

ARTICLE OPEN



A pan-sarcoma landscape of telomeric content shows that alterations in *RAD51B* and *GID4* are associated with higher telomeric content

Radwa Sharaf¹, Dexter X. Jin¹, John Grady^{2,3,4}, Christine Napier^{2,3,4}, Ericka Ebot¹, Garrett M. Frampton¹, Lee A. Albacker¹, David M. Thomas^{2,3,4} and Meagan Montesion¹✉

Tumor cells need to activate a telomere maintenance mechanism, enabling limitless replication. The bulk of evidence supports that sarcomas predominantly use alternative lengthening of telomeres (ALT) mechanism, commonly associated with alterations in *ATRX* and *DAXX*. In our dataset, only 12.3% of sarcomas harbored alterations in these genes. Thus, we checked for the presence of other genomic determinants of high telomeric content in sarcomas. Our dataset consisted of 13555 sarcoma samples, sequenced as a part of routine clinical care on the FoundationOne®Heme platform. We observed a median telomeric content of 622.3 telomeric reads per GC-matched million reads (TRPM) across all samples. In agreement with previous studies, telomeric content was significantly higher in *ATRX* altered and *POT1* altered sarcomas. We further observed that sarcomas with alterations in *RAD51B* or *GID4* were enriched in samples with high telomeric content, specifically within uterus leiomyosarcoma for *RAD51B* and soft tissue sarcoma (not otherwise specified, nos) for *GID4*. Furthermore, *RAD51B* and *POT1* alterations were mutually exclusive with *ATRX* and *DAXX* alterations, suggestive of functional redundancy. Our results propose a role played by *RAD51B* and *GID4* in telomere elongation in sarcomas and open research opportunities for agents aimed at targeting this critical pathway in tumorigenesis.

npj Genomic Medicine (2023)8:26; <https://doi.org/10.1038/s41525-023-00369-6>

INTRODUCTION

Sarcomas represent a rare and heterogeneous group of tumors with a mesenchymal origin, exhibiting different biological behavior and subsequent varying clinical manifestations. Due to their rarity, they have been relatively understudied. To allow for unchecked cellular proliferation, tumors must overcome the problem of telomere shortening¹. Telomeres are DNA-protein complexes that consist of 5–15 kb of repetitive hexamers forming protective caps at the ends of linear chromosomes and shorten by an average of 50–150 base pairs with every cell cycle^{2–5}. If telomeres reach a critical limit in length, cells go into replicative senescence and can no longer divide^{6,7}.

Currently, two mechanisms for telomere maintenance have been described. These include telomerase-dependent elongation, mediated by *TERT* activation^{8–11}, and alternative lengthening of telomeres (ALT), linked to the loss of *ATRX* or *DAXX*^{12–14}. Conflicting results have been published about telomerase activity in sarcomas, exemplified by the wide range of reported soft tissue sarcomas displaying telomerase activity, ranging from 7% up to 81%^{15–18}. Previous studies have reported a *TERT* promoter mutation prevalence rate of 6–11% across sarcomas^{19–21}. The bulk of evidence supports that 20–60% of sarcomas activate the ALT pathway for telomere elongation^{22–26}. However, in our dataset, only 12.3% of sarcomas harbored *ATRX/DAXX* alterations (11.4% for *ATRX* and 0.9% for *DAXX*), leaving the remaining cases unexplained.

Here we sought to investigate the genomic determinants of high telomeric content beyond the canonical *ATRX/DAXX* alterations within our sarcoma dataset. We characterized the telomeric landscape of 13555 sarcoma samples across 38 disease types sequenced on the FoundationOne®Heme platform. We also

validated our findings using samples sequenced on a different platform, FoundationOne®CDx, as well as samples from an independent academic cohort. To assess potential clinical relevance, we compared overall survival from date of metastatic diagnosis of sarcoma patients whose tumors harbored genomic alterations associated with high telomeric content with patients whose tumors lacked these alterations.

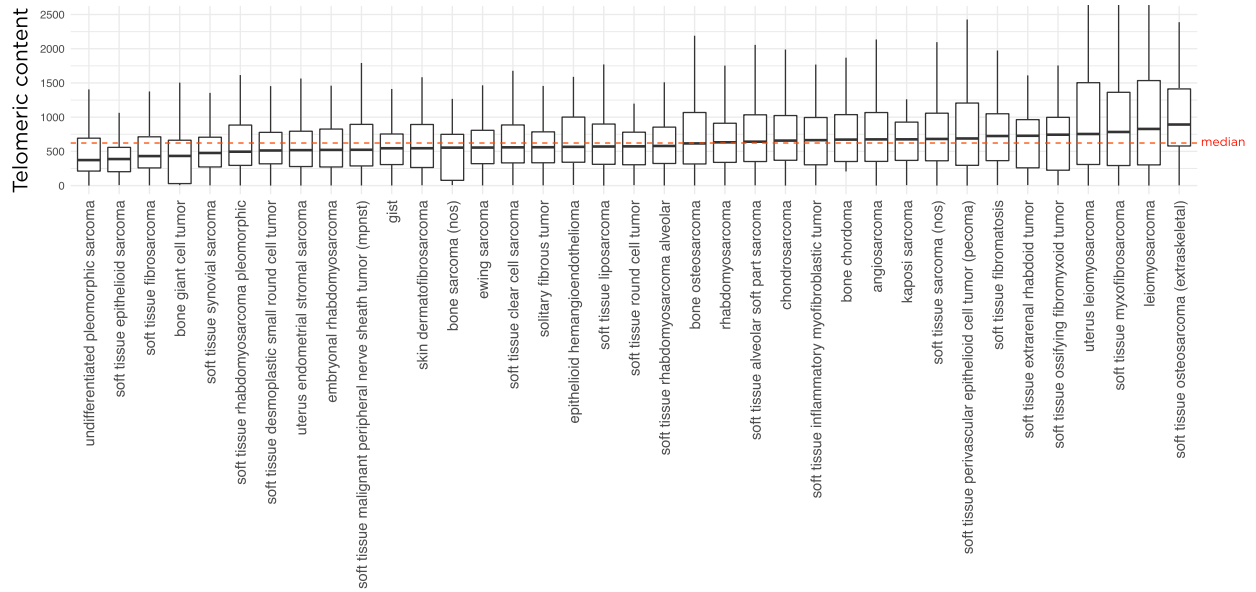
RESULTS

Telomeric content across FMI's sarcoma cohort

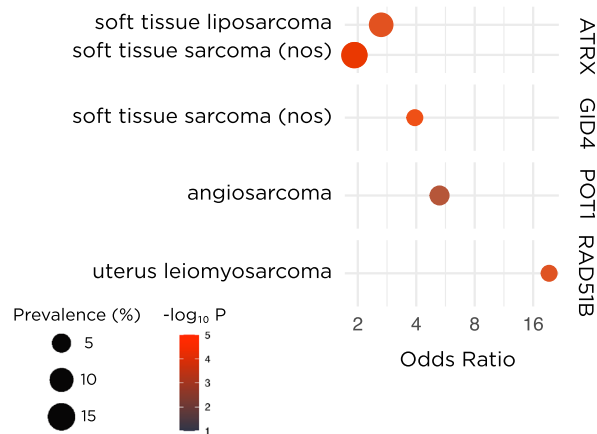
We measured the telomeric content for 13555 sarcoma samples, representing 38 unique sarcoma types, using TelomereHunter. Our results showed that telomeric content varied by sarcoma disease type, with a median telomeric content of 622.3 TRPM (telomeric reads per GC content-matched million reads) across all sarcoma types and a range of 519.4 (Fig. 1A and Table S1). The highest median telomeric content was observed in soft tissue osteosarcoma (extraskelatal) at 892.4 TRPM and the lowest in undifferentiated pleomorphic sarcoma at 373.0 TRPM. Based on previously-published criteria^{27,28}, we categorized sarcoma diseases broadly as either translocation-associated ($N = 2447$, 18.1%) or genomically complex/other which display complex karyotypes ($N = 11108$, 82.0%, Table S1). We hypothesized that the median telomeric content of the genomically complex sarcomas will be greater than the translocation-associated group, given the extra telomeres stemming from chromosome gain events. Our analysis showed that both the median telomeric content, as well as the median age, of the genomically complex/other group was significantly higher than the translocation-associated group (650 vs 542 TRPM, $p < 0.0001$ for telomeric content and 56 vs 36 years,

¹Foundation Medicine Inc., Cambridge, MA, USA. ²Omico Australian Genomic Cancer Medicine, Sydney, Australia. ³Garvan Institute of Medical Research, Sydney, Australia. ⁴St Vincent's Clinical School, University of New South Wales, Sydney, Australia. ✉email: mmontesion@foundationmedicine.com

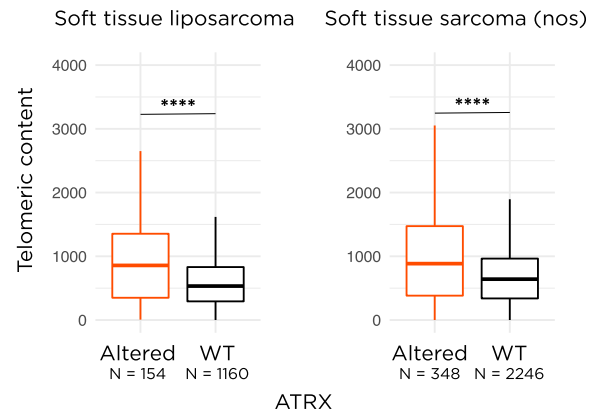
A.



B.



C.



D.

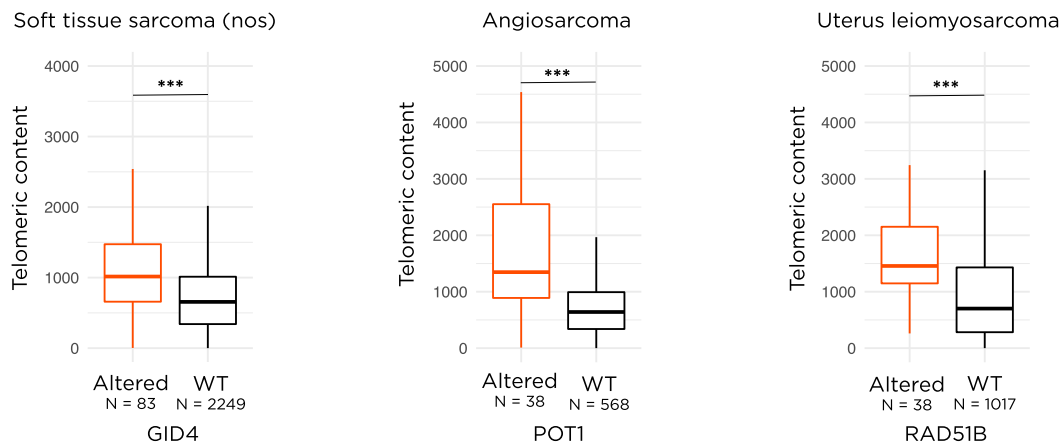


Fig. 1 Overview of telomeric content landscape in sarcomas and screen results. **A** A boxplot showing the landscape of telomeric content across sarcoma disease groups. The red dotted line depicts the median telomeric content across all samples. **B** Results from the screen displaying the 4 genetic associations identified within each disease ontology. The p -value is denoted by the color and alteration prevalence is denoted by the size of each circle. **C** Boxplots depicting the difference in telomeric content between *ATRX* altered and *ATRX* WT samples within soft tissue liposarcoma and soft tissue sarcoma nos. **D** Boxplots depicting the difference in telomeric content between *GID4* altered vs *GID4* WT samples within soft tissue sarcoma nos, *POT1* altered vs *POT1* WT within angiosarcoma, and *RAD51B* altered vs *RAD51B* WT within uterus leiomyosarcoma. The counts of samples within each group are indicated under the x-axis. **** $p < 0.0001$ and *** $p < 0.001$. nos, not otherwise specified. In all boxplots, the box extends from the first to the third quartile with a line in the middle that represents the median.

$p < 0.0001$ for age, Fig. S1A, B). This could explain our observation that sarcoma samples from older patients had a higher median telomeric content (Fig. S1C).

Impact of genetic alterations on telomeric content

To assess which genetic alterations were associated with high telomeric content, we compared the frequency of alterations observed in samples that fall in the top quartile of telomeric content to the frequency seen in the bottom quartile within each tumor type. In agreement with previous studies, telomeric content was significantly higher in *ATRX* altered samples. Specifically, alterations in *ATRX* were significantly enriched in the top quartile of soft tissue liposarcoma (Bonferroni-corrected $p < 0.01$, odds ratio (OR) = 2.7 [1.7–4.2]) and soft tissue sarcoma not otherwise specified (nos, Bonferroni-corrected $p < 0.01$, OR = 2.0 [1.4–2.6], Fig. 1B and Table S2). Median telomeric content of *ATRX* altered samples was significantly higher than WT samples (877 vs 534 TRPM for soft tissue liposarcoma and 889 vs 647 TRPM for soft tissue sarcoma nos, Fig. 1C). Intriguingly, *DAXX* did not show up in our screen, so we assessed the impact of its alterations on the telomeric content of three disease types with at least 20 *DAXX* altered samples. The median telomeric content was higher in two groups but did not reach statistical significance after Bonferroni's correction (1041 vs 730 TRPM for uterus leiomyosarcoma and 1132 vs 671 TRPM for soft tissue sarcoma nos, $p > 0.05$ for all comparisons, Fig. S2).

We further identified that alterations in 3 genes, *GID4*, *POT1*, and *RAD51B*, were enriched in sarcomas with high telomeric content, specifically within soft tissue sarcoma nos for *GID4* (Bonferroni-corrected $p < 0.01$, OR = 3.9 [2.0–7.7]), angiosarcoma samples for *POT1* (Bonferroni-corrected $p < 0.05$, OR = 5.3 [2.0–14.4]), and uterus leiomyosarcoma for *RAD51B* (Bonferroni-corrected $p < 0.01$, OR = 20.5 [2.6–146.0], Fig. 1B and Table S2). Interestingly, the odds ratio for the enrichment of alterations within these three newly-identified genes in high telomeric content samples exceeded the odds ratio observed for *ATRX*, a canonical telomere maintenance mechanism (TMM) gene, alterations in which are known to be associated with ALT. The impact of alterations in *RAD51B*, *GID4*, and *POT1* on telomeric content was proportional to tumor purity, but in samples WT for known TMM genes, telomeric content was not impacted by tumor purity (Fig. S3). Median telomeric content of *GID4* altered soft tissue sarcoma nos samples was significantly higher than *GID4* WT samples (1016 vs 662 TRPM, $p < 0.001$), *POT1* altered angiosarcoma was significantly higher than *POT1* WT (1485 vs 651 TRPM, $p < 0.001$), and *RAD51B* altered uterus leiomyosarcoma was significantly higher than *RAD51B* WT (1529 vs 710 TRPM, $p < 0.001$, Fig. 1D). Overall, the alteration prevalence for these genes was low, where *GID4* was altered in 3.6% (83/2332) of soft tissue sarcoma nos, *POT1* in 6.3% (38/606) of angiosarcoma, and *RAD51B* in 3.6% (38/1055) of uterus leiomyosarcoma samples (Table S3).

Types of genetic alterations in *RAD51B*, *GID4*, and *POT1*

Within uterus leiomyosarcomas, 84.2% of the identified *RAD51B* alterations were copy number deletions (Fig. 2A). Median size of the deleted segment was 565171 nucleotides, which is smaller than the size of *RAD51B* at 776243 nucleotides (Fig. 2B). Thus, we investigated which region of *RAD51B* was most frequently deleted and observed that in the majority of samples, the deletion spanned exons 4–11 corresponding to the ATPase domain (Fig. 2C)²⁹. We also observed that *RAD51B* alterations were commonly co-occurring with alterations in *FAF1* (26.5%), *CDKN2C* (23.5%), and *CIC* (14.7%, Fig. S4).

Within soft tissue sarcoma nos, all the *GID4* alterations were copy number amplifications with a median copy number of 7 (range 5–45 copies, Fig. 2D). Median size of the amplified segment was 4320468 nucleotides, which is smaller than the size of chr17p11.2 (6.2 Mb) where *GID4* lies, but larger than the size of *GID4* itself at 29108 nucleotides (Fig. 2E). We also assessed

whether the amplification of any of the cytobands on the p-arm of chr17 was associated with high telomeric content in soft tissue sarcoma nos. Our analysis showed that samples with amplifications in p11.2 on chr17 were enriched in samples with high telomeric content ($p = 0.05$, Fig. 2F). This association was not observed in any of the other cytobands on chr17p. We also observed that *GID4* altered samples were associated with alterations in *TP53* (83.3%) and *RB1* (47.8%, Fig. S4).

Within angiosarcoma, all the *POT1* alterations were short variant alterations, occurring throughout its length, where 70% of the mutations were observed only once and none were observed more than twice (Fig. 2G). For samples where zygosity could be assessed, the majority of samples (89%, 25/28) harbored a single-hit *POT1* alteration and in 11% (3/28) of samples, double-hit mutations were observed. One particular sample with a double-hit alteration had one mutation predicted to be somatic and another predicted to be of germline origin, while the remaining two samples with double-hit alterations had predicted homozygous somatic mutations. Median telomeric content was higher in the double-hit group vs single-hit but did not reach statistical significance (4266 vs 1570 TRPM, $p > 0.05$). However, both groups had significantly higher telomeric content compared to WT (median WT: 652 TRPM, $p < 0.01$ for double-hit, $p < 0.0001$ for single-hit, Fig. 2H). We also observed that *POT1* altered samples were enriched in samples with a high tumor mutational burden (TMB) and harboring a UV mutational signature (Fig. S4), similar to findings from an angiosarcoma landscape study which showed *POT1* alterations to be enriched in TMB high head and neck angiosarcomas³⁰.

Mutual exclusivity of genetic alterations and impact of co-occurrence

Prevalence of *RAD51B*, *GID4*, and *POT1* alterations varied widely across sarcoma subtypes and tended to mostly occur in diseases with high prevalence rates of alterations in the other telomere maintenance genes (Fig. 3A). Thus, we checked for mutual exclusivity and observed that alterations within *RAD51B* and *POT1* were predominantly mutually exclusive with alterations in *ATRX* and *DAXX* ($p = 8E-6$, Fig. 3B). This is suggestive of a redundancy between alterations in these genes, likely a result of affecting similar pathways. *GID4*, on the other hand, was co-altered with *ATRX* in 23% of *GID4* altered samples (Fig. 3B). Consequently, we queried the impact of a *GID4* alteration on top of an existing *ATRX* alteration on telomeric content. In soft tissue sarcoma nos, median telomeric content of the double-altered samples was higher than either single *GID4* altered or single *ATRX* altered samples, although neither reached statistical significance (1072 vs 1053 vs 874 TRPM, respectively, $p > 0.05$ for all comparisons, Fig. 3C). Similarly, in *ATRX* altered uterus leiomyosarcoma, *RAD51B* alterations exerted an additive effect on telomeric elongation (2226 TRPM in double-altered vs 1449 in *RAD51B* altered only and 812 in *ATRX* altered only, Fig. 3D). Across three different disease ontologies, when samples were ranked based on median telomeric content, alterations in the newly identified genes from our screen (*RAD51B*, *GID4*, and *POT1*) were seen in the groups with the highest or second highest telomeric content, or both (Fig. 3C–E).

Impact of alterations on TERRA levels

ALT has been linked to elevated levels of long noncoding telomeric repeat-containing RNA (TERRA), transcribed off telomeric repeats^{31–34}. Thus, we measured TERRA levels using TelomereHunter for select samples with available RNA sequencing data and compared between TERRA levels in samples with alterations in *RAD51B*, *GID4*, or *POT1* vs WT. We observed that the median level of TERRA expression was significantly higher in *GID4* altered soft tissue sarcoma nos vs WT (Fig. 4A, 22.7 vs 9.1, $p < 0.001$). We also observed a trend towards higher median TERRA levels in *POT1* altered angiosarcoma vs WT (Fig. 4B, 12.5 vs

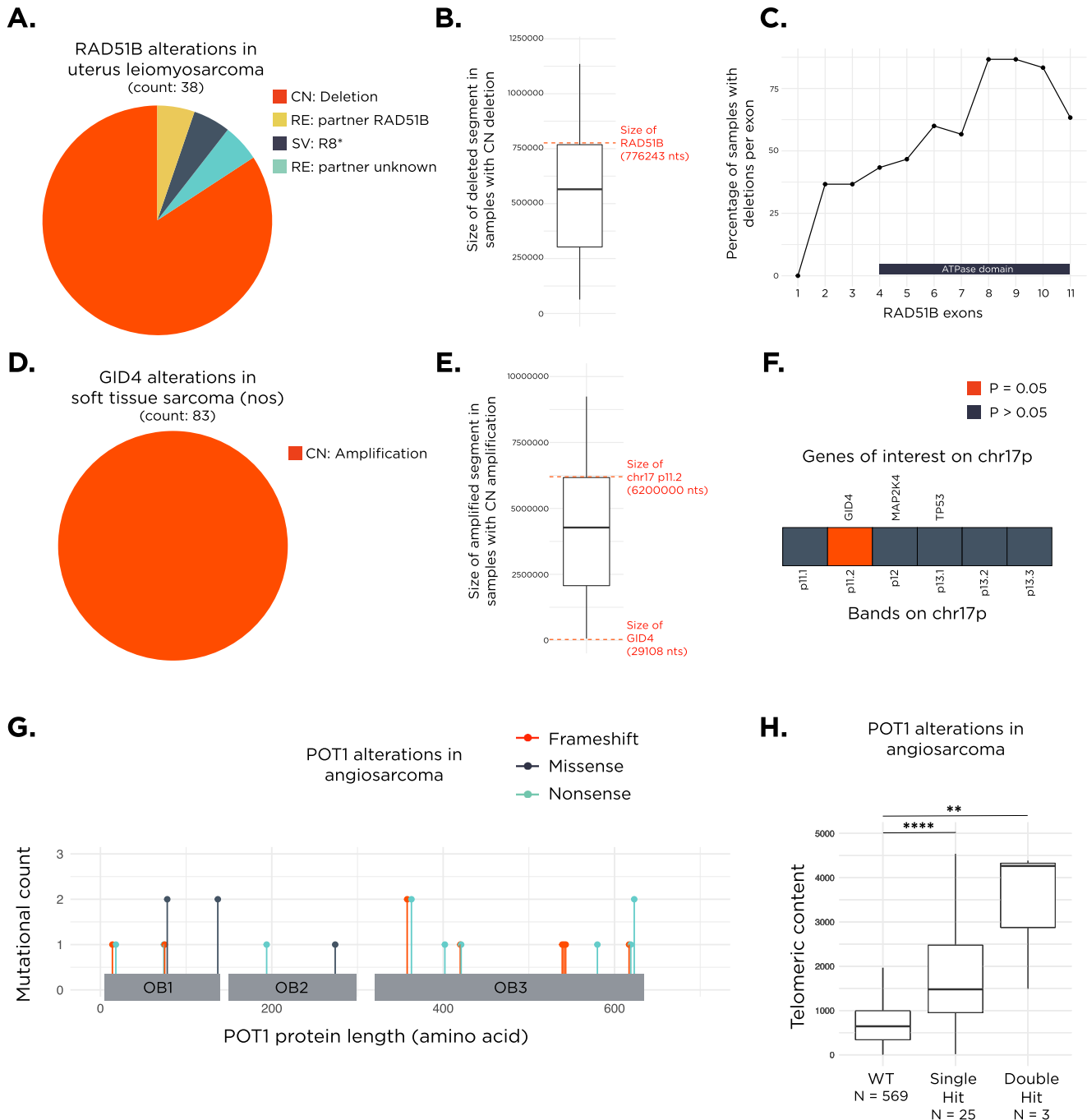


Fig. 2 Alteration types observed for *RAD51B*, *GID4*, and *POT1*. **A** Pie chart displaying the proportion of various *RAD51B* alterations observed across uterus leiomyosarcoma samples. **B** Within samples with a *RAD51B* copy-number deletion, the size of the deleted segment is displayed on the boxplot and a red dotted line denotes the size of the *RAD51B* gene in nucleotides. **C** A line plot displaying the proportion of samples with a *RAD51B* copy-number deletion harboring deletions within each exon of *RAD51B*. The position of the ATPase domain is indicated at the bottom. **D** Pie chart showing that all of the *GID4* alterations observed across soft tissue sarcoma nos samples are copy number amplifications. **E** Within samples with a *GID4* copy-number amplification, the size of the amplified segment is displayed on the boxplot and two red dotted lines indicate the size of the *GID4* gene and the size of the chr17p11.2 cytoband, respectively. **F** Analysis of enrichment for high telomeric content in samples with amplifications across chromosome bands in chr17p. Select genes of interest within certain cytobands are shown on the top. Orange indicates $p = 0.05$ and slate blue indicates $p > 0.05$. **G** Lollipop plot displaying the count of mutations observed across the Oligonucleotide/Oligosaccharide Binding (OB) fold domains of *POT1*. **H** Boxplot showing the distribution of telomeric content values across angiosarcoma samples with a *POT1* single hit mutation, double hit mutation or WT. CN copy number alteration, RE rearrangement, SV short variant alteration, nos, not otherwise specified, nts nucleotides, TPM transcripts per million. **** $p < 0.0001$ and ** $p < 0.01$. In all boxplots, the box extends from the first to the third quartile with a line in the middle that represents the median.

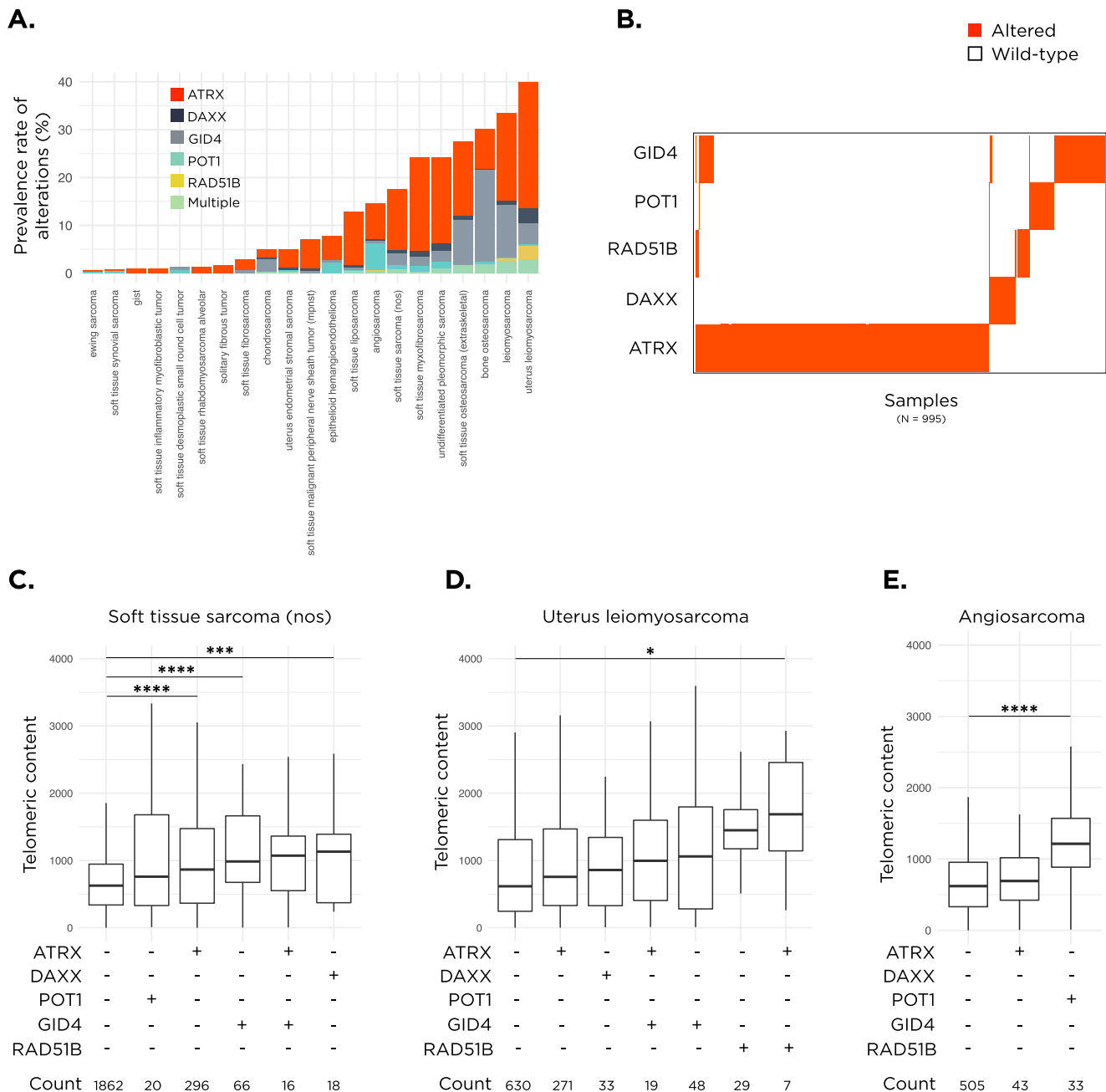


Fig. 3 Mutual exclusivity of RAD51B, POT1, ATRX, and DAXX. **A** Barplot indicating the prevalence rate of alterations within *ATRX*, *DAXX*, *GID4*, *POT1*, and *RAD51B* within each disease. Analysis restricted to diseases with at least 100 total samples. **B** Tile plot showing the distribution of alterations in *ATRX*, *DAXX*, *RAD51B*, *POT1*, and *GID4* within sarcoma samples in the FMI dataset. Plot depicts 995 sarcoma samples, which harbor at least one alteration in these genes. Altered samples are shown in orange and non-altered samples are shown as white. The impact of single and/or multiple alterations within these genes on telomeric content is shown for soft tissue sarcoma nos in **C**, uterus leiomyosarcoma samples in **D**, and angiosarcomas in **E**. Within each group, the symbol (+) means altered and the symbol (-) means non-altered. Groups are plotted if they contain >5 samples. Only comparisons against WT are shown. **** $p < 0.001$, *** $p < 0.001$, and * $p < 0.05$. In all boxplots, the box extends from the first to the third quartile with a line in the middle that represents the median.

5.8, $p > 0.05$) and a trend towards higher levels in *RAD51B* altered uterus leiomyosarcoma vs WT (Fig. 4C, 46.7 vs 30.1, $p > 0.05$).

Comparison of results to two independent datasets

Across all sarcoma samples within our screening cohort, samples with alterations in either *GID4*, *RAD51B*, *POT1*, or *ATRX* had significantly higher telomeric content than WT samples (Fig. 5A). All sarcoma samples in our screening cohort were tested using the FoundationOne®Heme platform. A minority of the sarcoma

samples received at our institution were run on a different testing platform, FoundationOne®CDx (F1CDx) and were not included in the screen. We thus used these 1739 F1CDx sarcoma samples for validation. We grouped samples based on the alteration status of their genes and compared between their telomeric content (Fig. 5B). Only groups with at least 10 samples are shown. Alterations in *ATRX* and *GID4* were associated with a significantly higher telomeric content compared to WT samples (median telomeric content 1263 TRPM in WT, 2146 TRPM in *GID4* altered, and 1646

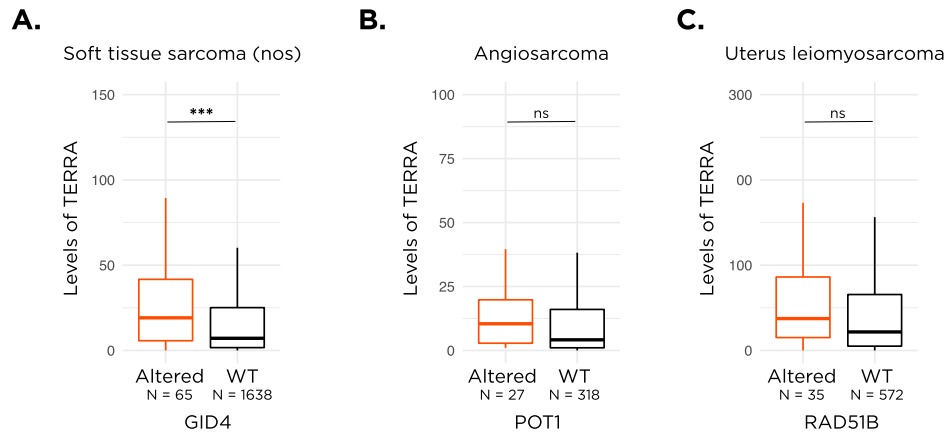


Fig. 4 Levels of TERRA expression. Boxplots depicting differences in levels of TERRA expression in *GID4* altered vs *GID4* WT soft tissue sarcoma nos samples (A), *POT1* altered vs *POT1* WT angiosarcoma samples (B), and *RAD51B* altered vs *RAD51B* WT uterus leiomyosarcoma samples (C). ns not significant. *** $p < 0.001$. In all boxplots, the box extends from the first to the third quartile with a line in the middle that represents the median.

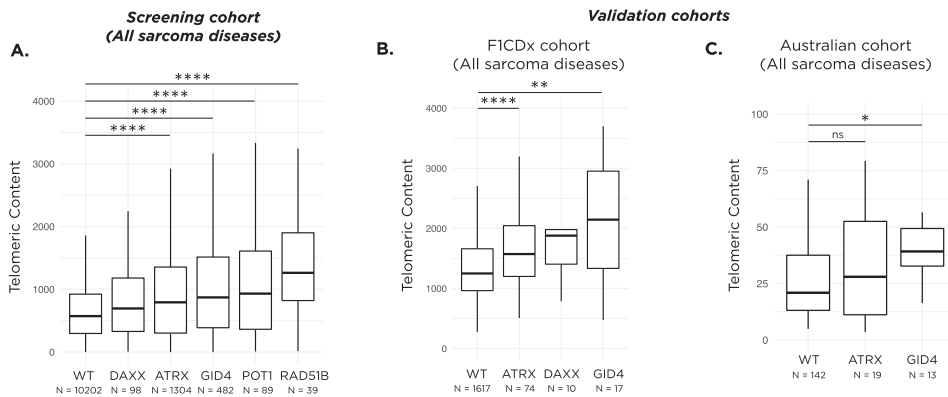


Fig. 5 Impact of alterations on telomeres pan-sarcoma. A Boxplot displaying the distribution of telomeric content values of samples in our screening cohort across all sarcoma diseases. **** $p < 0.0001$. Analysis was restricted to samples that are altered in only one of the telomere-maintenance mechanism genes. B Boxplot of the telomeric content values in the FICDx cohort, specifically for WT, *ATRX*, *DAXX*, and *GID4* altered soft tissue tumors. Analysis restricted to groups with at least 10 samples. ** $p < 0.01$ and **** $p < 0.0001$. C Boxplot of the telomeric content values in the Australian cohort, specifically for WT, *ATRX*, and *GID4* altered soft tissue tumors. Analysis restricted to groups with at least 10 samples. * $p < 0.05$; ns not significant. In all boxplots, the box extends from the first to the third quartile with a line in the middle that represents the median.

TRPM in *ATRX* altered). In addition, we conducted a disease-specific analysis for the FICDx cohort based on the sarcoma subtypes identified in our screen and observed that the results trended in the same direction (Fig. S5).

We also investigated the impact of genetic alterations on the telomeric content in an independent Australian dataset consisting of 174 sarcoma samples profiled on the TSO500 assay. *GID4* altered samples had significantly higher median telomeric content compared to WT (40.9 vs 22.6 TRPM, $p < 0.05$, Fig. 5C), while the median telomeric content of *ATRX* altered samples trended towards being higher (29.1 vs 22.6 TRPM, $p > 0.05$, Fig. 5C). Only groups with at least 10 samples were included in this analysis. Of note, panels from both independent validation datasets contained baiting for *TERC* and *TERT*. None of the samples altered in *RAD51B*, *GID4*, *ATRX*, or *DAXX* harbored alterations in *TERC* or *TERT*.

Impact of *GID4* alterations on overall survival in sarcoma

Next, we checked whether the presence of *GID4* alterations impacted overall survival in sarcoma patients with metastatic disease. To this purpose, we utilized two clinical cohorts. The first is the CGDB cohort, which consisted of 449 sarcoma patients with metastatic disease, 213 of which were soft tissue sarcoma patients.

The second is the Australian cohort, which consisted of 235 sarcoma patients with metastatic disease, 211 of which were soft tissue sarcoma patients. Patient demographics data is summarized in Table S4. No significant differences were observed in the mOS between *GID4* altered and *GID4* WT advanced sarcoma patients in either cohort (CGDB cohort: mOS of 5.6 months [3.0-NA] vs 9.8 months [7.4-12.2], $p = 0.4$; Australian cohort: mOS of NA months [19.7-NA] vs 41.3 months [30.9-59.9], $p = 0.6$, Fig. S6).

Similarly, when the analysis was restricted to patients with soft tissue sarcoma tumors, no significant differences were observed in the mOS between *GID4* altered and *GID4* WT patients in either cohort (CGDB cohort: mOS of 3.5 months [1.6-NA] vs 8.1 months [5.6-11.4], $p = 0.2$; Australian cohort: mOS of NA months [35.8-NA] vs 41.3 months [34.1-59.9], $p = 0.4$, Fig. S6).

DISCUSSION

In this study, we investigated the genetic determinants of high telomeric content across 13555 sarcoma samples in the Foundation Medicine dataset and found that samples with high telomeric content were enriched in deletions in *RAD51B*, amplifications in *GID4* (17p11.2), and short variant mutations in *POT1* within select disease ontologies. This finding was largely in agreement with

results from an additional dataset of sarcomas profiled on another Foundation Medicine platform, as well as in an independent academic dataset profiled on a different commercial platform.

Among sarcomas, multiple studies reported that a substantial proportion activates the ALT mechanism for telomere elongation^{23,25}, which is associated with the loss of *ATRX* and *DAXX*^{31,35}. However, it was also reported that *ATRX/DAXX* alterations were not present in all ALT-positive samples, with one study reporting that 45% of ALT-positive sarcoma samples were *ATRX* intact^{23,36}. In our dataset, only 12.3% of sarcomas harbored *ATRX/DAXX* alterations. Given the necessity to elongate telomeres for tumorigenesis, we hypothesized that the remaining 87.7% of sarcoma samples need to have either acquired mutations in other genes leading to ALT or activated non-ALT-mediated mechanisms of telomere lengthening.

Our screen identified three genes: *RAD51B*, *POT1*, and *GID4*, alterations in which were associated with higher telomeric content. *RAD51B* is a DNA-damage repair pathway gene. It is a paralog of *RAD51*, sharing a low level of sequence identity with *RAD51* (~25%), where most of the similarity is found within their ATPase domains^{29,37}. *RAD51*'s activity and its role in homologous recombination repair appear to be strictly regulated by a number of cofactors including *RAD51B*^{38,39}. Specifically, *RAD51B* is a component of the BCDX2 complex, that is responsible for the recruitment of *RAD51* in the early stages of homologous recombination⁴⁰. *RAD51* has been identified in ALT-associated promyelocytic leukemia (PML) bodies (APBs)⁴¹. Recently, in vitro studies showed that *RAD51* antagonizes the formation of C-circles⁴², which are extrachromosomal circular telomeric DNA considered a quantitative biomarker of the ALT mechanism^{43,44}. However the depletion of *RAD51* did not significantly reduce telomere DNA synthesis⁴².

In cell lines, *RAD51B* inactivation led to a comparatively mild phenotype, however the inactivation of other BCDX2 complex components resulted in a dramatic reduction in DNA double-strand break-induced homologous recombination^{27,28}. A recent cryo-EM study observed that entirely different surfaces of *RAD51B*, *RAD51C*, and *RAD51D* are involved in intermolecular interactions within BCDX2²⁹. To our knowledge, no studies have probed the impact of *RAD51B* deletion on telomere length. Our results show that its deletion is associated with higher telomeric content, suggesting that *RAD51B* may be a repressor of the BCDX2 complex. Further studies are needed to delineate the precise roles of *RAD51* and *RAD51B* in telomere elongation.

POT1 is a component of the shelterin complex and binds to single-stranded telomeric DNA at the ends of linear chromosomes⁴⁵. It was observed that heritable defects in *POT1* increase the risk of cancers, including sarcomas, and study participants with *POT1* mutations had significantly longer telomeres than age-matched controls^{46–52}. In vitro cell line studies showed that *POT1* modulates telomerase-mediated telomere elongation^{53,54}. In human embryonic stem cells, known to express telomerase⁵⁵, introducing mutations in *POT1* resulted in elongated telomeres⁵⁶. The induction of *POT1* ubiquitination and subsequent degradation induced cell death in ALT-positive cell lines, but was compatible with survival in telomerase-positive cells⁵⁷.

No prior studies have suggested a link between *GID4* and telomeres. *GID4* is a subunit of the GID ubiquitin ligase, which plays a role in glucose sensing and energy homeostasis^{58,59}. Others have reported that glucose levels in healthy individuals are negatively associated with telomere length⁶⁰. Furthermore, patients with type 2 diabetes harbor shorter telomeres^{61–64} and their telomere length was associated with glycemic progression and diabetic complications, such as nephropathy^{65–67}. In prospective studies where healthy non-diabetic participants were followed up for years, those with a lower telomere length at baseline were more likely to develop diabetes, with one study reporting a hazard ratio of 2.0 when comparing participants in the

bottom vs top quartile of telomere length^{68–70}. Consequently, it will be important to investigate the mechanisms mediating the interaction between telomere length and glucose metabolism. To our knowledge, no other genes on chr17p11.2 have been previously associated with telomere elongation either.

Overall, we observed a low alteration prevalence for the genes identified in our screen. *RAD51B* was altered in 3.6% (38/1055) of uterus leiomyosarcoma samples, lower than the previously reported *RAD51B* alteration prevalence of 7.4% (16/216)⁷¹ and 12.4% (13/105)⁷². Similarly, we observed a *POT1* alteration prevalence of 6.3% (38/606) in angiosarcoma samples, lower than previously reported at 15.7% (22/140) and 23.3% (20/86)^{30,73}. It is important to note that despite the low alteration prevalence we observed for these genes, the disease ontologies identified in our screen had the highest alteration prevalence rates for *RAD51B* (uterus leiomyosarcoma) and *POT1* (angiosarcoma) among all sarcoma diseases surveyed, both within our cohort, as well as in independent cohorts^{73,74}. In addition, we saw evidence that these alterations impacted telomeric content values when we analyzed the data pan-sarcoma. This suggests that the associations seen with telomere elongation within particular disease ontologies were observed due to the presence of sufficient sample counts and that the impact of these mutations may be extended to more diseases provided more samples are profiled, especially since the telomeric content values trended in the same direction in other sarcoma types (data not shown).

Overall, we found that alterations within *RAD51B*, *POT1*, *ATRX*, and *DAXX* were predominantly mutually exclusive, which frequently points to functional redundancy^{75,76}, providing an additional layer of evidence that all four genes promote telomere elongation. On the other hand, *GID4* was co-altered with *ATRX* in 23% of *GID4* altered samples. However the presence of a *GID4* alteration on top of an *ATRX* alteration led to even higher telomeric content, within both soft tissue sarcoma nos and uterus leiomyosarcoma samples. This suggests either that *GID4* activates an additional telomere elongation pathway or promotes the already-activated ALT pathway in those tumors.

Detecting telomere maintenance mechanisms may be beneficial in the prognosis of sarcomas. Specifically in osteosarcomas, the complete absence of a known telomere maintenance mechanism was strongly associated with improved survival⁷⁷. In some reports, identifying the mechanism driving telomere elongation was linked to overall survival. For example, in soft tissue sarcomas, ALT was associated with decreased survival compared to patients with telomerase activation^{26,78–82}, whereas in osteosarcomas, no difference in clinical outcomes was observed between ALT and telomerase-activated tumors⁷⁷. In our analysis, we identified no difference in median overall survival from the date of metastatic diagnosis in patients with *GID4* altered sarcoma tumors compared to their WT counterparts.

It is important to note the limitations of this study. The FoundationOne®Heme testing platform only captures certain regions of the genome, thus our analysis was limited to those baited regions. Specifically, testing on this platform does not include baiting for either *TERC* or *TERT*, hence telomerase-mediated lengthening could not be assessed across these samples. However, analyzing samples from the independent validation datasets, sequenced on different panels, revealed that none of the samples altered in *RAD51B*, *GID4*, *ATRX*, or *DAXX* harbored alterations in either *TERC* or *TERT*. Furthermore, our results demonstrate a strong association between the observed alterations and high telomeric content. Future wet-lab studies are needed to confirm causation and elucidate the telomere maintenance mechanisms activated in these altered samples.

METHODS

Foundation Medicine dataset

The Foundation Medicine sarcoma dataset consisted of 13555 sarcoma samples sequenced as a part of routine clinical care on the FoundationOne®Heme platform. All samples contained a minimum of 20% tumor nuclei. The pathologic diagnosis of each sarcoma case was first made by the referring institution and later confirmed by pathologists in Foundation Medicine on routine H&E staining. Samples represented 38 unique disease ontologies, with counts shown in Table S1. Counts are shown for sarcoma types with at least 5 samples. The top five disease ontologies present were soft tissue sarcoma (not otherwise specified, nos), soft tissue leiomyosarcoma, soft tissue liposarcoma, uterus leiomyosarcoma and bone osteosarcoma, accounting for 19.1%, 13.3%, 9.6%, 8.0% and 5.7% of the total sarcoma sample cohort, respectively. This is the main cohort referenced throughout the manuscript, unless otherwise indicated.

Approval for this study, including a waiver of informed consent and a HIPAA waiver of authorization, was obtained from the Western Institutional Review Board (Protocol No. 20152817) prior to study conduct. The Institutional Review Board granted a waiver of informed consent under 45 CFR § 46.116 based on review and determination that this research meets the following requirements: (i) the research involves no more than minimal risk to the subjects; (ii) the research could not practicably be carried out without the requested waiver; (iii) the waiver will not adversely affect the rights and welfare of the subjects. The work conforms to the principles of the Helsinki Declaration.

Sample sequencing

The main cohort of samples were sequenced using a targeted panel (FoundationOne®Heme)^{83,84}. Samples were assayed using baits for all coding exons of 465 cancer-related genes plus select introns from 31 genes frequently rearranged in cancer. Sequencing of captured libraries was performed using the Illumina sequencing platform to a median exon coverage depth for targeted regions of $\geq 250\times$, and subsequently analyzed for genomic alterations, including short variant alterations (base substitution and indels), copy number alterations (focal amplifications and homozygous deletions), and select gene fusions or rearrangements. *TERC* and *TERT* were not included in the FoundationOne®Heme panel. Tumor mutational burden (TMB) was called as previously described⁸⁴. Throughout this manuscript, “altered” refers to a sarcoma sample harboring known or likely pathogenic alterations in the indicated gene, whereas “WT” refers to a sample lacking these alterations or containing variants of unknown significance in the indicated gene.

Telomeric content

We used TelomereHunter 1.1.0, a publicly available tool, for the estimation of telomeric content from targeted sequencing data⁸⁵. TelomereHunter extracts telomeric reads and determines the telomeric content by normalizing the telomere read count by all reads in the sample with a comparable GC content (48–52%). We ran the tool on our sarcoma cohort using the default parameters and set the repeat threshold to 7 for 49 bp paired-end reads. For a subset of samples, RNA sequencing was also performed and TelomereHunter 1.1.0 was used for estimation of telomeric repeat-containing RNA (TERRA) content.

Independent sarcoma datasets

For validation purposes, we utilized two distinct datasets. The first validation dataset consists of 1739 sarcoma samples sequenced on another panel, namely the FoundationOne®CDx (F1CDx) platform^{83,84} and is referred to as the F1CDx cohort. This panel utilizes a baitset targeting all coding exons of 309 cancer-related

genes plus select introns from 34 genes frequently rearranged in cancer. Importantly, this panel includes baits for *RAD51B* and *GID4*, but not *POT1*. The second validation dataset, referred to as the Australian cohort, consists of 174 sarcoma samples profiled on the TSO500 panel, a targeted panel that contains baits for *RAD51B* and *GID4*, but not *POT1*⁸⁶. Samples selected passed filters set for tumor purity and sequencing depth.

Survival analysis of CGDB cohort

The retrospective clinical analysis utilized the nationwide (US-based) Foundation Medicine–Flatiron Health real-world clinico-genomic database (CGDB, data collected through December 31, 2021) which includes electronic health record (EHR)–derived deidentified data for patients in the Flatiron Health database who underwent comprehensive genomic profiling by Foundation Medicine, linked by de-identified deterministic matching⁸⁷. Institutional Review Board approval of the study protocol was obtained prior to study conduct and included a waiver of informed consent based on the observational, non-interventional nature of the study (WCG IRB, Protocol No. 420180044).

Clinical outcome data were available for 449 sarcoma patients, 213 of which were soft tissue sarcoma patients. Overall survival was calculated from date of metastatic diagnosis based on a composite mortality variable⁸⁸. To account for left truncation, patients were treated as at risk of death only after the later of their first sequencing report date and their second visit in the Flatiron Health network on or after January 1, 2011, as both are requirements for inclusion in the cohort. For the Kaplan–Meier analyses, the log-rank test was used to compare between *GID4* WT vs altered. Due to low sample count, survival analysis couldn't be performed for *RAD51B* or *POT1*.

Survival analysis of Australian cohort

Clinical outcome data were available for 235 sarcoma patients, 211 of which were soft tissue sarcoma patients. Overall survival was calculated from date of metastatic diagnosis. For the Kaplan–Meier analyses, the log-rank test was used to compare between *GID4* WT vs altered. Due to low sample count, survival analysis couldn't be performed for *RAD51B* or *POT1*.

Statistics

All statistical analyses were conducted using R (4.0.2)⁸⁹. Wilcoxon rank sum (two-sided) and Kruskal–Wallis were used to test for differences between two groups or between multiple groups, respectively. To screen for genetic alterations associated with high telomeric content, we first identified samples with telomeric content values in the top quartile of their respective disease and those with low content in the bottom quartile and then performed a Fisher's exact test to check for the enrichment of alterations across the baited genes. Results for all gene hits with a corrected *P* value < 0.05 after Bonferroni's correction and an odds ratio > 1 are plotted in Fig. 1B. For the analysis shown in Fig. 2F, we performed a Fisher's exact test to check for the enrichment of amplifications across chromosome bands of chr17p in samples with telomeric content in the top quartile. MEGSA (version beta 2) was used to test for mutual exclusivity in Fig. 3B⁹⁰. To denote significance, **p* < 0.05 , ***p* < 0.01 , ****p* < 0.001 , and *****p* < 0.0001 , while ns denotes non-significant.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

DATA AVAILABILITY

Academic researchers can gain access to Foundation Medicine data in this study by contacting the corresponding author and filling out a study review committee form. You and your institution will be required to sign a data transfer agreement. Survival data that support the findings of this study have been originated by Flatiron Health, Inc. and Foundation Medicine, Inc. These de-identified data may be made available upon request and are subject to a license agreement with Flatiron Health and Foundation Medicine; interested researchers should contact cgdb-fmi@flatiron.com to determine licensing terms.

Received: 17 May 2023; Accepted: 18 August 2023;

Published online: 14 September 2023

REFERENCES

- Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
- Huffman, K. E., Levene, S. D., Tesmer, V. M., Shay, J. W. & Wright, W. E. Telomere shortening is proportional to the size of the G-rich telomeric 3'-overhang. *J. Biol. Chem.* **275**, 19719–19722 (2000).
- Rahman, R., Forsyth, N. R. & Cui, W. Telomeric 3'-overhang length is associated with the size of telomeres. *Exp. Gerontol.* **43**, 258–265 (2008).
- Martens, U. M., Chavez, E. A., Poon, S. S., Schmoor, C. & Lansdorp, P. M. Accumulation of short telomeres in human fibroblasts prior to replicative senescence. *Exp. Cell Res.* **256**, 291–299 (2000).
- Wright, W. E., Tesmer, V. M., Huffman, K. E., Levene, S. D. & Shay, J. W. Normal human chromosomes have long G-rich telomeric overhangs at one end. *Genes Dev.* **11**, 2801–2809 (1997).
- Shay, J. W. & Wright, W. E. Senescence and immortalization: role of telomeres and telomerase. *Carcinogenesis* **26**, 867–874 (2005).
- D'Adda Di Fagagna, F. et al. A DNA damage checkpoint response in telomere-initiated senescence. *Nature* **426**, 194–198 (2003).
- Kim, N. W. et al. Specific association of human telomerase activity with immortal cells and cancer. *Science* **266**, 2011–2015 (1994).
- Shay, J. W. & Bacchetti, S. A survey of telomerase activity in human cancer. *Eur. J. Cancer* **33**, 787–791 (1997).
- Blackburn, E. H. et al. Recognition and elongation of telomeres by telomerase. *Genome* **31**, 553–560 (1989).
- Morin, G. B. The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. *Cell* **59**, 521–529 (1989).
- Bryan, T. M., Englezou, A., Dalla-Pozza, L., Dunham, M. A. & Reddel, R. R. Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. *Nat. Med.* **3**, 1271–1274 (1997).
- Henson, J. D., Neumann, A. A., Yeager, T. R. & Reddel, R. R. Alternative lengthening of telomeres in mammalian cells. *Oncogene* **21**, 598–610 (2002).
- Dunham, M. A., Neumann, A. A., Fasching, C. L. & Reddel, R. R. Telomere maintenance by recombination in human cells. *Nat. Genet.* **26**, 447–450 (2000).
- Lauer, N. K. et al. Absence of telomerase activity in malignant bone tumors and soft-tissue sarcomas. *Sarcoma* **6**, 43–46 (2002).
- Yoo, J. & Robinson, R. A. Expression of telomerase activity and telomerase RNA in human soft tissue sarcomas. *Arch. Pathol. Lab. Med.* **124**, 393–397 (2000).
- Schneider-Stock, R. et al. Prognostic significance of telomerase activity in soft tissue sarcomas. *Int. J. Oncol.* **15**, 775–780 (1999).
- Aogi, K. et al. Telomerase activity in soft-tissue and bone sarcomas. *Clin. Cancer Res.* **6**, 4776–4781 (2000).
- Koelsche, C. et al. TERT promoter hotspot mutations are recurrent in myxoid liposarcomas but rare in other soft tissue sarcoma entities. *J. Exp. Clin. Cancer Res.* **33**, 33 (2014).
- Campanella, N. C. et al. TERT promoter mutations in soft tissue sarcomas. *Int. J. Biol. Markers* **31**, e62–e67 (2016).
- Saito, T. et al. TERT promoter mutations are rare in bone and soft tissue sarcomas of Japanese patients. *Mol. Clin. Oncol.* **4**, 61–64 (2016).
- Dille, R. L. & Greenberg, R. A. Alternative telomere maintenance and cancer. *Trends Cancer* **1**, 145–156 (2015).
- Liau, J.-Y. et al. Comprehensive screening of alternative lengthening of telomeres phenotype and loss of ATRX expression in sarcomas. *Mod. Pathol.* **28**, 1545–1554 (2015).
- Heaphy, C. M. et al. Prevalence of the alternative lengthening of telomeres telomere maintenance mechanism in human cancer subtypes. *Am. J. Pathol.* **179**, 1608–1615 (2011).
- Henson, J. D. et al. A robust assay for alternative lengthening of telomeres in tumors shows the significance of alternative lengthening of telomeres in sarcomas and astrocytomas. *Clin. Cancer Res.* **11**, 217–225 (2005).

- Henson, J. D. & Reddel, R. R. Assaying and investigating alternative lengthening of Telomeres activity in human cells and cancers. *FEBS Lett.* **584**, 3800–3811 (2010).
- International Agency for Research on Cancer (IARC). *WHO Classification of Tumours of Soft Tissue and Bone* (World Health Organization, 2013).
- Chibon, F. et al. Validated prediction of clinical outcome in sarcomas and multiple types of cancer on the basis of a gene expression signature related to genome complexity. *Nat. Med.* **16**, 781–787 (2010).
- Kim, Y. M. & Choi, B.-S. Structural and functional characterization of the N-terminal domain of human Rad51D. *Int. J. Biochem. Cell Biol.* **43**, 416–422 (2011).
- Espejo-Freire, A. P. et al. Genomic landscape of angiosarcoma: a targeted and immunotherapy biomarker analysis. *Cancers* **13**, 4816 (2021).
- Lovejoy, C. A. et al. Loss of ATRX, genome instability, and an altered DNA damage response are hallmarks of the alternative lengthening of Telomeres pathway. *PLoS Genet.* **8**, 12–15 (2012).
- Barthel, F. P. et al. Systematic analysis of telomere length and somatic alterations in 31 cancer types. *Nat. Genet.* **49**, 349–357 (2017).
- Sieverling, L. et al. Genomic footprints of activated telomere maintenance mechanisms in cancer. *Nat. Commun.* **11**, 1–13 (2020).
- Sharaf, R. et al. A pan-cancer landscape of telomeric content shows that RAD21 and HGF alterations are associated with longer telomeres. *Genome Med.* **14**, 25 (2022).
- Dyer, M. A., Qadeer, Z. A., Valle-Garcia, D. & Bernstein, E. ATRX and DAXX: Mechanisms and mutations. *Cold Spring Harb. Perspect. Med.* **7**, a026567 (2017).
- de Nonneville, A. & Reddel, R. R. Alternative lengthening of telomeres is not synonymous with mutations in ATRX/DAXX. *Nat. Commun.* **12**, 1552 (2021).
- Miller, K. A., Sawicka, D., Barsky, D. & Albaladejo, J. S. Domain mapping of the Rad51 paralogs protein complexes. *Nucleic Acids Res.* **32**, 169–178 (2004).
- Takata, M. et al. The Rad51 Paralog Rad51B promotes homologous recombinational repair. *Mol. Cell Biol.* **20**, 6476–6482 (2000).
- Takata, M. et al. Chromosome instability and defective recombinational repair in knockout mutants of the five Rad51 paralogs. *Mol. Cell Biol.* **21**, 2858–2866 (2001).
- Chun, J., Buechelmaier, E. S. & Powell, S. N. Rad51 paralog complexes BCDX2 and CX3 act at different stages in the BRCA1-BRCA2-dependent homologous recombination pathway. *Mol. Cell Biol.* **33**, 387–395 (2013).
- Yeager, T. R. et al. Telomerase-negative immortalized human cells contain a novel type of promyelocytic leukemia (PML) body. *Cancer Res.* **59**, 4175–4179 (1999).
- Zhang, J.-M., Yadav, T., Ouyang, J., Lan, L. & Zou, L. Alternative lengthening of telomeres through two distinct break-induced replication pathways. *Cell Rep.* **26**, 955–968.e3 (2019).
- Henson, J. D. et al. DNA C-circles are specific and quantifiable markers of alternative lengthening-of-telomeres activity. *Nat. Biotechnol.* **27**, 1181–1185 (2009).
- Cesare, A. J. & Reddel, R. R. Alternative lengthening of telomeres: models, mechanisms and implications. *Nat. Rev. Genet.* **11**, 319–330 (2010).
- Baumann, P. & Cech, T. R. Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science* **292**, 1171–1175 (2001).
- Schooley, R. T. et al. AIDS clinical trials group 5197: a Placebo-controlled trial of immunization of HIV-1-infected persons with a replication-deficient adenovirus type 5 vaccine expressing the HIV-1 core protein. *J. Infect. Dis.* **202**, 705–716 (2010).
- Robles-Espinoza, C. D. et al. POT1 loss-of-function variants predispose to familial melanoma. *Nat. Genet.* **46**, 478–481 (2014).
- Calvete, O. et al. A mutation in the POT1 gene is responsible for cardiac angiosarcoma in TP53-negative Li-Fraumeni-like families. *Nat. Commun.* **6**, 8383 (2015).
- Bainbridge, M. N. et al. Germline mutations in shelterin complex genes are associated with familial glioma. *J. Natl Cancer Inst.* **107**, 384 (2015).
- Shi, J. et al. Rare missense variants in POT1 predispose to familial cutaneous malignant melanoma. *Nat. Genet.* **46**, 482–486 (2014).
- DeBoy, E. A. et al. Familial clonal hematopoiesis in a long telomere syndrome. *N. Engl. J. Med.* <https://doi.org/10.1056/NEJMoa2300503> (2023).
- Gong, Y., Stock, A. J. & Liu, Y. The enigma of excessively long telomeres in cancer: lessons learned from rare human POT1 variants. *Curr. Opin. Genet. Dev.* **60**, 48–55 (2020).
- Loayza, D. & De Lange, T. POT1 as a terminal transducer of TRF1 telomere length control. *Nature* **423**, 1013–1018 (2003).
- Wang, F. et al. The POT1-TPP1 telomere complex is a telomerase processivity factor. *Nature* **445**, 506–510 (2007).
- Hiyama, E. & Hiyama, K. Telomere and telomerase in stem cells. *Br. J. Cancer* **96**, 1020–1024 (2007).
- Kim, W.-T. et al. Cancer-associated POT1 mutations lead to telomere elongation without induction of a DNA damage response. *EMBO J.* **40**, e107346 (2021).
- Episkopou, H., Diman, A., Claude, E., Vicoconte, N. & Decottignies, A. TSPYL5 depletion induces specific death of ALT cells through USP7-dependent proteasomal degradation of POT1. *Mol. Cell* **75**, 469–482.e6 (2019).

58. Mohamed, W. I. et al. The human GID complex engages two independent modules for substrate recruitment. *EMBO Rep.* **22**, e52981 (2021).
59. Lampert, F. et al. The multi-subunit GID/CTLH E3 ubiquitin ligase promotes cell proliferation and targets the transcription factor Hbp1 for degradation. *Elife* **7**, e35528 (2018).
60. Grunnet, L. G. et al. Leukocyte telomere length is associated with elevated plasma glucose and HbA1c in young healthy men independent of birth weight. *Sci. Rep.* **9**, 7639 (2019).
61. Jeanclous, E. et al. Shortened telomere length in white blood cells of patients with IDDM. *Diabetes* **47**, 482–486 (1998).
62. Qi Nan, W., Ling, Z. & Bing, C. The influence of the telomere-telomerase system on diabetes mellitus and its vascular complications. *Expert Opin. Ther. Targets* **19**, 849–864 (2015).
63. Wang, J. et al. Association between telomere length and diabetes mellitus: A meta-analysis. *J. Int. Med. Res.* **44**, 1156–1173 (2016).
64. Tamura, Y. et al. β -cell telomere attrition in diabetes: inverse correlation between HbA1c and telomere length. *J. Clin. Endocrinol. Metab.* **99**, 2771–2777 (2014).
65. Verzola, D. et al. Accelerated senescence in the kidneys of patients with type 2 diabetic nephropathy. *Am. J. Physiol. Ren. Physiol.* **295**, F1563–F1573 (2008).
66. Fyhrquist, F. et al. Telomere length and progression of diabetic nephropathy in patients with type 1 diabetes. *J. Intern. Med.* **267**, 278–286 (2010).
67. Cheng, F. et al. Shortened leukocyte telomere length is associated with glycemic progression in type 2 diabetes: a prospective and mendelian randomization analysis. *Diabetes Care* **45**, 701–709 (2022).
68. Willeit, P. et al. Leucocyte telomere length and risk of type 2 diabetes mellitus: new prospective cohort study and literature-based meta-analysis. *PLoS One* **9**, e112483 (2014).
69. Zhao, J. et al. Short leukocyte telomere length predicts risk of diabetes in american indians: the strong heart family study. *Diabetes* **63**, 354–362 (2014).
70. Verhulst, S. et al. A short leukocyte telomere length is associated with development of insulin resistance. *Diabetologia* **59**, 1258–1265 (2016).
71. Astolfi, A. et al. Genomic database analysis of uterine leiomyosarcoma mutational profile. *Cancers* **12**, 2126 (2020).
72. Ciccarone, F. et al. Role of Homologous Recombination Repair (HRR) genes in uterine leiomyosarcomas: a retrospective analysis. *Cancers (Basel)* **14**, 1934 (2022).
73. Wu, Y., Poulos, R. C. & Reddel, R. R. Role of POT1 in human cancer. *Cancers (Basel)* **12**, 2739 (2020).
74. Shen, E. et al. POT1 mutation spectrum in tumour types commonly diagnosed among POT1-associated hereditary cancer syndrome families. *J. Med. Genet.* **57**, 664–670 (2020).
75. Ciriello, G., Cerami, E., Sander, C. & Schultz, N. Mutual exclusivity analysis identifies oncogenic network modules. *Genome Res.* **22**, 398–406 (2012).
76. Yeang, C.-H., McCormick, F. & Levine, A. Combinatorial patterns of somatic gene mutations in cancer. *FASEB J.* **22**, 2605–2622 (2008).
77. Ulaner, G. A. et al. Absence of a telomere maintenance mechanism as a favorable prognostic factor in patients with osteosarcoma. *Cancer Res.* **63**, 1759–1763 (2003).
78. Matsuo, T. et al. Alternative lengthening of telomeres as a prognostic factor in malignant fibrous histiocytomas of bone. *Anticancer Res.* **30**, 4959–4962 (2010).
79. Venturini, L., Motta, R., Gronchi, A., Daidone, M. & Zaffaroni, N. Prognostic relevance of ALT-associated markers in liposarcoma: a comparative analysis. *BMC Cancer* **10**, 254 (2010).
80. Liao, J.-Y. et al. Leiomyosarcoma with alternative lengthening of telomeres is associated with aggressive histologic features, loss of ATRX expression, and poor clinical outcome. *Am. J. Surg. Pathol.* **39**, 236–244 (2015).
81. Lee, Y.-K., Park, N.-H. & Lee, H. Prognostic value of alternative lengthening of telomeres-associated biomarkers in uterine sarcoma and uterine carcinosarcoma. *Int. J. Gynecol. Cancer* **22**, 434–441 (2012).
82. Costa, A. et al. Telomere maintenance mechanisms in liposarcomas: association with histologic subtypes and disease progression. *Cancer Res.* **66**, 8918–8924 (2006).
83. Frampton, G. M. et al. Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. *Nat. Biotechnol.* **31**, 1023–1031 (2013).
84. Chalmers, Z. R. et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med.* **9**, 34 (2017).
85. Feuerbach, L. et al. TelomereHunter—In silico estimation of telomere content and composition from cancer genomes. *BMC Bioinforma.* **20**, 1–11 (2019).
86. Kosco, K., Sickler, B., Chan, S., Alnaeeli, M. & Beruti, S. Abstract 4262: Analytical validation of TruSight Oncology 500, a comprehensive genomic profiling LDT for use in routine patient care. *Cancer Res.* **80**, 4262–4262 (2020).
87. Singal, G. et al. Association of patient characteristics and tumor genomics with clinical outcomes among patients with non-small cell lung cancer using a clinico-genomic database. *JAMA* **321**, 1391–1399 (2019).
88. Zhang, Q., Gossai, A., Monroe, S., Nussbaum, N. C. & Parrinello, C. M. Validation analysis of a composite real-world mortality endpoint for patients with cancer in the United States. *Health Serv. Res.* **56**, 1281–1287 (2021).
89. R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, Vienna, 2018).
90. Hua, X. et al. MEGSA: A powerful and flexible framework for analyzing mutual exclusivity of Tumor Mutations. *Am. J. Hum. Genet.* **98**, 442–455 (2016).

ACKNOWLEDGEMENTS

We thank patients who permitted the use of their samples for research. Research was supported through funding from the National Health and Medical Research Council for an L3 Investigator grant (GNT1195742) and from Omico Australia. This work was self-funded by Foundation Medicine, Inc.

AUTHOR CONTRIBUTIONS

RS, LAA, DMT and MM designed and conducted the study. RS, DXJ, EE, JG, CN and MM collected and analyzed the data. RS and MM wrote and edited the manuscript. GMF, LAA, DMT and MM supervised the work. All authors read and approved the final manuscript.

COMPETING INTERESTS

RS, DXJ, EE, GMF, LAA, and MM are employees of Foundation Medicine, Inc., a wholly owned subsidiary of Roche Holdings, Inc. and Roche Finance Ltd, and these employees have equity interest in an affiliate of these Roche entities. DT is the CEO of Omico, which provides national precision oncology services in Australia. DT has commercial and academic partnerships with multiple industry partners, including: Roche, Pfizer, Eisai, Astra Zeneca, Amgen, Abbvie, Eli Lilly, Beigene, Bayer, Illumina, Foundation Medicine, Microba, Merck, GMDx, Biotesse, InterVenn. DT acts in a consultant capacity to the Health Futures Fund of Australian Unity, and the Maine Cancer Genome Initiative. All other authors don't have conflicts of interest.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41525-023-00369-6>.

Correspondence and requests for materials should be addressed to Meagan Montesin.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023