Fusion of short telomeres in human cells is characterized by extensive deletion and microhomology, and can result in complex rearrangements

Boitelo T. Letsolo, Jan Rowson and Duncan M. Baird*

Department of Pathology, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, UK

Received September 8, 2009; Revised November 18, 2009; Accepted December 3, 2009

ABSTRACT
Telomere fusion is an important mutational event that has the potential to lead to large-scale genomic rearrangements of the types frequently observed in cancer. We have developed single-molecule approaches to detect, isolate and characterize the DNA sequence of telomere fusion events in human cells. Using these assays, we have detected complex fusion events that include fusion with interstitial loci adjacent to fragile sites, intra-molecular rearrangements, and fusion events involving the telomeres of both arms of the same chromosome consistent with ring chromosome formation. All fusion events were characterized by the deletion of at least one of the telomeres extending into the sub-telomeric DNA up to 5.6 kb; close to the limit of our assays. The deletion profile indicates that deletion may extend further into the chromosome. Short patches of DNA sequence homology with a G:C bias were observed at the fusion point in 60% of events. The distinct profile that accompanies telomere fusion may be a characteristic of the end-joining processes involved in the fusion event.

INTRODUCTION
By capping the ends of linear chromosomes and preventing fusion, telomeres play a key role in upholding chromosomal integrity. The loss of the end-capping function can occur as a consequence of gradual telomere erosion or stochastic telomeric deletion; both of which can create short dysfunctional telomeres (1–3). The end-capping function can also be abrogated by the experimental disruption of the telomeric protein structure, which creates telomeres that are dysfunctional, but full length (4). The loss of telomere function can result in the fusion of telomeres with other telomeric loci, or non-telomeric double-stranded DNA breaks; this can result in the formation of dicentric chromosomes. With ongoing cell division the resulting cycles of anaphase-bridging, breakage and fusion can result in genomic rearrangements such as non-reciprocal translocations (NRTs) and localized amplifications that are common in a broad spectrum of tumour types (5–8). It is thus considered that telomere dysfunction may facilitate progression in the early stages of tumorigenesis. Consistent with this, telomere length is reduced in some types of carcinoma compared to matched normal tissue (9–11) and an increase in anaphase-bridging, a surrogate marker of telomere fusion, is detected at the adenoma-carcinoma transition, a point that is characterized by increased chromosomal instability in colorectal carcinoma (12), breast (13) and oral squamous cell carcinomas (14). Furthermore, breakpoint frequency and telomere length is negatively correlated at specific chromosome ends in some carcinoma derived cells (15–17). These data are backed up by evidence from telomerase knockout mouse models, which show that telomere loss correlates with genomic instability which may drive a higher rate of tumour formation and a lower age of onset (18,19). Carcinomas from telomerase knockout mice contain high frequencies of chromosomes lacking telomeric signals, anaphase bridges, NRTs and end-to-end fusions (6).

If the telomere repeat array and the subtelomeric sequences remain intact, telomere fusion between sister chromatids will generate inverted repeats that may create fragile sites, such that a subsequent breakage may then occur close to the site of the original telomere fusion (20): a second round sister chromatid fusion of the broken chromosome, may then result in further amplification. Thus, in addition to large-scale genomic rearrangements,
such as NRTs, sister chromatid telomere fusion and cycles of anaphase-bridging, breakage and fusion can result in localized gene amplification (21). Hence, short dysfunctional telomeres have the potential to create the types of mutation that may drive progression in the early stages of cancer, it is therefore of interest to understand the mechanistic basis of telomere fusion in human cells.

Experimental intervention in the protein structure of the telomere, such as abrogating the function of TRF2, results in immediate telomeric de-protection and fusion that is dependent upon factors required for non-homologous end-joining (NHEJ), including ligase IV, DNA PKcs and Ku (22,23). However in contrast, the fusion of short dysfunctional telomeres in telomerase mutants of Yeast, Arabidopsis and mice is independent of Ku70, DNA-PKcs and Ligase IV (1,24–26). These data imply that critically shortened telomeres may be subjected to fusion that is mediated by alternative end-joining mechanisms. We have investigated fusion between short dysfunctional telomeres in human cells using a PCR-based assay that detects single-molecule telomere fusion events. This revealed that fusion was accompanied by deletion and short patches of homology at the fusion point (2). We have now expanded this assay, which was previously based on just two chromosome ends, to include at least 43% of telomeres in the human genome. Sequence analysis of a large set of telomere fusion events revealed more extensive sub-telomeric deletions that extend up to the limit of our assays (6.1 kb) that are consistent with error-prone DNA-PK independent end-joining (27–30). In addition, fusion events were characterized by short patches of homology at the fusion point that exhibited a G:C bias. We also observed events that may result in large-scale genomic rearrangements, including, fusion events that are consistent with chromosome circularization, fusion with interstitial telomeres, and intra-molecular fusion prior to chromosomal fusion.

MATERIALS AND METHODS

Cell culture

Telomerase expressing HEK293 cells that were originally derived from human embryonic kidney cells transformed with adenoviral DNA (31), were cultured in 88% Dulbecco’s Modified Eagle’s Medium (DMEM; containing 4500 mg/l D-glucose, 110 mg/l sodium pyruvate, and Dulbecco’s Modified Eagle’s Medium (DMEM; containing 4500 mg/l D-glucose, 110 mg/l sodium pyruvate, and 10% (v/v) Foetal Calf Serum, 1
cell culture

MATERIALS AND METHODS

Cell culture

Telomerase expressing HEK293 cells that were originally derived from human embryonic kidney cells transformed with adenoviral DNA (31), were cultured in 88% Dulbecco’s Modified Eagle’s Medium (DMEM; containing 4500 mg/l D-glucose, 110 mg/l sodium pyruvate, and non-essential amino acids but no L-glutamine), supple-

MATERIALS AND METHODS

Cell culture

Telomerase expressing HEK293 cells that were originally derived from human embryonic kidney cells transformed with adenoviral DNA (31), were cultured in 88% Dulbecco’s Modified Eagle’s Medium (DMEM; containing 4500 mg/l D-glucose, 110 mg/l sodium pyruvate, and non-essential amino acids but no L-glutamine), supple-

DNA extraction

Genomic DNA was extracted by standard proteinase K, RNase A and phenol/chloroform protocols (33). High molecular weight DNA was solubilized by either digesting with EcoRI or serial dilutions in 10 mM Tris–HCl (pH 8.0) to ~100 ng/µl, prior to quantifying the DNA in triplicate by Hoechst 33258 fluorometry.

Telomere fusion assay

Several primers were designed within the sub-telomeric repeat regions that are shared among multiple chromosomes. Primers that target some of the TelBamII and TelBam 3.4 classes of subtelomeric repeat regions (34–36) were designed using the 21q and 16p telomere-adjacent sequences respectively (35).

The telomere fusion assay was then carried out as described previously (2) with the following modifications. The EcoRI solubilized DNA was used in the assay with the 21q group primers and 100 ng/µl Tris diluted DNA for the 16p group primers. PCR reactions to detect fusion involving the 21q group telomeres were undertaken with XpYpM/17p6 and 21q1 and for the 16p group with XpYpM/17p6 and 16p1 primers. Fusion molecules were detected and the frequencies quantified by Southern blotting and hybridization with the XpYp and 17p telomere-adjacent probes as described previously (2). In order to determine the chromosomes participating in the fusion events for subsequent sequence characterization, further hybridizations were undertaken with the 21q or 16p telomere-adjacent probes generated by PCR with 21qseq1/21qseq1rev or the 16p2/16pseq1rev PCR primers (Figure 1A), these probes yield additional non-specific products and thus were not used for quantification. Any fusion products were then re-amplified for direct sequence analysis using nested PCR primers (XpYpO, 17p7, 21qseq1 and 16p2, Figure 1A) as determined by Southern hybridization, as described previously (2).

To provide an estimate of the frequency of fusion, the total number of bands generated from each DNA analysed was counted and the number of input molecules in each analysis was estimated from the amount of DNA used in each reaction together with the size of a diploid human genome. Thus, the frequency of fusion could be calculated from the number of positive molecules compared to the number of input molecules. This will represent an underestimate of the total telomere fusion frequency as the fusion assays described here are capable of detecting a small fraction of the total number of different pair wise telomere fusion combinations in the human genome.

Oligonucleotides

XpYp and 17p primers used during this study have been described previously (2). The chromosome ends that previously hybridized to the TelBam11 probe (21q, 1q, 2q, 5q, 6q, 6p, 8p, 10q, 13q, 17q, 19p, 19q, 22q and the 2q13 interstitial telomeric locus) and to the TelBam3.4 probe (16p, 1p, 9p, 12p, 15q, XqYq and the 2q14

MATERIALS AND METHODS

Cell culture

Telomerase expressing HEK293 cells that were originally derived from human embryonic kidney cells transformed with adenoviral DNA (31), were cultured in 88% Dulbecco’s Modified Eagle’s Medium (DMEM; containing 4500 mg/l D-glucose, 110 mg/l sodium pyruvate, and non-essential amino acids but no L-glutamine), supple-

MATERIALS AND METHODS

Cell culture

Telomerase expressing HEK293 cells that were originally derived from human embryonic kidney cells transformed with adenoviral DNA (31), were cultured in 88% Dulbecco’s Modified Eagle’s Medium (DMEM; containing 4500 mg/l D-glucose, 110 mg/l sodium pyruvate, and non-essential amino acids but no L-glutamine), supple-

DNA extraction

Genomic DNA was extracted by standard proteinase K, RNase A and phenol/chloroform protocols (33). High molecular weight DNA was solubilized by either digesting with EcoRI or serial dilutions in 10 mM Tris–HCl (pH 8.0) to ~100 ng/µl, prior to quantifying the DNA in triplicate by Hoechst 33258 fluorometry.

Telomere fusion assay

Several primers were designed within the sub-telomeric repeat regions that are shared among multiple chromosomes. Primers that target some of the TelBamII and TelBam 3.4 classes of subtelomeric repeat regions (34–36) were designed using the 21q and 16p telomere-adjacent sequences respectively (35).

The telomere fusion assay was then carried out as described previously (2) with the following modifications. The EcoRI solubilized DNA was used in the assay with the 21q group primers and 100 ng/µl Tris diluted DNA for the 16p group primers. PCR reactions to detect fusion involving the 21q group telomeres were undertaken with XpYpM/17p6 and 21q1 and for the 16p group with XpYpM/17p6 and 16p1 primers. Fusion molecules were detected and the frequencies quantified by Southern blotting and hybridization with the XpYp and 17p telomere-adjacent probes as described previously (2). In order to determine the chromosomes participating in the fusion events for subsequent sequence characterization, further hybridizations were undertaken with the 21q or 16p telomere-adjacent probes generated by PCR with 21qseq1/21qseq1rev or the 16p2/16pseq1rev PCR primers (Figure 1A), these probes yield additional non-specific products and thus were not used for quantification. Any fusion products were then re-amplified for direct sequence analysis using nested PCR primers (XpYpO, 17p7, 21qseq1 and 16p2, Figure 1A) as determined by Southern hybridization, as described previously (2).

To provide an estimate of the frequency of fusion, the total number of bands generated from each DNA analysed was counted and the number of input molecules in each analysis was estimated from the amount of DNA used in each reaction together with the size of a diploid human genome. Thus, the frequency of fusion could be calculated from the number of positive molecules compared to the number of input molecules. This will represent an underestimate of the total telomere fusion frequency as the fusion assays described here are capable of detecting a small fraction of the total number of different pair wise telomere fusion combinations in the human genome.

Oligonucleotides

XpYp and 17p primers used during this study have been described previously (2). The chromosome ends that previously hybridized to the TelBam11 probe (21q, 1q, 2q, 5q, 6q, 6p, 8p, 10q, 13q, 17q, 19p, 19q, 22q and the 2q13 interstitial telomeric locus) and to the TelBam3.4 probe (16p, 1p, 9p, 12p, 15q, XqYq and the 2q14
interstitial locus) probes (34,36) were termed 21q and 16p groups, respectively (Figure 1A). The primers for the telomere fusion assay were designed for both groups within the sub-telomeric regions of the 21q and 16p sequences (35,36) (Table 1).

**RESULTS**

**Telomere fusion assay**

We have previously developed a single-molecule PCR-based telomere fusion assay for two chromosome ends, XpYp and 17p. The sub-telomeric sequence of

---

**Figure 1.** Illustrating the single molecule telomere fusion assay. (A) Diagrammatic representation of oligonucleotide primers used for the PCR reaction; the XpYp and 17p primers are specific to those chromosome end; the 21q group primers are capable of amplifying the telomeres of 1q, 2q, 5q, 6q, 6p, 8p, 10q, 13q, 19p, 19q, 21q, 22q and 2q13; the 16p group amplifies the telomeres of 1p, 9p, 12p, 15q 16p and XqYq as well as the 2q14. (B) Examples of telomere fusion gels, showing the improvement in the sensitivity of the PCR reaction when the additional telomeres are included in the reaction, as indicated on the left. HEK293, and MRC5 E6E7 with population doublings (PD), are indicated above. Each reaction contains 100 ng of DNA, and the fusion products were detected with the XpYp telomere adjacent probe. The frequency at which fusion was detected is indicated below each analysis, molecular weight markers are indicated on the right.
these telomeres is sufficiently unique such that sequence-specific telomere length and fusion assays have been developed (2,37–39). However, the analysis of fusion between just two chromosome ends is limiting; increasing the scope of the assay to include more ends will improve the sensitivity and allow the detection of rare fusion events, which will be important in examining telomere fusion directly from human tissue in both normal and disease contexts. It will also allow the detection of a broader range of fusion structures and the identification of specific chromosome ends in the human genome that are subjected to fusion. We therefore chose to expand the scope of this assay by including PCR primers based within the complex subtelomeric sequences that are shared between multiple chromosome ends (40). Two subtelomeric sequence families were chosen for analysis; one family included the telomeres of 21q, 1q, 2q, 5q, 6q, 6p, 8p, 10q, 13q, 17q, 19p, 19q, 22q and the 2q13 interstitial locus representing some of the 2q, 5q, 6q, 6p, 8p, 10q, 13q, 17q, 19p, 19q, 22q and the 16p6 telomere groups, respectively.

A valuable aspect of these assays is the ability to characterize the DNA sequence of fusion events; this provides verification that the fusion events are genuine, and allows one to examine the underlying biology. However, fusion within the 21q or the 16p telomere family members is not practical to sequence these events directly. Thus, for the initial development of the assay we used PCR primers targeted to the 21q and 16p family in conjunction with the specific chromosome ends XpYp and 17p. To validate the assay, we used HEK293 cells which we had previously shown to exhibit fusion between the XpYp and 17p telomeres at a high frequency (2). HEK293 cells express telomerase but maintain telomere length distributions that extend to within the length range that we had previously detected fusion events (2). Combining the 17p, XpYp and 21q primers resulted in an increase (2.9×) in the number of fusions detected (Figure 1B); an increase in the number of fusion events was also observed when combining 17p, XpYp and 16p primers (data not shown). The combination of XpYp, 17p, 21q and 16p resulted in a further increase (4×) in the number of detectable fusion events (Figure 1B). Given the apparent success of these assays we then applied them to a clonal MRC5 fibroblast culture in which we had expressed HPV6E6E7 oncoproteins (2). We have previously described that, in this culture, as the cells continued to divide beyond the point of senescence the telomeres continue to erode to a threshold length at which fusion events were detected between the XpYp and 17p telomeres (2). Crucially, there was no detectable change in telomere dynamics (erosion rates or frequencies of sporadic deletion) following HPV6E6E7 expression (2). Using the 16p and 21q assays we observed a similar increase in the number of detectable fusion events with ongoing cell division through to the point of ‘crisis’, overall there was an increase in the number of detectable fusion events compared to that observed with the XpYp/17p fusion analysis alone (Figure 1B). The number of fusion events was so great in the MRC5 E6E7 cells undergoing ‘crisis’ that it appeared to interfere with the amplification creating some smearing on the gels, which was not apparent in samples with small numbers of fusions (Figure 1B).

The increase in the number of detectable fusion events with the modified assays indicated that we were detecting fusion between additional chromosome ends. In order to verify this and to examine the mutational spectrum, we undertook direct sequence analysis of single fusion events isolated from both HEK293 and MRC5 cells (Figure 2). The telomeres participating in each fusion event were identified by sequential Southern hybridization using telomere adjacent probes. Putative single fusion molecules were re-amplified with nested PCR primers and in order to identify the fusion points as well as the specific telomeres involved in each event, multiple sequencing reactions were undertaken for each fusion product. We sequenced a total of 298 fusion events; 125 events from HEK293 cells (of which n = 41 and n = 84 were from the 16p and 21q groups, respectively) and 173 events from MRC5 cells (of which n = 71 and n = 102 were from the 16p and 21q groups, respectively). Fusions were detected between either 17p or XpYp and several telomeres within the 21q and 16p families. This demonstrated that our assays were capable of detecting fusion between the multiple chromosome ends targeted in the assay.

### Table 1. Oligonucleotide sequences for fusion analysis

| Fusion PCR | 21q1: | 5’-CTTGGTTGTCGAGAGGTTGATG-3’ |
| 16p1: | 5’-TGGACTGTCCTTCTCAGGGT-3’ |
| Sequencing | 16pseq1: | 5’-TCGAGTTGAGCTAGGAGAAG-3’ |
| XqYqseq1: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 16pseq2: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 16pseq4: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 16pseqB: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 16pseq1: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 16pseq1rev: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 16pseq3: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 16pseq2rev: | 5’-TGGTTGTCGACAGACCTG-3’ |
| Subtel1: | 5’-TGGTTGTCGACAGACCTG-3’ |
| Subtel2: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 1pseqA: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 1pseq1: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 1pseq2: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 1pseq1rev: | 5’-TGGTTGTCGACAGACCTG-3’ |

### Fusion PCR

| 21q1: | 5’-CTTGGTTGTCGAGAGGTTGATG-3’ |
| 16p1: | 5’-TGGACTGTCCTTCTCAGGGT-3’ |
| Sequencing | 16pseq1: | 5’-TCGAGTTGAGCTAGGAGAAG-3’ |
| XqYqseq1: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 16pseq2: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 16pseq4: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 16pseqB: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 16pseq1: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 16pseq1rev: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 16pseq3: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 16pseq2rev: | 5’-TGGTTGTCGACAGACCTG-3’ |
| Subtel1: | 5’-TGGTTGTCGACAGACCTG-3’ |
| Subtel2: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 1pseqA: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 1pseq1: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 1pseq2: | 5’-TGGTTGTCGACAGACCTG-3’ |
| 1pseq1rev: | 5’-TGGTTGTCGACAGACCTG-3’ |

### Reference

1844 Nucleic Acids Research, 2010, Vol. 38, No. 6
Sub-telomeric deletion and microhomology

The sequence analysis revealed that all the fusion events involved the deletion of either one or both of the telomeres, these deletion events extended into the telomere-adjacent DNA. Using the 16p group assay, which extends the furthest from the start of the telomere, we detected fusion events close to the limit of the assay, 6.1 kb (Figures 2A and 3A). Both the MRC5 and HEK293 cells displayed similar deletion profiles \((P = 0.65)\) with combined mean deletion size of 2.443 kb ± 1.540 (SD) (Figure 3A). Similar deletion profiles were apparent at the other chromosome ends analysed in the 21q group, XpYp and 17p (Figure 3B–D). The highest frequency of fusion (mode) occurred within 0.2–2 kb of the telomeres; however these data were not normally distributed (Figures 3A–D, \(P < 0.0008\), Shapiro–Wilk normality test) and distal to the mode the...
frequency of fusion events appeared to remain constant per unit length. These profiles indicate that sub-telomeric deletion prior to the fusion may be more extensive than that detected here, extending up to and beyond the limit of our assays.

In addition to the characteristic deletion events that accompany telomere fusion, the majority of fusion events (60–74% in HEK293 and MRC5 respectively) exhibited short patches (1–12 nt, overall mean 1.7 nt) of 100% sequence homology (Figure 3E). Compared to the A:T/G:C content of the telomere-adjacent sequences studied here (48% A:T, 52% G:C) there was a significant bias towards G:C base pairs within the regions of microhomology at the fusion points (39% A:T, 61% G:C, \( P = 0.002 \) chi-square).

Telomere length and fusion

The frequency of fusion increased in the MRC5 E6E7 cells with ongoing cell division as the cells approached crisis (Figure 1B). Interestingly, these cells displayed fusion between either XpYp or 17p and just two of the nine telomeres within the 21q group (10q and 21q); this was in contrast to the HEK293 cells where fusion was detected involving seven different telomeres in the 21q group. This indicated that in MRC5 cells a specific subset of telomeres may have eroded to a length at which they became subjected to fusion, whereas the other chromosome ends within the 21q family had not.

The TTAGGG telomere repeat content at the fusion junctions was limited, with 38% to 60% of fusions in the HEK293 and MRC5 cells, respectively containing no telomere repeats (Figure 3F). Of the fusions that did contain TTAGGG repeats immediately adjacent to the fusion point, the mean length was 5.8 ± 0.7 (SE) repeats. Single examples of longer segments of contiguous TTAGGG repeats adjacent to the fusion point were also observed with the largest block of 43 repeats observed in MRC5 cells; further sampling may yield longer telomeres involved in fusion (Figure 3F). Just one fusion event of the total 298 sequenced revealed telomere repeats on each side of the fusion point, in this case 6.5 repeats from 17p and 3.5 repeats from 10q (Figure 2B). This was consistent with our previous observations, where the paucity of fusion events containing telomere repeats on either side of the fusion point was not accounted for by the resulting inverted repeats being refractory to PCR (2).
An example of duplication is indicated with an arrow, deletions and insertions are indicated with parenthesis. Gaps (−) of repeat units from the start of the telomere repeat array. The key indicates the lettering system for each type of telomere repeat variant identified.

The proximal 1–3 kb of human telomeres contains hyper-variable interspersion patterns of the canonical telomere repeat sequence TTAGGG with telomere repeat variants (37,43,44). Our large data set of single-molecule fusion sequences involving the XpYp and 17p telomeres, allowed us to examine in detail the telomere variant repeat (TVR) composition derived from single cells within the same cell strain. These regions contain numerous different types of telomere repeat variants interspersed with TTAGGG repeats, some of which are in contiguous blocks up to 16 repeats in length within the alleles described here (Figure 4 and Supplementary Figure S1). It was apparent from this analysis that the same allele sequenced from separate cells within the same cell strain, exhibited differences within the repeat distributions (Figure 4 and Supplementary Figure S1). These differences could be attributed to single base insertions or deletions, converting one pre-existing repeat to another, for example TTAGGG to TAGGG, and expansions or contractions in the size of contiguous repeat blocks which include duplications as well as larger deletions of up to 31 repeats (168 bp) (Figure 4). Many of the repeat changes were detected in more than one independent fusion event and must have arisen prior to any fusion event. Thus, these mutation events have not arisen as a consequence of the fusion event itself and therefore may reflect ongoing mutational processes within these regions of the telomere.

The TVR data also revealed a further aspect of mutation within telomere repeat arrays. We documented 26 different telomere variant repeat types at the XpYp and 17p telomeres in MRC5 and HEK293 cells. In addition, we observed that the same allele sequenced from separate cells within the same cell strain, exhibited differences within the repeat distributions (Figure 4 and Supplementary Figure S1). These differences could be attributed to single base insertions or deletions, converting one pre-existing repeat to another, for example TTAGGG to TAGGG, and expansions or contractions in the size of contiguous repeat blocks which include duplications as well as larger deletions of up to 31 repeats (168 bp) (Figure 4). Many of the repeat changes were detected in more than one independent fusion event and must have arisen prior to any fusion event. Thus, these mutation events have not arisen as a consequence of the fusion event itself and therefore may reflect ongoing mutational processes within these regions of the telomere. Interestingly, we also observed two fusion events which

Figure 4. Examples of the telomere variant repeat structure derived from sequence analysis of the telomeres adjacent to single telomere fusion events. The telomere codes run left to right (proximal to distal), the fusion point is at the right hand of each code, the numbering above indicates the number of repeat units from the start of the telomere repeat array. The key indicates the lettering system for each type of telomere repeat variant identified. Gaps (−) in the codes were introduced to improve the alignment, these gaps therefore indicate putative expansions or contractions of repeat blocks. An example of duplication is indicated with an arrow, deletions and insertions are indicated with parenthesis.
contained blocks of repeat variants, without an obvious origin, immediately adjacent to the fusion point (Figure 4).

Complex fusion events

We observed fusion between both the 17p and XpYp telomeres and the interstitial ancestral telomere–telomere fusion at 2q14, this locus is approximately 130 Mb from the 2q telomere (Figure 2C) (42). It could not be determined whether this represented fusion following a large-scale subtelomeric deletion process initiated from the 2q telomere, or fusion directly with a de novo DNA double-strand break that occurred near this locus; however, it was clear that dysfunctional telomeres are capable of fusion to non-telomeric loci. Indeed, we detected fusion events (34/298, 11%) that contained additional genomic sequences inserted between the fused telomeres (Figure 2D). The genomic locus of four insertions could be identified; one was derived from 8q24.3, one from 7p21.1 and two from Xp22.1. Interestingly, these loci, as well as 2q14, are all close to fragile chromosomal sites (45,46). Two of the insertion events contained sequences from the originating telomere inserted in the reverse orientation. In MRC5 cells where the XpYp telomeric alleles could be identified by numerous SNPs in the telomere adjacent DNA (2,37), it was clear that these inverted insertions were derived from the same allele involved in the fusion event, and were therefore, intra-allelic (Figure 2E). The remaining 28 insertion events were too small to identify the originating chromosomal locus (Figure 2F).

We also detected fusion events between the XpYp and XqYq telomeres in 23 out of the 112 fusion events sequenced involving the 16p group of telomeres (21%, Figures 2A and 2F); such events may result in the formation of ring chromosomes.

DISCUSSION

Here, we have described a comprehensive analysis of telomere fusion between multiple chromosome ends in human cells. These assays can detect sister chromatid-type fusion when there has been a deletion of at least one of the telomere repeat arrays creating an imperfect inverted repeat, they can also detect fusion between heterologous chromosomes containing substantial arrays of head-to-head telomere repeats that arise following TRF2 depletion (2). They cannot however, detect the presence of sister chromatid-type fusions that create perfect inverted repeat sequences (2). Interestingly, studies using telomere-adjacent I-SceI induced breaks, show that the majority (90%) of selected clones had undergone chromosomal ‘healing’ by the addition of new telomere (47). Of those clones that had not undergone chromosomal healing, inverted repeats close to the site of the original break was the most common event; this is consistent with sister chromatid fusion (48). The majority of events that are detected with our assays involved heterologous chromosomes, with 5% of events that were consistent with sister chromatid fusion (2). The difference between these two datasets may be related to the fact that our data is examining telomere fusion in the context of widespread telomeric dysfunction across multiple chromosomes, compared to the selection of events involving breakage at a single chromosome end in the context of otherwise functional telomeres at the other chromosome ends.

It is clear from our previous work (2), as well as that from other labs (48,49), together with the large data set described here, that the fusion of short dysfunctional telomeres is accompanied by telomeric and sub-telomeric deletion. This profile is also consistent with studies using I-SceI induced double-strand breaks where larger deletions were observed at telomeric loci (49). The size profile of these deletion events indicates that deletion is likely to extend beyond the proximal limit of our assays, which at the 16p group is 6.1 kb. Whilst the mechanism of deletion has not been characterized, studies using conditional cdc13 mutants in Saccharomyces cerevisiae show that the exo1 dependent single-stranded DNA can extend for at least 8 kb from the start of the telomere (50). Furthermore, breakpoint frequency within the terminal chromosomal bands observed in pancreatic carcinoma and osteosarcoma has been shown to negatively correlate with telomere length at specific chromosome ends (15). This data was determined from karyotype analysis and therefore the distance of the breakpoints within the terminal bands may be several mega base pairs from the chromosomal terminus. Thus, we consider that the subtelomeric deletion that follows the formation of short dysfunctional telomeres could be extensive, and may include terminal coding regions. Several genes have been identified close to the telomere, for example members of the Wiskott–Aldrich Syndrome Protein and Scar homolog (WASH) gene family lie within 5 kb from the start of the telomere repeat array (51). Therefore, irrespective of whether they undergo fusion, or healing by de novo telomere formation, the extensive deletion events that occur at short dysfunctional telomeres are likely to create a significant mutational burden. Indeed it is clear that terminal chromosomal deletions have been implicated in numerous clinical genetic conditions (52,53).

The extensive deletion that accompanies telomere fusion together with microhomology at the fusion junction may be indicative of the DNA repair processes that mediate fusion between short dysfunctional telomeres. The limited homology at the fusion junction, together with the extensive deletion, appear inconsistent with single-stranded annealing or NHEJ as being the predominant mechanisms underlying fusion; both these mechanisms have, in different contexts, been described to mediate telomere fusion (23,54). An alternative NHEJ pathway has been described in mammalian cells (55,56) that is DNA-PK independent. This pathway is considered to be error-prone resulting in extensive deletion (55,57) and is facilitated by microhomology at the fusion point with a bias toward G:C base pairs (58). The profile that accompanies DNA-PK-independent NHEJ is reminiscent of the extensive deletion and microhomology with a G:C bias described here following the fusion of short dysfunctional telomeres. This alternative NHEJ pathway
may account for double-strand break repair in the absence of functional classical NHEJ in some tumours (59), V(D)J recombination and IgH class switching in mouse models (60,61) and thus is physiologically relevant. This pathway is poorly characterized and the involvement, if any, of the key proteins required in this pathway such as PARP-1 and Ligase III (62–64), in telomere fusion, has yet to be established. It will be of interest to ascertain which pathway mediates the fusion between short dysfunctional telomeres.

We observed fusion events in which there was an insertion of additional telomere-adjacent DNA, which was derived from one of the two telomeres involved in the fusion event. Informative SNPs allowed us to establish that the inserted DNA was derived from the same telomeric allele involved in the fusion event and it was therefore an intra-allelic event. Furthermore, in all the examples of this type of fusion, the DNA was inserted in the reverse orientation with respect to the telomere repeat array. These observations lead us to consider that the fusion event involved an intra-molecular rearrangement prior to the final fusion between the two chromosomes. We propose the following sequence of events as illustrated in Figure 5, as a speculative model to account for both the presence of multiple fusion points within the same single molecule and the reverse orientation of the insertion. In this model short telomeres lose their end protection function, possibly because of a lack of TRF1 and

---

**Figure 5.** Illustrating a model, for the sequence of events, that may account for intra-allelic inverted insertion that accompanies a subset of telomere fusion events. This model is using the fusion sequence depicted in Figure 2F as an example, the deletion points are indicated as distances from the start of the telomere repeat array. (A–B) Short telomeres become de-protected and are subjected to nucleolytic resection, yielding a 3’-overhang within the telomere-adjacent DNA. (C) The 3’-end of the overhang (−111) folds back upon itself and anneals adjacent to the −596 position, this is ligated and endonucleolytic cleavage creates a 5’-overhang structure which can be rendered double-stranded by strand extension (D). Further resection (E) to the −356 position creates a substrate (F) for annealing and ligation of the 1p sub-telomere at position −1260 (G). Endonucleolytic cleavage removes the unpaired 3’-overhang from XpYp (G), strand extension and ligation complete the fusion (H–I).
DNA (Figure 5A–B). The 3'-end of this single-stranded DNA folds back upon itself and anneals at points of microhomology that are sufficiently stable to facilitate ligation (Figure 5C). The intra-molecular loop is then processed by endonucleolytic cleavage to create a 5'-overhang with the inversion and one fusion junction in place (Figure 5D). The 5'-overhang is then rendered double-stranded by strand extension, and subjected to further 5'-3' resection to create a substrate for fusion (Figure 5E–F) with another telomere or double-stranded DNA break (Figure 5G–I). An alternative mechanism as elaborated previously (20,21,49) could involve an initial sister chromatid fusion. This creates an imperfect inverted repeat which forms a fragile site that breaks following anaphase bridging close to the site of the first fusion, this broken end now contains an inversion which then fuses to the second chromosome involved in the fusion event.

We observed numerous examples of fusion between XpYp and XqYq; such events have the potential to create circular chromosomes. These ring chromosomes can occur constitutively, resulting in variable clinical consequences which is not consistent with a loss of genetic material in the originating ring chromosome (65,66). Some examples of ring chromosomes, such as ring chromosome 19, indicate they have formed as a consequence of telomere fusion without any cytogenetically detectable terminal deletion (67). Thus, it has been postulated that ‘ring syndromes’ result from the subsequent mitotic instability of the ring chromosome (66). Ring chromosomes are also frequently detected in tumour cells, and have been shown to be highly unstable and undergo continuous rearrangements (68); resulting in potentially oncogenic amplification and deletion (69). Interestingly, some tumour types for example liposarcomas and parosteal osteosarcoma exhibit ring chromosomes in over 90% of tumours (70–72). Whereas other cancer types, including many types of carcinoma, the frequency of ring chromosomes is much less <10% of cases (72).

Extensive variation in the human population in the distribution patterns of telomere repeat variants that occur in the proximal 2–3 kb of telomere repeat arrays has been described (37,44). Differences in the TVR distributions of related telomeric alleles indicate a pattern of intra-allelic mutation, which is remarkably similar to that as described here in the TVR distributions obtained from single cells of the same cell strain. The patterns of mutation, including repeat expansion and contraction, and mutation within the telomere repeats themselves, are consistent with errors during replication such as stuttering of the replication fork. Consistent with this, telomere repeat arrays, as well as other repetitive loci, present a problem for the replication fork, and as such, require additional processing to facilitate progression (73–75).

Analysis of the TVR distributions also revealed the presence of two fusions with additional repeat variants immediately adjacent to the fusion point; these did not appear to be derived from the progenitor allele. We considered that this may be related to the process of fusion, with the possibility of processing of a heteroduplex at the fusion point creating a scrambled repeat structure; however, this appeared unlikely as the general repeat structure was conserved and the DNA sequence of the participating chromosome did not contain the correct bases. Instead, we consider that these inserted repeats may represent an insertion of additional DNA derived from another telomere during the fusion event.

All the fusion events that involved non-telomeric loci, 2q14, 8q24.3, 7p21.1 and Xp22.1, where identified within the same cytogenetic location as previously documented fragile sites (45,46). In contrast to normal cells, fragile sites are involved in chromosomal rearrangements in tumour cells, these are typically deletions, but translocations involving fragile sites have also been identified (76–78). The fusion of short dysfunctional telomeres with fragile sites may thus provide one mechanism to generate large-scale genomic rearrangements in cancer cells.

Whatever the underlying mechanisms that generate the fusion structures described here, it is clear from this data, that short dysfunctional telomeres can be subjected to considerable processing prior to fusion, creating complex structures and substantial deletion.

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS
We thank Harold Riethman (Wistar institute) for providing telomere-adjacent sequence information. We also thank the staff at the sequencing facility of the Cardiff University Central Biotechnology Service.

FUNDING
Cancer Research UK (grant no. C1799/A6932; C1799/A5603). Funding for open access charge: Cancer Research UK (grant no. C1799/A6932).

Conflict if interest statement. None declared.

REFERENCES


