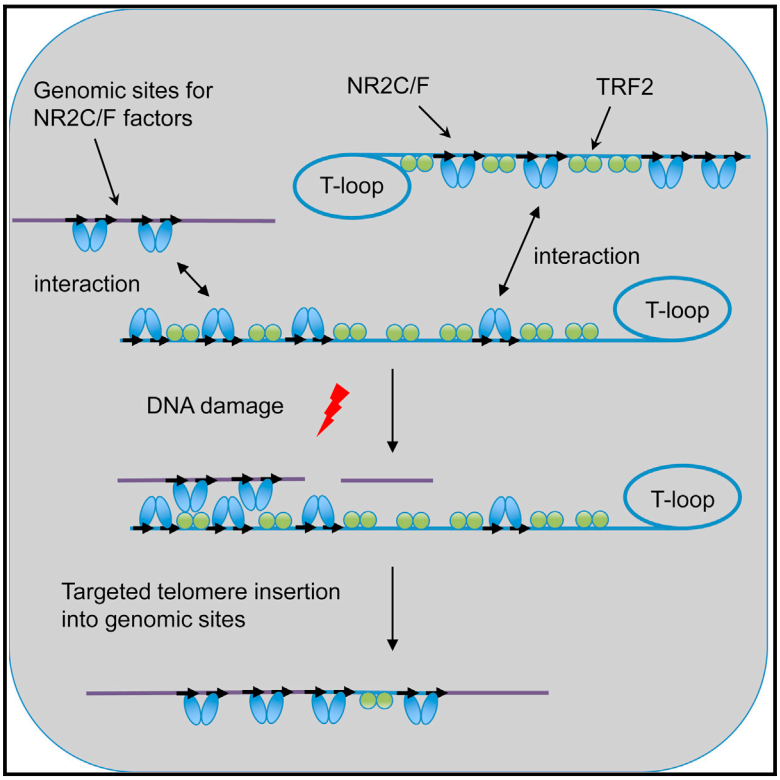
Article

Cell

Nuclear-Receptor-Mediated Telomere Insertion Leads to Genome Instability in ALT Cancers

Graphical Abstract



Authors

Paulina Marzec, Claudia Armenise, ..., Frédéric Chibon, Jérôme Déjardin

Correspondence

jerome.dejardin@igh.cnrs.fr

In Brief

GGGTCA variant repeats accumulate on ALT telomeres and lead to the aberrant recruitment of NR2C/F nuclear receptors. These receptors can then bridge to their native binding sites within chromatin and drive the inclusion of telomeric sequences within the genome, creating sites prone to breakage and translocations.

Highlights

- NR2C/F orphan nuclear receptors bind to ALT telomeres at GGGTCA direct repeats
- NR2C/F factors bridge their target loci, clustering and relocalizing ALT telomeres
- Telomeric sequences are inserted in ALT genomes in an NR2C/F-dependent manner
- Telomere insertion destabilizes ALT genomes and contributes to complex rearrangements

Marzec et al., 2015, Cell *160*, 913–927
February 26, 2015
http://dx.doi.org/10.1016/j.cell.2015.01.044

CellPress

Article

Cell

Nuclear-Receptor-Mediated Telomere Insertion Leads to Genome Instability in ALT Cancers

Paulina Marzec,[1] Claudia Armenise,[1] Gaëlle Pérot,[2,3] Fani-Marlen Roumelioti,[4] Eugenia Basyuk,[5] Sarantis Gagos,[4] Frédéric Chibon,[2,3] and Jérôme Déjardin[1,*]

[1]INSERM AVENIR Team, Institute of Human Genetics, CNRS UPR 1142, 141 rue de la Cardonille, 34396 Montpellier, France

[2]Translational Research, Department of Biopathology, Institut Bergonié, 229 cours de l'Argonne, 33076 Bordeaux, France

[3]INSERM U916, 229 cours de l'Argonne, 33076 Bordeaux, France

[4]Laboratory of Genetics and Gene Therapy, Center of Basic Research II, Biomedical Research Foundation of the Academy of Athens, 11527 Athens, Greece

[5]Institut de Génétique Moléculaire de Montpellier, CNRS–UMR5535, 1919 route de Mende, 34293 Montpellier Cedex 5, France

*Correspondence: jerome.dejardin@igh.cnrs.fr

http://dx.doi.org/10.1016/j.cell.2015.01.044

SUMMARY

The breakage-fusion-bridge cycle is a classical mechanism of telomere-driven genome instability in which dysfunctional telomeres are fused to other chromosomal extremities, creating dicentric chromosomes that eventually break at mitosis. Here, we uncover a distinct pathway of telomere-driven genome instability, specifically occurring in cells that maintain telomeres with the alternative lengthening of telomeres mechanism. We show that, in these cells, telomeric DNA is added to multiple discrete sites throughout the genome, corresponding to regions regulated by NR2C/F transcription factors. These proteins drive local telomere DNA addition by recruiting telomeric chromatin. This mechanism, which we name targeted telomere insertion (TTI), generates potential common fragile sites that destabilize the genome. We propose that TTI driven by NR2C/F proteins contributes to the formation of complex karyotypes in ALT tumors.

INTRODUCTION

Cancer is characterized by genomic alterations that lead to oncogene activation and/or tumor suppressor loss. These changes accumulate during tumor development and can be detected as translocations, amplifications, or deletions of chromosomal segments. Another major characteristic of cancer cells is their unlimited proliferative potential. This feature is dependent on the activation of a telomere maintenance mechanism upon exit from crisis (Hanahan and Weinberg, 2011). During cell crisis, genome instability and telomere dysfunction have been primarily linked through the classical mechanism of breakage-fusion-bridge cycle (BFB) first described by Barbara McClintock (McClintock, 1941). In this chain of events, unprotected or broken telomeres can fuse to another chromosomal extremity via non-homologous end joining. Fusions create di-centric chromosomes that eventually break at random positions during mitosis, generating deletions and amplifications of chromosomal segments and more unprotected chromosome ends (Murnane, 2012). This cycle persists until chromosomal extremities get stabilized by telomere addition via telomerase activation. In human, telomerase is activated in the majority of cancers. However, in a subset of tumors, mostly sarcomas, telomeres are maintained by a recombination/amplification mechanism termed alternative lengthening of telomeres (ALT) (Bryan et al., 1997). These tumors typically harbor highly heterogeneous and complex karyotypes (Taylor et al., 2011). The lack of apparent specific translocations makes it challenging to identify the mechanism driving tumorigenesis in these cancers. Efforts to characterize these tumors are thus currently limited to identifying specific gene expression signatures (Chibon et al., 2010). Similarly, the mechanism underlying ALT activation and maintenance in these tumors is unknown. Because tumors in which telomerase is inhibited can activate ALT in mouse models (Hu et al., 2012), it is critical to dissect this pathway to design efficient anti-cancer therapies targeting telomere maintenance. We previously showed that orphan nuclear receptors of the NR2C/F classes (TR2, TR4, COUP-TF1, COUP-TF2, and EAR2), which belong to the nuclear hormone receptor (NHR) family of transcription factors, are aberrantly associated with telomeres in a prototypic ALT(+) cell line (Déjardin and Kingston, 2009). This finding was unexpected, as transcription factors usually associate with gene regulatory regions, and telomeres do not contain classical genes. Here, we address the biological relevance of this finding. We identify a critical role for these proteins in the ALT process and in active destabilization of the genome. We dissect the mechanism leading to their aberrant recruitment to telomeres, and we show that these proteins have a major architectural role: NR2C/F proteins can bridge together bound loci in the nuclear space. By promoting spatial proximity, NR2C/F proteins favor the telomere-telomere recombination necessary for ALT maintenance. Surprisingly, NR2C/F-driven spatial proximity also induces the tethering of telomeric chromatin to hundreds of regular NR2C/F-binding sites throughout the genome. This abnormal organization triggers the insertion of telomeric material to these sites, and this process depends on NR2C/F proteins. Insertions of telomeric DNA throughout the genome lead to the creation of potential common fragile sites that are known to be prone to breakage

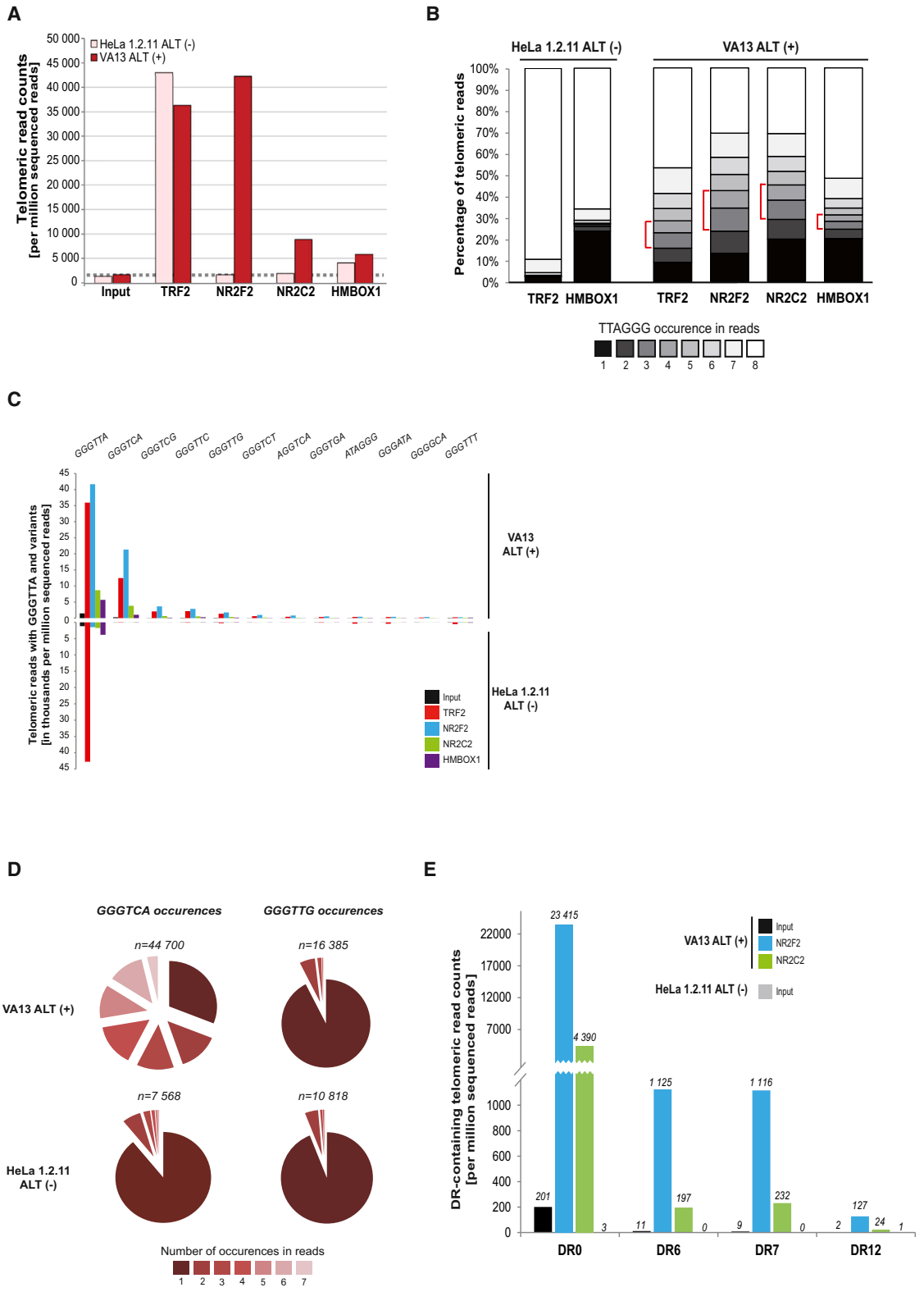


Figure 1. Analysis of the Binding Profile of NR2C/F Factors at Telomeres by ChIP Sequencing

(A) Normalized amount of telomeric reads in ALT(−) (pink) and ALT(+) (red) libraries prepared from input DNA, TRF2, NR2F/C2, and HMBOX1 IPs (dashed line displays normalized amount in input libraries).

(legend continued on next page)

and translocations. Because this mechanism of telomere-driven genome instability is fundamentally distinct from the BFB cycles and occurs as the consequence of the activation of a telomere maintenance mechanism, we name it targeted telomere insertions (TTI). In line with the role played by NR2C/F factors in TTI, we show that these proteins associate with telomeres in primary ALT tumors in situ and that this association correlates with the extent of karyotype rearrangements. Therefore, NR2C/F have a critical role in ALT and in the activation of TTI. We propose that this mechanism of telomere-driven genome instability induces heterogeneous genomes and contributes to the generation of complex karyotypes in ALT sarcomas (Taylor et al., 2011).

RESULTS

NR2C/F Factors Bind to Direct Repeats of a Variant Telomeric Motif

To get insights into the significance of NR2C/F binding to telomeres, we first examined how these factors are recruited. A mutant NR2C2 protein with point mutations disrupting DNA binding (Tanabe et al., 2007) fails to accumulate at telomeres (Figures S1A and S1B), suggesting that NR2C/F directly bind to DNA. NHR usually associate with DNA as dimers by binding to a composite sequence made of two half-sites (the 5′-A/GGGTCA-3′ motif). Depending on the mutual orientation and spacing of these half-motifs, the full binding site varies extensively (Sandelin and Wasserman, 2005). Since the NHR half-site is related to the canonical telomere basic repeat unit 5′-GGGTTA-3′, we hypothesized that NR2C/F could be recruited to ALT telomeres through binding to an iteration of the naturally occurring variant 5′-GGGTCA-3′ (Allshire et al., 1989).We thus analyzed by chromatin immunoprecipitation (ChIP) combined with high-throughput sequencing the DNA sequences associated with NR2C2 and NR2F2, those associated with the canonical telomere-binding protein TRF2 (de Lange, 2005), and with HMBOX1, another DNA-binding protein that we previously identified at telomeres regardless of the maintenance mechanism (Déjardin and Kingston, 2009). This analysis was performed in both ALT(+) and in ALT(−) cell lines (*WI-38 VA13 2RA* and *HeLa 1.2.11* cells, respectively) using validated antibodies (Figures S2A, S4A, and S7A). While TRF2 and HMBOX1 are enriched at both types of telomeres, NR2C2/F2 proteins bind only to ALT(+) telomeres (Figures 1A and S2A), consistent with our original findings (Déjardin and Kingston, 2009). Even when overexpressed, these factors cannot be detected at ALT(−) telomeres (Figure S1D), ruling out an effect due to differences in the expression level. To further characterize the binding mode of these factors, immunoprecipitated sequences were categorized according to their content in the canonical telomere motif GGGTTA (from one to eight occurrences in each 50-nt-long sequencing read). The canonical telomere motif has a similar distribution in the TRF2 and HMBOX1 libraries, implying similar binding specificity in both cell types (Figure 1B). In these libraries, the majority of enriched reads contains seven or eight occurrences of the canonical motif, suggesting that TRF2 and HMBOX1 bind to the canonical telomere sequence in vivo, as expected (Bilaud et al., 1997; Broccoli et al., 1997; Kappei et al., 2013). In contrast, reads associated with NR2C2/F2 showed a different distribution of GGGTTA occurrences, indicating a distinct binding specificity (Figure 1B). To characterize the motifs that allow specific NR2C2/F2 binding, we analyzed these reads further (red brackets in Figure 1B) and found that, among all possible sequences, the GGGTCA motif was specifically enriched in NR2C2/F2 reads (Figure 1C). This is in agreement with the classical sequence-specific binding mode for NHR and our original hypothesis (Benoit et al., 2006; Déjardin and Kingston, 2009; Conomos et al., 2012). Although less frequent (by ~6-fold), the GGGTCA motif is also present in ALT(−) telomeres, indicating that the simple presence of the motif is not sufficient to promote NR2C/F recruitment. Thus, we analyzed the occurrence of the NHR motif in telomeric reads. GGGTCA is essentially found as a multimer in ALT(+) sequences and as a monomer in ALT(−) reads (Figure 1D), suggesting that NR2C/F cannot be recruited to single GGGTCA motif at telomeres. Since these data suggest a classical binding mode for NR2C/F factors, we searched for the full binding sites for these proteins. We identified the direct repeats DR0, DR6, and DR7 (two half-sites in the same orientation and separated by 0, 6, or 7 nucleotides) as the major NR2C/F-binding sites at telomeres (Figure 1E), and these sites are, at least for DR0, ~80-fold enriched in ALT(+)compared to ALT(−) telomeric DNA. Therefore, NR2C/F recruitment is promoted by the presence of DR0, 6, and 7 motifs specifically in ALT telomeres.

The Telomere Protein TRF2 Binds to Hundreds of NR2C/F Regions throughout the Genome of ALT Cells

The aberrant recruitment of NR2C/F factors could suggest that telomeres potentially act as ''molecular sinks'' for these transcription factors in ALT cells. Titration could impinge on the binding and the regulation of NR2C/F targets, which would indirectly control ALT and/or tumorigenesis (Safe et al., 2014). Therefore, we analyzed the genome-wide binding profile of NR2C2 and

(B) TTAGGG content of telomeric reads in IPs from ALT(−) (left) and ALT(+) (right) cells. Histograms display for each library the percentages of telomeric reads containing one to eight TTAGGG occurrences. Red brackets highlight subsets of telomeric reads containing three or four TTAGGG occurrences that are strongly enriched in ALT(+) libraries.

(C) Percentage of telomeric reads containing the indicated repeat variant in ALT(+) and ALT(−) libraries prepared from input DNA, TRF2, NR2F/C, and HMBOX1 IPs. Variant repeats were identified within telomeric reads that contained three or four TTAGGG occurrences and were sorted based on their relative amount in telomeric reads from ALT(+) NR2C/F libraries.

(D) Pie charts showing that GGGTCA multimerization is specific for ALT(+) telomeres. The charts display the number of telomeric reads containing one to seven GGGTCA (left) or NHR unrelated variant ''GGGTTG'' (right) occurrences (n indicates the number of telomeric reads containing the GGGTCA or GGGTTG variants in ALT(+) and ALT(−) input libraries). GGGTTG is used as a control to show that multimerization is specific for the GGGTCA motif.

(E) Normalized number of telomeric reads containing GGGTCA DR0, DR6, DR7, and DR12 in libraries prepared from ALT(−) input and ALT(+) input and orphan receptors NR2F/C IPs.

See also Figures S1 and S2.

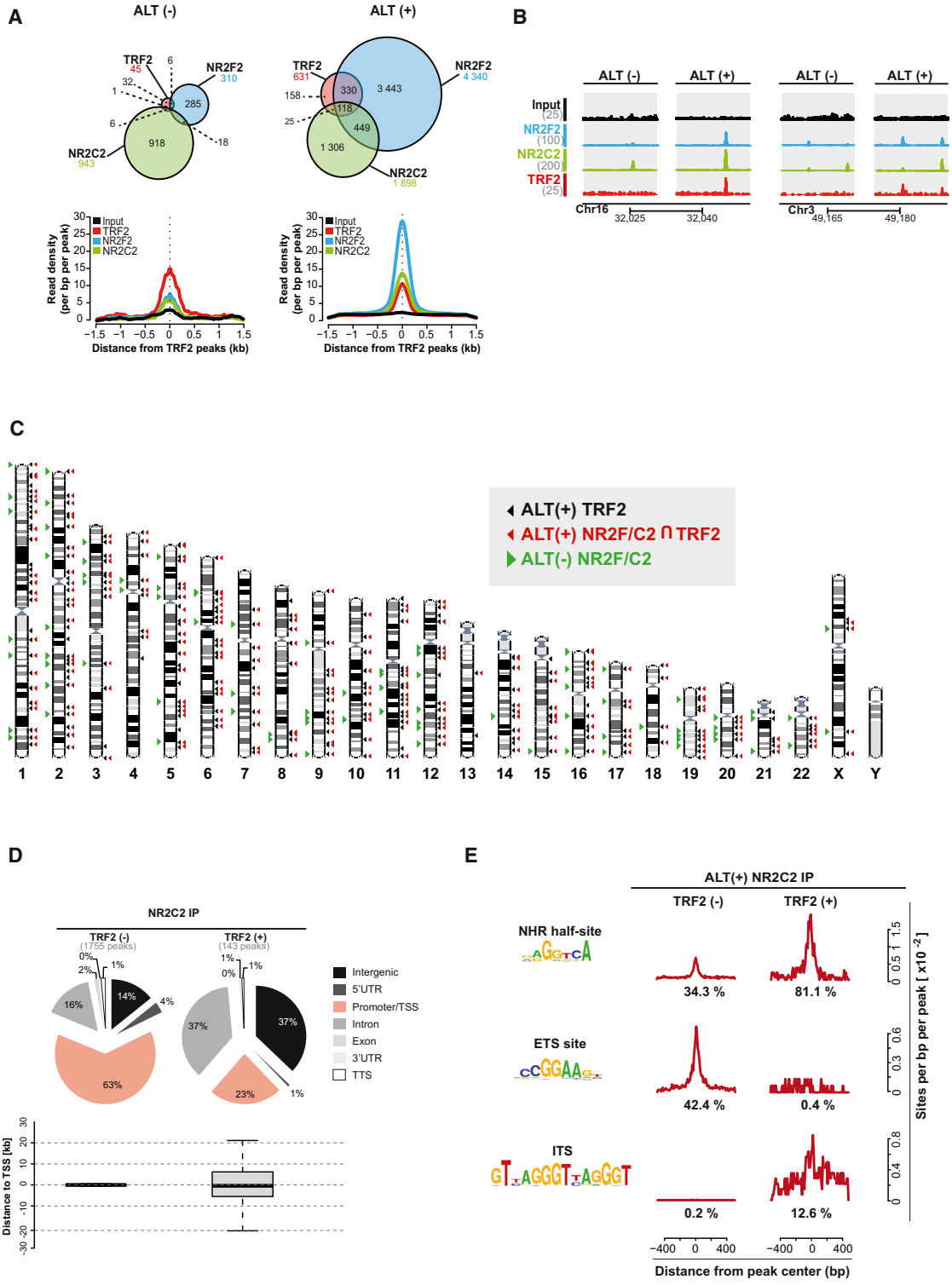


Figure 2. TRF2 Binds to Hundreds of Loci throughout the Genome in ALT Cells

(A) Overlap between TRF2- and NR2F/C-binding sites genome wide. (Top) Venn diagrams displaying peak overlap (values indicate the number of peaks). (Bottom) Average read densities in NR2F/C libraries relative to the TRF2 peaks.

(B) Density profiles of input, TRF2, and NR2F/C reads in two representative loci in ALT(−) and ALT(+). TRF2 is only bound in ALT(+).

(C) Chromosomal locations of ALT(+) TRF2 peaks, ALT(+) NR2C/F peaks overlapping with TRF2 peaks, and ALT(-)NR2C/F peaks.

(legend continued on next page)

NR2F2 and concluded that ALT(+) telomeres do not titrate these transcription factors since we identify more binding regions in ALT(+) than in ALT(−) genomes (Figure 2A). We also analyzed the genomic distribution of TRF2, as it supposedly binds not only to telomeres, but also to rare interstitial telomeric sequences (ITS) (Simonet et al., 2011). ITS contain a various number of iterations of the GGGTTA motif explaining TRF2 recruitment and are well-characterized common fragile sites (CFS) (Bosco and de Lange, 2012). Surprisingly in ALT(+) cells, we identified several hundreds of TRF2-binding sites (Figure 2A) that are not mapping to known ITS. However, most of these sites (75%) overlap with binding sites for NR2C2/F2 proteins (Figures 2A and 2B). Importantly, none of these regions were bound by TRF2 in ALT(−) cells, pointing to an ALT-specific binding mode for TRF2 (Figure 2B). In contrast, in ALT(−) cells, TRF2 binds only to 45 regions, 20 of which corresponded to known ITS often located in subtelomeric regions (Simonet et al., 2011) (Figures 2A–2C). Interestingly, when we looked at the position of these sites with respect to genes, we found that the NR2C2 regions that also recruit TRF2 have a broad distribution and are usually located far from gene promoters (Figure 2D). In contrast, the TRF2-negative NR2C2 regions are principally located at gene transcription start sites. Moreover, the sequence content of these two populations is different. The classical promoter-bound NR2C2 regions are enriched in the expected motifs, in particular the ETS binding sequence (O'Geen et al., 2010) (Figure 2E). On the other hand, the non-genic NR2C2 regions bound by TRF2 lack the ETS motif but are highly enriched in the GGGTCA motif (81% reads). Remarkably, most of these sites (~87%) also lack the canonical telomere sequence, excluding a classical DNA-mediated recruitment mechanism for TRF2. Thus, our data point to an unusual recruitment mode for TRF2 throughout the genome of ALT cells.

NR2C/F Proteins Induce Spatial Proximity of Their Binding Loci

A simple interaction between TRF2 and NR2C2/F2 only in ALT(+) cells is unlikely because a large number of NR2C2/F2 sites remain TRF2 free. Moreover, we failed to detect any interaction between TRF2 and NR2C/F proteins by coimmunoprecipitation (data not shown). Additionally, the absence of telomeric motifs at NR2C2/F2 sites likely excludes a direct TRF2 recruitment. Thus, we hypothesized that physical interactions between ALT telomeric material and endogenous NR2C/F-binding sites occur via NR2C/F proteins (Figure 3A). Accordingly, NR2C/F proteins would bridge ALT telomeres and extra-chromosomal telomeric material generated by the recombination process (Cesare and Griffith, 2004) together and to endogenous NR2C/F regions. By carrying over telomeric material, TRF2 is most probably cross-linked by formaldehyde "in *trans*" at endogenous NR2C/F regions, resulting in the appearance of enrichment peaks for TRF2 throughout the ALT genome at NR2C/F binding regions. The same bridging feature would explain why telomeres extensively interact with each other in ALT cells. Classical methods to measure locus interactions like chromosome conformation capture (Dekker et al., 2002) are challenging for repetitive sequences like telomeres. Thus, to test telomere bridging by these factors, we used super-resolution three-dimensional structured illumination microscopy (3D-SIM), which improves spatial resolution by a factor of eight (Gustafsson et al., 2008). We expressed the DNA-binding mutant NR2C2 protein in the *SaOS-2* ALT cell line. This mutant acts as a dominant-negative for endogenous NR2C2 (Tanabe et al., 2007), stripping it off telomeres (Figure S1C). In *SaOS-2* cells, NR2C2 is the only orphan nuclear receptor bound to telomeres, suggesting that no other NHR could compensate for NR2C2 loss. The expression of this mutant disrupts telomere-telomere interactions, as shown by super-resolution microscopy (Figure 3B). The number of telomeric clusters is reduced, and this is accompanied by an increase in the number of single detectable telomere signals (Figures 3B–3E, S3A, and S3B). On the other hand, expression of the mutant form of NR2C2 in *HeLa 1.2.11* ALT(−) line has no effect on telomere number (Figure S4C). Similar results were obtained upon simultaneously knocking down NR2C1, NR2C2, and NR2F2 (NR2C/F) proteins in the *WI38-VA13* ALT cell line (Figures S4A and S4B). Next, to evaluate bridging between telomeric material and non-telomeric NR2C/F regions, we developed an independent assay. In this approach, the sub-nuclear localization of a fluorescent plasmid DNA harboring NR2C/F DR0-binding sites can be tracked (Figure S3C). If our bridging hypothesis is valid, this plasmid should be targeted to ALT(+) telomeres, which concentrate NR2C/F proteins. In contrast, this plasmid should not co-localize with ALT(−) telomeres, which are devoid of NR2C/F factors. Indeed, despite the formation of large cytoplasmic aggregates in transfected cells, which were unavoidable, the plasmid is efficiently targeted to ALT(+), but not to ALT(−) telomeres (Figure S3C), whereas the control plasmid, not containing NR2C/F binding motifs, does not accumulate at telomeres. Moreover, this recruitment is NR2C/F dependent, as it disappears upon NR2C/F depletion by RNAi. Next, we measured whether NR2C/F tethering to a *LacO* transgenic locus non-homologous to telomeric sequences is sufficient to drive the proximity of that locus to ALT telomeres. We used a transgenic *U2-OS* ALT cell line containing a single *LacO* array (Robinett et al., 1996), in which we expressed a NR2C2-LacI fusion protein, able to bind to the *LacO* array in the absence of NHR binding motif. Tethering NR2C2-LacI to *LacO* leads to the extensive co-localization of the array to telomeric clusters (85% co-localization, Figure 3F), consistent with the bridging feature of this factor. Unexpectedly, this also leads to the appearance of multiple *LacO* signals co-localized with telomeric signals (82% of cells were showing, on average, ten independent *LacO* foci), suggesting a dramatic instability of the *LacO* array upon interaction with ALT telomeres. Neither the tethering of GFP alone nor the

(D) Distribution of NR2C2 peaks overlapping (+) or not (−) with TRF2 peaks in ALT(+) cells. (Top) Pie charts displaying categories according to the nearest transcriptional start site (promoter, TSS). (Bottom) Boxplots showing the distribution of NR2C2 peaks to the nearest TSS.

(E) Overrepresented motifs in NR2C2 peaks overlapping (+) or not (−) with TRF2 peaks in ALT(+) cells (red curves display the average location of the motif around central peak positions, and values indicate the percentage of peaks containing the motif).

See also Figures S1 and S2.

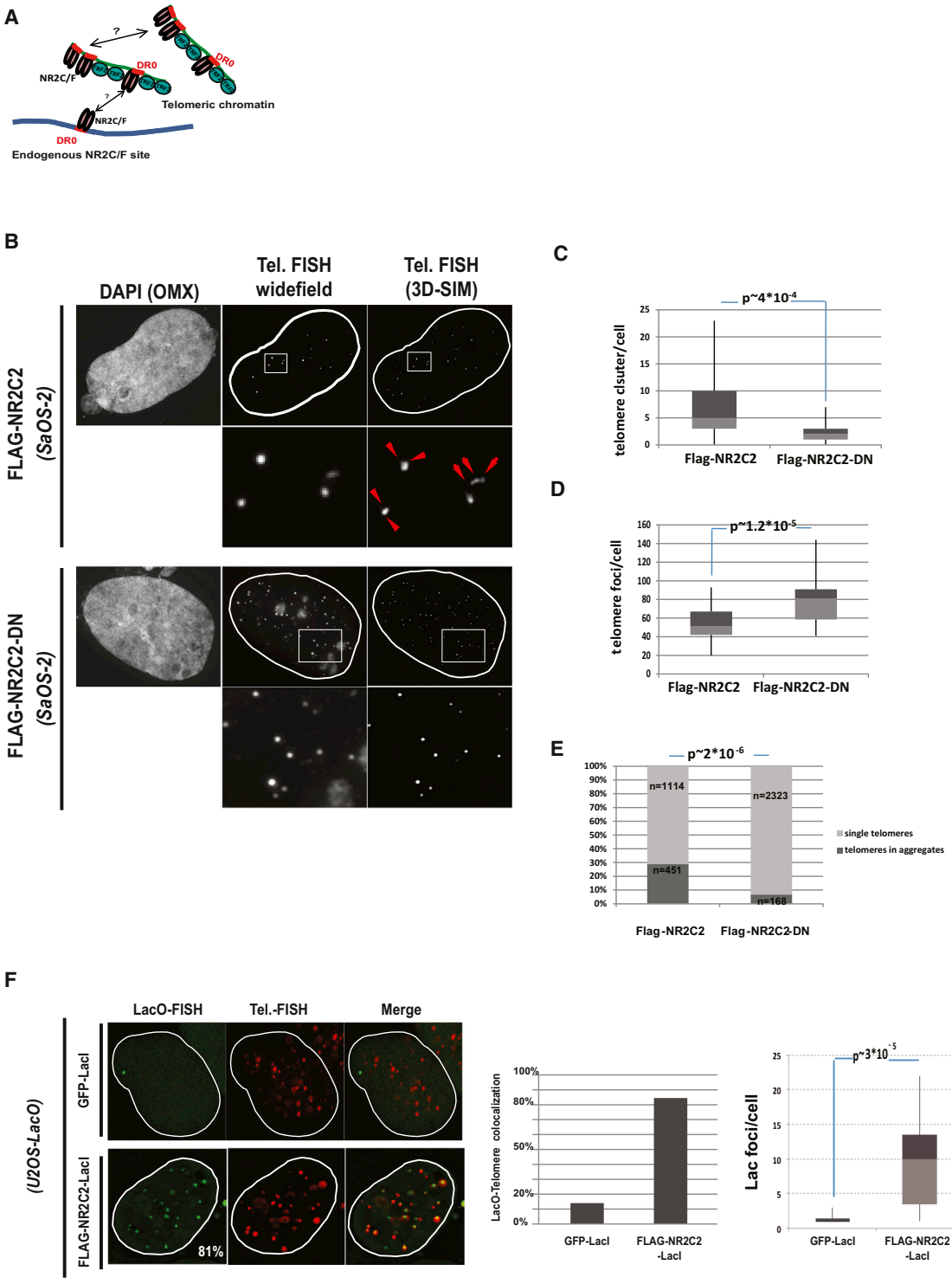


Figure 3. Locus Proximity Induced by NR2C/F Proteins

(A) Model of NR2C/F-induced proximity of telomeric and genomic sites. The same model applies to the telomere/telomere proximity necessary for recombination in ALT.

(B) Telomere FISH showing "de-clustering" of telomeres (visible by 3D-SIM super resolution microscopy) and increased number of telomeric foci upon NR2C2-DN expression in *SaOS-2* cells, suggesting that telomere/telomere interactions are dissociated. The arrowheads indicate single telomeric foci within clusters. Clusters were defined as single signals in the wide field mode, which could be resolved as at least two individual signals in super-resolution mode. Right panels

(legend continued on next page)

tethering of NR2C2-LacI in a transgenic *HeLa* ALT(−) cell line harboring a single *LacO* array leads to its telomeric re-localization, or an amplification of *LacO* (Figures 3F and S3D). Altogether, our data demonstrate that NR2C/F binding is necessary for telomere-telomere proximity and for bridging bound loci to telomeres in the nuclear space. The amplification of *LacO* in the NR2C2 tethering experiment also suggests that bridging to ALT telomeres likely induces instability of the co-localized locus.

Telomere DNA Insertions at NR2C/F Regions in the Genome of ALT Cells

Physical interactions between translocating loci is a major requirement for chromosomal translocations. In fact, translocations are primarily driven by the spatial organization of chromosomes in the nucleus (Misteli and Soutoglou, 2009). Chromosomal architecture must be highly perturbed in ALT cells because ALT genomes show complex chromosome rearrangements with multiple heterogeneous translocations (Guillou and Aurias, 2010; Jain et al., 2010; Lovejoy et al., 2012). Since NR2C/F proteins can drive the physical interaction of bound loci and their apparent instability, we asked whether this feature could trigger instability at endogenous NR2C/F loci. We hypothesized that the genomic NR2C/F-binding sites interacting with telomeric material could also be sites of telomere sequence insertion. Telomeric DNA insertions at discrete genomic sites should yield composite sequencing reads that fail to be aligned to the reference genome. Thus, we focused our analysis on sequencing reads that had mismatches. In both ALT(+) and ALT(−) cell lines, ~90% of the input libraries contain perfectly mappable reads, suggesting comparably high sequencing quality (Figure S2B). Likewise, reads mapping to repeated DNA are of comparably high quality (~88% aligned perfectly). However, while telomere reads from ALT(−) cells are also of high quality, telomere reads from ALT(+) cells are more degenerated and ~25% cannot be aligned without allowing mismatches (Figure S2B). To characterize this telomere specific discrepancy, we analyzed TRF2 libraries in both ALT(+) and ALT(−) cells and examined the sequence organization of the reads that could not be aligned to the reference genome. This showed that, among all possible random hexamers in rearranged sequences, there is a striking bias for the canonical GGGTTA and the variant GGGTCA motifs in ALT(+) samples (Figure 4A). This indicates that, specifically in ALT cells, reads containing these motifs are prone to rearrangements, regardless of their location in the genome. We extended this analysis to NR2C2, NR2F2, and HMBOX1 libraries. Reads in the NR2C2 and NR2F2 libraries are even more degenerated than in the TRF2 and HMBOX1 libraries (Figures S2C and S2D), suggesting that NHR binding regions are intrinsically more unstable. To get insights into the nature of these rearranged sequences, we analyzed their content in details using the strategy depicted in Figure 4B. Strikingly, these rearranged sequences are composed of a part mapping to unique genome regions and aberrant random additions of GGGTCA and/or GGGTTA sequences (Figure 4B, bottom), suggesting that they can result from insertions of telomeric DNA. These insertions do not occur at a precise position but are always located close to (<30 nt) DR0 motifs. The systematic presence of DR0 motifs in the non-rearranged portion of these reads points to the involvement of NR2C/F proteins in targeting the local rearrangement between ALT telomeric DNA and endogenous NR2C/F-binding sites. The few ALT(−) sequences that could not be mapped to the genome contain mostly single-nucleotide changes with no motif addition, suggesting that they did not arise from telomeric DNA insertion. In ALT cells, the rearranged reads map to 23 distinct genomic regions of which 19 (~82%) corresponded to TRF2-positive NR2C/F peaks (Figure S6D). This indicates that only a small subset (19/473, ~4%) of NR2C/F regions able to recruit telomeric material are in fact loci for targeted telomeric insertions (TTI) in ALT cells.

NR2C/F-Driven Telomeric DNA Insertions at DNA Double-Stranded Breaks Are Involved in Chromosomal Translocations in ALT Cells

Artificial insertion of telomeric DNA inside genomes creates ITS, and this has been shown to promote chromosome rearrangements (Kilburn et al., 2001). Because ITS are potential common fragile sites (Bosco and de Lange, 2012), addition of telomeric DNA throughout the genome by TTI can be viewed as a source of genome instability. As TTI parallels ALT, it must be an ongoing mechanism in proliferating cells. To demonstrate that TTI is an active process in ALT cells, we tried to provoke telomere sequence addition throughout the genome. To this aim, we induced DNA double-strand breaks (DSB) by drug treatments or γ irradiation and looked for telomere insertions at these sites by scoring the number of ITS signals on metaphase chromosome spreads. Detectable ITS are more frequent in untreated ALT(+) than ALT(−) chromosomes (~5-fold), in line with our sequencing data showing enrichment in telomeric sequences at genomic sites in ALT(+) cells (Figure 5A). Upon DSB induction, they double in ALT(+). No change was observed in ALT(−) chromosomes, suggesting that breaks are normally repaired without telomeric sequences added in this cell line (Murnane, 2012). In ALT cells, ~30% of both pre-existing and newly formed ITS sites are also bound by NR2C2 and 39% by TRF1, indicating their telomeric origin (Figures 5B and 5C). Consistent with previous data (Bosco and de Lange, 2012), these sites are potentially

show the boxplot quantification of this effect. Top and bottom boxes show the first and third quartile around the mean. p values are from a two-sided Student's t test.

(C) Loss of telomere clustering upon NR2C2 DN expression (aggregate).

(D) Increase in detectable single telomere number.

(E) Distribution of telomeres as individual or clustered signals as measured by super-resolution microscopy. p value from a two-sided Student's t test.

(F) (Left) FISH in *U2-OS* cells harboring the *LacO* transgenic array. Cells were transfected either with GFP-LacI (top) or Flag-NR2C2-LacI (bottom). (Middle) Chart displaying the co-localization of *LacO* with telomere signals. (Right) Chart measuring the extent of *LacO* signal amplification as counted by the number of individual Lac signals in transfected cells. Right panel shows the boxplot quantification of this effect. Top and bottom boxes show the first and third quartiles around the mean. p values are from a two-sided Student's t test.

See also Figures S3 and S4.

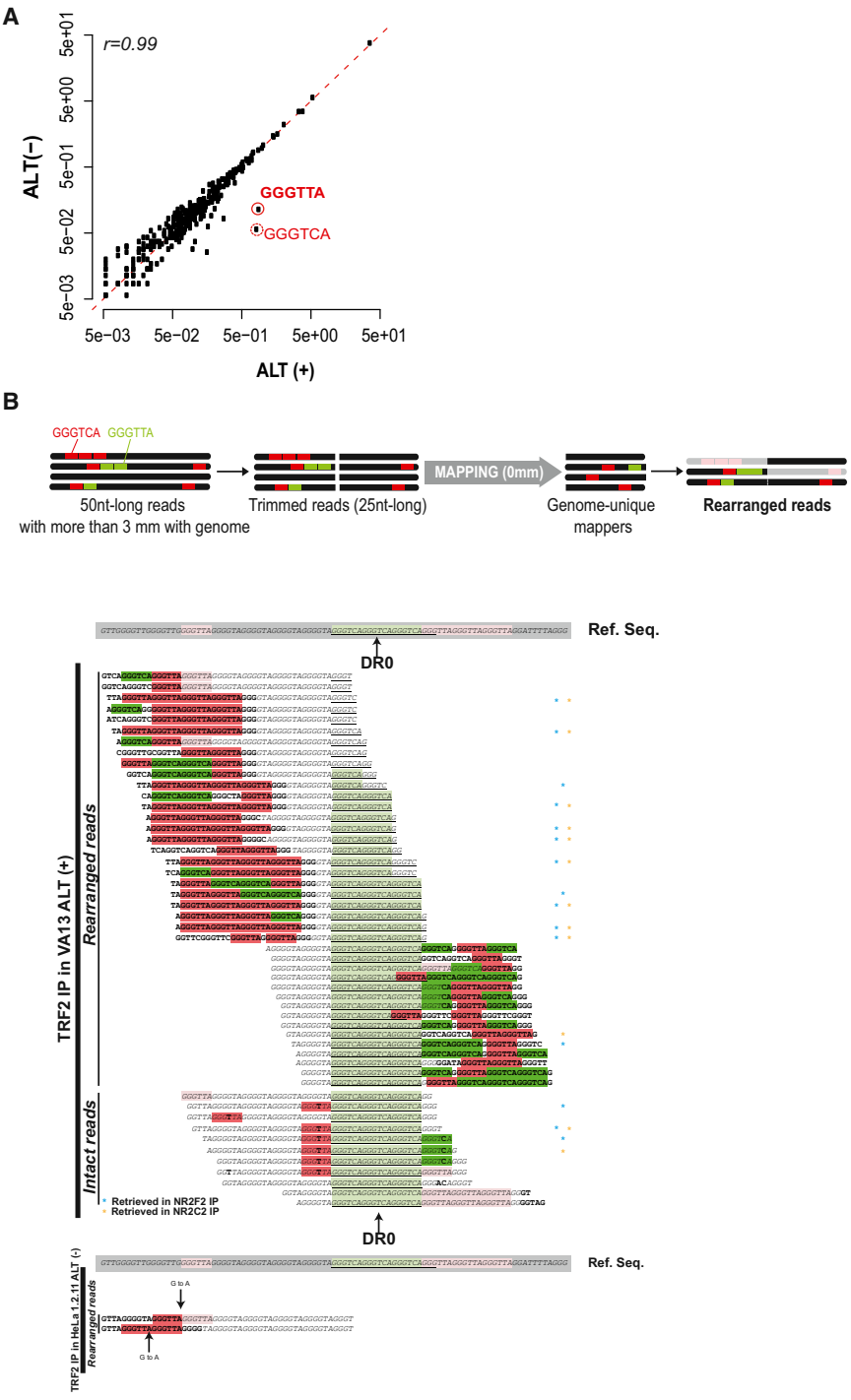


Figure 4. Degenerated Sequencing Reads in ALT Cells Contain Telomeric Motifs

(A) Plot showing the frequency of degenerated reads (≥3 mismatches) containing different hexamers in ALT(−) and ALT(+) TRF2 IPs. Each dot represents a different hexamer. Circled in red: hexamers that are abundant in ALT(+), but not in ALT(−).

(B) (Top) Flowchart displays the strategy to identify "rearranged reads" (see Experimental Procedures). Briefly, non-mappable 50-nt-long reads were split in two 25-nt-long "trimmed" reads that were then individually mapped onto the genome. Non-mappable reads with at least one "trimmed" read that mapped uniquely were considered as "rearranged." (Bottom) Representative examples of rearranged and intact reads mapping near a putative OR-binding site (DR0 underlined) of the Chr2 fusion ITS (reference genome sequence from the hg18 assembly on the top) retrieved from the TRF2 IP library in ALT (+) (top) and ALT(−) (bottom) cells. Bold letters highlight mismatched nucleotides relative to the reference sequence, and italic letters perfectly mapped nucleotides. Red and green boxes highlight GGGTTA and GGGTCA repeats, respectively, found in the reference genomic sequence (light color) or only in the rearranged reads (dark). Asterisks highlight rearranged reads identified also in the NR2F/C2 libraries. In the ALT(−) panel, the arrows indicate single-nucleotide substitutions.

fragile sites, as we observed increased breakage upon combined TRF1 knockdown and aphidicolin treatments (Figure S5A). TTI is NR2C/F dependent because no new ITS formed upon NR2C1, NR2C2, and NR2F2 silencing (Figure 5C). Moreover, NR2C/F-silenced cells have a significantly higher number of chromosomal fusions (arrowheads in Figure 5D), mostly without detectable telomeric signal. These fusions greatly increase upon DSB induction (up to 30% of fused chromosomes). This also suggests the involvement of NR2C/F proteins in preventing chromosomal fusions, a necessary condition for the maintenance of telomeric integrity in ALT cells. Similar results were also obtained in another ALT(+) cell line and with another DNA-damaging agent (Figure S5B), demonstrating that telomere sequence addition to broken chromosomal sites is common in ALT. Moreover, because NR2C/F depletion has no major effect on TRF1 and TRF2 levels (Figure S4A), this effect is unlikely due to a shelterin defect but suggests a protective role for NR2C/F factors on ALT telomeres. Since internal telomeric DNA has the inherent potential to form common fragile sites (CFS) in the human genome (Bosco and de Lange, 2012) and telomeres are fragile sites (Sfeir et al., 2009), we reasoned that TTI should have the potential to form translocations. Such translocations could leave telomeric DNA between two chromosomal segments. DSB induced by γ irradiation in the *VA13* cell line resulted in 88 unique random break points on 693 analyzed chromosomes. 33% (29) of scored translocations have detectable telomeric signals at the translocation points between segments from distinct chromosomes,

as measured by SKY-FISH and telomere FISH on metaphase spreads (Figures 5E and S6 and Table S1). Out of these 29 events, 13 do not involve terminal fusions, suggesting that TTI occurs frequently (~15% of scored translocations) when cells are challenged. Thus, telomeric DNA likely participates in chromosomal translocations in ALT cells.

Binding of NR2C/F Proteins to Telomeres Is a Hallmark of ALT in Tumors and Correlates with the Extent of Genome Rearrangements

To directly explore the link between NR2C/F association to telomeres, genome instability, and cancer in humans, we checked whether these proteins can also be found on telomeres in human primary tumors. We analyzed 180 primary sarcomas from the "complexity index in sarcoma" (CINSARC) signature collection (Chibon et al., 2010) by immunofluorescence and FISH on tissue microarrays (see Table S2 for tumors data). ALT nuclei contain structures called ALT-associated pro-myelocytic leukemia bodies (APB), in which the telomeric DNA is abnormally associated with the PML protein. The presence of APB is a diagnostic marker of ALT (Henson et al., 2005). Accordingly, we scored 54.4% primary tumors analyzed as ALT(+) because these tumors have detectable APB. This is in line with the average ALT occurrence in sarcomas (Henson and Reddel, 2010), validating our approach. The vast majority of ALT(+) tumors (~79%) also show telomeric accumulation of NR2F2 or NR2C2 (Figures 6A and 6B). We also analyzed healthy tissue sections surrounding 12 distinct tumors, and none (0/12) show NR2C2/F2 telomeric accumulation (Figure S7 and Table S3). Thus, NR2C/F telomeric accumulations are cancer specific and do not predate tumor development. Importantly, the extent of NR2C/F association to telomeres correlates with the tumor grade (25% in grades 1 and 2, versus 61% in grade 3) (Figure 6C). As tumor grade is a strong indicator of genome complexity in sarcomas (Chibon et al., 2010), increased telomeric accumulation of NR2C/F proteins mirrors increased genome rearrangements. This suggests the involvement of NR2C/F proteins in generating complex karyotypes in human sarcomas.

DISCUSSION

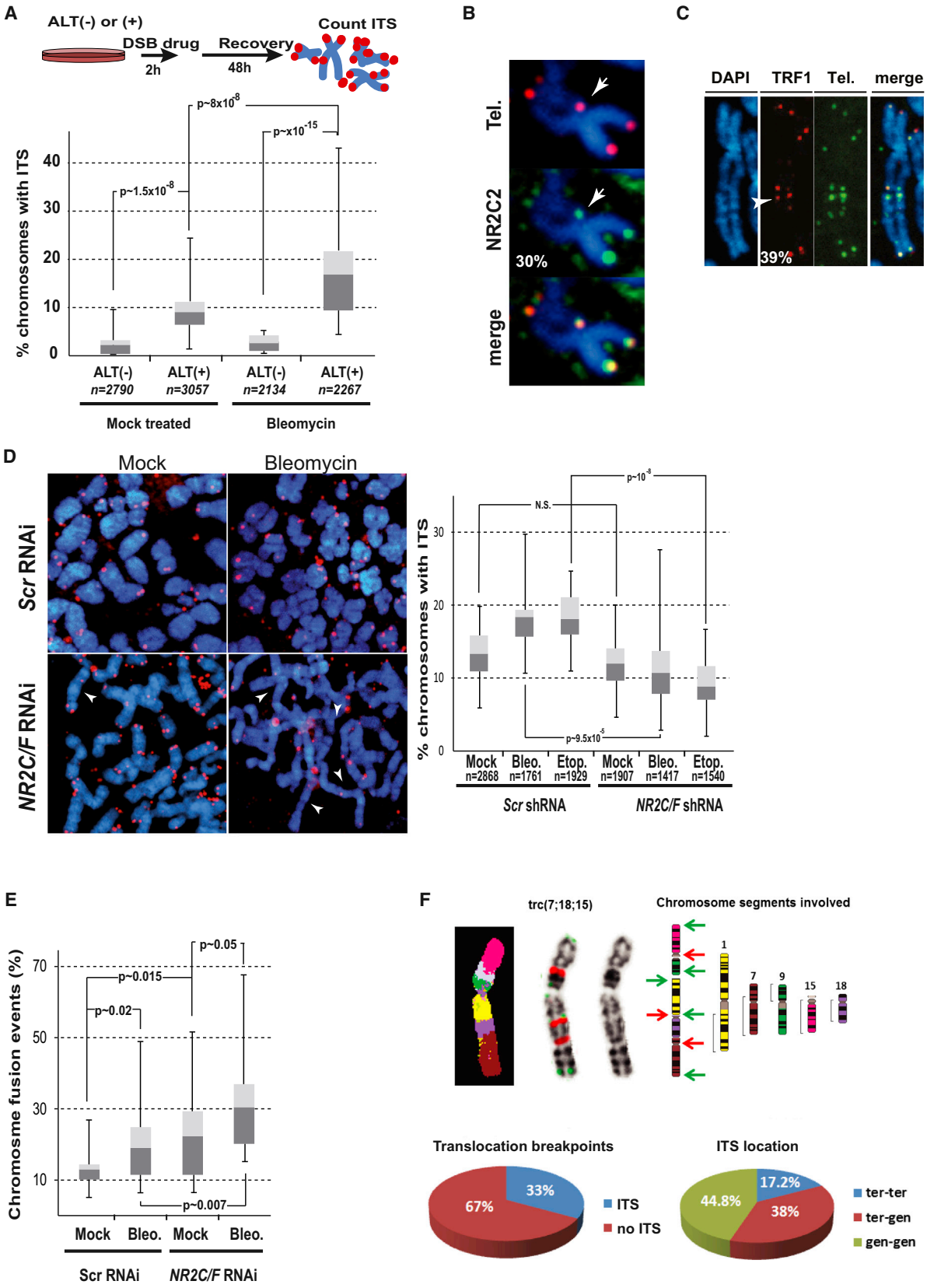
Telomeres are important protective chromosomal structures that safeguard the genome. When telomeres are deprotected, classical BFB cycles ensue and lead to genome alterations. These alterations include deletions and amplifications of chromosomal segments typically observed in tumors (Figure 7A). The BFB cycles, although probably arising as a consequence of the initial loss of important genome surveillance mechanisms, favor the acquisition of oncogenic mutations or accelerate the loss of surveillance pathways that characterize transformed cells (Artandi et al., 2000). Importantly, BFB cycles are stopped when broken chromosomal extremities are healed by the addition of telomeres. Chromosomal healing is usually achieved by the reactivation of telomerase, which is involved in the creation of functional telomeres. Therefore, it is thought that the acquisition of telomere maintenance by telomerase reactivation stabilizes the transformed genome and favors an unlimited proliferation of selected transformed cells. Here, we describe another mechanism of telomere-driven genome instability that actually occurs as a consequence of the activation of aberrant telomere maintenance (Figure 7B). In contrast to the genome stabilization conferred by telomerase, we show that ALT activation also directly destabilizes the genome, using an unexpected mechanism that we name targeted telomere insertion (TTI). We propose that TTI contributes to the complex karyotypes found in tumors or cell lines in which ALT is activated.

NR2C/F-Mediated Long-Distance Interactions

The massive recruitment of orphan nuclear receptors at telomeres in most ALT tumors or cell lines underlies a requirement to maintain a critical function. Consistently, loss of these proteins leads to defective telomere maintenance (Conomos et al., 2012; Déjardin and Kingston, 2009) and chromosomal fusions (Figure 5D). We found no evidence for a role of NR2C/F proteins in transcribing telomeres (data not shown), but we show here that this function is structural. By inducing the proximity of their binding loci, NR2C/F proteins promote physical interactions of telomeric material, a necessary requirement for recombination. An unexpected consequence of this bridging ability is that telomeric material is also able to physically interact with non-telomeric NR2C2/F2-binding sites throughout chromosomes. This represents a further confirmation that bridging is a major feature of NR2C/F proteins. Intriguingly, not all NR2C/F genomic sites have this ability. Telomere-genome interactions usually occur at NR2C/F regions located at a distance from genes, while promoters bound by NR2C/F do not seem to be involved. We believe that the regions able to contact telomeres might be enhancers because these elements are known to interact at long distance and organize local chromosomal architecture (Smallwood and Ren, 2013). How would regions bound by the same transcription factors be located in close physical proximity? We can think of at least two possibilities: either these proteins bind to a shared machinery/structure available in limiting amounts, or these factors have the ability to engage into homotypic interactions. In line with this, RXR proteins, which belong to the NR2B family of NHR, have been shown to be able to oligomerize in vitro (Chen and Privalsky, 1995). The biological significance of such architectural ability is not totally clear, but the clustering of co-regulated regions (presumably bound by the same factors) is a recurrent feature. A benefit of clustering/compartmentalizing nuclear transactions is to increase the local concentration of reactive species to ensure the robustness of biological processes (Déjardin, 2012).

Targeted Telomere Insertions: Implications for Genome Stability

Another NHR, the androgen receptor (AR), was shown to drive the proximity of a subset of its target genes upon hormone induction (Lin et al., 2009; Mani et al., 2009). Combined with genotoxic stress, proximity is required to promote cancer-specific translocations. We show here that the bridging function conferred upon NR2C/F binding drives telomere sequence additions throughout the genome upon genotoxic stress. Thus,



(legend on next page)

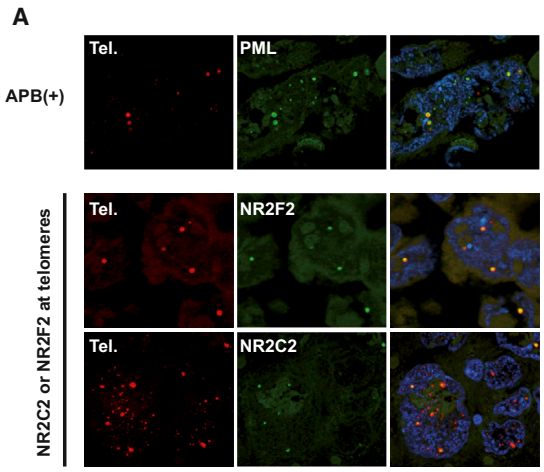
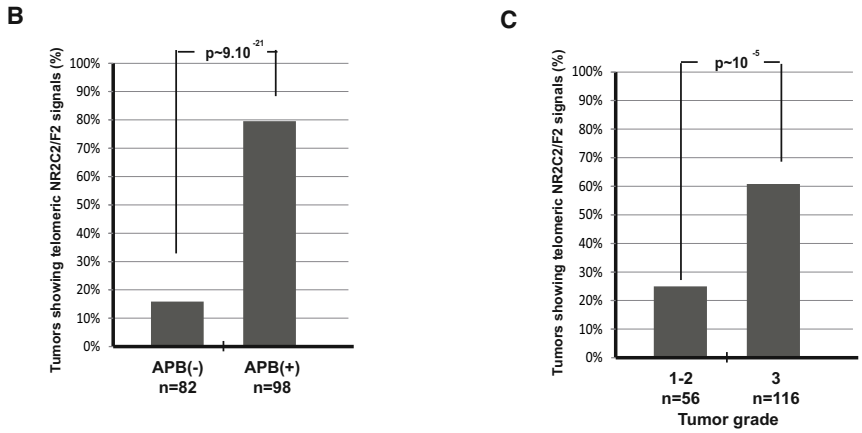


Figure 6. Telomeric NR2C/F Correlate with Sarcoma Grades

(A) Most APB-positive sarcomas samples show NR2C/F localization at telomeres. Immunofluorescence (IF)/telomere FISH in grade 2 leiomyosarcoma biopsies. (Top) APB scoring by co-localization of the FISH signal (red) to PML bodies (green). (Bottom) Localization of NR2C2 or NR2F2 (green) at telomeres (red) in the same tumors.

(B) Frequency of telomeric NR2C/F in APB(+) and in APB(−) tumors based on the IF/FISH analysis. p value from a two-sided t test.

(C) The frequency of NR2C2/F(+) telomeres increases with the tumor grade. See Extended Experimental Procedures for staining procedures. See also Figure S7 and Tables S2 and S3.



we believe that, in cancer cells, the bridging ability of transcription factors, ordinarily used to modulate gene expression, is frequently diverted to trigger chromosomal translocations. In contrast to the well-defined translocation that AR controls in prostate cancer, TTI yields heterogeneous rearrangements. We think that TTI contributes to the appearance and maintenance of complex mutator phenotypes at least at two levels: (1) the ongoing insertions of telomeric DNA at regulatory regions probably directly affect neighboring gene regulation, and (2) inserted telomeric DNA, which was shown to be prone to breakage, can contribute to elevated genomic instability. Our sequencing analysis could not allow us to measure the size of inserted telomeric DNA, but the detection of newly formed ITS by FISH (Figure 5C) suggests that these additions could extend to several kilobase pairs. Sarcomas in which ALT is active have been shown to systematically harbor complex karyotypes with non-specific translocations (Montgomery et al., 2004; Scheel et al., 2001). Although we cannot exclude that other mechanisms could be involved, we propose that TTI contributes to generating a subset of complex chromosomal rearrangements in these cancers.

EXPERIMENTAL PROCEDURES

Cell Culture

The *U2-OS*, *WI-38 VA13 2RA* (*VA13* throughout the text), and *Saos-2* cell lines were obtained from ATCC. The *HeLa 1.2.11* cell line was kindly provided by

Figure 5. Translocation Breakpoints Contain ITS in ALT Cells

(A) (Top) Outline of the ITS induction assay; (bottom) boxplot showing the percentage of chromosomes with ITS signals in different conditions. Top and bottom boxes show the first and third quartiles around the mean. p values are from a two-sided Student's t test.

(B) IF/FISH showing localization of NR2C2 (green) to a DSB-induced ITS signal (red) on a metaphase chromosome (*VA13* cells).

(C) IF/FISH showing localization of TRF1 (red) to a DSB-induced ITS signal (green) on a metaphase chromosome (*VA13* cells).

(D) (Left) Telomeric FISH on chromosome spreads of *Scr* RNAi and triple knockdown *VA13* cells after DSB drug treatment (bleomycin) showing chromosome fusions (arrows); mock, no treatment. (Right) Boxplot displaying induction of ITS sites in *Scr* RNAi and in triple of NR2C1, NR2C2, and NR2F2 knocked down *VA13* cells, treated or not with a DSB inducing drug. Top and bottom boxes show the first and third quartiles around the mean. p values are from a two-sided Student's t test.

(E) Boxplot quantification of the fusion events. Top and bottom boxes show the first and third quartiles around the mean. p values are from a two-sided Student's t test.

(F) (Upper-left) SKY-FISH combined with telomere (green) and centromere (red) FISH showing interstitial telomeric signals at the translocation points between chromosomes 1, 7, 9, 15, and 18 in the ALT+ *VA13* cell line. (Upper-right) Graphical representation of the rearranged chromosome. (Bottom) Distribution of translocation events with or without ITS sites upon γ irradiation. ter-ter, telomere-telomere translocations; ter-Cen, telomere-centomere translocations; gen-gen, telomere-genome translocations.

See also Figures S5 and S6 and Table S1.

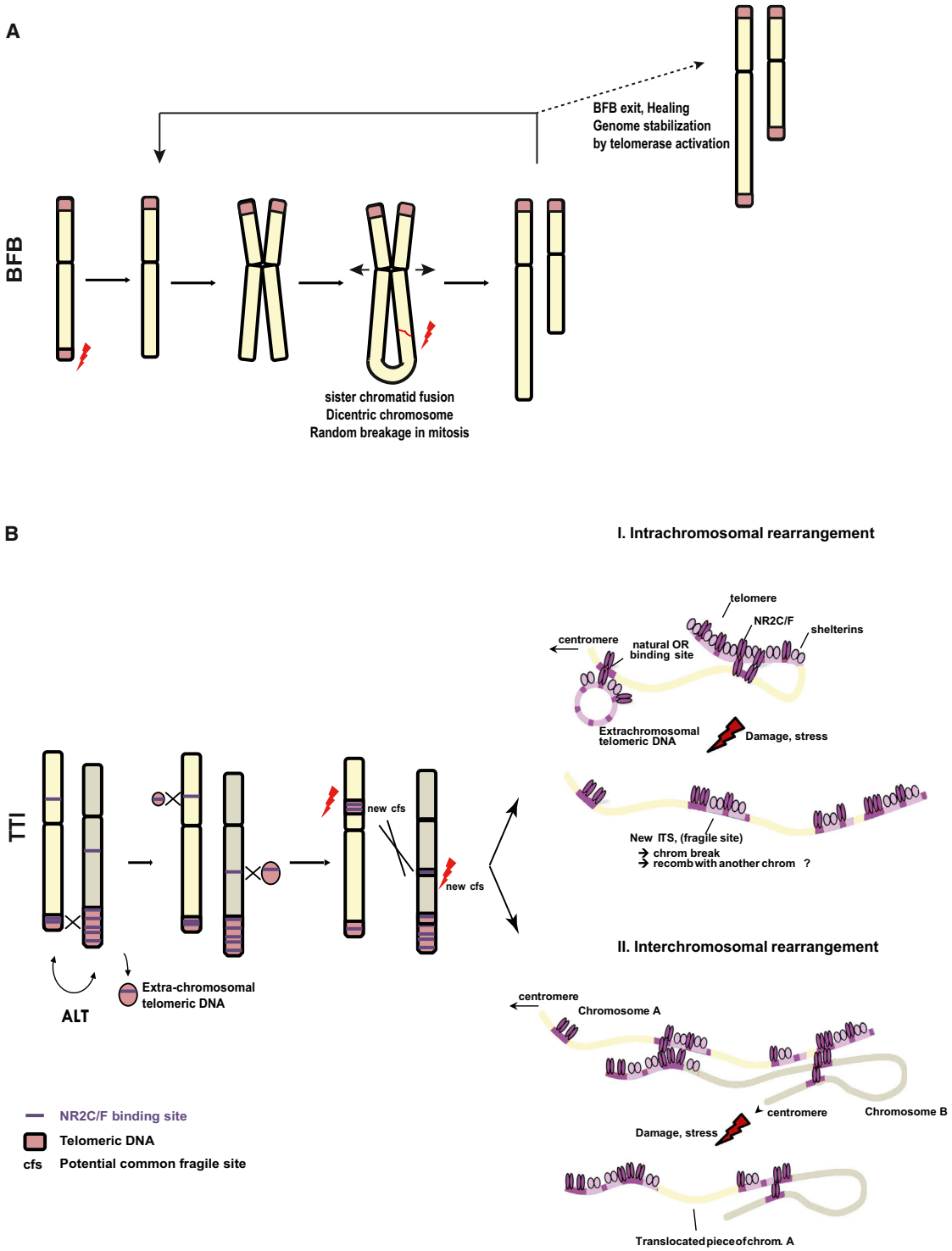


Figure 7. Comparison of TTI with BFB

(A) Outline of the BFB cycles. Instability is stopped by the acquisition of telomerase. (B) Outline of the TTI. Instability is further enhanced by ALT. Insertions of telomeric DNA lead to the creation of potential fragile sites at endogenous NR2C/F binding regions. Two possible outcomes are highlighted upon breakage of these sites, intra- or inter-chromosomal rearrangements.

Titia de Lange. All were cultured in DMEM Glutamax (Life Technologies) supplemented with 10% FBS (Eurobio). *U2-OS* and *HeLa* cell lines, harboring single genomic insertions of the *LacO* array, were cultured with hygromycin B.

Cell Immunofluorescence

Staining was performed as described before (Déjardin and Kingston, 2009), using the following antibodies: anti-TRF2 (Abcam [13579]); anti-PML (Santa Cruz Biotechnology [sc-5621/sc-966]); anti-NR2C2 (PPMX [PP-H0107B-00]); anti-HMBOX1 (Abcam [ab97643]); anti-NR2F2 (Abcam [ab50487]), anti-Flag (Sigma [F7425]). Secondary antibodies: Jackson ImmunoResearch (anti-rabbit DyLight 488 [711-485-152]; anti-mouse DyLight 488 [715-485-150]; anti-mouse DyLight549 [715-505-150]; anti-rabbit DyLight 549 [711-505-152]; anti-mouse DyLight649 [715-495-150]; anti-rabbit DyLight 549 [711-495-152]).

IF/FISH and FISH

After incubation with secondary antibodies, cells were cross-linked at 37°C in 3.6% formaldehyde for 20 min and then incubated at 75°C in 2×SSC for 1 hr and in 0.1M NaOH for 10 min and rinsed with 2×SSC. Then, cells were dehydrated by successive 75% and 100% ethanol baths and air dried. Slides were incubated at 82°C with a telomeric-C PNA probe coupled with FAM or Cy3 (Panagene [F1001/F1002]) for 2 min and then at 37°C for 12 hr. The probe was diluted in hybridization buffer (10% dextran sulfate, 50% deionized formamide, 2×SSC) to the final concentration of 50 mM. Then, cells were washed with 2×SSC at 42°C for 20 min. Finally, slides were mounted in ProlongGold (Life Technologies).

Structural-Illumination-Based Super-Resolution Microscopy and Telomere Clustering

Telomeres were labeled with a PNA probe using the FISH protocol described above. Full z stacks were acquired for each nucleus using the 100× objective. Telomeric signals were counted on Z projections of nuclei. If more than one telomeric signal could be distinguished within a telomeric focus, then such focus was scored as a telomeric cluster.

Targeting *LacO* to Telomeres

Cells were transfected either with GFP-LacI or with GFP-LacI and Flag-NR2C2-LacI, and the array was visualized either with the GFP or by *LacO*-specific FISH. Transfections were performed using the AMAXA nucleofector device using nucleofector reagent from Mirus according to manufacturers' instructions. Quantification of *LacO* localization to telomeres was performed 48 hr after transfection.

Transfections and Constructs

We performed triple NR2C/F knockdown because neither single nor double knockdowns had a strong impact on ALT telomeres, probably because of redundant function (Déjardin and Kingston, 2009). The triple knockdown was performed with Stealth RNAi (Life Technologies) using Lipofectamine RNAiMAX (Life Technologies). The oligonucleotides were transfected twice within 48 hr. RNAi oligonucleotides used in the experiment were: siRNA negative control Med GC, NR2F2 [NR2F2MSS235955], NR2C1 [NR2C1HSS110947], and NR2C2 [NR2C2HSS110950]. Flag-NR2C2 and Flag-NR2C2-DN were kindly provided by Osamu Tanabe. Constructs were transfected using the AMAXA Cell Line Nucleofector Kit R (*VA13* cells) or Kit V (*Saos-2* cells) according to the manufacturer's instructions.

ITS Induction

Asynchronously growing cells were incubated with bleomycin (Calbiochem, 30 mU/ml) or etoposide (Sigma, 10 μM). After 2 hr, the drug-containing medium was replaced with fresh medium, and colcemid was added 43 hr after drug release to prepare chromosome spreads. We quantified the interstitial telomeric FISH signal on single and (rare) fused chromosomes. In triple *Nr2c/f* knockdown, because of the high frequency of chromosomal fusions, we quantified the interstitial telomeric FISH signal only from single (not fused) chromosomes to avoid counting fused chromosomes with residual telomere signal as ITS.

Aphidicolin Treatment and TRF1 Knockdown

Cells were treated with low doses (0.3 μM) of Aphidicolin for 24 hr. Aphidicolin was added 24 hr after a second TRF1 RNAi transfection (two RNAi transfections within 48 hr). After 20 hr of Aphidicolin treatment, colcemid was added to medium to induce mitotic arrest. RNAi used for TRF1 knockdown was purchased from Dharmacon (SmartPool RNAi).

Western Blotting

Nuclear extracts from transfected cells were run on 12% Bis-Tris gels (Life Technologies [NP0341]) and transferred to PVDF membranes using a liquid transfer system for 2 hr at 300 mA. The membrane was blocked for 30 min at room temperature in 1× PBS containing 5% milk and then incubated with primary antibodies diluted 1/1,000 in the same buffer for 2 hr at room temperature, washed twice for 15 min in PBS-0.05%Tween 20, and incubated for 1 hr with secondary antibodies diluted 1/5,000 in 5% milk-PBS which was followed by two washes in PBS-0.05% Tween 20. Antibodies: αFlag (Sigma [F7425]), αNR2C2 (PPMX [PP-H0107B-00]), αHMBOX1 (Abcam [ab97643]), αTRF2 (Abcam [ab13579]), αTRF1 (Abcam [ab10579]), αNR2F2 (Abcam [ab50487]), αNR2C1 (Santa Cruz [sc-9087]), αPCNA (Santa Cruz [sc-25280]), anti-rabbit-HRP (Sigma [A0545]), anti-mouse-HRP (Sigma [A4416]).

ChIP Sequencing

Cells growing in monolayer were cross-linked in 1% formaldehyde/PBS for 10 min, washed twice in PBS, and scraped in PBS/0.05% Tween. Then cells were pelleted, washed in PBS, and incubated at 4°C in lysis buffer 1 for 10 min, then at room temperature in lysis buffer 2 for 10 min, dounced with a tight pestle, centrifuged, resuspended in lysis buffer 3 (1 ml/IP for ~2 × 10^7cells) and sonicated (12 pulses of 70% power, 15'' ON, 45'' OFF using a Misonix sonicator) to obtain chromatin fragments of 200 bp. Subsequently, chromatin was pre-cleared at 4°C with 10 μl/ml BSA-blocked Dynabeads (Life Technologies, a mix 1:1 of protein A and protein G beads) for 30 min and incubated at 4°C with 5 μg antibody/IP overnight. Chromatin was then incubated with magnetic beads at 4°C for 2 hr. Beads were washed five times with RIPA and once in TE with 50 mM NaCl. Chromatin was eluted from the beads by incubating in elution buffer with shaking for 30 min. Cross-linking was removed by overnight incubation at 65°C. After RNaseA and proteinase K treatments, the DNA was extracted with phenol:chloroform and ethanol precipitation. Isolated DNA was resuspended in water. ChIP experiments were performed four to six times independently for each antibody. Libraries were cloned and sequenced by Fasteris SA (Switzerland) using the Illumina strategy (HiSeq2000, single-end).

Antibodies

anti-TRF2 (SantaCruz Biotechnology, sc-9143); anti-NR2C2-PPMX (PP-H0107B-00); anti-HMBOX1 (Abcam, ab97643); anti-NR2F2 (Abcam, ab50487).

Buffers Composition

Buffers used for ChIP have been described previously (Lee et al., 2006).

Bioinformatic Analysis

Bioinformatic analysis is described in the Extended Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental information includes Extended Experimental Procedures, seven figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2015.01.044.

ACKNOWLEDGMENTS

We would like to thank Jan Karlseder and Cyril Ribeyre for critical reading of the manuscript and members of the Déjardin lab for stimulating discussions, Osamu Tanabe for the gift of the NR2C2 constructs, and Evi Soutoglou for the gift of the *HeLa LacO* transgenic line. Work in J.D.'s lab is funded by the European Research Council (StG#242934) and INSERM AVENIR. P.M. is funded by the Association pour la Recherche contre le Cancer.

Received: December 19, 2013
Revised: November 18, 2014
Accepted: January 27, 2015
Published: February 26, 2015

REFERENCES

Allshire, R.C., Dempster, M., and Hastie, N.D. (1989). Human telomeres contain at least three types of G-rich repeat distributed non-randomly. Nucleic Acids Res. *17*, 4611–4627.

Artandi, S.E., Chang, S., Lee, S.L., Alson, S., Gottlieb, G.J., Chin, L., and DePinho, R.A. (2000). Telomere dysfunction promotes non-reciprocal translocations and epithelial cancers in mice. Nature *406*, 641–645.

Benoit, G., Cooney, A., Giguere, V., Ingraham, H., Lazar, M., Muscat, G., Perlmann, T., Renaud, J.P., Schwabe, J., Sladek, F., et al. (2006). International Union of Pharmacology. LXVI. Orphan nuclear receptors. Pharmacol. Rev. *58*, 798–836.

Bilaud, T., Brun, C., Ancelin, K., Koering, C.E., Laroche, T., and Gilson, E. (1997). Telomeric localization of TRF2, a novel human telobox protein. Nat. Genet. *17*, 236–239.

Bosco, N., and de Lange, T. (2012). A TRF1-controlled common fragile site containing interstitial telomeric sequences. Chromosoma *121*, 465–474.

Broccoli, D., Smogorzewska, A., Chong, L., and de Lange, T. (1997). Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. Nat. Genet. *17*, 231–235.

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

Cesare, A.J., and Griffith, J.D. (2004). Telomeric DNA in ALT cells is characterized by free telomeric circles and heterogeneous t-loops. Mol. Cell. Biol. *24*, 9948–9957.

Chen, H., and Privalsky, M.L. (1995). Cooperative formation of high-order oligomers by retinoid X receptors: an unexpected mode of DNA recognition. Proc. Natl. Acad. Sci. USA *92*, 422–426.

Chibon, F., Lagarde, P., Salas, S., Pérot, G., Brouste, V., Tirode, F., Lucchesi, C., de Reynies, A., Kauffmann, A., Bui, B., et al. (2010). 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.

Conomos, D., Stutz, M.D., Hills, M., Neumann, A.A., Bryan, T.M., Reddel, R.R., and Pickett, H.A. (2012). Variant repeats are interspersed throughout the telomeres and recruit nuclear receptors in ALT cells. J. Cell Biol. *199*, 893–906.

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

Déjardin, J. (2012). How chromatin prevents genomic rearrangements: locus colocalization induced by transcription factor binding. BioEssays *34*, 90–93.

Déjardin, J., and Kingston, R.E. (2009). Purification of proteins associated with specific genomic Loci. Cell *136*, 175–186.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science *295*, 1306–1311.

Guillou, L., and Aurias, A. (2010). Soft tissue sarcomas with complex genomic profiles. Virchows Arch. *456*, 201–217.

Gustafsson, M.G., Shao, L., Carlton, P.M., Wang, C.J., Golubovskaya, I.N., Cande, W.Z., Agard, D.A., and Sedat, J.W. (2008). Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. Biophys. J. *94*, 4957–4970.

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

Henson, J.D., and Reddel, R.R. (2010). Assaying and investigating Alternative Lengthening of Telomeres activity in human cells and cancers. FEBS Lett. *584*, 3800–3811.

Henson, J.D., Hannay, J.A., McCarthy, S.W., Royds, J.A., Yeager, T.R., Robinson, R.A., Wharton, S.B., Jellinek, D.A., Arbuckle, S.M., Yoo, J., et al. (2005). 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.

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

Jain, S., Xu, R., Prieto, V.G., and Lee, P. (2010). Molecular classification of soft tissue sarcomas and its clinical applications. Int. J. Clin. Exp. Pathol. *3*, 416–428.

Kappei, D., Butter, F., Benda, C., Scheibe, M., Draškovič, I., Stevense, M., Novo, C.L., Basquin, C., Araki, M., Araki, K., et al. (2013). HOT1 is a mammalian direct telomere repeat-binding protein contributing to telomerase recruitment. EMBO J. *32*, 1681–1701.

Kilburn, A.E., Shea, M.J., Sargent, R.G., and Wilson, J.H. (2001). Insertion of a telomere repeat sequence into a mammalian gene causes chromosome instability. Mol. Cell. Biol. *21*, 126–135.

Lee, T.I., Johnstone, S.E., and Young, R.A. (2006). Chromatin immunoprecipitation and microarray-based analysis of protein location. Nat. Protoc. *1*, 729–748.

Lin, C., Yang, L., Tanasa, B., Hutt, K., Ju, B.G., Ohgi, K., Zhang, J., Rose, D.W., Fu, X.D., Glass, C.K., and Rosenfeld, M.G. (2009). Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. Cell *139*, 1069–1083.

Lovejoy, C.A., Li, W., Reisenweber, S., Thongthip, S., Bruno, J., de Lange, T., De, S., Petrini, J.H., Sung, P.A., Jasin, M., et al.; ALT Starr Cancer Consortium (2012). Loss of ATRX, genome instability, and an altered DNA damage response are hallmarks of the alternative lengthening of telomeres pathway. PLoS Genet. *8*, e1002772.

Mani, R.S., Tomlins, S.A., Callahan, K., Ghosh, A., Nyati, M.K., Varambally, S., Palanisamy, N., and Chinnaiyan, A.M. (2009). Induced chromosomal proximity and gene fusions in prostate cancer. Science *326*, 1230.

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

Misteli, T., and Soutoglou, E. (2009). The emerging role of nuclear architecture in DNA repair and genome maintenance. Nat. Rev. Mol. Cell Biol. *10*, 243–254.

Montgomery, E., Argani, P., Hicks, J.L., DeMarzo, A.M., and Meeker, A.K. (2004). Telomere lengths of translocation-associated and nontranslocation-associated sarcomas differ dramatically. Am. J. Pathol. *164*, 1523–1529.

Murnane, J.P. (2012). Telomere dysfunction and chromosome instability. Mutat. Res. *730*, 28–36.

O'Geen, H., Lin, Y.H., Xu, X., Echipare, L., Komashko, V.M., He, D., Frietze, S., Tanabe, O., Shi, L., Sartor, M.A., et al. (2010). Genome-wide binding of the orphan nuclear receptor TR4 suggests its general role in fundamental biological processes. BMC Genomics *11*, 689.

Robinett, C.C., Straight, A., Li, G., Willhelm, C., Sudlow, G., Murray, A., and Belmont, A.S. (1996). In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. J. Cell Biol. *135*, 1685–1700.

Safe, S., Jin, U.H., Hedrick, E., Reeder, A., and Lee, S.O. (2014). Minireview: role of orphan nuclear receptors in cancer and potential as drug targets. Mol. Endocrinol. *28*, 157–172.

Sandelin, A., and Wasserman, W.W. (2005). Prediction of nuclear hormone receptor response elements. Mol. Endocrinol. *19*, 595–606.

Scheel, C., Schaefer, K.L., Jauch, A., Keller, M., Wai, D., Brinkschmidt, C., van Valen, F., Boecker, W., Dockhorn-Dworniczak, B., and Poremba, C. (2001). Alternative lengthening of telomeres is associated with chromosomal instability in osteosarcomas. Oncogene *20*, 3835–3844.

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

Simonet, T., Zaragosi, L.E., Philippe, C., Lebrigand, K., Schouteden, C., Augereau, A., Bauwens, S., Ye, J., Santagostino, M., Giulotto, E., et al. (2011). The human TTAGGG repeat factors 1 and 2 bind to a subset of interstitial telomeric sequences and satellite repeats. Cell Res. *21*, 1028–1038.

Smallwood, A., and Ren, B. (2013). Genome organization and long-range regulation of gene expression by enhancers. Curr. Opin. Cell Biol. *25*, 387–394.

Tanabe, O., McPhee, D., Kobayashi, S., Shen, Y., Brandt, W., Jiang, X., Campbell, A.D., Chen, Y.T., Chang, Cs., Yamamoto, M., et al. (2007). Embryonic and fetal beta-globin gene repression by the orphan nuclear receptors, TR2 and TR4. EMBO J. *26*, 2295–2306.

Taylor, B.S., Barretina, J., Maki, R.G., Antonescu, C.R., Singer, S., and Ladanyi, M. (2011). Advances in sarcoma genomics and new therapeutic targets. Nat. Rev. Cancer *11*, 541–557.