Common Long Human Inversion Polymorphism on Chromosome 8*p*

Karl W. Broman, Naomichi Matsumoto, Sabrina Giglio, Christa Lese Martin, Jessica A. Roseberry, Orsetta Zuffardi, David H. Ledbetter and James L. Weber

Abstract

In an analysis of human crossover interference, we identified apparent triple recombination events, in a short region on chromosome 8*p*, on the maternally-derived chromosomes in four individuals (two from each of two families). While this may have indicated an error in marker order, the inverted order was inconsistent with recombination events in other individuals. We were thus led to the hypothesis of an inversion polymorphism in the region, which was subsequently confirmed by fluorescent *in situ* hybridization (FISH). The inversion spans approximately 12 cM on the female genetic map and 2.5 – 5.3 Mb on the physical map. The allele frequency of the inverted order (D8S1130 telomeric; D8S351 centromeric) in 50 individuals of European ancestry was 21%. This is only the second known common, long inversion polymorphism in the human genome.

Keywords: CEPH; FISH; inversion; polymorphism

1 Introduction

Inversions in gene order along chromosomes have frequently been observed by comparing related species [14, 24, 25], including great apes [16, 21, 22, 29]. Human inversion mutations occur at a low, but detectable frequency. Paracentric (not involving the centromere) inversions that are large enough to be detectable by standard cytogenetic analysis occur at a frequency of 1 – 5 per 10,000 individuals [23]. The frequency of human submicroscopic inversions is unknown, although inversions have been identified as the cause of specific heritable disorders (see, for example, [1, 8, 15, 19, 20]). Chromosomal inversions are of particular clinical interest because recombination within the inverted region in heterozygotes can lead to segmental aneusomies and concomitant abnormalities.

The only well characterized common human inversion polymorphism is the 48 kb inversion of the Emery-Dreifuss muscular dystrophy and filamin genes on the X chromosome [26]. This inversion is present in populations of European descent at a frequency of about 18%. Page and colleagues also recently made a preliminary report of a potentially common 3 Mb inversion polymorphism on chromosome Y*p* flanked

by inverted 300 kb repeats [27]. Here we describe a common, paracentric inversion polymorphism spanning > 2.5 Mb in chromosome band $8p23.1 - 8p22$.

2 Materials and Methods

We considered high-density genotype data on eight of the CEPH reference families [6]. These families, which were recruited in order to form the first human genetic maps, are largely three-generation families, with 10–15 siblings each. They have been genotyped at $>$ 8,000 short tandem repeat polymorphisms (STRPs, also known as microsatellites). The genotype data are publicly available (see the Marshfield web site, `http://research.marshfieldclinic.org/genetics`).

Initial marker order was taken from [3]. Haplotypes were constructed with use of the *chrompic* option of the CRI-MAP program [12]. The physical length of the inverted region was estimated based on the December 22, 2001, version of the University of California, Santa Cruz, draft human sequence (see `http://genome.ucsc.edu`).

Fluorescent *in situ* hybridization (FISH) was carried out as previously described [5]. A minimum of five spreads were examined for each individual. BAC clones were obtained from Genome Systems (St. Louis, Missouri, USA).

3 Results and Discussion

In an examination of the sites of meiotic recombination in eight of the CEPH reference families, as part of an analysis of human crossover interference [4], we observed that the maternally inherited chromosomes in two offspring from each of CEPH families 1362 and 1413 show similar and highly unlikely crossover patterns (Figure 1). Even in the absence of crossover interference, the probability of four triple crossovers within 12 cM is vanishingly small. All four of these chromosomes revert to single crossovers when the region between and including D8S351 and D8S1130 is inverted. However, the inverted order of markers is inconsistent with recombination events in other CEPH families (see Figure 1).

Fluorescent *in situ* hybridization (FISH) was used to confirm and extend the initial evidence for inversion. BAC clones encompassing D8S351 and D8S1130 near the ends of the inverted segment (see Figure 1) were used as probes. We arbitrarily defined the normal allele as having the marker order in Figure 1, and the inverted allele as having the markers between D8S351 and D8S1130 inverted. Shown in Figure 2 are representative metaphase results from two individuals with each of the three possible genotypes, including the mother (1362-02) (panel c) of CEPH family members of 1362-10 and 1362-11, who is homozygous for the inverted order (relative to the order shown in Figure 1). CEPH individuals 1362-10, 1413-03, and 1413-02 (mother of 1413-03 and 1413-09) were also found to be homozygous for the inverted order (data not shown).

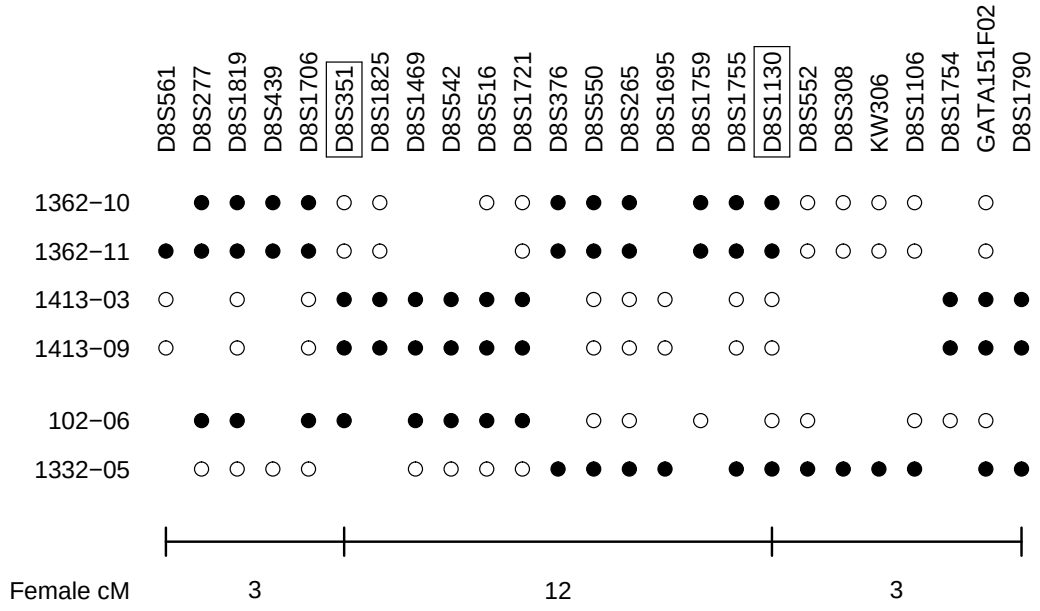

Figure 1: Maternal haplotypes for a small portion of chromosome 8*p* for six CEPH family children (identified by family – individual). Filled symbols indicate alleles from the maternal grandfather, open symbols alleles from the maternal grandmother, and blank spaces indicate missing data (due mostly to homozygous markers in the mother). The order of markers is telomeric (left) to centromeric (right). BACs encompassing the two markers shown in boxes were used in the FISH experiments.

Metaphase FISH carried out on 50 unrelated individuals of European ancestry revealed 33 homozygotes with the order shown in Figure 1, 13 heterozygotes, and 4 homozygotes for the inverted order (inversion frequency 21%; 95% confidence interval, assuming Hardy-Weinberg equilibrium, 13 – 30%). The genotype frequencies showed no significant deviation from Hardy-Weinberg equilibrium.

The inversion polymorphism appears to be either extremely old or the result of recurrent mutations. With only a single, relatively recent, inversion mutation event, and assuming no recombination in heterozygotes, there should only be one common "inverted" haplotype. With at least two inversion events occurring upon different haplotype backgrounds, recombination events in parents homozygous for the inversion could produce many different haplotypes. Although we don't know which orientation is ancestral, construction of haplotypes in the CEPH families using available genotyping data revealed several different haplotypes for each orientation. All three haplotypes for the order shown in Figure 1 were quite different with multiple (up to 17) repeat differences between alleles (data not shown). Similarly, all six haplotypes for the inverted order were very different. Since short tandem repeats (microsatellites) nearly always mutate by gain or loss of one or two repeat units [2, 28], it is unlikely that a single in-

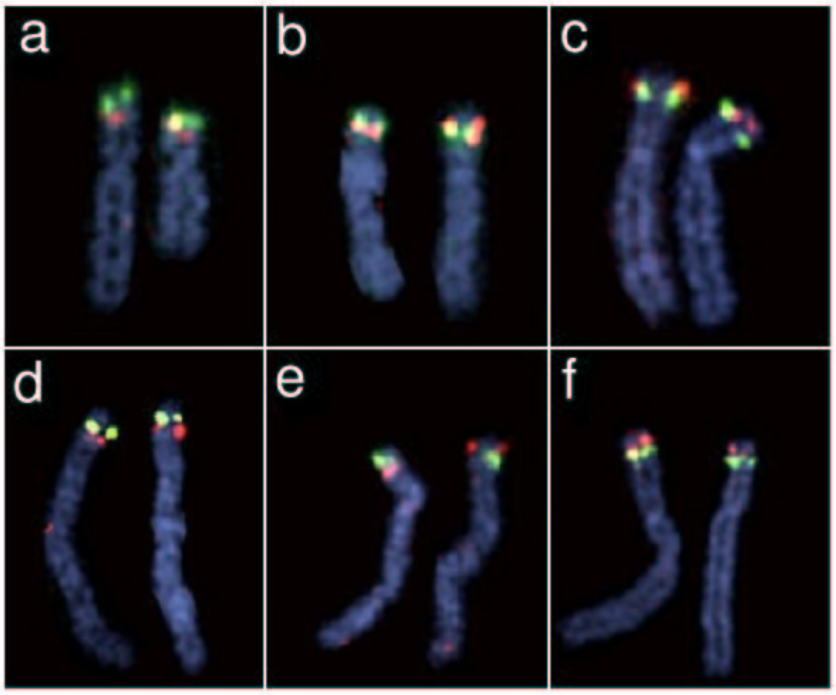

Figure 2: Metaphase FISH results from CEPH family individuals (lymphoblastoid cell lines) and other individuals (peripheral blood lymphocytes). Probes were DNA from BAC 173O4 (encompassing D8S351) labeled with Spectrum Green and BAC 257O3 (encompassing D8S1130) labeled with Spectrum Orange. a) CEPH individual 102-01, homozygous for the order shown in Figure 1. b) CEPH individual 1331-02, heterozygous for the inversion. c) CEPH individual 1362-02, homozygous for the inverted order. d) e) f) Individuals homozygous for the Figure 1 order, heterozygous, and homozygous for the inverted order, respectively.

version mutation event occurred within the last few hundred thousand years. Although we cannot rule out a single event that occurred longer ago, we favor the alternative that at least two inversion mutation events occurred relatively recently.

We further characterized the inversion through examination of relevant genome maps. The genetic length of the inverted region is approximately 12 and 2 cM on the female and male genetic maps, respectively [3]. Using the December 22, 2001, version of the University of California-Santa Cruz draft human sequence, the length of the inverted region was estimated to be at least 2.5 Mb and possibly as long as 5.3 Mb. The sequence assembly in this region of 8*p* is still crude with many gaps, both large and small, and other uncertainties. Sites of the inversion breakpoints are not yet precisely known. From both the CEPH and FISH results, the inversion breakpoints appear to be at similar locations in all individuals; however, the precision of these approaches is limited.

The inversion is likely mediated by two clusters of olfactory receptor genes that flank the inverted segment at both ends [9]. Olfactory receptor genes are found on nearly every human chromosome [11]. The flanking repeated sequences are apparently in inverted orientation (Matsumoto *et al.*, in preparation). The 48 kb emerin/filamin inversion on the X chromosome is also flanked by 11 kb inverted repeat sequences [26]. Intrachromatid recombination between inverted non-adjacent repeat sequences results in the inversion of the intervening segment. As the human genomic sequence becomes finished, it may be possible to identify additional inversion polymorphisms through searches for intrachromosomal inverted repeats with high sequence similarity.

The 8*p* inversion may have substantial clinical impact. For example, Giglio *et al*. [9] studied eight mothers of children with the inverted duplication 8it p rearrangement, and found that all were heterozygous for the inversion described herein. Inv dup (8*p*) is a well-known chromosomal abnormality of maternal origin that causes multiple abnormalities including mental retardation [7, 13, 17]. The frequency of inv dup (8*p*) has been estimated at 1/15,000 [9]. It may be that women heterozygous for the chromosome 8*p* inversion that we identified are more likely to bear children with the inv dup (8*p*) rearrangement. Giglio *et al*. [10] recently identified another inversion polymorphism, on chromosome 4p16, which is also flanked by clusters of olfactory receptor genes. These two chromosomal inversions, on chromosomes 4 and 8, appear to be involved in the recurrent t(4;8)(p16;p23) translocation.

Heterozygotes for the chromosome 8*p* inversion may also have slightly reduced fertility compared to homozygotes of either genotype due to unbalanced gametes produced through recombination within the inverted region. The rearrangement may also affect the expression of genes near the inversion breakpoint. Such effects are well known for translocations [18]. Genes within or adjacent to the inverted segment include several defensins, GATA-binding protein 4 (GATA4), cathepsin B (CTSB), tankyrase (TNKS), and methionine sulfoxide reductase A (MSRA).

Submicroscopic inversions are difficult to identify. Use of improbable meiotic products as an inversion signature (see Figure 1) becomes much more difficult as the size of the inversion decreases. For inversion of only two or three adjacent markers, the phase patterns will masquerade as genotyping errors or mutations. Also, a recombination event is required within the inverted region for detection, and it may be necessary for the parent to be homozygous for the inversion for recombination to occur. A better approach to detect inversion polymorphisms is likely to be comparison of various genome maps, especially including sequence assemblies, which are prepared using DNA from different donors. Our results clearly demonstrate that differences in marker order between various genome maps should not automatically be dismissed as errors.

Acknowledgments

This work was supported by NIH grants N01–HV–48141 to JLW and R01–HD–36111 to DHL.

Karl W. Broman, Center for Medical Genetics, Marshfield Medical Research Foundation (current address: Department of Biostatistics, Johns Hopkins University),
kbroman@jhsph.edu

Naomichi Matsumoto, Department of Human Genetics, University of Chicago (current address: Department of Human Genetics, Nagasaki University School of Medicine)

Sabrina Giglio, Department of Human Genetics, University of Chicago (current address: Diagnostica e Ricerca, San Raffaele, Milan)

Christa Lese Martin, Department of Human Genetics, University of Chicago

Jessica A. Roseberry, Department of Human Genetics, University of Chicago

Orsetta Zuffardi, Biologia Generale e Genetica Medica, Università di Pavia

David H. Ledbetter, Department of Human Genetics, University of Chicago

James L. Weber, Center for Medical Genetics, Marshfield Medical Research Foundation

References

[1] M. L. Bondeson, N. Dahl, H. Malmgren, W. J. Kleijer, T. Tonnesen, B. M. Carlberg, and U. Pettersson. Inversion of the IDS gene resulting from recombination with IDS-related sequences is a common cause of the Hunter syndrome. *Human Molecular Genetics*, 4:615–621, 1995.

[2] B. Brinkmann, M. Klintschar, F. Neuhuber, J. Huhne, and B. Rolf. Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. *American Journal of Human Genetics*, 62:1408–1415, 1998.

[3] K. W. Broman, J. C. Murray, V. C. Sheffield, R. L. White, and J. L. Weber. Comprehensive human genetic maps: individual and sex-specific variation in recombination. *American Journal of Human Genetics*, 63:861–869, 1998.

[4] K. W. Broman and J. L. Weber. Characterization of human crossover interference. *American Journal of Human Genetics*, 66:1911–1926, 2000.

[5] S. S. Chong, S. D. Pack, A. V. Roshke, A. Tanigami, R. Carrozzo, A. C. M. Smith, W. B. Dobyns, and D. H. Ledbetter. A revision of the lissencephaly and Miller-Dieker syndrome critical regions in chromosome 17p13.3. *Human Molecular Genetics*, 6:147–155, 1997.

[6] J. Dausset, H. Cann, D. Cohen, M. Lathrop, J. M. Lalouel, and R. White. Centre d'Etude du Polymorphisme Humain (CEPH): collaborative genetic mapping of the human genome. *Genomics*, 6:575–577, 1990.

[7] C. E. M. de Die-Smulders, J. J. Engelen, C. T. Schrander-Stumpel, L. C. Govaerts, B. de Vries, J. S. Vles, A. Wagemans, S. Schijns-Fleuren, G. Gillessen-Kaesbach, and J.-P. Fryns. Inversion duplication of the short arm of chromosome 8: clinical data on seven patients and review of the literature. *American Journal of Medical Genetics*, 59:369–374, 1995.

[8] Y. J. M. de Kok, G. F. M. Merkx S. M. van der Maarel, I. Huber, S. Malcolm, H. H. Ropers, and F. P. M. Cremers. A duplication/paracentric inversion associated with familial X-linked deafness (DFN3) suggests the presence of a regulatory element more than 400 kb upstream of the POU3F4 gene. *Human Molecular Genetics*, 4:2145–2150, 1995.

[9] S. Giglio, K. W. Broman, N. Matsumoto, V. Calvari, G. Gimelli, T. Neumann, H. Ohashi, L. Voullaire, D. Larizza, R. Giorda, J. L. Weber, D. H. Ledbetter, and O. Zuffardi. Olfactory receptor gene clusters, genomic-inversion polymorphisms, and common chromosome rearrangements. *American Journal of Human Genetics*, 68:874–883, 2001.

[10] S. Giglio, V. Calvari, G. Gregato, G. Gimelli, S. Camanini, R. Giorda, A. Ragusa, S. Guerneri, A. Selicorni, M. Stumm, H. Tonnies, M. Ventura, M. Zollino, G. Neri, J. Barber, D. Wieczorek, M. Rocchi, and O. Zuffardi. Heterozygous submicroscopic inversions involving olfactory recepter-gene clusters mediate the recurrent t(4;8)(p16;p23) translocation. *American Journal of Human Genetics*, 71:276–285, 2002.

[11] G. Glusman, I. Yanai, I. Rubin, and D. Lancet. The complete human olfactory subgenome. *Genome Research*, 11:685–702, 2001.

[12] P. Green, K. Falls, and S. Crooks. *Documentation for CRI-MAP, version 2.4*, 1990.

[13] W. J. Guo, F. Callif-Daley, M. C. Zapata, and M. E. Miller. Clinical and cytogenetic findings in seven cases of inverted duplication of 8p with evidence of a telomeric deletion using fluorescence *in situ* hybridization. *American Journal of Medical Genetics*, 58:230–236, 1995.

[14] M. A. Huynen, B. Snel, and P. Bork. Inversions and the dynamics of eukaryotic gene order. *Trends in Genetics*, 17:304–306, 2001.

[15] M. A. Jobling, G. Williams, K. Schiebel, A. Pandya, K. McElreavey, L. Salas, G. A. Rappold, N. A. Affara, and C. Tyler-Smith. A selective difference between human Y-chromosomal DNA haplotypes. *Current Biology*, 8:1391–1394, 1998.

[16] H. Kehrer-Sawatzki, B. Schreiner, S. Tanzer, M. Platzer, S. Muller, and H. Hameister. Molecular characterization of the pericentric inversion that causes differences between chimpanzee chromosome 19 and human chromosome 17. *American Journal of Human Genetics*, 71:375–388, 2002.

[17] A. Kleczkowska, J.-P. Fryns, F. D'Hondt, J. Jaeken, and H. van den Berghe. Partial duplication 8p due to interstitial duplication: inv dup (8)(p21.1 – p22). *Annals of Genetics*, 30:47–51, 1987.

[18] D.-J. Kleinjan and V. van Heyningen. Position effect in human genetic disease. *Human Molecular Genetics*, 7:1611–1618, 1998.

[19] K. Lagerstedt, S. L. Karsten, B. M. Carlberg, W. J. Kleijer, T. Tonnesen, U. Pettersson, and M. L. Bondeson. Double-strand breaks may initiate the inversion mutation causing the Hunter syndrome. *Human Molecular Genetics*, 6:627–633, 1997.

[20] D. Lakich, H. H. Kazazian, S. E. Antonarakis, and J. Gitschier. Inversions disrupting the factor VIII gene are a common cause of severe haemophilia A. *Nature Genetics*, 5:236–241, 1993.

[21] S. Muller, R. Stanyon, P. Finelli, N. Archidiacono, and J. Wienberg. Molecular cytogenetic dissection of human chromosomes 3 and 21 evolution. *Proceedings of the National Academy of Sciences USA*, 97:206–211, 2000.

[22] E. Nickerson and D. L. Nelson. Molecular definition of pericentric inversion breakpoints occurring during the evolution of humans and chimpanzees. *Genomics*, 50:368–372, 1998.

[23] M. J. Pettenati, P. N. Rao, M. C. Phelan, F. Grass, K. W. Rao, P. Cosper, A. J. Carroll, F. Elder, J. L. Smith, M. D. Higgins, J. T. Lanman, R. R. Higgins, M. G. Butler, F. Luthardt, E. Keitges, C. Jackson-Cook, J. Brown, S. Schwartz, D. L. Van Dyke, and C. G. Palmer. Paracentric inversions in humans: a review of 446 paracentric inversions with presentations of 120 new cases. *American Journal of Medical Genetics*, 55:171–187, 1995.

[24] M. T. Pletcher, T. Wiltshire, D. E. Cabin, M. Villanueva, and R. H. Reeves. Use of comparative physical and sequence mapping to annotate mouse chromosome 16 and human chromosome 21. *Genomics*, 74:45–54, 2001.

[25] J. M. Ranz, F. Casals, and A. Ruiz. How malleable is the eukaryotic genome? Extreme rate of chromosomal rearrangement in the genus Drosophila. *Genome Research*, 11:230–239, 2001.

[26] K. Small, J. Iber, and S. T. Warren. Emerin deletion reveals a common X-chromosome inversion mediated by inverted repeats. *Nature Genetics*, 16:96–99, 1997.

[27] C. A. Tilford, T. Kuroda-Kawaguchi, H. Skaletsky, S. Rozen, L. G. Brown, M. Rosenberg, J. D. McPherson, K. Wylie, M. Sekhon, T. A. Kucaba, R. H. Waterston, and D. C. Page. A physical map of the human Y chromosome. *Nature*, 409:943–945, 2001.

[28] X. Xu, M. Peng, Z. Fang, and X. Xu. The direction of microsatellite mutations is dependent upon allele length. *Nature Genetics*, 24:396–399, 2000.

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