ORIGINAL ARTICLE

Sequence family variant loss from the *AZFc* interval of the human Y chromosome, but not gene copy loss, is strongly associated with male infertility

N Machev, N Saut, G Longepied, P Terriou, A Navarro, N Levy, M Guichaoua, C Metzler-Guillemain, P Collignon, A-M Frances, J Belougne, E Clemente, J Chiaroni, C Chevillard, C Durand, A Ducourneau, N Pech, K McElreavey, M-G Mattei, M J Mitchell

J Med Genet 2004;**41**:814–825. doi: 10.1136/jmg.2004.022111

See end of article for authors' affiliations

Correspondence to: M J Mitchell, Inserm U. 491, Faculté de médecine, 13385 Marseille, France; mitchell@medecine.univ-mrs.fr

Revised version received 14 June 2004
Accepted for publication 16 June 2004

Background: Complete deletion of the complete *AZFc* interval of the Y chromosome is the most common known genetic cause of human male infertility. Two partial *AZFc* deletions (gr/gr and b1/b3) that remove some copies of all *AZFc* genes have recently been identified in infertile and fertile populations, and an association study indicates that the resulting gene dose reduction represents a risk factor for spermatogenic failure.

Methods: To determine the incidence of various partial *AZFc* deletions and their effect on fertility, we combined quantitative and qualitative analyses of the *AZFc* interval at the *DAZ* and *CDY1* loci in 300 infertile men and 399 control men.

Results: We detected 34 partial *AZFc* deletions (32 gr/gr deletions), arising from at least 19 independent deletion events, and found gr/gr deletion in 6% of infertile and 3.5% of control men ($p>0.05$). Our data provide evidence for two large *AZFc* inversion polymorphisms, and for relative hot and cold spots of unequal crossing over within the blocks of homology that mediate gr/gr deletion. Using SFVs (sequence family variants), we discriminate *DAZ1/2*, *DAZ3/4*, *CDY1a* (proximal), and *CDY1b* (distal) and define four types of *DAZ-CDY1* gr/gr deletion.

Conclusions: The only deletion type to show an association with infertility was *DAZ3/4-CDY1a* ($p=0.042$), suggesting that most gr/gr deletions are neutral variants. We see a stronger association, however, between loss of the *CDY1a* SFV and infertility ($p=0.002$). Thus, loss of this SFV through deletion or gene conversion could be a major risk factor for male infertility.

In 1976, the cytogenetic analysis of six azoospermic individuals mapped the *AZF* (*azoospermia factor*) locus, which is necessary for normal spermatogenesis, to the distal long arm euchromatin of the human Y chromosome.[1] It was not until 1992 that molecular analysis confirmed these findings with the discovery of two interstitial deletions situated in this region in azoospermic men.[2] In 1996, a third, independent, deletion interval was identified and the three intervals were named *AZFa*, *AZFb*, and *AZFc*.[3] Subsequent studies have demonstrated that deletions affecting the *AZFc* interval are by far the most common, affecting approximately 1 in 10 men with idiopathic azoospermia.[4–6]

The first *AZFc* candidate gene to be isolated from the *AZFc* interval was *DAZ* (d*eleted in* a*zoospermia*).[5] *DAZ* originated on the Y chromosome as a transposition of a DNA segment containing the autosomal germ-cell specific gene *DAZL*, and comparative analyses in primates show that the *DAZ* transposition occurred ~35 million years ago in a common ancestor of old world monkeys and humans.[7] There is at least one other functional gene within the *AZFc* interval: *CDY1*, a retroposon that codes a protein with proven histone acetyltransferase activity.[8] Although found only in primates, *CDY1* is believed to have been retroposed directly into MSY (*m*ale *s*pecific region of the *Y* chromosome) from a transcript of the autosomal gene *CDYL* more than 150 million years ago, making it one of the oldest genes on the Y chromosome.[9] At present, *DAZ* and *CDY1* are the only transcription units in the *AZFc* interval for which there is any evidence that function has been selectively maintained on the Y chromosome.

In almost all cases, deletions of the entire *AZFc* interval are associated with infertility characterised by a drastic reduction in sperm count due to the loss of active spermatogenesis from most testicular tubules. Spermatogenesis is completed, but on a very small scale. Most *AZFc* deleted men are severely oligozoospermic or azoospermic, but reported sperm counts range from 0 to 7 million per millilitre. There are, nevertheless, rare *AZFc* deleted men who have conceived multiple children naturally.[10–12] All of the sons of these men are infertile. The few spermatozoa produced by most *AZFc* deleted men can be successfully used for in vitro fertilisation by ICSI (intracytoplasmic sperm injection). Although *AZFc* deleted children conceived by ICSI are generally healthy, it is important to identify the *AZFc* genes involved, as this will improve the diagnosis of genetic causes of male infertility, and, thus, allow an informed evaluation of the health risks faced by an ICSI conceived child.

Despite the high frequency of the *AZFc* microdeletion, the key genes have not yet been identified. The detection and characterisation of smaller deletions within the *AZFc* interval are required, in order to understand how the *AZFc* interval contributes to human fertility. Initial evidence that partial *AZFc* deletions are present in a small percentage of infertile men came from several groups focusing on either sequence family variant (SFV) or fluorescent in situ hybridisation (FISH) analysis of the *DAZ* gene.[13–15] Recently, using locus specific PCR assays and dual colour FISH, a common class of partial deletion was described and was named the gr/gr deletion (g1/g2,r1/r3,r2/r4).[16] The gr/gr deletion is predicted to

Abbreviations: CEPH, Centre d'étude du polymorphisme humain; FISH, fluorescent in situ hybridisation; ICSI, intracytoplasmic sperm injection; SFV, sequence family variant

result in the loss of some, but not all, copies of all *AZFc* genes.[17] In that study, gr/gr deletion was found in 3.8% of infertile men with a sperm count of $<5\times10^6$/ml, and 2.2% of fertile men. However, its absence from a group of 148 normospermic men led to the conclusion that the gr/gr deletion is associated with decreased sperm production resulting from the reduced dose of *AZFc* genes.[16] Here, we present our characterisation of the *AZFc* interval in 300 infertile and 399 control men. Using quantitative and qualitative assays, we identified 34 cases of partial *AZFc* deletion: 32 lack some copies of both *DAZ* and *CDY1*, while two lack some copies of either *DAZ* or *CDY1*.

METHODS

Infertile men and controls

We studied five different groups: (i) 254 individuals with unexplained infertility and reduced or absent sperm counts of $0–5\times10^6$ spermatozoa (spz)/ml (composed of 94 individuals with $0–0.1\times10^6$ spz/ml, 73 with $>0.1–1\times10^6$ spz/ml, and 87 with $>1–5\times10^6$ spz/ml); (ii) 46 individuals with unexplained infertility and sperm counts of $5–131\times10^6$ spz/ml; (iii) 210 unselected men of unknown fertility status; (iv) 185 fertile men with one or more children; and (v) four normospermic men. Samples were collected using approved protocols and the informed consent of all individuals was obtained. The required approval for this study was obtained from the French committee for the Protection of Persons in Biomedical Research (CCPPRB). The infertile men were drawn from a larger group of 361 individuals recruited from two infertility clinics in Marseilles and one in Toulon, France. For all patients, urogenital examination, karyotype Y microdeletion status, and sperm parameters, according to the criteria of the WHO (World Health Organization), were determined. These analyses allowed us to exclude 61 individuals: 28 obstructive azoospermic cases (including 11 cases of bilateral agenesis of the vas deferens associated with mutations in both alleles of *CFTR*), 13 47,XXY and one 47,XXY/46,XX/46,XY Klinefelter cases, three 47,XYY men, 10 unrelated *AZFc* deletions, four *AZFb+c* deletions, one partial *AZFb* deletion, and one *AZFa* deletion.

The control samples are mostly from local sources, but include 57 unrelated fathers from CEPH (Centre d'étude du polymorphisme humain) reference families. The unselected men of unknown fertility were anonymous blood donors from the local blood transfusion service (Etablissement Français du Sang Alpes Mediterranée). The proportion of Y haplogroups in the CEPH fathers is similar to our local populations, except for Hg-DE, which is more frequent in local controls (9.5%–19% local *v* 5% CEPH). The CEPH father, 142001, with the gr/gr deletion has seven children.

Sequence family variant (SFV) analysis

We screened for SFVs in *DAZ* and *CDY1*. For *DAZ*, we chose the SFV at the sequence tagged site sY587 in intron 10,[18] also referred to as ''DAZ-SNV V''.[15] The sY587 SFV discriminates *DAZ1/2* from *DAZ3/4*, and its position outside the intragenic duplications that affect exons 2–8 means that sY587 is present in only one copy per *DAZ* gene. To confirm the sY587 results, we also sequenced all men with a partial *AZFc* deletion for an SFV 93 bp upstream from exon 2 of the *DAZ* gene. This SFV also differentiates *DAZ1/2* (A) and *DAZ3/4* (G). The results agree with those obtained for sY587. Since the Y chromosome reference sequence shows no SFV within the transcribed portion of the *CDY1* gene, we used a C/A SFV situated 7750 bp 5′ of the *CDY1* translation start codon (*CDY1-7750*). We also used an SFV that we discovered in the 5′ UTR of *CDY1* when sequencing *CDY1* in our patients with partial *AZFc* deletions, and it was the only polymorphism that we discovered (MJM and EC, unpublished data). SFVs were scored by PCR followed by restriction enzyme digestion (2 units for 2 h): *DAZ* sY587, *Dra*I (*DAZ*1/2 cut); *CDY1-7750*, *Pvu*II (*CDY1b* cut); and *CDY1*-84, *Alu*I (*CDY1a* cut). To prevent cross-amplification of *CDY2*, the *CDY1*-84 SFV was scored by two rounds of PCR, the first being specific for *CDY1*, followed by digestion. The SFV *DAZ* exon2–93 was scored by sequencing. Primer pairs: sY587, o912/o913; *CDY1*-7750, o1025/o1026; *DAZ* exon2-93, o1226/o1227; *CDY1*-84, 1° o1187/o1188 (30 cycles), 2° o1187/o1195 (10 cycles).

Cases of SFV absence were confirmed by sequencing, and no additional SNPs were detected.

Fluorescent in situ hybridisation

We performed FISH using the *DAZ* cosmid 18E8[18] to probe interphase nuclei from lymphoblastoid lines. In the majority of interphase nuclei, which are at G1, 18E8 gives a single hybridisation signal for each *DAZ* gene pair present. In most men, where there are two *DAZ* gene pairs, cosmid 18E8 gives two signals on the majority of interphase nuclei.[14]

FISH was performed on interphase nuclei obtained using standard cytogenetic techniques on non-synchronised lymphoblastoid cell lines. The *DAZ* cosmid 18E8 was labelled with biotin-14dCTP (Bioprime DNA labelling system, Invitrogen, France), hybridised using a standard protocol and revealed by means of fluorescein isothiocyanate-conjugated avidin. Over 100 nuclei were screened for each patient. Nuclei were counterstained with propidium iodide diluted in anti-fade solution. Preparations were observed using an axioplan-2 Zeiss fluorescent microscope and the images were captured with a SenSys CCD camera (Photometrics, Tucson, AZ, USA).

PCR based quantification of copy number

In order to quantify copy number using PCR, we simultaneously amplified the *AZFc* locus to be quantified and a homologous locus outside the *AZFc* interval, as an internal standard, using a single primer pair in a standard PCR reaction of 28–32 cycles. The primers flank an insertion/deletion difference of 3–5 bp, which allowed the products amplified from the *AZFc* loci and the control loci to be separated by polyacrylamide gel electrophoresis. One of the primers was labelled at its 5′ end with a fluorochrome (IRDye 800, LI-COR, Lincoln, NE, USA) and the other carried a GTTTCTT tag at its 5′ end. This short tag has been reported to stimulate the *Taq* polymerase to add an adenosine residue to the end of the complementary (labelled) strand, resulting in the adenylation of almost all labelled fragments.[19] This overcame the problem of doublet bands caused by heterogeneous adenosine addition, which had initially made quantification difficult. Following PCR, the reaction was mixed with 95% formamide, denatured at 95°C for 5 min, and the different sized loci separated on an automatic sequencer Gene ReadIR 4200 (LI-COR). Quantification was performed by ONE-DScan (Scanalytics, Fairfax, VA, USA).

We developed five quantitative tests of this type. The positions of the loci screened are shown in figs 1 and 2.

CDY1-cds v CDY2 (primers oMJ953a/o1023)

There are two identical copies each of *CDY1* and *CDY2*, which share 98% nucleotide identity.[20] *CDY2* is not affected by the complete *AZFc* deletion. We amplified *CDY1* and *CDY2* across a 3 bp indel difference in the coding region, to give fragments of 134 bp for *CDY1* and 137 bp for *CDY2*. All men were screened with this test.

CDY1-9825 v CDY2 (primers o1207/o1208)

A further CDY test that was used to screen all cases of SFV loss at *CDY1-7750*.

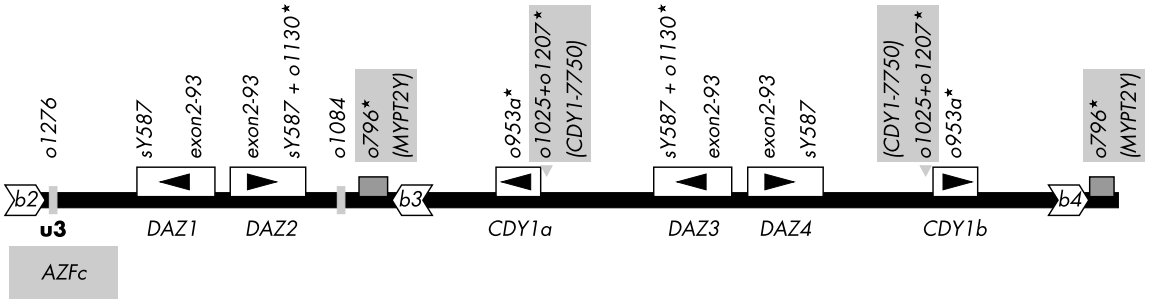

Figure 1A Qualitative and quantitative analyses of the *AZFc* interval. Schematic representation of the *AZFc* interval, showing the relative position of the loci analysed. Tests developed in this study are designated by their forward primer (see Methods) and dose tests are indicated with an asterisk.

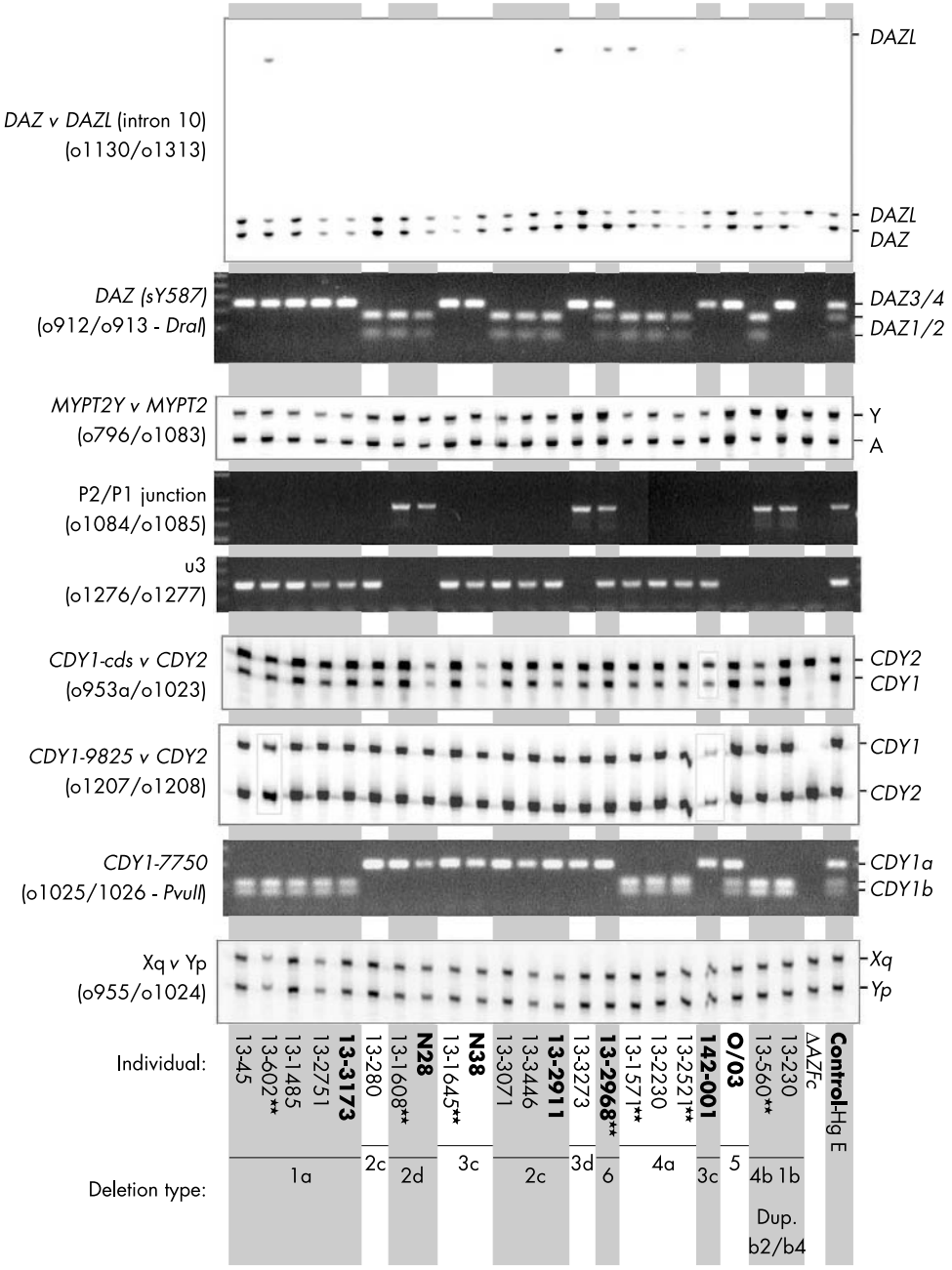

Figure 1B The results obtained with 22 individuals with a partial *AZFc* deletion are shown, together with a case of b2/b4 deletion (complete *AZFc* deletion) and an undeleted case. Identifiers of individuals from control populations are shown in **bold**. **Results confirmed by FISH with *DAZ* cosmid 18E8. The extra fragment observed in the *DAZ v DAZL* test corresponds to a 40 bp insertion polymorphism in intron 10 (see Methods). In these cases, the fragments amplified from each *DAZL* allele have the same intensity.

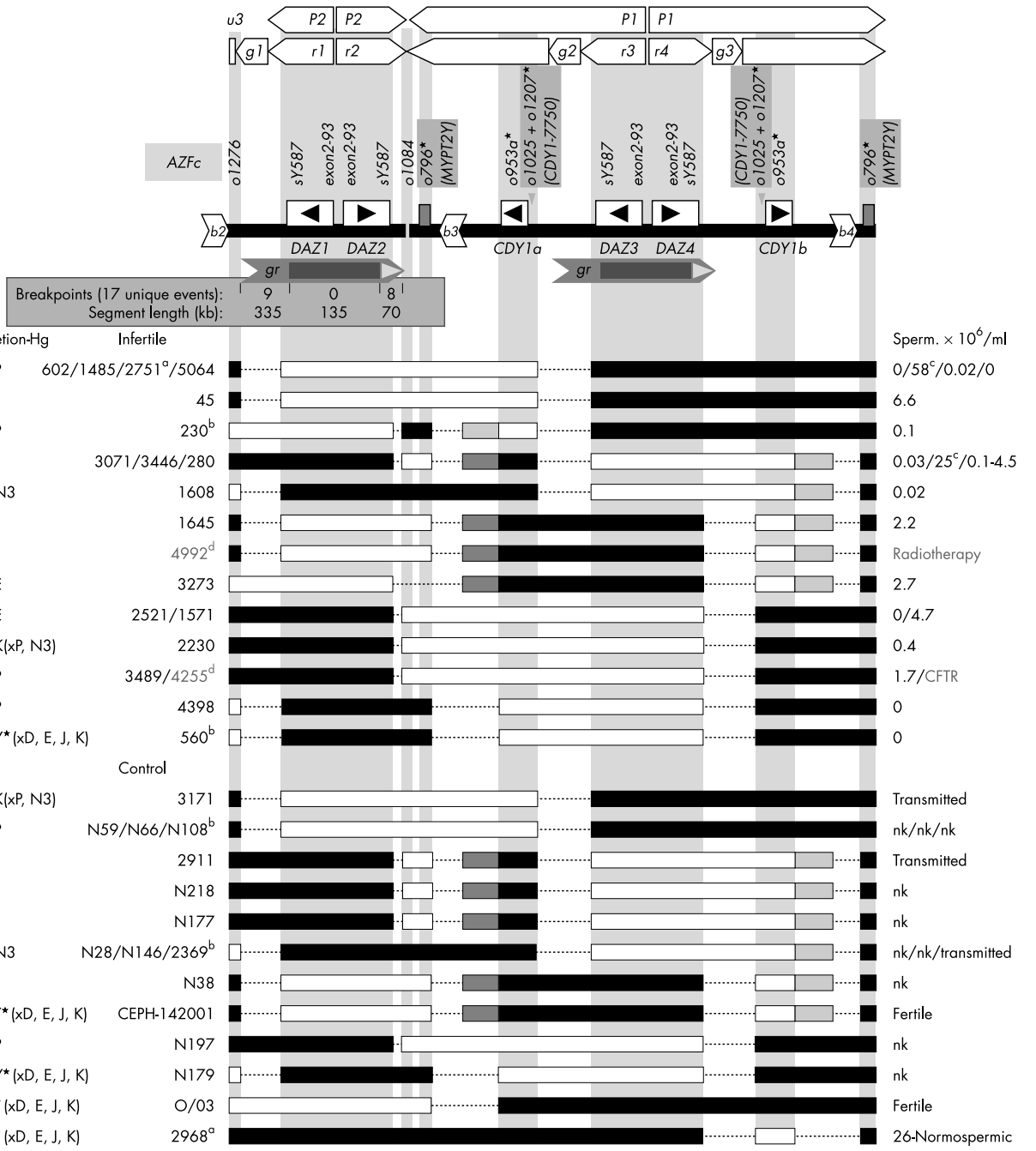

Figure 2 The six types of *DAZ/CDY1* partial deletions of the *AZFc* interval found in the fertile and infertile populations. The deletions are shown relative to a schematic representation of the *AZFc* interval.[4] Large arrows, labelled ''gr'', under the gene map indicate the position of the large direct gr amplicon repeats, between which unequal crossing over is proposed to mediate gr/gr deletions.[16] No breakpoints were found within the region of the *DAZ* genes indicated by the dark grey box within the ''gr'' arrow, whereas the pale grey triangle indicates a hotspot of unequal crossing over. Amplicons b2 and b4 mark the boundaries of the *AZFc* interval. The various SFVs and copy number tests (*) used in this study are shown above the *AZFc* gene map. Deleted segments are represented by white bars, and retained segments by black bars. Pale grey and dark grey bars indicate, respectively, that the segment is designated deleted, or retained, based on the assumption that the chromosome carried an inversion polymorphism (fig 4) prior to deletion. The dotted lines represent regions that were not tested directly and whose presence or absence could not be extrapolated from the results obtained with flanking markers. Deletion type 1–6, and inversion type a, b, c, or d (fig 4) are indicated to the left of each deletion, together with the Y haplogroup and the identifier of the individuals. The sperm count, where known, is shown to the right, in the same order as the identifiers. ''nk'' means fertility status is not known. [a]Father has been tested and has the same deletion; [b]cases of b2/b4 duplication following gr/gr deletion; [c]although 13-1485 and 13-3446 produce normal numbers of spermatozoa, 13-1485 showed a moderate asthenospermia and >90% teratospermia, while 13-3446 showed a severe asthenospermia and a moderate teratospermia; [d]indentifiers in grey refer to men who were excluded from the study because they had obstructive azoospermia: 13-4992 (radiotherapy for testicular cancer) and 13-4255 (bilateral agenesis of the vas deferens). Diagram is not to scale.

DAZ v DAZL (primers o1130/o1313)

In order to dose *DAZ*, we co-amplified a fragment of intron 10 from *DAZ* and *DAZL*. This intron contains the SFV at sY587 and is present in one copy per *DAZ* gene. Thus, from a 46,XY DNA with an intact *AZFc* interval, we would expect to amplify four copies of *DAZ* for every two copies of *DAZL* (66%:33%). The expected fragment sizes are 184 bp for *DAZ* and 187 bp for *DAZL*. When we performed the test, an extra band at 230 bp was observed in 144 out of 575 individuals. We determined, by sequencing, that this fragment results from a 40 bp insertion polymorphism in the *DAZL* intron 10, since in a given individual it has the same intensity as the 187 bp *DAZL* fragment, which has half its expected intensity. Furthermore, there are eight individuals with the 230 bp band who lack the 187 bp *DAZL* band altogether, and their observed frequency (0.014) is close to the expected homozygote frequency (0.018) for a biallelic locus with a minor allele frequency of 0.134 (154 out of 1152 chromosomes tested).

MYPT2Y v MYPT2 (primers o796/o1083)

We derived a test for the ends of the palindrome P1. This locus is present in two copies on the Y chromosome, where the proximal copy is situated between the amplicons b3 and r2 (*DAZ2*) and the distal copy is situated distal to amplicon b4 and, thus, outside the *AZFc* interval. This locus is part of a block with 98% identity to the *MYPT2* gene on chromosome 1. The gene on the Y chromosome is a pseudogene that we refer to as *MYPT2Y*-ps. This test amplifies *MYPT2Y*-ps (173 bp), together with *MYPT2*, on chromosome 1 (168 bp), and was used to screen patients with a suspected deletion.

Yp v Xq (primers o955/o1024)

This test is derived from the *PCDHX* and *PCDHY* genes on Yp and Xq. It controls for high levels of 46,XY/45,X mosaicism which, if present, could lead to false deletion detection by the other dose tests. This possibility was excluded for all individuals except two males of unknown fertility who do not have a partial *AZFc* deletion.

Locus specific PCR flanking *DAZ1/2*

We performed two locus specific PCR assays. We derived primers (oMJ1084/oMJ1085) from either side of the P2/P1 junction fragment situated 70 kb distal to the 3′ end of *DAZ2*. The presence of this site specific junction fragment was tested for in all groups. Only one fertile Hg-J man who had not previously shown evidence of deletion failed to amplify this fragment, indicating that there is a low frequency of Y chromosomes (<0.2%) which lack this evolutionary junction fragment. This test corresponds to sY1291.[16]

The junction situated proximal to *DAZ1* is defined by a sequence block that is duplicated on Yp. There is, however, a deletion/insertion difference between the Yp and the *AZFc* homologue that allowed us to develop a specific PCR test for *AZFc* with primers oMJ1276/oMJ1277. This region is part of u3 on the published *AZFc* map[4] and is situated close to the 50f2/C locus.[21] It corresponds to sY1191[16] or sY1192.[4] To determine the frequency of deletion polymorphisms at this locus in our population, we tested 242 non-deleted men (132 Hg-P, 61 Pg-Y*{xD,E,J,P}, 20 Hg-J, and 29 Hg-DE). No deletions were found (data not shown).

PCR primer pairs

The 5′ to 3′ sequences of the unpublished oligonucleotide primers used in this study are: o796: ctcactacatgacattcagg; o912: tgtatttaaaatgtgcacttcactgt; o913: ccagtcacaaaatgccacat; o953a: tattgagacccttgcacctg; o955: aatttcttcacagccagacagc; o1023: gtttcttggagtttcccttctgtcacc; o1024: gtttcttgcactattcattg cacagac; o1025: gaaatgccataatgtgctaacactg; o1026: aagga gagtgttaatacataccctg; o1083: gtttcttcccagtatctagtacagtgc; o1084: agaactgccaggtctgtgtc; o1085: ttgccacaaagagatagctctgg; o1130: ttaagtactactgtagacac; o1187: gttcaggcacatggcatagc; o1188: ccactcacagttttgtggtc; o1195: cttcaacctcaaactcctgg; o1207: gtttcttccactgtagaaattcacctcc; o1208: gaagtttgcatagtg gacagc; o1226: aactgagctgcactggtgac; o1227: ccccgaatgaccag cagcc; o1276: agtctgagtggctagtgcac; o1277: gaagcaaagtt cagctgtgc; o1313: gtttcttgtataatgtagaagagtagagc.

The following annealing temperatures were used: o796/o1083: 56℃; o912/o913: 54℃; o953a/o1023: 58℃; o955/o1024: 54℃; o1025/o1026: 58℃; o1084/o1085: 60℃; o1130/o1313: 54℃; o1187/o1188: 60℃; o1187/o1195: 58℃; o1207/o1208: 56℃; o1226/o1227: 60℃; o1276/o1277: 60℃.

Y haplotyping

Y haplotyping was performed as previously published for the YAP,[22] M9,[23 24] SRY4064,[25 26] and 92R7[25 26] polymorphisms. The Tat polymorphism[27] was assayed by PCR digestion with primers o1489: atgtatatagtacgtctgtagg and o1490: gtaagca taattgagaaggtgcc (annealing at 54℃). The 12f2 assay was performed as a duplex with *SRY* primers 3′SRY15 and 3′SRY16 and 12f2 primers 12f2D and 12f2F.[26] These primer pairs were used at a concentration of 300 nM for each *SRY* primer and 600 nM for each 12f2 primer and the annealing temperature used was 52℃. Polymorphisms were visualised by restriction enzyme digest for M9 (*Hin*fI), SRY4064 (*Bsr*BI), Tat (*Hpy*CH4IV), and 92R7 (*Hin*dIII).

All men were haplotyped for YAP, 12f2, and 92R7 polymorphisms (table 1), and individuals with partial AZF deletions were further genotyped with the Tat, M9, and SRY-4064 SNPs, defining five haplogroups, E, J, K(xN3,P), N3, and P, and one paragroup Y*(xD,E,J,K).

Table 1 Y haplogroup distribution in the four study populations and numbers of microdeletions (complete *AZFc*, or b2/b4, deletions) and partial *AZFc* deletions

	Y Haplogroup											
	P			J			DE			Y*(xD,E,J,P)		
Population	(micro)	n	(part)	(micro)	n	(part)	(micro)	n	(part)	(micro)	n	(part)
⩽5 M/ml (n=263)	(7)	141	(6)	(1)	28	(1)	(0)	47	(5)	(2)	47	(3)
% of total		54			11			18			18	
>5 M/ml (n=46)	(0)	23	(1)	(0)	4	(1)	(0)	10	(1)	(0)	9	(0)
% of total		50			9			22			19	
Unknown fertility (n=210)	(0)	121	(6)	(0)	14	(1)	(0)	20	(0)	(0)	55	(3)
% of total		58			7			10			26	
Fertile/ normospermic (n=189)	(0)	93	(0)	(0)	21	(0)	(0)	25	(1)	(0)	50	(5)
% of total		49			11			13			26	

micro, microdeletion; part, partial deletion.

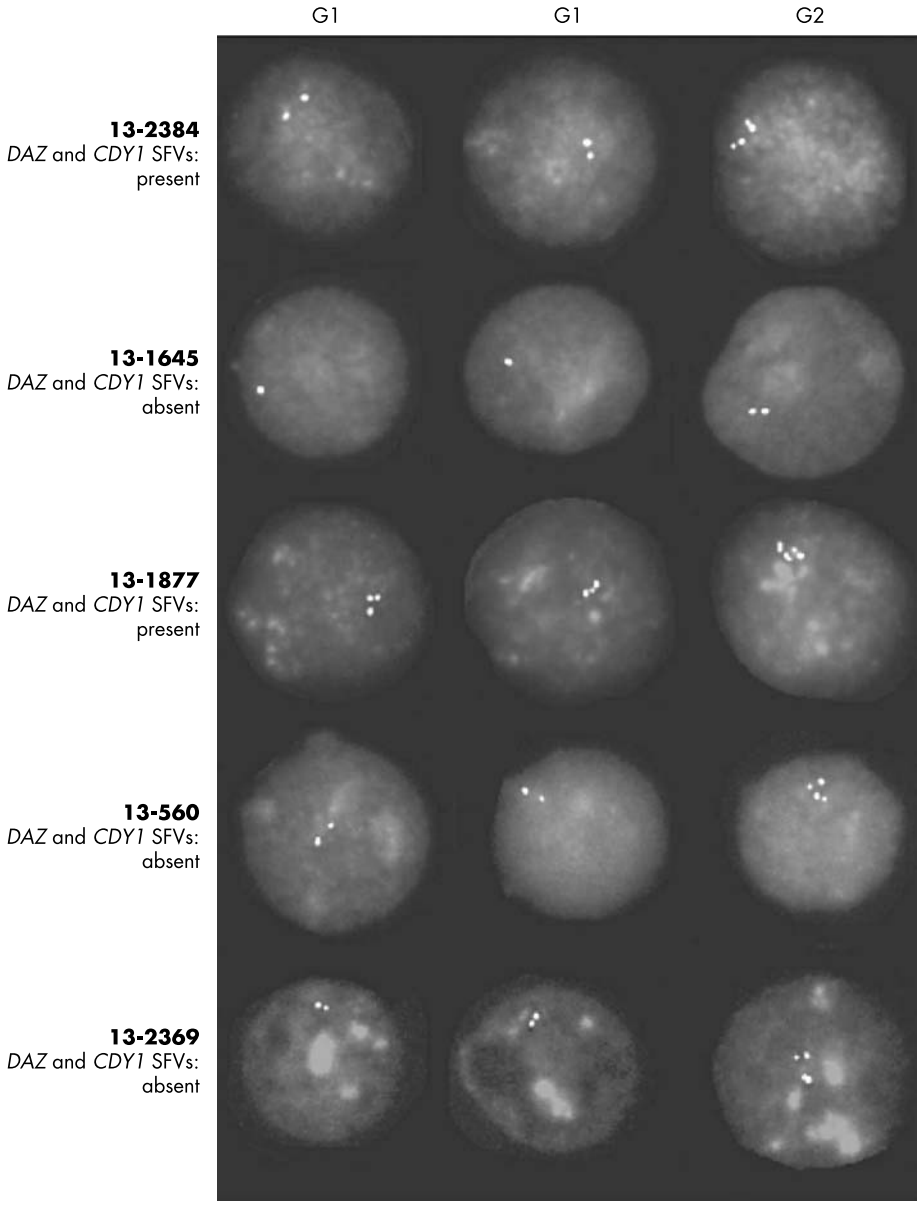

Figure 3 Quantification of *DAZ* gene pairs in interphase nuclei by FISH, using the *DAZ* cosmid 18E8 as probe. Each spot corresponds to a *DAZ* gene pair. Representative G1 nuclei are shown, as well as nuclei considered to be at G2 on the basis of hybridisation pattern.

Statistical analysis

We tested the significance of the observed difference in the incidence of gr/gr deletion or SFV absence between our infertile and control populations by means of Fisher's exact test, using the program developed by Ø Langsrud (http://www.matforsk.no/ola/fisher.htm). Our null hypothesis was that incidence is the same in infertile and control populations. Given the exploratory nature of this study, all associations with $p<0.05$ are considered as significant, and worth further study. In this study, however, we have performed 11 tests of association with infertility: four *DAZ+CDY1* gr/gr deletion types, loss of four SFV sites at *DAZ* and *CDY1*, loss of a *CDY1-7750* SFV (a or b) with, and without, copy loss, and presence of a gr/gr deletion. Applying the Bonferroni correction for multiple testing ($\alpha = 0.05$/number of tests performed), an α of 0.0045 is obtained. Thus, $p<0.0045$ is considered the true statistical significance for this study.

RESULTS

Screening for partial *AZFc* deletions

In order to investigate the incidence, extent, and impact on fertility of partial *AZFc* deletions in a large number of fertile and infertile individuals, we characterised the apparently intact Y chromosomes in five groups of men in four different ways: (i) quantification of five *AZFc* loci, including *DAZ* and *CDY1*, using quantitative PCR and FISH; (ii) detection of nucleotides (SFVs) that discriminate between members of the *DAZ* and *CDY1* gene families; (iii) the specific detection of unique fragments flanking the *DAZ1/2* gene pair at the u3 segment (proximal) and the P2/P1 palindrome junction (distal), corresponding, respectively, to sY1191/sY1192 and sY1291 used in previous studies[16 28]; and (iv) haplotyping with MSY binary polymorphisms. Some of these results, and a summary of the Y chromosome variants defined in this study, are presented in figs 1–3.

Most partial *AZFc* deletions of *DAZ* or *CDY1* are gr/gr deletions

This combination of tests was initially used to screen 487 men. All 487 men were screened with three quantitative tests: *DAZ v DAZL*, *CDY1*-cds *v CDY2*, and Yp *v* Xq. This latter test controls for cases of 46,XY/45,X mosaicism. Some 18 men showed copy loss at both *DAZ* and *CDY1*, but not Yp. These men all showed SFV loss at both *DAZ* (sY587-[18]) and

CDY1-7750 and were deleted for either one of the unique markers flanking the *DAZ1/2* locus. We conclude that these 18 men carry partial *AZFc* deletions of the Y chromosome that correspond to the gr/gr deletions recently described.[16]

The original definition of the gr/gr deletion did not predict the deletion of the u3 segment.[17] However a more recent study has shown that the Y chromosome of haplogroup Hg-N3 probably carries a gr/gr deletion.[28] The deletion of u3 was explained in this case by invoking the existence of an inversion around *DAZ1/2*, mediated by recombination between elements b2 and b3. Our results confirm this finding and, moreover, show that u3 deletion is not limited to Hg-N3 but is found on other types of Y chromosome (fig 2). We conclude that deletion of either u3 or the P2/P1 junction, but not both, is diagnostic of gr/gr deletion.

Only three men who did not show copy loss for *DAZ* or *CDY1* showed SFV loss at both these loci. These men were, however, deleted for the u3 locus specific marker indicating that their Y chromosome carries a gr/gr deletion, but that the *DAZ* and *CDY1* copy number has been restored by a subsequent b2/b4 duplication.[16] We did not find any cases with SFV loss at sY587 in *DAZ* and *CDY1* that were not also deleted for one of the unique sites, u3 or P2/P1, that flank *DAZ1/2*, suggesting that in most cases either of these approaches can be used to detect gr/gr deletions. Indeed, the loss of the *DAZ3/4* variant at sY587 is diagnostic of gr/gr deletion, while the loss of the *DAZ1/2* variant was found in only three undeleted men. Loss of the *CDY1* SFV is a much less reliable indicator of deletion, since, of the 49 cases detected, only 22 have a partial *AZFc* deletion. Copy number status was verified at *CDY1-7750* using a further quantitative test, *CDY1-9825 v CDY2*, and the results obtained at these two loci are in complete accord. This effectively excludes the possibility that small deletions, specific to the 5′ end of *CDY1*, are the cause of SFV absence at this site.

Two exceptional individuals showed deletions that affect either *DAZ* or *CDY1*, but not both: the normospermic man, 13-2968, (26×10^6 spz/ml) is deleted for *CDY1b* only, while the fertile control, O/03, is deleted for the *DAZ1/2* gene pair only. It is likely that O/03 is a second case of the b1/b3 deletion reported recently,[16] but confirmation of this will require the mapping of his proximal boundary. The deletion carried by 13-2968 has not been previously reported, and shows that two copies of *CDY1* are not required for normospermia.

Our results show that the vast majority of partial *AZFc* deletions that result in the loss of some copies of *DAZ* and *CDY1* are gr/gr deletions, mediated by unequal crossing over between two large (~540 kb) directly repeated gr amplicons (fig 2). We also show that our combination of SFV tests and locus specific tests provides a stringent method for their detection. We therefore screened a further 212 individuals with SFV and locus specific tests only. Quantitative tests were performed in gr/gr deletion cases, as a means of detecting the b2/b4 duplication. In this group we found a further 11 gr/gr deletions, one of which had a b2/b4 duplication.

All men in whom we detected a partial *AZFc* deletion were tested with a further quantitative test: *MYPT2Y v MYPT2* situated at the end of palindrome P1. Copy loss at *MYPT2Y* was observed in men with a gr/gr deletion who are deleted for the P2/P1 junction, but not in cases of u3 deletion. Indeed, in the three men predicted to have a b2/b4 duplication, whose deletions must have arisen on a b2/b3 inverted Y chromosome, since they include u3 and not the P2/P1 junction, we saw increased copy number of the *MYPT2Y*-ps locus (shown for 13-560 and 13-230; fig 1B). The proximal copy of *MYPT2Y*-ps would not be deleted from a b2/b3 inverted Y chromosome, and would therefore be duplicated by a b2/b4 duplication, to give three copies instead of two. This provides further evidence for the existence of the b2/b3 mediated inversion polymorphism[28] and b2/b4 duplications.[16]

Our cumulative results are summarised in fig 2 and tables 1 and 2. Overall we detected 32 gr/gr deletions, four of which carry b2/b4 duplications. Based on SFV loss and Y haplogroup, we are able to discern 17 independent gr/gr deletions.

Quantification of *DAZ* gene pairs by FISH

FISH analysis with the DAZ cosmid 18E8 was performed on four infertile individuals without SFV loss at either sY587 or *CDY1-7750* and on one fertile and six infertile individuals (** in fig 1B) with SFV absence at both these loci. This confirmed our PCR based analyses. Representative results are shown in fig 3.

One infertile man without SFV loss, 13-1877, showed 62% of nuclei with three or more signals, and only 35% of nuclei with two signals. We conclude that 13-1877 has three pairs of *DAZ* genes, and this was corroborated by PCR based dose tests (data not shown). A similar *DAZ* gene pair duplication has recently been described in a fertile sperm donor.[29] This duplication is most likely the result of the event reciprocal to the gr/gr deletion.

Loss of *DAZ1/2* or *DAZ3/4*

Using sets of SFVs that discriminate between each *DAZ* gene, other studies have concluded that partial *AZFc* deletions can result in the loss of one, two, or three *DAZ* genes.[14 15] Our combined quantification and SFV analysis of the *DAZ* gene family in 487 men allowed us to address this question.

Out of 487 men, we identified 19 individuals with reduced *DAZ* copy number, who all lack the SFV at sY587. Since the SFV at sY587 differentiates *DAZ1* and *DAZ2* from *DAZ3* and *DAZ4*, this shows that the vast majority of partial *AZFc* deletions remove either the *DAZ1/2* gene pair or the *DAZ3/4* gene pair, but not combinations of two *DAZ* genes from each gene pair. Since there are normally four *DAZ* genes and two *DAZL* genes, *DAZ*:*DAZL* ratios of 66%:33%, 50%:50%, or 33%:66% are expected, respectively, in cases where no, two, or three *DAZ* genes have been deleted. Since we never observed the 33%:66% ratio, we conclude that deletions of three *DAZ* genes are extremely rare. Although there is no obvious mechanism by which single *DAZ* gene deletions could arise, we cannot formally exclude their existence, since the expected difference in *DAZ*:*DAZL* dose ratios, 60%:40% as compared to 66%:33%, is too slight to measure reliably.

Inversion polymorphism in the *DAZ3/4* palindrome P1

In the Y chromosome reference sequence,[20 30] *CDY1b* lies outside the region that is predicted to be deleted by gr/gr recombination. We nevertheless detected 14 partial *AZFc* deletions that result in the loss of *CDY1b*, but that otherwise have all the characteristics of ''classic'' gr/gr deletions. These deletions involving *CDY1b* can, therefore, be explained either by invoking a second inversion polymorphism of the P1 palindrome, or the gene conversion of *CDY1a* by *CDY1b* prior to deletion. This latter possibility is very unlikely, as *CDY1b* deletions represent 14/32, or 48%, of the gr/gr deletions, and the observed frequency of this gene conversion event is <1%. It therefore seems likely that there is a further inversion polymorphism that may be mediated by crossing over between the homologous inverted segments b3 and b4. This suggests that there are at least four possible orders for sequences in this interval, and we find evidence for all four possibilities in the current study (fig 4). In light of these inversion variants, which place amplicon g3 within a gr amplicon (fig 4), we suggest that the original definition of gr/gr deletion[16] should be extended to include any deletion that has arisen through unequal crossing over within g or r

Table 2A Percent SFV absence

	Population							
	⩽5 M/ml		>5 M/ml		Unknown		Fertile/normo	
	n=254		n=46		n=210		n=189	
SFV absent	%	(n)	%	(n)	%	(n)	%	(n)
DAZ1/2	2.4	(6)	4.3	(2)	1.9	(4)	3.2	(6)
DAZ3/4	3.5	(9)	2.2	(1)	2.8	(6)	1.1	(2)
CDY1a	9.4	(24)	6.5	(3)	4.3	(9)	2.7	(5)
CDY1b	3.1	(8)	4.3	(2)	2.4	(5)	2.1	(4)
DAZ1/2+CDY1a	1.6	(4)	4.3	(2)	1.4	(3)	0.5	(1)
DAZ1/2+CDY1b	0.8	(2)	0	(0)	0.5	(1)	0.5	(1)
DAZ3/4+CDY1a	2.4	(6)	0	(0)	1.0	(2)	0	(0)
DAZ3/4+CDY1b	1.2	(3)	2.2	(1)	1.9	(4)	1.1	(2)

normo, normospermic.

Table 2B Percent SFV absence with deletion

	Population							
	⩽5 M/ml		>5 M/ml		Unknown		Fertile/normo	
	n=254		n=46		n=210		n=189	
SFV absent	%	(n)	%	(n)	%	(n)	%	(n)
DAZ1/2	2.4	(6)	4.3	(2)	1.9	(4)	1.6	(3)
DAZ3/4	3.5	(9)	2.2	(1)	2.9	(6)	1.1	(2)
CDY1a	3.9	(10)	4.3	(2)	2.4	(5)	0.5	(1)
CDY1b	2.0	(5)	2.2	(1)	2.4	(5)	2.1	(4)
DAZ1/2+CDY1a	1.6	(4)	4.3	(2)	1.4	(3)	0.5	(1)
DAZ1/2+CDY1b	0.8	(2)	0	(0)	0.5	(1)	0.5	(1)
DAZ3/4+CDY1a	2.4	(6)	0	(0)	1.0	(2)	0	(0)
DAZ3/4+CDY1b	1.2	(3)	2.2	(1)	1.9	(4)	1.1	(2)

normo, normospermic.

amplicons. The fact that even in only 32 men with a gr/gr deletion each haplogroup defined contains at least two inversion variants suggests that these inversions are relatively frequent events.

Genetic exchange within the non-recombining Y chromosome

A previous study concluded that SFV absence was the result of gene conversion between the arms of the large palindromes that make up most of the *AZFc* interval, but internal *AZFc* deletions in the fertile population were not considered as a possible alternative explanation, and no quantification was performed at the loci studied.[31] Here, we establish that the majority of cases of SFV absence at *CDY1-7750* are not associated with gr/gr deletion or copy loss. We therefore further investigated this question for two SFVs at the *CDY1* locus, in our population of individuals known to carry two copies of *CDY1*. The SFVs used were *CDY1-7750* and *CDY1-84*, from the 5′ UTR of *CDY1*. This latter polymorphism has been published independently.[31]

A comprehensive phylogeny of human Y chromosomes, based on the genotyping of MSY binary polymorphisms, is now available.[25] The genotyping of our study population allowed us to assign individuals to specific branches of the phylogeny and determine the relative position of SFV loss or acquisition.

The results for 465 undeleted men are summarised in fig 5. Our analysis shows that the *CDY1* SFVs used are present on

Table 2C Percent SFV absence without deletion

	Population							
	⩽5 M/ml		>5 M/ml		Unknown		Fertile/normo	
	n=254		n=46		n=210		n=189	
SFV absent	%	(n)	%	(n)	%	(n)	%	(n)
DAZ1/2	0	(0)	0	(0)	0	(0)	1.6	(3)
DAZ3/4	0	(0)	0	(0)	0	(0)	0	(0)
CDY1a	6.7	(14)	2.2	(1)	1.9	(4)	2.1	(4)
CDY1b	1.4	(3)	2.2	(1)	0	(0)	0	(0)
DAZ1/2+CDY1a	0	(0)	0	(0)	0	(0)	0	(0)
DAZ1/2+CDY1b	0	(0)	0	(0)	0	(0)	0	(0)
DAZ3/4+CDY1a	0	(0)	0	(0)	0	(0)	0	(0)
DAZ3/4+CDY1b	0	(0)	0	(0)	0	(0)	0	(0)

normo, normospermic.

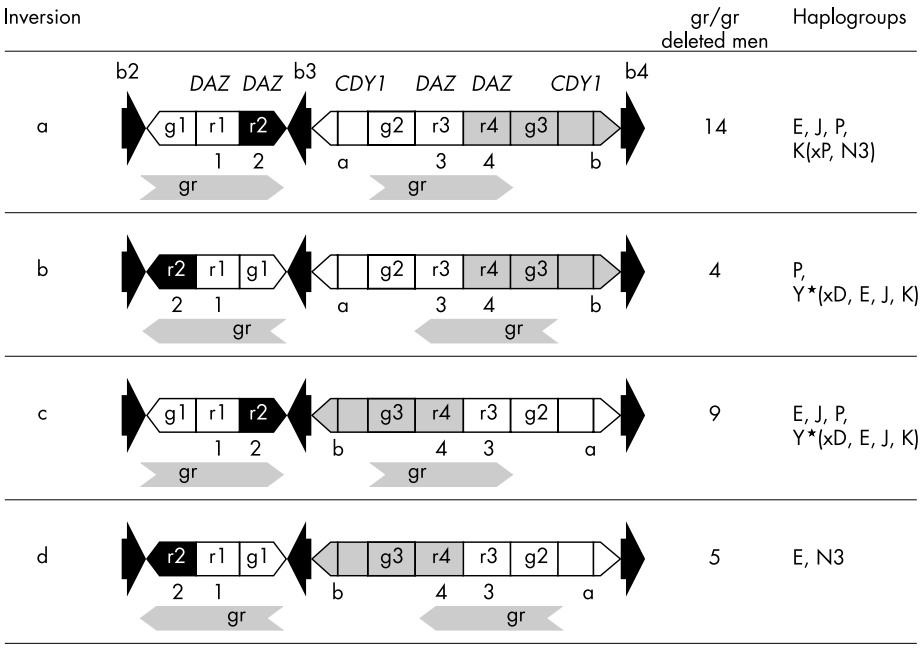

Figure 4 Inversion polymorphisms of the *AZFc* interval inferred from SFV analysis of gr/gr deleted Y chromosomes. The position of the gr amplicons, between which unequal crossing over would result in a gr/gr deletion, are shown.

all four Y chromosomes defined, demonstrating that the presence of these two *CDY1* SFVs is the ancestral state of all the derived Y chromosomes. Furthermore, SFV loss is not limited to the non-derived Y chromosomes, but is found in all four groups. We conclude that the majority of SFV loss at the *CDY1* loci tested does not result from deletion, but that, as previously proposed,[31] it is frequently a consequence of intrapalindrome gene conversion.

It is clear from our data that for an SFV to be a useful indicator of partial deletion, the incidence of its absence from intact Y chromosomes must be estimated for the Y haplotypes that compose the study group. This has not been done in other studies.[14 15]

DAZ and *CDY1* copy number on the gr/gr deleted Hg-N3 Y chromosome

It has been proposed that b2/b4 duplication may be a way in which full spermatogenic activity can be restored to an individual carrying a gr/gr deleted Y chromosome.[16] If gr/gr deletions have an effect on spermatogenesis by reducing gene

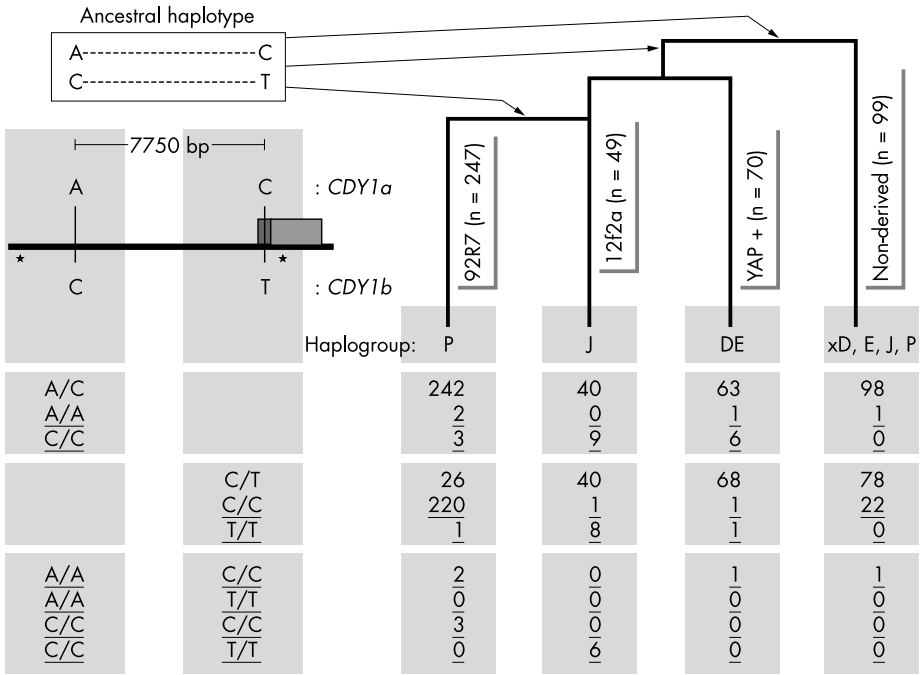

Figure 5 SFV loss on intact Y chromosomes by intrachromosomal genetic exchange at two sites 5′ to *CDY1* in the *AZFc* interval. The number of individuals observed for each genotype at the SFV site is shown for each Y haplogroup determined. Genotypes with SFV absence are underlined. Asterisks indicate the positions of the two dose tests used. Also shown are the numbers of individuals who have the SFV at neither site. The presence of the SFV is found on each of the four Y chromosomes defined, showing that the SFV is ancestral and that its absence from a Y chromosome carrying two copies of *CDY1* must result from genetic exchange between the *CDY1* genes.

dosage, then a strong selection for these re-duplicated gr/gr deleted chromosomes would be expected, and consequently gr/gr deleted Y chromosomes present at a high percentage in a population would be the duplicated, rescued form. Two examples of gr/gr deleted Y chromosomes that have attained a high prevalence in certain populations have so far been described: the Hg-D2b Y chromosome present in 20–30% of Japanese men,[16 32] and the Hg-N3 Y chromosome present in 50% of Finnish men.[21 27 28] The gr/gr deletion on the Hg-D2b Y chromosome is associated with a sperm count of $>40\times10^6$ spz/ml in 66% of fertile carriers.[33] It has been shown by FISH that the D2b gr/gr deleted Y chromosome does not carry a b2/b4 duplication.[16] This information is, however, not yet known for the Hg-N3 Y chromosome.

Using the Tat bi-allelic marker, we identified three control men and one infertile man with an Hg-N3 Y chromosome, all of whom carry the same gr/gr deletion, type 4b, deleted for *CDY1b*, *DAZ3/4*, and u3. Our results confirm that the Y chromosome on which the Hg-N3 deletion arose carried a b2/b3 inversion around *DAZ1/2*,[28] but show in addition that it carried a second inversion of the P1 palindrome around *DAZ3/4*. Dose tests revealed that only one fertile Hg-N3 man carries a b2/b4 duplication, and thus has four copies of *DAZ* and two copies of *CDY1*, while three others have only two copies of *DAZ* and one copy of *CDY1*. There would therefore appear to have been no strong selection for the b2/b4 duplication among men with either Hg-D2b or Hg-N3 Y chromosomes. Indeed, we and others[16] have found that b2/b4 duplications are generally rare among the gr/gr deleted chromosomes, indicating that the majority of gr/gr deletions have little effect on male fertility.

DAZ and *CDY1* SFVs define four deletion types

Partial *AZFc* deletions of both *DAZ* and *CDY1* were not found with significantly different frequencies in infertile and control populations: 18 infertile men (6%) and 14 control men (3.5%) ($p = 0.085$) had such deletions. We therefore investigated the possibility that certain types of gr/gr deletion are associated with male infertility. The incidence of SFV loss without deletion at *DAZ* (sY587) or *CDY1-7750* loci was low (table 2C), indicating that these SFVs can be used to define deleted loci with a low error rate. On the basis of the *DAZ* and *CDY1* SFVs retained, we were able to define four gr/gr deletion types (fig 2 and table 2B).

Similar frequencies of gr/gr deletions of *CDY1b* with either *DAZ1/2* (0.7% *v* 0.5%) or *DAZ3/4* (1.3% *v* 1.5%) were observed for infertile versus control populations. Deletion of *CDY1a* with *DAZ1/2* is slightly more frequent in infertile populations (2.3% *v* 1.0%; $p = 0.14$). However, a statistically significant, but weak, association with infertility was observed only with the type 4 gr/gr deletion characterised by *DAZ3/4* and *CDY1a* SFV loss: 6/254 infertile men with $<5\times10^6$ spz/ml and 2/399 control men ($p = 0.042$) had this deletion. No fertile men carried this deletion (0/189). We did, however, find this deletion in a man with agenesis of the vas deferens who carries two mutations of the *CFTR* gene (13-4255; fig 2), further weakening this association between type 4 gr/gr deletion and infertility. Collectively, the incidence of gr/gr deletion of *CDY1a* shows a weak association with infertility: 12/300 infertile men *v* 6/399 control men ($p = 0.035$) had this deletion.

Loss of the *CDY1-7750* SFV is associated with infertility

Comparing our infertile and control populations, we found similar frequencies of SFV absence for *DAZ1/2*, *DAZ3/4*, or *CDY1b* (distal), with and without deletion (table 2A). There was, however, a significantly higher frequency in the infertile compared to the control populations for the absence of the *CDY1a* SFV (proximal) at *CDY1-7750* (27/300 *v* 14/399; $p = 0.002$). This result is significant even when the number of tests is taken into account (significance: $p<0.0045$; see Methods). This indicates that the absence of this SFV is associated with infertility. This increase is found for all four haplogroups defined (table 3A) and is at the limit of statistical significance in the largest, Hg-P (10/157 *v* 5/214; $p = 0.047$).

Table 2C shows the incidence of SFV absence without deletion. In men in whom there is no evidence of a partial *AZFc* deletion, absence of a *CDY1-7750* variant (a or b) is significantly more frequent in the infertile than the control population, even when multiple testing is taken into account (19/282 *v* 8/383; $p = 0.0026$). Absence of each variant at *CDY1-7750* is also significantly higher in infertile men compared to control men: 15/282 *v* 8/383 for *CDY1a* ($p = 0.02$) and 4/282 *v* 0/383 for *CDY1b* ($p = 0.03$). The higher incidence of *CDY1-7750* absence is seen in all three derived haplogroups (table 3B), suggesting that it is not an artefact of population stratification. This indicates that loss of a segment of the *AZFc* interval defined by *CDY1-7750* has a deleterious effect on male fertility. The mechanism for this loss is not clear and will require further study. Since SFV loss is not associated with copy loss in these cases, a deletion of a gr/gr duplicated chromosome, such as we describe in patient 13-1877, is one possibility. In this case, the absence of a locus flanking *CDY1-7750* could underlie the infertility. If, on the other hand, gene conversion were responsible, the SFV at *CDY1-7750* would be required for full male fertility.

DISCUSSION

Most gr/gr deletions have little effect on male fertility

Here, we show that the majority of partial *AZFc* deletions that affect copies of *CDY1* or *DAZ* are probably mediated by gr/gr amplicon recombination. Based on the *DAZ* and *CDY1* genes deleted, we sub-divided the gr/gr deletions found into four types. Deletions of *CDY1a* show a weak association with infertility, suggesting that some may cause infertility. On the other hand, deletions of *CDY1b*, representing 48% of the total

Table 3A *CDY-7750* SVF loss with and without copy loss

	Y Haplogroup											
	P			J			DE			Y*(xD,E,J,P)		
	CDY1-7750			*CDY1-7750*			*CDY1-7750*			*CDY1-7750*		
Population	(a)	Total	(b)	(a)	Total	(b)	(a)	Total	(b)	(a)	Total	(b)
⩽5 M/ml	(9)	134	(2)	(7)	28	(1)	(6)	47	(3)	(2)	47	(2)
>5 M/ml	(1)	23	(0)	(1)	4	(0)	(1)	10	(2)	(0)	9	(0)
Unknown fertility	(5)	121	(2)	(3)	14	(1)	(0)	20	(0)	(1)	55	(2)
Fertile/normospermic	(0)	93	(0)	(3)	21	(0)	(1)	25	(1)	(1)	50	(3)
Total	(15)	371	(4)	(14)	67	(2)	(8)	102	(6)	(4)	161	(7)

Table 3B *CDY-7750* SFV loss without copy loss

	Y Haplogroup											
	P			J			DE			Y*(xD,E,J,P)		
	CDY1-7750			*CDY1-7750*			*CDY1-7750*			*CDY1-7750*		
Population	(a)	Total	(b)	(a)	Total	(b)	(a)	Total	(b)	(a)	Total	(b)
⩽5 M/ml	(3)	128	(2)	(7)	27	(0)	(4)	42	(0)	(0)	41	(1)
>5 M/ml	(0)	22	(0)	(0)	3	(0)	(1)	9	(1)	(0)	9	(0)
Unknown fertility	(1)	115	(0)	(3)	13	(0)	(0)	20	(0)	(0)	52	(0)
Fertile/normospermic	(0)	93	(0)	(3)	21	(0)	(1)	24	(0)	(0)	45	(0)
Total	(4)	358	(2)	(13)	64	(0)	(6)	95	(1)	(0)	148	(1)

number of gr/gr deletions, are present in 2% of both infertile and control populations, suggesting that they represent variants of the Y chromosome with little effect on fertility. Our results do not exclude the possibility that certain gr/gr deletions, occurring on particular Y chromosomes, could be the cause of male infertility, but they do strongly suggest that most gr/gr deletions do not have a significant effect on male fertility. Our study therefore challenges the conclusion of a recent study,[16] that the majority of gr/gr deletions affect male fertility through a reduction in sperm production. It remains possible that all gr/gr deletions reduce sperm numbers. This question could best be addressed by comparing sperm counts in groups of fertile men carrying the same gr/gr deletion with and without b2/b4 duplications.

gr/gr recombination is not evenly distributed in the gr amplicon

It was originally assumed that the gr/gr unequal cross over frequency must be comparable to that of the b2/b4 event that generates the complete *AZFc* deletion because, at ~540 kb, the gr amplicon is more than twice the length of the b2 or b4 amplicon.[16] On this basis, it was estimated that gr/gr deletions must have a negative effect on infertility, otherwise they would be present on >40% of contemporary Y chromosomes. The fact that in 487 men we did not observe any *DAZ* gene copy loss without sY587 SFV loss shows that most gr/gr deletions do not occur through unequal cross over in the 135 kb segment situated between the two sY587 sites of each *DAZ* gene pair. In contrast, the type 1b, 2c, 3d, and 4a deletions (fig 2), representing 15 of the 32 gr/gr deletions (eight of 17 independent events), must have their proximal breakpoint in the 70 kb situated between sY587 in the 3′ end of *DAZ2* and the distal end of the P2 palindrome, identifying this region as a relative hotspot of unequal crossing over. Our results show that the gr/gr deletion frequency is independent of the length of homology. Thus, the incidence of gr/gr deletions in the general population provides no support for their negative effect on fertility.

gr/gr deletions and normospermia

In the study that found a significant association between gr/gr deletions and spermatogenic failure (9/246 infertile men and 0/148 normospermic men; $p<0.014$, Fisher's exact test, one-sided), the men in the normospermic population were characterised by a total sperm count of $>40\times10^6$. In our study, we identified two infertile men with $>40\times10^6$ spermatozoa in their ejaculate who carry distinct gr/gr deletions (13-1485 and 13-3446; fig 2). These two men were part of a group of 24 men, from infertile couples, who had sperm counts of $>20\times10^6$ per millilitre. This high incidence of gr/gr deletion (8.3%) in a group characterised by a high sperm count indicates that most gr/gr deletions have little or no effect on sperm production.

In the published association study, gr/gr deletions were initially identified by a negative result for the P2/P1 junction marker, sY1291.[16] This strategy was chosen, prior to the discovery of the b2/b3 inversion polymorphism,[28] on the basis of an earlier theoretical description of gr/gr deletions.[17] When a gr/gr deletion occurs on a b2/b3-inverted chromosome, however, u3 is deleted instead of the P2/P1 junction. Although a u3 marker was used in that study,[16] all 27 gr/gr deletions identified were negative for the P2/P1 marker, and none for the u3 marker, and it is unclear whether u3 was screened for in all men or only in men negative for the P2/P1 junction. This same study did not identify the u3-gr/gr deletion in two Hg-N3 individuals, demonstrating that the authors' screening strategy did not detect all gr/gr deletions. In contrast, we found nine u3-deleted chromosomes out of 32 gr/gr deletions. Thus, the study showing a link between gr/gr deletion and reduced sperm production may not include 25% of the total number of gr/gr deletions present in the groups of men studied. The proposed link between gr/gr deletion and reduced sperm count should, therefore, be disregarded until the u3-gr/gr deletions, which remove more genes than the P1/P2-gr/gr deletions,[28] are taken into account.

AZFc genes and human fertility

It has been proposed that partial *AZFc* deletions may affect male fertility by reducing the dose of certain *AZFc* genes.[16] In consequence, any gr/gr deletion should reduce fertility and should be subject to negative reproductive selection. The results of our study reveal weaknesses in the main lines of evidence advanced in support of this model, suggesting that gene copy loss through gr/gr deletion often has little effect on human fertility.

We suggest, therefore, that an alternative model in which most gr/gr deletions are neutral, and the rate of gr/gr deletion is much lower than that of complete *AZFc* deletion, should also be considered. The effects of gr/gr deletions would be determined by the presence of sequence dependent loss-of-function differences between *AZFc* genes, and would only be seen when the functional copy is deleted and the mutated copy retained. No effect of gr/gr deletion would be seen on Y chromosomes with identical gene copies, or if the mutated copy was deleted. The effect of gr/gr deletion would, therefore, be expected to vary on Y chromosomes of different haplogroups.

The relatively strong association observed in our study between infertility and SFV loss at *CDY1-7750* without a reduction in copy number, represents the first evidence that, in conjunction with partial *AZFc* deletion or gene conversion, sequence dependent loss-of-function differences between *AZFc* gene copies may frequently be the basis for spermatogenic failure. The further study of these variants could lead to the identification of the *AZFc* genes that are necessary for efficient spermatogenesis in man.

ACKNOWLEDGEMENTS

We are grateful to all the individuals who agreed to participate in this study. We also owe thanks to Professors Marc Gamerre and Agnes Noizet of the Centre de Procréation Médicalement Assistée in Marseilles for their help in obtaining patient material and clinical information, to Professor Philippe De Micco of the Etablissement Français du Sang Alpes-Méditerranée for his help in obtaining anonymous control samples, and to Howard Kahn for the CEPH samples. We thank David Page and Laura Brown for sending cosmid 18E8. We thank Alice Sansonetti for technical assistance, and Margaret Mitchell for proof reading the manuscript.

ELECTRONIC-DATABASE INFORMATION

The program developed by Ø Langsrud is available at http://www.matforsk.no/ola/fisher.htm (accessed 19 August 2004).

Authors' affiliations

N Machev, N Saut, G Longepied, A Navarro, N Levy, J Belougne, E Clemente, C Durand, A Ducourneau, M-G Mattei, M J Mitchell, Inserm U.491, Faculté de médecine, 13385 Marseille, France
P Terriou, Institut de Médecine de la Reproduction, 13008 Marseille, France
M Guichaoua, C Metzler-Guillemain, Laboratoire de spermiologie, Hôpital de la Conception, 13385 Marseille, France
P Collignon, A-M Frances, Service de Génétique, Centre Hospitalier de Toulon, 83056 Toulon, France
J Chiaroni, Etablissement Français du Sang Alpes-Mediterranee, 13005 Marseille, France
C Chevillard, Inserm U.399, Faculté de médecine, 13385 Marseille, France
N Pech, E.A. Biodiversité, Université de Provence, 13331 Marseille, France
K McElreavey, Reproduction, Fertility and Populations, Institut Pasteur, 75015 Paris, France

This work was supported by funding from the French national medical research body, Inserm. NS was supported by a studentship from the Ministère de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche (MENESR) and a grant from the Fondation Médicale pour la Recherche.

Conflict of interest: none declared.

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint first authors.

REFERENCES

1 **Tiepolo L**, Zuffardi O. Localization of factors controlling spermatogenesis in the nonfluorescent portion of the human Y chromosome long arm. *Hum Genet* 1976;**34**:119–24.
2 **Ma K**, Sharkey A, Kirsch S, Vogt P, Keil R, Hargreave TB, McBeath S, Chandley AC. 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* 1992;**1**:29–33.
3 **Vogt PH**, Edelmann A, Kirsch S, Henegariu O, Hirschmann P, Kiesewetter F, Kohn FM, Schill WB, Farah S, Ramos C, Hartmann M, Hartschuh W, Meschede D, Behre HM, Castel A, Nieschlag E, Weidner W, Grone 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.
4 **Kuroda-Kawaguchi T**, Skaletsky H, Brown LG, Minx PJ, Cordum HS, Waterston RH, Wilson RK, Silber S, Oates R, Rozen S, Page DC. The AZFc region of the Y chromosome features massive palindromes and uniform recurrent deletions in infertile men. *Nat Genet* 2001;**29**:279–86.
5 **Reijo R**, Lee TY, Salo P, Alagappan R, Brown LG, Rosenberg M, Rozen S, Jaffe T, Straus D, Hovatta O, Delachapelle A, Silber S, Page DC. Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene. *Nat Genet* 1995;**10**:383–93.
6 **Vogt PH**. Human chromosome deletions in Yq11, AZF candidate genes and male infertility: history and update. *Mol Hum Reprod* 1998;**4**:739–44.
7 **Seboun E**, Barbaux S, Bourgeron T, Nishi S, Agulnik A, Egashira M, Nikkawa N, Bishop C, Fellous M, McElreavey K, Kasahara M, Algonik A. Gene sequence, localization, and evolutionary conservation of DAZLA, a candidate male sterility gene. *Genomics* 1997;**41**:227–35.
8 **Lahn BT**, Tang ZL, Zhou J, Barndt RJ, Parvinen M, Allis CD, Page DC. Previously uncharacterized histone acetyltransferases implicated in mammalian spermatogenesis. *Proc Natl Acad Sci U S A* 2002;**99**:8707–12.
9 **Dorus S**, Gilbert SL, Forster ML, Barndt RJ, Lahn BT. The CDY-related gene family: coordinated evolution in copy number, expression profile and protein sequence. *Hum Mol Genet* 2003;**12**:1643–50.
10 **Chang PL**, Sauer MV, Brown S. Y chromosome microdeletion in a father and his four infertile sons. *Hum Reprod* 1999;**14**:2689–94.
11 **Gatta V**, Stuppia L, Calabrese G, Morizio E, Guanciali-Franchi P, Palka G. A new case of Yq microdeletion transmitted from a normal father to two infertile sons. *J Med Genet* 2002;**39**:e27.
12 **Saut N**, Terriou P, Navarro A, Levy N, Mitchell MJ. The human Y chromosome genes *BPY2, CDY1* and *DAZ* are not essential for sustained fertility. *Mol Hum Reprod* 2000;**6**:789–93.
13 **Bienvenu T**, Patrat C, McElreavey K, de Almeida M, Jouannet P. Reduction in the DAZ gene copy number in two infertile men with impaired spermatogenesis. *Ann Genet* 2001;**44**:125–8.
14 **de Vries JW**, Hoffer MJ, Repping S, Hoovers JM, Leschot NJ, van der Veen F. Reduced copy number of DAZ genes in subfertile and infertile men. *Fertil Steril* 2002;**77**:68–75.
15 **Fernandes S**, Huellen K, Goncalves J, Dukal H, Zeisler J, Rajpert De Meyts E, Skakkebaek NE, Habermann B, Krause W, Sousa M, Barros A, Vogt PH. High frequency of DAZ1/DAZ2 gene deletions in patients with severe oligozoospermia. *Mol Hum Reprod* 2002;**8**:286–98.
16 **Repping S**, Skaletsky H, Brown L, van Daalen SK, Korver CM, Pyntikova T, Kuroda-Kawaguchi T, de Vries JW, Oates RD, Silber S, van der Veen F, Page DC, Rozen S. Polymorphism for a 1.6-Mb deletion of the human Y chromosome persists through balance between recurrent mutation and haploid selection. *Nat Genet* 2003;**35**:247–51.
17 **Yen P**. The fragility of fertility. *Nat Genet* 2001;**29**:243–4.
18 **Saxena R**, de Vries JW, Repping S, Alagappan RK, Skaletsky H, Brown LG, Ma P, Chen E, Hoovers JM, Page DC. Four DAZ genes in two clusters found in the AZFc region of the human Y chromosome. *Genomics* 2000;**67**:256–67.
19 **Brownstein MJ**, Carpten JD, Smith JR. Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. *Biotechniques* 1996;**20**:1004–6.
20 **Skaletsky H**, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T, Chinwalla A, Delehaunty A, Delehaunty K, Du H, Fewell G, Fulton L, Fulton R, Graves T, Hou SF, Latrielle P, Leonard S, Mardis E, Maupin R, McPherson J, Miner T, Nash W, Nguyen C, Ozersky P, Pepin K, Rock S, Rohlfing T, Scott K, Schultz B, Strong C, Tin-Wollam A, Yang SP, Waterston RH, Wilson RK, Rozen S, Page DC. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* 2003;**423**:825–37.
21 **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.
22 **Hammer MF**, Horai S. Y chromosomal DNA variation and the peopling of Japan. *Am J Hum Genet* 1995;**56**:951–62.
23 **Hurles 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.
24 **Underhill PA**, Jin L, Lin AA, Mehdi SQ, Jenkins T, Vollrath D, Davis RW, Cavalli-Sforza LL, Oefner PJ. Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res* 1997;**7**:996–1005.
25 **Y Chromosome Consortium**. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res* 2002;**12**:339–48.
26 **Rosser ZH**, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, Amos W, Armenteros M, Arroyo E, Barbujani G, Beckman G, Beckman L, Bertranpetit J, Bosch E, Bradley DG, Brede G, Cooper G, Corte-Real HB, de Knijff P, Decorte R, Dubrova YE, Evgrafov O, Gilissen A, Glisic S, Golge M, Hill EW, Jeziorowska A, Kalaydjieva L, Kayser M, Kivisild T, Kravchenko SA, Krumina A, Kucinskas V, Lavinha J, Livshits LA, Malaspina P, Maria S, McElreavey K, Meitinger TA, Mikelsaar AV, Mitchell RJ, Nafa K, Nicholson J, Norby S, Pandya A, Parik J, Patsalis PC, Pereira L, Peterlin B, Pielberg G, Prata MJ, Previdere C, Roewer L, Rootsi S, Rubinsztein DC, Saillard J, Santos FR, Stefanescu G, Sykes BC, Tolun A, Villems R, Tyler-Smith C, Jobling MA. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am J Hum Genet* 2000;**67**:1526–43.
27 **Zerjal T**, Dashnyam B, Pandya A, Kayser M, Roewer L, Santos FR, Schiefenhovel W, Fretwell N, Jobling MA, Harihara S, Shimizu K, Semjidmaa D, Sajantila A, Salo P, Crawford MH, Ginter EK, Evgrafov OV, Tyler-Smith C. Genetic relationships of Asians and Northern Europeans, revealed by Y-chromosomal DNA analysis. *Am J Hum Genet* 1997;**60**:1174–83.
28 **Fernandes S**, Paracchini S, Meyer LH, Floridia G, Tyler-Smith C, Vogt PH. A large AZFc deletion removes DAZ3/DAZ4 and nearby genes from men in Y haplogroup N. *Am J Hum Genet* 2003;**74**:180–7.
29 **de Vries JW**, Repping S, van Daalen SK, Korver CM, Leschot NJ, van der Veen F. Clinical relevance of partial AZFc deletions. *Fertil Steril* 2002;**78**:1209–14.
30 **Tilford CA**, Kuroda-Kawaguchi T, Skaletsky H, Rozen S, Brown LG, Rosenberg M, McPherson JD, Wylie K, Sekhon M, Kucaba TA, Waterston RH, Page DC. A physical map of the human Y chromosome. *Nature* 2001;**409**:943–5.
31 **Rozen S**, Skaletsky H, Marszalek JD, Minx PJ, Cordum HS, Waterston RH, Wilson RK, Page DC. Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. *Nature* 2003;**423**:873–6.
32 **Ewis AA**, Lee J, Shinka T, Nakahori Y. Microdeletions of a Y-specific marker, Yfm1, and implications for a role in spermatogenesis. *J Hum Genet* 2002;**47**:257–61.
33 **Kuroki Y**, Iwamoto T, Lee J, Yoshiike M, Nozawa S, Nishida T, Ewis AA, Nakamura H, Toda T, Tokunaga K, Kotliarova SE, Kondoh N, Koh E, Namiki M, Shinka T, Nakahori Y. Spermatogenic ability is different among males in different Y chromosome lineage. *J Hum Genet* 1999;**44**:289–92.