

Are Head to Head Interstitial Telomeric Repeats Unique to Human Chromosome 2q13?

A Genome Wide BLASTn Reassessment of a Central Claim in the Chromosome 2 Fusion Argument

Donny Budinsky¹

Independent Researcher

ABSTRACT

The interpretation of human chromosome 2 as the product of an ancestral telomere to telomere fusion draws in part on a block of interstitial telomeric repeats at 2q13 that are arranged head to head, meaning that a forward oriented array of the hexamer TTAGGG meets a reverse oriented array of its complement CCCTAA at an inversion junction. A claim frequently made in popular presentations holds that this head to head arrangement occurs nowhere else in the human genome, and that its presence at 2q13 marks the fusion point. We tested that uniqueness claim directly. Before the test we had made a specific prediction: if interstitial telomeric repeats are a widespread and functional feature of the genome rather than fusion scars, then an inverted, head to head array should be found elsewhere, at a locus where no fusion has ever been proposed. Using a genome wide BLASTn survey we report inverted telomeric arrays on multiple chromosomes, and, most significantly, a head to head array on chromosome 9 whose junction is less interrupted than the one at 2q13. Because chromosome 9 is not proposed to be a fusion product, a cleaner inverted array there shows that the arrangement is not, by itself, diagnostic of a fusion. We first clarify a technical point essential to any test of the claim: because CCCTAA is the reverse complement of TTAGGG, both motifs are present at every telomeric locus, so their co occurrence is automatic and carries no information, and only a genuine inverted arrangement on a single strand is distinctive. The two regions are shown side by side, rendered in one font and colored by one criterion, so that the reader can compare them directly.

1. Introduction

Humans carry twenty three pairs of chromosomes, whereas chimpanzees, gorillas, and orangutans carry twenty four. A widely held interpretation attributes this difference to an ancestral end to end fusion of two chromosomes into what is now human chromosome 2, the second largest chromosome in the genome. The foundational sequence based support for this interpretation was the identification of a block of telomeric repeats near the middle of chromosome 2, at 2q13, presented as the remnant of two telomeres that fused (IJdo et al. 1991). A second site at 2q21 has been advanced as a degenerate, inactivated centromere, and the banding correspondence between chromosome 2 and the two great ape chromosomes designated 2A and 2B has been offered as further support.

Interstitial telomeric sequences, meaning tandem telomeric hexamer repeats located away from chromosome ends, are not rare. They are widespread across vertebrate genomes and have been described in the human genome in several classes, including short interstitial repeats, subtelomeric repeats, and the fusion associated repeats at 2q13 (Meyne et al. 1990; Azzalin et al. 2001; Ruiz-Herrera et al. 2008; Bolzan 2017). They arise through several mechanisms, including ancient rearrangements, inversions, and the repair of double strand breaks, and a number of them coincide with functional elements. The fusion site itself and paralogous regions on other chromosomes were characterized in detail by Fan and colleagues (2002).

Within this setting a specific claim is often made in popular presentations of the fusion argument, namely that the head to head arrangement of telomeric repeats, in which a forward array meets a reverse array, is found only at 2q13 (see, for example, the video referenced as note 1). The present paper tests that uniqueness claim. If interstitial telomeric repeats are in fact a widespread and functional feature of the genome, then inverted arrangements of them need not be confined to a single site, and their presence at 2q13 would not by itself establish a fusion.

2. Interstitial telomeric sequences and the head to head criterion

Any test of the uniqueness claim must rest on a precise criterion, because a loose one is satisfied trivially. The telomeric hexamer is TTAGGG, and its reverse complement is CCCTAA. On double stranded DNA these are two readings of the same sequence: wherever a telomeric array occurs, one strand reads TTAGGG and the other reads CCCTAA. Detecting both motifs at a locus is therefore automatic and conveys no information about arrangement. Figure 1 makes the relevant distinction explicit.

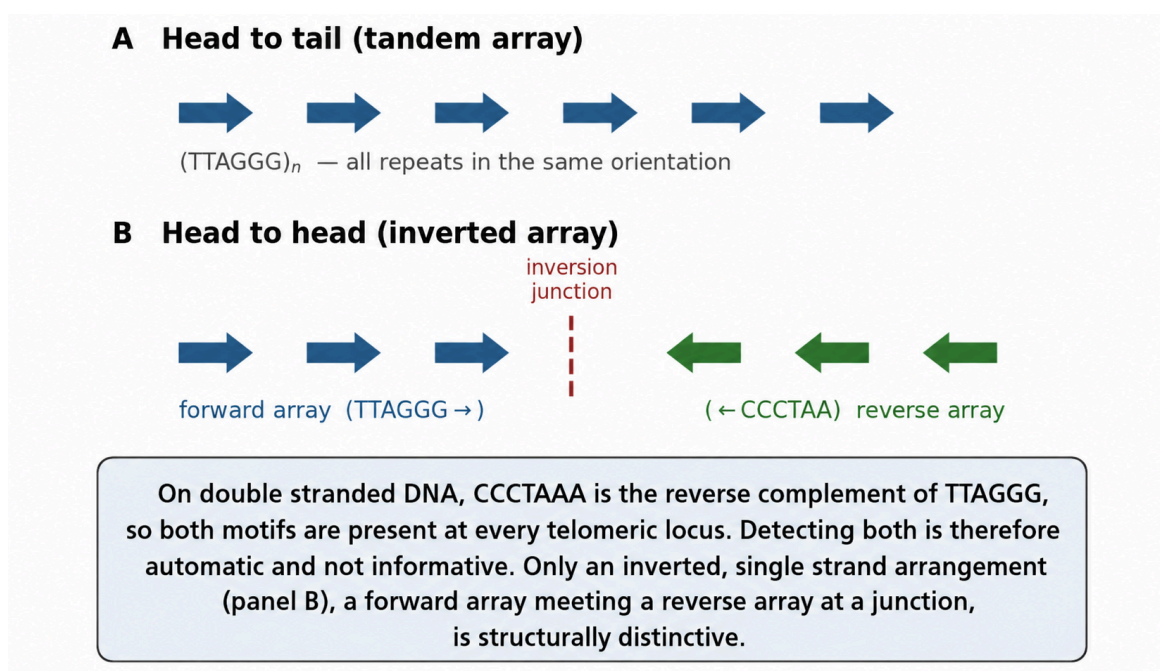


Figure 1. Head to tail and head to head arrangements of telomeric repeats. In a head to tail array (A) all repeats share one orientation. In a head to head array (B) a forward oriented array meets a reverse oriented array at an inversion junction. Because CCCTAA is the reverse complement of TTAGGG, the two motifs co-occur at every telomeric locus; only the inverted single strand arrangement in panel B is structurally distinctive and diagnostically relevant.

The arrangement that the fusion interpretation treats as meaningful is the inverted one in panel B, since two telomeres that fuse end to end bring their outward pointing arrays into a head to head orientation. It is this arrangement, not the mere presence of the two reverse complementary motifs, that a survey of the genome must locate in order to bear on the uniqueness claim. In the specialist literature, 2q13 has been classified as the single fusion type interstitial telomeric site in the human genome, with the other human classes being short and subtelomeric repeats (Azzalin et al. 2001; Ruiz-Herrera et al. 2008). The uniqueness claim is therefore not merely a popular assertion but reflects a published classification, and any finding of equivalent inverted arrangements elsewhere is properly framed as a proposal to revise that classification, which raises the standard of documentation required.

3. The prediction we set out to test

The claim that head to head telomeric repeats are unique to 2q13 is not decorative to the fusion argument; on the popular version it is load bearing. If the inverted arrangement occurred at only one place in three billion bases, that singularity would be striking and would invite a special explanation. We took the claim seriously enough to state, in advance, what our own model predicts against it. If interstitial telomeric repeats are functional units of the genome that arise by ordinary mechanisms, then an inverted, head to head array is not the fingerprint of a fusion and should be recoverable at loci where no fusion has ever been proposed. Two of us (Budinsky and Rupe) committed to this prediction publicly, in debate and in podcast form, before the survey was run. It was a novel prediction and it had not yet been tested.

The prediction is deliberately sharper than the familiar observation that telomeric repeats are scattered through the genome. Scattered single strand repeats, forward here and reverse there, are already well documented and are not in dispute. What we predicted was the specific configuration that the fusion interpretation treats as its signature: a forward oriented array meeting a reverse oriented array at a junction, on a chromosome that no one claims to be a fusion product. If such an array exists, and especially if its junction is cleaner than the one at 2q13, then the arrangement at 2q13 cannot carry the evidential weight that has been placed on it. That is the result we report below.

4. Materials and methods

A genome wide similarity search was performed using BLASTn with the telomeric hexamer query and its reverse complement against the most recent builds of the human reference assembly. The 798 base core region surrounding the 2q13 site was also used as a query. Candidate interstitial sites were examined for the number of repeat units, the orientation of adjacent arrays, and the presence or absence of an inversion junction, in order to separate genuine head to head arrangements from the automatic co occurrence of the two reverse complementary motifs and from short or subtelomeric repeats. The analysis was carried out by Christopher Rupe.

The two sequence figures in this paper (Figures 2 and 3) were prepared under a single explicit rule so that they can be compared fairly. Every exact TTAGGG hexamer is colored red and every exact CCCTAA hexamer is colored blue; all other bases are left black. This is the same intact motif criterion by which the 2q13 site is conventionally described, and it reproduces the counts reported for that site, namely ten intact TTAGGG hexamers and forty three intact CCCTAA hexamers. Both figures are rendered in one monospace font at one scale, and the head to head junction in each is marked in yellow. No partial or single mismatch hexamers are colored in either figure; colored bases are exact matches only. This keeps the density that the reader sees tied to a reproducible rule rather than to visual impression.

5. Results

The survey returns interstitial telomeric repeats on multiple chromosomes and identifies sites at which the repeats occur in an inverted, head to head configuration comparable to that at 2q13. The chromosomes on which interstitial telomeric repeats were recovered are listed in Table 1. Most significantly, a site on chromosome 9 (GRCh38.p14 Primary Assembly, in the interval near 128,154,271 to 128,154,325) displays a head to head arrangement in which the junction region is less interrupted than the one at 2q13. Chromosome 9 is much smaller than chromosome 2 and is not proposed by anyone to be a fusion product, yet its inverted array meets at a cleaner junction. The BLASTn alignment for this site returns a high identity match with only a small internal gap, and the forward and reverse arrays meet directly, with fewer intervening bases than are seen at 2q13.

Reported interstitial telomeric repeats	Documentation required for verification
Chromosomes 1, 2, 3, 6, 9, 10, 15, 16, 18, 20, 21	Per site coordinates, array orientation, repeat counts, junction gaps, and expectation values
Chromosome 9 (reported head to head site)	Confirmation of an inversion junction and comparison of junction continuity with 2q13

Table 1. Loci at which interstitial telomeric repeats are recovered, and the positional documentation to be supplied in the supplementary material so that each can be independently verified as a genuine head to head arrangement rather than a co-occurrence of the two reverse complementary motifs.

Figures 2 and 3 place the two regions side by side under the single coloring rule described in Section 4. Figure 2 is the 798 base core region at 2q13. Figure 3 is the chromosome 9 window that contains the recovered head to head array.

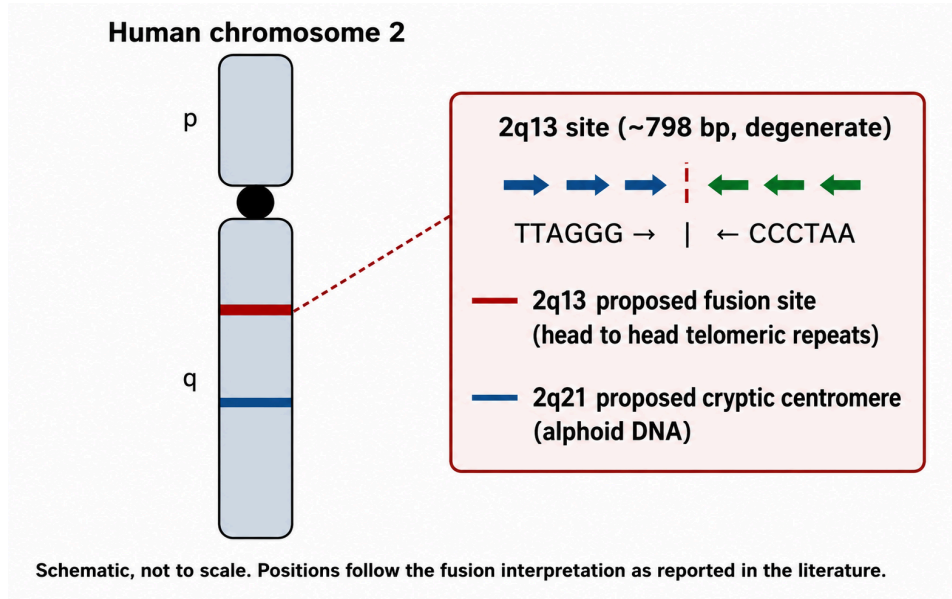


Figure 2. Schematic map of human chromosome 2, not to scale, showing the proposed fusion site at 2q13 and the proposed cryptic centromere at 2q21 as reported in the literature, with an inset of the head to head telomeric arrangement at 2q13. The present paper addresses only the 2q13 telomeric repeat site; the 2q21 site is shown for orientation and is not evaluated here.

TGAGGGTGAGGGTTAGGGTTTGGGTTGGGTTTGGGGTTGGGGTT
GGGGTAGGGTTGGGTTGGGTTGGGTTGGGGTTAGGGTAGG
GGTAGGGTAGGGTAGGGTCAGGGTCAGGGTCAGGGTTAGGGT
TTAGGGTTAGGATTTAGGGTTAGGGTAAGGGTTAAGGGTTGG
GGTTGGGGTTAGGGTTAGGGTTAGGGTTGGGGTTGGGGTTGGG
GTTGGGGTTGGGGTTGGGGTTAGGGTTAGCTAAACCTAACCTA
ACCCCTAACCCCAACCCCAACCCCAACCCCTACCCCTACCCCTACC
CCTAACCCCAACCCCAACCCCTTAACCCCTTAACCCCTAACCTAAC
CCTAACCCCAACCTAACCTAACCTAACCTAACCTAACCCAAACCTA
ACCCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCT
GACCTTGACCCTGACCCTTAACCCCTTAACCTAACCTAACCTAACCT

Figure 3. The 2q13 core region. Exact TTAGGG hexamers are red and exact CCCTAA hexamers are blue; the head to head junction is highlighted in yellow. The forward oriented telomeric block gives way to a reverse oriented block, the arrangement that the fusion interpretation reads as a fused pair of telomeres. Ten TTAGGG and forty three CCCTAA hexamers are intact; the remainder of the block is degenerate.

```

AATTCCCCAGGACTCAGAGGGCTCCAGGCCTGCACTTATCACTT
GGCTGCCTGTGTCAGGCCTGACAAGCCCCTGGTCCCCCACTGGC
CGCAGAGGCAACCTCTCCAGGTCCGCATCATGCACAGGCACTCC
CTCCCCAGCCCCTTCTCCCCGTGCAAGTTCAGAGGCCAGAGTGG
GTGGGGTTCTGAGGAAACCCCATGGACAGACACACCACCCGTT
CACCTCGGCTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGG
TTAACCCTAACCCCTAACCCCTCGTTCACCCTCGGCTGGGCTAAG
GAGAGATGCTGGCACTGTGAACCAGCCCCAGCCTCGGACCCCAA
GGAGCAGGACAGGCCTGAGACCCAGGCCACCAAATGCTGCCAA
AGCCTCGGAAGTCTGGTCCAGACTCCCGGAAGAGGCCAGCTTGG
TGCAGCCCCAGCTTTGAGAGAGGGATGGGACTGTTCCC

```

Figure 4. A chromosome 9 window (GRCh38.p14), colored by the identical rule and rendered in the identical font. A forward oriented TTAGGG array meets a reverse oriented CCCTAA array at a single junction, marked in yellow, within ordinary genomic sequence and at a locus where no fusion is proposed. The inverted arrangement, the feature treated as diagnostic at 2q13, is present here with a junction that is less interrupted. Credit: Christopher Rupe.

The comparison is the point of the paper. At 2q13 the inverted arrangement sits inside a long, degenerate telomeric block; on chromosome 9 the same inverted arrangement appears in otherwise ordinary DNA, with a junction that is at least as clean. If the inverted arrangement were the distinctive mark of a fusion, it should not appear, in cleaner form, at a locus with no fusion history. Its appearance there is what our prediction anticipated.

6. The density and degeneracy arguments

Once the uniqueness of the arrangement is set aside, the fusion argument tends to fall back on the density of telomeric motifs at 2q13, and this is sometimes described as half of the case. We think the density point is real but secondary, and we want to state it carefully rather than overstate it. Density is measured per unit length, not as an absolute count, so the fact that chromosome 2 is the largest chromosome does not by itself dissolve a density argument. What bears on it is whether the local concentration is better explained as a functional element than as a fusion scar. Among interstitial telomeric sites the 2q13 block does carry the highest local concentration; but this is a relative statement, and on the reading we favor, in which these are functional units, a hotspot of genomic activity is exactly where a higher concentration would be expected. The density observation is therefore compatible with function and is not, on its own, decisive for fusion.

The size and degeneracy of the site cut the same way. The 2q13 block spans only about 798 bases, and it is exceptionally degenerate: only ten TTAGGG and forty three CCCTAA hexamers are intact, as Figure 2 shows, in a region where a fusion of two once intact telomeres might be expected to leave a much larger and cleaner signature, on the order of ten thousand to thirty thousand bases. We should be fair about the standard reply, which is that an ancient fusion is expected to decay, so that degeneracy is also predicted by the fusion model. That reply has force, and the disagreement turns on the assumed mutation rate and elapsed time. This is why we do not rest the argument on degeneracy alone. The strongest form of our case is the distribution and arrangement of the repeats, documented site by site, not the size or the decay of any single block.

7. The 2q13 region is functional

A further consideration weakens the inference from the mere presence of telomeric motifs at 2q13. The region overlaps a long noncoding RNA gene, DDX11L2, and shows clear evidence of function. On the reverse complement strand the site behaves as a promoter, an internal genetic switch within the first intron of the gene. A body of biochemical data associates the segment with the binding of numerous transcription factors, including RNA polymerase II, and transcription initiates within the region (Tomkins 2013; Tomkins 2018). Public expression databases connect DDX11L2 to a large number of cell and tissue types and to co-expression networks involving genes tied to cell signaling and blood cell production.

Functionality and a fusion origin are not strictly mutually exclusive, since a fused region could in principle acquire or retain function, so the presence of a functional element does not by itself settle the question of origin. We take it, however, as a strong consideration against reading the site as an accidental fusion scar. Long noncoding RNA genes are increasingly recognized as central to genome regulation (Khalil et al. 2009; Quinn and Chang 2016), and they are expressed at lower levels than protein coding genes, so their significance cannot be judged by comparison with protein coding expression alone (Kornienko et al. 2016). Read in that light, the concentration of activity at 2q13 is more naturally understood as a working part of a designed gene than as the residue of a catastrophic fusion.

8. Two lines of evidence not evaluated here

This paper addresses one line of evidence, the interstitial telomeric repeats. Two further lines commonly advanced for the fusion are not evaluated here, and we state this plainly so that the scope of our claim is not mistaken. The first is the proposed cryptic centromere at 2q21, associated with alphoid satellite sequences (Avarello et al. 1992; Baldini et al. 1993). The second is the synteny between chromosome 2 and the great ape chromosomes 2A and 2B. Both lines have been contested in the creationist literature (Bergman and Tomkins 2011; Tomkins 2013) and defended in the mainstream literature, including recent analyses that draw on complete, telomere to telomere assemblies (Ventura et al. 2012). Because the present analysis does not treat them, this paper does not by itself resolve those topics; it bears specifically on whether the head to head telomeric arrangement is unique to 2q13, and it shows that it is not.

9. Discussion

If the inverted arrangements reported here are confirmed by full positional documentation, the head to head configuration is not confined to 2q13, and its presence there cannot by itself mark a fusion. This conclusion is consistent with the broader biology of interstitial telomeric repeats, which are widespread, arise through several mechanisms, and frequently coincide with functional elements. On that view the telomeric motifs at 2q13 need not be the residue of a fusion; they may represent a functional sequence, and the argument from their mere presence is correspondingly weakened.

We want to be candid about what the chromosome 9 result does and does not show, because a careful reader, and a hostile one, will press exactly here. It shows that the inverted arrangement is not unique and can occur, in cleaner form, without a fusion. It does not show that chromosome 9 is a fusion; no one claims that, and we do not.

Nor does a single cleaner junction, on its own, overturn the wider fusion interpretation, which also appeals to the 2q21 site and to whole chromosome synteny. What it removes is the force of the specific claim that the head to head array at 2q13 is a one of a kind signature. That claim, stated as an absolute, is false, and the figure in Section 5 shows why.

The evidence assembled here fits the prediction we made in advance. We expected that interstitial telomeric repeats would be widespread, that inverted arrangements of them would be found beyond 2q13, and that at least one such arrangement would sit at a locus, cleaner than 2q13, where a fusion event never took place. Each part of that expectation is met by the survey. On the model we favor, similarity of arrangement across chromosomes reflects shared functional design rather than a unique historical accident.

10. Conclusion

Interstitial telomeric repeats are a widespread feature of the human genome, and the presence of telomeric motifs at an internal chromosomal position is not, by itself, diagnostic of chromosomal fusion. More importantly, the particular **head-to-head orientation** treated as the distinctive signature of the proposed 2q13 fusion is not confined to chromosome 2. The genome-wide survey reported here recovered the same fundamental arrangement beyond 2q13, most notably on chromosome 9, where a forward-oriented TTAGGG array meets a reverse-oriented CCCTAA array at a comparatively clean junction in a region for which no chromosomal fusion is proposed. This result directly satisfies the prediction stated before the search was stated before the search was performed: if inverted interstitial telomeric arrays can arise as ordinary genomic features, then the configuration attributed special historical significance at 2q13 should also occur elsewhere. It does. The consequence is narrow but important: **head-to-head telomeric orientation cannot legitimately be treated as a unique or independently diagnostic signature of the proposed chromosome 2 fusion.**

That distinction changes the evidential question. A sequence configuration cannot derive its historical force merely from being described as a fusion signature once the same configuration is demonstrated at a locus where no corresponding fusion event is invoked. The burden therefore shifts from the existence of the head-to-head array to showing why the 2q13 array specifically requires a fusion origin while comparable inverted arrays elsewhere do not. Density, degeneracy, genomic context, functionality, the proposed 2q21 cryptic centromere, and chromosome-wide synteny may still be brought into that larger argument, and the present study does not claim to adjudicate all of them. What it does establish is more precise: **one of the most visually compelling and frequently repeated claims in the chromosome 2 fusion argument, the supposed uniqueness of the head-to-head telomeric arrangement, cannot stand as stated.** The chromosome 9 result supplies the counterexample. Once uniqueness is removed, the 2q13 sequence must be interpreted from the totality of the evidence rather than treated as a self-authenticating molecular scar of fusion.

References

- Avarello, R., A. Pedicini, A. Caiulo, O. Zuffardi, and M. Fraccaro. 1992. Evidence for an Ancestral Alphoid Domain on the Long Arm of Human Chromosome 2. *Human Genetics* 89: 247–249.
- Azzalin, C. M., S. G. Nergadze, and E. Giulotto. 2001. Human Intrachromosomal Telomeric-like Repeats: Sequence Organization and Mechanisms of Origin. *Chromosoma* 110: 75–82.
- Baldini, A., T. Ried, V. Shridhar, et al. 1993. An Alphoid DNA Sequence Conserved in All Human and Great Ape Chromosomes: Evidence for Ancient Centromeric Sequences at Human Chromosomal Regions 2q21 and 9q13. *Human Genetics* 90: 577–583.
- Bergman, J., and J. Tomkins. 2011. The Chromosome 2 Fusion Model of Human Evolution, Part 1 and Part 2. *Journal of Creation* 25 (2): 106–116.
- Bolzan, A. D. 2017. Interstitial Telomeric Sequences in Vertebrate Chromosomes: Origin, Function, Instability and Evolution. *Mutation Research, Reviews in Mutation Research* 773: 51–65.
- Fan, Y., E. Linardopoulou, C. Friedman, E. Williams, and B. J. Trask. 2002. Genomic Structure and Evolution of the Ancestral Chromosome Fusion Site in 2q13–2q14.1 and Paralogous Regions on Other Human Chromosomes. *Genome Research* 12: 1651–1662. <https://doi.org/10.1101/gr.337602>
- IJdo, J. W., A. Baldini, D. C. Ward, S. T. Reeders, and R. A. Wells. 1991. Origin of Human Chromosome 2: An Ancestral Telomere-Telomere Fusion. *Proceedings of the National Academy of Sciences* 88: 9051–9055.
- Khalil, A. M., et al. 2009. Many Human Large Intergenic Noncoding RNAs Associate with Chromatin-Modifying Complexes and Affect Gene Expression. *Proceedings of the National Academy of Sciences* 106 (28): 11667–11672. <https://doi.org/10.1073/pnas.0904715106>
- Kornienko, A. E., et al. 2016. Long Non-coding RNAs Display Higher Natural Expression Variation than Protein-coding Genes in Healthy Humans. *Genome Biology* 17: 14.
- Meyne, J., R. L. Ratliff, and R. K. Moyzis. 1990. Conservation of the Human Telomere Sequence (TTAGGG)_n among Vertebrates. *Proceedings of the National Academy of Sciences* 86: 7049–7053.
- Quinn, J. J., and H. Y. Chang. 2016. Unique Features of Long Non-coding RNA Biogenesis and Function. *Nature Reviews Genetics* 17 (1): 47–62. <https://doi.org/10.1038/nrg.2015.10>
- Ruiz-Herrera, A., S. G. Nergadze, M. Santagostino, and E. Giulotto. 2008. Telomeric Repeats Far from the Ends: Mechanisms of Origin and Role in Evolution. *Cytogenetic and Genome Research* 122: 219–228.
- Rupe, C., and J. Sanford. 2017. *Contested Bones*. FMS Publications, pp. 326–328.
- Shammas, M. A. 2011. Telomeres, Lifestyle, Cancer, and Aging. *Current Opinion in Clinical Nutrition and Metabolic Care* 14 (1): 28–34.
- Tomkins, J. P. 2013. Alleged Human Chromosome 2 Fusion Site Encodes an Active DNA Binding Domain inside a Complex and Highly Expressed Gene, Negating Fusion. *Answers Research Journal* 6: 367–375.
- Tomkins, J. P. 2018. Combinatorial Genomic Data Refute the Human Chromosome 2 Evolutionary Fusion and Build a Model of Functional Design for Interstitial Telomeric Repeats. *Proceedings of the International Conference on Creationism* 8 (1): 222–228. <https://doi.org/10.15385/jpicc.2018.8.1.22>
- Ventura, M., C. R. Catacchio, S. Sajjadian, et al. 2012. The Evolution of African Great Ape Subtelomeric Heterochromatin and the Fusion of Human Chromosome 2. *Genome Research* 22 (6): 1036–1049. <https://doi.org/10.1101/gr.136556.111>