Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
GENOMICS
Genomics 86 (2005) 767 – 770
www.elsevier.com/locate/ygeno

Short Communication

The complex history of distal human chromosome 1q

David Haig *

Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA

Received 8 August 2005; accepted 26 September 2005

Abstract

Human chromosome 1 has been claimed to be a conserved ancestral chromosome of eutherian mammals. However, two small regions from distal 1q (with orthology to mouse chromosome 11) appear to have a different history. These two regions are proposed to have been added to the ancestor of human chromosome 1 as a single block that was subsequently disrupted by a paracentric inversion. The translocation and inversion appear to have occurred at some time after the primate lineage diverged from a common ancestor with rodents. Reconstruction of the history of distal human chromosome 1q is complicated by the "reuse" of breakpoints in different mammalian lineages and by coincidental shared synteny between humans and cats.
© 2005 Elsevier Inc. All rights reserved.

Keywords: HSA1; SCAN domain; NALP; Comparative gene mapping

Whole-genome sequences are providing a wealth of new information about mammalian linkage groups. If a reliable phylogeny is available, these data can be used to make inferences about evolutionary changes in linkage groups. Recent phylogenetic studies [1,2] place primates together with rodents, lagomorphs, tree shrews, and flying lemurs in a clade known as Euarchontoglires. Complete genome sequences are currently available for two primates, *Pan troglodytes* and *Homo sapiens* (HSA), and two rodents, *Mus musculus* (MMU) and *Rattus norvegicus* (RNO). Euarchontoglires is the sister group to Laurasiatheria (a diverse group whose members include artiodactyls and carnivores). The best annotated laurasiatherian genome is currently that of *Canis familiaris* (CFA). Extensive sequence data are also available for *Bos taurus* (BTA) and *Sus scrofa* (SSC). Euarchontoglires and Laurasiatheria together comprise Boreoeutheria. Genomic data are limited for nonboreoeutherian mammals. *Gallus gallus* (GGA) is the only nonmammalian tetrapod for which a nearly complete genome sequence is available. These genomic data allow inferences about linkage in the genomes of three human ancestors. These are the most recent common ancestors (MRCAs) of rodents and humans, of laurasiatherians and humans, and of birds and humans.

Murphy and colleagues [3] concluded that HSA1 is a conserved linkage group that was present in the MRCA of eutherian mammals. My purpose, in this brief note, is to argue that a small region of distal HSA1q is probably a recent addition to this ancestral linkage group.

Results

The distal 100 Mb of human chromosome 1q has six breaks in synteny (A-F) between human and mouse genomes that define six blocks (I-VI) of apparently conserved synteny (Tables 1A and 1B; Fig. 1).

Breakpoints A and E

Breakpoints A and E separate genes with orthologs on MMU11 from genes with orthologs on MMU1. The genes of Blocks I and V are contiguous on MMU11 and CFA14, whereas the genes of Blocks II and VI are contiguous on MMU1 and CFA7. The simplest interpretation is that these are the breakpoints of a pericentric inversion that has occurred in either the rodent or the human lineage since both diverged from their MRCA. There are two possibilities:

(1) The current arrangement of Blocks I-VI was present in the MRCA and has been retained in the human lineage. If

* Fax: +1 617 495 5667.
E-mail address: dhaig@oeb.harvard.edu.

0888-7543/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2005.09.016

Table 1A
Flanking markers for breakpoints in human/mouse synteny from distal human chromosome 1q with chromosomal location of mouse orthologs

Breakpoint	Distal marker (mouse location)	Proximal marker (mouse location)
A	*ZNF496* (MMU11)	*ELYS* (MMU1)
B	*FMN2* (MMU1)	*CHRM3* (MMU13)
C	*ARID4B* (MMU13)	*TOMM20* (MMU8)
D	*RHOU* (MMU8)	*HIST3H2BA* (MMU11)
E	*JMJD4* (MMU11)	*CDC42BPA* (MMU1)
F	*SPTA1* (MMU1)	*CD1E* (MMU3)

so, an inversion occurred in the rodent lineage that juxtaposed Block II with Block VI and Block I with Block V. This was followed by a fission that separated (II+VI) and (I+V) onto separate rodent chromosomes.

(2) Blocks I and V were contiguous in the MRCA, as were Blocks II and VI. An inversion occurred in the human lineage to give the current human arrangement. Either Blocks (II+VI) and (I+V) were linked in the MRCA, in which case a fission occurred in the rodent lineage, or they were unlinked in the MRCA, in which case a fusion occurred in the human lineage.

I favor interpretation (2) because Blocks II and VI both contain genes that map to BTA16 and CFA7, whereas Blocks I and V contain genes mapping to BTA7 and CFA14. One of the breakpoints for this inversion would have occurred between *ZNF496* and *JMJD4* and the other between *ELYS* and *CDC42BPA* (Fig. 1). The latter breakpoint corresponds to a break in synteny between the human and the chicken genomes: the ortholog of *ELYS* maps to GGA3, whereas the ortholog of *CDC42BPA* maps to GGA14. Therefore, this short interval appears to have been involved in at least two independent evolutionary rearrangements.

Breakpoint B

Breakpoint B separates genes with orthologs on MMU13 from genes with orthologs on MMU1. *CHRM3* (ortholog on MMU13) is separated from *RGS7* (ortholog on MMU1) by 0.9 Mb on HSA1. *MACHR* and *LOC395620,* the presumed orthologs of these genes in the chicken genome, are separated by 0.3 Mb on GGA3. If the chromosome assemblies are accurate, this close juxtaposition in chickens and humans strongly suggests that *CHRM3* and *RGS7* have maintained conserved synteny since the MRCA of humans and chickens.

Table 1B
Human/mouse synteny blocks from distal human chromosome 1q

Block	MMU	Size (Mb)	Distal boundary	Proximal boundary
I	11	≈3	1q telomere	*ZNF496*
II	1	≈7	*ELYS*	*FMN2*
III	13	≈4	*CHRM3*	*ARID4B*
IV	8	≈6.5	*TOMM20*	*RHOU*
V	11	<1	*HIST3H2BA*	*JMJD4*
VI	1	≈70	*CDC42BPA*	*SPTA1*

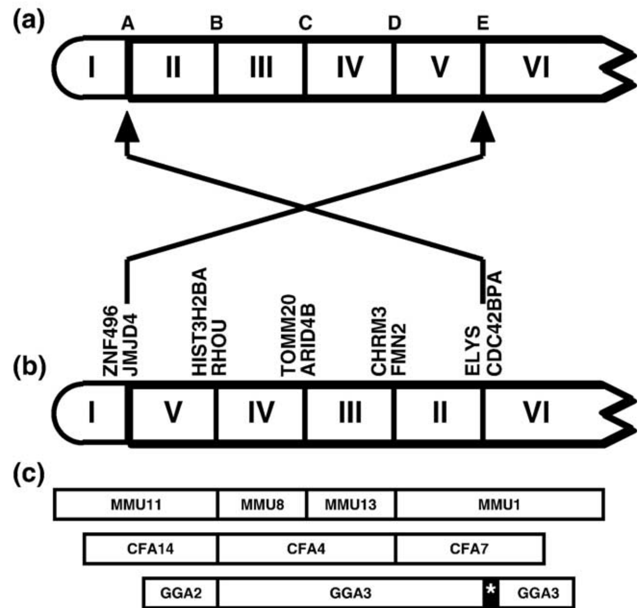

Fig. 1. Comparative gene mapping of distal HSA1q. (a) Breaks in human/mouse synteny (A-E) on HSA1 define syntenic Blocks I-VI (not drawn to scale). (b) The inferred ancestral order of these blocks in primates before the paracentric inversion that gave rise to the current gene order. (c) Syntenic relations between this ancestral order and mouse (MMU), dog (CFA), and chicken (GGA) chromosomes. The asterisk marks a small block of synteny between HSA1 and GGA14 that separates loci on HSA1 with orthologs on two noncontiguous regions of GGA3.

However, *CHRM3* and *RGS7* are unlinked in dogs, cattle, rats, and mice (*CHRM3* on CFA4, BTA28, RNO17, MMU13; *RGS7* on CFA7, BTA16, RNO13, MMU1). Since the MRCA of rodents and primates is believed to have lived more recently than the MRCA of Euarchontoglires and Laurasiatheria, conserved linkage in chickens and humans implies that there have been independent disruptions of synteny at almost identical sites in an ancestor of rats and mice and an ancestor of dogs and cattle.

Breakpoint C

Breakpoint C separates genes with orthologs on MMU13 from genes with orthologs on MMU8. Genes from Blocks III and IV, on either side of the breakpoint, have been mapped to a single bovine chromosome (BTA28), a single porcine chromosome (SSC14), and a single chicken chromosome (GGA3). Genes from both sides of the breakpoint have also been found on a single contig in *Fugu* [4]. These data strongly suggest that Blocks III and IV were linked in the MRCAs of rodents and primates, of Euarchontoglires and Laurasiatheria, and of birds and mammals. Comparative mapping between human and chicken argues strongly for this interpretation: *LYST* (on MMU13) and *IRF2BP2* (on MMU8) are separated by 1 Mb on HSA1 and the orthologous loci *LOC421514* and *LOC428585* are separated by only 140 kb on GGA3. The genes of Block IV map to RNO19, whereas the genes of Block III map to RNO17. Therefore, the disruption of conserved synteny occurred at some time in the rodent

lineage after it diverged from the primate lineage but before the divergence of rats and mice.

Breakpoint D

Breakpoint D separates genes with orthologs on MMU8 (Block IV) from genes with orthologs on MMU11 (Block V). Genes from Block IV map to BTA28, CFA4, and GGA3. Genes from Block V map to BTA7, CFA14, and GGA2. Thus, there is no evidence for an association of Blocks IV and V in any of these species.

The evidence summarized above suggests that breakpoints A and E represent a paracentric inversion that occurred in the human lineage, whereas breakpoints B and C represent "fissions" that occurred in the rodent lineage, after the two lineages diverged from their most recent common ancestor. Thus, the inferred order of these blocks prior to the paracentric inversion was centromere-VI-II-III-IV-V-I-telomere. Furthermore, the observation that genes from GGA3 span breakpoints B and C argues for conserved ancestral synteny of genes from Blocks II, III, and IV in humans and chickens. Blocks I and V both contain genes from MMU11. Therefore, Blocks I and V appear to have been linked in the MRCA of rodents and primates.

For how long have Blocks VI and V been linked? Comparative mapping on both sides of breakpoint D provides little evidence that Block (II + III + IV + VI) was linked to Block (I + V) in the MRCA of mice and humans. Moreover, comparative mapping suggests an alternative location for Block (I + V) in the MRCA of Euarchontoglires and Laurasiatheria. Genes from Block (I + V) map to MMU11, BTA7, and GGA2. Several loci from these three chromosomes also have orthologs on HSA5. Therefore, it is possible that Block (I + V) was linked to genes from HSA5 in the MRCAs of mice and humans and of Euarchontoglires and Laurasiatheria but was translocated to HSA1 at some time in the primate lineage. This translocation would have occurred before the divergence of Old World monkeys and apes because *GUK1* (from Block V) has been mapped to the equivalent of HSA1 in the rhesus macaque [5].

One piece of evidence may speak for an older association of genes from either side of breakpoint D. Genes from Block IV (*ACTA1*), Block V (*GUK1*), and Block I (*FLJ20531*) map to cat chromosome A1 [3]. However, the shared linkage of *ACTA1* and *GUK1* in cats and humans may be an example of coincidental shared synteny:

(1) Loci from HSA1 and HSA10q appear to have been linked in an early member of Laurasiatheria because genes from BTA28, SSC14, and CFA4 map to HSA10q, as well as to Blocks III and IV of HSA1 [6,7].
(2) Block (I + V) of HSA1 may have been linked to loci from HSA5 in an ancestral laurasiatherian (see above).
(3) HSA10q and HSA5 both contain loci with orthologs on CFA4, an association that appears to have arisen in the carnivore lineage.

Therefore, a carnivore-specific rearrangement may explain why *GUK1* from Block V and *ACTA1* from Block IV both map to cat chromosome A1, without their shared linkage in cats and humans being evidence of an ancestral association.

Discussion

Murphy and colleagues [3] concluded that HSA1 is a conserved linkage group that was present in the last common ancestor of eutherian mammals. This conclusion is correct for the bulk of HSA1, but my analysis suggests that the genes of Blocks I and V (≈4 Mb with orthologs on MMU11) may be a recent addition to this linkage group.

Block I contains a SCAN domain zinc finger protein (*ZNF496*) adjacent to a member of the NALP family (*CIAS1*). Tight linkage of genes for a SCAN domain zinc finger protein and a NALP protein also occurs on HSA11 (*ZNF215*, *NALP14*) and HSA19 (*ZNF444*, *NALP5*) and a less close association on HSA17 (*ZNF287*, *NALP1*). The large SCAN domain zinc finger family is absent from the chicken genome [8]. This suggests that an amplicon has dispersed linked SCAN and NALP genes to multiple sites in mammalian genomes since the mammalian lineage diverged from its MRCA with birds. The current assembly of CFA14 contains orthologs of flanking markers from Blocks I and V but does not contain orthologs of *ZNF496* or *CIAS1*. If this absence is real, rather than a problem of the assembly, it may provide evidence that *ZNF496* and *CIAS1* have been translocated into Block I since Euarchontoglires diverged from their MRCA with Laurasiatheria.

My analysis of the history of a small region of HSA1 revealed two kinds of obstacles to the reconstruction of ancestral chromosomes, namely the "reuse" of breakpoints and coincidental shared synteny. Although only five breaks in synteny were investigated in a small number of species, I found evidence for independent rearrangements at two of these five sites. First, a paracentric inversion in the human lineage involved a break between *ELYS* and *CDC42BPA* (breakpoint E). Synteny between human and chicken chromosomes is also disrupted at this point. Second, synteny appears to have been independently disrupted at breakpoint B in the lineages of rodents and of ferungulates (carnivores + artiodactyls). The evolutionary reuse of breakpoints has been previously noted [7,9,10]. Detection of repeated use of a breakpoint requires comparative mapping data from multiple species with a well-resolved phylogeny.

Murphy and colleagues [7] have recently attempted an ambitious reconstruction of the ancestral karyotypes of various mammalian ancestors. Their favored reconstruction of the ancestor of Boreoeutheria has the loci of HSA1 on two chromosomes (*contra* Murphy and colleagues [3]). The bulk of HSA1 is located on boreoeutherian ancestral chromosome 1. The remainder is associated with HSA10 on boreoeutherian ancestral chromosome 14. The same configuration is present in their reconstructed karyotype of the ancestral ferungulate. This division of HSA1 corresponds to breakpoint B in my analysis. The association of loci from distal HSA1q (Blocks III and IV) with loci from

HSA10 in an ancestral ferungulate concurs with my analysis. However, I propose that breakpoint B is not ancestral but occurred in the lineage leading to ferungulates. The difference in conclusions arises because my analysis uses an outgroup (*G. gallus*) that allowed me to detect the reuse of breakpoint B, whereas Murphy and colleagues' karyotype of the ancestral boreoeutherian did not use an outgroup, but instead used the midpoint (on an unrooted tree) between the reconstructed karyotypes of an ancestral ferungulate and the MRCA of mice and humans [7]. I favor a scenario in which Blocks III and IV became associated with loci of HSA10 in the ferungulate lineage *after* divergence from the boreoeutherian ancestor, because I can detect no continuity of chicken synteny between HSA1 and HSA10 (i.e., I am hypothesizing that *this* breakpoint has not been reused).

Haig [6] contrasted *conserved synteny* (markers are linked in two species because they were linked in a common ancestor and have remained linked in both lineages since this ancestor) and *coincidental shared synteny* (markers are syntenic in two species for reasons other than uninterrupted ancestral linkage). Coincidental shared synteny, if undetected, can lead to erroneous conclusions about ancestral linkage groups. For example, Haig [6] proposed that HSA7 is a conserved ancestral chromosome of primates, but it is now clear that this hypothesis is false [11]. I was misled because HSA7 contains two independent blocks of conserved synteny with MMU5, and these blocks have become jumbled together on HSA7 by a series of inversions. My analysis, in this paper, suggests that the linkage of *GUK1, ACTA1*, and *FLJ20531* in cattle and humans is another example of coincidental shared synteny that has contributed to a mistaken conclusion that Blocks I and V were ancestrally linked to the rest of HSA1 [3].

Materials and methods

The NCBI Mapviewer (http://www.ncbi.nlm.nih.gov/mapview/) was consulted to determine relative gene order in humans, chickens, dogs, and mice. NCBI Entrez Gene (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db = gene) provided information on additional gene locations. Orthologies between species were determined by common gene symbols or by similar gene descriptions with similar flanking markers. My conclusions are based on the state of genome assemblies in mid-2005. I hope that these were reasonably accurate.

References

[1] W.J. Murphy, et al., Resolution of the early placental mammal radiation using Bayesian phylogenetics, Science 294 (2001) 2348–2351.

[2] M. Scally, et al., Molecular evidence for the major clades of placental mammals, J. Mamm. Evol. 8 (2001) 239–277.

[3] W.J. Murphy, L. Frönicke, S.J. O'Brien, R. Stanyon, The origin of human chromosome 1 and its homologs in placental mammals, Genome Res. 13 (2003) 1880–1888.

[4] M.S. Taylor, R.S. Devon, J.K. Millar, D.J. Porteous, Evolutionary constraints on the Disrupted in Schizophrenia locus, Genomics 81 (2003) 67–77.

[5] S.J. O'Brien, J.A.M. Graves, Report of the committee on comparative gene mapping, Cytogenet. Cell Genet. 58 (1991) 1124–1151.

[6] D. Haig, A brief history of human autosomes, Philos. Trans. R. Soc. London B 354 (1999) 1447–1470.

[7] W.J. Murphy, et al., Dynamics of mammalian chromosome evolution inferred from multispecies comparative maps, Science 309 (2005) 613–617.

[8] International Chicken Genome Sequencing Consortium, Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution, Nature 432 (2004) 695–716.

[9] P. Pevzner, G. Tesler, Human and mouse genomic sequences reveal extensive breakpoint reuse in mammalian evolution, Proc. Natl. Acad. Sci. USA 100 (2003) 7672–7677.

[10] Y. Yue, B. Grossmann, M. Ferguson-Smith, F. Yang, T. Haaf, Comparative cytogenetics of human chromosome 3q21.3 reveals a hot spot for ectopic recombination in hominoid evolution, Genomics 85 (2004) 36–47.

[11] F. Richard, M. Lombard, B. Dutrillaux, Phylogenetic origin of human chromosomes 7, 16, and 19 and their homologs in placental mammals, Genome Res. 10 (2000) 644–651.