Telomeres in ICF syndrome cells are vulnerable to DNA damage due to elevated DNA:RNA hybrids

Shira Sagie1, Shir Toubiana1,*, Stella R. Hartono2,*, Hagar Katzir1, Aya Tzur-Gilat1, Shany Havazelet1, Claire Francastel3, Guillaume Velasco3, Frédéric Chédin2 & Sara Selig1

DNA:RNA hybrids, nucleic acid structures with diverse physiological functions, can disrupt genome integrity when dysregulated. Human telomeres were shown to form hybrids with the IncRNA TERRA, yet the formation and distribution of these hybrids among telomeres, their regulation and their cellular effects remain elusive. Here we predict and confirm in several human cell types that DNA:RNA hybrids form at many subtelomeric and telomeric regions. We demonstrate that ICF syndrome cells, which exhibit short telomeres and elevated TERRA levels, are enriched for hybrids at telomeric regions throughout the cell cycle. Telomeric hybrids are associated with high levels of DNA damage at chromosome ends in ICF cells, which are significantly reduced with overexpression of RNase H1. Our findings suggest that abnormally high TERRA levels in ICF syndrome lead to accumulation of telomeric hybrids that, in turn, can result in telomeric dysfunction.
Telomeres are nucleoprotein complexes that maintain the integrity and stability of eukaryotic chromosome ends1–3. T-loop formation, binding of shelterin and non-shelterin proteins and a unique chromatin structure, all protect the chromosome ends from deleterious events such as degradation and end-to-end fusions3. Human telomeres undergo gradual attrition with each cell division, and when they reach a critical length, telomere shortening elicits replicative senescence4, with attrition with each cell division, and when they reach a critical length, telomere shortening elicits replicative senescence4, with attrition with each cell division, and when they reach a critical length, telomere shortening elicits replicative senescence4, with attrition with each cell division, and when they reach a critical length, telomere shortening elicits replicative senescence4, with.

TERRA has been implicated in numerous telomeric roles, such as regulation of telomere length, replication and heterochromatinization14–19 (reviewed in refs 2, 13). Evidence is emerging that the function and regulation of TERRA are telomere state dependent such that telomere length, telomerase expression and ALT pathway activity can influence the role that TERRA has at telomeres (reviewed in ref. 20).

R-loops, three-stranded nucleic acid structures that consist of a DNA:RNA hybrid and a displaced single-stranded DNA loop31, are predisposed by strand asymmetry in the distribution of guanines and cytosines, termed GC-skewing. These structures form mainly co-transcriptionally when positive GC skew is present such that DNA:RNA hybrids form between the G-rich RNA strand and the C-rich complementary DNA strand32. Although various studies indicate that DNA:RNA hybrids have a positive effect on gene transcription and are beneficial to the cell22–25, these structures have also been shown to mediate genome instability and replication stress26. R-loops have been implicated in human diseases, including trinucleotide expansion diseases, neurological diseases and cancer (reviewed in ref. 27). Telomeric DNA and TERRA transcripts are predicted to form hybrids, with the G-rich (UUAGGG)₉ TERRA transcript annealing to the C-rich (CCCTAA), DNA template. Indeed, recent studies support the existence of such hybrids at telomeres in *S. cerevisiae* (whose telomeres are comprised of a different G-rich repeat)14,28,29 and suggest that, in the absence of a telomere-maintenance mechanism, TERRA-telomeric DNA hybrids may promote accelerated telomere loss in *cis*14. In addition, telomeric DNA:RNA hybrids were found in various human cancer cells, both telomerase-positive and ALT cancerous cells30. In the latter, DNA:RNA hybrids were suggested to have a role in facilitating telomere recombination.

ICF (Immunodeficiency, Centromeric instability and Facial anomalies) syndrome type I is a rare autosomal-recessive disease, caused by hypomorphic mutations in the *ICF2* gene31,32, the major DNA methyltransferase involved in *de novo* methylation of repetitive sequences in mammalian cells during development32. Subtelomeres, as other repetitive sequences, are severely hypomethylated in ICF type I syndrome cells35,36. We detected accelerated telomere shortening and significant telomere loss, premature replicative senescence and significantly elevated levels of TERRA transcripts in both ICF fibroblast and lymphoblastoid cells (LCLs)33,35,36. Although it was proposed that TERRA has a causative role in the generation of telomeric abnormalities in ICF syndrome14,17,33–35, the underlying mechanism by which this occurs is as yet unclear.

Here we further investigate the occurrence of human telomeric hybrids in various cell types. Furthermore, we address the question of whether all telomeres are equally competent in generating these hybrids and whether the subtelomeric regions may affect this capacity. Our findings establish that telomeric DNA:RNA hybrids occur also in primary human cells and that subtelomeric sequences have an effect on generation of telomeric hybrids. We demonstrate that elevated TERRA levels are associated with higher levels of telomeric hybrids in ICF syndrome and suggest a role for these DNA:RNA hybrids in promoting damage and instability at telomeric regions in this disease.

**Results**

**Human subtelomeres are predicted to form DNA:RNA hybrids.** Human telomere-hexameric (TTAGGG)₉ repeats are predicted to form DNA:RNA hybrids, with the C-rich template annealing to the G-rich TERRA transcript. We validated this capacity in *vitro* and demonstrated, as in a previous study30, that these hybrids are formed only in a specific direction and are sensitive to RNase H, an enzyme that specifically degrades RNA strands within DNA:RNA hybrids (Supplementary Fig. 1).

The majority of TERRA transcripts initiate at the last few hundred base-pairs (bps) of the subtelomeric region, although some TERRA species may start 5–10 kb upstream of the telomere tract38. As most DNA:RNA hybrids are assumed to form co-transcriptionally22,39, we speculated that subtelomeric sequences might facilitate the formation of telomeric hybrids. To test this hypothesis, we first analysed the sequence of the distal 2 kb region adjacent to the telomere tract at both chromosome ends for CpG density, GC content and GC skew30. Regions with a strong GC skew downstream of the TERRA promoter may be prone to DNA:RNA hybrid formation. For this analysis, we utilized the previously described subtelomeric sequences8,10, focussing on high-confidence subtelomeric regions whose sequence is available in the UCSC GRCh38/hg38 release with a clearly defined telomeric region or at least three consecutive TTAGGG repeats at the 3′ end. These subtelomeric regions were overlaid with the predicted TERRA promoters and transcription start sites (TSSs), as determined by the Genomatix software40.

Most human subtelomeric regions exhibit high CpG density and GC content in regions corresponding to the predicted promoters for TERRA (Fig. 1a), thus closely resembling CpG island promoters. This is consistent with a similar analysis of a subgroup of TERRA promoters8 and reinforced by the findings that TERRA transcribing telomeres show higher GC content in comparison to the non-transcribing ones38. Examination of GC skew revealed that its levels are variable over the entire 2 kb region. However, a substantial number, but not all, of the analysed subtelomeres possess positive GC skew in the region located immediately downstream of the predicted TERRA TSSs (Fig. 1a). Thus a subgroup of subtelomeres appears prone to form DNA:RNA hybrids at the end of chromosome arms, which could then extend into the neighbouring telomeric regions.

**Subtelomeric DNA:RNA hybrids form in human cell lines.** The analysis above prompted us to test whether DNA:RNA hybrids indeed form at distal human subtelomeric regions in *vivo*. Quantification of DNA:RNA hybrids is classically carried out by DNA:RNA hybrid immunoprecipitation (DRIP), a procedure that involves immunoprecipitation of these hybrids with the S9.6 antibody22, and subsequently quantitative PCR (qPCR) or sequencing. For this, we analysed available genome-wide DRIP-
Figure 1 | Sequence characteristics and potential for DNA:RNA hybrid formation in human distal subtelomeres. (a) 2 kb distal subtelomeric regions immediately adjacent to the telomere repeat tract were analysed for CpG density, GC content and GC skew over tiled overlapping 200 bp windows. The value for each window is depicted using a colour heatmap, as indicated below; each tick mark corresponds to a window. The analysed subtelomeres, indicated on the right, are clustered by the GC skew downstream of the annotated TERRA promoter, when available (grey: no annotation). Predicted Pol2 promoters and putative TSSs are indicated by boxes and black vertical lines, respectively. (b) DRIP-seq signal over each subtelomeric region is indicated by a colour heatmap reflecting signal enrichment over input for each 2 kb window (colour scheme is indicated at the right). A, B, and C refer to three distinct DRIP-seq datasets in human fibroblast (Fibro - A) and human Ntera2 cells (NT2 - B and C) measured relative to input. Each dataset included two independent technical replicates. The presence of consistent signal peaks in each replicate within a dataset is noted by an asterisk (*), as indicated. D corresponds to C but measured relative to an RNase H-treated control. Subtelomeric regions are arranged in the same order as a. See ‘Results’ for details. (c) Representative screenshots of DRIP-seq data over three distinct subtelomeric regions (22q, 15q and 10q, as indicated). Normalized DRIP-seq signal densities are displayed for each region over two distinct replicates (indicated by (I) and (II)) as well as input and, when available, RNase H-treated controls. The position of the Pol2 promoters and putative TSSs is shown at the top. Vertical dashed lines indicate the position of restriction enzyme sites used to fragment the genomic DNA before DRIP-seq. The grey shaded area highlights a TRF encompassing the TERRA promoter and/or downstream regions showing significant DRIP-seq signal.
Telomeric DNA:RNA hybrids are elevated in ICF cells. Primary human fibroblasts were shown previously to express low levels of TERRA\(^3\), however, these levels appear to be sufficient for hybrid formation at subtelomeric regions, as demonstrated above. As hybrid formation is correlated with transcript levels\(^26,42\), we next asked whether the levels of DNA:RNA subtelomeric/telomeric hybrids would be influenced by TERRA levels. Cells from ICF type I syndrome patients provide an excellent platform to answer this question owing to their abnormally elevated TERRA levels. ICF cells were shown previously to display elevated levels of TERRA\(^3,33,34,37\) and here we validate this finding for individual chromosome ends (Supplementary Fig. 4). We compared TERRA levels of five LCLs generated from ICF syndrome patients with four LCLs from normal individuals and three heterozygous carriers of ICF type I syndrome. We did not detect statistically significant differences between the normal and carrier LCLs throughout the various analyses; therefore, these samples were pooled together for further analyses and designated as the wild-type (WT) control group. Quantitative reverse transcription–PCR (qRT-PCR) analysis demonstrated that, while TERRA levels varied among the different subtelomeres, they were consistently higher in ICF compared with WT LCLs.

We then proceeded to determine the levels of DNA:RNA hybrid formation in ICF versus WT cells by DRIP analysis. Digested DNA from ICF and WT LCLs (Fig. 2a) was subjected to DRIP using the S9.6 hybrid-specific antibody, and the amount of immunoprecipitated material was compared with input using qPCR for 11 subtelomeric regions and 3 control non-telomeric regions (Fig. 2b,c and Supplementary Fig. 5a). No statistically significant differences were observed in the enrichment of hybrids at control regions between the WT and ICF groups, in agreement with similar transcript levels of these genes (Supplementary Fig. 5b). Treatment of the DNA samples prior to DRIP with bacterial RNase H tested the specificity of the DRIP procedure, and, as expected, RNase H treatment significantly reduced or eliminated DRIP-qPCR signals emanating from both control and subtelomeric regions (Supplementary Fig. 5c,d). This analysis reveals that subtelomeric DNA:RNA hybrids form also in LCLs and that hybrid levels vary among subtelomeres both in the WT and the ICF groups (Fig. 2b,c). One-way analysis of variance for repeated measurements revealed a significant difference (P value < 0.05) in the DRIP enrichment between the subtelomeres, for both the WT and ICF groups. Importantly, the amount of DNA:RNA hybrids at many of the tested subtelomeres was higher for ICF cells than for WT cells (Fig. 2b,c), suggesting that high TERRA expression leads to elevated formation of DNA:RNA hybrids at chromosome ends. DRIP enrichment levels of each individual LCL segregated the ICF from the WT group and six of the tested subtelomeric regions showed a statistically significant difference in hybrid formation between both groups (Fig. 2d). Five of the subtelomeric regions tested here did not show significant differences in DNA:RNA hybrid loads between WT and LCLs, even though lower average hybrid levels in WT cells were apparent for 10 of the 11 examined subtelomeres. Notably, the three subtelomeres that most substantially differed in TERRA levels between ICF and WT LCLs (2p, 10q and 15q) were also those that displayed significantly higher hybrid formation in ICF versus WT (Fig. 2d and Supplementary Fig. 4). Interestingly, the cumulative GC skew calculated downstream of the TERRA TSS/promoter segregated subtelomeres whose hybrid formation was affected by high TERRA levels in ICF cells from those that were not affected (Fig. 2e). Collectively, these data suggest that the formation of DNA:RNA hybrids at chromosome ends is influenced by the sequence characteristics of the region adjacent to the telomeric tract, including the TERRA promoter and TSS, and is promoted by elevated levels of TERRA and/or by perturbation of the normal chromatin structure in these regions in ICF syndrome cells.

(TTAGGG)\(_n\) repeats are crucial for forming chromosome end hybrids. As demonstrated above, the ability of human telomeric repeats to form DNA:RNA hybrids was confirmed \textit{in vitro} by us (Supplementary Fig. 1) and another study\(^10\). Direct study of telomere repeat hybrids \textit{in vivo} is impeded by the repetitive nature of the regions, which prevents their amplification by PCR. Therefore, DRIP analysis of terminal restriction fragments (TRFs), utilizing subtelomeric primers, does not distinguish whether hybrids are formed at telomere-repeats or subtelomeric regions or both (Fig. 2a). In order to allow this distinction, we searched for a common restriction site that would separate the telomeric tract from the subtelomeric region downstream to the TERRA promoter in several chromosome ends. We detected five chromosome ends (7q, 8p, 9p, 13q and 21q) in which \textit{Hinfl} restriction sites are properly positioned (Fig. 3a) and performed DRIP on these subtelomeres. In this set of experiments, we first digested all samples with the standard enzyme cocktail and then set aside one half of each sample for further digestion with \textit{Hinfl}. We then validated that the samples were digested efficiently, as shown in Supplementary Fig. 6 (boxed section of the figure). We analysed the paired samples (+ / − \textit{H}) by DRIP-qPCR for four ICF LCLs (Fig. 3b,c) and four WT LCLs (Fig. 3d,e), including subtelomeres that encompass a \textit{Hinfl} site (+ \textit{Hinfl}) (Fig. 3b,d) and subtelomeres lacking a distal \textit{Hinfl} site between the hexameric repeat and the PCR-amplified region (− \textit{Hinfl}) (Fig. 3c,e and Supplementary Fig. 7). The DRIP enrichment values obtained in the absence of \textit{Hinfl} digestion (− \textit{H}) were
ICF samples, this decrease was not statistically significant subtelomeres that were affected in ICF; however, in contrast to
NATURE COMMUNICATIONS | DOI: 10.1038/ncomms14015 | www.nature.com/naturecommunications
decrease in hybrid enrichment after hybrids in these regions. In WT samples, we noticed an average TRF, are not responsible for the generation of the DNA:RNA that the upstream regions, now disconnected from the shorter digestion, the hybrid enrichment levels were higher, indicating HinfI Analysis of subtelomere 22q revealed that, following results in the release of a shorter TRF (Supplementary Fig. 7). The subtelomere are displayed for each LCL (WT samples—blue triangles, ICF samples—red triangles). Two-tailed Student’s t-tests were performed to determine statistical differences between WT and ICF samples (** = P value < 0.01, * = P value < 0.05). Six out of the 11 subtelomeres studied showed significant differences. (e) GC skew downstream of TERRA promoter is predictive of increased DNA:RNA hybrid formation in ICF compared with WT LCLs. GC skew information is as shown in Fig. 1a. The telomeres that appear in this figure are those that contain a putative TSS.

arbitrarily set at 1, and the enrichment values obtained following HinfI digestion (+ H) represent relative enrichment in comparison to the − H samples (Fig. 3b–e) (absolute percentage of input data are shown in Supplementary Fig. 6 and Supplementary Table 1).

In the ICF samples, at all of the subtelomeres containing a HinfI site close to the telomere tract, with the exception of 7q, a significant reduction in hybrid enrichment (P value < 0.001) was observed following removal of the telomeric sequence by HinfI digestion (Fig. 3b). This reduction was not observed at subtelomeres lacking a HinfI restriction site between the telomere repeat and the region analysed by amplification (Fig. 3c and Supplementary Figs 6 and 7), indicating that the additional HinfI digestion per se did not affect DRIP efficiency. In the case of subtelomere 22q, the enzyme cocktail generates a relatively long TRF, and the additional digestion with HinfI, whose site is present in close proximity upstream to the amplified region, results in the release of a shorter TRF (Supplementary Fig. 7).

Analysis of subtelomere 22q revealed that, following HinfI digestion, the hybrid enrichment levels were higher, indicating that the upstream regions, now disconnected from the shorter TRF, are not responsible for the generation of the DNA:RNA hybrids in these regions. In WT samples, we noticed an average decrease in hybrid enrichment after HinfI digestion in the same subtelomeres that were affected in ICF; however, in contrast to ICF samples, this decrease was not statistically significant (Fig. 3d). The subtelomeres lacking a distal HinfI site failed to demonstrate consistent differences in hybrid enrichment between the HinfI–digested and non-digested samples (Fig. 3e). Altogether, these results indicate that, while the hybrids are predicted to initiate at subtelomeres in the vicinity of TSSs, they extend into telomere-hexameric repeats, which constitute a crucial component of the hybrids at chromosome ends.

High TERRA levels are present in ICF cells during S phase. A mechanism by which TERRA could form DNA:RNA hybrids concomitantly with telomere replication, leading to fork stalling and telomere loss events, has been suggested previously for high TERRA-expressing yeast cells, as well as for ICF cells. Such a mechanism would require TERRA to be present at high levels during S phase. Measurement of TERRA levels at various cell-cycle stages in the telomerase-positive HeLa, HT1080 and SJSA1 cell lines demonstrated a decline in TERRA levels in late S and G2. On the other hand, in ALT-positive cells, in which TERRA levels are much higher, no decline from S to G2 phases was found. As shown previously and here (Supplementary Fig. 4), TERRA levels are markedly elevated in cycling ICF cell populations; however, TERRA levels at specific cell-cycle stages in ICF cells had not been determined previously. To this end, avoiding drug-induced cell-cycle synchronization, we used live cell DNA staining (Vybrant DyeCycle Violet stain) and sorted
three ICF and two WT LCLs into G1, S and G2 phases and examined fraction purity based on propidium iodide staining (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner). Following cell sorting, RNA was extracted from each cell-cycle fraction, and cDNA was generated (Fig. 4, upper left corner).

Levels of TERRA were assessed in four ICF LCLs (pCor, pG, pY, pH) and four WT LCLs (GM08729, GM19116c, FY, 3125) were digested with the enzyme cocktail and then each sample was split into two fractions. One of these fractions was further digested with HinfI, and DRIP was performed on both fractions. The enrichment values with HinfI digestion as 1. Each line connects between the enrichment values of the same sample with or without additional HinfI digestion. (b,d) Analysis of subtelomeres that contain a HinfI restriction site in close proximity to the telomere as described in panel a. (c,e) Analyses of subtelomeres that lack a HinfI restriction site and serve as controls. (b,c) Analyses of various subtelomeres, with and without HinfI sites, in four ICF samples. (d.e) Analysis of various subtelomeres, with and without HinfI sites, in four WT samples. *** = P values < 0.001, two-tailed Student’s t-test.
obtained from G1 fractions were set at 1, and enrichment values obtained from the S and G2 fractions were compared relative to the G1 values. In agreement with the relatively high TERRA levels detected at all cell-cycle stages in ICF LCLs (Fig. 4), DNA:RNA hybrids were detected throughout the cell cycle. The enrichment in S phase was significantly higher in comparison to G1 (P value < 0.05, Wilcoxon signed-rank test) for the analysed subtelomeres in samples ICF1 and ICF2, however, not so for ICF3. In WT cells, no consistent pattern of cell-cycle stage-specific enrichment was apparent. Altogether, these findings demonstrate that telomeric DNA:RNA hybrids are present both in ICF and WT cells throughout the cell cycle with no distinct and reproducible pattern in all the examined cell lines and suggest that, when present in S-phase, these DNA:RNA hybrids could affect telomere replication.

**Telomeres in ICF cells display RNase H1-sensitive DNA damage.**

Several observations suggest that DNA double-strand breaks arise from a collision between the replication machinery and an R-loop (reviewed in refs 49 and 50). Double-strand breaks resulting from such encounters at telomeric regions could lead to telomere loss. To further explore this possibility in ICF cells, we tested whether ICF cells display DNA damage at telomeric regions. LCLs grow in suspension and are not readily amenable for 3D-interphase analysis of telomere-dysfunction-induced foci. We therefore proceeded to examine the occurrence of DNA damage signals (DDSs) at chromosome ends in mitotic cells in ICF and WT LCLs. To enrich for cells in metaphase, the LCLs were briefly treated with colcemid, then cytospun on slides and subjected to immunofluorescence with an antibody against γ-H2AX, a marker for DNA damage. DDSs were detectable both along the chromosome arms and at the extreme chromosome ends, with signals emanating from either one or from both sister chromatids (Fig. 6a,b). Scoring the percentage of γ-H2AX-positive chromosome ends (either from one or two chromatids) in three ICF LCLs and three WT LCLs indicated that ICF LCLs exhibited significantly higher DNA damage signals at these loci (Fig. 6c and Supplementary Table 2). In contrast, no significant difference in

---

**Figure 4 | TERRA is abundant in ICF LCLs during all stages of the cell cycle.** Relative TERRA levels were determined in various LCLs sorted to G1, S and G2/M phases. FACS analysis of cell-cycle distribution prior and after sorting was performed to validate the purity of the sorted cell populations. The upper left boxed region displays a typical example of the cell-cycle distribution of an LCL population prior to sorting (depicted as ‘Total’) stained with vibrant dye cycle violet, and propidium iodide staining of the enriched populations of G1, S and G2/M after sorting. qRT-PCR analysis of TERRA in the sorted cell-cycle fractions was carried out for five telomeres (9p, 10q, 15q, 16p and 22q) in three ICF LCLs (pCor, pG and pY designated as ICF1, 2 and 3, respectively) and two WT LCLs (GM08729 and GM19116c designated WT1 and 2, respectively). For each subtelomere, TERRA expression in the G1 sample of WT1 was set at 1 and all other expression levels are described relatively to this sample. Bars and error bars represent means and s.e.m. of two experimental repeats. To be noted—the y axis is presented in a logarithmic scale.
Discussion

DNA:RNA hybrids at telomeric regions have been demonstrated previously in yeast and in human telomerase-positive and ALT cancer cells14,28,30. The data presented here demonstrate for the first time that DNA:RNA hybrids form at chromosome ends also in primary WT fibroblasts and in LCL lines derived from normal individuals. Using the whole-genome DRIP-seq approach, these hybrids are detected in the majority, but not in all, of chromosome ends in human fibroblasts and in the NT2 human embryonal carcinoma cell line (Fig. 1b,c). Our analysis focused on the last 2 kb of subtelomeric regions, where TERRA promoters and TSSs have either been described7 or are predicted to reside40. The presence of such hybrids in normal cells, albeit in low levels, suggests that the low TERRA levels present in normal cells are sufficient to elicit hybrid formation.

Although the human telomeric (TTAGGG)$_n$ repeat was shown by in vitro studies to form hybrids (ref. 30 and Supplementary Fig. 1), no direct in vivo evidence for the involvement of these sequences in hybrid generation was available. The protocol utilized in the previous study30, as well as in our study, cannot discriminate whether the subtelomeric regions detected following the DRIP procedure were pulled down by the S9.6 antibody based on hybrids present in telomeric regions or in subtelomeric regions, or both. Here we demonstrate clearly that the enrichment for hybrids is significantly reduced when subtelomeric regions are subjected to DRIP following the removal of the adjacent telomeric repeats (Fig. 3). Hence, the perfectly GC-skewed telomeric repeat, which comprises the majority of the TERRA molecules, contributes greatly to the hybrids formed at chromosome ends. The stability of hybrids forming at such regions could be further enhanced by the propensity of the displaced G-rich telomeric DNA to form G-quadruplexes51.

All mammalian chromosomes end with the canonical TTAGGG repeat, and human subtelomeres share many common characteristics, such as a high GC content, enrichment for the dinucleotide CpG, telomere-like repeats and families of larger repeats7,8. Despite these similarities, subtelomeric sequences vary noticeably10. Based on several sequence characteristics of the distal chromosome ends, we predicted that the capability to generate DNA:RNA hybrids would differ among various subtelomeres. In particular, this expectation was supported by the occurrence of GC-skewed regions, which were shown to be highly prone to hybrid formation22, and which are present in close vicinity to the putative TERRA promoters (Fig. 1). Indeed, our findings from both whole-genome DRIP-seq data and from probing individual subtelomeres suggest that subtelomeres differ in their capacity to form hybrids (Figs 1b,c and 2). If indeed this is the case, the subtelomeric regions that comprise the 5'-end of the TERRA molecules may be the regions that ‘seed’ the formation of hybrids, which then could extend further into the telomere-repeat region. The difference between human telomeres in their capacity to instigate formation of DNA:RNA hybrids suggests that the integrity of specific chromosome ends may be more at risk than others. Further studies will determine whether and how a specific subtelomeric sequence can influence the function of the telomere.

Figure 5 | Telomeric DNA:RNA hybrids form during all stages of the cell cycle in ICF LCLs. Three ICF LCLs (pCor, pG and pY, designated as ICF1, 2 and 3, respectively) and two WT LCLs (GM08729 and GM19116c, designated WT1 and 2, respectively) were sorted to G1, S and G2 phases, after which DRIP was performed and probed for the relative enrichment of six subtelomeric regions at the various cell-cycle phases. For each subtelomere, the enrichment in G1 was set at one, and the enrichment in S and G2 was compared relatively with the G1 values. Bars and error bars represent means and s.e.m. of two experimental repeats.

The percentage of the DDS positioned along the chromosome arms was evident between the ICF and WT groups (Supplementary Fig. 10).

We next explored whether the chromosome-end DDSs are related to the observed higher levels of DNA:RNA hybrids at telomeric regions in ICF cells. To address this question, we ectopically expressed RNase H1 in two ICF and two WT LCLs. Lentiviral constructs containing RNase H1-GFP (green fluorescent protein) and a control cytoplasmic-GFP were used to infect the LCLs, and GFP-positive cells were separated by a fluorescence-activated cell sorter (FACS) and propagated in culture. We found a eightfold increase in those expressing RNase H1-GFP in comparison to the control-GFP cells. In addition, we verified that hybrids at a non-telomeric and telomeric regions were reduced in the RNase H1-treated cells in comparison to the control-only GFP-expressing cells (Supplementary Fig. 11). Approximately a week after the infection, the cells were subjected to immunofluorescence on metaphase chromosomes to detect γ-H2AX, as described above. In WT LCLs, which display lower basal levels of chromosome-end DDS, the expression of RNase H1 led to increased levels of chromosome-end DDS in one WT LCL, whereas the other displayed a slight reduction. However, strikingly, in both ICF LCLs, overexpression of RNase H1 resulted in a significant reduction in γ-H2AX-positive chromosome ends (Fig. 6d and Supplementary Table 3), down to the levels observed in WT cells. These findings support the notion that, in ICF cells, DNA:RNA hybrids lead to DNA damage at telomeric regions, which, if not repaired, may lead to telomere loss.

Discussion

DNA:RNA hybrids at telomeric regions have been demonstrated previously in yeast and in human telomerase-positive and ALT cancer cells14,28,30. The data presented here demonstrate for the first time that DNA:RNA hybrids form at chromosome ends also in primary WT fibroblasts and in LCL lines derived from normal individuals. Using the whole-genome DRIP-seq approach, these hybrids are detected in the majority, but not in all, of chromosome ends in human fibroblasts and in the NT2 human embryonal carcinoma cell line (Fig. 1b,c). Our analysis focused on the last 2 kb of subtelomeric regions, where TERRA promoters and TSSs have either been described7 or are predicted to reside40. The presence of such hybrids in normal cells, albeit in low levels, suggests that the low TERRA levels present in normal cells are sufficient to elicit hybrid formation. Although the human telomeric (TTAGGG)$_n$ repeat was shown by in vitro studies to form hybrids (ref. 30 and Supplementary Fig. 1), no direct in vivo evidence for the involvement of these sequences in hybrid generation was available. The protocol utilized in the previous study30, as well as in our study, cannot discriminate whether the subtelomeric regions detected following the DRIP procedure were pulled down by the S9.6 antibody based on hybrids present in telomeric regions or in subtelomeric regions, or both. Here we demonstrate clearly that the enrichment for hybrids is significantly reduced when subtelomeric regions are subjected to DRIP following the removal of the adjacent telomeric repeats (Fig. 3). Hence, the perfectly GC-skewed telomeric repeat, which comprises the majority of the TERRA molecules, contributes greatly to the hybrids formed at chromosome ends. The stability of hybrids forming at such regions could be further enhanced by the propensity of the displaced G-rich telomeric DNA to form G-quadruplexes51.

All mammalian chromosomes end with the canonical TTAGGG repeat, and human subtelomeres share many common characteristics, such as a high GC content, enrichment for the dinucleotide CpG, telomere-like repeats and families of larger repeats7,8. Despite these similarities, subtelomeric sequences vary noticeably10. Based on several sequence characteristics of the distal chromosome ends, we predicted that the capability to generate DNA:RNA hybrids would differ among various subtelomeres. In particular, this expectation was supported by the occurrence of GC-skewed regions, which were shown to be highly prone to hybrid formation22, and which are present in close vicinity to the putative TERRA promoters (Fig. 1). Indeed, our findings from both whole-genome DRIP-seq data and from probing individual subtelomeres suggest that subtelomeres differ in their capacity to form hybrids (Figs 1b,c and 2). If indeed this is the case, the subtelomeric regions that comprise the 5'-end of the TERRA molecules may be the regions that ‘seed’ the formation of hybrids, which then could extend further into the telomere-repeat region. The difference between human telomeres in their capacity to instigate formation of DNA:RNA hybrids suggests that the integrity of specific chromosome ends may be more at risk than others. Further studies will determine whether and how a specific subtelomeric sequence can influence the function of the telomere.
whether in physiological aging or in other situations where TERRA is dysregulated, such as in ICF syndrome, human telomeres differ in their propensity to shorten. ICF type I cells are severely hypomethylated at subtelomeric regions, display a short telomere phenotype and express TERRA at extremely high levels. Previous studies have demonstrated that TERRA is upregulated at short telomeres and the scenario that the short telomere length in ICF syndrome influences TERRA levels cannot be ruled out. However, the high TERRA levels are an inherent characteristic of these cells independent of telomere length, as high TERRA persists in ICF cells also when telomeres are substantially elongated either by ectopic expression of hTERT or by reprogramming to iPSCs. As TERRA is the main candidate RNA for forming DNA:RNA hybrids at subtelomere/telomeric regions, we studied whether such hybrids will form preferentially in ICF cells. However, the involvement of additional telomeric transcript species in telomeric hybrids cannot be excluded. Indeed, formation of DNA:RNA hybrids at human telomeric regions is enhanced in ICF LCLs, and clearly, the group of subtelomeres that differs most substantially in their TERRA levels between ICF and WT LCLs are those that also display significantly elevated hybrid formation in ICF cells. We cannot preclude, however, that additional characteristics of the subtelomeric and telomeric region such as the hypomethylation or aberrant chromatin modifications contribute to hybrid formation. Accelerated telomere shortening in ICF syndrome could be a direct consequence of the subtelomeric/telomeric DNA:RNA hybrids, as unscheduled collisions between replication and transcription machinery have been shown to evoke double-strand breaks via DNA:RNA hybrid formation. We postulate that, in the scenario of abnormally elevated TERRA levels, particularly during the phase at which telomeres are replicated, RNase H and additional factors that normally handle the levels of DNA:RNA hybrids would fail to deal with the hybrid overload. This would result in the accumulation of telomeric hybrids, interference with the telomere replication process and generation of DNA damage. To this end, we demonstrate that ICF LCLs display significantly higher levels of TERRA at several

![Figure 6](https://example.com/figure6.png)

**Figure 6** | Chromosome end DNA damage signals in ICF LCLs are RNase H1 sensitive. (a–c) Cytospun metaphase spreads of three ICF LCLs pCor, pG and pY (designated ICF 1, 2 and 3, respectively) and three WT LCLs GM08729, GM19116c and 3125 (designated WT 1, 2 and 3, respectively) were stained with an antibody for γ-H2AX and were then scored for the percentage of chromosome ends displaying γ-H2AX signals (either at one or both sister chromatids). (a,b) Representative stained metaphases of ICF1 and ICF3. Red arrows point to a single chromatid-stained chromosome end, white arrows point to a double-chromatid stained chromosome end and the yellow arrow points to a γ-H2AX signal positioned along the chromosome arm. The chromosome boxed by the dotted line is enlarged in the upper left corner. Bar equals 10 μM. (c) All three ICF samples display a significantly higher percentage of γ-H2AX signals at their chromosome ends in comparison to each of the WT samples (P value < 0.001, Proportions test. All P values appear in Supplementary Table 2). At least 400 chromosomes were scored for each sample. Bars and error bars represent percentages and s.e.m.

(d) RNase H1-GFP or cytoplasmic GFP were expressed in ICF1, ICF3, WT1 and WT2 LCLs, sorted based on GFP expression and analysed for chromosome end γ-H2AX signals. A minimum of 180 chromosomes was scored for each sample. ‘RNaseH’ refers to cells expressing the fused RNase H1-GFP. ‘Control’ refers to cells expressing the cytoplasmic-GFP. The levels of γ-H2AX signals at chromosome ends were significantly decreased in the ICF samples expressing RNase H1-GFP (P value < 0.05 Proportions test. All P values appear in Supplementary Table 3). Bars and error bars represent percentages and s.e.m.
An RNAse H1-expressing plasmid was prepared by amplification of a fragment containing the human RNAse H1 starting at M27, fused to enhanced GFP (eGFP), from the pEGFP-M27-H1 plasmid (a gift from Robert Crouch)58. The primers used for amplification were: Forward—5'-CTCTAGATCCGAGCTCAACG-3', containing a BglII restriction site (underlined) and Reverse—5'-AGCTGACATTACTGTGACGGC-3', containing a SalI restriction site (underlined) and a 1.58 kb amplified fragment was cloned into pcR2.1 (Invitrogen). Sanger sequencing confirmed the sequence of a pcR2.1-RNase H1-eGFP-positive clone. The pcTKH1-eGFP plasmid was generated by subcloning the BglII-Sall fragment from the pcR2.1-RNase H1-eGFP plasmid into the BamHI and XhoI sites of the lentiviral vector pTK208, containing the cytomegalovirus promoter. Plasmid pTK113 contains the eGFP gene cloned into the pTK208 plasmid. Both pTK208 and pTK113 are gifts from Tal Kafri.

Calculation of CpG density and GC content and GC skew. CpG density, GC content and GC skew of 2 kb subtelomeric regions adjacent to the telomere tract were calculated using standard equations with sliding window method59. Heatmaps and clusters were created using heatmap.3 function from R package ‘GMD’. Analysis of these regions was carried out on the GenoMatix software suite (PromoTheInspector program for identification of putative Pol II promoter regions60, TSSs were predicted based on Nergadze et al.61.

DRIIP analyses. The S9.6 antibody was produced from the HB-8730 cell line either by recovery from ascites fluid and purification to homogeneity by Antibodies Inc. (Davis, CA, USA)62 or as follows: the HB-8730 cells were transferred to growth in serum-deprived medium (0.5% FCS in RPMI) for 48 h. Ammonium sulfate was added to the medium to 29.1 g per 100 ml, and the samples were stirred for 3 h at 4 °C. Then the mixture was centrifuged for 30 min at 34,155 g for 20 min, resuspended in a few millilitres of 10% glycerol/1× PBS and subjected to overnight dialysis in the same buffer. The concentration and efficiency of the antibody prepared in this way were validated by comparison to the commercial S9.6 antibody (ENH001, KeraFAST).

DRIIP was carried out based on the detailed here: Five million cells were washed in PBS, centrifuged, and resuspended without pipetting in 1.6 ml TE with 83 µl of 10% SDS and 5 µl of 20 mg per ml protease K and incubated overnight at 37 °C. The next morning, DNA was extracted with phenol/chloroform/isoamylalcohol (25:24:1) (Affymetrix, no. 70853) using phase lock tubes (Spreme, no. 2302840). After partial resuspension in TE, DNA was fragmented with XhoI, Spatient, HindIII, EcoRI and BglII in 2.1 ml digestion buffer (NEB) containing 1 mg ml−1 2 mM spermidine. Twenty five units of RNase H (NEB, no. M0297) were added to half of each sample to serve as a negative control, and the samples were incubated overnight at 37 °C. Following extraction in phenol/chloroform/isoamylalcohol (25:24:1), 2 ml phase lock tubes, DNA concentration was determined using a NanoDrop ND-1000, and an aliquot was separated by agarose gel electrophoresis to validate efficient digestion. After setting aside 1% for input control, 4 µg of DNA were used for immunoprecipitation overnight with binding buffer (10 mM NaPO4, pH 7.0/0.14 M NaCl/0.05% Triton X-100) and 10 µg of the S9.6 antibody on a rotisserie shaker at 4 °C. The following morning, 50 µl of agarose A/G beads (Pierce, no. 26804) were precooled 2 °C with binding buffer with bound DNA/antibody complex and incubated for 2 h at 4 °C on a rotisserie shaker after three washes with binding buffer at room temperature for 10 min each, the beads were eluted with 250 µl elution buffer (50 mM Tris pH 8.0, 10 mM EDTA, 0.5 M NaCl, 0.1% SDS) and 7 µl proteinase K (20 mg ml−1) and incubated at 55 °C for 2 h. Finally, the eluted DNA was extracted with phenol/chloroform/isoamylalcohol (25:24:1), precipitated with ethanol and resuspended in 50 µl of TE for analysis by qPCR.

When DRIIP experiments were performed with an additional HinfI digestion, DNA was first digested overnight with the standard enzyme cocktail and then split into two equal aliquots. HinfI was added to one of the aliquots, and both aliquots were incubated at 37 °C for 3 h and then processed as described above. Primers used for qPCR reactions were designed using the Primer Express Software (Thermo Fisher Scientific) in regions as close as possible to the telomere tract, based on the reported subtelomeric sequences63. Owing to the repetitive nature of the subtelomeric regions, some of the primer pairs amplify most to the DNA/antibody complex and incubated for 2 h at 4 °C on a rotisserie shaker. After three washes with binding buffer at room temperature for 10 min each, the beads were eluted with 250 µl elution buffer (50 mM Tris pH 8.0, 10 mM EDTA, 0.5 M NaCl, 0.1% SDS) and 7 µl proteinase K (20 mg ml−1) and incubated at 55 °C for 2 h. Finally, the eluted DNA was extracted with phenol/chloroform/isoamylalcohol (25:24:1), precipitated with ethanol and resuspended in 50 µl of TE for analysis by qPCR.

DRIIP-seq Mapping. Input, untreated and bacterial RNAse H-treated DRIIP-seq data sets from WT human primary fibroblasts64 and the human testicular carcinoma cell line, Ntera2 (RNA), were quality/adapter trimmed using fastx-mq wymapped and mapped using Bowtie2 v2.1.0 (ref. 67) to human genome version GRCh37/hg19 supplemented by DNA sequences representing missing subtelomeric regions, when necessary. PCR duplicates were removed using SAMtools65. Peaks were called using MACS v2.1 (ref. 69) from input or RNAse H-treated samples as controls. For validation of putative Pol II promoter regions, the most distal 2 kb of subtelomeric region of each chromosome were aggregated from peaks in that region.

Methods

Cell lines and cell culture. ICF LCLs pCor, pG, pH, pT and pY30 (designated as ICF1, 2, 4, 5 and 3, respectively) were grown in RPMI supplemented with 20% FCS, 2 mM Glutamine, 100 U ml−1 penicillin and 100 mg ml−1 streptomycin. Control WT LCLs included the following: 3125 and 3133 (ref. 62), GM08729, GM08728 and GM08731, respectively. ALL cells30. ICF telomeres, contrary to ALT cancer cells, do not evolve as telomere hybrids. Nevertheless, ICF syndrome belongs to the emerging group of human diseases in which dysregulation of a long non-coding RNA is strongly associated with an abnormal phenotype. Cancer cells that employ the ALT pathway to maintain telomeres are also hypomethylated at subtelomeric regions and transcribe high levels of cell-cycle dysregulated TERRA66–69. In ALT cells, the high TERRA levels are suggested to contribute to the elevated levels of hybrids, and accumulation of telomeric hybrids could lead to the homologous recombination typical to these cells30. ICF telomeres, contrary to ALT cancer cells, do not engage in recombination33. These different outcomes of high TERRA and telomeric hybrids may relate to the fact that ICF cells are untransformed and normally suppress telomere recombination, in contrast to ALT cells that carry mutations that unleash telomere recombination64,65,60.

Regardless of the outcome, abnormally high levels of TERRA and the resulting enrichment of hybrids at telomeric regions are likely to endanger the integrity of human telomeres. Future studies that manipulate TERRA levels in various cell types, including ICF cells, will further establish the causal role of TERRA and the mechanisms by which telomeric DNA/RNA hybrids impede the normal maintenance of human telomeres.

Plasmids. The telomere-repeat containing plasmids were prepared as following: A 240 bp PstI fragment containing 40 telomere repeats was released from the pHuR93 (ATCC 61076) plasmid58 and cloned into the pBlueScript vector, generating pBS-240bpT. An 810 bp HindIII-KpnI fragment containing 135 telomere repeats was released from pXneo 270(T2A)G3 (Addgene plasmid no. 12403; a gift from Titia de Lange41) and cloned into the pFC53 plasmid generating pFC53-800bpT pl. A pFC53 plasmid containing mAIR1 served as a positive control for DNA:RNA hybrids (pFC53-mAIR122).

NATURE COMMUNICATIONS | DOI: 10.1038/ncomms14015 | www.nature.com/naturecommunications
Cell sorting. Cell sorting by DNA content was carried out as following: One
microlitre of Vybrant DyeCycle Violet Stain (V35003, Life Technologies) was
added to the cells and incubated for 15 min. The fixed cells were then
collected by sorting. The cells were then stained with Propidium Iodide
for 15 min and analysed on a FACSCalibur to validate the purity of the sorted
100,000 cells from each fraction were fixed in EtOH, stained with Propidium Iodide
and analysed by flow cytometry. Sorting of GFP-positive LCLs was
carried out under aseptic conditions on the FACS Aria IIIu Cell Sorter using
the 130 nm nozzle.

RNA extraction and qRT-PCR. ICF LCLs, whose growth is considerably atten-
uated, required sorting of approximately 40 million cells in order to obtain a
sufficient number of cells from S and G2 phases for expression analysis. Sorted cells
intended for RNA extraction were collected into 5 ml tubes containing 1 ml
RNA Later solution (AM7020, Thermo Fisher Scientific). When the total volume
in the tubes reached 4.5 ml, cells were centrifuged at 840g and resuspended in 1 ml
RNA Later. RNA was extracted using the RNeasy Mini Kit (no. 74004, QIAGEN)
and an additional DNase treatment with TURBO DNase (AM2238, Thermo Fisher
Scientific) was carried out in order to eliminate any trace of DNA in the pre-
paration. RNA concentrations were determined by Qubit fluorometric quantita-
tion. Total RNA was reverse transcribed at 55°C. The cDNA was then transcribed at 37°C with SuperScript III reverse transcriptase (Invitrogen) using a β-actin-specific primer (5′-AGTCCG
CCTAGAAGCCATTG-3′) and a TERRA-specific primer composed from five
telomere-hexamer repeat (CCCTAA)₅. cDNAs were then analysed with the same primers used for DRIP (Supplementary Table 4). qRT-PCR was carried out on
an Applied Biosystems StepOnePlus Real-Time PCR system with Fast SYBR
Green Master Mix (AB-4385612, Applied Biosystems). Analysis was carried out by the delta delta Ct method using the β-actin gene as the reference.

Lentiviral infections. Each LCL was infected serially twice with virions produced
from either pTK-H1-eGFP or from pTK131 plasmids in the HEK293FT packaging
cell line. Virus-containing supernatant was collected 48–72 h after transfection and
polybrene was added to a final concentration of 8 µg/ml. Ten millilitres of
virus-containing media were added to 1 × 10⁵ cells in a 15 ml tube, and the cells were
spun for 90 min at 450g and resuspended at 1 × 10⁶ cells/ml. Lentiviral infections
were carried out using the 130 nm nozzle and stained, the cells were incubated with 20 ng/ml
1% Triton in PBS (PBST) at room temperature and incubated with a secondary
anti-mouse antibody (no. 7074, Cell Signaling Technology) and visualized
using a BX51 microscope (Olympus), and captured with an Olympus DP73 camera controlled by the cellSens
Entry Microscopy Imaging software (Olympus).

Statistical analysis. For each statistical analysis, we examined normality by
Shapiro–Wilk Normality test and examined homoscedasticity between groups by
Levene’s test. When the data did not meet the assumptions of the parametric tests,
we calculated the statistics by non-parametric tests as Wilcoxon signed-rank test
for paired samples or the Wilcoxon signed rank-sum test (Mann–Whitney U-test) for
non-paired samples.

Data availability. DRIP-seq data sets were published in Ginno et al.41 and Lim et al.42 The data sets are available in the NCBI Gene Expression Omnibus (GEO;
http://www.ncbi.nlm.nih.gov/geo) under accession number GSE45530 and
GSE57353, respectively.

References
1. Hug, N. & Lingner, J. Telomere length homeostasis. Chromosoma 115, 413–425
(2006).
2. Giraud-Panis, M. J. et al. One identity or more for telomeres? Front. Oncol. 3,
48 (2013).
3. Palm, W. & de Lange, T. How shelterin protects mammalian telomeres. Annu.
Rev. Genet. 42, 301–334 (2008).
4. Allsopp, R. C. & Harley, C. B. Evidence for a critical telomere length in
senescent human fibroblasts. Exp. Cell Res. 219, 130–136 (1995).
5. Hermany, M. T., Strong, M. A., Hao, L. Y. & Greider, C. W. The shortest
telomere, not average telomere length, is critical for cell viability and
chromosome stability. Cell 107, 67–77 (2001).
6. Zou, Y., Steir, A., Gryznow, S. M., Shay, J. W. & Wright, W. E. Does a sentinel
or a subset of short telomeres determine replicative senescence? Mol. Biol. Cell
15, 3709–3718 (2004).
7. Nergadze, S. G. et al. CpG island promoters drive transcription of human
telomeres. RNA 15, 2186–2194 (2009).
8. Ambrosini, A., Paul, S., Hu, S. & Riethman, H. Human subtelomeric duplicon
structure and organization. Genome Biol. 8, R151–R151.113 (2007).
9. Galati, A., Micheli, E. & Cacchione, S. Chromatin structure in subtelomeric
dynamics. Front. Oncol. 3, 46 (2013).
10. Stong, N. et al. Subtelomeric CTCF and cohesion binding site organization
improved subtelomeric assemblies and a novel annotation pipeline. Genome Res.
24, 1039–1050 (2014).
11. Azzalin, C. M., Reichenbach, P., Khorliati, L., Giulotto, E. & Lingner, J.
Telomeric repeat containing RNA and RNA surveillance factors at mammalian
telomeres. Sci. Rep. 3, 316–325 (2013).
12. Schoeftner, S. & Blasco, M. A. Developmentally regulated transcription of
mammalian telomeres by DNA-dependent RNA polymerase II. Nat. Cell Biol.
10, 228–236 (2008).
13. Azzalin, C. M. & Lingner, J. Telomere functions grounding on TERRA firma.
Trends Cell Biol. 25, 39–46 (2015).
14. Balk, B. et al. Telomeric RNA-DNA hybrids affect telomere-length dynamics
and senescence. Nat. Struct. Mol. Biol. 20, 1199–12052 (2013).
15. Cusanelii, E., Romero, C. A. & Charrand, P. Telomeric noncoding RNA
tERRA is induced by telomere shortening to nucleate telomerase molecules at
short telomeres. Mol. Cell 51, 780–791 (2013).
16. Maicher, A., Kastner, I. & Luke, B. Telomeres and disease: enter TERRA. RNA
9, 843–849 (2012).
17. Pfeiffer, V. & Lingner, J. TERRA promotes telomere shortening through
exonuclease 1-mediated resection of chromosomal ends. PLoS Genet. 8, e1002747
(2012).
18. Redon, S., Zemp, I. & Lingner, J. A three-state model for the regulation of
clonal TERRA expression by TERRA and TERRA-binding proteins. Mol. Biol. Cell
13, 5191–5201 (2002).
19. Vitelli, V. et al. More on the lack of correlation between TERRA expression
and telomere length. Front. Oncol. 3, 245 (2013).
20. Rippe, K. & Luke, B. TERRA and the state of the telomere. Nat. Struct. Mol.
Biol. 22, 853–858 (2015).
21. Santos-Pereira, J. M. & Aguilera, A. R loops: new modulators of genome
dynamics and function. Nat. Rev. Genet. 16, 583–597 (2015).
22. Ginno, P. A., Lott, P. L., Christensen, H. C., Korf, I. & Chedin, F. R-loop
formation is a distinctive characteristic of unmethylated human CpG island
promoters. Mol. Cell 45, 814–825 (2012).
23. Hartono, S. R., Korf, I. F. & Chedin, F. GC skew is a conserved property of
unmethylated CpG island promoters across vertebrates. Nucleic Acids Res. 43,
9729–9741 (2015).
24. Lombrana, R., Almeida, R., Alvarez, A. & Gomez, M. R-loops and initiation of
DNA replication fork progression. Nat. Immunol. 4, 442–451 (2003).
25. Yu, K., Chedin, F., Hsieh, C. L., Wilson, T. E. & Lieber, M. R. R-loops at
telomere, not average telomere length, is critical for cell viability and
chromosome stability. Exp. Cell Res. 219, 215–224 (2006).
26. Yu, T. Y., Kao, Y. W. & Lin, J. J. Telomeric transcripts stimulate telomere
senescence and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. Nature 402, 187–191 (1999).
32. Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99, 247–257 (1999).
33. Yehezkeli, S., Segov, Y., Viegas-Pequignot, E., Skorecki, K. & Selig, S. Hypomethylation of subtelomeric regions in ICF syndrome is associated with abnormally short telomeres and enhanced transcription from telomeric regions. Hum. Mol. Genet. 17, 2776–2789 (2008).
34. Deng, Z., Campbell, A. E. & Lieberman, P. M. TERRA, CpG methylation and telomere heterochromatin: lessons from ICF syndrome cells. Cell Cycle 9, 69–74 (2010).
35. Yehezkeli, S. et al. Characterization and rescue of telomeric abnormalities in ICF syndrome type I fibroblasts. Front. Oncol. 3, 35 (2013).
36. Luke, B. & Lingner, J. TERRA: telomeric repeat-containing RNA. EMBO J. 28, 2503–2510 (2009).
37. Sage, S. et al. Induced pluripotent stem cells as a model for telomeric abnormalities in ICF type I syndrome. Hum. Mol. Genet. 23, 3629–3640 (2014).
38. Porro, A. et al. Functional characterization of the TERRA transcriptome at damaged telomeres. Nat. Commun. 5, 3579 (2014).
39. Huertas, P. & Aguilera, A. Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. Mol. Cell 12, 711–721 (2003).
40. Diman, A. et al. Nuclear respiratory factor 1 and endurance exercise promote human telomere transcription. Sci. Adv. 2 (2016).
41. Ginno, P. A., Lim, Y. W., Lott, P. L., Korf, I. & Chedin, F. GC skew at the 5′ ends of human genes links R-loop formation to epigenetic regulation and transcription termination. Genome Res. 23, 1590–1600 (2013).
42. Lim, Y. W., Sanz, L. A., Xu, X., Hartono, S. R. & Chedin, F. Genome-wide DNA hypomethylation and RNA:DNA hybrid accumulation in Acicardi-Goutières syndrome. eLife 4, 80087 (2015).
43. Loomis, E. W., Sanz, L. A., Chedin, F. & Hagerman, P. J. Transcription-associated R-loop formation across the human FMR1 CGG-repeat region. PLoS Genet. 10, e1004294 (2014).
44. Maicher, A., Kastner, L., Dees, M. & Luke, B. Deregulated telomere transcription causes replication-dependent telomere shortening and promotes cellular senescence. Nucleic Acids Res. 40, 6649–6659 (2012).
45. Porro, A., Feuerhahn, S., Reichenbach, P. & Lingner, J. Molecular dissection of telomeric repeat-containing RNA biogenesis unveils the presence of distinct and multiple regulatory pathways. Mol. Cell 30, 4808–4817 (2010).
46. Flynn, R. L. et al. Alternative lengthening of telomeres renders cancer cells hypersensitive to ATR inhibitors. Science 347, 273–277 (2015).
47. Flynn, R. L. et al. TERRA and hnrRNPA1 orchestrate an RPA-to-POT1 switch on telomeric single-stranded DNA. Nature 471, 532–536 (2011).
48. Aronesty, E., Bah, A., Wischnewski, H., Shchepachev, V. & Azzolin, C. M. The telomeric transcriptome of Schizosaccharomyces pombe. Nucleic Acids Res. 40, 2995–3005 (2012).
49. Helmrich, A., Ballarino, M., Nuessler, E. & Tora, L. Transcription-replication encounters, consequences and genomic instability. Nat. Struct. Mol. Biol. 20, 412–418 (2013).
50. Cusanelli, E. & Charrand, P. Telomeric repeat-containing RNA TERRA: a noncoding RNA connecting telomere biology to genome integrity. Front. Genet. 6, 143 (2015).
51. Bah, A., Wischnewski, H., Shchepachev, V. & Azzalin, C. M. The telomeric transcriptome of Schizosaccharomyces pombe. Nucleic Acids Res. 42, 4391–4405 (2014).
52. Cusanelli, E. & Charrand, P. Telomeric repeat-containing RNA TERRA: a noncoding RNA connecting telomere biology to genome integrity. Front. Genet. 6, 143 (2015).
53. Lalevée, S. & Feil, R. Long noncoding RNAs in human disease: emerging mechanisms and therapeutic strategies. Epigenetics 7, 877–879 (2015).
54. Episkopou, H. et al. Alternative lengthening of telomeres is characterized by reduced compaction of telomeric chromatin. Nucleic Acids Res. 42, 4391–4405 (2014).
55. Vera, E., Canela, A., Fraga, M. F., Esteller, M. & Blasco, M. A. Epigenetic regulation of telomeres in human cancer. Oncogene 27, 6817–6833 (2008).
56. Tilman, G. et al. Subtelomeric DNA hypomethylation is not required for telomeric sister chromatid exchanges in ALT cells. Oncogene 28, 1682–1693 (2009).
57. Lee, M. E., Rha, S. Y., Jeung, H. C., Chung, H. C. & Oh, B. K. Subtelomeric DNA methylation and telomere length in human cancer cells. Cancer Lett. 281, 82–91 (2009).
58. Clynes, D. et al. Suppression of the alternative lengthening of telomere pathway by the chromatin remodelling factor ATRX. Nat. Commun. 6, 7538 (2015).
59. Velasco, G. et al. Germline genes hypomethylation and expression define a molecular signature in peripheral blood of ICF patients: implications for diagnosis and etiology. Orphanet. J. Rare Dis. 9, 56 (2014).
60. Ofrir, R., Wong, A. C., McDermid, H. E., Skorecki, K. L. & Selig, S. Position effect of human telomeric repeats on replication timing. Proc. Natl. Acad. Sci. USA 96, 11434–11439 (1999).
61. Moyzis, R. K. et al. A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. Proc. Natl. Acad. Sci. USA 85, 6622–6626 (1988).
62. Hanish, J. P., Yanowitz, J. L. & de Lange, T. Stringent sequence requirements for the formation of human telomeres. Nat. Cell Biol. 6, 8861–8865 (1994).
63. Cerritelli, S. M. et al. Failure to produce mitochondrial DNA results in embryonic lethality in Nasehi1 null mice. Mol. Cell 11, 807–813 (2003).
64. Arosney, E. Command-line tools for processing biological sequencing data. https://github.com/ExpressionAnalysis/ea-utils (2011).
65. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359 (2012).
66. Li, H. et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079 (2009).
67. Zheng, Y. et al. Model-based analysis of ChiP-Seq (MACS). Genome Biol. 9, R137 (2008).
68. Cesare, A. J. et al. Spontaneous occurrence of telomeric DNA damage response in the absence of chromosome fusions. Nat. Struct. Mol. Biol. 16, 1244–1251 (2009).

Acknowledgements
We thank Harold Reithman for sharing unpublished subtelomeric sequences and to Anabelle Decottignies for sharing the TERRA promoter and TSS prediction data with us and for helpful advice and comments. We thank Dale Frank and Miriam Gagliardi for helpful discussions and Daniel Kornitzer, Omer Schwartzman and Maty Tsukerman for comments on the manuscript. Thanks to the BCF unit members, Rap- paport Faculty of Medicine: Yaakov Salouky and Amir Gruz from the FACS unit, and Lit Linde and Shirmat Mamrut from the genomic center for help in technical issues. This research was supported by The Israel Science Foundation (grant no. 883/12, to S. Selig), The Legacy Heritage Bio-Medical Program of the Israel Science Foundation (grant no 657/15, to S. Selig) and The National Institutes of Health (GM094299 to F.C.). S. Sagie is grateful to the Arzieli Foundation for the award of an Arzieli Fellowship. S.R.H. is a Howard Hughes Medical Institute International Student Research fellow.

Author contributions
S. Selig, S. Sagie and F.C. conceived and designed the experiments, S. Sagie, S. Selig, S.T., H.K., A.T.-G. and S.H. performed the experiments, S. Sagie, S. Selig, F.C., S.R.H. and S.T. analysed and interpreted the data, S. Selig, S. Sagie and F.C. wrote the paper, S.T., S.R.H., C.F. and G.V. contributed to preparing the manuscript. C.F. and G.V. contributed reagents.

Additional information
Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://www.nature.com/reprintsandpermissions/

How to cite this article: Sagie, S. et al. Telomeres in ICF syndrome cells are vulnerable to DNA damage due to elevated DNA:RNA hybrids. Nat. Commun. 8:14015 (2017). doi: 10.1038/ncomms14015

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

© The Author(s) 2017