Human Molecular Genetics, 2000, Vol. 9, No. 15 2291–2296

Deletion of *azoospermia factor a* (*AZFa*) region of human Y chromosome caused by recombination between HERV15 proviruses

Chao Sun+, Helen Skaletsky, Steve Rozen, Jörg Gromoll[1], Eberhard Nieschlag[1], Robert Oates[2] and David C. Page§

Howard Hughes Medical Institute, Whitehead Institute, and Department of Biology, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, MA 02142, USA, [1]Institute of Reproductive Medicine of the University, Domagkstrasse 11, D-48129 Münster, Germany and [2]Department of Urology, Boston University School of Medicine, Boston, MA 02118, USA

Received 31 May 2000; Revised and Accepted 2 August 2000 DDBJ/EMBL/GenBank accession nos: G49213, G64723–G64733, G64861 and G64905

Deletion of any of three regions of the human Y chromosome results in spermatogenic failure and infertility. We previously sequenced one of these regions, *azoospermia factor a* (*AZFa*) and found that it spanned ~800 kb. By sequence-tagged site (STS) content mapping, we roughly defined deletion breakpoints in two unrelated, azoospermic men with *AZFa* deletions. The positions of proximal and distal breakpoints were similar in the two men. Hypothesizing that the deletions might have been generated by homologous recombination, we searched electronically for DNA sequence similarities between the proximal and distal breakpoint regions. These comparisons revealed the most striking sequence similarities anywhere along or near the *AZFa* region. In the proximal breakpoint region, we identified a 10 kb provirus of the recently defined HERV15 class of endogenous retroviruses. In the distal breakpoint region, we found a second HERV15 provirus, 94% identical in DNA sequence to the first and in the same orientation. (A partial LINE insertion in this distal provirus proved to be the basis of the previously described *DYS11*/p12f polymorphism.) The *AZFa* deletions in the two men differed slightly, but all breakpoints fell within the HERV15 proviruses. Indeed, sequencing of deletion junctions from the two men revealed that homologous recombination had occurred within large domains of absolute sequence identity between the proximal and distal proviruses. When combined with published STS mapping data for other *AZFa*-deleted men, our findings suggest that recombination between these two HERV15 proviruses could account for most *AZFa* deletions.

INTRODUCTION

Most Y-chromosomal mutations of medical importance appear to involve deletions or other rearrangements. The known exceptions are *SRY* point mutations that cause XY sex reversal (1,2) and one case of a *USP9Y* point mutation that caused spermatogenic failure (3). Perhaps the most frequent Y chromosome defects are submicroscopic, interstitial Yq deletions that disrupt spermatogenesis (4–8). Indeed, interstitial Yq deletions are the most common molecularly defined cause of spermatogenic failure in human populations. Deletion of any of three *azoospermia factor* regions on Yq (*AZFa*, *AZFb* or *AZFc*) dramatically impairs spermatogenesis (4–8). Laboratories throughout the world have identified many such interstitial Yq deletions, but little is known about the molecular mechanisms by which the deletions arise. Prior to the present study, no breakpoints of interstitial Yq deletions had been analyzed at the level of nucleotide sequence.

We previously sequenced the *AZFa* region and found that it spanned ~800 kb (3). The *AZFa* region appears to contain only two genes, *USP9Y* and *DBY* (9,10). *USP9Y* is required for male germ cell development and indirect evidence suggests that *DBY* may also play a critical role in this process (3). At least 10 unrelated men with deletions of the *AZFa* region have been reported in the literature (3,6,7,10–12). In all 10 cases, STS mapping results providing rough localization of proximal and distal breakpoints have been published. In 8 of the 10 cases, the size and position of the *AZFa* deletion appear to be similar. Within the broad limits of resolution of the published sequence-tagged site (STS) mapping data, the deletions in these eight unrelated men have indistinguishable proximal breakpoints; they also have indistinguishable distal breakpoints. These comparative observations led us to entertain the hypothesis that some particular molecular mechanism, operating on defined hot spots in this region of Yq, might be responsible for the recurrence of similar *AZFa* deletions. We set out to explore this hypothesis by precisely defining and sequencing the proximal and distal breakpoints in two unrelated, azoospermic men with typical *AZFa* deletions.

+Present address: Department of Molecular Technologies, Biogen Inc., 14 Cambridge Center, Cambridge, MA 02142, USA
§To whom correspondence should be addressed. Tel: +1 617 258 5203; Fax: +1 617 258 5578; Email: dcpage@wi.mit.edu

RESULTS

Our laboratories identified two individuals with *AZFa* deletions by testing a total of 1307 men with spermatogenic failure (azoospermia or severe oligospermia) for the presence or absence of Yq STSs (13). Of these 1307 men, 860 were screened in Cambridge, MA, and azoospermic man WHT2996 was shown to be deleted for *AZFa*, as previously reported (3). An additional 447 men were screened in Münster, Germany, where azoospermic man WHT3667 was found to be deleted for *AZFa*. In each of these two men, we mapped the proximal breakpoint to the 281 kb region between STSs sY82 and sY86 and the distal breakpoint to the 595 kb region between sY87 and sY88 (Fig. 1). These STS mapping results were similar to those reported by other laboratories in unrelated cases of *AZFa* deletion (6,7,10–12). As reported previously (3), the *AZFa* deletion in WHT2996 had arisen *de novo*; all STSs tested were present in his father (Fig. 1). No paternal sample was available in the case of WHT3667.

Electronic search of *AZFa* sequence suggests candidate sites for homologous recombination

With the complete nucleotide sequence of the *AZFa* region (3) in hand, we set out to define more precisely the proximal and distal breakpoints in WHT2996 and WHT3667. Throughout the human genome, many gene deletions appear to result from homologous recombination between direct repeat sequences (14). With this generalization in mind, we searched electronically for DNA sequences that were present in the same orientation in both the proximal and distal *AZFa* breakpoint regions. Specifically, we searched for sequence identities of at least 100 bp. [In mammalian cells, efficient DNA recombination between two sequences appears to require at least 200 bp of nucleotide identity; identities of <95 bp appear to be poor substrates for recombination (15,16).] In this electronic analysis, we compared all 100 bp segments of the *AZFa* region with all other 100 bp segments, using the BLASTN search tool (17). The arched lines in Figure 1 depict all 100 bp, same-orientation identities within the *AZFa* region. As is evident in Figure 1, most such identical segments are clustered within two blocks located ~800 kb apart, one within the proximal breakpoint region and the other within the distal breakpoint region.

By searching a sequence database of repetitive elements using RepeatMasker (http://ftp.genome.washington.edu/RM/RepeatMasker.html), we discovered that the two blocks of sequence identity are two different members of a recently defined class of human endogenous retrovirus, HERV15 (18). The proximal block corresponds to a single HERV15 provirus and the distal block corresponds to a second HERV15 provirus. The two proviruses are in the same orientation on the Y chromosome. The proximal provirus (provirus A) is 10 kb in length. The distal copy (provirus B) is 1.5 kb longer because of a partial LINE insertion, which divides provirus B into a 5′ segment of 8.2 kb and a 3′ segment of 1.8 kb (Fig. 2). The two proviruses exhibit >96% identity to the HERV15 consensus sequence (18). Apart from the LINE insertion in provirus B, proviruses A and B are 94% identical in DNA sequence. Indeed, the sequences of proviruses A and B are identical in a 1.3 kb region 5′ of their centers and they differ at only one nucleotide in a 1.7 kb region near their 3′ ends (Fig. 2).

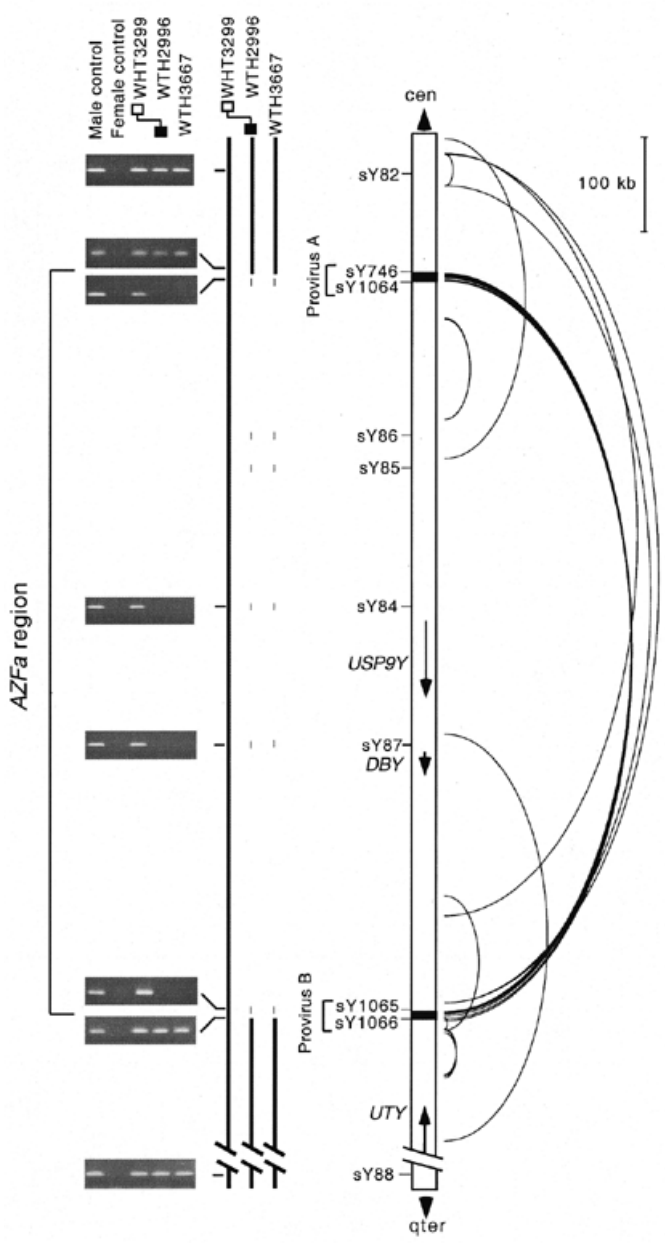

Figure 1. *AZFa* region of the human Y chromosome. The vertical bar at the center of the figure represents this portion of the chromosome, oriented with respect to the centromere (above) and the long-arm telomere (below) (based on Fig. 1 in ref. 3). Within the vertical open bar, three arrows depict the direction of the transcription and the positions of the *USP9Y*, *DBY* and *UTY* genes (only the 3′ portion of *UTY* is shown). Solid black bars represent HERV15 proviruses A and B. Immediately to the left of the chromosome are listed 10 STSs used to characterize deletions in azoospermic individuals WHT2996 and WHT3667. Results of testing the genomic DNA from WHT2996, his father (WHT3299) and WHT3667 for the presence or absence of these STSs are summarized further to the left. Solid black bars encompass STSs found to be present, and minus signs indicate absence of those STSs. Gel images of eight of these tests are shown at the extreme left. Some data for WHT2996 have been reported previously (3). To the right of the chromosome are depicted all same-orientation, 100 bp nucleotide identities within the 1 Mb region shown. Each arc connects two 100 bp segments with identical nucleotide sequences. Note the concentration

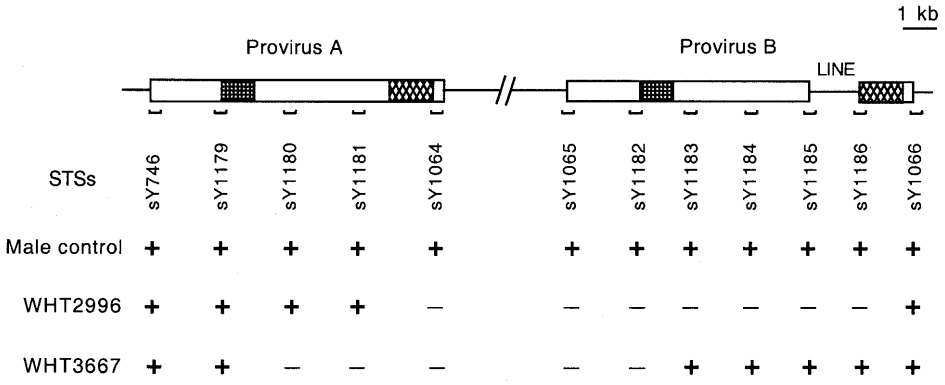

Figure 2. STS mapping of deletion breakpoints within proviruses A and B. At the top is a schematic diagram of proviruses A and B. Long regions of the virtual sequence identity between proviruses A and B are shaded. Shown immediately below are the positions of 12 STSs, specific to either provirus A or provirus B. Results of testing *AZFa*-deleted men WHT2996 and WHT3667 for presence (+) or absence (–) of these 12 STSs are shown at the bottom.

These electronic analyses suggested that the HERV15 proviruses might be favorable substrates for homologous recombination giving rise to *AZFa* deletions. To test this prediction, we designed STS markers (sY746 and sY1064) that immediately flank the proximal provirus and other STS markers (sY1065 and sY1066) that immediately flank the distal provirus (Fig. 1). We then tested the *AZFa*-deleted men, WHT2996 and WHT3667, for these four markers. The results indicate that in each of these men, the proximal breakpoint falls within provirus A and the distal breakpoint falls within provirus B (Fig. 1).

Recombination between proximal and distal proviruses occurred within long sequence identities

To localize breakpoints more precisely within the proviruses, we designed three STS markers (sY1179–sY1181) specific to and distributed along the length of provirus A and five STS markers (sY1182–sY1186) specific to and distributed along the length of provirus B (Fig. 2). Specificity of these PCR assays for their respective proviruses was confirmed on bacterial artificial chromosome (BAC) clones containing either provirus A or B (data not shown). We tested both *AZFa*-deleted men for these eight STSs. These tests revealed that in WHT2996, the proximal breakpoint is located near the 3′ end of provirus A; the distal breakpoint is positioned similarly in provirus B (Fig. 2). In WHT3667, the proximal breakpoint is located much more 5′ in provirus A; the distal breakpoint is positioned similarly in the 5′ portion of provirus B (Fig. 2). These results demonstrated that the breakpoints differed by several kb in the two men. In both cases, however, the breakpoints fell within or near regions of extensive sequence identity between proviruses A and B (Fig. 2).

Finally, in WHT2996 and WHT3667, we PCR amplified and sequenced the provirus A/provirus B junction fragments created by *AZFa* deletion. Because the deletion breakpoints differed in WHT2996 and WHT3667 (Fig. 2), PCR analysis of the junction fragments was customized to each case. Of the two patients, WHT2996 had the more 3′ breakpoints (Fig. 2). Here we employed a PCR primer pair that in normal males amplified a 658 bp segment of provirus A. The forward primer was specific to provirus A and was located 5′ of the proximal breakpoint of WHT2996. The reverse primer was common to proviruses A and B and was located 3′ of the distal breakpoint of WHT2996. If the *AZFa* deletion in WHT2996 had been created by homologous recombination, then we would also expect a PCR product of 658 bp in WHT2996. This was observed (data not shown). Sequencing of the PCR product from WHT2996 revealed that homologous recombination between proviruses A and B had occurred within a 447 bp region of absolute identity between the two proviruses (Fig. 3A).

The second patient, WHT3667, had the more 5′ breakpoints (Fig. 2). Here we employed a forward PCR primer that was specific to provirus A and was located 5′ of the proximal breakpoint of WHT3667. The reverse primer was specific to provirus B and was located 3′ of the distal breakpoint of WHT3667. In normal males, this primer pair yielded no product, as expected given the 800 kb separating proviruses A and B when the *AZFa* region is intact. In WHT3667, however, we expected and observed a PCR product of 1685 bp. Sequencing of the PCR product from WHT3667 demonstrated that homologous recombination between proviruses A and B had occurred within a 1285 bp region of absolute identity between the two proviruses (Fig. 3B).

Polymorphism for LINE insertion in provirus B

While studying the HERV15 proviruses that bound the *AZFa* region, we observed that the LINE sequence in provirus B was absent in some men tested as controls. We developed positive PCR assays for both the presence and absence of the LINE sequence (see Materials and Methods).

Four types of evidence suggested that the presence or absence of this LINE element corresponds to a previously reported polymorphism at *DYS11*—a polymorphism detected by hybridizing plasmid p12f to Southern blots of human genomic DNAs (19). First, plasmid p12f was the source of STS sY83 (13) (GenBank accession no. G64734), which is 88% identical to sequence located only 200 bp from the LINE insertion in provirus B. (sY83 and p12f derive from the homologous region of provirus A.) Second, the reported sizes of the polymorphic *Taq*I and *Eco*RI fragments detected by p12f (19) agreed reasonably well with those of the polymorphic

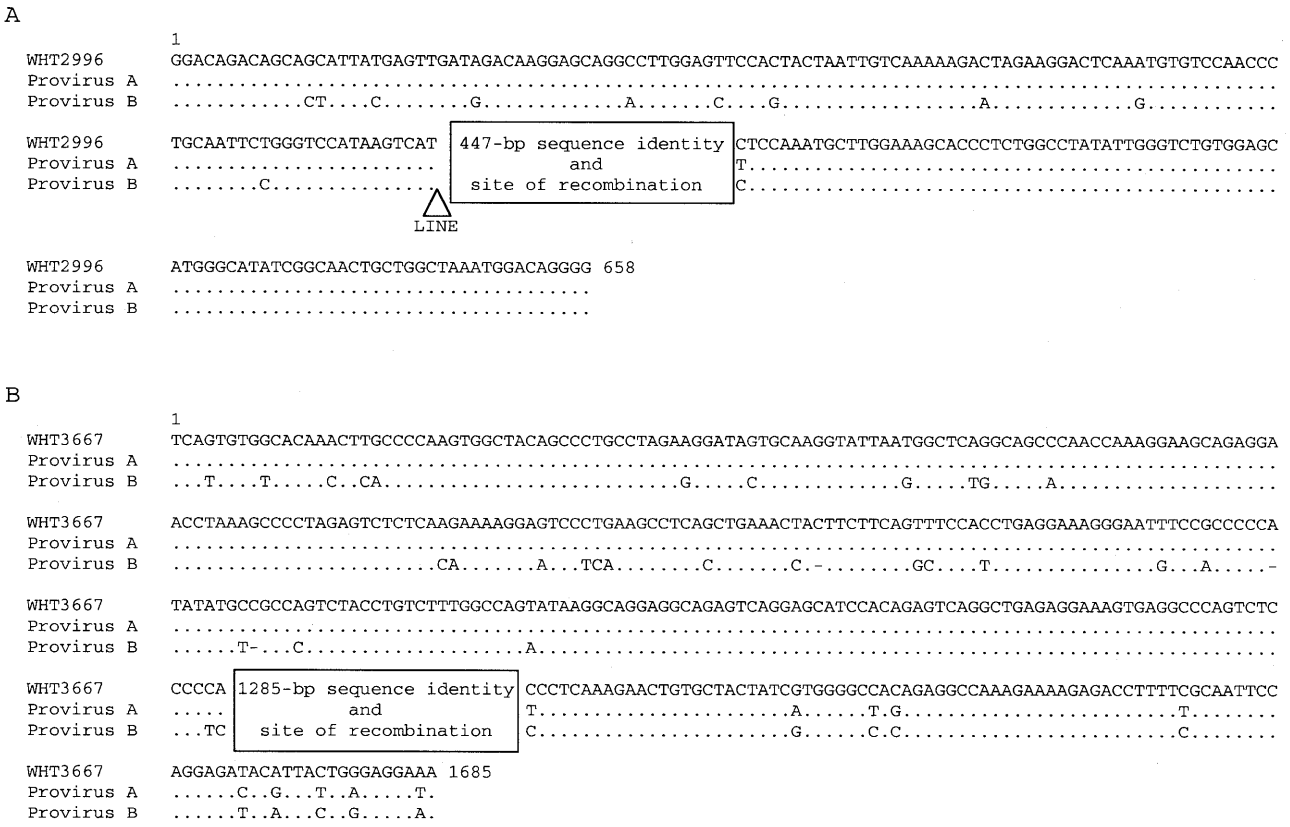

Figure 3. Nucleotide sequence analysis of provirus A/provirus B junction fragments in *AZFa*-deleted men (A) WHT2996 and (B) WHT3667. Dots indicate the identity to the sequence immediately above. (**A**) Sequence of 658 bp PCR product from WHT2996 aligned with homologous portions of proviruses A and B. The box represents a 447 bp region of absolute identity among the three sequences. (**B**) Sequence of 1685 bp PCR product from WHT3667 aligned with homologous portions of proviruses A and B. The box represents a 1285 bp region of absolute identity among the three sequences.

LINE-containing fragments predicted by electronic analysis of bacterial artificial chromosomes CTC-494G17 and CTC-203M13 (GenBank accession nos. AC005820 and AC002992, respectively) (3). The electronically predicted *Eco*RI fragments are 4704 bp (LINE-positive provirus B), 3162 bp (LINE-negative provirus B) and 5101 bp (homologous region in provirus A). The electronically predicted *Taq*I fragments are 10 357 bp (LINE-positive provirus B), 8815 bp (LINE-negative provirus B) and 8754 bp (homologous region in provirus A). Third, LINE-negative men exhibited Y chromosome haplotypes that were consistent with identity between the LINE polymorphism and the previously reported *DYS11*/p12f polymorphism. Here we typed 28 LINE-negative men for the Y-linked polymorphisms M9 (20), YAP (21) and SRY10831 (22). All 28 LINE-negative men tested proved to have the following haplotype: M9-C, YAP-negative, SRY10831-G, which is consistent with the haplotype previously reported for men with the derived allele at *DYS11*/p12f (23). Fourth, by testing 20 unrelated men, we directly correlated the *Taq*I polymorphism detected by p12f with the presence or absence of the LINE sequence, as detected by PCR. Of the 20 men, 16 exhibited a 10 kb *Taq*I fragment, and all 16 tested positive for the LINE sequence. The remaining four men exhibited no 10 kb *Taq*I fragment, and all four tested negative for the LINE sequence.

Taken together, these findings demonstrate that the presence or absence of the partial LINE element in provirus B is identical to the previously reported restriction fragment length polymorphism at *DYS11*/p12f. When combined with previous studies of the p12f polymorphism (23), our results imply that the provirus B LINE element was present in the common ancestor of extant human Y chromosomes and that LINE-negative Y chromosomes descend from a single deletion event. This deletion may have been caused either by homologous recombination between 14 bp direct repeats that flank the partial LINE element, or by recombination (double crossing-over or gene conversion) between proviruses A and B. As described in Materials and Methods, we now report PCR assays for typing this valuable and widely studied polymorphism.

DISCUSSION

Several human diseases are known to result from gene deletions that are caused by homologous recombination between direct repeats flanking the critical gene(s) (14). Previously reported examples include X-linked ichthyosis (24) and numerous autosomal disorders, including Charcot–Marie–Tooth disease (14). Our findings suggest that spermatogenic failure due to deletion

of the Y chromosome *AZFa* region can be added to this growing list of disorders.

Our laboratories identified two azoospermic men, WHT2996 and WHT3667, with deletions of the *AZFa* region. Both deletions proved to be the result of homologous recombination between two HERV15 proviruses that on an intact Y chromosome are located 800 kb apart. The precise sites of homologous recombination differed in the two men, but in both cases, recombination occurred within lengthy (447 and 1285 bp) regions of identity between the proviruses. In seven of nine other published cases of *AZFa* deletion, STS mapping data (6,7,11,12,25) are consistent with provirus–provirus recombination as described here for WHT2996 and WHT3667. Two published cases (SAYER and AZ539) have smaller deletions within the *AZFa* region (10,12) and cannot be explained on this basis. Taken together, past and present findings suggest that recombination between HERV15 proviruses may account for most, but not all cases of *AZFa* deletion. The PCR assays described here should facilitate testing of this hypothesis by other investigators.

To our knowledge, this is the first report of provirus–provirus recombination causing deletion, duplication or disruption of a human gene. Given the ubiquity of retroviral proviruses in the human genome (26,27), we anticipate that additional examples will be found. An entirely speculative question at present is whether the biological characteristics of human endogenous retroviruses might predispose their proviruses to become sites of recombination. In the present cases, we do not know whether *AZFa* deletion via homologous recombination occurred in the paternal germ line (during mitotic or meiotic cell cycles) or in the early embryo.

As described in Results, electronic analysis of the nucleotide sequence of the *AZFa* region played a critical role in our efforts to efficiently define recombination sites in WHT2996 and WHT3667. Indeed, it was electronic analysis that initially identified the HERV15 proviruses as potential sites of homologous recombination. This electronic analysis was possible because the complete nucleotide sequence of the *AZFa* region had been determined (3). With knowledge of the DNA sequence of all human chromosomes increasing rapidly, electronic prediction of recombination sites could soon play an important role in analyzing any chromosomal rearrangement in which homologous recombination is suspected.

MATERIALS AND METHODS

STS analysis of genomic DNA

We tested for the presence or absence of Y-chromosomal STSs with PCR primers selected using Primer3 (28). Primer sequences and thermocycling conditions for all STSs studied have been deposited in GenBank, where accession numbers are as follows: sY746, G49213; sY1064, G64723; sY1065, G64724; sY1066, G64725; sY1179, G64726; sY1180, G64727; sY1181, G64728; sY1182, G64729; sY1183, G64730; sY1184, G64731; sY1185, G64732; sY1186, G64861; sY1187, G64733; sY1193, G64905.

The LINE polymorphism in provirus B (the *DYS11*/p12f polymorphism) was scored using two complementary PCR assays. STS sY1187 is present in LINE-positive men but absent in LINE-negative men. STS sY1193 is present in LINE-negative men but absent in LINE-positive men.

Sequence analysis of provirus A/provirus B junction fragments

The junction fragment in WHT2996 was PCR amplified using the primers 5′-GGACAGACAGCAGCATTATGAG-3′ and 5′-CCCCTGTCCATTTAGCCA-3′. The junction fragment in WHT3667 was amplified using the primers 5′-TCAGTGTGGC-ACAAACTTGC-3′ and 5′-TTTCCCCCAGTAATGTATC-TCC-3′. The resulting PCR products were sequenced using a ThermoSequenase II kit and an ABI 370 machine (Perkin Elmer, Foster City, CA).

ACKNOWLEDGEMENTS

We thank Laura Brown for maintaining patient archives, Raaji Alagappan, Loreall Pooler and Marta Velez-Stringer for preparing genomic DNA, Bruce Lahn for support and advice, and Debbie Berry, Richa Saxena and Tomoko Kawaguchi for comments on the manuscript. C.S. was the recipient of an NIH post-doctoral fellowship. This work was supported in part by the NIH.

REFERENCES

1. Berta, P., Hawkins, J.R., Sinclair, A.H., Taylor, A., Griffiths, B.L. and Goodfellow, P.N. (1990) Genetic evidence equating *SRY* and the testis-determining factor. *Nature*, **348**, 448–450.
2. Jäger, R.J., Anvret, M., Hall, K. and Scherer, G. (1990) A human XY female with a frame-shift mutation in the candidate testis-determining gene *SRY*. *Nature*, **348**, 452–454.
3. Sun, C., Skaletsky, H., Birren, B., Devon, K., Tang, Z., Silber, S., Oates, R. and Page, D.C. (1999) An azoospermic man with a *de novo* point mutation in the Y-chromosomal gene *USP9Y*. *Nature Genet.*, **23**, 429–432.
4. Ma, K., Sharkey, A., Kirsch, S., Vogt, P., Keil, R., Hargreave, T.B., McBeath, S. and Chandley, A.C. (1992) Towards the molecular localisation of the AZF locus: mapping of microdeletions in azoospermic men within 14 subintervals of interval 6 of the human Y chromosome. *Hum. Mol. Genet.*, **1**, 29–33.
5. Reijo, R., Lee, T.Y., Salo, P., Alagappan, R., Brown, L.G., Rosenberg, M., Rozen, S., Jaffe, T., Strauss, D., Hovatta, O. *et al.* (1995) Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene. *Nature Genet.*, **10**, 383–393.
6. Vogt, P.H., Edelmann, A., Kirsch, S., Henegariu, O., Hirschmann, P., Kiesewetter, F., Köhn, F.M., Schill, W.B., Farah, S., Ramos, C. *et al.* (1996) Human Y chromosome azoospermia factors (AZF) mapped to different subregions in Yq11. *Hum. Mol. Genet.*, **5**, 933–943.
7. Qureshi, S.J., Ross, A.R., Ma, K., Cooke, H.J., McIntyre, M.A., Chandley, A.C. and Hargreave, T.B. (1996) Polymerase chain reaction screening for Y chromosome microdeletions: a first step towards the diagnosis of genetically-determined spermatogenic failure in men. *Mol. Hum. Reprod.*, **2**, 775–779.
8. Simoni, M., Gromoll, J., Dworniczak, B., Rolf, C., Abshagen, K., Kamischke, A., Carani, C., Meschede, D., Behre, H.M., Horst, J. *et al.* (1997) Screening for deletions of the Y chromosome involving the DAZ (Deleted in AZoospermia) gene in azoospermia and severe oligozoospermia. *Fertil. Steril.*, **67**, 542–547.
9. Lahn, B.T. and Page, D.C. (1997) Functional coherence of the human Y chromosome. *Science*, **278**, 675–680.
10. Brown, G.M., Furlong, R.A., Sargent, C.A., Erickson, R.P., Longepied, G., Mitchell, M., Jones, M.H., Hargreave, T.B., Cooke, H.J. and Affara, N.A. (1998) Characterisation of the coding sequence and fine mapping of the human *DFFRY* gene and comparative expression analysis and mapping to the *Sxr^b* interval of the mouse Y chromosome of the *Dffry* gene. *Hum. Mol. Genet.*, **7**, 97–107.

11. Pryor, J.L., Kent-First, M., Muallem, A., Van Bergen, A.H., Nolten, W.E., Meisner, L. and Roberts, K.P. (1997) Microdeletions in the Y chromosome of infertile men. *N. Engl. J. Med.*, **336**, 534–539.
12. Sargent, C.A., Boucher, C.A., Kirsch, S., Brown, G., Weiss, B., Trundley, A., Burgoyne, P., Saut, N., Durand, C., Levy, N. *et al.* (1999) The critical region of overlap defining the *AZFa* male infertility interval of proximal Yq contains three transcribed sequences. *J. Med. Genet.*, **36**, 670–677.
13. Vollrath, D., Foote, S., Hilton, A., Brown, L.G., Beer-Romero, P., Bogan, J.S. and Page, D.C. (1992) The human Y chromosome: a 43-interval map based on naturally occurring deletions. *Science*, **258**, 52–59.
14. Lupski, J. (1998) Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet.*, **14**, 417–422.
15. Waldman, A. and Liskay, R. (1988) Dependence of intrachromosomal recombination in mammalian cells on uninterrupted homology. *Mol. Cell. Biol.*, **8**, 5350–5357.
16. Liskay, R.M., Letsou, A. and Stachelek, J.L. (1987) Homology requirement for efficient gene conversion between duplicated chromosomal sequences in mammalian cells. *Genetics*, **115**, 161–167.
17. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.
18. Kapitonov, V.V. and Jurka, J. (1999) Humrep.ref, Repbase Update (http://www.girinst.org).
19. Casanova, M., Leroy, P., Boucekkine, C., Weissenbach, J., Bishop, C., Fellous, M., Purrello, M., Fiori, G. and Siniscalco, M. (1985) A human Y-linked DNA polymorphism and its potential for estimating genetic and evolutionary distance. *Science*, **230**, 1403–1406.
20. Underhill, P.A., Jin, L., Lin, A.A., Mehdi, S.Q., Jenkins, T., Vollrath, D., Davis, R.W., Cavalli-Sforza, L.L. and Oefner, P.J. (1997) Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res.*, **7**, 996–1005.
21. Hammer, M.F. (1994) A recent insertion of an Alu element on the Y chromosome is a useful marker for human population studies. *Mol. Biol. Evol.*, **11**, 749–761.
22. Whitfield, L.S., Sulston, J.E. and Goodfellow, P.E. (1995) Sequence variation of the human Y chromosome. *Nature*, **378**, 379–380.
23. Jobling, M.A., Pandya, A. and Tyler-Smith, C. (1997) The Y chromosome in forensic analysis and paternity testing. *Int. J. Legal Med.*, **110**, 118–124.
24. Yen, P.H., Li, X.-M., Tsai, S.-P., Johnson, C., Mohandas, T. and Shapiro, L.J. (1990) Frequent deletions of the human X chromosome distal short arm result from recombination between low copy repetitive elements. *Cell*, **61**, 603–610.
25. Vogt, P., Chandley, A.C., Hargreave, T.B., Keil, R., Ma, K. and Sharkey, A. (1992) Microdeletions in interval 6 of the Y chromosome of males with idiopathic sterility point to disruption of AZF, a human spermatogenesis gene. *Hum. Genet.*, **89**, 491–496.
26. Löwer, R., Löwer, J. and Kurth, R. (1996) The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences. *Proc. Natl Acad. Sci. USA*, **93**, 5177–5184.
27. Patience, C., Wilkinson, D.A. and Weiss, R.A. (1997) Our retroviral heritage. *Trends Genet.*, **13**, 116–120.
28. Rozen, S. and Skaletsky, H. (1999) In Misener, S. and Krawetz, S.A. (eds), *Bioinformatics Methods and Protocols*. Humana Press, Totowa, NJ, pp. 365–386.