary statistics were generated using the Researchpy library's summary object.

RNA sequencing and pathway enrichment scores

Samples were grouped into four sample categories; baseline, pre-summit, post-summit and descent (Figure 3.8A). Differential expression was called using DESeq2 against the baseline using all blood samples together in a multivariate model correcting for subject and sample category and using the off-mountain subjects as controls (Love et al., 2014). Pathway enrichment

was calculated using log2 fold-change values for the three comparisons by running GSEA against the MSigDB database and custom signatures from TelNet database consisting of genes positively related to Alternative Lengthening of Telomeres (ALT) or telomerase- mediated telomere elongation pathways (Braun et al., 2018; Liberzon et al., 2011; Subramanian et al., 2005).

REFERENCES

Aubert, G., and Lansdorp, P.M. (2008). Telomeres and aging. Physiol Rev *88*, 557-579.

Azzam, E.I., Jay-Gerin, J.P., and Pain, D. (2012). Ionizing radiation-induced metabolic oxidative stress and prolonged cell injury. Cancer Lett *327*, 48-60.

Berardinelli, F., Antoccia, A., Buonsante, R., Gerardi, S., Cherubini, R., De Nadal, V., Tanzarella, C., and Sgura, A. (2013). The role of telomere length modulation in delayed chromosome instability induced by ionizing radiation in human primary fibroblasts. Environ Mol Mutagen *54*, 172-179.

Blackburn, E.H., and Epel, E.S. (2012). Telomeres and adversity: Too toxic to ignore. Nature *490*, 169-171.

Braun, D.M., Chung, I., Kepper, N., Deeg, K.I., and Rippe, K. (2018). TelNet - a database for human and yeast genes involved in telomere maintenance. BMC Genet *19*, 32.

Bryan, T.M., Englezou, A., Dalla-Pozza, L., Dunham, M.A., and Reddel, R.R. (1997). Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. Nat Med *3*, 1271-1274.

Caradonna, F. (2015). Nucleoplasmic bridges and acrocentric chromosome associations as early markers of exposure to low levels of ionising radiation in occupationally exposed hospital workers. Mutagenesis *30*, 269-275.

Cawthon, R.M. (2009). Telomere length measurement by a novel monochrome multiplex quantitative PCR method. Nucleic Acids Res *37*, e21.

Cherkas, L.F., Hunkin, J.L., Kato, B.S., Richards, J.B., Gardner, J.P., Surdulescu, G.L., Kimura, M., Lu, X., Spector, T.D., and Aviv, A. (2008). The association between physical activity in leisure time and leukocyte telomere length. Arch Intern Med *168*, 154-158.

Cohen, S., Janicki-Deverts, D., Turner, R.B., Casselbrant, M.L., Li-Korotky, H.S., Epel, E.S., and Doyle, W.J. (2013). Association between telomere length and experimentally induced upper respiratory viral infection in healthy adults. Jama *309*, 699-705.

Coluzzi, E., Buonsante, R., Leone, S., Asmar, A.J., Miller, K.L., Cimini, D., and Sgura, A. (2017). Transient ALT activation protects human primary cells from chromosome instability induced by low chronic oxidative stress. Sci Rep *7*, 43309.

Crucian, B., Stowe, R.P., Mehta, S., Quiriarte, H., Pierson, D., and Sams, C. (2015). Alterations in adaptive immunity persist during long-duration spaceflight. NPJ Microgravity *1*, 15013.

Crucian, B.E., Chouker, A., Simpson, R.J., Mehta, S., Marshall, G., Smith, S.M., Zwart, S.R., Heer, M., Ponomarev, S., Whitmire, A.*, et al.* (2018). Immune System Dysregulation During Spaceflight: Potential Countermeasures for Deep Space Exploration Missions. Front Immunol *9*, 1437.

Crucian, B.E., Zwart, S.R., Mehta, S., Uchakin, P., Quiriarte, H.D., Pierson, D., Sams, C.F., and Smith, S.M. (2014). Plasma cytokine concentrations indicate that in vivo hormonal regulation of immunity is altered during long-duration spaceflight. J Interferon Cytokine Res *34*, 778-786.

Cucinotta, F.A. (2014). Space radiation risks for astronauts on multiple International Space Station missions. PLoS One *9*, e96099.

Cucinotta, F.A., Hamada, N., and Little, M.P. (2016). No evidence for an increase in circulatory disease mortality in astronauts following space radiation exposure. Life Sciences in Space Research *10*, 53-56.

de Lange, T. (2009). How telomeres solve the end-protection problem. Science *326*, 948-952.

Delp, M.D., Charvat, J.M., Limoli, C.L., Globus, R.K., and Ghosh, P. (2016). Apollo Lunar Astronauts Show Higher Cardiovascular Disease Mortality: Possible Deep Space Radiation Effects on the Vascular Endothelium. Sci Rep *6*, 29901.

Dunham, M.A., Neumann, A.A., Fasching, C.L., and Reddel, R.R. (2000). Telomere maintenance by recombination in human cells. Nat Genet *26*, 447-450.

Eitan, E., Tamar, A., Yossi, G., Peleg, R., Braiman, A., and Priel, E. (2016). Expression of functional alternative telomerase RNA component gene in mouse brain and in motor neurons cells protects from oxidative stress. Oncotarget *7*, 78297-78309.

Epel, E.S., Blackburn, E.H., Lin, J., Dhabhar, F.S., Adler, N.E., Morrow, J.D., and Cawthon, R.M. (2004). Accelerated telomere shortening in response to life stress. Proceedings of the National Academy of Sciences of the United States of America *101*, 17312-17315.

Gardner, M., Bann, D., Wiley, L., Cooper, R., Hardy, R., Nitsch, D., Martin-Ruiz, C., Shiels, P., Sayer, A.A., Barbieri, M.*, et al.* (2013). Gender and telomere length: Systematic review and meta-analysis. Exp Gerontol.

Garrett-Bakelman, F.E., Darshi, M., Green, S.J., Gur, R.C., Lin, L., Macias, B.R., McKenna, M.J., Meydan, C., Mishra, T., Nasrini, J.*, et al.* (2019). The NASA Twins Study: A multidimensional analysis of a year-long human spaceflight. Science *364*.

George, K., Rhone, J., Beitman, A., and Cucinotta, F.A. (2013). Cytogenetic damage in the blood lymphocytes of astronauts: effects of repeat long-duration space missions. Mutat Res *756*, 165-169.

Greider, C.W., and Blackburn, E.H. (1985). Identification of a Specific Telomere Terminal Transferase-Activity in Tetrahymena Extracts. Cell *43*, 405-413.

Gu, Y., Honig, L.S., Schupf, N., Lee, J.H., Luchsinger, J.A., Stern, Y., and Scarmeas, N. (2015). Mediterranean diet and leukocyte telomere length in a multi-ethnic elderly population. Age (Dordr) *37*, 24.

Hada, M., Cucinotta, F.A., Gonda, S.R., and Wu, H. (2007). mBAND analysis of chromosomal aberrations in human epithelial cells exposed to low- and high-LET radiation. Radiat Res *168*, 98-105.

Harley, C.B., Futcher, A.B., and Greider, C.W. (1990). Telomeres Shorten During Aging of Human Fibroblasts. Nature *345*, 458-460.

Hartwig, F.P., Nedel, F., Collares, T.V., Tarquinio, S.B., Nor, J.E., and Demarco, F.F. (2012). Telomeres and tissue engineering: the potential roles of TERT in VEGF-mediated angiogenesis. Stem Cell Rev Rep *8*, 1275-1281.

Haycock, P.C., Heydon, E.E., Kaptoge, S., Butterworth, A.S., Thompson, A., and Willeit, P. (2014). Leucocyte telomere length and risk of cardiovascular disease: systematic review and meta-analysis. BMJ *349*, g4227.

Honig, L.S., Kang, M.S., Cheng, R., Eckfeldt, J.H., Thyagarajan, B., Leiendecker-Foster, C., Province, M.A., Sanders, J.L., Perls, T., Christensen, K.*, et al.* (2015). Heritability of telomere length in a study of long-lived families. Neurobiol Aging *36*, 2785-2790.

Honig, L.S., Kang, M.S., Schupf, N., Lee, J.H., and Mayeux, R. (2012). Association of shorter leukocyte telomere repeat length with dementia and mortality. Arch Neurol *69*, 1332-1339.

Hou, M., Xu, D., Bjorkholm, M., and Gruber, A. (2001). Real-time quantitative telomeric repeat amplification protocol assay for the detection of telomerase activity. Clin Chem *47*, 519-524.

Hu, J., Hwang, S.S., Liesa, M., Gan, B., Sahin, E., Jaskelioff, M., Ding, Z., Ying, H., Boutin, A.T., Zhang, H.*, et al.* (2012). Antitelomerase therapy provokes ALT and mitochondrial adaptive mechanisms in cancer. Cell *148*, 651-663.

Jung, S.N., Yang, W.K., Kim, J., Kim, H.S., Kim, E.J., Yun, H., Park, H., Kim, S.S., Choe, W., Kang, I.*, et al.* (2008). Reactive oxygen species stabilize hypoxia-inducible factor-1 alpha protein and stimulate transcriptional activity via AMP-activated protein kinase in DU145 human prostate cancer cells. Carcinogenesis *29*, 713-721.

Kim, N.W., Piatyszek, M.A., Prowse, K.R., Harley, C.B., West, M.D., Ho, P.L.C., Coviello, G.M., Wright, W.E., Weinrich, S.L., and Shay, J.W. (1994). Specific Association of Human Telomerase Activity With Immortal Cells and Cancer. Science *266*, 2011-2015.

Lee, S.M.C., Ribeiro, L.C., Martin, D.S., Zwart, S.R., Feiveson, A.H., Laurie, S.S., Macias, B.R., Crucian, B.E., Krieger, S., Weber, D.*, et al.* (2020). Arterial structure and function during and after long-duration spaceflight. J Appl Physiol (1985) *129*, 108-123.

Lerman, O.Z., Greives, M.R., Singh, S.P., Thanik, V.D., Chang, C.C., Seiser, N., Brown, D.J., Knobel, D., Schneider, R.J., Formenti, S.C.*, et al.* (2010). Low-dose radiation augments vasculogenesis signaling through HIF-1-dependent and -independent SDF-1 induction. Blood *116*, 3669-3676.

Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdottir, H., Tamayo, P., and Mesirov, J.P. (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics *27*, 1739-1740.

Liu, B., Sun, Y., Xu, G., Snetselaar, L.G., Ludewig, G., Wallace, R.B., and Bao, W. (2019a). Association between Body Iron Status and Leukocyte Telomere Length, a Biomarker of Biological Aging, in a Nationally Representative Sample of US Adults. J Acad Nutr Diet *119*, 617-625.

Liu, H., Yang, Y., Ge, Y., Liu, J., and Zhao, Y. (2019b). TERC promotes cellular inflammatory response independent of telomerase. Nucleic Acids Res *47*, 8084-8095.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods *25*, 402-408.

Lloyd, D.C., Purrott, R.J., and Dolphin, G.W. (1973). Chromosome aberration dosimetry using human lymphocytes in simulated partial body irradiation. Phys Med Biol *18*, 421-431.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.

Maeda, T., Nakamura, K., Atsumi, K., Hirakawa, M., Ueda, Y., and Makino, N. (2013). Radiation-associated changes in the length of telomeres in peripheral leukocytes from inpatients with cancer. Int J Radiat Biol *89*, 106-109.

McInnes, L., Healy, John, Melville, James (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction.

McKenna, M.J., and Bailey, S.M. (2017). Chromosomal and Telomeric biomarkers of normal tissue injury to evaluate risk of degenerative health effects (secondary malignancy, cardiovascular disease) post radiation therapy. Transl Cancer Res *6*, S789-S794.

McKenna, M.J., Robinson, E., Taylor, L., Tompkins, C., Cornforth, M.N., Simon, S.L., and Bailey, S.M. (2019). Chromosome Translocations, Inversions and Telomere Length for Retrospective Biodosimetry on Exposed U.S. Atomic Veterans. Radiat Res *191*, 311-322.

Miller, D., Reynolds, G.E., Mejia, R., Stark, J.M., and Murnane, J.P. (2011). Subtelomeric regions in mammalian cells are deficient in DNA double-strand break repair. DNA Repair (Amst) *10*, 536-544.

Minamino, T., Mitsialis, S.A., and Kourembanas, S. (2001). Hypoxia extends the life span of vascular smooth muscle cells through telomerase activation. Mol Cell Biol *21*, 3336-3342.

Miri, M., Nazarzadeh, M., Alahabadi, A., Ehrampoush, M.H., Rad, A., Lotfi, M.H., Sheikhha, M.H., Sakhvidi, M.J.Z., Nawrot, T.S., and Dadvand, P. (2019). Air pollution and telomere length in adults: A systematic review and meta-analysis of observational studies. Environ Pollut *244*, 636-647.

Mitelman, F. (2000). Recurrent chromosome aberrations in cancer. Mutat Res *462*, 247-253.
Moyzis, R.K., Buckingham, J.M., Cram, L.S., Dani, M., Deaven, L.L., Jones, M.D., Meyne, J., Ratliff, R.L., and Wu, J.R. (1988). A Highly Conserved Repetitive Dna-Sequence ; (Ttaggg)N ; Present At the Telomeres of Human-Chromosomes. Proceedings of the National Academy of Sciences of the United States of America *85*, 6622-6626.

Murnane, J.P., Sabatier, L., Marder, B.A., and Morgan, W.F. (1994). Telomere Dynamics in an Immortal Human Cell-Line. Embo Journal *13*, 4953-4962.

Nishi, H., Nakada, T., Kyo, S., Inoue, M., Shay, J.W., and Isaka, K. (2004). Hypoxia-inducible factor 1 mediates upregulation of telomerase (hTERT). Mol Cell Biol *24*, 6076-6083.

Ohno, H., Shirato, K., Sakurai, T., Ogasawara, J., Sumitani, Y., Sato, S., Kmaizumi, K., Ishida, H. Kizaki, T. (2012). Effect of exercise on HIF-1 and VEGF signaling. J Phys Fitness Sports Med *1*, 5-16.

Olovnikov, A.M. (1971). Principle of marginotomy in template synthesis of polynucleotides. Dokl Akad Nauk SSSR *201*, 1496-1499.

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods *14*, 417-419.

Paul, L., Cattaneo, M., D'Angelo, A., Sampietro, F., Fermo, I., Razzari, C., Fontana, G., Eugene, N., Jacques, P.F., and Selhub, J. (2009). Telomere length in peripheral blood mononuclear cells is associated with folate status in men. J Nutr *139*, 1273-1278.

Potapova, T.A., Unruh, J.R., Yu, Z., Rancati, G., Li, H., Stampfer, M.R., and Gerton, J.L. (2019). Superresolution microscopy reveals linkages between ribosomal DNA on heterologous chromosomes. J Cell Biol *218*, 2492-2513.

Protsenko, E., Rehkopf, D., Prather, A.A., Epel, E., and Lin, J. (2020). Are long telomeres better than short? Relative contributions of genetically predicted telomere length to neoplastic and non-neoplastic disease risk and population health burden. PLoS One *15*, e0240185.

Ray, F.A., Robinson, E., McKenna, M., Hada, M., George, K., Cucinotta, F., Goodwin, E.H., Bedford, J.S., Bailey, S.M., and Cornforth, M.N. (2014). Directional genomic hybridization: inversions as a potential biodosimeter for retrospective radiation exposure. Radiat Environ Biophys *53*, 255-263.

Ray, F.A., Zimmerman, E., Robinson, B., Cornforth, M.N., Bedford, J.S., Goodwin, E.H., and Bailey, S.M. (2013). Directional genomic hybridization for chromosomal inversion discovery and detection. Chromosome Res *21*, 165-174.

Robinson, E., McKenna, M.J., Bedford, J.S., Goodwin, E.H., Cornforth, M.N., Bailey, S.M., and Ray, F.A. (2019). Directional Genomic Hybridization (dGH) for Detection of Intrachromosomal Rearrangements. Methods Mol Biol *1984*, 107-116.

Rochette, P.J., and Brash, D.E. (2010). Human telomeres are hypersensitive to UV-induced DNA Damage and refractory to repair. PLoS Genet *6*, e1000926.

Seabold, S.a.P., J. (2010). Statsmodels: econometric and statistical modeling with python. Proceeding of the 9th Phthon in Science Conference.

Shay, J.W. (2013a). Are Short Telomeres Hallmarks of Cancer Recurrence? Clin Cancer Res.

Shay, J.W. (2013b). Are short telomeres predictive of advanced cancer? Cancer Discov *3*, 1096-1098.

Shim, G., Ricoul, M., Hempel, W.M., Azzam, E.I., and Sabatier, L. (2014). Crosstalk between telomere maintenance and radiation effects: A key player in the process of radiation-induced carcinogenesis. Mutat Res Rev Mutat Res.

Silva, R., Folgosa, F., Soares, P., Pereira, A.S., Garcia, R., Gestal-Otero, J.J., Tavares, P., and Gomes da Silva, M.D. (2013). Occupational cosmic radiation exposure in Portuguese airline pilots: study of a possible correlation with oxidative biological markers. Radiat Environ Biophys *52*, 211-220.

Sishc, B.J., Nelson, C.B., McKenna, M.J., Battaglia, C.L., Herndon, A., Idate, R., Liber, H.L., and Bailey, S.M. (2015). Telomeres and Telomerase in the Radiation Response: Implications for Instability, Reprograming, and Carcinogenesis. Front Oncol *5*, 257.

Smith, M.G., Kelley, M., Basner, M. (2020). A brief history of spaceflight from 1961 to 2020: an analysis of missions and astronaut demographics. Acta Astronautica *175*, 290-299.

Smith, S.M., Heer, M., Shackelford, L.C., Sibonga, J.D., Spatz, J., Pietrzyk, R.A., Hudson, E.K., and Zwart, S.R. (2015). Bone metabolism and renal stone risk during International Space Station missions. Bone *81*, 712-720.

Smith, S.M., Heer, M., Wang, Z., Huntoon, C.L., and Zwart, S.R. (2012). Long-duration space flight and bed rest effects on testosterone and other steroids. J Clin Endocrinol Metab *97*, 270-278.

Sridharan, D.M., Asaithamby, A., Bailey, S.M., Costes, S.V., Doetsch, P.W., Dynan, W.S., Kronenberg, A., Rithidech, K.N., Saha, J., Snijders, A.M.*, et al.* (2015). Understanding cancer development processes after HZE-particle exposure: roles of ROS, DNA damage repair and inflammation. Radiat Res *183*, 1-26.

Stone, R.C., Horvath, K., Kark, J.D., Susser, E., Tishkoff, S.A., and Aviv, A. (2016). Telomere Length and the Cancer-Atherosclerosis Trade-Off. PLoS Genet *12*, e1006144.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S.*, et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

Valdes, A.M., Andrew, T., Gardner, J.P., Kimura, M., Oelsner, E., Cherkas, L.F., Aviv, A., and Spector, T.D. (2005). Obesity, cigarette smoking, and telomere length in women. Lancet *366*, 662-664.

von Zglinicki, T. (2000). Role of oxidative stress in telomere length regulation and replicative senescence. Ann N Y Acad Sci *908*, 99-110.

Watson, J.D. (1972). Origin of concatameric T7 DNA. Nature: New Biology *239*, 197-201.
Wilson, D.M., 3rd, and Thompson, L.H. (2007). Molecular mechanisms of sister-chromatid exchange. Mutat Res *616*, 11-23.

Wong, H.P., and Slijepcevic, P. (2004). Telomere length measurement in mouse chromosomes by a modified Q-FISH method. Cytogenet Genome Res *105*, 464-470.

Wu, J.Z., and Ho, P.C. (2009). Comparing the relative oxidative DNA damage caused by various arsenic species by quantifying urinary levels of 8-hydroxy-2'-deoxyguanosine with isotope-dilution liquid chromatography/mass spectrometry. Pharm Res *26*, 1525-1533.

Yu, X., Deng, L., Wang, D., Li, N., Chen, X., Cheng, X., Yuan, J., Gao, X., Liao, M., Wang, M.*, et al.* (2012). Mechanism of TNF-alpha autocrine effects in hypoxic cardiomyocytes: initiated by hypoxia inducible factor 1alpha, presented by exosomes. J Mol Cell Cardiol *53*, 848-857.
Yuan, X., Larsson, C., and Xu, D. (2019). Mechanisms underlying the activation of TERT transcription and telomerase activity in human cancer: old actors and new players. Oncogene *38*, 6172-6183.

Zhang, J., Rane, G., Dai, X., Shanmugam, M.K., Arfuso, F., Samy, R.P., Lai, M.K., Kappei, D., Kumar, A.P., and Sethi, G. (2016). Ageing and the telomere connection: An intimate relationship with inflammation. Ageing Res Rev *25*, 55-69.

Zhao, Y., Lai, K., Cheung, I., Youds, J., Tarailo, M., Tarailo, S., and Rose, A. (2006). A mutational analysis of Caenorhabditis elegans in space. Mutat Res *601*, 19-29.

Zheng, Q., Liu, P., Gao, G., Yuan, J., Wang, P., Huang, J., Xie, L., Lu, X., Di, F., Tong, T.*, et al.* (2019). Mitochondrion-processed TERC regulates senescence without affecting telomerase activities. Protein Cell *10*, 631-648.

Zwart, S.R., Booth, S.L., Peterson, J.W., Wang, Z., and Smith, S.M. (2011). Vitamin K status in spaceflight and ground-based models of spaceflight. J Bone Miner Res *26*, 948-954.

Zwart, S.R., Morgan, J.L., and Smith, S.M. (2013). Iron status and its relations with oxidative damage and bone loss during long-duration space flight on the International Space Station. Am J Clin Nutr *98*, 217-223.

CHAPTER 4: AN EXPLORATORY STUDY: MACHINE LEARNING PREDICTIONS OF TELOMERE LENGTH AND CHROMOSOME REARRANGEMENTS POST-RADIATION THERAPY[3]

Chapter summary

Our investigations into telomeric and cytogenetic changes in astronauts as a result of spaceflight highlights the informative nature of these biomarkers for modifying disease risk. Through these works we explored the relevant question of whether predicting telomeric and cytogenetic responses to a given exposure, namely IR, had been previously attempted; we were surprised to find it had not. Being able to predict such events would provide significant and material improvements to predicting an individual's modified disease risk, through changes in telomere length and incidence of chromosome aberrations. Such predictive tools could provide huge benefits to scenarios where personalized monitoring of health is critical, such as for astronauts during spaceflight or patients after radiotherapy. Here, we longitudinally measured telomere length and chromosome aberrations in a cohort of 15 prostate cancer patients undergoing Intensity Modulated Radiation Therapy (IMRT). We demonstrate that telomeric responses to IMRT are highly predictable when utilizing a machine learning framework (XGBoost), and attempt but fail to predict changes in chromosome rearrangements. This work indicates that telomeric responses to given exposures may be generally predictable.

[3] This chapter is as draft of a first author manuscript invited for publication at the Journal of Personalized Medicine. Anticipated publication date is Jan 2020.

Authors:

Jared J. Luxton[1,2], Miles J. McKenna[1,2], Aidan M. Lewis[1], Lynn E. Taylor[1], Sameer G. Jhavar[3], Gregory P. Swanson[3], and Susan M. Bailey[1,2]

Affiliations:

[1]Department of Environmental and Radiological Health Sciences

[2]Cell and Molecular Biology Program, Colorado State University, Fort Collins, CO, USA

[3]Baylor Scott & White Medical Center, Temple, TX, USA

Overview

The ability to predict a cancer patient's response to radiotherapy and risk of developing adverse late health effects would greatly improve personalized treatment regimens and individual outcomes. Telomeres represent a compelling biomarker of individual radiosensitivity and risk, as exposure can result in dysfunctional telomere pathologies that coincidentally overlap with many radiation-induced late effects, ranging from degenerative conditions like fibrosis and cardiovascular disease to proliferative pathologies like cancer. Here, telomere length was longitudinally assessed in a cohort of fifteen prostate cancer patients undergoing Intensity Modulated Radiation Therapy (IMRT) utilizing Telomere Fluorescence *in situ* Hybridization (Telo-FISH). To evaluate genome instability and enhance predictions for individual patient risk of secondary malignancy, chromosome aberrations were also assessed utilizing directional Genomic Hybridization (dGH) for high-resolution inversion detection. We present the first implementation of individual telomere length data in a machine learning model, XGBoost, trained on pre-radiotherapy (baseline) and *in vitro* exposed (4 Gy γ-rays) telomere length measures, to predict post-radiotherapy telomeric outcomes, which together with chromosomal instability provide insight into individual radiosensitivity and risk for radiation-induced late effects.

Introduction

Radiation late effects are a broad class of negative and often permanent health effects experienced by cancer patients long after radiation therapy (Barnett et al., 2009; Bentzen, 2006), which can include cardiovascular disease (Yusuf et al., 2017), pulmonary and arterial fibrosis (Carver et al., 2007), cognitive deficits (Greene-Schloesser and Robbins, 2012), bone fractures

(Yaprak et al., 2018), and secondary cancers (Suit et al., 2007). Such late effects are of particular concern for pediatric patients (Armstrong et al., 2010), and risks for radiation late effects are highly dependent on patient-intrinsic factors as well, including genetics, age, sex, and lifestyle (Barnett et al., 2009; Bentzen, 2006; Rajaraman et al., 2018). Therefore, identifying a patient's specific risks for radiation late effects prior to radiotherapy is important for improving individual treatment planning and overall patient outcomes. A number of strategies for predicting risks for radiation late effects have been employed, which tend to irradiate patient-derived samples *in vitro* for monitoring of biomarker(s) to infer *in vivo* cellular and normal tissue *in vivo* responses to exposure (Habash et al., 2017); e.g., evaluation of γ-H2AX foci kinetics (Banáth et al., 2004; Redon et al., 2009), apoptosis in normal blood lymphocytes (Schmitz et al., 2007), and chromosome aberration frequencies (Baeyens et al., 2002; Baria et al., 2001; Huber et al., 2011). Additionally, Genome Wide Association Studies (GWAS) (Kerns et al., 2010, 2016), sequencing (Young et al., 2014), and imaging studies (i.e radiogenomics (Bodalal et al., 2019)) have revealed promising putative markers that show promise for predicting risks for late effects. However, accurately predicting an individual patient's response to radiotherapy and associated risk of developing adverse late health effects remains challenging in terms of cost-effectiveness, throughput, and predictive power, therefore new approaches are needed.

Telomeres are protective features of chromosomal termini that guard against genome degradation and inappropriate activation of DNA damage responses (DDRs) (de Lange, 2009; Moyzis et al., 1988). It is well established that telomeres shorten with cell division, oxidative stress (von Zglinicki, 2000), and aging (Aubert and Lansdorp, 2008). Telomeres also shorten with a host of lifestyle factors (e.g., nutrition (Vidaček et al., 2017), exercise (Arsenis et al., 2017), stress (Epel et al., 2004)) and environmental exposures (e.g. air pollution (Miri et al.,

2019), UV (Stout and Blasco, 2013)) as well. Telomere length is a highly heritable trait, as is telomere length regulation (Broer et al., 2013; Delgado et al., 2019; Honig et al., 2015; Weng et al., 2016), supportive of individual variation in telomeric response to specific stressors. Interestingly, short telomeres have been proposed as hallmarks of radiosensitivity (Ayouaz et al., 2008), and ionizing radiation (IR) exposure has been shown to evoke both shortening and lengthening of telomeres (Bains et al., 2019; Berardinelli et al., 2013; De Vitis et al., 2019; F et al., 2014; Maeda et al., 2013; Sgura et al., 2006). Short telomeres are biomarkers and even effectors for a range of aging-related pathologies (Armanios and Blackburn, 2012), including cardiovascular disease (CVD) (Martínez Paula and Blasco Maria A., 2018), pulmonary fibrosis (Martínez Paula and Blasco Maria A., 2018), and aplastic anemia (Calado et al., 2012), degenerative conditions also regarded as radiation late effects (Adams et al., 2003; Green and Rubin, 2014; Tsoutsou and Koukourakis, 2006). On the other hand, longer telomeres are associated with increased cancer risk, particularly for leukemias (McNally et al., 2019), a common cancer following IR-exposure (Dracham et al., 2018). Thus, patients with shorter telomeres after radiation therapy are more likely to develop short telomere (degenerative) pathologies, while patients with longer telomeres following radiotherapy are at higher risk for developing proliferative pathologies (cancer).

Given that telomere length is influenced by a variety of genetic factors (Broer et al., 2013; Delgado et al., 2019; Honig et al., 2015; Weng et al., 2016) and exposures including IR exposure(Bains et al., 2019; Berardinelli et al., 2013; De Vitis et al., 2019; F et al., 2014; Maeda et al., 2013; Sgura et al., 2006), we reasoned that a patient's telomeric outcome post-radiation therapy, rather than their pre-treatment (baseline) measures, would be most informative for assessing individual risks for radiation late effects and long-term health consequences.

Furthermore, since patient-derived pre-radiation therapy samples irradiated *in vitro* provide an informative proxy for individual patient radiosensitivity and response *in vivo* (Alsner et al., 2007; Andreassen et al., 2013; Paul et al., 2011), an effective means to accurately predict an individual patient's telomeric outcome post-radiation therapy could be developed, thereby improving personalized treatment strategies and individual outcomes.

Chromosome aberrations (CAs) are well-established biomarkers of IR-exposure (Huang et al., 2003), associated with virtually all cancers (Willis et al., 2015), and highly informative indicators of risk for radiation late effects, in particular, secondary cancers (Baeyens et al., 2002; Baria et al., 2001; Huber et al., 2011). Ionizing radiation is exceptional in its ability to induce prompt double strand breaks (DSBs) (Ward, 1988), damage that obligates a cellular response to address and resolve. Chromosome rearrangements result from the misrepair of such damage, and so provide a quantitative measure of cellular capacity for DNA repair (Huang et al., 2003). In general, IR-induced CAs negatively impact cell survival and genome stability, resulting in senescence, apoptosis, and cancer (Huang et al., 2003), respectively. Notably, chromosomal inversions and deletions have previously been proposed as signatures of radiation-induced secondary cancers (Behjati et al., 2016). Cytogenetic analysis however, is both time and labor intensive, often requiring that hundreds or even thousands of cells be scored, limiting its clinical utility (Mosesso and Cinelli, 2019). We speculated that inclusion of an additional type of CA, specifically inversions, now possible using the strand-specific cytogenomic methodology of directional Genome Hybridization (dGH) (Ray et al., 2013), might serve to reduce the number of cells required, while also informing potential risks for secondary cancers.

Significant advancements have also been made in the application of machine learning (ML) to a variety of scenarios, including predictions related to acute radiation toxicity (Pella et

al., 2011), treatment planning (Fan et al., 2019), and secondary cancer risk post radiation therapy (Lee et al., 2020). Extreme Gradient Boosting (XGBoost) is a powerful ML model that uses a gradient boosted ensemble of decision trees to learn complex relationships (linear and nonlinear) within data (Chen and Guestrin, 2016). XGBoost has many translational applications, such as predicting future gastric cancer risk (Taninaga et al., 2019), lung cancer detection (Yu et al., 2020), and radiation-related fibrosis (Wang et al., 2019). One potentially limiting caveat to ML is the requirement for extraordinarily large amounts of data to create robust, generalizable models. Telomere Fluorescence *in situ* Hybridization (Telo-FISH) is a cell-by-cell imaging-based approach for measuring telomere length capable of generating sufficient volumes of data for developing ML models; average experiments generate 200,000 - 1,000,000 individual telomere length measurements (Poon and Lansdorp, 2001). Interestingly and to date, individual telomere length measurements (Telo-FISH, Q-FISH, flow-FISH, etc.) have not been utilized in ML models for risk predictions, despite the informative nature of such an approach.

Here we provide a proof-of-principle demonstration utilizing longitudinal analysis of telomere length and chromosomal instability in fifteen (15) prostate cancer patients undergoing Intensity Modulated Radiation Therapy (IMRT). We present the first implementation of individual telomere length (Telo-FISH) data in a ML model - XGBoost - and evaluate its ability to predict post-IMRT telomeric outcomes using individual patient's pre-IMRT (baseline) and *in vitro* irradiated telomere lengths. Overall, results provide insight into predicting individual radiosensitivity and risk for radiation-induced late effects.

Results

Longitudinal analyses of telomere length associated with radiation therapy

Blood was collected from 15 prostate cancer patients undergoing IMRT at baseline (pre-IMRT), immediately post-IMRT (conclusion of treatment regimen), and three months post-IMRT. Baseline blood samples were split, half serving as the non-irradiated control (0 Gy), and the other half irradiated *in vitro* (4 Gy, Cs^{137} γ-rays) as a proxy for individual radiation response. The lengths of thousands of individual telomeres (n = 50 cells/patient/time point) were measured on metaphase chromosomes (lymphocytes stimulated from whole blood) by Telo-FISH at all time points (1: pre-therapy non-irradiated; 2: *in vitro* irradiated (4 Gy); 3B: immediately post-IMRT; and 4 C: three months post-IMRT) (**Figure 4.1A)**. For the overall cohort, differences in mean telomere length (MTL) between samples approached, but did not reach statistical significance (p = 0.059, repeated measures ANOVA). Relative to the pre-IMRT non-irradiated samples, overall MTL modestly increased after 4 Gy *in vitro* irradiation, and showed an even greater increase immediately after completion of the IMRT regimen, suggesting that increased MTL is an overall response to radiation exposure in this cohort. At three months IMRT, MTL for the cohort approached pre-IMRT levels.

Complete blood counts (CBC) were also evaluated in the same samples, and longitudinal changes in patients' MTL were negatively correlated (R^2 = -0.126) with total peripheral white blood cell (WBC) counts (**Figure 4.2A**). Longitudinal correlations between numbers of WBC types and MTL (all time points, for each patient) revealed a positive relationship with basophils (R^2 = 0.278) and a negative relationship with lymphocytes (R^2 = -0.294) (**Figure 4.2B**). Furthermore, longitudinal correlations between MTL and the proportions of lymphocyte sub-groups (all time points, for each patient) revealed positive relationships with natural killer (NK) and CD4 cells (R^2 = 0.408, 0.282), and negative relationships with CD8 and CD19 cells (R^2 = -0.251, -0.288) (**Figure 4.2C**). These results support the notion that the overall changes in MTL

associated with radiation exposure, specifically apparent telomere elongation, could be at least partially due to cell killing and shifts in lymphocyte populations, as previously proposed (Maeda et al., 2013).

Telomere length dynamics revealed individual differences in radiation response

We hypothesized that groups of patients would cluster based on differential telomeric responses to radiation therapy, with sub-groups displaying either shorter or longer MTL post-IMRT. Clustering patients by longitudinal changes in MTL revealed two broad trends over time (**Figure 4.1B**). Patients that clustered in group 1 (n=3) had relatively longer MTL at baseline (pre-IMRT), and showed a dramatic, persistent *decrease* in MTL post-IMRT (**Figure 4.1C**). Those patients that clustered in group 2 (n=11) had relatively shorter MTL at baseline, and showed a dramatic, sustained *increase* in MTL post-IMRT (**Figure 4.1C**). Reduced MTL three months post-IMRT suggests increased risks for degenerative radiation late effects (Calado et al., 2012; Martínez Paula and Blasco Maria A., 2018), while increased MTL suggests increased risks for proliferative secondary cancers (McNally et al., 2019).

In addition to MTL, Telo-FISH provides measures for many hundreds of individual telomeres, enabling generation of telomere length distributions and longitudinal analysis of shifts in populations of short and long telomeres (Poon and Lansdorp, 2001). For the overall cohort, numbers of short telomeres (yellow) decreased and numbers of long telomeres (red) dramatically increased three months post-IMRT (**Figure 4.3A**). When individual telomeres from patients in the MTL clustered group 1 (n=3) were combined, dramatic and persistent increases in the numbers of short telomeres post-IMRT were observed (**Figure 4.3B**), while MTL clustered group 2 patients (n=11) showed dramatic and persistent increases in numbers of long telomeres

post-IMRT (**Figure 4.3C**). Again, patients with increased numbers of short telomeres are presumed to have increased risks for degenerative radiation late effects (Calado et al., 2012; Martínez Paula and Blasco Maria A., 2018), while those with increased numbers of long telomeres are at increased risk of secondary cancers (McNally et al., 2019). Numbers of short and long telomeres were feature engineered (see Materials and Methods) from each patient's individual telomere length data for further analysis.

Differences in the average number of short and long telomeres between samples approached but did not reach statistical significance for the overall cohort ($p<0.1$; repeated measures ANOVA) (**Figure 4.4A**). We speculated that clustering patients by numbers of short or long telomeres would reveal longitudinal trends similar to those observed when clustering patients by MTL (**Figure 4.1B/C**). Clustering patients by longitudinal changes in numbers of short or long telomeres (**Figure 4.4B/D**) revealed two broad trends over time (**Figure 4.4C/E**). Clustered group 1 (n=3) showed a dramatic, sustained increase in numbers of short telomeres post-IMRT, with a corresponding decrease in numbers of long telomeres (**Figure 4.4C/E**). Clustered group 2 (n=11) showed a dramatic, nearly uniform decrease in numbers of short telomeres post-IMRT, with a corresponding increase in long telomeres (**Figure 4.4C/E**). Importantly, clustering patients either by MTL or by numbers of short or long telomeres post-IMRT identified the same three patients with shorter telomeres, and eleven with longer telomeres (**Figure 4.1B**, **Figure 4.4B/D**).

Linear regression poorly predicted post-IMRT telomeric outcomes

Based on the two distinct groups identified by MTL and numbers of short and long telomeres three months post-IMRT (**Figure 4.1, 4.3**), we hypothesized that pre-IMRT

measurements of MTL and numbers of short and long telomeres could predict their respective post-IMRT outcomes using linear regression. For MTL, two linear regression models were created. The first used only MTL from pre-IMRT (baseline) non-irradiated samples as the independent variable, and the second used MTL from both the non-irradiated and *in vitro* irradiated pre-IMRT samples as independent variables for predicting post-IMRT MTL (**Figure 4.5A**). The R^2 values for the two models were 0.161 and 0.165 respectively (**Figure 4.5A**), evidence that linear regression poorly captured the relationship between pre- and post-IMRT MTL. For numbers of short and long telomeres, two linear regression models were similarly created. The models for short telomeres yielded R^2 values of 0.433 and 0.554, and the models for long telomeres yielded R^2 values of 0.046 and 0.208 (**Figure 4.5B/C**). While the models for numbers of short telomeres had modestly higher R^2 values than those for MTL or long telomeres, all linear regression models performed too poorly to confidently predict telomeric outcomes.

Development of XGBoost machine learning models for accurate prediction of post-IMRT telomeric outcomes

The fact that linear regression poorly predicted post-IMRT telomeric outcomes could be due to the low number of observations (n=14), and/or the nonlinearity of telomere length dynamics (changes over time) in response to radiation exposure (**Figure 4.1-.4).** We sought an alternative approach that could effectively utilize our vast dataset of pre-IMRT individual telomere length measurements (n=128,800), and also capture the nonlinearity of telomeric responses. Considering that XGBoost had recently been used to predict cancer risk and radiation-induced fibrosis using patient data (Chen and Guestrin, 2016; Taninaga et al., 2019; Wang et al.,

2019; Yu et al., 2020), we hypothesized that XGBoost models could be trained with pre-IMRT individual telomere length measurements to accurately predict post-IMRT telomeric outcomes.

Pre-IMRT (baseline) telomere length data required extensive preprocessing prior to training the XGBoost model for predicting three-month post-IMRT MTL (**Figure 4.6**). Data was reshaped into a matrix consisting of 128,800 rows (one for each individual telomere measurement) and four columns: patient ID, individual telomere length value, label denoting pre-IMRT sample of origin (non-irradiated or *in vitro* irradiated), and three-month post-IMRT MTL (**Table 4.1A**). Reshaped data was randomly shuffled and stratified by patient ID and sample of origin, then split into training (80% of total) and test (20% of total) data sets. Shuffling guarded against order of measurement bias (Telo-FISH image acquisition), while stratifying ensured equivalent numbers of individual telomeres from each patients' pre-IMRT samples (non-irradiated vs. *in vitro* irradiated) in the training and test data sets. Patient IDs were stripped from the training and test data sets, and individual telomeres from the non-irradiated and *in vitro* irradiated samples were encoded as 0 and 1 to denote sample origin (**Table 4.1B**). XGBoost model hyperparameters were optimized using a randomized hyperparameter search (Bergstra and Bengio, 2012).

XGBoost model performance was evaluated across the training data set using five-fold cross validation (Stone, 1974). Mean absolute error (MAE), the mean of all differences between predicted and actual values of mean telomere length, was used to assess the model's performance and ability to generalize to new data (**Table 4.2A**). Five-fold cross validation on the full training data set yielded an average MAE of 3.233 with a standard deviation of 0.052 (**Table 4.2A**), suggesting that the model was not overfitting to portions (folds) of the training data and that it could generalize to new data. Model performance was also evaluated when training across

variable numbers of individual telomere measurements (n=100 to 103,040) (**Table 4.2A**). After training the XGBoost model on the full training data set, the model was challenged to predict three-month post-IMRT MTL using new data - the test data set. The XGBoost model predictions for MTLs in the test set matched the true values with an R^2 value of 0.882 (**Table 4.2A**). Averaging predictions per patient for three-month post-IMRT MTL in the test set increased the R^2 value to 0.931 (**Figure 4.7A/D**).

XGBoost models were then challenged to predict post-IMRT telomere length on patients whose pre-IMRT individual telomeres they were not trained on. We iteratively trained 14 XGBoost models, where in each model we "left out" (i.e round robin approach) one patient's individual telomeres from XGBoost training, but included the telomeres during model testing. Generally speaking, the XGboost models were extremely performant in predicting post-IMRT telomere length on entirely new patients (**Figure 4.8A-N**). Some deviations in performance were observed – we attribute these deviations in performance to low sample size. We also attempted a "leave two" and "leave three" patients out training and testing approach to understand the limits of generalizability for our XGBoost models (**Figure 4.9A-N; Figure 4.10A-N**). We again found evidence of strong generalizability by the models to new patients. Together, these results demonstrate that the XGBoost model learned the nonlinear relationships between pre-IMRT individual telomere length data and three-month post-IMRT MTLs (training data set), and generalized to new data (test data set) and new patients with highly accurate predictions.

Pre-IMRT individual telomere length data was also processed and reshaped for training separate XGBoost models to predict numbers of short or long telomeres three months post-IMRT (**Figure 4.7B/C**, **Table 4.1C/E**). Reshaped data was split into training (80%) and test (20%) data sets and shuffled and stratified in an identical manner as described for MTL (**Table 4.1D/F**).

Hyperparameters of the XGBoost models were optimized using a randomized search (Bergstra and Bengio, 2012), and the model's performance and generalizability were analyzed using five-fold cross validation (Stone, 1974) with a MAE error metric. For XGBoost models for short telomeres, five-fold cross validation on the full training data set yielded an average MAE of 232.3 with a standard deviation of 5.870 (**Table 4.2B**), while XGBoost models for long telomeres yielded an average MAE of 326.0 and standard deviation of 3.93, suggesting that both models were reasonably good at fitting the data and likely to generalize to new data (**Table 4.2C**). Model performance was also evaluated using variable numbers of training data (n=100 to 103,040). Fully trained XGBoost models were challenged with predicting three-month post-IMRT numbers of short or long telomeres in the test set, and predictions matched the true values with an R^2 value of 0.814 and 0.827, respectively (**Figure 4.7B/D, Table 4.2B/C**). Averaging predictions per patient for post-IMRT numbers of short or long telomeres increased the R^2 value to 0.877 and 0.890, respectively (**Fig 4.7E/F**). These results suggested that the XGBoost models learned the relationships between pre-IMRT individual telomere length data and three-month post-IMRT numbers of short or long telomeres (training data set), and effectively generalized to new data (test data set).

Longitudinal analyses of chromosomal instability associated with radiation therapy

Directional Genomic Hybridization (dGH) is a cytogenomics, fluorescence-based methodology for high-resolution detection of chromosome aberrations (CAs) missed even by sequencing (Cornforth et al., 2018), particularly inversions (Garrett-Bakelman et al., 2019; Ray et al., 2013). We hypothesized that the increased efficiency of dGH for detecting inversions would facilitate scoring fewer metaphase spreads (n=30/time point/patient) than traditional

cytogenetic techniques (Mosesso and Cinelli, 2019), while still retaining superb sensitivity to individual chromosomal instability, and thus the ability to infer patients at higher risks for secondary cancers. Many significant differences in frequencies of IR-induced rearrangements were observed (**Figure 4.11A-D**), with inversions occurring at the highest frequencies, consistent with expectations (Garrett-Bakelman et al., 2019; Ray et al., 2013). Interestingly, overall average frequencies of inversions at three months post-IMRT were comparable to the *in vitro* irradiated samples (**Figure 4.11A**). Frequencies of translocations, dicentrics, and chromosome fragments (deletions) were highest after *in vitro* irradiation, and remained relatively high immediately post-IMRT (**Figure 4.11B-D)**. High frequencies of translocations and chromosome fragments persisted post-IMRT (**Figure 4.11B/D**), while dicentrics decreased somewhat (**Figure 4.11C**). Frequencies of sister chromatid exchanges (SCE) did not significantly change over time, consistent with expectation and exposure. Notably, elevated frequencies of CAs at three months post-IMRT suggested ongoing genomic instability in the overall cohort (Behjati et al., 2016; Huang et al., 2003; Willis et al., 2015) (**Figure 4.11A-D**).

Significant changes in frequencies of IR-induced rearrangements were also coincident with numbers of peripheral blood lymphocytes. Longitudinal correlations between patients' average frequencies of CAs and numbers of peripheral blood lymphocytes (all time points) revealed strongly negative correlations (**Figure 4.12A-D**). Frequencies of inversions and dicentrics had the highest negative correlations (R^2 = -0.752, -0.751), indicating they were highly informative - and similar - markers for cell death. These results suggest that patients demonstrating chromosomal instability (specifically, elevated frequencies of inversions or dicentrics), also experience higher levels of cell killing (i.e., greater radiosensitivity) consistent with previous reports (Ballarini et al., 2015; Carrano, 1973).

Next, we hypothesized that clustering patients by longitudinal changes in CA frequencies (all samples) would reveal groups of patients with lower or higher frequencies of CAs, which would be indicative of individual chromosomal instability and radiosensitivity. When clustering patients by CA type, we observed groups of patients with differential responses only for inversions and chromosome fragments (deletions), which displayed increased frequencies immediately post-IMRT, suggesting increased chromosomal instability (**Figure 4.13A/D**, **Figure 4.14A/B**). We note that the two patients with the highest post-IMRT frequencies of inversions (ID #16) and chromosome fragments (ID #6), also had very high post-IMRT MTLs; both biomarkers suggestive of increased risks for secondary cancers (Baeyens et al., 2002; Baria et al., 2001; Huber et al., 2011; McNally et al., 2019) **(Figure 4.1C, Figure 4.13A/D**, **Figure 4.14A/B**).

Other CA types had longitudinal responses that were relatively uniform between patients and did not cluster patients (**Figure 4.13B/C**, **Figure 4.15A/B**). We hypothesized that while individual types of CAs failed to cluster patients into groups, individual patients may show lower or higher frequencies of CAs. To determine if some patients showed a general susceptibility to chromosomal instability, we feature engineered 'aberration index' by summing all types of CAs (less sister chromatid exchanges) (**Figure 4.13A-D**) per cell for all time points. As indicated by the aberration index, groups of patients with lower or higher total CA frequencies were not observed (**Figure 4.13E**, **Figure 4.15C**). These results in conjunction with the telomere length data, identified two patients (ID #s 6, 16) at potentially increased risks for secondary cancers (Baeyens et al., 2002; Baria et al., 2001; Huber et al., 2011; McNally et al., 2019), and are supportive of inversions and deletions being more informative than other CA types for predicting IR-induced secondary cancers, consistent with prior reports (Behjati et al., 2016). These results

also indicate that the numbers of cells scored were too low (n=30) to detect significant differences in individual patient susceptibility to chromosomal instability in general.

Linear regression poorly predicted radiation-induced chromosomal instability

We speculated that pre-IMRT CA frequencies could be predictive of post-IMRT frequencies. Two linear regression models were made for each CA type to predict post-IMRT frequencies; the first used only the pre-IMRT (baseline) non-irradiated sample CA frequency, and the second used CA frequencies from both pre-IMRT non-irradiated and *in vitro* irradiated samples. The models showed poor predictive power overall, and although inclusion of the *in vitro* irradiated sample data improved performance overall, both models were insufficient for predicting post-IMRT CA frequencies with confidence (**Figure 4.16A-E**). The model for dicentrics performed best, with an R^2 score of 0.514 when using data from both irradiated and non-irradiated baseline samples. These results suggest that while *in vitro* irradiated sample data added predictive power, the number of cells scored per time point/patient (n=30) was too low to enable accurate predictions of individual patient outcomes regarding CAs frequencies post-IMRT using linear regression.

XGBoost machine learning models poorly predicted radiation-induced chromosomal instability

We attempted training XGBoost models using pre-IMRT (baseline) CA data to predict post-IMRT CA frequencies. Rather than using CA data per patient, which would be insufficient for model training (n=15), we used pre-IMRT CA frequencies on a per cell basis (n=840) to predict three-month post-IMRT average CA frequencies. Pre-IMRT CA frequency data was

extensively processed prior to XGBoost model training (**Figure 4.18, Table 4.3A-D**), in a nearly identical manner as described for pre-IMRT telomere length data. The key difference was that CA data was reshaped to train XGBoost models with pre-IMRT CA count data per cell (n=672 cells) in order to predict three-month post-IMRT average CA frequencies. Separate datasets and XGBoost models were created for each type of CA.

XGBoost models for each type of CA were evaluated across their respective training sets using five-fold cross validation (Stone, 1974) with a MAE metric. The cross-validation metrics for all XGBoost models with CA data suggested a failure of the models to learn relationships between pre-IMRT CA count data per cell and three-month post-IMRT average CA frequencies (**Table 4.4A**). Furthermore, dramatic fluctuations in model performance were noted when running multiple iterations of cross-validation, again suggesting that the models failed to learn the relationships between the pre- and post-IMRT CA frequencies (**Table 4.4A-C**). We attempted to improve model performance with many types of feature engineering (e.g boolean features), numerical transformations, and adjustments to model hyperparameters, none of which yielded meaningful improvements in any combination (data not shown). Regardless of poor model performance in cross-validation, we challenged the XGBoost models to predict post-IMRT average CA frequencies using pre-IMRT CA count data per cell in the test set (n=168 cells). In XGBoost model predictions for three month post-IMRT CA frequencies in the test set, none of the predictions matched the true values, with an R^2 above 0.1 (**Figure 4.17A-E, Table 4.4A-C**). These results indicate that either the amount of data was insufficient for training XGBoost models (n=840 cells at pre-IMRT), or the strategy of predicting post-IMRT average CA frequencies using pre-IMRT CA count data per cell was inherently faulty.

Discussion

Better prediction of a cancer patient's individual response to radiation therapy and risk for developing adverse late health effects remains a prime objective for the treatment modality in general (Barnett et al., 2009; Bentzen, 2006; Carver et al., 2007; Greene-Schloesser and Robbins, 2012; Suit et al., 2007; Yaprak et al., 2018; Yusuf et al., 2017), and particularly in regard to pediatric patients (Armstrong et al., 2010). Over recent years, a variety of approaches for predicting radiation late effects have been developed (Baeyens et al., 2002; Banáth et al., 2004; Baria et al., 2001; Bodalal et al., 2019; Habash et al., 2017; Huber et al., 2011; Kerns et al., 2010, 2016; Redon et al., 2009; Schmitz et al., 2007; Young et al., 2014), albeit with varying degrees of compromise between cost-effectiveness, throughput, and predictive power. One notable and extremely promising exception is the use of ML models, which can leverage extensive amounts of patient data to make accurate predictions of treatment outcomes (Fan et al., 2019; Lee et al., 2020; Pella et al., 2011; Taninaga et al., 2019; Wang et al., 2019; Yu et al., 2020).

Predicting a patient's telomeric response to radiation therapy is of clinical interest for predicting risks of radiation late effects, as shorter telomeres confer radiosensitivity (Ayouaz et al., 2008) and increase risk of degenerative late effects (CVD (Martínez Paula and Blasco Maria A., 2018), pulmonary fibrosis (Martínez Paula and Blasco Maria A., 2018), aplastic anemi a (Calado et al., 2012)), while longer telomeres increase risk for secondary cancers, particularly leukemias (McNally et al., 2019). Given that telomeric responses to radiation exposure can be highly dynamic (Bains et al., 2019; Berardinelli et al., 2013; De Vitis et al., 2019; F et al., 2014; Maeda et al., 2013; Sgura et al., 2006) and vary between individuals (**Figure 4.1-3**), a framework for predicting a patient's particular telomeric responses to radiation therapy is critical for utilizing telomere length as a biomarker for radiation late effects. Here, we demonstrate the

feasibility of using ML to accurately predict an individual patient's telomeric response to radiation therapy. We successfully implemented individual telomere length data in a machine learning model, XGBoost, for highly accurate predictions of post-IMRT telomeric outcomes (**Fig 5/6, Table 4.3**). The ML models and Telo-FISH methods used are fully available, providing a valuable resource for continued research into telomere length as a biomarker for radiation late effects associated with any manner of exposure.

The possibility of improving assessment of chromosomal instability and associated risk for development of secondary cancers following radiation therapy (Baeyens et al., 2002; Baria et al., 2001; Huber et al., 2011) was also explored utilizing dGH, which facilitated inversion detection at higher resolution than traditional cytogenetic assays (Garrett-Bakelman et al., 2019; Ray et al., 2013). Indeed, inversions were observed at higher frequencies than other types of CAs both before and after radiation therapy (**Figure 4.11A**), consistent with prior reports (Garrett-Bakelman et al., 2019; Ray et al., 2013). Groups of patients with increased frequencies of chromosomal inversions and fragments (deletions), previously proposed signatures of radiation-induced cancers (Behjati et al., 2016), were also observed three months post-IMRT (**Figure 4.13A/D**). Two patients from these groups had very high MTLs three months post-IMRT as well, also supportive increased risks for secondary cancers (Baeyens et al., 2002; Baria et al., 2001; Huber et al., 2011; McNally et al., 2019). We attempted to derive some predictive value from CA data with linear regression and XGBoost implementations, but both efforts were summarily unsuccessful; the low numbers of cells scored per patient likely subverted successful predictions from the data.

Although we were unable to predict post-IMRT changes in CA frequencies, our observations of strong correlations between patients' average frequencies of CAs and changes in

peripheral blood lymphocyte counts associated with IMRT indicated that our approach of detecting inversions for evaluating chromosomal instability was fundamentally correct. Patients with higher levels of radiation therapy-induced chromosomal instability also experienced increased levels of cell death; i.e., they exhibited individual radiosensitivity (Ballarini et al., 2015; Carrano, 1973). Inversions and dicentrics in particular had strong, negative correlations with lymphocytes cell counts (R^2 = -0.752, -0.751) (**Figure 4.12A/B**).

Relationships between peripheral blood cell count data and MTL were also observed. Counts of peripheral white blood cells (WBCs) were negatively correlated with MTL associated with IMRT (R^2 = -0.126), supportive of shorter telomeres contributing to cell killing/radiosensitivity (**Figure 4.2A**). When parsing WBCs by sub-type, a stronger negative relationship between MTL and lymphocyte counts was seen (R^2 = -0.294). When parsing lymphocytes by sub-type and correlating MTL with the proportions of cell-types, we observed positive correlations with NK and CD4 cells (R^2 = 0.408, 0.282), and negative correlations with CD8 and CD19 cells (R^2 = -0.251, -0.288). These results support our previously proposed supposition that the observed changes in MTL associated with radiation exposure could be partially due to changes in peripheral blood lymphocyte cell populations (Maeda et al., 2013).

Longitudinal assessment of individual telomere length by Telo-FISH in cancer patients undergoing IMRT facilitated demonstration of XGBoost as the ML model of choice for predicting telomeric outcomes post-IMRT. Given the notion that risks for radiation late effects occur on a spectrum (Armstrong et al., 2010; Barnett et al., 2009; Bentzen, 2006; Carver et al., 2007; Greene-Schloesser and Robbins, 2012; Suit et al., 2007; Yaprak et al., 2018; Yusuf et al., 2017), and the differential telomeric responses between individuals and radiation modalities, we posit that the true range of telomeric responses for radiation therapy patients in general is much

broader than those observed here in this prostate cancer cohort (**Figure 4.1-3**). Thus, while our XGBoost models effectively generalized to new data within our experimental design (similar patient sex, radiation modality, cancer type, etc.) (**Fig 6**, **Table 4.3**), it's unlikely that our trained models, in their current iteration, would generalize to data collected under different experimental parameters. Moreover, with regard to measurement of individual telomere lengths for training XGBoost models, Telo-FISH could readily be interchanged with comparable assays (Q-FISH, flow-FISH), which may provide higher throughput. Additionally, the ML approaches described here were not strictly dependent upon XGBoost, and could be conducted using other machine learning models and frameworks (e.g., random forests, kNN). Our paradigm of training ML models with individual telomere length data for prediction of post-IMRT telomeric outcomes provides improved predictive power and novel insight into individual radiosensitivity and risk of late effects, as well as a general framework that could be deployed for radiation therapy patients regardless of cancer type, radiation modality, or individual patient sex or genetic susceptibilities.

Figures

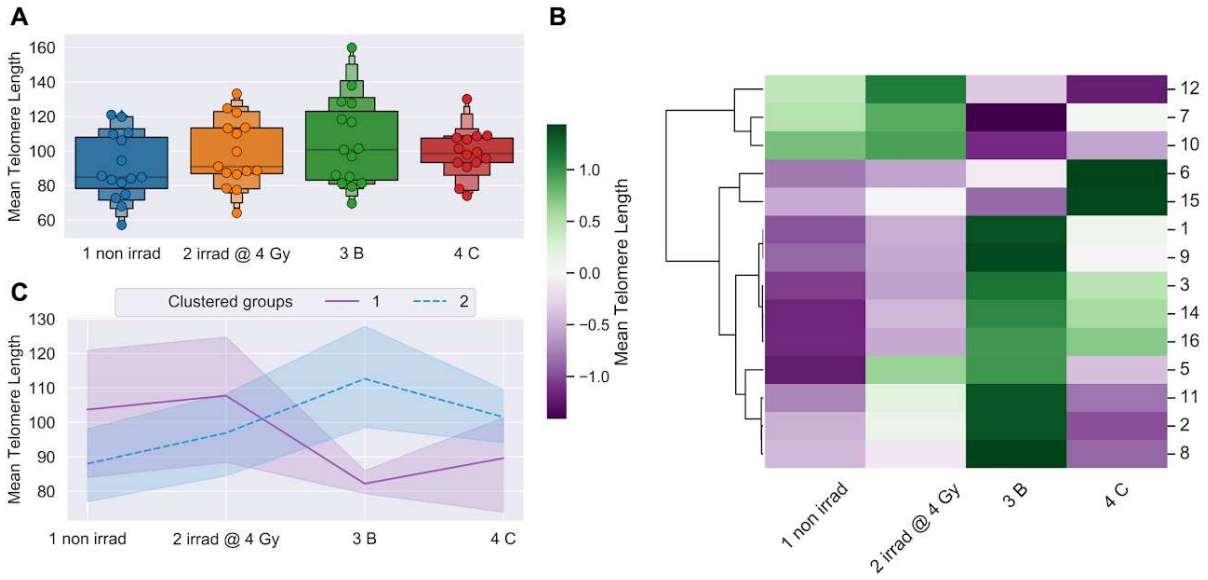

Figure 4.1. Telomere length dynamics (Telo-FISH). Mean telomere length expressed as relative fluorescence intensity. A) Time-course for all patients (n=15; 50 cells/patient/time point): 1 non irrad: pre-IMRT non-irradiated; 2 irrad @ 4 Gy: pre-IMRT in vitro irradiated; 3 B: immediate post-IMRT; and 4 C: 3 months post-IMRT. Boxes denote quantiles, horizontal grey lines denote medians. Telomere values were standardized using a pair of BJ1/BJ-hTERT control samples. B) Hierarchical clustering of patients by longitudinal changes in mean telomere length (z-score normalized). C) Time-course for clustered groups of patients (n=3, purple; n=11, blue); center lines denote medians, lighter bands denote confidence intervals. Patient ID 13 not clustered; 3 months post-IMRT sample failed to culture. Significance was assessed using a repeated measures ANOVA and post-hoc Tukey's HSD test.

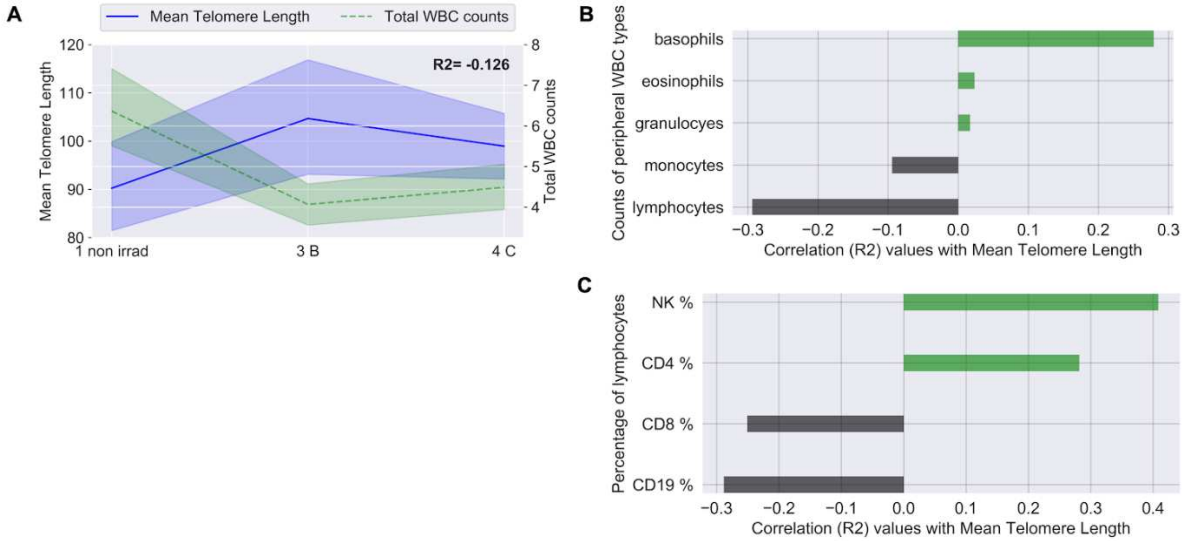

Figure 4.2. Correlations between telomere length, peripheral white blood cells, and lymphocytes. Mean telomere length (Telo-FISH) plotted longitudinally against peripheral white blood cell (WBC) counts (thousands per microliter) from complete blood count tests for all patients; and longitudinal correlations between mean telomere length and counts of WBC types, and proportions of lymphocyte cell types. 1 non irrad: pre-IMRT non-irradiated; 3 B: immediate post-IMRT; 4 C: 3 months post-IMRT. Pearson correlation R2 values were calculated between longitudinal values, as shown bolded in (A), on a per patient basis. Correlations between mean telomere length and WBC counts A); center lines denote medians, lighter bands denote confidence intervals. Correlations between mean telomere length and five main WBC types B), and proportions of lymphocyte cell types C).

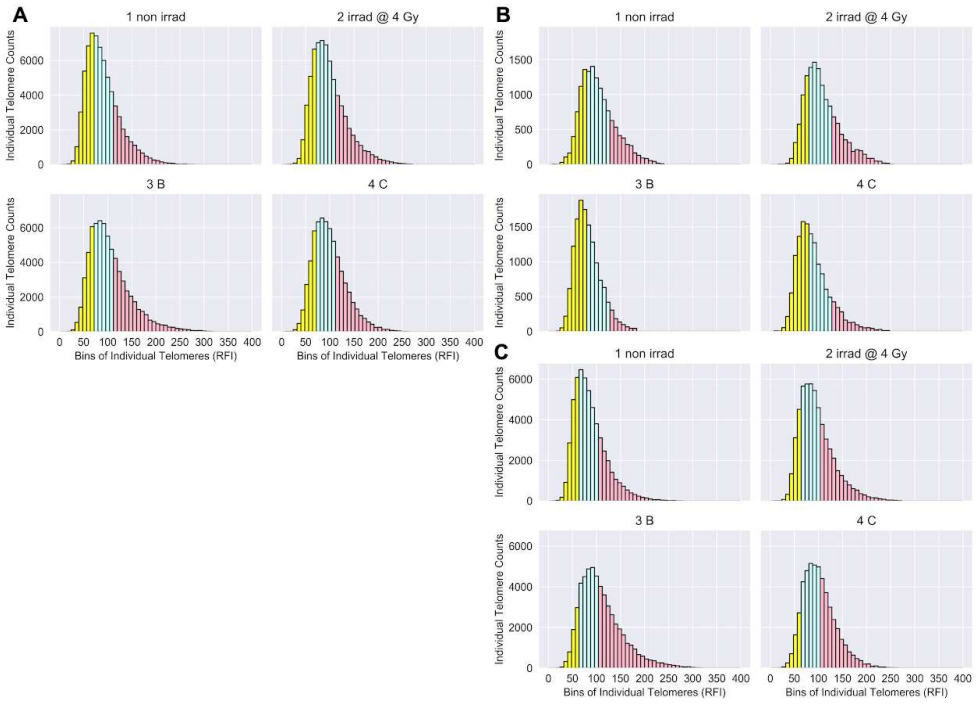

Figure 4.3. Telomere length distributions (Telo-FISH). Individual telomere length distributions of patients (n=15): 1 non irrad: pre-IMRT non-irradiated; 2 irrad @ 4 Gy: pre-IMRT in vitro irradiated; 3 B: immediate post-IMRT; 4 C: 3 months post-IMRT. RFI: Relative Fluorescence Intensity. Individual telomeres from the pre-therapy non-irradiated time point were split into quartiles, designating telomeres in bottom 25% in yellow, the middle 50% in blue, and top 25% in red. Quartile cut-off values, established by the distribution of the pre-therapy non-irradiated time point, were applied to subsequent time points to feature engineer the relative shortest (yellow), mid-length (blue), and longest (red) individual telomeres per time point. A) Individual telomeres for all patients (averaged) per time point. B) Individual telomeres for patients in mean telomere length clustered group 1 (n=3; aggregated) and C) group 2 (n=11; aggregated).

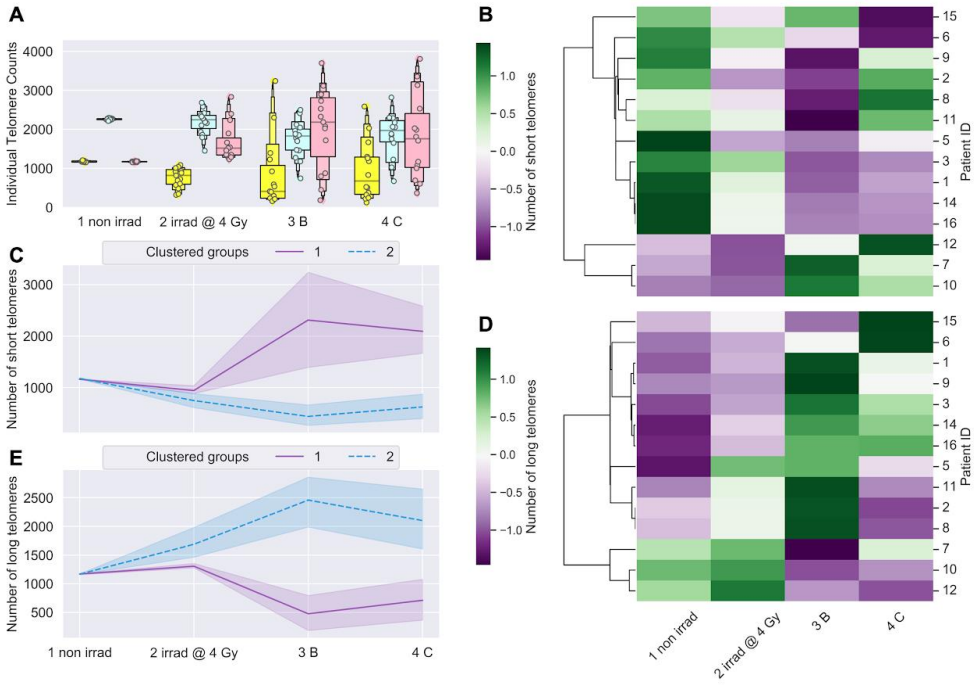

Figure 4.4. Longitudinal shifts in numbers of short and long telomeres (Telo-FISH). Numbers of short and long telomeres feature engineered from individual telomere length distributions per patient. 1 non irrad: pre-IMRT non-irradiated; 2 irrad @ 4 Gy: pre-IMRT in vitro irradiated; 3 B: immediate post-IMRT; 4 C: 3 months post-IMRT. Shortest (yellow), mid-length (blue), and longest (red) telomeres were feature engineered per patient (n=15) (see Materials and Methods). A) Counts of short, medium, and long telomeres, 4600 individual telomeres per patient (n=15) per time point. Significance was assessed for counts of short and long telomere using a square-root transformation and a repeated measures ANOVA with post-hoc Tukey's HSD test. Hierarchical clustering of patients by longitudinal changes in numbers of short telomeres B) and long telomeres D) (z-score normalized). Time-courses of patient groups (n=3, purple; n=11, blue) clustered by numbers of short C) and long E) telomeres; center lines denote medians and lighter bands denote confidence intervals. Patient ID 13 not clustered; 3 months post-IMRT sample failed to culture.

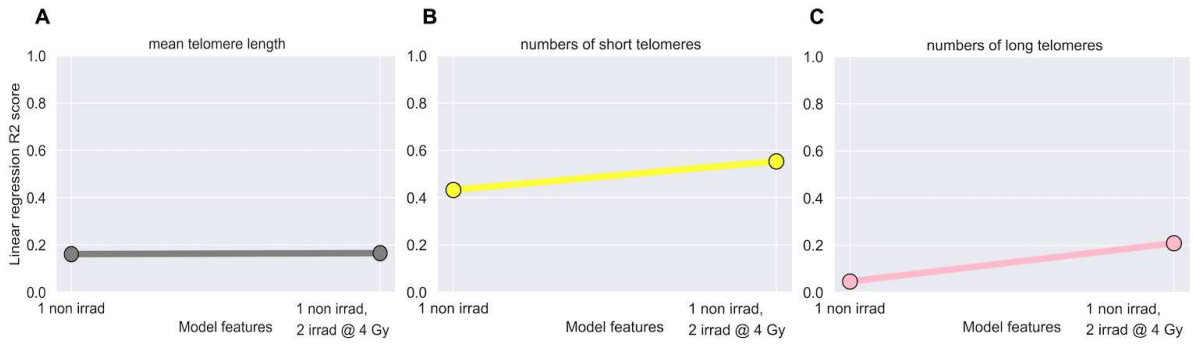

Figure 4.5. Linear regression models failed to predict post-IMRT telomeric outcomes. Ordinary least squares linear regression models were made using pre-IMRT telomeric data (Telo-FISH) from only the non-irradiated (1 non irrad) or also the in vitro irradiated (2 irrad @ 4 Gy) samples to predict 3 months post-IMRT telomeric outcomes. R2 values indicate the amount of variance in 3 months post-IMRT telomeric outcomes explained by the pre-therapy sample data. Models made using mean telomere length A) and numbers of short B) and long C) telomeres.

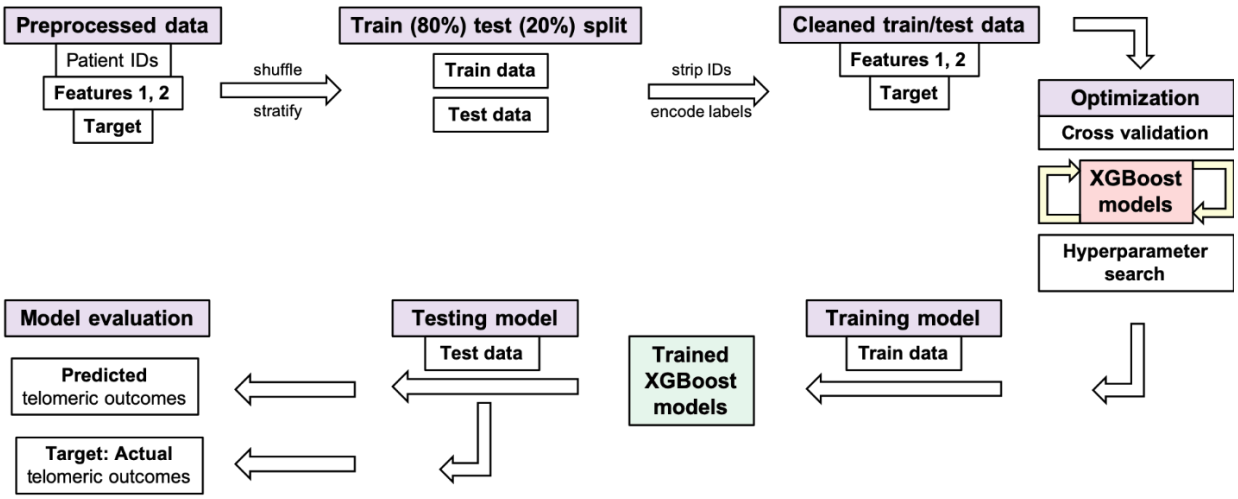

Figure 4.6. Processing of Telo-FISH data for training and testing XGBoost models. Schematic for our machine learning pipeline with individual telomere length data (Telo-FISH). Preprocessed data: Feature 1: pre-IMRT individual telomere lengths measurements (n=128,800); Feature 2: pre-IMRT sample labels (non-irradiated, *in vitro* irradiated, encoded as 0/1); Target: 3 months post-IMRT telomeric outcomes (mean telomere length or numbers of short and long telomeres). Data is randomly shuffled and stratified (by patient ID and pre-therapy sample origin) and split into training (80%) and testing (20%) datasets; patient IDs are stripped after splitting. Five-fold cross validation was used, and models were evaluated with Mean Absolute Error (MAE) and R2 scores between predicted and true values in the test set. See Materials and Methods and Code availability for model hyperparameters and implementations in Python.

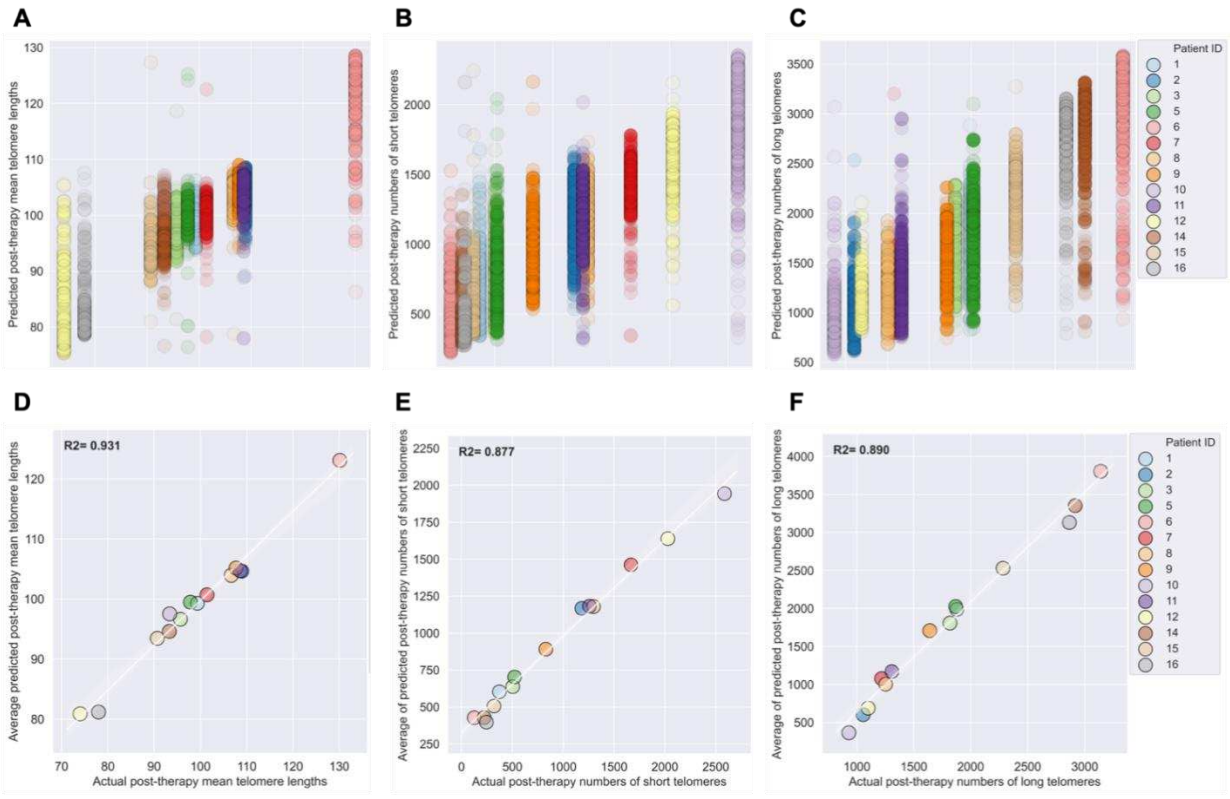

Figure 4.7. High performance of XGBoost models for predicting post-IMRT telomeric outcomes. Three separate XGBoost models were trained on pre-IMRT individual telomere length measurements (n=103,040, Telo-FISH) to predict 3 months post-IMRT telomeric outcomes. Trained XGBoost models were challenged with the test set (new data, n=25,760 individual telomeres) to predict 3 month post-IMRT telomeric outcomes for A) mean telomere length and numbers of B) short and C) long telomeres. XGBoost predictions were averaged on a per patient basis for mean telomere length D) and numbers of short E) and long F) telomeres; dark (white) line represents a simple regression line (X/Y), lighter bands the 95% confidence interval, R2 values (coefficient of determination) are noted in bold.

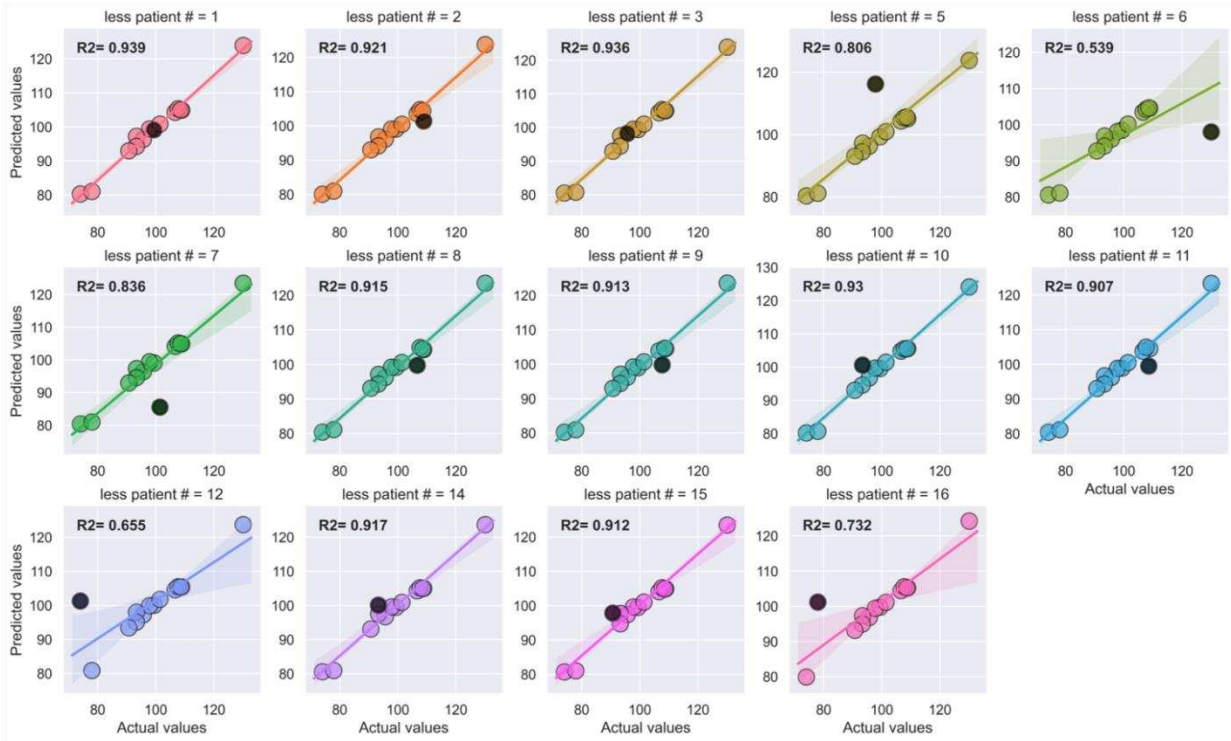

Figure 4.8. Strong generalizability of XGBoost models to new patient data (leave one out approach). A – N) Fourteen separate XGBoost models were iteratively trained on pre-IMRT individual telomere length measurements (n=93,840, Telo-FISH), excluding one patient, and tested to predict 3 months post-IMRT mean telomere length, with the inclusion of the patient excluded during training. Each panel is one model; patients excluded during training for that model are noted in the panel headers and plotted in black. Dark lines represents a simple regression line (X/Y), lighter bands the 95% confidence interval, R2 values (coefficient of determination) are noted in bold.

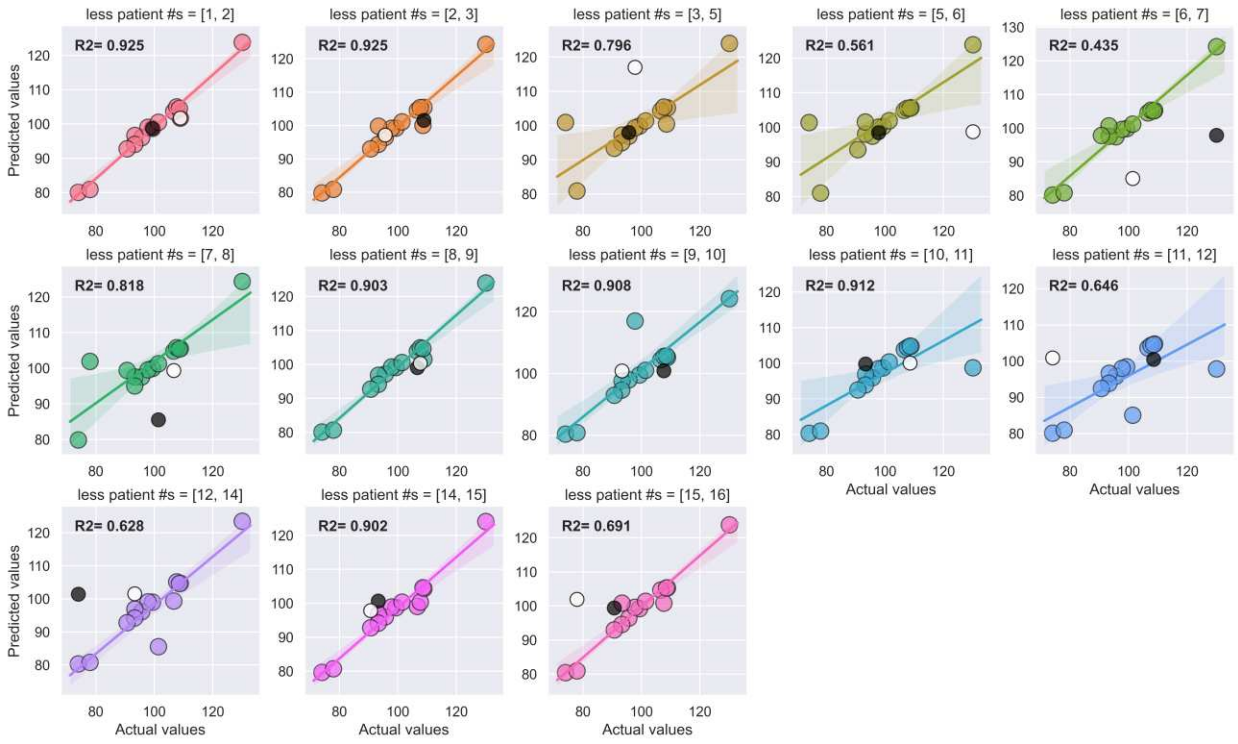

Figure 4.9. Strong generalizability of XGBoost models to new patient data (leave two out). **A – M**) Thirteen separate XGBoost models were iteratively trained on pre-IMRT individual telomere length measurements (n=84,640, Telo-FISH), excluding two patients, and tested to predict 3 months post-IMRT mean telomere length, with the inclusion of the two patients excluded during training. Each panel is one model; patients excluded during training for that model are noted in the panel headers and plotted in (black: index 1, white: index 2). Dark lines represents a simple regression line (X/Y), lighter bands the 95% confidence interval, R2 values (coefficient of determination) are noted in bold.

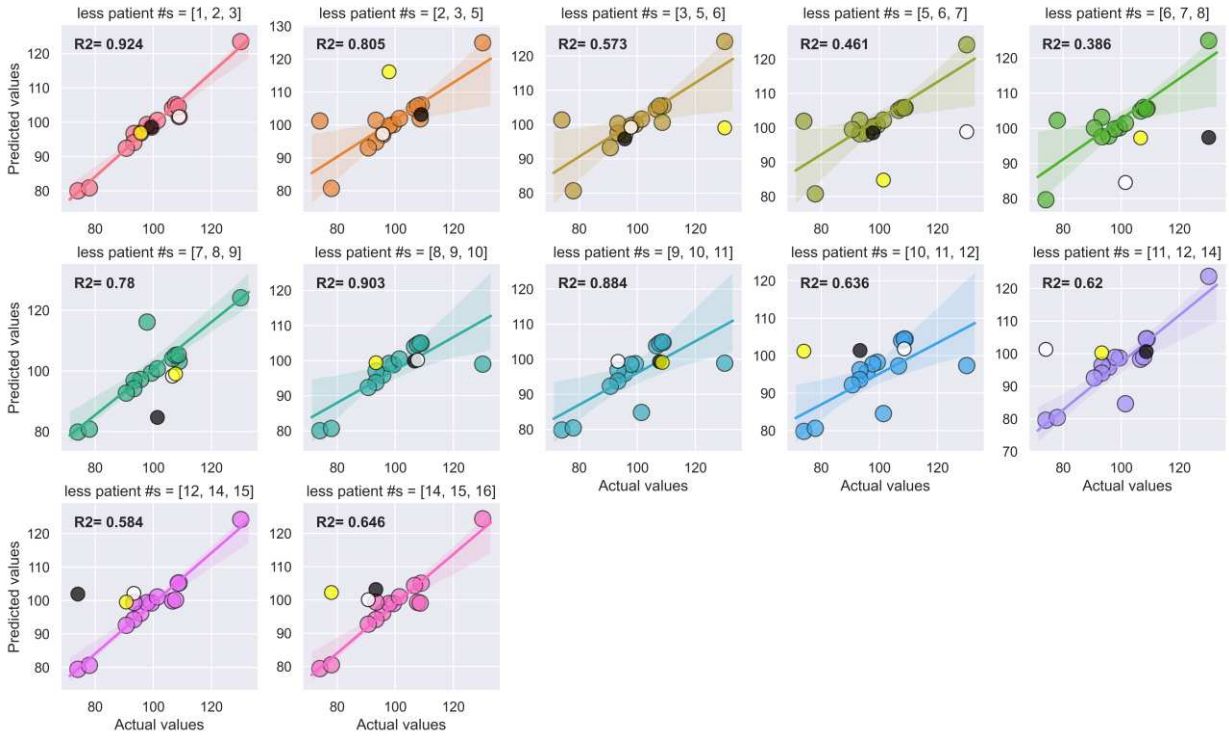

Figure 4.10. Strong generalizability of XGBoost models to new patient data (leave three out). A – L) Twelve separate XGBoost models were iteratively trained on pre-IMRT individual telomere length measurements (n=75,440, Telo-FISH), excluding three patients, and tested to predict 3 months post-IMRT mean telomere length, with the inclusion of the three patients excluded during training. Each panel is one model; patients excluded during training for that model are noted in the panel headers and plotted in (black: index 1, white: index 2, yellow: index 3). Dark lines represents a simple regression line (X/Y), lighter bands the 95% confidence interval, R2 values (coefficient of determination) are noted in bold.

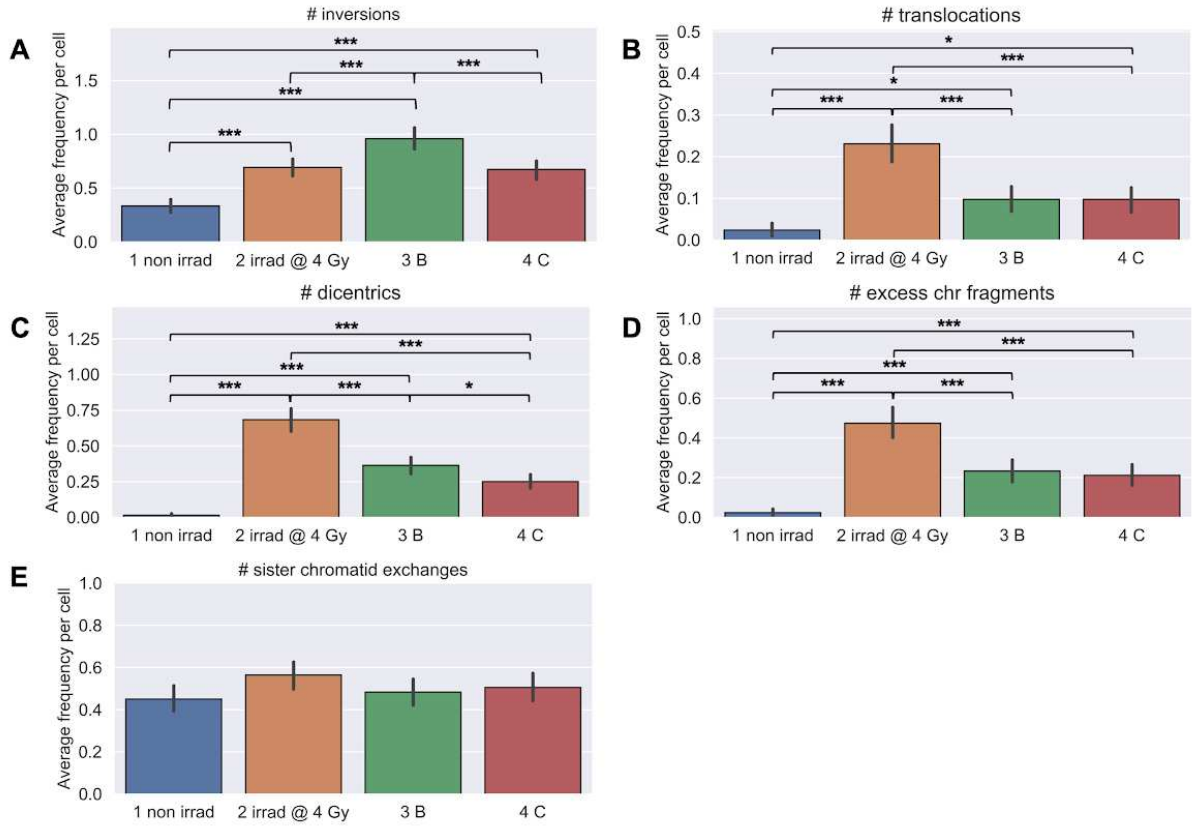

Figure 4.11. Longitudinal analyses of chromosomal instability (dGH). Chromosome aberrations were scored using directional Genomic Hybridization (dGH) in cultured T-cells harvested in metaphase (n=30/patient/timepoint) from whole blood of patients (n=15). 1 non irrad: pre-IMRT non-irradiated; 2 irrad @ 4 Gy: pre-IMRT in vitro irradiated; 3 B: immediate post-IMRT; 4 C: 3 months post-IMRT. Counts of inversions and translocations (A/B) were adjusted for clonality, where identical aberrations between cells are noted but scored only once. Excess chr fragments: counts of chromosome fragments per cell after subtracting 1 count per n observed dicentrics. A) inversions, B), translocations, C) dicentrics, D) chromosome fragment, and E) sister chromatid exchanges. Significance was assessed for average aberration frequencies using a repeated measures ANOVA and post-hoc Tukey's HSD test. $p<0.05$, $p<0.01$, $p<.001$ = *, **, ***.

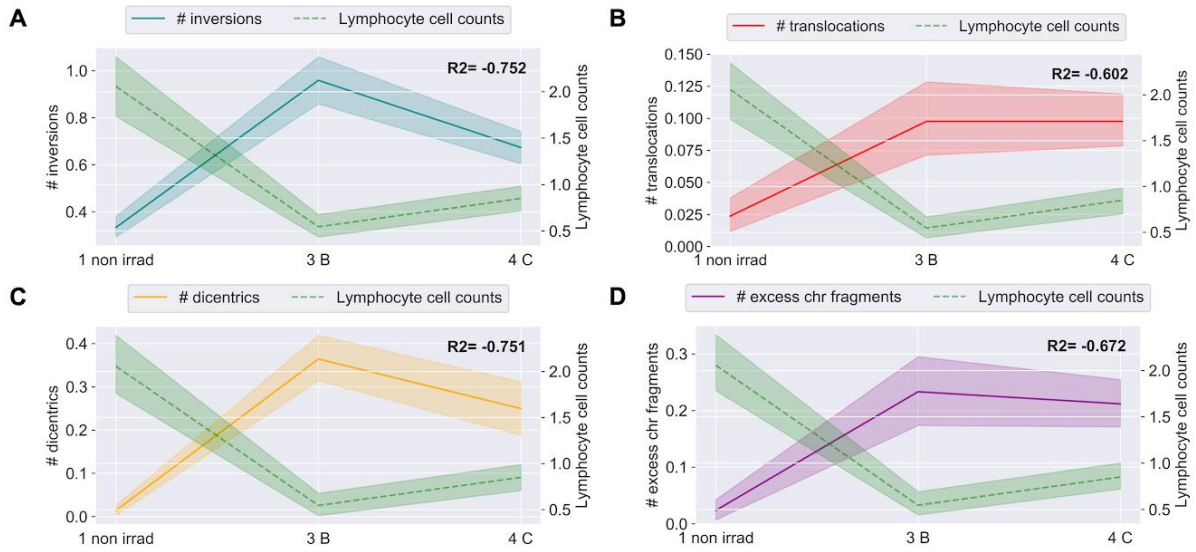

Figure 4.12. Correlations between chromosome aberrations and peripheral blood lymphocytes. Average frequencies of chromosome aberrations plotted longitudinally against lymphocyte cell counts (thousands per microliter) from complete blood count tests for all patients. 1 non irrad: pre-IMRT non-irradiated; 3 B: immediate post-IMRT; 4 C: 3 months post-IMRT. Excess chr fragments: counts of chromosome fragments per cell after subtracting 1 count per n observed dicentrics. Center lines denote medians, lighter bands denote confidence intervals. Pearson correlation R2 values were calculated between plotted values on a per patient basis and noted in bold on each graph. A) Inversions, B) translocations, C) dicentrics, D) chromosome fragments and lymphocyte cell counts.

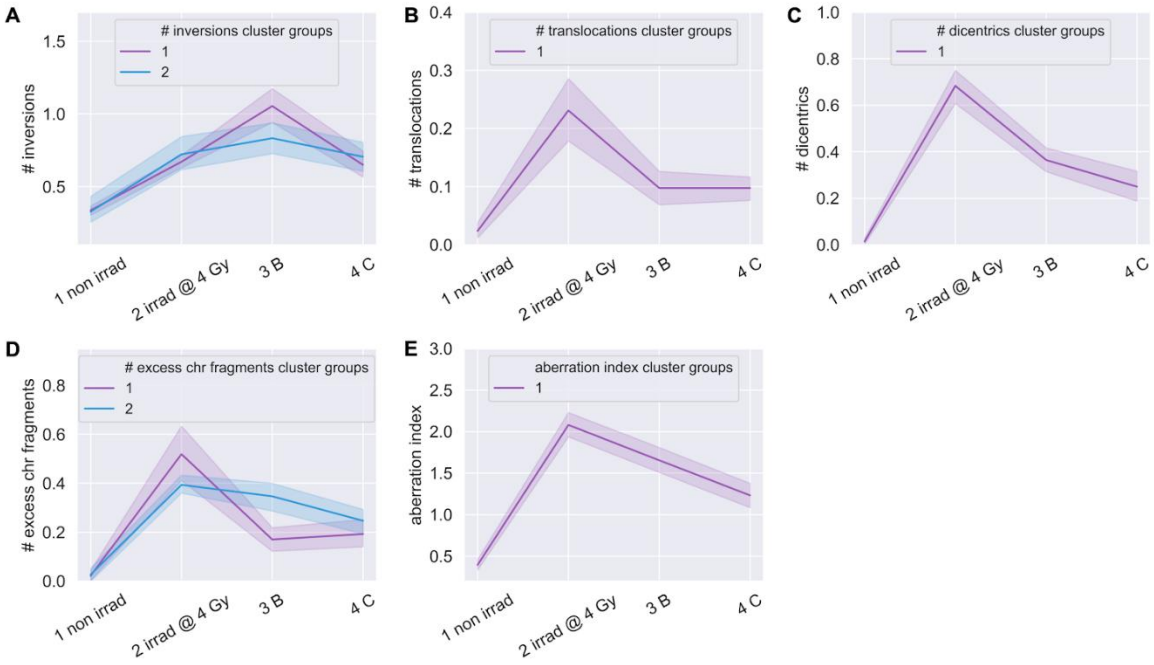

Figure 4.13. Clustering of patients by chromosome aberration frequencies. Time-courses for groups of patients hierarchically clustered into discrete groups (blue/purple) per aberration type. 1 non irrad: pre-IMRT non-irradiated; 2 irrad @ 4 Gy: pre-IMRT in vitro irradiated; 3 B: immediate post-IMRT; 4 C: 3 months post-IMRT. Excess chr fragments: counts of chromosome fragments per cell after subtracting 1 count per n observed dicentrics. Aberration index is created by summing all aberrations (A-D) per cell. Center lines denote medians and lighter bands denote confidence intervals. Clustered groups of patients for inversions A), translocations B), dicentrics C), chromosome fragments D), and aberration index E).

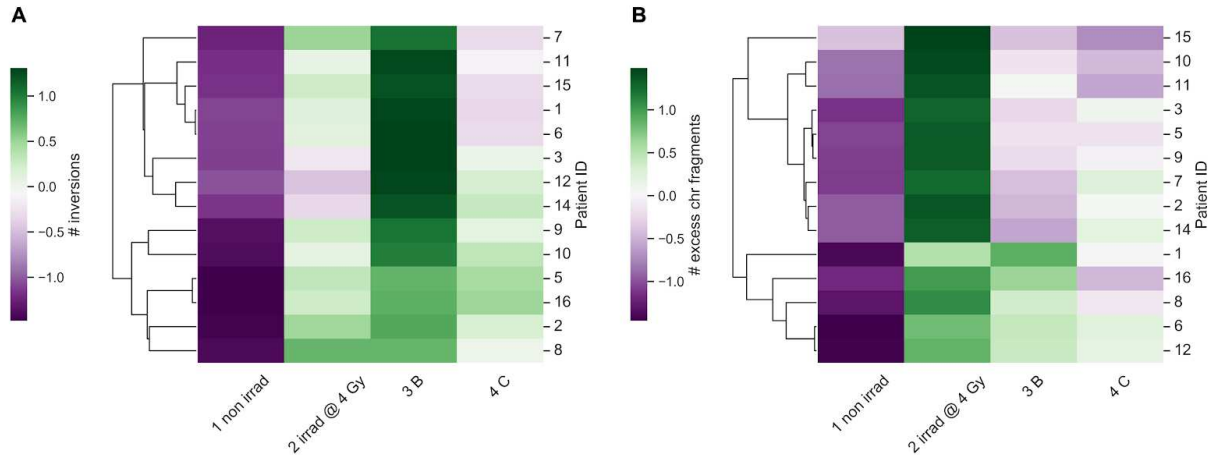

Figure 4.14. Clustering of patients by inversions and chromosome fragments (deletions). Hierarchical clustering of patients by longitudinal changes in chromosome aberrations scored by directional Genomic Hybridization (dGH). 1 non irrad: pre-IMRT non-irradiated; 2 irrad @ 4 Gy: pre-IMRT in vitro irradiated; 3 B: immediate post-IMRT; 4 C: 3 months post-IMRT. Excess chr fragments: counts of chromosome fragments per cell after subtracting 1 count per n observed dicentrics. Patients were clustered by inversions A) and chromosome fragments B) (z-score normalized). Patient ID 13 not clustered; 3 months post-IMRT sample failed to culture.

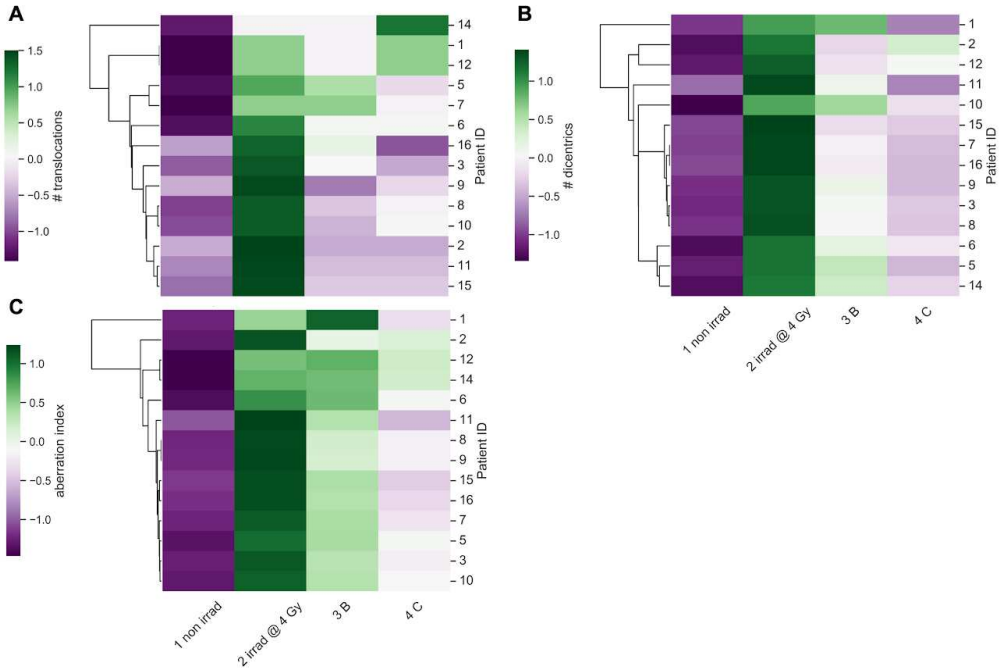

Figure 4.15. Chromosome aberrations generally failed to cluster patients. Hierarchical clustering of patients by longitudinal changes in chromosome aberrations scored by directional Genomic Hybridization (dGH). 1 non irrad: pre-IMRT non-irradiated; 2 irrad @ 4 Gy: pre-IMRT in vitro irradiated; 3 B: immediate post-IMRT; 4 C: 3 months post-IMRT. Aberration index is created by summing all aberrations (inversions, translocations, dicentrics, chromosome fragments) per cell. Patients were clustered by translocations A), dicentrics B), and aberration index C) (z-score normalized). Patient ID 13 not clustered; 3 months post-IMRT sample failed to culture.

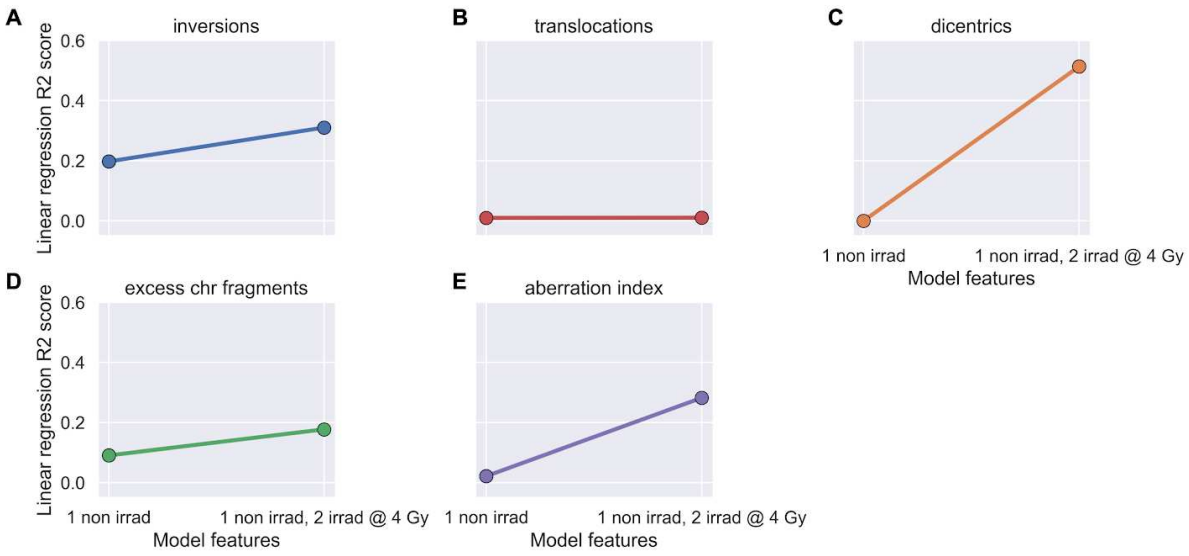

Figure 4.16. Linear regression models failed to predict post-IMRT chromosome aberration frequencies. Ordinary least squares linear regression models were made using pre-IMRT average aberration frequencies from only the non-irradiated (1 non irrad) or also the in vitro irradiated (2 irrad @ 4 Gy) samples to predict 3 months post-IMRT average aberration frequencies. Excess chr fragments: counts of chromosome fragments per cell after subtracting 1 count per n observed dicentrics. Aberration index is created by summing all aberrations (A-D) per cell. R2 values indicate the amount of variance in late post-IMRT outcomes explained by the pre-IMRT sample data. Models made with inversions A), translocations B), dicentrics C), chromosome fragments D), and aberration index E).

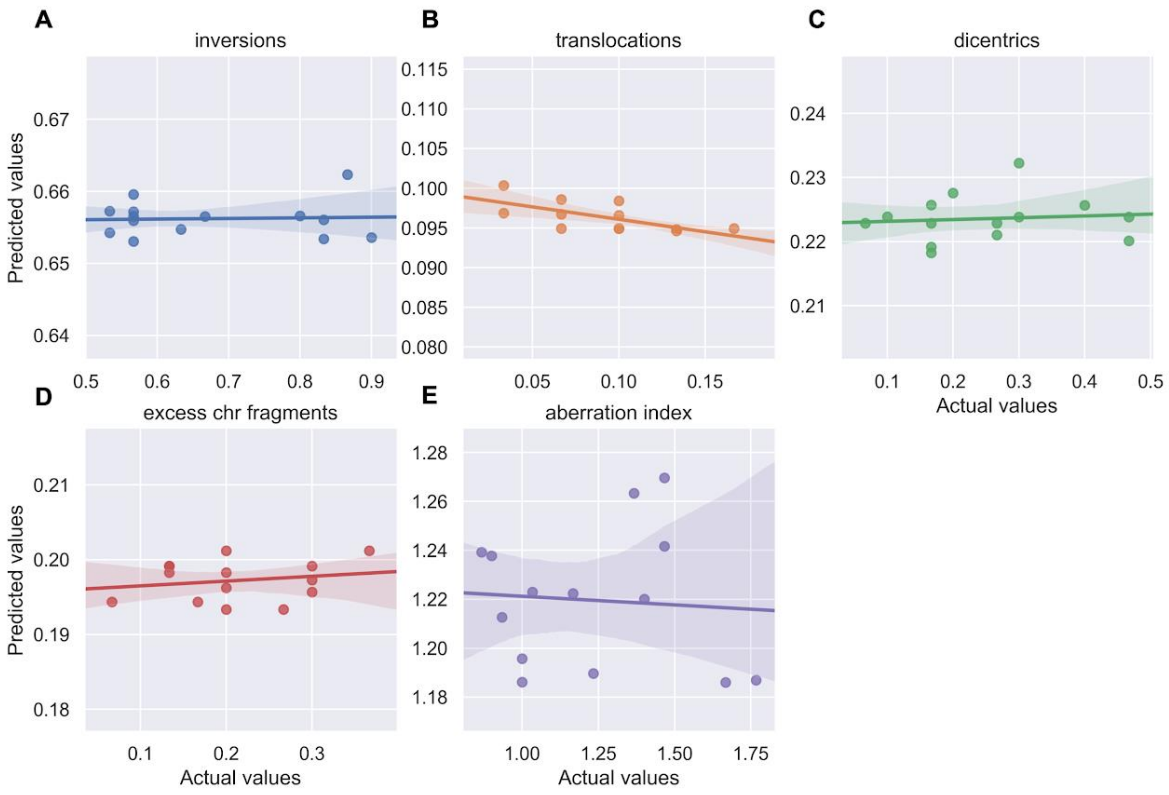

Figure 4.17. XGBoost models failed to predict post-IMRT chromosome aberration frequencies. XGBoost models were trained on pre-IMRT counts of different chromosome aberration types per cell (n=672) to predict 3 months post-IMRT average chromosome aberration frequencies. Trained XGBoost models were challenged with the test set (new data, n=168 cells) to predict 3 months post-IMRT average chromosome aberration frequencies. Excess chr fragments: counts of chromosome fragments per cell after subtracting 1 count per n observed dicentrics. Aberration index is created by summing all aberrations (A-D) per cell. XGBoost predictions were averaged on a per patient basis for inversions A), translocations B), dicentrics C), chromosome fragments D), and aberration index E). For all models, R2 values between averaged predictions and actual values did not exceed 0.100.

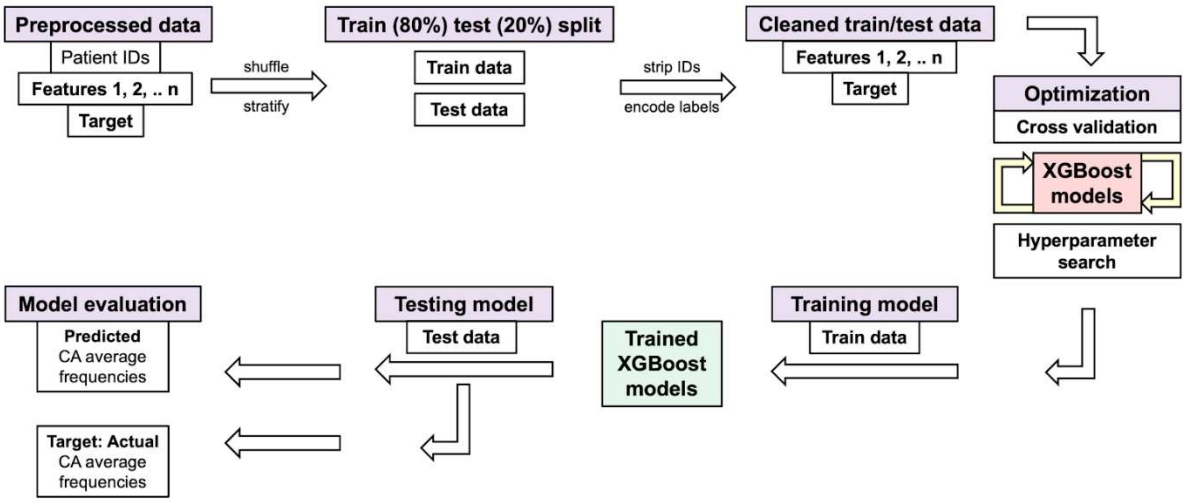

Figure 4.18. Processing of chromosome aberration data for XGBoost models. Schematic for machine learning pipeline using chromosome aberration (CAs) data. Preprocessed data: Feature 1: pre-IMRT counts of scored CAs; Feature 2: pre-IMRT sample labels (non-irradiated, in vitro irradiated, encoded as 0/1); Feature n: represents pre-IMRT counts of multiple types of CAs (for aberration index). Target: Late post-IMRT average frequencies of CAs (either specific aberration type or aberration index). Data is randomly shuffled and stratified (by patient ID and pre-IMRT sample origin) and split into training (80%) and testing (20%) datasets; patient IDs are stripped after splitting. Five-fold cross validation was used, and models were evaluated with Mean Absolute Error (MAE) and R^2 between predicted and true values in the test set. See Materials and Methods and Code availability for model parameters and implementations in Python.

Tables

Table 4.1. Example views of individual telomere length data matrices used to train XGBoost models. XGBoost models were trained on 103,040 individual telomere length measurements (one telomere per row) (Telo-FISH) from pre-IMRT non-irradiated (1 non irrad) and in vitro irradiated (2 irrad @ 4 Gy) samples to predict 3 months post-IMRT (4 C) telomeric outcomes. Matrices represent examples of pre- (A/C/E) and post-processed (B/D/F) training data. Patient IDs are stripped after data is shuffled and stratified. The 'encoded sample origin' column contains numerical encodings denoting individual telomeres' pre-IMRT sample of origin (0: non-irradiated, 1: in vitro irradiated). XGBoost models were trained to predict mean telomere length (A/B) and numbers of short (C/D) and long (E/F) telomeres at 3 months post-IMRT with data in the format as shown.

A

patient id	pre-therapy sample origin	individual telomeres (RFI)	4 C telo means
1	1 non irrad	52.79329603949808	99.34629891451401
1	2 irrad @ 4 Gy	100.30726247504634	99.34629891451401
1	1 non irrad	59.12849156423784	99.34629891451401
1	2 irrad @ 4 Gy	106.64139157520613	99.34629891451401
1	1 non irrad	69.68715077213746	99.34629891451401
1	2 irrad @ 4 Gy	107.69724693733689	99.34629891451401

B

encoded sample origin	individual telomeres (RFI)	4 C telo means
1.0	71.84704355757034	90.6803515449468
0.0	58.01948086775996	108.91532697997721
0.0	125.05216035895008	93.35225326745208
0.0	99.84003125432304	93.35225326745208
0.0	157.34096506511176	108.91532697997721
1.0	59.127900279322205	99.34629891451401

C

patient id	pre-therapy sample origin	individual telomeres (RFI)	4 C # short telos
1	1 non irrad	52.79329603949808	372
1	2 irrad @ 4 Gy	100.30726247504634	372
1	1 non irrad	59.12849156423784	372
1	2 irrad @ 4 Gy	106.64139157520613	372
1	1 non irrad	69.68715077213746	372
1	2 irrad @ 4 Gy	107.69724693733689	372

D

encoded sample origin	individual telomeres (RFI)	4 C # short telos
0.0	39.80714575487005	319.0
0.0	84.7669312523909	2028.0
0.0	48.569832356338225	372.0
1.0	99.34779587017763	829.0
1.0	104.85784735429183	319.0
1.0	92.25878757956735	124.0

E

patient id	pre-therapy sample origin	individual telomeres (RFI)	4 C # long telos
1	1 non irrad	52.79329603949808	1987
1	2 irrad @ 4 Gy	100.30726247504634	1987
1	1 non irrad	59.12849156423784	1987
1	2 irrad @ 4 Gy	106.64139157520613	1987
1	1 non irrad	69.68715077213746	1987
1	2 irrad @ 4 Gy	107.69724693733689	1987

F

encoded sample origin	individual telomeres (RFI)	4 C # long telos
0.0	56.551567152220926	2026.0
0.0	103.18180673387677	2026.0
0.0	69.58478047400733	365.0
0.0	56.18104859876975	1078.0
1.0	137.72889825629034	1002.0
1.0	84.46927366319693	1987.0

Table 4.2. Metrics of XGBoost models for predicting post-IMRT telomeric outcomes. XGBoost models were trained on pre-IMRT individual telomere length measurements (Telo-FISH) to predict 3 months post-IMRT telomeric outcomes. Metrics assess model performance during (five) cross-fold validation (CV) (columns 1-2 from left) and when challenged with the test set (test) (columns 3-4 from left). Model performance was evaluated with mean absolute error (MAE) (std dev: standard deviation) across a range of samples in the training data (n=100 to 103,040). R^2: correlation metric. Metrics of XGBoost models for predicting 3 months post-IMRT (4 C) mean telomere length A), numbers of short B) and long C) telomeres.

A

Average MAE of CV folds	Std dev of MAE of CV folds	MAE predicted vs. test values	R2 predicted vs. test values	N samples training data
11.4602	1.6502	13.4903	-0.8393	100.0
10.6657	0.4454	10.3646	-0.2049	500.0
8.0423	0.486	7.9009	0.1788	1000.0
6.7089	0.3895	6.0449	0.5126	2000.0
4.8488	0.2224	4.642	0.7094	4000.0
3.9282	0.0988	3.7677	0.8215	8000.0
3.6385	0.0447	3.5413	0.851	16000.0
3.3792	0.0626	3.3483	0.8755	32000.0
3.2944	0.051	3.2521	0.881	64000.0
3.233	0.052	3.2596	0.8817	103040.0

B

Average MAE of CV folds	Std dev of MAE of CV folds	MAE predicted vs. test values	R2 predicted vs. test values	N samples training data
705.0956	48.4789	680.9499	-0.5887	100.0
573.3922	25.5422	521.0982	-0.0162	500.0
440.9283	22.7264	425.5251	0.2572	1000.0
386.4338	19.0126	326.1635	0.5396	2000.0
315.2925	5.9607	292.0579	0.6593	4000.0
269.2991	6.6633	260.4209	0.7433	8000.0
257.6623	3.5097	247.6769	0.7747	16000.0
243.5729	4.1386	241.8505	0.7987	32000.0
233.7408	5.251	231.1663	0.803	64000.0
236.2825	2.0593	234.1744	0.8112	103040.0

C

Average MAE of CV folds	Std dev of MAE of CV folds	MAE predicted vs. test values	R2 predicted vs. test values	N samples training data
1056.6558	219.1554	953.2471	-0.4405	100.0
763.7998	38.9092	727.2706	0.0447	500.0
629.6607	49.9928	627.9304	0.2945	1000.0
548.353	24.8756	481.5782	0.5641	2000.0
409.3232	4.8234	415.0895	0.674	4000.0
382.1325	11.974	376.8821	0.7505	8000.0
353.0249	6.234	348.5064	0.7981	16000.0
343.0401	4.5386	329.2967	0.8128	32000.0
331.0765	3.7999	331.8519	0.813	64000.0
330.3521	2.0857	335.931	0.8191	103040.0

Table 4.3. Example views of chromosome aberration data matrices used to train XGBoost models. XGBoost models were trained on chromosome aberration count data (one cell per row, n=672) from pre-IMRT non-irradiated (1 non irrad) and in vitro irradiated (2 irrad @ 4 Gy) samples to predict 3 months post-IMRT (4 C) chromosome aberration frequencies. Matrices represent pre- (A/C) and post-processed (B/D) training data. Patient IDs are stripped after data is shuffled and stratified. The 'encoded sample origin' column contains numerical encodings denoting cells' pre-IMRT sample of origin (0: non-irradiated, 1: in vitro irradiated). XGBoost models shown were trained to predict average inversion frequencies (A/B) and aberration index frequencies (C/D).

A

patient id	pre-therapy sample origin	# inversions	4 C # inversions
5	1 non irrad	0	0.7083333333333334
11	2 irrad @ 4 Gy	2	0.4583333333333333
1	1 non irrad	0	0.5
9	1 non irrad	0	0.7083333333333334
11	1 non irrad	0	0.4583333333333333
16	1 non irrad	0	0.7916666666666666

B

encoded sample origin	# inversions	4 C # inversions
0.0	0.0	0.7083333333333334
1.0	0.0	0.5
1.0	0.0	0.5
0.0	1.0	0.7083333333333334
1.0	0.0	0.5
1.0	1.0	0.7916666666666666

C

patient id	pre-therapy sample origin	# inversions	# translocations	# dicentrics	# excess chr fragments	4 C aberration index
9	1 non irrad	0	0	0	0	1.125
7	2 irrad @ 4 Gy	1	0	1	1	0.8333333333333334
11	2 irrad @ 4 Gy	0	1	0	0	0.6666666666666666
1	1 non irrad	0	0	0	0	0.9583333333333334
16	1 non irrad	0	0	0	0	1.0833333333333333
6	1 non irrad	0	0	0	0	1.375

D

encoded sample origin	# inversions	# translocations	# dicentrics	# excess chr fragments	4 C aberration index
0.0	0.0	0.0	0.0	0.0	1.0833333333333333
0.0	0.0	0.0	0.0	0.0	0.9583333333333334
1.0	0.0	0.0	0.0	0.0	1.4166666666666667
0.0	0.0	0.0	0.0	0.0	1.2083333333333333
1.0	2.0	0.0	0.0	0.0	1.2083333333333333
0.0	0.0	0.0	0.0	0.0	0.9166666666666666

Table 4.4. Metrics of trained XGBoost models for predicting post-IMRT average frequencies of chromosome aberrations. Multiple iterations of XGBoost models (A-C) were trained on pre-IMRT chromosome aberration counts per cell (n=672 cells) to predict late post-IMRT average chromosome aberration frequencies. Time points for pre-IMRT data were encoded (0/1: non-irradiated, in vitro irradiated). Metrics assess model performance during (five) cross-fold validation (CV) and when challenged with the test set (test). Model performance was evaluated with mean absolute error (MAE) (std dev: standard deviation). R^2: correlation metric. Performance of models with identical initializations and hyperparameters for predicting average frequencies of inversions , translocations, dicentrics, chromosome fragments, and aberration index are shown (A-C).

A

Features	Target	Average MAE of CV folds	Std dev of MAE of CV folds	MAE predicted vs. test values	R2 predicted vs. test values
# inversions, encoded samples	4 C # inversions	0.1746	0.0445	0.2724	-0.213
# translocations, encoded samples	4 C # translocations	0.0412	0.0188	0.1327	-0.3905
# dicentrics, encoded samples	4 C # dicentrics	0.1171	0.0461	0.2508	0.0019
# excess chr fragments, encoded samples	4 C # excess chr fragments	0.0939	0.0334	0.1787	-0.1228
all aberrations, encoded samples	4 C aberration index	0.2541	0.0496	0.5137	-0.05

B

Features	Target	Average MAE of CV folds	Std dev of MAE of CV folds	MAE predicted vs. test values	R2 predicted vs. test values
# inversions, encoded samples	4 C # inversions	0.1759	0.0332	0.2187	-1.5965
# translocations, encoded samples	4 C # translocations	0.0375	0.0096	0.1096	-0.004
# dicentrics, encoded samples	4 C # dicentrics	0.1167	0.0463	0.2023	-0.0215
# excess chr fragments, encoded samples	4 C # excess chr fragments	0.084	0.0103	0.202	-0.1071
all aberrations, encoded samples	4 C aberration index	0.3294	0.1258	0.3596	-0.0256

C

Features	Target	Average MAE of CV folds	Std dev of MAE of CV folds	MAE predicted vs. test values	R2 predicted vs. test values
# inversions, encoded samples	4 C # inversions	0.15	0.0573	0.1977	-0.0709
# translocations, encoded samples	4 C # translocations	0.0448	0.0145	0.0977	-0.0346
# dicentrics, encoded samples	4 C # dicentrics	0.1123	0.034	0.2177	-0.0558
# excess chr fragments, encoded samples	4 C # excess chr fragments	0.0681	0.0155	0.194	-0.0372
all aberrations, encoded samples	4 C aberration index	0.3255	0.0528	0.5078	-0.0259

Data availability

Raw and processed individual telomere length (Telo-FISH) data files and chromosome aberration score sheets (dGH) are available for download at https://github.com/Jared-Luxton/. All data processing pipelines and code was written in Python and stored in Jupyter notebooks at https://github.com/Jared-Luxton/. The Jupyter notebooks can be run within a web browser and are available for download.

Materials and Methods

Patient consent, IMRT therapy information

With informed consent as per the institutional review board, 16 consecutive patients that were receiving pelvis and prostate or prostate fossa radiation therapy were asked to participate. No patient had received androgen ablation or chemotherapy to avoid confounding factors. One patient was found to have metastatic disease after consent and was removed from further study. A total of 15 patients provided consent and blood was obtained at pre-IMRT (baseline), immediately post-IMRT (the last week) and 3 months post-IMRT. Blood was subject to complete blood counts, and telomere length and chromosome aberration analyses. Radiation consisted of 54 Gy to the pelvic lymphatics, with a total of 70 Gy (n=11) or 78 Gy (n=3) to the prostate fossa. One patient underwent brachytherapy boost.

Sample collection and processing for Telo-FISH and dGH

Peripheral blood was drawn and shipped in 10 mL sodium heparin tubes (Becton, Dickinson, and Co #367874) under ambient conditions to Colorado State University and received within 24 hours of blood draw. All heparinized blood samples were cultured in T-25 tissue culture flasks,

at 1 parts blood per 9 parts Gibco PB-Max Karyotyping Medium (ThermoFisher #12557013), with 5.0 mM 5-bromo-deoxyuridine (BrdU) and 1.0 mM 5-bromo-deoxycytidine (BrdC) added to the medium as previously described (Ray et al., 2013). Pre-IMRT blood samples were split into two fractions (non-irradiated and *in vitro* irradiated) with identical culturing conditions as other time point samples, and one fraction was irradiated in a Cs137 irradiator *in vitro* at a dose rate of 2.5 Gy/min for a total dose of 4 Gy (γ-rays). 48 hours after stimulation, KaryoMax Colcemid (ThermoFisher #15210040) was added (0.1 μg per mL of medium) for four hours of incubation, then metaphase chromosome spreads were harvested with standard cytogenetic protocols (Howe et al., 2014). Prior to Telo-FISH and dGH, slides with metaphase chromosome spreads were subject to CO-FISH for removal of BrdU/BrdC incorporated DNA as previously described (Williams and Bailey, 2009).

Telomere Fluorescence *in situ* Hybridization (Telo-FISH), imaging, quantifications

Protocol: Slides with metaphase chromosome spreads were prepared and hybridized with a fluorescently labeled telomere probe as previously described (Poon and Lansdorp, 2001). Briefly, slides were washed in 1x PBS for 5 min, dehydrated with an ice-cold ethanol series (75%, 85% and 100%) for 2 min each, air dried, and denatured in 70% formamide in 2x saline sodium citrate (SSC) at 75°C for 2 min, followed by a second ice-cold ethanol series, and air dried again. Probe hybridization mixture consisted of G-rich (TTAGGG-'3) peptide nucleic acid (PNA) telomere probe labeled with Cyanine-3 (Cy3; Biosynthesis) at a 5nM concentration in 36 μL of formamide, 12 μL of 0.5 M Tris-HCl, 2.5 μL of 0.1 M KCl, and 0.6 μl of 0.1 M MgCl2. Hybridization mixture was incubated at 75°C for 5 min and cooled on ice for 10 min, then 50 μL of mix was applied to each slide. Slides were coverslipped and hybridized at 37°C for 4 h. After

hybridization, slides were washed five times at 43.5°C for three min each: washes one and two: 50% formamide in 2xSSC; washes three and four: 100% 2xSSC; and washes five and six: 2xSSC plus 0.1% Nonidet P-40. After washing, slides were counterstained with one drop of DAPI in Prolong Gold Antifade (ThermoFisher #P36931), coverslipped, and stored at 4°C for 24 h prior to imaging.

Image acquisition: Metaphase spreads (50 per patient/time point) were imaged at 100x mag on a Zeiss Axio Imager.Z2, Cool SNAP ES2 camera, and X-cite 120 LED lamp lightsource.

Individual telomere quantifications: Relative fluorescence intensity of individual telomeres was quantified using the ImageJ (Schneider et al., 2012) plugin Telometer (https://demarzolab.pathology.jhmi.edu/telometer/). Variation in Telo-FISH was controlled by assigning each patient a pair of slides made from BJ1 primary cells and BJ-hTERT cell lines. For each patient the slide preparation, Telo-FISH protocol, image acquisition and telomere quantifications were performed on the full time-course of samples and a pair of BJ1/BJ-hTERT controls (50 metaphases per control) at the same time and on the same respective days. Mean telomere length was quantified for each pair of control samples yielding a ratio for standardizing patients' telomere values as previously described (Wong and Slijepcevic, 2004).

Telo-FISH data processing, feature engineering of short and long telomeres

Processing individual telomere length data: For each patient, outliers were removed from individual telomere length data per sample by omitting measurements three standard deviations from the mean. For samples with fewer individual telomere length measurements than the theoretical number (human cells, 50 metaphase spreads), missing telomere values were imputed by randomly sampling measurements from the observed distribution of individual telomeres;

randomly sampled telomeres were added up to the theoretical number of telomeres per sample.

Feature engineering short and long telomeres: Individual telomeres from the pre-IMRT non-irradiated time point were split into quartiles, designating telomeres in the bottom 25% in yellow, the middle 50% in blue, and top 25% in red. Quartile cut-off values, established by the pre-IMRT non-irradiated sample's distribution (values that separate quartiles), were applied to subsequent time points to feature engineer the relative shortest (yellow), mid-length (blue), and longest (red) individual telomeres per time point.

Statistical and clustering analyses of Telo-FISH data

Statistical and clustering analyses were conducted with Python in Jupyter notebooks (see Code availability). With the statsmodels library (Seabold and Perktold, 2010), mean telomere length and numbers of short and long telomeres were analyzed with a repeated measures ANOVA and post-hoc Tukey's HSD test (two-tailed p values for both tests). Analyses were performed on all patients (n=14, less patient ID 13; 3 months post-IMRT sample failed to culture) and all four time course samples. A square root transformation was performed on numbers of short and long telomeres prior to statistical analysis. Ordinary least squares linear regression was performed with the scikit-learn LinearRegression tool. Hierarchical clustering analyses were performed on z-score normalized data using the scipy library with a single linkage method and Pearson correlation metric. Pearson correlations between patients' longitudinal measurements of telomere length and complete blood count data was done with Python.

XGBoost models with individual telomere length data, randomized hyperparameter search, cross validation

XGBoost models, model hyperparameter tuning, and cross validation tools were performed in Python through the scikit-learn API (Pedregosa et al., 2018). XGboost model features were individual telomere length values and sample labels denoted pre-IMRT sample origin (non-irradiated, *in vitro* irradiated), which were encoded as 0/1. Model hyperparameters were tuned using a randomized search with RandomizedSearchCV. For models predicting mean telomere length at late post-IMRT, final model hyperparameters were modified as follows: n_estimators=200, max_depth=7, learning_rate=0.2, objective ='reg:squarederror', random_state=1. For models predicting short and long telomeres at late post-IMRT, final model hyperparameters were similar as for mean telomere length, with max_depth=6. Five-fold cross validation was performed with cross_val_score and a negative mean absolute error metric.

directional Genomic Hybridization (dGH), image acquisition, data processing

Protocol: High-resolution detection of chromosome aberrations (inversions, translocations) was performed with directional Genomic Hybridization (dGH) whole chromosome (Cy3) and subtelomere (Cy5) paints to chromosomes 1, 2, and 3 (KromaTiD Inc.) as previously described(Ray et al., 2013). Briefly, slides were submersed in Hoechst 33258 (Millipore Sigma #B1155) for 15 min, photolyzed for 35 min using a SpectroLinker UV Crosslinker (365 nm UV), and treated with exonuclease III (New England Biolabs #M0206L) for 30 min. Paint hybridization mixture was applied to slides, which were then coverslipped, sealed with rubber cement, and denatured at 70°C for three min. Slides were hybridized for 24 h at 37°C, followed by five washes in 2xSSC at 43.5°C. After washing, slides were counterstained with one drop of DAPI in Prolong Gold Antifade (ThermoFisher #P36931), coverslipped, and stored at 4°C for 24 h prior to imaging.

Image acquisition: Metaphase spreads (30 per patient/time point) were imaged/scored at 63x mag on a Zeiss Axio Imager.Z2, Cool SNAP ES2 camera, and X-cite 120 LED lamp lightsource. **Data processing:** Counts of chromosome aberrations were adjusted for clonality, where identical aberrations between cells for a patient's given time point were noted but scored only once.

Statistical and clustering analyses of chromosome aberrations (dGH)

Statistical and clustering analyses were conducted with Python in Jupyter notebooks (see Code availability). With the statsmodels library, average chromosome aberration frequencies were analyzed with a repeated measures ANOVA and post-hoc Tukey's HSD test (two-tailed p values for both tests). Analyses were performed on all patients (n=14, less patient ID 13; 3 months post-IMRT sample failed to culture) and all time course samples (4). Ordinary least squares linear regression was performed with the scikit-learn LinearRegression tool. Hierarchical clustering analyses were performed on z-score normalized data using the scipy library with a single linkage method and Pearson correlation metric. Pearson correlations between patients' longitudinal measurements of average chromosome aberration frequencies and complete blood count data was done with Python.

XGBoost model design with chromosome aberrations

XGBoost models, model hyperparameter tuning, and cross validation were accessed in Python via the same manner as described for Telo-FISH data above. XGboost model features were counts of scored chromosome aberrations per cell, with sample labels denoting pre-IMRT sample origin (non-irradiated, *in vitro* irradiated; encoded as 0/1). Model hyperparameters were tuned using a randomized search with RandomizedSearchCV; models were ultimately non-

performant. Final model hyperparameters (used with all chromosome aberrations) were: n_estimators=200, max_depth=15, learning_rate=0.1, objective='reg:squarederror', random_state=0. Five-fold cross validation was performed with a negative mean absolute error metric.

REFERENCES

Adams, M.J., Hardenbergh, P.H., Constine, L.S., and Lipshultz, S.E. (2003). Radiation-associated cardiovascular disease. Crit. Rev. Oncol. Hematol. *45*, 55–75.

Alsner, J., Rødningen, O.K., and Overgaard, J. (2007). Differential gene expression before and after ionizing radiation of subcutaneous fibroblasts identifies breast cancer patients resistant to radiation-induced fibrosis. Radiother. Oncol. *83*, 261–266.

Andreassen, C.N., Overgaard, J., and Alsner, J. (2013). Independent prospective validation of a predictive test for risk of radiation induced fibrosis based on the gene expression pattern in fibroblasts irradiated in vitro. Radiother. Oncol. J. Eur. Soc. Ther. Radiol. Oncol. *108*, 469–472.

Armanios, M., and Blackburn, E.H. (2012). The telomere syndromes. Nat. Rev. Genet. *13*, 693–704.

Armstrong, G.T., Stovall, M., and Robison, L.L. (2010). Long-Term Effects of Radiation Exposure among Adult Survivors of Childhood Cancer: Results from the Childhood Cancer Survivor Study. Radiat. Res. *174*, 840–850.

Arsenis, N.C., You, T., Ogawa, E.F., Tinsley, G.M., and Zuo, L. (2017). Physical activity and telomere length: Impact of aging and potential mechanisms of action. Oncotarget *8*, 45008–45019.

Aubert, G., and Lansdorp, P.M. (2008). Telomeres and aging. Physiol. Rev. *88*, 557–579.
Ayouaz, A., Raynaud, C., Heride, C., Revaud, D., and Sabatier, L. (2008). Telomeres: hallmarks of radiosensitivity. Biochimie *90*, 60–72.

Baeyens, A., Thierens, H., Claes, K., Poppe, B., Messiaen, L., De Ridder, L., and Vral, A. (2002). Chromosomal radiosensitivity in breast cancer patients with a known or putative genetic predisposition. Br. J. Cancer *87*, 1379–1385.

Bains, S.K., Chapman, K., Bright, S., Senan, A., Kadhim, M., and Slijepcevic, P. (2019). Effects of ionizing radiation on telomere length and telomerase activity in cultured human lens epithelium cells. Int. J. Radiat. Biol. *95*, 54–63.

Ballarini, F., Altieri, S., Bortolussi, S., Carante, M., Giroletti, E., and Protti, N. (2015). The role of DNA cluster damage and chromosome aberrations in radiation-induced cell killing: a theoretical approach. Radiat. Prot. Dosimetry *166*, 75–79.

Banáth, J.P., MacPhail, S.H., and Olive, P.L. (2004). Radiation Sensitivity, H2AX Phosphorylation, and Kinetics of Repair of DNA Strand Breaks in Irradiated Cervical Cancer Cell Lines. Cancer Res. *64*, 7144–7149.

Baria, K., Warren, C., Roberts, S.A., West, C.M., and Scott, D. (2001). Chromosomal radiosensitivity as a marker of predisposition to common cancers? Br. J. Cancer *84*, 892–896.
Barnett, G.C., West, C.M.L., Dunning, A.M., Elliott, R.M., Coles, C.E., Pharoah, P.D.P., and Burnet, N.G. (2009). Normal tissue reactions to radiotherapy. Nat. Rev. Cancer *9*, 134–142.

Behjati, S., Gundem, G., Wedge, D.C., Roberts, N.D., Tarpey, P.S., Cooke, S.L., Van Loo, P., Alexandrov, L.B., Ramakrishna, M., Davies, H., et al. (2016). Mutational signatures of ionizing radiation in second malignancies. Nat. Commun. *7*.

Bentzen, S.M. (2006). Preventing or reducing late side effects of radiation therapy: radiobiology meets molecular pathology. Nat. Rev. Cancer *6*, 702–713.

Berardinelli, F., Antoccia, A., Buonsante, R., Gerardi, S., Cherubini, R., Nadal, V.D., Tanzarella, C., and Sgura, A. (2013). The role of telomere length modulation in delayed chromosome instability induced by ionizing radiation in human primary fibroblasts. Environ. Mol. Mutagen. *54*, 172–179.

Bergstra, J., and Bengio, Y. (2012). Random search for hyper-parameter optimization. J. Mach. Learn. Res. *13*, 281–305.

Bodalal, Z., Trebeschi, S., Nguyen-Kim, T.D.L., Schats, W., and Beets-Tan, R. (2019).

Radiogenomics: bridging imaging and genomics. Abdom. Radiol. *44*, 1960–1984.

Broer, L., Codd, V., Nyholt, D.R., Deelen, J., Mangino, M., Willemsen, G., Albrecht, E., Amin, N., Beekman, M., de Geus, E.J.C., et al. (2013). Meta-analysis of telomere length in 19 713 subjects reveals high heritability, stronger maternal inheritance and a paternal age effect. Eur. J. Hum. Genet. *21*, 1163–1168.

Calado, R.T., Cooper, J.N., Padilla-Nash, H.M., Sloand, E.M., Wu, C.O., Scheinberg, P., Ried, T., and Young, N.S. (2012). Short telomeres result in chromosomal instability in hematopoietic cells and precede malignant evolution in human aplastic anemia. Leukemia *26*, 700–707.

Carrano, A.V. (1973). Chromosome aberrations and radiation-induced cell death: II. Predicted and observed cell survival. Mutat. Res. Mol. Mech. Mutagen. *17*, 355–366.

Carver, J.R., Shapiro, C.L., Ng, A., Jacobs, L., Schwartz, C., Virgo, K.S., Hagerty, K.L., Somerfield, M.R., Vaughn, D.J., and ASCO Cancer Survivorship Expert Panel (2007). American Society of Clinical Oncology clinical evidence review on the ongoing care of adult cancer survivors: cardiac and pulmonary late effects. J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol. *25*, 3991–4008.

Chen, T., and Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. - KDD 16 785–794.

Cornforth, M.N., Anur, P., Wang, N., Robinson, E., Ray, F.A., Bedford, J.S., Loucas, B.D.,

Williams, E.S., Peto, M., Spellman, P., et al. (2018). Molecular Cytogenetics Guides Massively Parallel Sequencing of a Radiation-Induced Chromosome Translocation in Human Cells. Radiat. Res. *190*, 88–97.

De Vitis, M., Berardinelli, F., Coluzzi, E., Marinaccio, J., O'Sullivan, R.J., and Sgura, A. (2019). X-rays Activate Telomeric Homologous Recombination Mediated Repair in Primary Cells. Cells *8*.

Delgado, D.A., Zhang, C., Gleason, K., Demanelis, K., Chen, L.S., Gao, J., Roy, S., Shinkle, J., Sabarinathan, M., Argos, M., et al. (2019). The contribution of parent-to-offspring transmission of telomeres to the heritability of telomere length in humans. Hum. Genet. *138*, 49–60.

Dracham, C.B., Shankar, A., and Madan, R. (2018). Radiation induced secondary malignancies: a review article. Radiat. Oncol. J. *36*, 85–94.

Epel, E.S., Blackburn, E.H., Lin, J., Dhabhar, F.S., Adler, N.E., Morrow, J.D., and Cawthon, R.M. (2004). Accelerated telomere shortening in response to life stress. Proc. Natl. Acad. Sci. U. S. A. *101*, 17312–17315.

F, B., A, S., A, D.M., S, L., Ga, C., F, R., C, T., and A, A. (2014). Radiation-induced telomere length variations in normal and in Nijmegen Breakage Syndrome cells. Int. J. Radiat. Biol. *90*, 45–52.

Fan, J., Wang, J., Chen, Z., Hu, C., Zhang, Z., and Hu, W. (2019). Automatic treatment planning based on three-dimensional dose distribution predicted from deep learning technique. Med. Phys. *46*, 370–381.

Garrett-Bakelman, F.E., Darshi, M., Green, S.J., Gur, R.C., Lin, L., Macias, B.R., McKenna, M.J., Meydan, C., Mishra, T., Nasrini, J., et al. (2019). The NASA Twins Study: A multidimensional analysis of a year-long human spaceflight. Science *364*.

Green, D.E., and Rubin, C.T. (2014). Consequences of irradiation on bone and marrow phenotypes, and its relation to disruption of hematopoietic precursors. Bone *0*, 87–94.

Greene-Schloesser, D., and Robbins, M.E. (2012). Radiation-induced cognitive impairment-from bench to bedside. Neuro-Oncol. *14*, iv37–iv44.

Habash, M., Bohorquez, L.C., Kyriakou, E., Kron, T., Martin, O.A., and Blyth, B.J. (2017). Clinical and Functional Assays of Radiosensitivity and Radiation-Induced Second Cancer. Cancers *9*.

Honig, L.S., Kang, M.S., Cheng, R., Eckfeldt, J.H., Thyagarajan, B., Leiendecker-Foster, C., Province, M.A., Sanders, J.L., Perls, T., Christensen, K., et al. (2015). Heritability of telomere length in a study of long-lived families. Neurobiol. Aging *36*, 2785–2790.

Howe, B., Umrigar, A., and Tsien, F. (2014). Chromosome Preparation From Cultured Cells. J.

Vis. Exp. JoVE.

Huang, L., Snyder, A.R., and Morgan, W.F. (2003). Radiation-induced genomic instability and its implications for radiation carcinogenesis. Oncogene *22*, 5848–5854.

Huber, R., Braselmann, H., Geinitz, H., Jaehnert, I., Baumgartner, A., Thamm, R., Figel, M., Molls, M., and Zitzelsberger, H. (2011). Chromosomal radiosensitivity and acute radiation side effects after radiotherapy in tumour patients - a follow-up study. Radiat. Oncol. Lond. Engl. *6*, 32.

Kerns, S.L., Ostrer, H., Stock, R., Li, W., Moore, J., Pearlman, A., Campbell, C., Shao, Y., Stone, N., Kusnetz, L., et al. (2010). Genome-wide association study to identify single nucleotide polymorphisms (SNPs) associated with the development of erectile dysfunction in African-American men after radiotherapy for prostate cancer. Int. J. Radiat. Oncol. Biol. Phys. *78*, 1292–1300.

Kerns, S.L., Dorling, L., Fachal, L., Bentzen, S., Pharoah, P.D.P., Barnes, D.R., Gómez-Caamaño, A., Carballo, A.M., Dearnaley, D.P., Peleteiro, P., et al. (2016). Meta-analysis of Genome Wide Association Studies Identifies Genetic Markers of Late Toxicity Following Radiotherapy for Prostate Cancer. EBioMedicine *10*, 150–163.

de Lange, T. (2009). How Telomeres Solve the End-Protection Problem. Science *326*, 948.

Lee, S., Liang, X., Woods, M., Reiner, A.S., Concannon, P., Bernstein, L., Lynch, C.F., Boice, J.D., Deasy, J.O., Bernstein, J.L., et al. (2020). Machine learning on genome-wide association studies to predict the risk of radiation-associated contralateral breast cancer in the WECARE Study. PLOS ONE *15*, e0226157.

Maeda, T., Nakamura, K., Atsumi, K., Hirakawa, M., Ueda, Y., and Makino, N. (2013).

Radiation-associated changes in the length of telomeres in peripheral leukocytes from inpatients with cancer. Int. J. Radiat. Biol. *89*, 106–109.

Martínez Paula, and Blasco Maria A. (2018). Heart-Breaking Telomeres. Circ. Res. *123*, 787–802.

McNally, E.J., Luncsford, P.J., and Armanios, M. (2019). Long telomeres and cancer risk: the price of cellular immortality. J. Clin. Invest. *129*, 3474–3481.

Miri, M., Nazarzadeh, M., Alahabadi, A., Ehrampoush, M.H., Rad, A., Lotfi, M.H., Sheikhha, M.H., Sakhvidi, M.J.Z., Nawrot, T.S., and Dadvand, P. (2019). Air pollution and telomere length in adults: A systematic review and meta-analysis of observational studies. Environ. Pollut. *244*, 636–647.

Mosesso, P., and Cinelli, S. (2019). In Vitro Cytogenetic Assays: Chromosomal Aberrations and Micronucleus Tests. In Genotoxicity Assessment: Methods and Protocols, A. Dhawan, and M.

Bajpayee, eds. (New York, NY: Springer), pp. 79–104.

Moyzis, R.K., Buckingham, J.M., Cram, L.S., Dani, M., Deaven, L.L., Jones, M.D., Meyne, J., Ratliff, R.L., and Wu, J.R. (1988). A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. Proc. Natl. Acad. Sci. U. S. A. *85*, 6622–6626.

Paul, S., Barker, C.A., Turner, H.C., McLane, A., Wolden, S.L., and Amundson, S.A. (2011). Prediction of In Vivo Radiation Dose Status in Radiotherapy Patients using Ex Vivo and In Vivo Gene Expression Signatures. Radiat. Res. *175*, 257–265.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., et al. (2018). Scikit-learn: Machine Learning in Python. ArXiv12010490 Cs.

Pella, A., Cambria, R., Riboldi, M., Jereczek-Fossa, B.A., Fodor, C., Zerini, D., Torshabi, A.E., Cattani, F., Garibaldi, C., Pedroli, G., et al. (2011). Use of machine learning methods for prediction of acute toxicity in organs at risk following prostate radiotherapy. Med. Phys. *38*, 2859–2867.

Poon, S.S.S., and Lansdorp, P.M. (2001). Quantitative Fluorescence In Situ Hybridization (Q-FISH). Curr. Protoc. Cell Biol. *12*, 18.4.1-18.4.21.

Rajaraman, P., Hauptmann, M., Bouffler, S., and Wojcik, A. (2018). Human individual radiation sensitivity and prospects for prediction. Ann. ICRP *47*, 126–141.

Ray, F.A., Zimmerman, E., Robinson, B., Cornforth, M.N., Bedford, J.S., Goodwin, E.H., and Bailey, S.M. (2013). Directional genomic hybridization for chromosomal inversion discovery and detection. Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol. *21*, 165–174.

Redon, C.E., Dickey, J.S., Bonner, W.M., and Sedelnikova, O.A. (2009). γ-H2AX as a biomarker of DNA damage induced by ionizing radiation in human peripheral blood lymphocytes and artificial skin. Adv. Space Res. Off. J. Comm. Space Res. COSPAR *43*, 1171–1178.

Schmitz, A., Bayer, J., Dechamps, N., Goldin, L., and Thomas, G. (2007). Heritability of susceptibility to ionizing radiation-induced apoptosis of human lymphocyte subpopulations. Int. J. Radiat. Oncol. Biol. Phys. *68*, 1169–1177.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of Image Analysis. Nat. Methods *9*, 671–675.

Seabold, S., and Perktold, J. (2010). Statsmodels: Econometric and Statistical Modeling with Python. (Austin, Texas), pp. 92–96.

Sgura, A., Antoccia, A., Berardinelli, F., Cherubini, R., Gerardi, S., Zilio, C., and Tanzarella, C.

(2006). Telomere length in mammalian cells exposed to low- and high-LET radiations. Radiat. Prot. Dosimetry *122*, 176–179.

Stone, M. (1974). Cross-Validatory Choice and Assessment of Statistical Predictions. J. R. Stat. Soc. Ser. B Methodol. *36*, 111–133.

Stout, G.J., and Blasco, M.A. (2013). Telomere length and telomerase activity impact the UV sensitivity syndrome xeroderma pigmentosum C. Cancer Res. *73*, 1844–1854.

Suit, H., Goldberg, S., Niemierko, A., Ancukiewicz, M., Hall, E., Goitein, M., Wong, W., and Paganetti, H. (2007). Secondary carcinogenesis in patients treated with radiation: a review of data on radiation-induced cancers in human, non-human primate, canine and rodent subjects. Radiat. Res. *167*, 12–42.

Taninaga, J., Nishiyama, Y., Fujibayashi, K., Gunji, T., Sasabe, N., Iijima, K., and Naito, T. (2019). Prediction of future gastric cancer risk using a machine learning algorithm and comprehensive medical check-up data: A case-control study. Sci. Rep. *9*, 1–9.

Tsoutsou, P.G., and Koukourakis, M.I. (2006). Radiation pneumonitis and fibrosis: Mechanisms underlying its pathogenesis and implications for future research. Int. J. Radiat. Oncol. *66*, 1281–1293.

Vidaček, N.Š., Nanić, L., Ravlić, S., Sopta, M., Gerić, M., Gajski, G., Garaj-Vrhovac, V., and Rubelj, I. (2017). Telomeres, Nutrition, and Longevity: Can We Really Navigate Our Aging? J. Gerontol. Ser. A.

Wang, J., Yang, P., Zhao, Y., Elhalawani, H., Liu, R., Zhu, H., Mohamed, A.S., Fuller, C.D., and Zhu, H. (2019). A Predictive model of radiation-related fibrosis based on radiomic features of Magnetic Resonance Imaging. Int. J. Radiat. Oncol. • Biol. • Phys. *105*, E599.

Ward, J.F. (1988). DNA Damage Produced by Ionizing Radiation in Mammalian Cells: Identities, Mechanisms of Formation, and Reparability. In Progress in Nucleic Acid Research and Molecular Biology, W.E. Cohn, and K. Moldave, eds. (Academic Press), pp. 95–125.

Weng, Q., Du, J., Yu, F., Huang, T., Chen, M., Lv, H., Ma, H., Hu, Z., Jin, G., Hu, Y., et al. (2016). The known genetic loci for telomere length may be involved in the modification of telomeres length after birth. Sci. Rep. *6*, 1–7.

Williams, E.S., and Bailey, S.M. (2009). Chromosome orientation fluorescence in situ hybridization (CO-FISH). Cold Spring Harb. Protoc. *2009*, pdb.prot5269.

Willis, N.A., Rass, E., and Scully, R. (2015). Deciphering the Code of the Cancer Genome: Mechanisms of Chromosome Rearrangement. Trends Cancer *1*, 217–230.

Wong, H.-P., and Slijepcevic, P. (2004). Telomere length measurement in mouse chromosomes by a modified Q-FISH method. Cytogenet. Genome Res. *105*, 464–470.

Yaprak, G., Gemici, C., Temizkan, S., Ozdemir, S., Dogan, B.C., and Seseogullari, O.O. (2018). Osteoporosis development and vertebral fractures after abdominal irradiation in patients with gastric cancer. BMC Cancer *18*, 972.

Young, A., Berry, R., Holloway, A.F., Blackburn, N.B., Dickinson, J.L., Skala, M., Phillips, J.L., and Brettingham-Moore, K.H. (2014). RNA-seq profiling of a radiation resistant and radiation sensitive prostate cancer cell line highlights opposing regulation of DNA repair and targets for radiosensitization. BMC Cancer *14*.

Yu, D., Liu, Z., Su, C., Han, Y., Duan, X., Zhang, R., Liu, X., Yang, Y., and Xu, S. (2020). Copy number variation in plasma as a tool for lung cancer prediction using Extreme Gradient Boosting (XGBoost) classifier. Thorac. Cancer *11*, 95–102.

Yusuf, S.W., Venkatesulu, B.P., Mahadevan, L.S., and Krishnan, S. (2017). Radiation-Induced Cardiovascular Disease: A Clinical Perspective. Front. Cardiovasc. Med. *4*.

von Zglinicki, T. (2000). Role of oxidative stress in telomere length regulation and replicative senescence. Ann. N. Y. Acad. Sci. *908*, 99–110.

CHAPTER 5: DISCUSSION and FUTURE DIRECTIONS

Overview

Telomeres erode with each cellular division and thus with aging. If telomere maintenance mechanisms are not present the erosion will continue until critically short lengths are reached at which point cells enter senescence or undergo apoptosis (Aubert and Lansdorp, 2008). As senescent cells accumulate, tissue function declines and manifests as aging and aging-related diseases, such as cancer and cardiovascular disease; individuals with longer telomeres are predisposed to the former while those with shorter telomeres the latter (Aubert and Lansdorp, 2008). Telomere length is thus both a biomarker of healthy aging and also a causal agent in disease progression (Aubert and Lansdorp, 2008). Throughout aging and the gradual attrition of telomere length, genetic factors, lifestyle choices, and environmental exposures can positively or negatively modulate telomere length, thus modifying risks for disease (Aubert and Lansdorp, 2008; Shammas, 2011; Shay, 2016; Xu et al., 2017; Zhan and Hägg, 2019; Zhan et al., 2017).

Telomere length thus can be utilized as a highly informative biomarker, which integrates an individual's health during and after specific exposures, providing material information for predicting future disease risk (Mirjolet et al., 2015; Protsenko et al., 2020; Vodenkova et al., 2020; Zhan and Hägg, 2019). CAs, like telomeres, are another highly informative biomarker, specifically for developing cancer after IR exposure (e.g. IR therapies, space radiation) (Baskar et al., 2012, 2014; Vodenkova et al., 2020). CAs are involved in virtually all cancers and are a standard measure of IR exposure. In the studies reported here, we used telomere length and CAs to interrogate the impacts of spaceflight on human health for astronauts aboard the ISS. Specifically, we aimed to elucidate mechanisms underpinning previous indications of space-

flight specific telomere elongation. We follow up these studies by attempting to develop novel machine learning frameworks for predicting telomeric and chromosomal responses to IR. We succeed in creating a novel framework for predicting an individual's telomeric responses to IR, which leverages individual telomere length measurements. The studies reported here clarify the short- and long-term impacts of space radiation in the context of telomere length and CAs, and provide a novel framework for predicting telomeric responses to IR using machine learning.

Discussion

Human telomere maintenance in extreme environments - ALT

In the recent NASA Twins study, dramatic telomere elongation was observed for the space-twin during a one-year mission aboard the ISS, which decreased rapidly upon return to below pre-flight levels. Elevated CAs (specifically inversions) were also observed during flight, concurrent with telomere length, and persisted post-flight (Garrett-Bakelman et al., 2019). In the first NASA Twins companion paper reported here, we aimed to explore mechanisms underlying spaceflight-specific telomere elongation. Here, we analyzed telomere length and CAs in two additional, unrelated astronauts on 6-month missions aboard the ISS. For the two additional astronauts, dramatic spaceflight-specific telomere elongation was again observed, and rapidly returned to or slightly below baseline levels at post-flight. Concurrent with spaceflight-specific elongation, we observed increased rates of heterogeneous sister telomere foci, and SCEs involving the telomere. Observations of elongated telomeres of heterogeneous length, paired with increased rates of SCEs, especially involving telomeric regions, are indicators of bona fide ALT.

We propose that long-duration spaceflight, specifically chronic oxidative stress from space radiation, induces transient activation of ALT to maintain telomere length, and that upon return from spaceflight the ALT activity is silenced. Spaceflight, namely space radiation, poses unique challenges for cellular biology (Chancellor et al., 2014, 2018; Cucinotta, 2015). The space radiation environment provides chronic, low doses of IR to astronauts aboard the ISS, punctuated by intermittent moments of higher dose (GCRs, SPEs) (Chancellor et al., 2014, 2018; Cucinotta, 2015). Chronic exposure of IR, even of low dose, is anticipated to induce chronic oxidative stress within cells (Azzam et al., 2012). With regards to telomere maintenance during chronic oxidative stress conditions, telomerase is known to be exported from the nucleus and localize to the mitochondria until the oxidative stress subsides (Haendeler Judith et al., 2009). Regardless of its mitochondrial activities, export from the nucleus would necessarily prevent telomerase from maintaining or elongating telomere length during chronic oxidative stress conditions, necessitating alternative means to maintain telomere length (Ahmed et al., 2008; Green et al., 2019). Knowing that telomerase activity is downregulated and or relegated to the mitochondria during periods of chronic oxidative stress provides important contextual evidence for all our telomerase activity measurements during or immediately after spaceflight. We previously proposed that heat inactivation from transport on the Soyuz capsule ablated telomerase activity; it could well be that those very low measurements of telomerase were actually “true”. Indeed, we have observed very low telomerase activity at the immediate post-flight timepoint (+7 days after) for all 3 astronauts in this study.

Notably, previous *in vitro* studies have reported that chronic oxidative stress, either induced chemically via low doses of hydrogen peroxide or via IR (X-rays), causes transient activation of ALT in normal human primary fibroblasts (Coluzzi et al., 2017; De Vitis et al.,

2019). Both of these *in vitro* studies observed elongated telomeres of heterogenous length paired with increased rates of T-SCE (Coluzzi et al., 2017; De Vitis et al., 2019). Both reports also indicated a lag between the initial induction of oxidative stress and activation of ALT, suggesting that the chronic nature of the oxidative stress is important for ALT activation. Importantly, these *in vitro* studies provide precedent for our observations of ALT activation during spaceflight.

In the second NASA Twins companion paper reported here, we aimed to further explore and characterize potential mechanisms contributing to ALT activation during spaceflight through analysis of biochemical indicators of oxidative stress and cell proliferation for 11 unrelated astronauts. As well, we sought data on human exposure to other extreme environments – namely Mt. Everest – to explore if transient ALT activation during chronic oxidative stress, resulting from the extreme environment, is part of a general adaptive response. First, we observed significant variation in the long-term effects of spaceflight on telomere length; some astronauts experienced significant shortening, others lengthening, and a few astronauts did not exhibit any changes. The variables accounting for the differences in these observations are not clear. However, these observations highlight the clear and important need for personalized monitoring in the context of spaceflight, given that those astronauts who experienced shortened telomeres have been predisposed to increased risks for CVD, while those with lengthened telomeres are at increased risks for cancer (Protsenko et al., 2020). Again, these observations highlight the importance of personalized monitoring of health outcomes in the context of spaceflight.

In terms of changes in biochemical analytes in relation to telomere length during spaceflight, we observed significant upregulation of many factors strongly associated with oxidative stress and inflammatory responses. Such factors included interleukins (e.g IL-4, IL-10, IL-5), TNF- α, and VEGF-1, and all were positively correlated with telomere length throughout

spaceflight, suggesting that with increased responses to oxidative stress (activation of necessary compensatory mechanisms) comes increased telomere length (Sallam and Laher, 2016). High upregulation of ferritin with a strong correlation to telomere length was observed, as would be expected for individuals with chronic oxidative stress (Sallam and Laher, 2016).

Acquiring human samples from spaceflight is inherently difficult, hence we turned to other extreme environments on Earth to ascertain if transient activation of ALT is a general adaptive response to chronic oxidative stress. We received many blood samples from human climbers during a long (30+ days) expedition on Mt. Everest. With these samples we again observed elongated telomeres during, but not after, the expedition on Mt. Everest; notably, the elongation was not immediate, suggesting a lag between exposure to the extreme environment (induction of oxidative stress) and activation of telomere elongation.

To probe mechanisms responsible for telomere elongation we turned to RNA sequencing experiments. We first observed that the expression of telomerase (TERT) was flat or completely ablated during but not before the Mt. Everest expedition. This result indicates that telomerase could likely not account for the elongation observed in the climbers. This result is reminiscent of telomerase activity measurements at the immediate post-flight timepoint (+7 days) in the second NASA companion paper, where telomerase activity exhibited a remarkable flattening in variation and activity levels for all astronauts. After observing the telomerase expression measurements in the Mt. Everest climbers we noted that TERC (telomerase RNA) and RAD50 were highly upregulated during the Mt. Everest climb. TERC has been evidenced to serve as a retrograde signal (to the nucleus) for mitochondrial function and is anticipated to have elevated expression during chronic oxidative stress (Zheng et al., 2019). RAD50 is the classic protein indicating DNA damage repair(Kinoshita et al., 2009). Pathway analyses revealed that, in line with gene

expression data, telomerase expression was non-responsive, not upregulated, and in-fact ablated during the Mt. Everest climb. In contrast, from the pathway analysis we observed significant upregulation of pathways involving telomere maintenance by repair as well as ALT during the climb. Altogether, we observed evidence for telomere elongation facilitated by transient ALT activation during exposure to the extreme environment Mt. Everest. These results suggest that transient activation of ALT may be a general adaptive response to chronic oxidative stress for human cells. Again, during periods of chronic oxidative stress telomerase is exported from the nucleus and relegated to the mitochondria; transient activation of ALT during long periods of chronic oxidative stress would enable cells to continue maintenance of telomeres while telomerase is absent from the nucleus (Green et al., 2019; Haendeler Judith et al., 2009).

Chromosomal DNA damage responses to spaceflight

With regards to spaceflight, we observed transient ALT activation, as indicated by elongated telomeres of heterogenous lengths paired with increased rates SCEs involving the telomere (Blagoev et al., 2010; Londoño-Vallejo et al., 2004; Rudd et al., 2007). Concurrent with telomere elongation during spaceflight, and also persisting afterwards, was significantly increased incidences of CAs. First, elongated telomeres will naturally enable cells greater proliferative capacity and thus increased numbers of divisions (Aubert and Lansdorp, 2008). If cells are granted increased proliferative capacity while also being subject to increased rates of CAs, as in the context of the space radiation environment, it is wholly likely if not certain that this combination poses increased risks for cancer (Aubert and Lansdorp, 2008; Chancellor et al., 2014, 2018; Cucinotta, 2015). Greater proliferative capacity enables more time to acquire

oncogenic mutations; the significant increases of CAs during spaceflight could indeed drive mutations tumorigenesis (Little, 2000).

With regards to the specific types of CAs increased via spaceflight, we observed increased rates of inversions during spaceflight which persisted at post-flight. Inversions, especially small inversions, have historically been difficult to observe (Weckselblatt and Rudd, 2015). Array-based techniques don't capture inversions as they don't affect copy number; short-read sequencing fails to capture inversions (unlikely for sufficient reads to extend over the breakpoints); and inversions are simply difficult to observe via traditional cytogenetic techniques (i.e. banding) (Weckselblatt and Rudd, 2015). The recently developed cytogenetic technique, directional Genomic Hybridization (dGH), efficiently captures inversions thanks to the development of sequence and strand specific fluorescent probes (Ray et al., 2013, 2014). We utilized dGH to score CAs in astronauts during spaceflight, enabling an improved assessment of CAs resulting from the space radiation environment owing to improved capture of inversions.

Although we did observe significant increases in inversions during (and persisting after) spaceflight, we notably did not observe markedly increased rates of other types of CAs. Dicentrics were indeed observed at higher rates during spaceflight; we propose however that these increases were not of material significance, as the observed dicentrics did not persist after spaceflight. Regarding translocations, we did not observe significant elevations during or after spaceflight. Altogether, it is apparently the case that inversions occur at higher background frequencies than both dicentrics and translocations; however, inversions did significantly elevate, nearly doubling during and after spaceflight, indicating material concerns for genomic instability during spaceflight as well as increased risks for cancer (Chancellor et al., 2014, 2018; Cucinotta, 2015; Little, 2000).

Though SCEs are not strictly induced by radiation they do result as secondary effects markers of radiation exposure and replication stress, i.e. SCE events can be triggered via DNA damage leading to repair events and thus provide an indication of genomic instability (Conrad et al., 2011; Nagasawa and Little, 1992). Along with inversions, we generally observed elevated or significantly elevated levels of SCEs during spaceflight, with trends back to baseline at post-flight. We notably observed elevated rates of SCEs involving the sub-telomere and indications of T-SCE (heterogenous sister telomeres); both SCEs importantly involve the telomere and are well regarded as markers for ALT (Blagoev et al., 2010; Londoño-Vallejo et al., 2004; Rudd et al., 2007). While SCEs cannot directly contribute to mutagenesis like inversions or CAs in general, they do provide important markers of replication stress and as such, indicated high levels of genomic instability during spaceflight, which generally returned to baseline levels at post-flight.

Predicting telomeric and chromosomal responses to IR

The highly informative nature of telomere length with respect to cardiovascular disease and cancer underscored the need for developing a framework to predict how an individual's telomere length will change to a given exposure and thus modify risks for disease (Mirjolet et al., 2015; Protsenko et al., 2020; Vodenkova et al., 2020; Zhan and Hägg, 2019). Following our observations of telomeric responses to spaceflight and knowing the importance of telomere length in modifying disease risk, we explored the possibility of developing a machine learning framework that leverages individual telomere length measurements to predicting changes in telomere length following IR exposure. Given that telomere length is modulated by genetic, lifestyle, and environmental factors, we hypothesized that how an individual's telomere length responses to a given exposure would be essentially determined by these factors and the type of

the exposure. If an individual's telomeric responses to an exposure, namely IR, was the result of that individual's predispositions, then it should follow that telomere length responds to IR in a manner that can be learned, and thus predicted, by a machine learning framework. In the context of our study with prostate cancer patients undergoing Intensity Modulated Radiation Therapy (IMRT), we specifically aimed to predict changes in telomere length to infer modified risk for cancer and cardiovascular disease, respectively.

Machine learning models by nature require large amounts of data for training before they can "learn" patterns and successfully make predictions using new data (Li et al., 2018). We found success in leveraging the hundreds of thousands of individual telomere length measurements generated by Telomere Fluorescence *in situ* Hybridization (Telo-FISH) for training our machine learning models to predict mean telomere length after IMRT in prostate cancer patients. We specifically utilized the XGBoost machine learning model, which is essentially a large ensemble of many smaller models, for our framework (Chen and Guestrin, 2016).

The intuition behind our approach is as follows: we train the XGBoost model to look at thousands of individual telomeres (Telo-FISH) from a patient and make "guesses", based on each individual telomere, for the patient's mean telomere length post-IMRT. We then average the thousands of individual predictions for the patient's post-IMRT mean telomere length, which yields the model's prediction for the patient's mean telomere length at post-IMRT. In our approach any given prediction based on an individual telomere is extremely poor, but by averaging all predictions for a given patient, the performance of the prediction becomes extremely high and accurate. Indeed, we demonstrated that our XGBoost models were able to predict post-IMRT mean telomere length, numbers of short telomeres, and numbers of long

telomeres when presented with new individual telomere data (test data) as well as new patients (round-robin training). Altogether, we provided the first framework for predicting an individual's telomeric response to a given exposure. Importantly, our machine learning framework was extremely performant and able to generalize to "new" patients, ascertained via the "leave one/two/three out" approach (Beirami et al., 2017), suggesting that our approach could be performant in a clinical setting where predicting telomere length post-IMRT (or post IR) is of interest.

In contrast to our work with telomeres, our experiences with creating machine learning models that predict an individual's chromosomal aberrations after IMRT did not bear success. We specifically aimed to create models that could evaluate the incidence of CAs in a patient at baseline, and from that baseline predict the incidence of CAs after IR therapy. There are numerous reasons that could account for our inability to effectively predict incidence of CAs post-IMRT from pre-IMRT counts of patient CAs. We believe the core reasons for the non-performant machine learning models from CA data came down to the amount and type of data.

With dGH we scored CAs in metaphase spreads from patients undergoing IMRT. Scoring CAs is unfortunately both a labor and time intensive process. Between the four timepoints involved this study, we scored 30 cells per timepoint for each of the 15 patients, yielding 450 cells per timepoint. While this number is more than sufficient for statistical analyses, it unfortunately may not have been facilitative for training XGBoost models. Indeed, the anticipated difference in performance between a machine learning model trained on 450 and 300,000 cells is stark. Another potential issue with using scored counts of CAs is that inherently, many values in your dataset will be zero; a dataset inflated with zeros presents unique challenges for effective training. We attempted various and numerous data transformation approaches to

circumvent the inflated-zero issue which were ultimately unsuccessful. Altogether, we do believe that predicting the incidence of CAs following IR exposure is possible, but likely we did not have sufficient data to train performant models.

Future directions

ALT activation during exposure to extreme environments

Altogether, from our studies with NASA astronauts aboard the ISS we have provided telomeric and cytogenetic evidence for transient activation of ALT during spaceflight; we propose this occurs due to chronic oxidative stress resulting specifically from chronic low dose exposure to space radiation but also from the general stressors of spaceflight (Azzam et al., 2012; Coluzzi et al., 2017; De Vitis et al., 2019). We have also provided telomeric and RNA sequencing evidence for transient ALT activation in the extreme environment of Mt. Everest. In combination, we propose that in humans ALT is transiently activated during chronic exposure to extreme environments, such as spaceflight and Mt. Everest, due to the chronic oxidative stress resulting from exposure to the extreme environment. The specific genes, timing, and biological implications of transient ALT activation, especially in conjunction with increased rates of CAs (inversions) remain unclear.

Future directions should interrogate the specific timing and genes involved in the transient activation of ALT. These pursuits should include cell proliferation studies to ascertain the increase in proliferative capacity in addition to mutagenesis studies examining increased risks for tumorigenesis from the elevated levels of inversions observed during spaceflight. Answering when, how, and to what extent the transient activation of ALT in these cells contributes to tumorigenesis should be thoroughly investigated. These investigations should also include

considerations for the differences in space radiation dose that would be anticipated for astronauts on the ISS and for those pursuing deep space travel. Indeed, as we move away from the Earth and into deep space, the dosages of space radiation to which astronauts are exposed will increase; how these increased doses will interact with telomeres, CAs, tumorigenesis, and astronaut health remains to be answered (Chancellor et al., 2014, 2018; Cucinotta, 2015).

Machine learning to predict telomeric exposures

We have provided a viable framework for utilizing individual telomere length measurements to train XGBoost machine learning algorithms to predict an individual's mean telomere length at post-IMRT. The successful demonstration of predicting an individual's telomeric response to radiotherapy naturally begs the question of "what else?". There are numerous contexts in which predicting an individual's telomeric response to a specific exposure or event would be of significant interest (Choi et al., 2014; Mirjolet et al., 2015; Protsenko et al., 2020; Vodenkova et al., 2020). As well, there are many functional and theoretical improvements that could be made to our machine learning approach.

Future experiments should consider not only predicting how an exposure will modify an individual's telomere length at a specific point in time, but also how the exposure will modify their telomere length over time. Longitudinal studies with an intermediary exposure (e.g. physical activity, IR therapy, spaceflight missions) and multiple subsequent timepoints, combined with training individual models for each timepoint, would achieve this goal. Future experiments should explore interventions anticipated to induce positive (diet, physical activity) and other negative changes (IR, depression) in telomere length. Machine learning models alternative to XGBoost should be considered; while XGBoost was state of the art at the time of

our manuscript's writing, numerous advancements and improved models are released seemingly yearly and fast become widely available. Such models and approaches to explore would include LightGM and CatBoost (Ke et al., 2017; Prokhorenkova et al., 2019); deep learning/neural networks and ensembling of different models (Boldrini et al., 2019; Li et al., 2018). As of this writing, we specifically recommend exploration of a deep learning approach enhanced with the Adam or AdaBelief optimizers and also sparse evolutionary networks (Kingma and Ba, 2017; Mocanu et al., 2018; Zhuang et al., 2020).

REFERENCES

Ahmed, S., Passos, J.F., Birket, M.J., Beckmann, T., Brings, S., Peters, H., Birch-Machin, M.A., von Zglinicki, T., and Saretzki, G. (2008). Telomerase does not counteract telomere shortening but protects mitochondrial function under oxidative stress. J. Cell Sci. *121*, 1046–1053.

Aubert, G., and Lansdorp, P.M. (2008). Telomeres and aging. Physiol. Rev. *88*, 557–579.

Azzam, E.I., Jay-Gerin, J.-P., and Pain, D. (2012). Ionizing radiation-induced metabolic oxidative stress and prolonged cell injury. Cancer Lett. *327*, 48–60.

Baskar, R., Lee, K.A., Yeo, R., and Yeoh, K.-W. (2012). Cancer and Radiation Therapy: Current Advances and Future Directions. Int. J. Med. Sci. *9*, 193–199.

Baskar, R., Dai, J., Wenlong, N., Yeo, R., and Yeoh, K.-W. (2014). Biological response of cancer cells to radiation treatment. Front. Mol. Biosci. *1*.

Beirami, A., Razaviyayn, M., Shahrampour, S., and Tarokh, V. (2017). On Optimal Generalizability in Parametric Learning. ArXiv171105323 Cs Stat.

Blagoev, K.B., Goodwin, E.H., and Bailey, S.M. (2010). Telomere sister chromatid exchange and the process of aging. Aging *2*, 727–730.

Boldrini, L., Bibault, J.-E., Masciocchi, C., Shen, Y., and Bittner, M.-I. (2019). Deep Learning: A Review for the Radiation Oncologist. Front. Oncol. *9*.

Chancellor, J.C., Scott, G.B.I., and Sutton, J.P. (2014). Space Radiation: The Number One Risk to Astronaut Health beyond Low Earth Orbit. Life Open Access J. *4*, 491–510.

Chancellor, J.C., Blue, R.S., Cengel, K.A., Auñón-Chancellor, S.M., Rubins, K.H., Katzgraber, H.G., and Kennedy, A.R. (2018). Limitations in predicting the space radiation health risk for exploration astronauts. Npj Microgravity *4*, 1–11.

Chen, T., and Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. 785–794.

Choi, D.K., Helenowski, I., and Hijiya, N. (2014). Secondary malignancies in pediatric cancer survivors: Perspectives and review of the literature. Int. J. Cancer *135*, 1764–1773.

Coluzzi, E., Buonsante, R., Leone, S., Asmar, A.J., Miller, K.L., Cimini, D., and Sgura, A. (2017). Transient ALT activation protects human primary cells from chromosome instability induced by low chronic oxidative stress. Sci. Rep. *7*, 43309.

Conrad, S., Künzel, J., and Löbrich, M. (2011). Sister chromatid exchanges occur in G2-irradiated cells. Cell Cycle *10*, 222–228.

Cucinotta, F.A. (2015). A New Approach to Reduce Uncertainties in Space Radiation Cancer Risk Predictions. PLoS ONE *10*.

De Vitis, M., Berardinelli, F., Coluzzi, E., Marinaccio, J., O'Sullivan, R.J., and Sgura, A. (2019). X-rays Activate Telomeric Homologous Recombination Mediated Repair in Primary Cells. Cells *8*.

Garrett-Bakelman, F.E., Darshi, M., Green, S.J., Gur, R.C., Lin, L., Macias, B.R., McKenna, M.J., Meydan, C., Mishra, T., Nasrini, J., et al. (2019). The NASA Twins Study: A multidimensional analysis of a year-long human spaceflight. Science *364*.

Green, P.D., Sharma, N.K., and Santos, J.H. (2019). Telomerase Impinges on the Cellular Response to Oxidative Stress Through Mitochondrial ROS-Mediated Regulation of Autophagy. Int. J. Mol. Sci. *20*.

Haendeler Judith, Dröse Stefan, Büchner Nicole, Jakob Sascha, Altschmied Joachim, Goy Christine, Spyridopoulos Ioakim, Zeiher Andreas M., Brandt Ulrich, and Dimmeler Stefanie (2009). Mitochondrial Telomerase Reverse Transcriptase Binds to and Protects Mitochondrial DNA and Function From Damage. Arterioscler. Thromb. Vasc. Biol. *29*, 929–935.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In Advances in Neural Information Processing Systems 30, I. Guyon, U.V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, eds. (Curran Associates, Inc.), pp. 3146–3154.

Kingma, D.P., and Ba, J. (2017). Adam: A Method for Stochastic Optimization. ArXiv14126980 Cs.

Kinoshita, E., van der Linden, E., Sanchez, H., and Wyman, C. (2009). RAD50, an SMC family member with multiple roles in DNA break repair: how does ATP affect function? Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol. *17*, 277–288.

Li, Y., Wu, F.-X., and Ngom, A. (2018). A review on machine learning principles for multi-view biological data integration. Brief. Bioinform. *19*, 325–340.

Little, J.B. (2000). Radiation carcinogenesis. Carcinogenesis *21*, 397–404.

Londoño-Vallejo, J.A., Der-Sarkissian, H., Cazes, L., Bacchetti, S., and Reddel, R.R. (2004). Alternative Lengthening of Telomeres Is Characterized by High Rates of Telomeric Exchange. Cancer Res. *64*, 2324–2327.

Mirjolet, C., Boidot, R., Saliques, S., Ghiringhelli, F., Maingon, P., and Créhange, G. (2015). The role of telomeres in predicting individual radiosensitivity of patients with cancer in the era of personalized radiotherapy. Cancer Treat. Rev. *41*, 354–360.

Mocanu, D.C., Mocanu, E., Stone, P., Nguyen, P.H., Gibescu, M., and Liotta, A. (2018). Scalable training of artificial neural networks with adaptive sparse connectivity inspired by network science. Nat. Commun. *9*, 2383.

Nagasawa, H., and Little, J.B. (1992). Induction of sister chromatid exchanges by extremely low doses of alpha-particles. Cancer Res. *52*, 6394–6396.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V., and Gulin, A. (2019). CatBoost: unbiased boosting with categorical features. ArXiv170609516 Cs.

Protsenko, E., Rehkopf, D., Prather, A.A., Epel, E., and Lin, J. (2020). Are long telomeres better than short? Relative contributions of genetically predicted telomere length to neoplastic and non-neoplastic disease risk and population health burden. PLOS ONE *15*, e0240185.

Ray, F.A., Zimmerman, E., Robinson, B., Cornforth, M.N., Bedford, J.S., Goodwin, E.H., and Bailey, S.M. (2013). Directional genomic hybridization for chromosomal inversion discovery and detection. Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol. *21*, 165–174.

Ray, F.A., Robinson, E., McKenna, M., Hada, M., George, K., Cucinotta, F., Goodwin, E.H., Bedford, J.S., Bailey, S.M., and Cornforth, M.N. (2014). Directional genomic hybridization: inversions as a potential biodosimeter for retrospective radiation exposure. Radiat. Environ. Biophys. *53*, 255–263.

Rudd, M.K., Friedman, C., Parghi, S.S., Linardopoulou, E.V., Hsu, L., and Trask, B.J. (2007). Elevated Rates of Sister Chromatid Exchange at Chromosome Ends. PLoS Genet *3*, e32.

Sallam, N., and Laher, I. (2016). Exercise Modulates Oxidative Stress and Inflammation in Aging and Cardiovascular Diseases. Oxid. Med. Cell. Longev. *2016*.

Shammas, M.A. (2011). Telomeres, lifestyle, cancer, and aging: Curr. Opin. Clin. Nutr. Metab. Care *14*, 28–34.

Shay, J.W. (2016). Role of Telomeres and Telomerase in Aging and Cancer. Cancer Discov. *6*, 584–593.

Vodenkova, S., Kroupa, M., Polivkova, Z., Musak, L., Ambrus, M., Schneiderova, M., Kozevnikovova, R., Vodickova, L., Rachakonda, S., Hemminki, K., et al. (2020). Chromosomal damage and telomere length in peripheral blood lymphocytes of cancer patients. Oncol. Rep. *44*, 2219–2230.

Weckselblatt, B., and Rudd, M.K. (2015). Human Structural Variation: Mechanisms of Chromosome Rearrangements. Trends Genet. TIG *31*, 587–599.

Xu, J., Songyang, Z., Liu, D., and Kim, H. (2017). Analysis of Average Telomere Length in Human Telomeric Protein Knockout Cells Generated by CRISPR/Cas9. In Telomeres and Telomerase, Z. Songyang, ed. (New York, NY: Springer New York), pp. 15–28.

Zhan, Y., and Hägg, S. (2019). Telomere length and cardiovascular disease risk. Curr. Opin. Cardiol. *34*, 270–274.

Zhan, Y., Karlsson, I.K., Karlsson, R., Tillander, A., Reynolds, C.A., Pedersen, N.L., and Hägg, S. (2017). Exploring the Causal Pathway From Telomere Length to Coronary Heart Disease: A Network Mendelian Randomization Study. Circ. Res. *121*, 214–219.

Zheng, Q., Liu, P., Gao, G., Yuan, J., Wang, P., Huang, J., Xie, L., Lu, X., Di, F., Tong, T., et al. (2019). Mitochondrion-processed TERC regulates senescence without affecting telomerase activities. Protein Cell *10*, 631–648.

Zhuang, J., Tang, T., Tatikonda, S., Dvornek, N., Ding, Y., Papademetris, X., and Duncan, J.S. (2020). AdaBelief Optimizer: Adapting Stepsizes by the Belief in Observed Gradients. ArXiv201007468 Cs Stat.