The chromosome 2 fusion model of human evolution—part 1: re-evaluating the evidence

Jerry Bergman and Jeffrey Tomkins

One of the leading molecular arguments for human evolution from a shared common ancestor with apes, particularly chimpanzees, is the 'chromosome 2 fusion model'. This scenario involves the claim that the fusion of two small chimpanzee-like chromosomes (2A and 2B) formed one stable chimera chromosome in humans, leading to the difference in diploid chromosome numbers between humans and great apes. A majority of the data for the fusion model is based on DNA hybridization and chromosomal staining experiments conducted prior to the sequencing of the human and chimpanzee genomes. In the present paper, we present a new analysis of the scientific literature, and in a companion paper (part 2 in this issue) a re-analysis of the available DNA sequence data that calls into question the validity of the fusion model.

One of the most popularized molecular arguments for human-primate evolution is the hypothetical prehistoric head-to-head fusion of two primate chromosomes (corresponding to 2A and 2B in chimpanzee) to form human chromosome number 2. Much of the research supporting this hypothetical model is based on indirect evidence derived from DNA hybridization and chromosomal staining techniques. These techniques provide only approximate estimates of sequence similarity, with hybridization-based analyses being more accurate than the analysis of stained chromosomal bands. This type of initial evidence, along with some targeted DNA sequencing of small genomic regions in human, seemed to indicate support for the fusion model.[1,2]

While the chromosome 2 fusion model is routinely touted as dogma, very little new genomic data, although readily available for analysis, has been presented as evidence. In addition, several science authors have recently published books for the general public popularizing this hypothetical model as one of the supposedly strongest arguments for human evolution from a shared common ancestor with apes, particularly chimpanzees.[3,4]

Popular reviews on this subject often include a simplified drawing depicting how the putative fusion of two small acrocentric[5] ape-like precursor chromosomes could have fused end-to-end to form the larger human chromosome 2, as shown in figure 1. In support of this hypothetical model of chromosome fusion, it is claimed that human chromosome 2 contains two key features that support the model. The first feature purportedly depicts the fusion event and contains genomic sequences representing a head-to-head fusion of telomeres, the highly specific end-cap DNA repeat motifs $(TTAGGG)_n$ located at the termini of linear mammalian chromosomes.[6]

The second key site purportedly represents a cryptic, non-functional centromere that was silenced following the fusion event (because a single functional centromere is required for chromosome stability and function). According to these claims, this fusion event accounts for the fact that humans have only 46 (2N) chromosomes and the great apes 48 (2N). Actually, the diploid genomes of gorilla, chimpanzee and orangutan have 48 but some gibbons have 44, and one Malaysian ape has 50.[7] The fusion model scenario involves a hominid evolved from a shared common ancestor with a diploid genome of 48 chromosomes, and, in some early human ancestor, two chromosomes fused, reducing the diploid chromosome complement to 46.

Examining the existing genomic evidence for fusion

Of the two genomic features that are claimed to support the fusion model, the primary evidence used is the presence of a reputed fusion site. This site is located in a pericentric region (meaning it is close to the present functional centromere) on the long arm of human chromosome 2. The DNA sequence at this location is supposed evidence of a head-to-head telomeric fusion of two acrocentric chromosomes.

In his recently published book, the *Nature* article Miller cites as proof for the fusion states only that "Human chromosome two is unique to the human lineage in being the product of a head-to-head fusion of two intermediate-sized ancestral chromosomes" and provides no evidence for this conclusion. However, Fairbanks offers more detail and claims that there appears to exist a fusion site involving a set of 158 telomere sequences, and, of the 158 repeats, he notes only 44 sets can be manipulated to achieve perfect telomere consensus sequences. Another example of the same claim from a book popularizing human evolution is as follows:

> "The DNA sequences in the human chromosome are exactly as expected from this scenario. Telomeres consist of many repeats of the nucleotide

sequence TTAGGG, and at the fusion point of the human chromosome, where the two telomeres fused, this sequence is found 'head to head'. The functional centromere in chromosome 2 lines up with the chimpanzee chromosome 2p13 chromosomal centromere. The remains of the redundant centromere from one of the ancestral ape chromosomes can also be found."[3]

As we will document, these popular claims of resounding evidence for a telomere fusion producing human chromosome 2 are, for the most part, unsupported by the scientific literature and actual DNA sequence information (see companion paper). However, we will first briefly clarify the structure and nature of telomeres as to what would be expected if such a fusion event occurred. In so doing, we will take into account the accepted evolutionary presuppositions and timelines related to such an event.

Telomeres are typically found at the ends of linear eukaryotic chromosomes and confer stability by preventing fusion via a 'capping' function. The telomere region involves a complex and dynamic framework of DNA motif repeats, structural loops, structural and functional RNAs and a wide variety of proteins.[6] In a fusion event as described in the human chromosome 2 model, the end result should produce two identifiable telomeres characterized by a specific repeat motif and oriented in a head-to-head configuration. A certain genomic landscape must be present if two chromosomes fused head-to-head as claimed by the current evolutionary model. The consensus 5' to 3' telomere motif in humans, chimps, apes, and mammals in general, is $(TTAGGG)_n$ and typically occurs in perfect tandem for stretches of DNA from about 10 to 15 kb (10,000 to 15,000 bases) and contains 1,667 to 2,500 telomere repeats at each chromosome end. In a head-to-head fusion of two chromosomes, we would expect at least 5,000 bases of $(TTAGGG)_n$ repeats in tandem, albeit in a slightly degenerate state, given a supposed ~1 to 5 million years of evolution since the fusion event occurred. At the point of fusion, we would also expect the orientation of the plus-strand repeat to change to the reverse complement $(CCCTAA)_n$, which should also occur in near-perfect tandem for approximately 5,000 or more bases.

In reality, the putative fusion site is but a vague shadow of what should be present given the model in question. One of the major problems with the fusion model is that, within the 10 to 30 kb window of DNA sequence surrounding the hypothetical fusion site, a glaring paucity of telomeric repeats exist that appear mostly as independent monomers, not tandem repeats. Based on the predicted model, thousands of intact motifs in tandem should exist. For the TTAGGG repeat to the left of the fusion site, less than 35 motifs exist, a normal human telomere would typically have 1667 to 2500.[6] For the CCCTAA reverse complement sequence, to the right of the fusion site, less than 150 telomere motifs can be found. Another problem with these two morifs, that we document in our companion research paper, is that their occurences are found scattered throughout both sides of the fusion site where they would not be expected. In other words, both the forward and reverse complement of the telomere motif populate both sides of the fusion site.

Besides the extreme paucity of telomeric repeats, their largely monomeric condition and their ubiquitous presence on both sides of the purported fusion site, there is very little to indicate that they once formed 10- to 15-kb stretches of perfect tandem 6-base repeats. If a fusion occurred, the alleged sequence no longer resembles telomeric repeats, a problem explained away by fusion supporters by claiming the telomeric repeat area is incredibly degenerate. Nor is the location of the alleged telomeric repeats in the region where it should be, judging by our analysis of the chromosome it is claimed to be fused to (covered in part II of the study).

The only evolutionary research group to seriously analyze the actual fusion site DNA sequence data in detail were confounded by the results which showed a lack of evidence for fusion—a genomic condition for this region

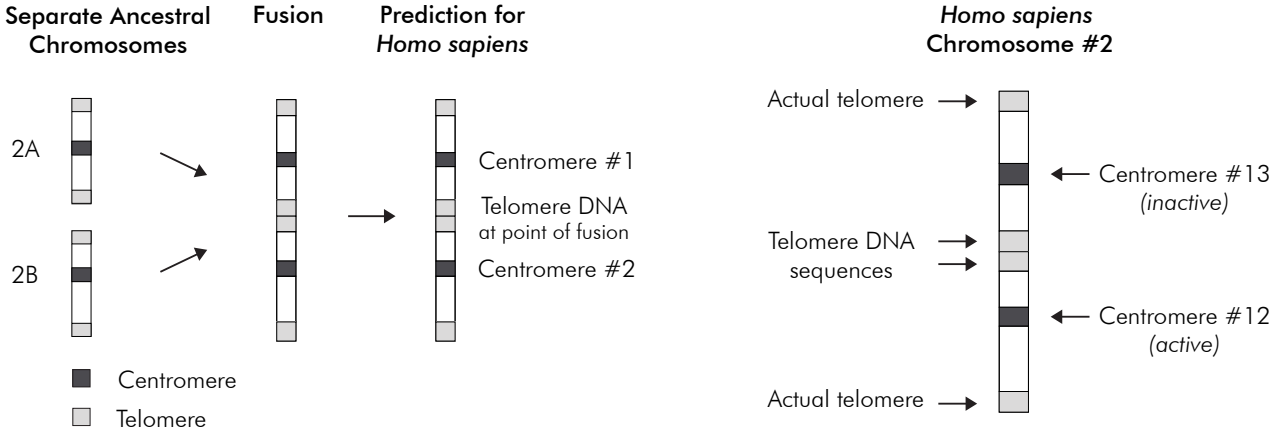

Figure 1. Depiction of a hypothetical scenario where chimpanzee chromosomes 2A and 2B supposedly fuse to form human chromosome 2. The prediction is on the left and the results, according to Miller ref. 4, pp. 106, 107, on the right. Miller's prediction was falsified, and thus the diagram on the left does not fit with the facts shown in both parts 1 and 2 of this study.

which they termed 'degenerate'.[8] In attempting to correlate rates of evolutionary change with the extreme degeneracy observed in the putative fusion region, they claimed that the "head-to-head arrays of repeats at the fusion site have degenerated significantly from the near perfect arrays of $(TTAGGG)_n$ found at telomeres." They also stated, "if the fusion occurred within the telomeric repeat arrays less than ~6 Ma, why are the arrays at the fusion site so degenerate?" The actual data indicates that perhaps the only thing that is degenerate is the evolutionary dogma surrounding the fusion model.

Because of the 'degenerate' nature of the DNA sequence data in this region, a variety of creative and manipulative approaches have been used to make the data look more telomere-like than it actually is. For example, Fairbanks claims that 44 out of 158 repeats match (28%) and that the rest of the sequences are 'close'.[3] The problem is, to obtain even this low match level, the consensus reading frame is entirely ignored and ambiguous matches are contrived by assuming many insertion and deletion mutations of varying sizes. In addition, Fairbanks' data include several additional perfect motifs immediately surrounding the fusion site that do not actually appear in the current GenBank accessions for this region.[3] Unfortunately, Fairbanks did not cite the accession number(s) for his fusion site sequence printed in his text. When the reading frame is corrected at various motifs near the fusion site, hardly any telomere sequences can be obtained. Fairbanks assumes that major differences between a perfect telomere and the existing sequence are the result of the accumulation of numerous insertions, deletions and other mutations, a post-hoc explanation that lacks strong DNA evidence.

Another problem for the fusion theory is the presence of a wide variety of genes throughout the fusion region. At present, no known protein coding genes have been found in the 10 to 15 kb tandem 6-base (TTAGGG) repeat terminal region of human telomeres.[9] In an analysis of a 614 kb area encompassing the postulated chromosome fusion site, Fan *et al*. found evidence of "at least 24 potentially functional genes and 16 pseudogenes".[10] In the 30-kb region directly encompassing the fusion site, which should definitely be devoid of any genes, there exists two actively transcribed genes, each in a flanking position in regard to the fusion site (one on each side). There are also at least two other genes in the immediate vicinity of the fusion site thought to be inactive due to frame shift mutations. However, research related to the human ENCODE (Encyclopedia of DNA Elements) project has shown that many genes thought to be inactive (pseudogenes) are actually functional due to a variety of newly discovered regulatory mechanisms.[11]

If the telomere motifs that populate internal areas of chromosomes serve some important, yet unknown function, the chromosome fusion model actually impedes research aimed at determining possible function in these regions. This type of reasoning is not without precedent. For example, the widely held concept of the genome consisting of mostly 'junk DNA' has now been discredited.[11,12]

Assuming that two telomeres exist in a head-to-head fusion produces another major problem, namely that telomeres are designed to prevent fusion. Broken chromosomes at any location immediately invoke the cell's double-stranded DNA repair machinery where the aberrant fusion of fragments actually triggers cell fault tolerance mechanisms.[6] In the case of an aberrant fusion, a senescence response or programmed cell death (apoptosis) cascade is normally triggered, effectively eliminating the damaged cell from the system.

A cell with telomeres that have progressively shortened over time and reached a threshold length will also activate the double-stranded DNA repair machinery; inducing cell senescence and/or death. When in certain types of germ-line cells, telomerase adds telomere repeats to shortened telomeres, chromosomes are 'healed' and can again become stable. The telomeres cap the ends of linear chromosomes and effectively prevent fusion or trigger cell elimination if the telomere is shortened to a certain point, damaged, or aberrantly fused.[6] According to the fusion model, this protective process was somehow bypassed in early humans.

Examining the evidence for a cryptic centromere

Yet another major problem with the fusion model is the lack of evidence for a cryptic second centromere site. Immediately following the supposed head-to-head telomere fusion, there would have existed two centromeres in the newly-formed chimeric chromosome, one from each of the two fused chromosomes. This type of event had to occur in a cell lineage of the germ-line to be heritable, and one of the centromeres would have had to be immediately eliminated or at least functionally silenced for cell division to progress normally. Evolutionists explain the lack of a clearly distinguishable non-functional secondary centromere by arguing that two centromeres would result in major instability when chromosomes pair up during cell division and consequently would be rapidly selected against. According to the evolutionary model, selection would continue until the second centromere was completely non-functional.

However, the evidence for a second remnant centromere at any stage of sequence degeneracy is negligible. As Fairbanks noted, "Fusion at the telomeres should have left two centromeres in the ancient fused chromosome, but there is only one now."[3] He then evaluated the "evidence that a centromere was once present at a second site". The supposed evidence includes the finding that "every human and great-ape chromosome centromere contains a highly variable DNA sequence that is repeated over and over, a 171 base-pair sequence called the Alphoid sequence."[3] Fairbanks

adds that scientists have "searched for Alphoid sequences in human chromosomes and found them at every centromere, as expected. They also found Alphoid sequences at the site in human chromosome 2 where the remnants of this second centromere should be. These remnants are evidence of a now-defunct centromere."[3]

The main problem with Fairbank's claim is that alpha-satellite DNA or alphoid DNA, although found in centromeric areas, is not unique to centromeres and is also highly variable. Because highly variable alphoid DNA is also commonly found in non-centromeric regions of human chromosomes, their presence does not indicate the remnants of a degenerate centromere.

Based on the reasoning of Fairbanks and others promoting the human chromosome 2 fusion model, one could conclude that human chromosomes contain literally hundreds of degenerate centromeres. As a result, locating a candidate alphoid region to erroneously support the presence of a degenerate centromere on chromosome 2 would not be unexpected or difficult to do, and does not support a cryptic centromere claim. In the companion research paper supporting this review, we show that the alphoid sequence in question does not align closely with known functional centromeric human DNA.

Despite the variation, there is enough sequence similarity for fluorescently labeled alphoid probes to hybridize to most classes of alphoid sequence in stretched chromosomal fibers. This explains many of the early chromosome fusion reports which relied on this technology (as shown in part II alphoid sequences are found throughout large sections of the chromosome and are not by themselves evidence of a centromere.).[1,2] Another problem is that, although research has been done on some primates, no systematic study of centromeres exists to determine how common alphoid DNA is in mammals.[13]

Multiple reports of alphoid/centromere similarity between humans and apes, involving both hybridization and sequence-based research, find that there is virtually no apparent evolutionary homology, except for moderate similarity on the X-chromosome centromere.[13] Baldini *et al.* found that the "highest sequence similarity between human and great ape alphoid sequences is 91%, much lower than the expected similarity for selectively neutral sequences."[13] Alphoid regions, in contrast to many classes of DNA sequences, are not well-conserved among mammalian taxa and even show high levels of diversity between chromosomes in the same genome.[14]

Cytogenetic anomalies argue against fusion

Other problems with the fusion theory include the fact that standard cytogenetic techniques, such as C-banding, have detected significantly less heterochromatic centromeric DNA on the long arm of human chromosome 2 than predicted by the fusion model. Evolutionists claim this is because the "bulk of the centromeric repetitive DNA has been lost".[13] Conversely, it is more likely that the so-called cryptic centromeric DNA never existed.

Not only does the DNA sequence at the putative cryptic centromere site argue against fusion, but a comparison of the chimp and human chromosomes reveals that the centromere in human chromosome 2 is in a very different location than predicted by a fusion event as shown in part two of this study. This necessitates an implausible series of events, including the loss of both chimp centromeres when chromosomes 2A and 2B fused, and the rapid evolution of a new centromere to provide functionality to human chromosome 2.

Mutations of the magnitude needed to support a fusion event pose serious cytogenetic problems both for the organism during regular somatic cell growth related to mitosis and during the meiotic events occurring in the germ-line tissues. Proper alignment requires the near-identical structure of each pair so that each chromosome aligns only with its sister chromosome. Chromosomal fusion is one major common cause of infertility. If meiosis does occur despite the aberration, the embryo produced from fertilization of these gametes typically self-aborts.[15]

Do fusions lead to new species?

Evolutionary scientists believe an ape-like ancestor evolved into a new species, called *Homo sapiens*, along with a major genomic fusion event. While the order of genes and their spatial relationship in the nucleus can affect gene expression, no new information or genes are added by fusing two existing chromosomes, because only the gene packaging is altered. However, the information content of the genome can still be strongly affected. Chromosomal fusion has been identified in a variety of animal taxa, such as ruminants (sheep, goats and cattle) that were phenotypically similar compared to normal animals in their genera although reproductive isolation did occur.[16] Evolutionists postulate that such an event may have contributed to a reproductive barrier in early evolving humans who, although they may have had a new karyotype, were still closely related to apes.

Actually, the fusion theory creates problems for Darwinists due to the fact that a complete absence of humans with 48 chromosomes exists. Although very rare, chromosome fusions do occasionally occur in humans but are not easily passed to offspring. If a chromosomal split occurred during human evolution, then two distinct human groups would result. One evolutionary explanation for this problem is that the entire population of 48-chromosome proto-humans became extinct. Altered animal karyotypes that are not detrimental are rare, but produce populations representing both karyotypes.

Recent genomic problems for fusion

A common claim for fusion is that "the DNA sequence of the rest of human chromosome 2 closely matches very precisely the sequences of the two separate chimpanzee

chromosomes."[17] This claim is unsupported by a lack of detailed comparative DNA sequence data. It is noteworthy that the chimp rough-draft DNA sequence assembly was largely based on the human genome as a framework for its construction.[18] One of the first published attempts at reporting a genome assembly based on a physical framework[19] constructed for the chimp genome was the recently reported chimp Y chromosome project.[20] The end result was a completely different and unexpected genomic landscape showing extreme DNA sequence dissimilarities (30% or greater difference) between the human and chimp Y chromosomes. Why comparisons have not been reported for the other chromosomes is unknown, considering the technology to provide a chimp genome assembly based on a chimp contig-based physical map has been available since 2006.[20–21]

Conclusion

The purportedly overwhelming DNA evidence for a fusion event involving two primate chromosomes to form human chromosome 2 does not exist, even without the aid of new analyses. In this report, our review of only the reported data by evolutionary scientists shows that the sequence features encompassing the purported chromosome-2 fusion site are far too ambiguous to infer a fusion event. In addition to a lack of DNA sequence data for a head-to-head chromosomal fusion, there also exists a decided paucity of data to indicate a cryptic centromere. In a companion paper (part 2) to this, we report the results of additional data analyses using a variety of bioinformatic tools and publicly available DNA sequence resources that further refute the hypothetical chromosome fusion model.

References

1. Yunis, J J. and Prakash, O., The origin of man: a chromosomal pictorial legacy, *Science* **215**:1525–1530, 1982.
2. Ijdo, J.W. *et al.*, Origin of human chromosome 2: an ancestral telomere-telomere fusion, *Proc. Natl. Acad. Sci.* **88**:9051–9055, 1991.
3. Fairbanks, D.J., *Relics of Eden*, Quotations and analyses are in reference to Chapter 1, Fusion, Prometheus Books, Amherst, N.Y., pp. 17–30, 2007.
4. Miller, K.R., *Only a Theory: Evolution and the Battle for America's Soul*, Viking, New York, 2008.
5. Acrocentric chromosome has one arm that is considerably shorter than the other arm. The 'acro-' in 'acrocentric' refers to the Greek word for 'peak'. The human genome contains five acrocentric chromosomes: 13, 14, 15, 21 and 22. The naming system of chimpanzee chromosomes '2A' and '2B' was first used by McConkey, E.H., Orthologous numbering of great ape and human chromosomes is essential for comparative genomics, *Cytogenet Genome Research* **105**:157–158, 2004. Prior to that, the chimpanzee chromosomes were numbered according to size, as is the system used for all other species.
6. Tomkins, J. and Bergman, J., Telomeres: implications for aging and evidence for intelligent design, *J. Creation* **25**(1):86–97, 2011.
7. Jauch, A. *et al.*, Reconstruction of genomic rearrangements in great apes and gibbons by chromosome painting, *PNAS* **89**:8611–8615, 1992.
8. Fan, Y. *et al.*, 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, 2002.
9. Devanshi, J. and Promisel J.P., Telomeric strategies: means to and end, *Annu. Rev. Genet.* **44**:243–69, 2010.
10. Fan, Y. *et al.*, Gene content and function of the ancestral chromosome fusion site in human chromosome 2q13-2q14.1 and paralogous regions, *Genome Research* **12**:1663–1672, 2002.
11. The ENCODE Project Consortium, Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project, *Nature* **447**:799–816, 2007. Also see the June 2007 (vol 17) issue of *Genome Research.*
12. Wells, J., *The Myth of Junk DNA*, Discovery Institute Press, Seattle, WA, 2011.
13. Baldini, A. *et al.*, 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, 1993.
14. Alkan, C., *et al.*, Genome-wide characterization of centromeric satellites from multiple mammalian genomes, *Genome Research* **21**:137–145, 2011.
15. Alberts, B. *et al.*, The cell cycle; in: *Molecular Biology of the Cell*, 5th ed., Garland Science, New York, pp. 1053–1114, 2008.
16. Lightner, J.K., Karyotype variability within the cattle monobaramin, *Answers Research J.* **1**:77–88, 2008.
17. Alexander, D., *Creation or Evolution: Do We Have to Choose?* Monarch, Oxford, UK, p. 212, 2008.
18. The Chimpanzee Sequencing and Analysis Consortium, Initial sequence of the chimpanzee genome and comparison with the human genome, *Nature* **437**:69–87, 2005.
19. Warren, R.L., Physical map assisted whole-genome shotgun assemblies, *Genome Res.* **16**:768–775, 2010.
20. Hughes, J.F. *et al.*, Chimpanzee and human Y chromosomes are remarkably divergent in structure and gene content, *Nature* **463**:536–539, 2010.
21. genome.wustl.edu/genomes/view/pan_troglodytes/.

Jerry Bergman has nine academic degrees including two Ph.Ds. His major areas of study for his graduate work were in biology, chemistry, psychology, and evaluation and research. He graduated from Wayne State University in Detroit, Medical University of Ohio in Toledo, University of Toledo and Bowling Green State University. A prolific writer, Dr Bergman has taught biology, chemistry and biochemistry at Northwest State in Archbold, Ohio for over 24 years. He is now an adjunct associate professor at The University of Toledo Medical College.

Jeffrey Tomkins has a Ph.D. in Genetics from Clemson University, a M.S. in Plant Science from the University of Idaho, Moscow and a B.S. in Agriculture from Washington State University. He was on the Faculty in the Dept of Genetics and Biochemistry, Clemson University, for a decade and is now a Staff Scientist at ICR. He has 56 publications in peer reviewed scientific journals and 7 book chapters in scientific books.

The chromosome 2 fusion model of human evolution—part 2: re-analysis of the genomic data

Jeffrey Tomkins and Jerry Bergman

A major argument for human evolution from a shared common ancestor with the great apes, particularly chimpanzees, is the 'chromosome 2 fusion model'. This molecular model involves the hypothetical fusion of two small acrocentric chimpanzee-like chromosomes (2A and 2B) at some ancient point in the human evolutionary lineage. Our analysis of the available genomic data shows that the sequence features encompassing the purported chromosome 2 fusion site are too ambiguous to accurately infer a fusion event. The data actually suggest that the core ~800 bp region containing the fusion site is not a unique cryptic and degenerate head-to-head fusion of telomeres, but a distinct motif that is represented throughout the human genome with no orthologous counterpart in the chimpanzee genome on either chromosome 2A or 2B. The DNA sequence evidence for a purported inactivated cryptic centromere site on chromosome 2, supposedly composed of centromeric alphoid repeats, is even more ambiguous and untenable than the case for a fusion site. The alphoid sequences in this region are quite variable and do not cluster with known functional human centromeric sequences. In addition, no ortholog for a cryptic centromere homologous to the alphoid sequence at human chromosome 2 exists on chimpanzee chromosomes 2A and 2B.

One of the most cited DNA-based arguments for human evolution is the hypothetical head-to-head fusion of two small ape-like chromosomes to form human chromosome 2.[1] The corresponding chromosomes supposedly represented in the great apes are 2A and 2B in the chimpanzee genome. A majority of the research that undergirds this model utilized indirect methods of DNA analysis. These data were derived from DNA probe hybridization, chromosomal banding (staining), and limited DNA sequencing techniques that were available prior to the advent of high-throughput DNA sequencing technology.[1,2]

Chromosome staining and hybridization techniques do not provide detailed DNA sequence information, but rather indicate putative areas of homology. Chromosome staining used to achieve visible banding markers yields information related to GC base content, repeat content, CpG island density, and degree of condensation over large areas rather than specific sequence homology.[3,4] Probe (DNA) hybridization is a more direct and accurate method for detecting DNA homology, but is subject to lab protocol variability and does not provide actual DNA sequences. Early DNA sequencing projects were largely limited to small, isolated regions of eukaryote genomes, a scenario that changed with the introduction of large-insert DNA cloning (bacterial artificial chromosomes; BACs) and BAC contig-based physical mapping strategies.

The advent of high-throughput DNA sequencing and it's accompanying technologies has largely replaced these earlier technologies for comparing both chromosomes and genomes. The first working draft of the human genome generated in both the public and private sectors was available in 2001 and a more complete draft of the public human genome sequence became available in 2003.[5–7] The chimpanzee genome project also received funding, and a 5-fold redundant shotgun sequence coverage was published in 2005.[8] Another 1.5-fold coverage was completed after this along with the construction of a BAC contig-based physical map for chimpanzee.[7]

While the chromosome 2 fusion model has been routinely discussed in reviews of human evolution, very little new supporting genomic data, although readily available for analysis, has been forthcoming. For the purpose of propagating the dogma surrounding human evolution, several science authors have recently published novice-level science books promoting the hypothetical chromosome 2 model.[10,11] This so-called factual data is routinely used as one of the leading arguments for human evolution from a shared common ancestor with apes.

The general model involves the hypothetical fusion of two small, acrocentric,[12] ape-like precursor chromosomes thought to have fused end-to-end, forming the single large human chromosome 2, as illustrated in figure 1. From a DNA sequence perspective, it is claimed that human chromosome 2 contains two key regions in its landscape. The first region of interest is thought to depict the actual head-to-head fusion of telomeres. Telomeres are end-cap DNA repeat motifs $(TTAGGG)_n$ located at the termini of linear mammalian chromosomes, recently reviewed by Tomkins and Bergman.[13] The second key region supposedly represents a cryptic non-functional centromere that was inactivated following the fusion event. For each chromosome, a single functional centromere is required for proper stability and function because a dual centromere situation created by such a fusion would cause cellular instability and destruction. Although there are no well-defined mechanisms for inactivating human centromeres,

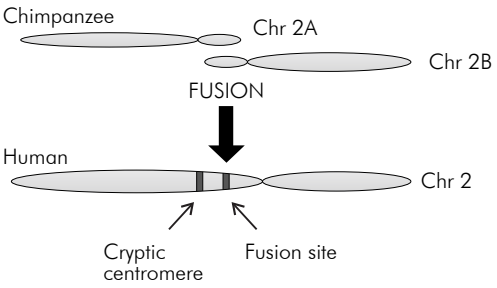

Figure 1. Depiction of a hypothetical scenario where chimpanzee chromosomes 2A and 2B supposedly fuse to form human chromosome 2. The two sites showing where the fusion occurred and an inactivated cryptic centromere are depicted.

it is believed that one of the two resulting centromeres was somehow silenced as a result of fusion. The chromosome 2 fusion is thought to account for the fact that humans have only 46 (2N) chromosomes and the great apes, including chimpanzee have 48 (2N). Modern humans supposedly evolved from a shared common ancestor with a diploid genome of 48 chromosomes, requiring a fusion event.

Examining the genomic evidence for fusion

Of the two genomic regions that are claimed to support the fusion model, the primary evidence is the purported fusion site. This site is located in a region close to the present functional centromere on the long arm of human chromosome 2. This particular area containing the 'fusion region' is often called 2qfus or 2chr2fus and occupies the genomic area between 2q13 and 2q14.1.[14] The two small chimpanzee chromosomes that supposedly contributed to the fusion event are currently identified as 2A and 2B.

The human 2qfus region has been sequenced and annotated for telomeric repeats, a variety of important functional genes, processed pseudogenes, and various open reading frames (ORFs). A fairly thorough and complete 614 kb (614,000 bases) annotated genomic landscape was constructed that encompasses the fusion site and was published by a lab in several related reports shortly after the initial first working draft of the human genome project.[15,16] The primary substrate for the effort relied on the assembled sequence from five overlapping, large-insert DNA clones (bacterial artificial chromosomes; BACs). As a result of this effort, a 177 kb region of contiguous sequence directly surrounding the 2qfus site corresponding to BAC clone RP11-395L14 (accession number AL078621) is available for public access and download. For the purpose of clarifying claims related to the fusion site, we subjected the complete BAC sequence of RP11-395L14 to a variety of telomere motif analyses (see Materials and Methods).

Fusion site DNA sequence analysis

Our DNA sequence analysis confirmed conclusions reached by Fan *et al*. The putative fusion site is 'highly degenerate' and a vague shadow of what should be present given the model proposed.[15] One of the major problems with the fusion model is that, within the 20- to 30-kb window of DNA sequence surrounding the hypothetical fusion site, there is a glaring paucity of telomeric repeats, and those that are present are mostly independent monomers, not tandem repeats. In fact, many of the motifs in the 30-kb region surrounding the putative 2qfus site are not only isolated monomers, but are separated by up to several thousand bases of DNA.

Even while completely disregarding a consensus 6-base reading frame when iterating through the repeats, for the left (plus strand) side of the fusion site, there are only 34 intact TTAGGG motifs (table 1). This analysis uses a generous allowance of 92,690 bases to the left of the fusion site where the first TTAGGG repeat is found on BAC RP11-395L14, well beyond the size of any normal human telomere. Based on the predicted model, thousands of intact TTAGGG motifs in tandem should exist. This is true even if allowing for an extremely high rate of degeneracy, which is an unreasonable expectation because meiotic recombination is suppressed in pericentric DNA due to its close proximity to the centromere. Recombination, the most likely theoretical source of sequence shuffling leading to the fusion site degeneration would therefore be less of consideration. Also, based on the predicted model, little, if any TTAGGG motifs should exist on the plus strand to the right of the fusion site. However, 18 intact TTAGGG motifs are found on the right of the fusion site; 35% of the total number of TTAGGG motifs located within a generous 156,911 base window surrounding the fusion site.

The reverse complement telomere sequence (CCCTAA) should be present in near-perfect tandem to the right of the fusion site. Like the TTAGGG motif, one would expect approximately 1667 to 2500 CCCTAA motifs if an end-to-end fusion occurred. However, only 136 intact motifs exist to the right of the fusion site, with the last CCCTAA on the BAC clone terminating at 64,221 bases to the right of the fusion (table 1). Again, this very generous stretch of sequence is much longer than a normal human telomere, and contains a paucity of motifs. In similar fashion to the TTAGGG forward motif, the CCCTAA motif was also located on both sides of the fusion site. Our analysis located a total of 18 occurrences of the CCCTAA motif (12% of the total) scattered throughout the opposite side of the fusion site, where it would not be expected to be found. In other words, both the forward and reverse complement of the telomere motif populate both sides of the fusion site. As a side note, the GC content of the 177 kb region encompassing the putative fusion site is significantly higher (45%) than the average (40%) for chromosome 2 (table 2).

A complete scan of the 237+ million bases of the assembled euchromatic sequence of chromosome 2 using the Skittle Genome Viewer software package showed that

Table 1. Telomere DNA sequence data for the 177 Kb BAC containing the fusion site.

Parameter	Number of occurrences		
	Left of fusion site	Right of fusion site	Total
TTAGGG motifs	34	18	52
CCCTAA motifs	18	136	154
Total bases DNA	108.569	68.167	176.736
GC content	4.56	313	~ 8000

the entire landscape, from end to end, is populated with TTAGGG and CCCTAA motifs. Small, isolated dense clusters of telomere motifs occured in at least 5 internal locations (data not shown). A complete iteration of the entire plus strand sequence of chromosome 2 (Per script written by Tomkins) indicated a total number of 'TTAGGG' and 'CCCTAA' occurences at 45,450 and 45,770, respectively (table 2). These numbers are roughly equal, indicating that both the forward and reverse orientation of the telomere motif occurs quite frequently at internal sites across the length of chromosome 2. These numbers indicate that a total of at least 547,320 internal bases on chromosome 2 are composed of widely distributed intact telomere motifs.

An important attribute associated with these internal telomere motifs is that they are largely monomeric. Of the 52 intact TTAGGG motifs on both sides of the fusion site, only three tandem occurrences were found, with the rest existing as independent monomers. Of the 154 intact CCCTAA motifs on both sides of the fusion site, eighteen tandem motifs were found, with the rest appearing as independent monomers. Although the density of motifs and dimeric repeats increases somewhat within the immediate vicinity of the putative fusion region, their positions in the reading frame from one 6-bp telomeric repeat to the next are erratic (not in frame).

Because of the extreme paucity of telomeric repeats, their largely monomeric condition, and their ubiquitous presence on both sides of the supposed fusion site, there exists little data to indicate that they may have once formed 10- to 15-kb stretches of perfect, tandem 6-base repeats. The 2qfus sequence is clearly degenerate beyond the point of indicating that intact telomeres once existed. Given the location in a region of suppressed pericentric recombination, one would expect a consideraby higher amount of telomere sequence preservation if the model was tenable.

In attempting to correlate rates of evolutionary change with the extreme degeneracy observed in the putative fusion region, one research group concluded that "the head-to-head repeat arrays at the RP11-395L14 fusion site have significantly degenerated from the near perfect $(TTAGGG)_n$ arrays found in telomeres."[15] This caused them to raise the question, "Why are the arrays at the fusion site so degenerate if the fusion occurred within the telomeric repeat arrays less than ~6 Mya?"[15]

A more valid explanation for the telomere-like features present at the putative fusion site is that they may represent some form of a distinct genomic motif. To test this idea, a 798-bp fragment (figure 2) encompassing the fusion site and the region where the telomeric motifs are more densely populated was used as a query subject in a BLAT[17] search on the most recent build of the human genome (v 37.1; www.genome.usc.edu) with masking disabled. The results revealed a total of 159 significantly placed hits throughout the genome on human chromosomes 1–11, 15, 18–20, X and Y. The homologous regions for these hits included areas near telomeres, pericentric areas, and a wide variety of internal euchromatic sites. Identity values ranged from 80.5 to 100%, supporting the conclusion that the telomere fusion site core sequence is not unique to its pericentric location on chromosome 2, and instead represents a sequence feature (motif) scattered throughout the human genome.

To verify the BLAT results and to identify homologous sites in the chimpanzee genome, the BLASTN algorithm was used (with no masking or gap extension) for comparisons between the 798-bp core 2qfus sequence and the most recent builds of the human (v 37.1) and chimp (v 2.1) genomes maintained at NCBI (www.ncbi.nlm.nih.gov/). Although the BLASTN query against the human genome was more data intensive than the index-based BLAT search, the results produced a total of 85 significantly placed hits on all human chromosomes except chromosomes 13, 16 and 17 (1–12, 14, 15, 18–22, X and Y). While the number of hits was reduced, compared to BLAT, more chromosomes with homologous sites were identified with the BLASTN search

Table 2. Telomere DNA sequence data for the for the assembled euchromatic sequence of human chromosome 2.

DNA base total	237.5 million bases
Number of TTAGGG motifs	45 450
Number of CCCTAA motifs	45 770
GC content	40.2%*

*The average GC content we calculated for chromosome 2 agrees with that published previously (Hillier *et al.*, Generation and annotation of the DNA sequences of human chromosomes 2 and 4, *Nature* **434**:724–731, 2005.)

because of the more direct nature of the algorithm (figure 3). Interestingly, human chromosomes 2, 16, 21 and 22 were peppered with the 'fusion site' sequence over the length of their entire euchromatic landscape (figure 3).

When the 798-bp core fusion sequence was BLASTN queried against the chimpanzee genome, the significantly placed hit count was reduced to 19, only 22% of the amount observed in the human genome. This is a startling find in light of the wide-spread claims that the human and chimpanzee genomes contain DNA sequence that is supposedly 96 to 98% similar, a claim perhaps related to the fact that the human genome was used as a scaffold to build the chimpanzee genome.[8] In addition, the human-chimp hit locations did not show strong synteny, as only 13 of the 19 hits (68%) shared visually similar locations in the genome (on chimpanzee chromosomes 1, 2B, 8, 9, 12, 14, 15, 18, 20 and 22).

The most startling outcome of this analysis is that the fusion site did not align with chimp chromosome 2A, one of the supposed pre-fusion precursors. Furthermore, the alignment at two locations on chromosome 2B, an internal euchromatic site and the telomere region of its long arm, did not match predicted fusion-based locations based on the fusion model. If the fusion model was credible, this should have produced an alignment with the telomeric region on chimpanzee 2B on the short arm.

There is, therefore, no real evidence for DNA homology between human and chimpanzee for the 798-bp core fusion sequence. The alignment data also severely calls into question claims of high overall sequence similarity of 96 to 98% between the genomes. Our results are indirectly supported by the exceptionally high levels of dissimilarity observed in a recent study of a section of the Y chromosome landscape between human and chimpanzee.

TGAGGGTGAGGG**TTAGGG**TTTGGGTTGGGTTTGGG
GTTGGGGTTGGGGTAGGGGTGGGGTTGGGGTTGGGGT
TGGGG**TTAGGG**GTAGGGGTAGGGGTAGGGGTAGGGTC
AGGGTCAGGGTCAGGG**TTAGGG**TT**TTAGGG**TTAGGATT
TTAGGGTTAGGGTAAGGGTTAAGGGTTGGGGTTGGGG
TTAGGGTTAGGGG**TTAGGG**TTGGGGTTGGGGTTGGGG
TTGGGGTTGGGGTTGGGG**TTAGGG**TTAGCTAAACCTAA
CCCTAAC***CCCTAA***CCCCAACCCCAACCCCAACCCTACCC
CTACCCCTAC***CCCTAA***CCCCAACCCCCACCCTTAACCCTT
AACCCTTA***CCCTAACCCTAA***CCCAAA***CCCTAA***CCCTA***CCC***
TAACCCTAACCCAA***CCCTAACCCTAA***CCCTA***CCCTAACCC***
TAACA***CCCTAA***AACCGTGACCCTGACCTTGACCCTGACC
CTTAACCCTTAA***CCCTAA***CCATAA***CCCTAA***A***CCCTAACCCT***
AAA***CCCTAACCCTAA***A***CCCTAACCCTAA***CACTACCCTA***CC***
CTAACCCCAAC***CCCTAA***C***CCCTAACCCTAA***CCCTAC***CCCT***
AACCCCAACCCCAGCCCCAACCCTTA***CCCTAA***CCCTA***CCC***
TAACCCTTAA***CCCTAA***C***CCCTAACCCTAA***C***CCCTAACCCT***
AACCCTACCCCAACCCCAAACCCAA***CCCTAA***CCCAA***CCC***
TAAC***CCCTAACCCTAA***CCCCTA***CCCTAA***CCCCTAGCCCTA
GCCCTAG***CCCTAACCCTAA***CCCTCG***CCCTAA***CCCTCA***CCC***
TAACCCTCACCCTCA***CCCTAA***

Figure 2. The 798 bp core sequence surrounding the fusion site on human chromosome 2 used for BLASTN searches against the most recent builds of the human and chimpanzee genomes. Intact telomeric motifs are highlighted in grey bold and bold italics for $TTAGG_n$ and it's reverse complement $(CCCTAA)_n$, respectively. The hypothetical head-to-head fusion site is underlined.

Examining DNA sequence for a cryptic centromere

Following the supposed head-to-head telomere-based fusion of two smaller chromosomes, two centromeres would have had to exist in the newly formed chimeric chromosome, one from each of the two fused chromosomes. According to the evolutionary model, sequence degeneration plus selection would continue until the second centromere was completely non-functional. The DNA evidence in question is based on the fact that human, great-ape, and other mammalian centromeres are composed of a highly variable class of DNA sequence that is repeated over and over called alpha-satellite or alphoid DNA.[18] Alphoid DNA, although found in centromeric areas, is not unique to centromeres and is even highly variable between homologous regions throughout the same mammalian genome.[18]

The basic human alphoid monomer is a 171-base motif represented by a patented synthetic consensus sequence in Genbank (Acc. # CS444613). There also exists two small sequenced clones representing alphoid repeats with proven cellular centromere function.[19] Nine different alphoid fragments in the cryptic centromere site associated with the purported chromosome 2 fusion event were also sequenced and submitted to GenBank by an Italian laboratory (see figure 4 for accession numbers). In total, we downloaded and analyzed all 12 of these sequences for similarity to each other and individually for genome-wide homology.

Using the BLAT tool on the most recent version (v 3.7) of the human genome assembly, the nine italian lab alphoid sequences elicited the strongest hits at the chromosome 2 putative cryptic centromere site for all accessions. This confirmed that they were cloned from this region of the genome. The consensus 171-bp alphoid sequence aligned at the cryptic centromere site with 90.6% identity, supporting the conclusion that the site contains alphoid-like sequences.

However, the concern is not if this location contains alphoid sequences that are known to be ubiquitous in the human genome, but how similar these sequences are to each other and to known functional centromeric alphoid repeats. Alphoid sequences located at centromeres form long series of repeat patterns that are very homogeneous in their repetitive structure, producing distinctive higher-order patterns. Alphoid regions that are non-centromeric are more diverse in their monomer content and form higher order patterns with different characteristics compared to centromeres.[20] At present, there are five known supra-classes of human alphoid monomers that combine in various combinations.[21] There is also evidence from research in

progress that alphoid monomer classes themselves can be broken down further into specific sub-fragments that may be present in the genome by themselves or as a sub-fragment in an alphoid repeat region (Tomkins, unpublished data).

In a human alphoid multiple-sequence alignment analysis, we combined the two functional centromeric alphoid sequences with the set of nine Italian alphoid sequences along with the consensus 171-base alphoid sequence in our data set (figure 4). We also created tandem repeats of the consensus 171-base alphoid sequence representing repeats of 2X to 4X in length as individual sequences. Alignments were conducted using the MUSCLE software package[22] then refined using the Gblocks program.[23]

The human alphoid alignments clearly revealed dissimilarity between alphoid sequences and distinct patterns of clustering. Patterns of similarity were computationally evaluated using PhyML[24] with tree rendering performed by TreeDyn (figure 4).[25] Four major groups were distinguished by the PhyML analysis with the functional centromere sequences clustering by themselves and not with the alphoid sequences located at the purported cryptic centromere site on chromosome 2. The sequences at the cryptic centromere site are clearly a diverse mixture of alphoid monomers, forming three separate groups and not distinctly representative of functional centromeric DNA. In a structural comparison of both the functional centromere and cryptic centromere sites on chromosome 2 with the genome visualization tool, Skittle,[26] the putative cryptic centromere site was considerably more sequence-diverse and structurally unordered compared to the functional centromere on chromosome 2 (data not shown). The complex higher-order architecture of this Alphoid-diverse site is clearly unique and not characteristic of a silenced degenerate centromere.

Multiple reports involving both hybridization and sequence-based research of alphoid/centromere similarity between humans and apes have found virtually no apparent evolutionary homology, except for moderate similarity on the X-chromosome centromere.[20,27] Baldini *et al*. found that the "highest sequence similarity between human and great ape alphoid sequences is 91%, much lower than the expected similarity for selectively neutral sequences."[28] Alphoid regions, in contrast to many classes of DNA sequences, are not well-conserved among taxa and even show high levels of diversity between chromosomes in the same genome.[18] When the human alphoid sequences in our data set were queried against the chimpanzee genome using both BLAT and BLASTN, we were unable to obtain a single significant hit, verifying the extreme dissimilarity observed in alphoid motifs between taxa. These data corresponded

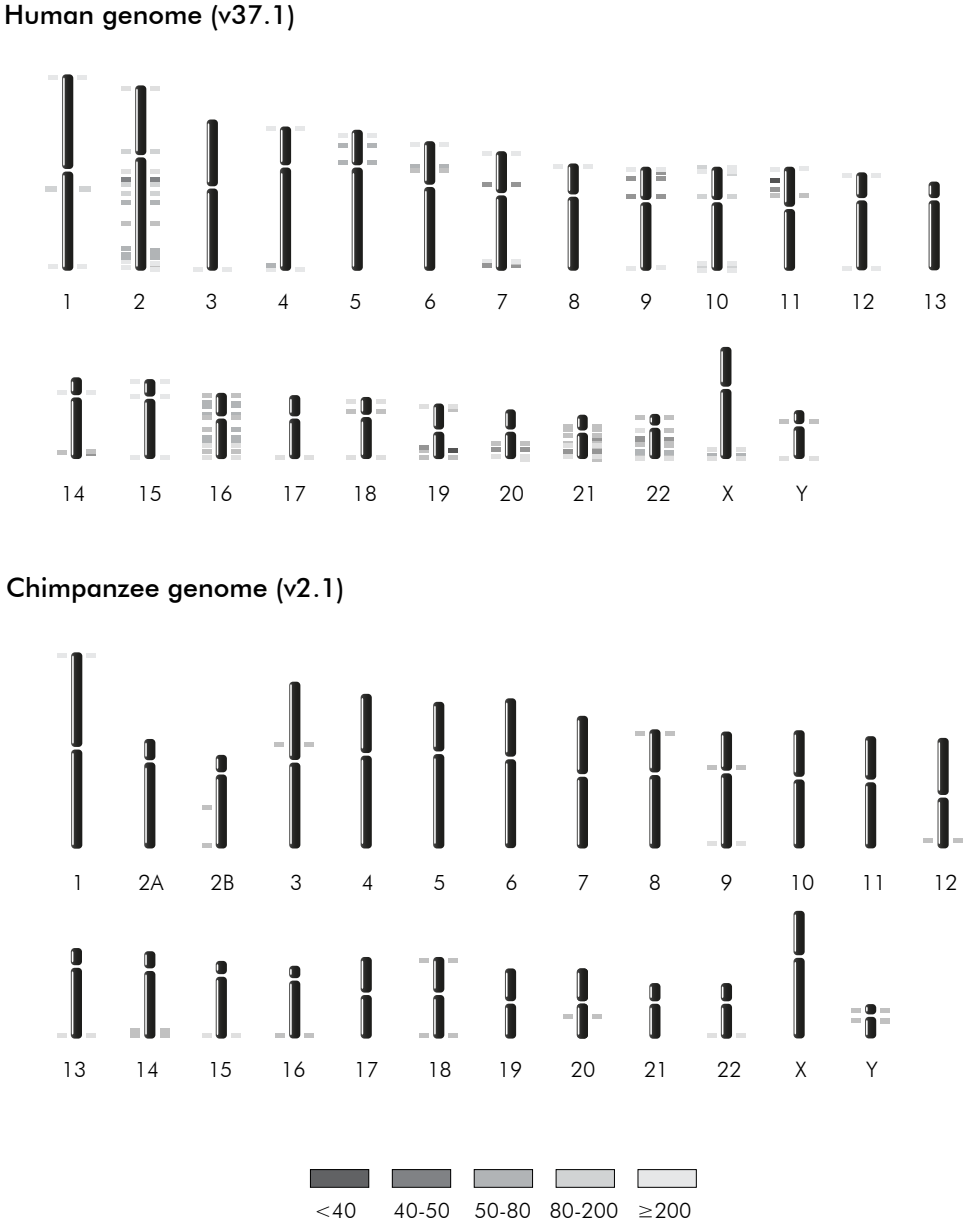

Figure 3. BLASTN results against the most recent builds of the human and chimpanzee genomes using a 798 bp human query sequence representing the core of the chromosome 2 fusion region.

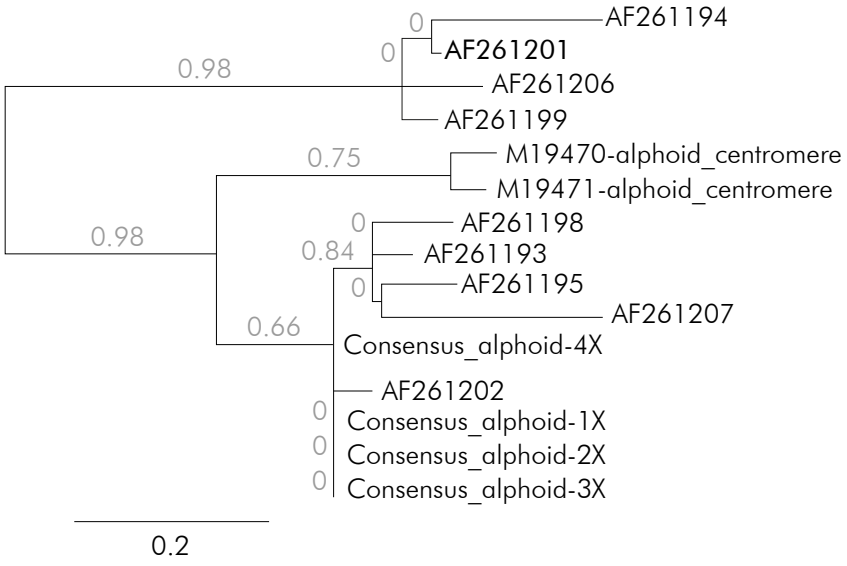

Figure 4. PhyML result with tree rendering by TreeDyn involving the nine Chromosome 2 alphoid sequences (prefix = AF) identified by accession numbers submitted to GenBank by Lonoce *et al*. (2000; unpublished—see genbank accessions at www.ncbi.nlm.nih.gov). The 171 bp consensus alphoid is included as a monomer and as repeats (2X, 3X, 4X). Two human alphoid sequences representing functional centromeric fragments identified by accession number are also included (prefix = M).

well with several decades of previous research by multiple labs, discussed above.

Summarized findings

1. The reputed fusion site is located in a peri-centric region with suppressed recombination and should exhibit a reasonable degree of tandem telomere motif conservation. Instead, the region is highly degenerate—a notable feature reported by a previous investigation.

2. In a 30 kb region surrounding the fusion site, there exists a paucity of intact telomere motifs (forward and reverse) and very few of them are in tandem or in frame.

3. Telomere motifs, both forward and reverse (TTAGGG and CCTAAA), populate both sides of the purported fusion site. Forward motifs should only be found on the left side of the fusion site and reverse motifs on the right side

4. The 798-base core fusion-site sequence is not unique to the purported fusion site, but found throughout the genome with 80% or greater identity internally on nearly every chromosome; indicating that it is some type of ubiquitous higher-order repeat.

5. No evidence of synteny with chimp for the purported fusion site was found. The 798-base core fusion-site sequence does not align to its predicted orthologous telomeric regions in the chimp genome on chromsomes 2A and 2B.

6. Queries against the chimp genome with the human alphoid sequences found at the purported cryptic centromere site on human 2qfus produced no homologous hits using two different algorithms (BLAT and BLASTN).

7. Alphoid sequences at the putative cryptic centromere site are diverse, form three separate sub-groups in alignment analyses, and do not cluster with known functional human centromeric alphoid elements.

Materials and Methods

DNA sequences described in this paper were downloaded from the National Center for Biotechnology (NCBI) web site in FASTA format text files.[29] Results from online BLAT (Blast-Like Alignment Tool)[17] searches were downloaded from the Genome Browser at the UCSC Genome Bioinformatics web site (genome.ucsc.edu/) as plain text files and parsed using a POSIX shell script written by J.P. Tomkins. Analyses for telomere motif occurence and GC content were performed using a Perl script written by J.P. Tomkins. Bioinformatic scripts developed and utilized in this study may be requested by contacting author Tomkins at jtomkins@icr.org. Figures depicting genome-view BLASTN (nucleotide BLAST) alignments were obtained using online software available at NCBI. For alphoid sequence alignments, the MUSCLE (Multiple Sequence Comparison by Log-Expectation)[22] program (v 3.7; www.ebi.ac.uk/Tools/muscle/index.html) followed by curation with Gblocks (v 0.91b; molevol.cmima.csic.es/castresana/Gblocks.html)[23] was used to evaluate alignments and select conserved blocks for analysis with PhyML (v 3.0; atgc.lirmm.fr/phyml/).[24] Tree data from PhyML was rendered with TreeDyn (v 198; www.treedyn.org/).[25] Sequence visualization of repeats and motif patterns were performed using the genome viewer software program Skittle[26] and the entire consensus sequence of human chromosome 2 downloaded as a compressed fasta file from NCBI.

Acknowledgments

We wish to thank John Sanford, Robert Carter, Clifford Lillo, and Mary Ann Stuart for their reviews of an earlier draft of this manuscript, and Josiah Seaman and John Sanford for their guidance in the usage of the 'Skittle' genome viewer software.

References:

1. Yunis, J.J. and Prakash, O., The origin of man: a chromosomal pictorial legacy, *Science* **215**:1525–1530, 1982.

2. Ijdo, J.W. *et al.*, Origin of human chromosome 2: an ancestral telomere-telomere fusion, *Proc. Natl. Acad. Sci.* **88**:9051–9055, 1991.

3. Furey, T.S and Haussler, D., Integration of the cytogenetic map with the draft human genome sequence, *Human Mol. Genet.* **12**:1037–1044, 2003.
4. Niimura, Y. and Gojobori, T., In silico chromosome staining: reconstruction of giemsa bands from the whole human genome sequence, *Proc. Natl. Acad. Sci.* **99**:797–802, 2002.
5. International Human Genome Sequencing Consortium, Initial sequencing and analysis of the human genome, *Nature* **409**:861–920, 2001.
6. Venter, J.C. *et al.*, The sequence of the human genome, *Science* **291**:304–1351, 2001.
7. International Human Genome Sequencing Consortium, Finishing the euchromatic sequence of the human genome, *Nature* **431**:931–945, 2004.
8. The Chimpanzee Sequencing and Analysis Consortium, Initial sequence of the chimpanzee genome and comparison with the human genome, *Nature* **437**:69–87, 2005.
9. Warren, R.L., Physical map assisted whole-genome shotgun assemblies, *Genome Res.* **16**:768–775, 2010.
10. Fairbanks, D.J., *Relics of Eden*, Prometheus Books, Amherst, NY, 2007.
11. Miller, K.R., *Only a Theory: Evolution and the Battle for America's Soul*, Viking, New York, 2008.
12. Acrocentric chromosome has one arm that is considerably shorter than the other arm. The 'acro' in acrocentric refers to the Greek word for 'peak'. The human genome contains five acrocentric chromosomes: 13, 14, 15, 21 and 22.
13. Tomkins, J. and Bergman, J., Telomeres: implications for aging and evidence for intelligent design, *J. Creation* **25**(2):86–97, 2011.
14. This system uses p to represent the small arm of the chromosome and q to represent the long arm. The numbers refer to chromosome bands – sections of the chromosome that can be visualized by stains and dyes or other techniques. Thus, 2q13 refers to section (band) 13 on the long arm of chromosome 2.
15. Fan, Y. *et al.*, Genomic structure and evolution of the ancestral chromosome fusion site in 2q13-2q14.1 and Paralogous Regions on Other Human Chromosomes, *Genome Res.* **12**:1651–1662, 2002.
16. Fan, Y. *et al.*, Gene content and function of the ancestral chromosome fusion site in human chromosome 2q13-2q14.1 and paralogous regions, *Genome Res.* **12**:1663–1672, 2002.
17. Kent, W.J., The BLAST-Like Alignment Tool, *Genome Res.* **12**:656–664, 2002.
18. Alkan, C. *et al.*, Genome-wide characterization of centromeric satellites from multiple mammalian genomes, *Genome Research* **21**:137–145, 2011.
19. Heartlein, M.W. *et al.*, Chromosome instability associated with human alphoid DNA transfected into the chinese hamster genome, *Molec Cell. Biol.* **8**:3611–3618, 1988.
20. Alkan C. *et al.*, Organization and evolution of primate centromeric DNA from whole-genome shotgun sequence data, *PLoS Comput Biol* **3**:1807–1818, 2007.
21. Roizes G., Human centromeric alphoid domains are periodically homogenized so that they vary substantially between homologues. Mechanism and implications for centromere functioning, *Nucleic Acids Res* **34**:1912–1924, 2006.
22. Edgar, C.E., MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Res.* **32**:1792–1797, 2004.
23. Talavera, G. and Castresana, J., Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments, *Systematic Biol.* **56**:564–577, 2007.
24. Guindon, S. *et al.*, PHYML Online—a web server for fast maximum likelihood-based phylogenetic inference, *Nucleic Acids Res.* **33**:W557–W559, 2004.
25. Chevenet, F. *et al*., TreeDyn: towards dynamic graphics and annotations for analyses of trees, *BMC Bioinformatics* **7**:439, 2006.
26. Seaman, J.D. and Sanford, J.C., Skittle: A 2-dimensional genome visualization tool, *BMC Bioinformatics* **10**:452, 2009.
27. Archidiacono, N. *et al.*, Comparative mapping of human alphoid sequences in great apes using fluorescence *in situ* hybridization, *Genomics* **25**: 477–484, 1995.
28. Baldini, A. *et al.*, 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, 1993.
29. www.ncbi.nlm.nih.gov/.

Jeffrey Tomkins has a Ph.D. in Genetics from Clemson University, a M.S. in Plant Science from the University of Idaho, Moscow and a B.S. in Agriculture from Washington State University. He was on the Faculty in the Dept of Genetics and Biochemistry, Clemson University, for a decade and is now a Staff Scientist at ICR. He has 56 publications in peer reviewed scientific journals and 7 book chapters in scientific books.

Jerry Bergman has nine academic degrees including two Ph.Ds. His major areas of study for his graduate work were in biology, chemistry, psychology, and evaluation and research. He graduated from Wayne State University in Detroit, Medical University of Ohio in Toledo, University of Toledo and Bowling Green State University. A prolific writer, Dr Bergman has taught biology, chemistry and biochemistry at Northwest State in Archbold, Ohio for over 24 years. He is now an adjunct associate professor at The University of Toledo Medical College.