

# Divergent outcomes of intrachromosomal recombination on the human Y chromosome: male infertility and recurrent polymorphism

Patricia Blanco, Maria Shlumukova, Carole A Sargent, Mark A Jobling, Nabeel Affara, Matthew E Hurles

## Abstract

**The Y chromosome provides a unique opportunity to study mutational processes within the human genome, decoupled from the confounding effects of interchromosomal recombination. It has been suggested that the increased density of certain dispersed repeats on the Y could account for the high frequency of causative microdeletions relative to single nucleotide mutations in infertile males. Previously we localised breakpoints of an *AZF<sub>a</sub>* microdeletion close to two highly homologous complete human endogenous retroviral sequences (HERV), separated by 700 kb. Here we show, by sequencing across the breakpoint, that the microdeletion occurs in register within a highly homologous segment between the HERVs. Furthermore, we show that recurrent double crossovers have occurred between the HERVs, resulting in the loss of a 1.5 kb insertion from one HERV, an event underlying the first ever Y chromosomal polymorphism described, the 12f2 deletion. This event produces a substantially longer segment of absolute homology and as such may result in increased predisposition to further intrachromosomal recombination. Intrachromosomal crosstalk between these two HERV sequences can thus result in either homogenising sequence conversion or a microdeletion causing male infertility. This represents a major subclass of *AZF<sub>a</sub>* deletions.**

(J Med Genet 2000;37:752-758)

Keywords: HERV; *AZF<sub>a</sub>*; infertility; Y chromosome

Increasing numbers of genetic disorders are being shown to result from the recombinogenic effect of flanking repeated sequences. Microdeletions, and more recently their reciprocal recombination product, duplications, have been reported to cause neurofibromatosis type 1,<sup>1</sup> Smith-Magenis syndrome,<sup>2</sup> and Charcot-Marie-Tooth disease, among other diseases.<sup>3</sup> Unequal exchange between repeats on homologous chromosomes (UCE) or sister chromatids (USCE) are the proposed mechanisms by which these rearrangements may occur. The pathogenic role of dispersed repeats in sponsoring illegitimate recombination, though documented, has been overshadowed by that of longer region specific paralogous duplications.<sup>3</sup> Investigations into the length of sequence

identity required to initiate homologous recombination in meiosis has led to the concept of a minimum efficient processing segment (MEPS). In mammalian cells, the MEPS is thought to be at least 200 bp in length.<sup>4</sup>

Human endogenous retroviral sequences (HERV) are a major subclass of dispersed repeats, accounting for about 1% of the human genome.<sup>5</sup> The remnants of ancient germ cell retroviral infections that have been multiply transposed after their original integration, partial and complete HERV proviral sequences are widely distributed throughout the genome.<sup>6</sup> The degree of sequence divergence between superfamilies of HERV varies markedly and is thought to correlate with the time since major waves of amplification.<sup>7</sup> Until now HERV sequences have not been associated with pathogenic illegitimate recombination.

The human Y chromosome has been recently shown to contain certain dispersed repeats, including Alu and HERV sequences, at a significantly higher frequency than the autosomes.<sup>8</sup> It has been suggested that deletions mediated by this higher frequency of dispersed repeats may be the major mode of pathogenic mutation on the Y chromosome.<sup>9</sup> It has also been suggested that the history of inversions on the Y chromosome that characterised its divergence from the X also results from this higher density of repeats.<sup>10 11</sup>

The *AZF<sub>a</sub>*, *AZF<sub>b</sub>*, and *AZF<sub>c</sub>* (azoospermic factor) regions on the long arm of the human Y chromosome have been defined by deletion intervals determined by screening azoospermic and oligozoospermic patients with a variety of probes and STSs from the Y chromosome<sup>12</sup> (fig 1). These three loci are involved in infertility phenotypes ranging from oligozoospermia to Sertoli cell only syndrome.<sup>12</sup> The *AZF<sub>a</sub>* region (primarily associated with Sertoli cell only syndrome) comprises around 1100 kb that have been fully sequenced and contain at least three genes, *DFFRY* (also known as *USP9Y*), *DBY*, and *UTY*.<sup>13</sup> *UTY* has been excluded from involvement in the phenotype, whereas it has been suggested that both *DFFRY* and *DBY* are involved in infertility, with severe impairment of spermatogenesis occurring when the two genes are defective or deleted.<sup>13</sup> A de novo point mutation in *DFFRY* has been described in a single oligozoospermic patient, but microdeletions remain by far the most common pathogenic lesion.<sup>14</sup>

The haploid Y chromosome does not recombine at meiosis over most of its length,

Human Molecular Genetics Group, University of Cambridge, Department of Pathology, Tennis Court Road, Cambridge CB2 1QP, UK

P Blanco  
C A Sargent  
N Affara

Department of Genetics, University of Leicester, University Road, Leicester LE1 7RH, UK  
M Shlumukova\*  
M A Jobling

Molecular Genetics Laboratory, McDonald Institute for Archaeological Research, University of Cambridge, Downing Street, Cambridge CB2 3ER, UK  
M E Hurles

Correspondence to:  
Dr Hurles,  
meh32@cam.ac.uk

\*Present address: School of Biomedical Sciences, Queens Medical Centre, Nottingham, NG7 2UH, UK.

Revised version received  
30 June 2000  
Accepted for publication 11  
July 2000

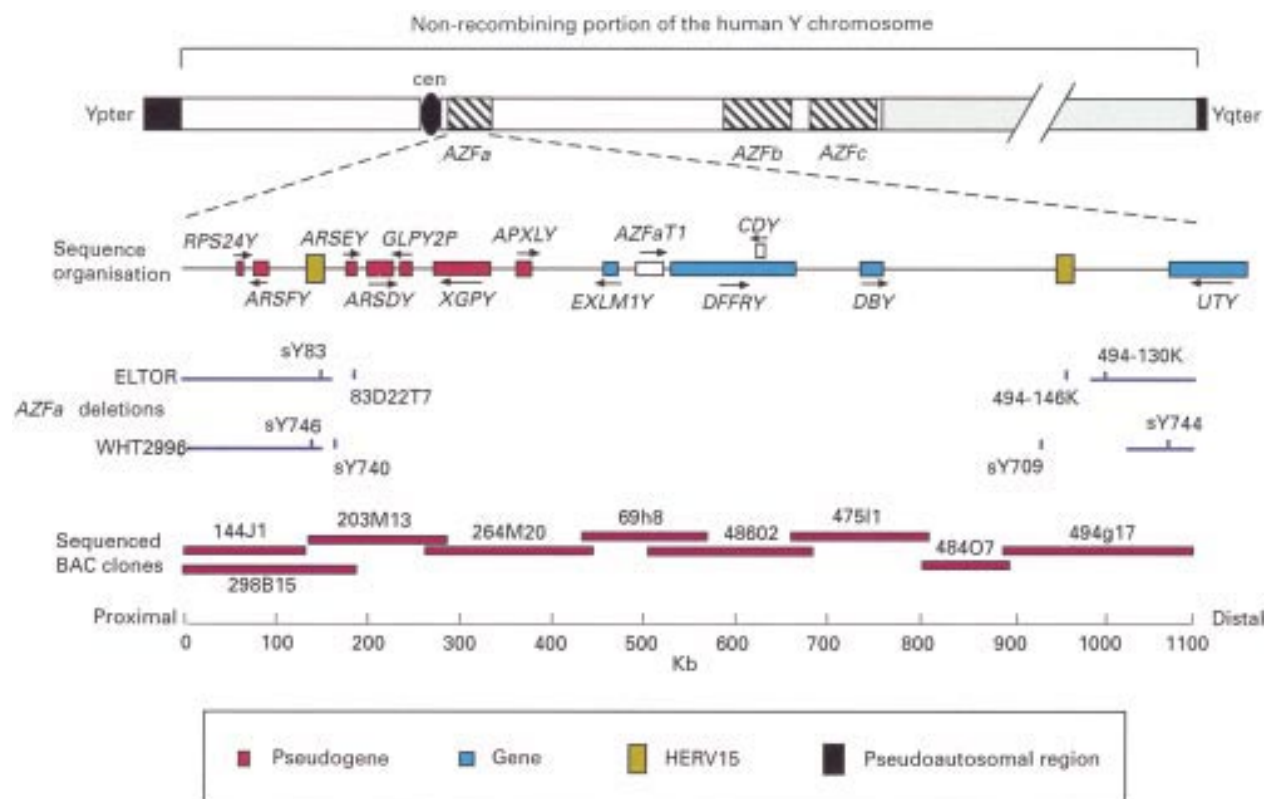


Figure 1 This figure illustrates the AZFa region of the long arm of the Y chromosome. Underneath the labelled ideogram of the human Y chromosome is an expanded view of the sequence organisation of the AZFa region including the genes and pseudogenes mapped previously<sup>13</sup> and the HERV15 sequences identified by use of the NIX program (see Materials and methods). Beneath are positioned the previously determined AZFa deletion breakpoints from two patients (ELTOR<sup>13</sup> and WHT2996<sup>14</sup> (see Discussion)) and the STS markers that span the breakpoints. Finally, the organisation of sequenced BAC clones covering 800 kb of the AZFa region<sup>14</sup> is superimposed on a kilobase scale.

and thus it contains within it haplotypes that are changed solely by mutation. Such haplotypes contain a simple record of our evolutionary past and have been exploited to investigate many facets of human population prehistory.<sup>15</sup> Non-recombining haplotypes constructed from unique mutations can be used to construct a perfect phylogeny, without reticulation. In the case of the Y chromosome, in contrast to mitochondrial DNA, single nucleotide polymorphisms (SNPs) and certain insertion/deletion polymorphisms (indels) can be considered to be unique on the basis of their producing just such a phylogeny.<sup>16</sup> Subsequently, putative recurrences of mutational events can be investigated by mapping them onto this single most parsimonious phylogeny.<sup>17, 18</sup> In addition, the Y chromosome has proven to be a useful model system for investigating intra-

chromosomal recombination mechanisms in isolation.<sup>19</sup>

The first Y linked polymorphism was discovered in 1985, by probing restriction enzyme digests of genomic DNA with a 2.3 kb *Bgl*II fragment from the p12f clone.<sup>20</sup> From Southern hybridisations of *Taq*I and *Eco*RI digests it was hypothesised that a deletion of roughly 2 kb was the event underlying the polymorphism. The Y chromosomal lineage defined by the 12f2 deletion is of interest for studies of prehistoric migrations and is found at highest frequencies (greater than 25%) in Middle Eastern, southern European, North African, and Ethiopian populations.<sup>21</sup>

Here we show that a major subclass of AZFa microdeletions results from a single crossover between two well separated, highly homologous HERV sequences. We also show that a double crossover event between these two HERV copies also underlies the 12f2 deletion, which, contrary to previous thought, can be shown to be recurrent in human evolution. We hypothesise that the sequences generated by these double crossovers may predispose to further single crossovers.

## Materials and methods

### SEQUENCE ANALYSIS

The BACs containing the deletion breakpoint intervals identified previously in ELTOR (AC002992 and AC005820) were analysed using the NIX programme at the HGMP resource centre (<http://www.hgmp.mrc.ac.uk/>)

Table 1 PCR primers used in this study

Primer	Sequence (5'-3')	T <sub>anneal</sub> (°C) long PCR/nested PCR
494-130kP1	gaggaaatacggatttgggg	58
SY83F	cttgaatcaagaagccct	58/60
494-146kP2	atggctcatccaactgag	58
ELT2	cctgtaacagataatgaatggg	58/60
12f2E	gttcttgactgggaagatctcag	58/65
12f2J	gtgccttccaagcattggag	65
12f2L	ttaattcagccctctgagcg	58
12f2M	tatggtctctatggaagtagg	58
12f2N	aaagcctgtgctcagcctgag	58
12f2D	ctgactgatcaaaatgcttacagatc	See Materials and methods
12f2F	tcctctagaattcttcacagaattg	See Materials and methods
3'Sry15	cttgattttctgtagaacaag	See Materials and methods
3'Sry16	tgctgttacataaatgggcac	See Materials and methods

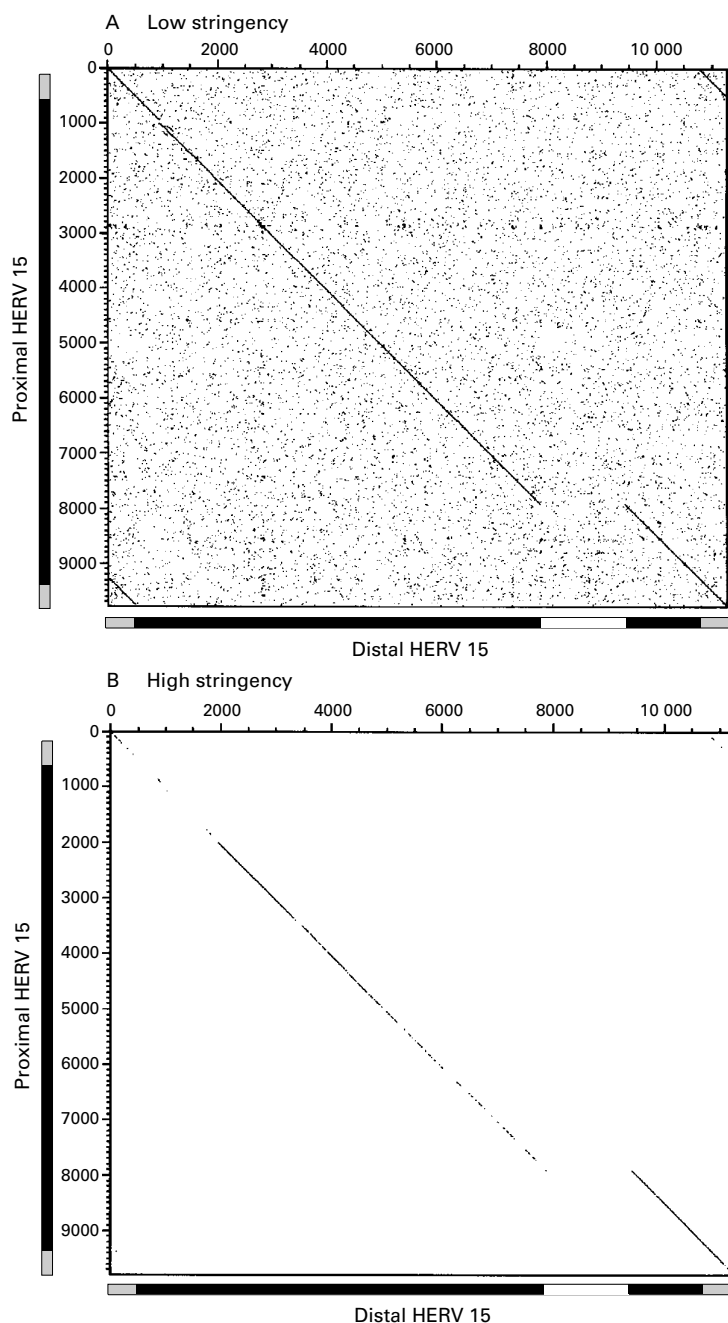


Figure 2 Dot plots of the proximal HERV (X axis) against the distal HERV (Y axis). The LTR15 sequences are shown in grey, HERV15 in black, and the L1 insertion in white. (A) Low stringency dot plot, (B) high stringency dot plot showing regions of greatest homology.

and the ENSEMBL database of automatically annotated eukaryotic sequence (<http://ensembl.ebi.ac.uk>). The consensus sequence of HERV15 was obtained from the latest RepBase update (<http://www.girinst.org>). Alignments were performed using CLUSTALW (<http://www.ebi.ac.uk/clustalw/>). Restriction analysis was performed using TACG (<http://genzi.virus.kyoto-u.ac.jp/tacg/tacg.form.html>) or Webcutter (<http://www.ccsi.com/firstmarket/cutter/cut2.html>). Nucleotide sequence databases were searched with BLAST 2.0 (<http://www.ncbi.nlm.nih.gov/BLAST>). Dot plots were produced by the program Dotter at the HGMP resource centre.

#### PCR

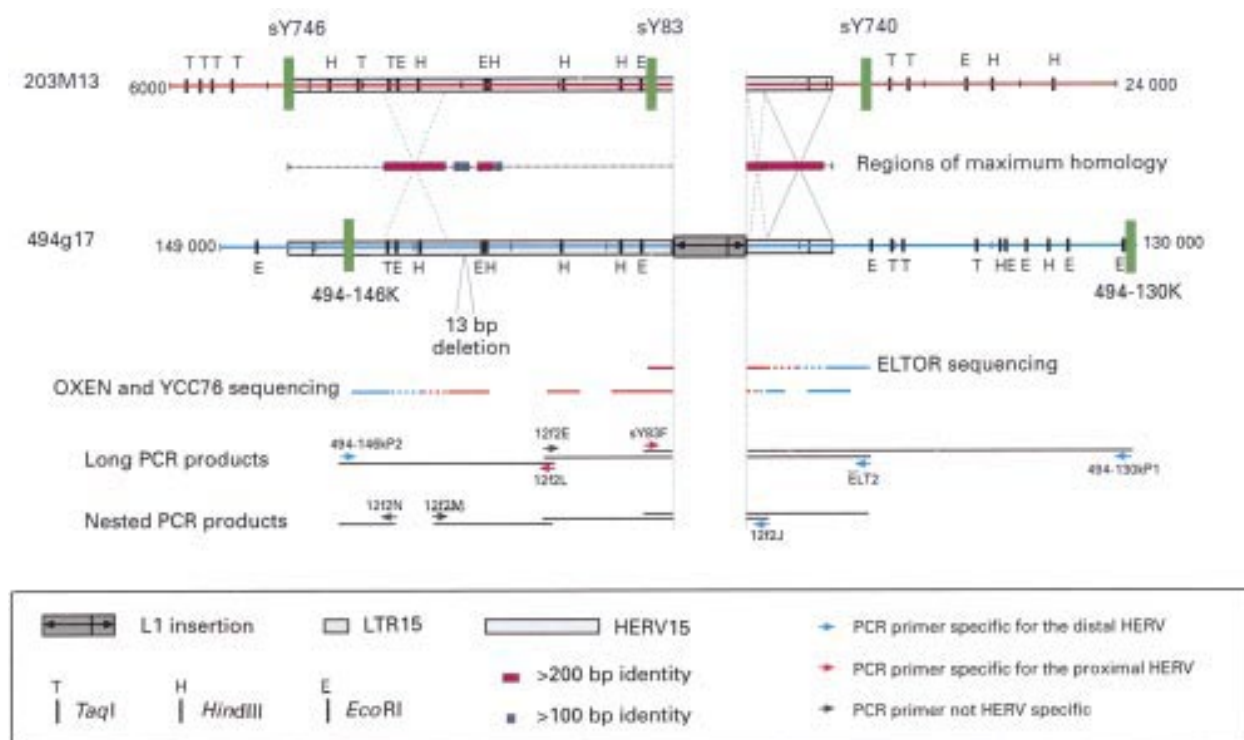
Long PCR was performed on 50 ng of genomic DNA either using the Dynazyme EXT kit from Flowgen or the Extensor long PCR system kit from Advanced Biotechnologies using conditions stipulated by the manufacturer. The primer pairs (MWG-Biotech) used are listed in table 1. Cycling conditions used with the Dynazyme EXT kit were 94°C for two minutes followed by 10 cycles of 94°C for 10 seconds, 58°C for 30 seconds, 68°C for 1-1.5 minutes per kb, followed by 24 cycles of 94°C for 10 seconds, 58°C for 30 seconds, 68°C for 1-1.5 minutes per kb plus 20 seconds per cycle, and a final extension step of 68°C for seven minutes. Cycling conditions for the breakpoint PCR with ELTOR DNA using the Advanced Biotechnologies kit were 94°C for two minutes followed by 10 cycles of 94°C for 10 seconds, 58°C for 30 seconds, 60°C for eight minutes, followed by 20 cycles of 94°C for 10 seconds, 58°C for 30 seconds, 68°C for eight minutes plus 20 seconds per cycle, and a final extension step of 68°C for seven minutes.

Nested PCR to generate sequencing template was performed using the primers in table 1 and 1 U of *Taq* polymerase (Geneo BioProducts). PCR was carried out in a volume of 20  $\mu$ l with 1 pmol/ $\mu$ l of each primer, 50 mmol/l Tris-HCl (pH 9.0), 15 mmol/l  $(\text{NH}_4)_2\text{SO}_4$ , 0.1% (v/v) Triton X-100, 2 mmol/l of  $\text{MgCl}_2$ , and 360  $\mu$ mol/l of each dNTP. The cycling conditions were 94°C for two minutes followed by 14-24 cycles of 94°C for 10 seconds, 58-65°C for 30 seconds, 68°C for 1-1.5 minutes per kb, followed by a final extension step of 68°C for seven minutes. Template for nested PCR was generated from long PCR products by running them on a 0.9% agarose (Geneo BioProducts) gel, cutting out the relevant band and incubating the band in 400  $\mu$ l of water overnight at 37°C; 2  $\mu$ l of the eluate were then used in each PCR.

#### 12F2 ASSAY

The 12f2 deletion polymorphism was typed using a newly developed assay. The primers are listed in table 1, and were designed from partial sequence data of the cosmid M13A12, isolated using 12f2 as a probe, from the Y chromosome specific library LL0YNC03. Primers 12f2D and 12f2F generate a specific 500 bp amplicon and primers 3'Sry15 and 3'Sry16 a 820 bp control amplicon which is present in all Y chromosomes. PCR conditions were 33-35 cycles of 94°C for 30 seconds, 59°C for 30 seconds, and 72°C for 45 seconds. The concordance of the PCR assay with the original hybridisation assay was indicated by showing that 23 chromosomes known from previous analysis to carry *TaqI*/8 kb alleles lacked the 12f2 test amplicon in this assay, whereas 23 subjects known to carry *TaqI*/10 kb alleles did not. The previously unknown recurrence of the deletion polymorphism was discovered during the screening of a large panel of Y chromosomes which had previously been typed for both the YAP and SRY-1532 polymorphisms (Z H Rosser, M E Hurles, M A Jobling, unpublished observations).





**Figure 3** Schematic alignment of HERV15 sequences. The 18 kb sequence encompassing the proximal HERV15 sequence is shown in red, aligned to the 19 kb sequence encompassing the distal HERV15 sequence in blue. Between them, indicated by a dashed line, is the region of homology upon which intervals of identity greater than 100 bp are shown in dark blue, and intervals of identity greater than 200 bp in purple. Positions of the displayed sequences within their BACs are shown at the ends. The STS discussed in the text are shown as vertical green bars. Digestion sites for three restriction enzymes are shown as short vertical black lines. The position of the 13 bp deletion discussed in the text is also shown. Sequence encompassing the distal conversion breakpoint is deposited in GenBank under accession number AJ278655. Diagonal lines between the sequences indicate the positions of crossovers: dashed, 12f2 deletion; solid, AZFa microdeletion. The PCR products and sequences discussed in Results are shown below. Sequence is colour coded for which HERV it is specific to and dashed lines indicate regions of identity where sequence cannot be assigned to either HERV. Sequence encompassing the distal conversion breakpoint is deposited in GenBank under accession number AJ278654 and sequence containing the AZFa deletion breakpoint is under accession number AJ278656. PCR primer sequences are indicated by small arrows and their sequences given in table 1.

#### CLONING

Restriction digests were performed with *EcoRI* from Roche according to the manufacturer's protocol. The ~9 kb long range PCR fragment spanning the ELTOR deletion breakpoint was gel purified and cloned into pGEM T-EASY (Promega) before sequencing.

#### SEQUENCING

Sequencing template was generated from nested PCR products using spin columns from Qiagen. All sequencing reactions were carried out using the BigDye terminator cycle sequencing kit from Applied Biosystems according to the manufacturer's protocol. Sequencing products were run on an ABI377 sequencer (PE Biosystems).

#### Results

We previously localised the breakpoints of the *AZFa* microdeletion in the patient ELTOR to two tightly defined STS intervals within Yq, on the proximal side between sY83 and 83D22T7 and on the distal side between 494-130k and 494-146k (fig 1). Sequence analysis of these two intervals shows the presence of homologous copies of a human endogenous retroviral sequence (HERV15). The STS sY83 lies towards the end of the proximal copy. The average degree of homology is 94%, but varies markedly along an alignment of the two HERVs, as shown by the dot plots in fig 2. The

distal HERV is interrupted by an insertion of L1 material consisting of two smaller L1 fragments lying back to back (fig 3). The presence of 14 bp direct repeats flanking the entire insertion suggests that these fragments have integrated as a unit. The region of the two HERV copies distal to sY83 is particularly highly homologous and we hypothesised that the breakpoint may lie within this region. To test this hypothesis, long range PCR was used to span the breakpoint using primers from the retained flanking STS on either side of the deletion (sY83 and 494-130k). A 9 kb amplicon was obtained. This amplicon was digested with *EcoRI* to localise the breakpoint further. Five of the six fragments were of the size expected from an in silico digest of the region 3' to the distal HERV. The remaining 3.1 kb fragment represents the proximal end of this 9 kb amplicon and must contain the breakpoint. This fragment was cloned and fully sequenced. As hypothesised, the breakpoint is perfectly in register within an interval of complete identity between the HERVs of 1242 bp. This interval includes the 3' portion of the HERV and the 5' portion of the 3' LTR (fig 3). As well as removing the genes *DFFRY* and *DBY*, the resulting ~700 kb deletion also removes a single HERV15 copy, a composite of the two original HERVs.

The 12f2 polymorphism was typed using a newly developed PCR assay, which was shown

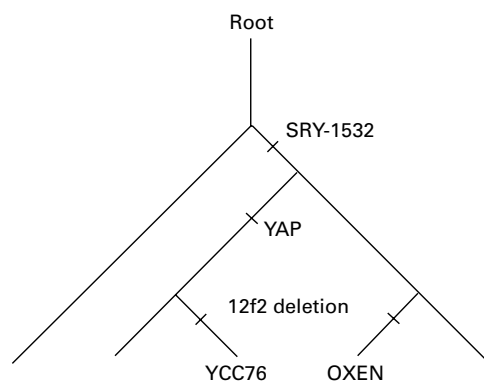


Figure 4 A rooted Y chromosomal phylogeny indicating the recurrent origins of the 12f2 deletion and the lineage affiliations of the two 12f2 deleted subjects sequenced.

to be in complete concordance with previous hybridisation analysis.<sup>22</sup> A screen of diverse Y chromosomes for the 12f2 deletion using the assay above showed that the 12f2 amplicon was absent from two sets of Y chromosomes with different haplotype backgrounds and geographical locations (data not shown). One of these haplotype backgrounds is defined primarily by the presence of the YAP insertion<sup>23</sup> and the derived form of the SRY-1532 polymorphism (also known as SRY10381 of Whitfield *et al*<sup>24</sup>), the other defined solely by the latter (fig 4). The YAP insertion has been shown to have occurred once in human evolution<sup>23</sup> with the ancestral state being the lack of an insertion. The ancestral state of the 12f2 deletion can be inferred to be the undeleted form from the published unambiguous rooting of the Y chromosomal phylogeny.<sup>16, 25</sup> Therefore, the 12f2 deletion must have occurred at least twice during human evolution. This information is summarised in the phylogeny in fig 4. This observation of recurrent deletion suggested that an event more complex than a simple deletion underlies the polymorphism(s).

The primers used in the 12f2 assay were used in a BLAST search of Y chromosomal contigs that identified the same distal HERV15 sequence identified above. Restriction analysis (*in silico*) of this sequence showed that the perfect excision of the L1 material described above could generate the polymorphic banding patterns observed in Southern hybridisations of genomic DNA digested with three different enzymes<sup>20</sup> (C Tyler-Smith, personal communication). However, no mechanism has been described by which perfect L1 excision can occur recurrently.

Consequently, two overlapping amplicons covering roughly 9 kb of the distal HERV were generated by long range PCR in two subjects, one from each of the two 12f2 deleted Y chromosomal lineages (YAP+, YCC76; YAP-, OXEN). A primer flanking the L1 insertion was used to generate sequence from the more distal of these nested PCR products in each subject. It was found that the L1 material was precisely excised together with one of the flanking 14 bp direct repeats in both 12f2 deleted lineages. In addition, however, sequence 5' to the site of excision was converted to be identical to the proximal copy of the

HERV. There was no conversion of sequence 3' to the site of L1 excision, thus defining the distal end of the converted sequence in both lineages. Further sequencing was undertaken to isolate the proximal end of the converted sequence. Alignment of the two HERV sequences showed hundreds of base substitutions and indels that differentiate between the two HERV sequences. A total of 238 of these differentiating sites were assayed by sequencing: 112 were found to be proximal to, four distal to, and 122 within the converted sequence. In addition, 3' to the distal four unconverted differentiating sites lies 110 bp of sequence flanking the HERV that is specific to the distal BAC. The proximal end of the converted sequence was localised to a 1285 bp interval of identity between the two HERV copies in both of the 12f2 deleted lineages, thus defining a minimum intervening interval of 4.6 kb of converted sequence. There was a single difference between the sequences generated from the two 12f2 deleted lineages, a deletion of 13 bp within the converted region of YCC76 (position 144024-144036 within clone 494g17). This deletion could either have occurred after the original conversion event in the recipient HERV or before the conversion event in the donor HERV.

## Discussion

### AN IMPORTANT SUBCLASS OF *AZF*<sub>a</sub> MICRODELETION

Here we define an *AZF*<sub>a</sub> microdeletion breakpoint at the sequence level for the first time and add Y chromosomally based infertility to a growing list of diseases resulting from illegitimate recombination between flanking repeats.<sup>3</sup> Furthermore we show that interchromosomal recombination need not be invoked to explain these events. We note that two recent studies defining *AZF*<sub>a</sub> breakpoints have included samples with remarkably similar breakpoints to those defined here.<sup>14, 26-28</sup> Most strikingly, the proximal breakpoint of the patient WHT2996<sup>14</sup> lies between the markers sY746 and sY740 (figs 1 and 3). Thus, we conclude that recombination between these HERV copies represents an important subclass of *AZF*<sub>a</sub> deletions.

### DOUBLE CROSSOVERS BETWEEN HERV SEQUENCES CAUSE THE 12F2 DELETION

The region of sequence conversion present in the distal HERV copy from subjects from the 12f2 deleted lineages is far longer than gene conversion events observed in any organism, to the best of the authors' knowledge. Thus, we conclude that this sequence conversion results from double crossover events. This mechanism is supported by the positioning of the conversion endpoints in intervals of identity between the HERV sequences longer than 400 bp. In common with other studies,<sup>4</sup> the length of identity rather than degree of homology is the more important determinant of homologous recombination. At first sight, the precise removal of a L1 insertion recurrently during human evolution seems highly unlikely. However, the proposed double crossover provides a plausible mechanism by which this can occur.

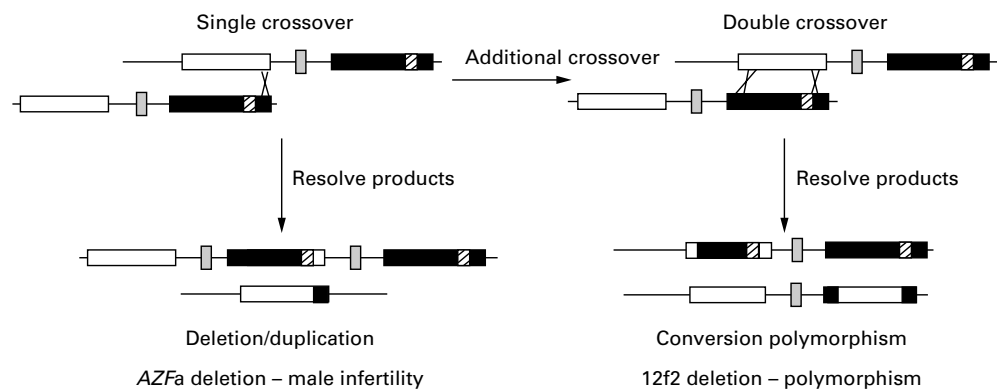


Figure 5 A mechanistic model for the divergent outcomes of single and double crossovers between misaligned sister chromatids (USCE). Proximal HERV is shown in white, the distal HERV in black, the L1 inserted material is shown hatched, and the intervening genes are represented in grey.

#### A MECHANISM OF INTRACHROMOSOMAL RECOMBINATION WITH DIVERGENT OUTCOMES

Exchange of material between the two copies is non-reciprocal as evidenced by the absence of a size change for the fragment from the proximal HERV in Southern hybridisations of 12f2 deleted subjects with the 12f2 probe. Thus, recombination would seem to occur because of misalignment between sister chromatids (USCE) rather than the alignment of repeats *in cis*. In light of this finding, we propose the mechanistic model for recombination between the two HERV shown in fig 5. Crosstalk between these HERV sequences can result in two outcomes, depending on the number of crossovers. Only a single product of each recombinant event shown in fig 5 is observed here. Duplications are difficult to detect using STS analysis alone, but we know that they exist on Yq.<sup>17</sup> As yet, the pathogenic effect of duplicating *AZF*a genes is unknown, though in the light of recent findings with autosomal syndromes<sup>2</sup> worth investigating.

Counting the number of bands in high stringency filter hybridisations using the probe 12f2 in both *TaqI* and *EcoRI* digests (data not shown) suggests that there are between 10 and 30 copies of HERV15 in the genome. Whether there is crosstalk between other copies of this family of dispersed repeats remains to be investigated.

#### PREMUTATIONS TO MICRODELETIONS?

We note that the two Y chromosomal lineages defined by the double crossover events reported here contain HERV sequences between which there is a substantially longer length of identity (6 kb), which could act as a better substrate for homologous recombination and as such may predispose to *AZF*a microdeletions. If true, 12f2 deletions could be considered as premutations, analogous to those in triplet repeat disorders.<sup>29</sup> Although further work is required to confirm this hypothesis this raises a number of important considerations.

(1) Given the fact that Y chromosomal lineages exhibit greater geographical differentiation than any other locus,<sup>30</sup> rates of *AZF*a deletion may vary substantially between different ethnic groups. Consequently, populations in which the two 12f2 deleted lineages occur at highest

frequency should be investigated. Neither of the two 12f2 deleted lineages is found at reasonable frequencies (less than 5%) within the north western European populations on which infertility research has focused.<sup>31</sup>

(2) If double recombinants can recurrently homogenise sequence between two ~10 kb repeats on the Y chromosome, the likelihood is that many of the pathogenic microdeletion causing repeats on the autosomes are also homogenised by double recombinant events. The likelihood of these homogenisations being polymorphic on autosomes is higher than for the Y chromosome given that the time since a most recent common ancestor of an autosome is on average four times that of the Y chromosome.<sup>15</sup> Thus, the homogenised repeat premutation (HRP) concept can readily be extended to many autosomal and X chromosomal microdeletions.

(3) If single crossovers occur within regions previously homogenised by undetected double crossovers, all breakpoints will appear to have occurred at one end of the homogenised tract. This would produce a false clustering of deletion breakpoints that may be misinterpreted. We recommend that parental repeats of deleted subjects be sequenced to exclude such possibilities when making assertions on breakpoint clustering. Clearly in the case presented here the breakpoint of the *AZF*a microdeletion lies outside the converted region, and so this is not an issue.

(4) The assumption that markers such as the 12f2 deletion are neutral is required for their use in reconstructing human prehistory. This assumption may be violated if these markers represent premutations, though only if negative selection outweighs drift.<sup>32 33</sup>

One of the hallmarks of the Y chromosome is the high frequency of amplified repeat sequences distributed throughout the euchromatic and heterochromatic regions. The findings reported in this paper illustrate the importance of flanking repeat sequences that may be disposed to intrachromosomal recombination in creating deletions or insertions. It may be that the majority of deletions of the Y chromosome can be explained by a similar model but involving different repeat sequences. It would be profitable to examine the location

of the common breakpoints in *AZFc* and *AZFc* patients to assess the involvement of flanking repeats in these deletions.

Sequences discussed in the text have been submitted to GenBank with the accession numbers AJ278654, AJ278655, and AJ278656. The authors would like to thank Chris Tyler-Smith for unpublished information and DNA samples. We are grateful to the Y Chromosome Consortium for the supply of genomic DNA and Zoe Rosser for unpublished data. We thank Pieter de Jong for the cosmid library LL0YNC03, constructed at the HGC, LLNL, Livermore, CA 94550, USA under the auspices of the National Laboratory Gene Library Project sponsored by the United States DOE. We would like to thank Professor Tim Hargreaves and Professor Howard Cooke for genomic DNA from patient ELTOR. Maria Shlumukova was funded by a Nuffield Foundation Undergraduate Research Bursary. MAJ is a Wellcome Trust Senior Fellow in Basic Biomedical Science (grant No 057559). This work was supported by the McDonald Institute for Archaeological Research and grants from the Wellcome Trust and the BBSRC. Patricia Blanco was supported by the British Council.

- 1 Lopez Correa C, Brems H, Lazaro C, Marynen P, Legius E. Unequal meiotic crossover: a frequent cause of NF1 microdeletions. *Am J Hum Genet* 2000;66:1969-74.
- 2 Potocki L, Chen KS, Park SS, Osterholm DE, Withers MA, Kimonis V, Summers AM, Meschino WS, Anyane-Yebo K, Kashork CD, Shaffer LS, Lupski JR. Molecular mechanism for duplication 17p11.2 - the homologous recombination reciprocal of the Smith-Magenis microdeletion. *Nat Genet* 2000;24:84-7.
- 3 Lupski J. Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet* 1998;14:417-22.
- 4 Lukacovich T, Waldman AS. Suppression of intrachromosomal gene conversion in mammalian cells by small degrees of sequence divergence. *Genetics* 1999;151:1559-68.
- 5 Lower R, Lower J, Kurth R. The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences. *Proc Natl Acad Sci USA* 1996;93:5177-84.
- 6 Patience C, Wilkinson DA, Weiss RA. Our retroviral heritage. *Trends Genet* 1997;13:116-20.
- 7 Sverdlov ED. Retroviruses and primate evolution. *BioEssays* 2000;22:161-71.
- 8 Erlandsson R, Wilson JF, Paabo S. Sex chromosomal transposable element accumulation and male-driven evolution. *Mol Biol Evol* 2000;17:804-12.
- 9 McElreavey K, Krausz C. Male infertility and the Y chromosome. *Am J Hum Genet* 1999;64:928-33.
- 10 Schwartz A, Chan DC, Brown LG, Alagappan R, Pettay D, Disteche C, McGillivray B, de la Chapelle A, Page DC. Reconstructing hominid Y evolution: X-homologous block, created by X-Y transposition, was disrupted by Yp inversion through LINE-LINE recombination. *Hum Mol Genet* 1998;7:1-11.
- 11 Lahn BT, Page DC. Four evolutionary strata on the human X chromosome. *Science* 1999;286:964-7.
- 12 Vogt PH, Edelmann A, Kirsch S, Henegariu O, Hirschmann P, Kiesewetter F, Köhn FM, Schill WB, Farah S, Ramos C, Hartmann M, Hartschuh W, Meschede D, Behre HM, Castel A, Nieschlag E, Weidner W, Gröne HJ, Jung A, Engel W, Haidl G. Human Y chromosome azoospermia factors (AZF) mapped to different subregions in Yq11. *Hum Mol Genet* 1996;5:933-43.
- 13 Sargent CA, Boucher CA, Kirsch S, Brown G, Weiss B, Trundle A, Burgoyne P, Saut N, Durand C, Levy N, Terriou P, Hargreave T, Cooke H, Mitchell M, Rappold GA, Affara NA. The critical region of overlap defining the AZFa male infertility interval of proximal Yq contains three transcribed sequences. *J Med Genet* 1999;36:670-7.
- 14 Sun C, Skaletsky H, Birren B, Devon K, Tang Z, Silber S, Oates R, Page DC. An azoospermic man with a de novo point mutation in the Y-chromosomal gene USP9Y. *Nat Genet* 1999;23:429-32.
- 15 Jobling MA, Tyler-Smith C. Fathers and sons: the Y chromosome and human evolution. *Trends Genet* 1995;11:449-56.
- 16 Hammer MF, Karafet T, Rasanayagam A, Wood ET, Altheide TK, Jenkins T, Griffiths RC, Templeton AR, Zegura SL. Out of Africa and back again: nested clastic analysis of human Y chromosome variation. *Mol Biol Evol* 1998;15:427-41.
- 17 Jobling MA, Samara V, Pandya A, Fretwell N, Bernasconi B, Mitchell RJ, Gerelsaikhan T, Dashnyam B, Sajantila A, Salo PJ, Nakahori Y, Disteche CM, Thangaraj K, Singh L, Crawford MH, Tyler-Smith C. Recurrent duplication and deletion polymorphisms on the long arm of the Y chromosome in normal males. *Hum Mol Genet* 1996;5:1767-75.
- 18 Hurler ME, Irven C, Nicholson J, Taylor PG, Santos FR, Loughlin J, Jobling MA, Sykes BC. European Y-chromosomal lineages in Polynesians: a contrast to the population structure revealed by mtDNA. *Am J Hum Genet* 1998;63:1793-806.
- 19 Bouzekri N, Taylor PG, Hammer MF, Jobling MA. Novel mutation processes in the evolution of a haploid minisatellite, MSY1: array homogenization without homogenization. *Hum Mol Genet* 1998;7:655-9.
- 20 Casanova M, Leroy P, Boucekine C, Weissenbach J, Bishop C, Fellous M, Purrello M, Fiori G, Siniscalco M. A human Y-linked DNA polymorphism and its potential for estimating genetic and evolutionary distance. *Science* 1985;230:1403-6.
- 21 Passarino G, Semino O, Quintana-Murci L, Excoffier L, Hammer M, Santachiara Benerecetti AS. Different genetic components in the Ethiopian population, identified by mtDNA and Y-chromosome polymorphisms. *Am J Hum Genet* 1998;62:420-34.
- 22 Pandya A. *Human Y-chromosomal DNA variation*. DPhil thesis, University of Oxford, 1998.
- 23 Hammer MF. A recent insertion of an Alu element on the Y chromosome is a useful marker for human population studies. *Mol Biol Evol* 1994;11:749-61.
- 24 Whitfield LS, Sulston JE, Goodfellow PN. Sequence variation of the human Y chromosome. *Nature* 1995;378:379-80.
- 25 Vogt PH, Affara N, Davey P, Hammer M, Jobling MA, Lau YFC, Mitchell M, Schempp W, Tyler-Smith C, Williams G, Yen P, Rappold GA. Report of the third international workshop on Y chromosome mapping 1997. *Cytogenet Cell Genet* 1997;79:2-16.
- 26 Foresta C, Ferlin A, Moro E. Deletion and expression analysis of AZFa genes on the human Y chromosome revealed a major role for DBY in male infertility. *Hum Mol Genet* 2000;9:1161-9.
- 27 Qureshi SJ, Ross AR, Ma K, Cooke HJ, McIntyre MA, Chandley AC, Hargreave TB. Polymerase chain reaction screening for Y chromosome microdeletions: the first step towards the diagnosis of genetically-determined spermatogenic failure in men. *Mol Hum Reprod* 1996;2:775-9.
- 28 Pryor JL, Kent-First M, Mualen A, Van Bergen AH, Noltzen WE, Meisner L, Roberts KP. Microdeletions in the Y chromosome of infertile men. *N Engl J Med* 1997;336:534-9.
- 29 Montermini L, Andermann E, Labuda M, Richter A, Pandolfo M, Cavalcanti F, Pianese L, Iodice L, Farina G, Monticelli A, Turano M, Filla A, de Michele G, Coccozza S. The Friedreich ataxia GAA triplet repeat: pre-mutation and normal alleles. *Hum Mol Genet* 1997;6:1261-6.
- 30 Seielstad MT, Minch E, Cavalli-Sforza LL. Genetic evidence for a higher female migration rate in humans. *Nat Genet* 1998;20:278-80.
- 31 Semino O, Passarino G, Brega A, Fellous M, Santachiara-Benerecetti AS. A view of the Neolithic demic diffusion in Europe through two Y chromosome-specific markers. *Am J Hum Genet* 1996;59:964-8.
- 32 Jobling MA, Williams G, Schiebel K, Pandya A, McElreavey K, Salas L, Rappold GA, Affara NA, Tyler-Smith C. A selective difference between human Y-chromosomal DNA haplotypes. *Curr Biol* 1998;8:1391-4.
- 33 Jobling MA, Tyler-Smith C. New uses for new haplotypes: the human Y chromosome, disease and selection. *Trends Genet* 2000;16:356-62.