Phenotypic variation within European carriers of the Y-chromosomal gr/gr deletion is independent of Y-chromosomal background.

Phenotypic variation within European carriers of the Y-chromosomal gr/gr deletion is independent of Y-chromosomal background

C Krausz, C Giachini, Y Xue, M K O'Bryan, J Gromoll, E Rajpert-de Meyts, R Oliva, I Aknin-Seifer, E Erdei, N Jorgensen, M Simoni, J L Ballescà, R Levy, G Balercia, P Piomboni, E Nieschlag, G Forti, R McLachlan and C Tyler-Smith

J. Med. Genet. 2009;46;21-31; originally published online 9 Sep 2008; doi:10.1136/jmg.2008.059915

Phenotypic variation within European carriers of the Y-chromosomal gr/gr deletion is independent of Y-chromosomal background

C Krausz,[1] C Giachini,[1] Y Xue,[2] M K O'Bryan,[3] J Gromoll,[4] E Rajpert-de Meyts,[5] R Oliva,[6] I Aknin-Seifer,[7] E Erdei,[8] N Jorgensen,[5] M Simoni,[4,9] J L Ballescà,[10] R Levy,[11] G Balercia,[12] P Piomboni,[13] E Nieschlag,[4] G Forti,[1] R McLachlan,[14,15] C Tyler-Smith[2]

► Supplementary tables and figure are published online only at http://jmg.bmj.com/content/vol46/issue1

For numbered affiliations see end of article

Correspondence to:
Dr Csilla Krausz, Andrology Unit, Department of Clinical Physiopathology, University of Florence, Viale Pieraccini, 6 Florence 50139, Italy;
c.krausz@dfc.unifi.it

Received 21 April 2008
Revised 7 August 2008
Accepted 17 August 2008
Published Online First
8 September 2008

ABSTRACT

Background: Previous studies have compared sperm phenotypes between men with partial deletions within the *AZF*c region of the Y chromosome and non-carriers, with variable results. In this study, a separate question was investigated, the basis of the variation in sperm phenotype within gr/gr deletion carriers, which ranges from normozoospermia to azoospermia. Differences in the genes removed by independent gr/gr deletions, the occurrence of subsequent duplications or the presence of linked modifying variants elsewhere on the chromosome have been suggested as possible causal factors. This study set out to test these possibilities in a large sample of gr/gr deletion carriers with known phenotypes spanning the complete range.

Results: In total, 169 men diagnosed with gr/gr deletions from six centres in Europe and one in Australia were studied. The *DAZ* and *CDY1* copies retained, the presence or absence of duplications and the Y-chromosomal haplogroup were characterised. Although the study had good power to detect factors that accounted for ⩾5.5% of the variation in sperm concentration, no such factor was found. A negative effect of gr/gr deletions followed by b2/b4 duplication was found within the normospermic group, which remains to be further explored in a larger study population. Finally, significant geographical differences in the frequency of different subtypes of gr/gr deletions were found, which may have relevance for the interpretation of case control studies dealing with admixed populations.

Conclusions: The phenotypic variation of gr/gr carriers in men of European origin is largely independent of the Y-chromosomal background.

The long arm of the human Y chromosome hosts a number of genes involved in spermatogenesis, and several types of recurrent Yq deletions are firmly associated with spermatogenic failure[1–3] (fig 1). Apart from the classic *AZF* deletions, new types of Yq rearrangements have recently attracted the attention of geneticists and andrologists.[4] The complex duplicated structure of the *AZF*c region predisposes to a series of rearrangements mediated by non-allelic homologous recombination, including the formation of partial deletions, deletion/duplications, and partial duplications.[5–7] The most clinically relevant mutation is termed the "gr/gr" deletion[8] and removes half of the *AZF*c gene content. This deletion is a significant risk factor for spermatogenic failure in some populations but apparently not in others.[9] These contradictory results are likely to derive from methodological differences between studies and recruitment biases. In only a minority of studies have the controls been normozoospermic men; controls and patients were not matched for ethnic background in highly admixed populations from Paris,[10] southern France[11] and Brazil[12] used in some studies, and the proportion of subjects with azoospermia or oligozoospermia differed between studies.

Although the debate about the significance of the gr/gr deletion as a risk factor for impaired sperm production is ongoing and unresolved, and was not further investigated here, there is a firm consensus about the heterogeneous phenotype associated with gr/gr deletions.[10–16] Even in studies in which gr/gr deletion is clearly shown as a risk factor for impaired spermatogenesis, the sperm phenotype ranges from azoospermia to normozoospermia.[9] This phenomenon is in sharp contrast to the classic *AZF* deletion phenotypes, which are invariably associated with impaired sperm production. Given that the number of genes removed by the gr/gr deletion is half that of the classic *AZF*c deletion, it is not unexpected that the effect of the deletion is milder and it is therefore a cofactor for spermatogenic impairment with variable penetrance. Thus, it is also possible that other Y-linked or non-Y genetic factors influence the pathogenicity of the deletion. The significance of the presence of polymorphisms or mutations in the autosomal homologue of the *DAZ* gene, *DAZL*[17] has been examined in one study,[14] but it was concluded that the polymorphic Thr12Ala change (T12A) was unlikely to contribute and no new mutations in the entire coding region of the *DAZL* gene were identified. No other autosomal factors modulating the deletion phenotype have yet been reported.

For Y-related factors, interpretation is complicated by the lack of recombination in the male-specific part of the chromosome, which results in complete linkage between variants throughout most of the chromosome. Their likely importance is illustrated by certain Y haplogroups (Y chromosome backgrounds defined by Y chromosome single-nucleotide polymorphisms (SNPs)), which carry fixed gr/gr deletions but are nevertheless present at high frequency in some populations, indicating that compensatory mechanisms may exist on these backgrounds.[18–20] The specific genes removed can vary between gr/gr deletions and this

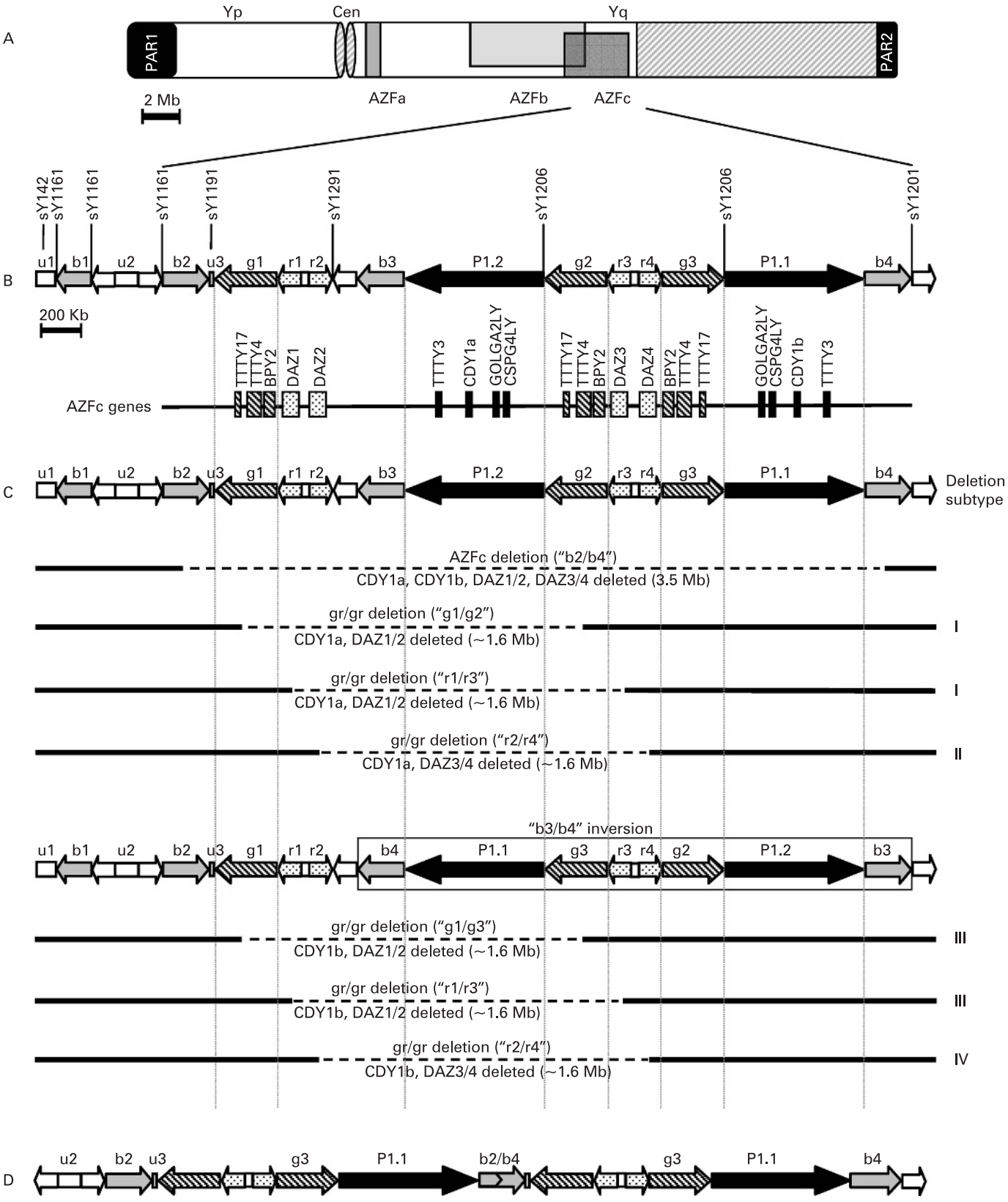
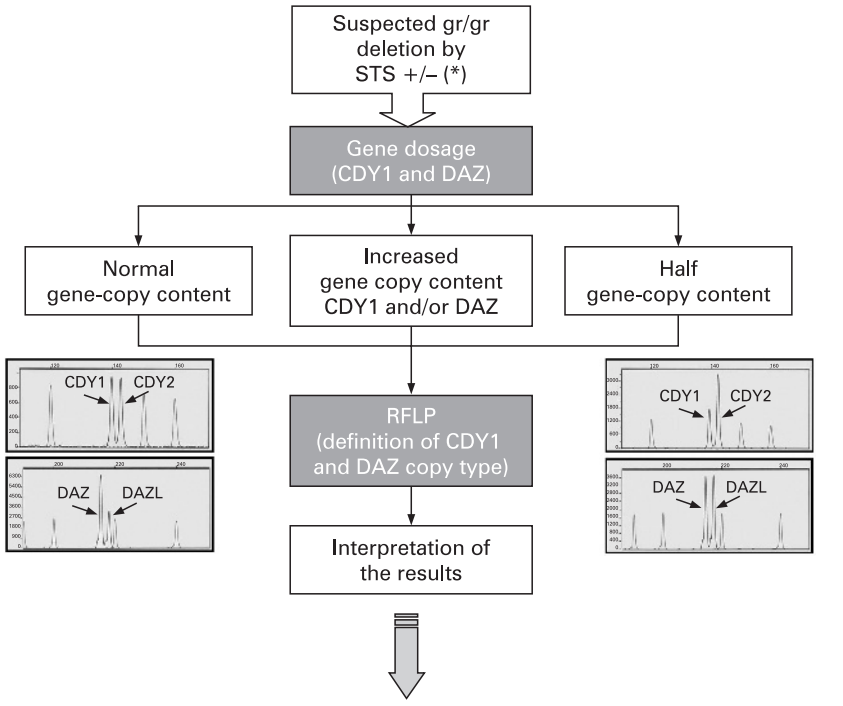


Figure 1 Schematic representation of the Y chromosome showing the structures, genes, sequence-tagged sites and rearrangements relevant to this study. (A) Deletions of the *AZFa*, *AZFb* and *AZFc* regions result in spermatogenic failure. *AZF*b deletions (two subtypes) overlap with the *AZF*c region. (B) The *AZF*c region presented in more detail showing the location of multicopy genes and transcription units in the reference sequence published by Skaletsky *et al*.[3] This region contains a number of repeated sequences with the same orientation (matching arrows) which through intrachromosomal recombination may lead to deletions. (C) The clinically relevant "b2/b4" (top) and "gr/gr" deletions (remainder) are shown. Half of the *AZF*c gene content is removed by gr/gr deletions, but these can vary in breakpoints (deletion subtypes I and or according to the presence of a b3/b4 inversion (deletion subtypes III and IV). (D) gr/gr deletion of subtype I or II followed by b2/b4 duplication. cen, centromere; PAR, pseudoautosomal region; Yp, short arm of the Y chromosome; Yq, long arm of the Y chromosome.

A



B

			Number of copies		Type of copies	
	sY 1291	Other STS*	DAZ	CDY1	DAZ	CDY1
Reference sequence	+	+	4	2	1/2 and 3/4	a and b
gr/gr del	–	+	2	1	1/2 or 3/4§	a or b
gr/gr del–b2/b4 dupl	–	+	4	2	1/2 or 3/4	a or b
gr/gr del–b2/b4 multiple dupl	–	+	⩾6	⩾3	1/2 or 3/4	a or b
gr/gr del–CDY1 ampl	–	+	2	⩾2	1/2 or 3/4	a or b
gr/gr del–DAZ ampl	–	+	⩾4	1	1/2 or 3/4	a or b
False deletion	–	+	4	2	1/2 and 3/4	a and b

Figure 2 Schematic representation of the multistep gr/gr deletion screening procedure and categories of gr/gr subtypes defined. (A) Gene dosage and restriction fragment length polymorphism analysis was used to distinguish between "simple" gr/gr deletions, false deletions and deletion/duplications events and to define the missing gene copies, ie gr/gr deletion subtypes (examples of electropherograms shown; see Methods for further details). *STS, sequence tagged site-based PCR method, according to Repping *et al*.[8] (B) Molecular characterisation of the various gr/gr rearrangement types. Gene dosage of *CDY1* and *DAZ* and the possible combinations of missing *DAZ* (*DAZ1/DAZ2/DAZ3/DAZ4*) and *CDY1* (*CDY1a/CDY1b*) copy types are shown for each rearrangement. *In the *AZFc* region: sY142, sY1258, sY1161, sY1197, sY1191, sY1206, sY1201. §Six cases showed a combination of *DAZ* copies from each *DAZ* duplex, possibly *DAZ1/3* or *DAZ3/4*. +, P positive; −, negative; del, deletion; dupl, duplication; ampl, amplification.

has been suggested as the most direct modulating factor. The *DAZ* and *CDY1* copies, which lie within the deleted region, have been examined as possible predictors for pathogenicity. The loss of *DAZ1/DAZ2* and *CDY1a* has been reported as more deleterious than loss of *DAZ3/DAZ4* and *CDY1b*.[14 18 21] Concerning the *DAZ* gene copies, different members of the *DAZ* gene family have different number of RNA recognition motifs (RRMs) and *DAZ* repeats, which may confer different functional activity on the four *DAZ* copies. However, it is also known that both elements show interindividual variability, which in theory might be linked to a particular Y lineage. Similarly, the loss of *CDY1a* or *CDY1b* copies might reflect a different Y structure. In fact, in the reference sequence, *CDY1b* lies outside the region that is predicted to be deleted by gr/gr recombination. This implies that the *CDY1b* loss may occur only in Y chromosomes that carry an inversion polymorphism in the *AZFc* region,[11] or must arise independently. Such an inversion polymorphism may affect the transcription of the remaining *AZFc* genes or may be linked to other Y structural variations. An alternative explanation for the gr/gr deletions found in normospermic men could be that a compensatory duplication, such as a gr/gr deletion followed by b2/b4 duplication, has restored normal gene number and function. Furthermore, Y variants outside the *AZFc* region may influence spermatogenesis and these will be associated with particular haplogroups, so may be recognised indirectly using a phylogenetic analysis. A well-resolved Y chromosome tree of 311 haplogroups is now available,[22] as well as extensive information about the haplogroups present in different regions of the world including Europe.[23–25] Consequently, a small number of informative Y-SNPs can be tested to identify the most prevalent haplogroups in Europe.

Most studies of gr/gr deletions have used sequence tag site (STS) +/− PCR analysis to identify the deletion and thus have been unable to provide information about the type of missing gene copies or about deletion/duplication events. For a fuller understanding, a method such as combined analysis based on a first step of STS +/− screening, followed by a confirmatory dosage analysis and gene copy characterisation of the samples with suspected deletions, is needed and has been described by Giachini *et al*.[14]

The aim of the present study was to determine whether any of the suggested Y-related factors can account for the different semen phenotypes found in gr/gr deletion carriers, rather than to re-investigate possible differences between carriers and non-carriers. For this purpose, we collected 169 DNA samples from seven countries that were previously defined by STS +/− analysis as gr/gr deleted men. All subjects were further characterised by gene dosage analysis, gene copy type and Y haplogroup, and these factors were related to spermatogenic

phenotype to search for a genetic profile specific for spermatogenic failure.

METHODS

Patients and controls were recruited according to local ethics committee policy and all subjects signed an informed consent form.

Subjects analysed

In total, 169 DNA samples were analysed for classic *AZF* deletions and for partial *AZFc* region deletions in the participating laboratories from the following countries: Australia (n = 50), Denmark (n = 25), France (n = 13), Germany (n = 35), Hungary (n = 4), Italy (n = 24) and Spain (n = 18). All 169 samples showed absence of amplification for sY1291 but presence of the other specific *AZFc* markers (sY142, sY1258, sY1161, sY1197, sY1191, sY1206 andsY1201; see GenBank for PCR primers) indicating the presence of a gr/gr deletion. Some of the patients/controls with gr/gr deletions (based on STS analysis) have been included in previous publications.[13–16]

The composition of the study population (n = 169) was as follows: 152 infertile and 17 fertile or normospermic men sent as "controls" (supplementary table 1). All subjects were determined to be of European ancestry, either on the basis of their surname or on direct questioning during recruitment about the origin of their parents. Of the 152 infertile patients, semen parameters were available for 134. In an additional 7 patients, only the semen phenotype (azoospermia, oligozoospermia) was given by the laboratories, and for the remaining 11 patients the only information given was that they were infertile. The group of 141 patients with known semen phenotype comprised 27 azoospermic, 18 cryptozoospermic (<1 million spermatazoa/ml), 82 oligozoospermic and 14 asthenoteratozoospermic subjects. The mean (SD) values of the three principal sperm parameters in the infertile group for which semen parameters were provided were as follows: sperm concentration: 6.9 (14.7)×10^6/ml (n = 134); total sperm count: 24.8 (60.6)×10^6 (n = 114); percentage progressive motility: 16.2 (16.8)% (n = 122); percentage normal morphology: 8.9 (8.1)% (n = 82). Motility and/or morphology were not determined in azoospermic and severely oligozoospermic men (<1×10^6 spermatozoa/ml). The mean sperm concentrations are reported for each country in supplementary table 2 (online).

Semen parameters were available for all 17 controls (except two (D255; B11) for whom morphology was not provided and one (D255) for whom motility was not provided). The control population consisted of 10 subjects presenting all the three major sperm parameters (concentration, motility and morphology) above the normal range;[26] and 7 subjects presenting normal sperm concentration but with motility and morphology below the normal range. The mean (SD) values of the sperm parameters in the control group were: sperm concentration: 76.8 (44.9)×10^6/ml (n = 17); total sperm count: 328.4 (396.7)×10^6 (n = 13); percentage progressive motility: 49.2 (13.3) (n = 16); percentage normal morphology: 20.8 (11.8) (n = 15).

The 169 samples were analysed further at the Andrology Unit (Department of Clinical Physiopathology, University of Florence) where a fine-scale molecular characterisation of the deletions using *AZFc* gene dosage and gene copy analysis was carried out as described by Giachini *et al*.[27] Y chromosome haplogroup analysis was performed using the multiplexed primers previously described[4] adapted for SNaPshot single base extension (Applied Biosystems, Foster City, California, USA) at the Wellcome Trust Sanger Institute (Cambridge, UK) for153 subjects. The markers *RPS4Y*$_{711}$, M145, M96, M89, M9 and M45 were typed on all samples and M123, M78, V6, M35 and M81 (haplogroup E, derived for M96), M201, M170, M52 and 12f2 (haplogroup F, derived for M89), M106, M61, M147, M214, M27, M76 and M70 (haplogroup K, derived for M9) or M17, M343, M369 and M18 (haplogroup P, derived for M45) on appropriate subsamples, but data were combined into the major haplogroups (E, F*(xK), K*(xP) and P) to provide numbers suitable for most statistical analyses. Full data are available for 150 samples (supplementary table 1 online).

Gene dosage

To measure the copy number of these genes, quantitative analysis of *CDY1* and *DAZ* was performed using a modified PCR-based method.[11] We simultaneously amplified the *AZFc* locus to be quantified (*CDY1* or *DAZ*) and a homologous locus outside the *AZFc* interval (*CDY2* and *DAZL*, respectively) as an internal standard with a known number of copies, using a single primer pair in a PCR reaction with a maximum of 24 cycles (end point of the exponential phase). 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 (FAM). The reaction was then mixed with formamide, denatured at 95°C for 5 min and the differently sized loci separated on an automatic sequencer (ABI Prism 310 Genetic Analyzer; PE Applied Biosystems, Foster City, California, USA). Quantification was performed by comparing the peak area of the *AZFc* locus with its homolog.

Table 1 Frequency of each gr/gr rearrangement in the whole study population

Rearrangement type	Total, n	Total, n with semen phenotypes	Normal sperm count, n			Abnormal sperm count, n		
			Normo	AT	Total 1	Az	Oz	Total 2
gr/gr del	128	122	9	16	25	18	79	97
gr/gr del–b2/b4 dupl	23	22	—	4	4	7	11	18
gr/gr del–b2/b4 multiple dupl	3	3	—	1	1	—	2	2
gr/gr del–*CDY1* ampl	4	1	—	—	—	—	1	1
gr/gr del–*DAZ* ampl	2	1	—	—	—	1	—	1
Total (n)	160*	149†	9	21	30	26	93	119

Ampl, amplification; AT, asthenozoospernic and/or teratozoospermic; Az, azoospermic; del, deletion; dupl, duplication; Normo, normozoospermic; Oz, oligozoospermic; Total 1, normospermic plus pure AT; Total 2, AZ plus OZ.
*False deletions were excluded (n = 9).
†In 11 cases no information about the semen phenotype was given by the laboratories; subjects with known semen phenotypes were divided into two major categories (normal or abnormal sperm) and further divided into subcategories.

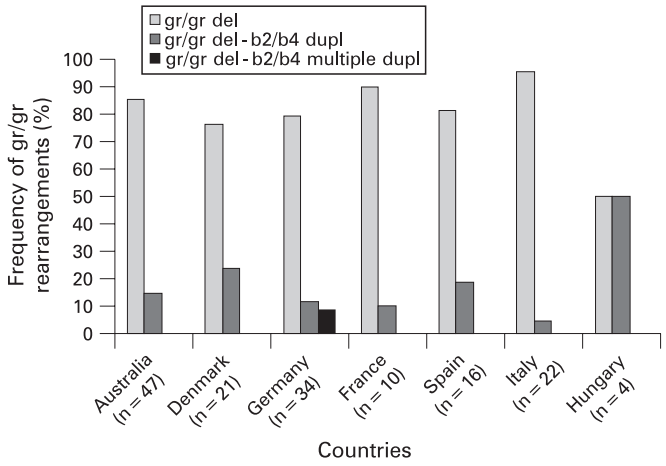


Figure 3 Comparison of the frequencies of the different gr/gr rearrangement types in each participating country. False deletions and gr/gr deletion–*DAZ* or *CDY1* amplification were excluded from this histogram.

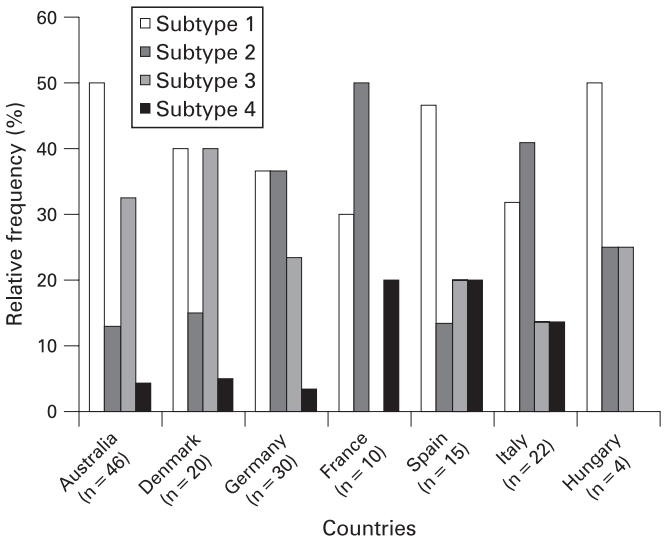


Figure 4 Comparison of the frequencies of the different gr/gr subtypes in each participating country. False deletions and gr/gr deletion–*DAZ* or *CDY1* amplification were excluded from this histogram. Deletion subtypes removed: (1) *DAZ*1/*DAZ*2+*CDY1a*; (2) *DAZ*3/*DAZ*4+*CDY1a*; (3) *DAZ*1/*DAZ*2+*CDY1b*; (4) *DAZ*3/*DAZ*4+*CDY1b*.

CDY1 vs *CDY2* (primers: oMY953a/o1023)

There are two identical copies each of *CDY1* and *CDY2*, which share 98% nucleotide identity. 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*.

DAZ vs *DAZL* (primers: o1130/01313)

We coamplified a fragment of intron 10 from *DAZ* (214 bp) and *DAZL* (217 bp). This intron is present in one copy per *DAZ* gene (in the reference 46,XY man there were four copies of *DAZ* and two copies of *DAZL*). Some samples showed a 40 bp insertion polymorphism in the *DAZL* intron 10, resulting in an extra band at 260 bp, which could be heterozygous or, more rarely, homozygous.

Gene copy type

Qualitative analysis for *CDY1* and *DAZ*, in order to determine which copies of these genes have been removed by the gr/gr deletion, was performed according to the method of Machev *et al*.[11] For *DAZ*, we chose the sequence family variant (SFV) at STS sY587 in intron 10, which discriminates *DAZ1/2* from *DAZ3/4*. For *CDY1*, we used a C/A SFV situated 7750 bp 5′ of the *CDY1* translation start codon (*CDY*7750), which distinguishes *CDY1a* from *CDY1b*. SFVs were scored by PCR followed by restriction enzyme digestion (5 U for at least 4 hours): *DAZ* sY587, DraI (*DAZ1/2* cut); *CDY1*-7750, PvuII (*CDY1b* cut). Digestion products were then analysed by electrophoresis at 100 V in 4% agarose gels containing ethidium bromide and visualised under ultraviolet light. Primer pairs: sY587, o912/o913; *CDY1*-7750, o1025/o1026.[11]

Statistical analysis

Median values between groups were compared using a non-parametric Mann–Whitney U test or Student's *t* test after normalisation of the distribution by log transformation, as shown by a one-sample Kolmogorov–Smirnov test; $p = 0.023$. We also performed a regression analysis on the log-transformed sperm counts to determine whether variation in sperm count (the dependent variable) could be explained by any of the other variables. Analyses were carried out using SPSS 14.0, except for power calculations, which were performed using the website http://www.danielsoper.com/statcalc/calc09.aspx.

RESULTS

Molecular analysis of the AZFc region

Definition of different gr/gr rearrangements

The molecular characterisation of samples sharing the absence of the sY1291 was necessary in order to distinguish between different gr/gr rearrangements that may not be found by the STS analysis used by the participating laboratories and to confirm the effective loss of genetic material—that is, to exclude "false" deletions (fig 2).

We defined the gene dosage of *CDY1* and *DAZ* and the type of missing *DAZ* (*DAZ*1/*DAZ*2/*DAZ*3/*DAZ*4) and *CDY1* (*CDY1a*/*CDY1b*) copies. The gene dosage and gene copy types are reported (fig 2) for each gr/gr rearrangement that could be identified by this combined two-step method.

In nine subjects (one normospermic for all three parameters and eight infertile with reduced sperm count) we found a pattern compatible with "false" deletions, showing normal *DAZ* and *CDY1* gene dosage and the presence of both *CDY1* copies (*CDY1a*/*CDY1b*) and all four *DAZ* copies (*DAZ*1/*DAZ*2/*DAZ*3/*DAZ*4). False deletions using STS analysis can be due either to polymorphisms/rearrangements at the annealing site of the primers or to PCR artefacts. The nine samples were reanalysed using simplex PCR with sY1291 at different annealing temperatures, which confirmed that these were undeleted in all cases, indicating that the failure of amplification was due to suboptimum PCR conditions for those specific samples (DNA quality may be responsible for such an amplification failure) and the false deletion rate was 5.3% (9/169). The nine samples originated from four different laboratories and were excluded from further analyses.

In the remaining 160 subjects we found a heterogeneous situation showing both "true" simple gr/gr deletions (n = 128), gr/gr deletion followed by b2/b4 duplication (n = 23), gr/gr deletion followed by multiple b2/b4 duplications (n = 3), confirmed gr/gr deletion but a higher copy number of *CDY* genes than the reference sequence (gr/gr deletion–*CDY1* amplification) (n = 4) and confirmed gr/gr deletion but a higher

Table 2 The frequency distribution of different subcategories defined on the basis of the missing *DAZ* and *CDY1* copies

	Normal sperm count, n (%)			Abnormal sperm count, n (%)			
	gr/gr del	gr/gr del–b2/b4 dupl and MD	Total 1	gr/gr del	gr/gr del–b2/b4 dupl and MD	Total 2	Total 1+2, n (%)
Type of missing *DAZ* copies							
DAZ1/2	14 (58.3)	4 (80)	18 (62.1)	62 (66.7)	15 (79)	77 (68.8)	95 (67.4)
DAZ3/4	10 (41.7)	1 (20)	11 (37.9)	31 (33.3)	4 (21)	35 (31.2)	46 (32.6)
Total	24	5	29	93	19	112	141*
Type of missing *CDY1* copy							
CDY1a	15 (60)	3 (60)	18 (60)	66 (68.8)	15 (75)	81 (69.8)	99 (67.8)
CDY1b	10 (40)	2 (40)	12 (40)	30 (31.2)	5 (25)	35 (30.2)	47 (32.2)
Total	25	5	30	96	20	116	146*
Deletion subtypes							
1	8 (33.3)	2 (40)	10 (34.5)	37 (40.2)	11 (57.9)	48 (43.2)	58 (41.4)
2	6 (25)	1 (20)	7 (24.1)	25 (27.2)	4 (21.05)	29 (26.1)	36 (25.7)
3	6 (25)	2 (40)	8 (27.6)	23 (25)	4 (21.05)	27 (24.3)	35 (25)
4	4 (16.7)	–	4 (13.8)	7 (7.6)	–	7 (6.3)	11 (7.8)
Total	24	5	29	92	19	111	140*

Del, deletion; dupl, duplication; MD, multiple duplication; Total 1, n. gr/gr del plus n. gr/gr del–b2/b4 dupl and MD for subjects with normal sperm count; Total 2, n. gr/gr del plus n. gr/gr del–b2/b4 dupl and MD for subjects with abnormal sperm count.
*Differences in total numbers (Total 1, Total 2 and Total 1+ 2) are due to six cases with unclassified *DAZ* deletions (one normal and five abnormal) and one case with an undefined *CDY1* copy deletion (one from the abnormal sperm count group).
The two cases with gr/gr del-*CDY1* ampl and *DAZ* ampl were excluded from the table.
Deletion subtypes are: (1) *DAZ*1/*DAZ*2+*CDY1a*; (2) *DAZ*3/*DAZ*4+*CDY1a*; (3) *DAZ*1/*DAZ*2+*CDY1b*; (4) *DAZ*3/*DAZ*4+*CDY1b*.
The percentages are calculated from the totals in a vertical manner.

copy number of *DAZ* genes than the reference sequence (gr/gr deletion–*DAZ* amplification) (n = 2) (fig 2, table 1).

Geographical distribution of gr/gr rearrangements

The most frequent gr/gr rearrangement type in all countries was the simple gr/gr deletion, which ranged from 50% (Hungary) to 95% (Italy). The number of Hungarian gr/gr deletion carriers was low (n = 4) and 50% presented gr/gr deletion–b2/b4 duplication. The frequency of gr/gr deletion–b2/b4 duplication varied between countries, being highest in Denmark (24%) and lowest in Italy (4.5%) with a comparable number of total subjects analysed from the two countries (21 and 22, respectively). Interestingly, we found gr/gr deletion–b2/b4 multiple duplication only in the German study population (8.8%) (fig 3). Deletions with amplified *DAZ* or *CDY* copies are not included in fig 3 because of their low number (n = 6). Three of four subjects with gr/gr deletion –*CDY1* amplification were French and one was Australian. Both subjects with gr/gr deletion –*DAZ* amplification were of Danish origin. Given the relatively low number of subjects in each group, differences in the distribution of gr/gr rearrangement types were not significant.

Definition of gr/gr deletion patterns (subtypes) based on the missing copies of *DAZ* and *CDY1* genes

Our restriction fragment length polymorphism (RFLP) method is based on the use of an SFV at sY587 which differentiates *DAZ*1 and *DAZ*2 from *DAZ*3 and *DAZ*4. In the majority of cases (n = 154), we found deletion of either the *DAZ*1/*DAZ*2 (n = 106, 69%) or the *DAZ*3/*DAZ*4 (n = 48, 31%) gene pairs. In six cases belonging either to the group of "true" gr/gr deletions (n = 5) or to gr/gr deletion–b2/b4 duplications (n = 1) the RFLP analysis identified a peculiar pattern characterised by the combination of one *DAZ* gene from each gene pair. It can be hypothesised that the recombination site was between the two red amplicons (fig 3), removing either *DAZ*1 with *DAZ*3 or *DAZ*2 with *DAZ*4. According to previous publications[11 14 27] gr/gr deletions may present either the removal of *CDY1a* or *CDY1b* (fig 1). The position of the *CDY1b* copy on the *AZFc* reference sequence implies that in those cases in which *CDY1b* is deleted, an inversion polymorphism in the *DAZ3/4* palindrome P1 may have occurred. The alternative hypothesis of gene conversion of *CDY1a* by *CDY1b* before deletion was considered highly unlikely by Machev *et al*[11] because *CDY1b* deletions were found at high frequency in the gr/gr deleted subjects (48%) whereas *CDY1b* gene conversion events were found only at low frequency (<1%). In our study population (n = 159), we found deletion of *CDY1a* in 107 (67%) subjects and of *CDY1b* in 52 (33%). In one case, there was insufficient DNA for RFLP analysis of *CDY1*.

The combination of the different copies of the *DAZ* and *CDY* genes gave four major gr/gr deletion patterns characterised by the loss of: *DAZ*1/*DAZ*2+*CDY1a* (subtype 1), *DAZ*3/*DAZ*4+*CDY1a* (subtype 2), *DAZ*1/*DAZ*2+*CDY1b* (subtype 3) and *DAZ*3/*DAZ*4+*CDY1b* (subtype 4). The frequency of the four subtypes in the whole study population (n = 153, excluding the six cases showing the combination of two *DAZ* gene for each gene pair and the case with inconclusive *CDY1*-RFLP results) was 65/153 (42.5%) for subtype 1; 37/153 (24.2%) for subtype 2; 39/153 (25.5%) for subtype 3; 12/153 (7.8%) for subtype 4.

Geographical distribution of gr/gr subtypes

There were significant differences between countries for frequency of gr/gr subtypes 2 and 3, whereas subtypes 1 and 4 were more similarly distributed (fig 4). The highest frequency of subtype 1 was found in Australia (50%), whereas the lowest was in France (30%). Subtype 4 was more frequent in France, Spain and Italy (13–20%) and less frequent in Australia, Germany and Denmark (3.3–5%). This subtype was absent in the four Hungarian gr/gr deletion carriers. Significant differences were found for the frequencies of subtype 2:Australia with the lowest value (13%) versus Germany (36.7%), France (50%) and Italy (41%) gave p values of 0.016, 0.008 and 0.01, respectively. The difference was also significant between France (the highest with 50%) and Spain (the second lowest, 13.3%)

(p = 0.05). Subtype 3 was completely absent from the French samples and differences were significant when France (0%) was compared with Australia (32.6%) and Denmark (40%), (p = 0.036 and p = 0.022, respectively).

Genotype–phenotype correlations

To evaluate the effect of different gr/gr rearrangements and gr/gr deletion subtypes on sperm production, we divided the study population into two groups on the basis of the sperm count: group 1 with abnormal sperm count (azoospermia or oligozoospermia) and group 2 with normal sperm count (including those subjects who presented either all three sperm parameters above the normal range or only normal sperm concentration independently of their fertility status). This subdivision meant that in the group of subjects with "normal sperm count" we also included 13 originating from the patient group with normal sperm concentration but reduced sperm motility and/or morphology (asthenozoospermia and/or teratozoospermia). The mean (SD) of the sperm parameters in this combined control group was: sperm concentration 59 (41.4) $\times 10^6$ spermatazoa/ml; percentage progressive motility 37.2 (20)%; percentage normal morphology 14.6 (10.7)%.

Rearrangements of gr/gr and semen phenotype categories

Given that different gr/gr rearrangements are associated with different gene dosage, we tested the hypothesis that differences in gene dosage might explain the heterogeneous semen phenotype found in gr/gr deletion carriers. However, the comparison of the two groups defined on the basis of sperm count showed a similar distribution of the different gr/gr rearrangements. The most prevalent rearrangement type in both groups was the gr/gr deletion (81.5% in group 1, 83.3% in group 2). Restoration of the original or higher gene dosage by single or multiple duplications did not seem to influence the phenotype as 17% of oligozoospermic/azoospermic subjects carried deletion/duplication events and this rearrangement was present at a similar frequency in the normozoospermic group (17%). Data are reported in table 1. The nine "fully normospermic" controls who presented all three sperm parameters above the normal range were all carriers of a simple gr/gr deletion without duplications.

Rearrangements of gr/gr and mean sperm concentration

In subjects for whom sperm parameters were available (n = 142), we compared the mean sperm concentration within each rearrangement type. There was no significant difference between subjects with half the *AZFc* gene dosage (gr/gr deletion) compared with normal (deletion/duplication) and higher copy number (deletion/multiple duplication): mean 16 (SD 32.2) $\times 10^6$/ml (n = 117) versus 9.3 (SD 13.7) $\times 10^6$/ml (n = 24). The corresponding comparison was also performed separately in subjects with abnormal sperm count (n = 111) and in subjects with normal sperm concentrations (n = 30). The comparison of gr/gr deleted and deleted/duplicated subjects in the first group showed no difference, 2.9 (SD 3.8) $\times 10^6$/ml (n = 92) versus 3.1 (SD 3.9) $\times 10^6$/ml (n = 19). Even after the removal of azoospermic men (given that gr/gr deletions are highly unlikely to be a definitive cause of azoospermia), the mean values remained similar.

In contrast, the group of subjects with normal sperm concentration showed nominally significant differences in gr/gr deletion versus deletion/duplication and multiple duplication: 64.2 (SD 43.4) $\times 10^6$/ml (n = 25) versus 32.8 (SD 11.8) $\times 10^6$/ml (n = 5), respectively (*t* test of $\log_n$ spermatozoa/ml p = 0.042). As all nine fully normospermic controls were carriers of gr/gr deletions, such a comparison could not be performed in this group. If we considered separately only the 16 subjects with normal sperm concentration and were sent as controls, we found a similar effect of duplications in this subgroup: 81.6 (SD 12.1) $\times 10^6$/ml (pure gr/gr deletions in 14 subjects) and 29.7 (SD 6.7) $\times 10^6$/ml (deletion—b2/b4 duplication and deletion–multiple duplications in two subjects). This interesting observation could have several explanations, including pure chance, as multiple tests were performed and replication in an independent sample is needed as the next step.

These data suggest that gene dosage differences are not a critical determinant of sperm output in the group of oligospermic/azoospermic men (although the mean sperm concentration was lower when duplications occurred), whereas it may reduce the spermatogenic potential in subjects with >20 million spermatozoa/ml.

Deletion subtypes and semen phenotype categories

The most frequently deleted *DAZ* copy was *DAZ1/2*, whereas the predominant missing *CDY1* copy was *CDY1a*, in both subjects with normal and abnormal sperm concentration (table 2). We and others have previously hypothesised that certain deletion subtypes based on the definition of the missing *DAZ* (1/2 or 3/4) and *CDY1* (a or b) copies (four possible combinations) are enriched in patients and thus could be more pathogenic than others. To test this hypothesis, we compared the frequency of the four different gr/gr subtypes between subjects with abnormal and normal sperm concentration. As stated above, in six cases the RFLP analysis of the *DAZ* gene showed a pattern which was different from the four classic combinations and in one case the RFLP analysis of the *CDY1* gene gave inconclusive result. These seven cases, together with those presenting gr/gr deletion and amplified *DAZ* or *CDY1* copy number, were excluded from the statistical analysis, so we selected only those subjects who had either gr/gr deletion or gr/gr deletion–b2/b4 duplication and multiple duplication (table 2).

In accordance with our previous data,[14] the most frequent subtype found in oligoazoospermic men was subtype 1, whereas the least frequent was subtype 4. However, men with normal sperm concentration also showed a similar frequency distribution for the four subtypes. It is nevertheless worth noting that, in the present study, subtype 4 (the only gr/gr subtype found in normospermic controls in the Italian sample[14 28]) was more frequent in the normospermic control group (16.7%) than in the oligospermic/azoospermic group (7.6%) (3.8% versus 6.3% if we consider both gr/gr deletions and gr/gr deletion–b2/b4 duplication (single or multiple)). The phenotype associated with subtype 4 therefore requires further investigation. The semen phenotype of the six patients with the unusual DAZ deletion pattern was heterogeneous, comprising azoospermia (n = 1), cryptozoospermia (n = 2), severe oligoasthenoteratozoospermia (n = 2) and asthenoteratozoospermia (n = 1). These data indicate that this rare pattern is mainly associated with abnormal spermatogenesis.

Deletion subtypes and mean sperm concentration

We compared the mean sperm concentration in subjects with *CDY1a* deletion versus *CDY1b* deletion in the whole study population and then separately in the group of subjects with normal and abnormal sperm count. Data are reported in table 3, in which we again excluded the six cases with unclassified *DAZ*

copies and gr/gr deletions with *DAZ* and *CDY1* copy amplification. Therefore, for this analysis we selected only those subjects who had either gr/gr deletion or gr/gr deletion–b2/b4 duplication or multiple duplications and with known sperm concentrations. Although in both groups with normal and abnormal sperm count, the mean values of sperm concentration were lower in cases of loss of *DAZ1*/*DAZ2* or *CDY1a* copy loss, the differences did not reach significance.

Y haplogroups and their relationship to AZFc rearrangements and to the phenotypic expression of gr/gr deletions and deletions/duplications

Y haplogroup analysis showed that haplogroups E, I, R1a and R1b, here combined with related chromosomes into the broader categories E, F*(xK), K*(xP) and P, made up most of the sample, as expected for a predominantly western European mixed sample (supplementary table 1). Haplogroup P was the most frequent haplogroup in men with both normal and abnormal sperm counts and the frequencies of the other haplogroups followed the same order P>F*(xK)>E>K*(xP) in the two groups (table 4). Similarly, within each of the classes (normal sperm concentration, abnormal sperm concentration or total), there were no significant differences in mean sperm concentration between the haplogroups (table 5). Duplications (including multiple duplications) following deletions were more rare in haplogroup P chromosomes than in other haplogroups and more frequent in haplogroup E chromosomes, but these differences were not significant ($p = 0.08$, $p = 0.08$ respectively; Fisher exact test). These duplications were more frequent in Denmark and Germany than in the other countries examined, but this enrichment could not be accounted for by the haplogroup distributions as haplogroup E is more common in south-eastern European countries.[24]

Because removal of both *CDY1b* and sY1291 as a single event implies the presence of an inversion polymorphism with respect to the reference sequence, it was of interest to analyse the haplogroup distribution in subjects with *CDY1a* versus *CDY1b* deletion. We found that haplogroup P was significantly more frequent when there was copy loss of *CDY1a* (56.7%) rather than *CDY1b* (14.6%) $p<0.001$. On the other hand, 73% of subjects with *CDY1b* copy loss have haplogroup F*(xP) or E (56.3% and 16.7%), whereas in *CDY1a* loss carriers the cumulative frequency of the same haplogroups was lower at 33% (haplogroup F*(xP) = 25% and haplogroup E = 7.7%). Within haplogroup F*(xP), there was a highly significant difference within haplogroup I which made up 81% of the *CDY1b* loss chromosomes but only 19% of the *CDY1a* loss chromosomes ($p = 0.00001$; data in supplementary table 1).

Finally, we used linear regression analysis on the same samples to investigate whether the variation in sperm concentration of oligospermic or normospermic men, considered as a continuous variable, could be explained by any of the factors measured: gr/gr versus deletion/duplication, *CDY1* deletion, *DAZ* deletion or Y haplogroup. None of these factors explained a significant proportion of the variation in sperm count (all $p>0.05$).

DISCUSSION

This is the largest study to date to investigate the contribution of Y-chromosome factors to the extensive and puzzling phenotypic variation exhibited by gr/gr deletion carriers, which ranges from normal spermatogenesis to azoospermia. The factors examined included both the known *AZFc* structural variants associated with this deletion (removal of various *DAZ* and *CDY1* gene copies, deletion followed by duplication) and the more general Y chromosome background, measured as the Y-SNP-defined haplogroup, which could reveal the influence of factors located anywhere on the male-specific part of the chromosome. We found that none of these factors accounted for a significant proportion of the spermatogenic variation associated with gr/gr deletions. We now consider whether any aspects of our study design and execution could have led us to a false negative conclusion, the power of the study to detect the effects we were seeking and the implications of the conclusion for biological understanding and clinical practice.

The study was designed to investigate a relatively homogeneous group of subjects with western European descent. Genetic ancestry was confirmed by the presence of the expected European Y haplogroups. All molecular characterisations were

Table 3 Sperm concentrations ($\times 10^6$/ml) of gr/gr deleted and gr/gr deleted–b2/b4 duplicated (single or multiple) men, classified on the basis of the type of missing *DAZ* and *CDY1* copies

	Normal sperm count, mean (SD) (n)			Abnormal sperm count, mean (SD) (n)			
	gr/gr del	gr/gr del–b2/b4 dupl and MD	gr/gr del and gr/gr del–b2/b4 dupl and MD	gr/gr del	gr/gr del–b2/b4 dupl and MD	gr/gr del and gr/gr del–b2/b4 dupl and MD	Total
Type of missing *DAZ* copies							
DAZ1/2	62.2 (42.0) (14)	31.9 (13.4) (4)	55.5 (39.4) (18)	3.1 (3.9) (60)	3.3 (3.6) (14)	3.2 (3.8) (74)	13.4 (27.1) (92)
DAZ3/4	71.4 (46.7) (10)	36.5) (1)	68.2 (45.5) (11)	2.8 (3.5) (28)	2.9 (5.7) (4)	2.8 (3.7) (32)	19.6 (36.6) (43)
Total	66.0 (43.3) (24)	32.8 (11.8) (5)	—	3.0 (3.8) (88)	3.2 (3.9) (18)	—	—
Type of missing *CDY1* copy							
CDY1a	52.7 (34.2) (15)	40.2 (8.6) (3)	50.6 (31.6) (18)	2.6 (3.4) (61)	3.4 (3.9) (14)	2.8 (3.5) (75)	12.0 (23.6) (93)
CDY1b	81.38 (51.4) (10)	21.7 (1.8) (2)	71.4 (52.0) (12)	3.5 (4.4) (30)	2.0 (4.1) (5)	3.3 (4.4) (35)	20.7 (39.5) (47)
Total	64.20 (43.3) (25)	32.8 (11.8) (5)	—	2.9 (3.8) (91)	3.1 (3.9) (19)	—	—
Deletion subtypes							
1	48.3 (28.2) (8)	42.0 (11.2) (2)	47.0 (25.3) (10)	2.6 (3.5) (35)	3.7 (3.3) (10)	2.8 (3.5) (45)	10.9 (20.3) (55)
2	64.2 (42.0) (6)	36.5) (1)	60.2 (39.7) (7)	3.1 (3.4) (22)	2.9 (5.7) (4)	3.1 (3.7) (26)	15.2 (29.5) (33)
3	80.8 (52.5) (6)	21.7 (1.8) (2)	66.0 (52.1) (8)	4.1 (4.5) (23)	2.5 (4.6) (4)	3.9 (4.5) (27)	18.1 (35.7) (35)
4	82.2 (57.8) (4)	(0)	82.2 (57.8) (4)	1.5 (3.7) (7)	(0)	1.5 (3.7) (7)	30.9 (51.7) (11)
Total	66.0 (43.3) (24)	32.8 (11.8) (5)	—	3.0 (3.8) (87)	3.2 (3.9) (18)	—	—

Del, deletion; dupl, duplication; MD, multiple duplication
The means of sperm concentration in these gr/gr rearrangements, further divided on the basis of the deletion subtype, are also shown.
Deletion subtypes are: (1) *DAZ1*/*DAZ2*+*CDY1a*; (2) *DAZ3*/*DAZ4*+*CDY1a*; (3) *DAZ1*/*DAZ2*+*CDY1b*; (4) *DAZ3*/*DAZ4*+*CDY1b*.

Table 4 Comparison of the frequency of gr/gr deletion and gr/gr deletion–b2/b4 duplication (single or multiple) in men with normal and abnormal sperm count, divided on the basis of their Y haplogroup

	Normal sperm count, n (%)			Abnormal sperm count, n (%)			
	gr/gr del	gr/gr del–b2/b4 dupl and MD	Total 1	gr/gr del	Gr/gr del–b2/b4 dupl and MD	Total 2	Total 1+ 2
Y haplogroup							
F*(xK)	6 (25)	1 (20)	7 (24.1)	34 (37.4)	8 (42.1)	42 (38.2)	49 (35.2)
P	14 (58.3)	2 (40)	16 (55.2)	42 (46.2)	6 (31.6)	48 (43.6)	64 (46)
K*(xP)	2 (8.3)	1 (20)	3 (10.3)	7 (7.7)	1 (5.3)	8 (7.3)	11 (7.9)
E	2 (8.3)	1 (20)	3 (10.3)	8 (8.8)	4 (21.1)	12 (10.9)	15 (10.8)
Total	24	5	29	91	19	110	139

Del, deletion; dupl, duplication; MD, multiple duplication; Total 1, n. gr/gr del plus n. gr/gr del–b2/b4 dupl and MD for subjects with normal sperm count; Total 2 = number of gr/gr del plus number of gr/gr del–b2/b4 dupl and MD for subjects with abnormal sperm count

conducted in a single centre (eliminating ~5% false positives from the original screen in the process), ensuring that uniform standards of genotype calling were applied to the entire set of samples. Our study was well-powered. For example, with a sample size of 134 and the standard requirement for $p = 0.05$, we would have 80% power to detect a variable that explained ⩾5.5% of the variation in sperm count. We therefore conclude that Y-chromosomal factors do not explain any substantial proportion of the variation in sperm phenotype of western European gr/gr deletion carriers.

The presence of certain Y haplogroups with constitutive deletions, both gr/gr and b2/b3, at high frequency in populations such as the Japanese and Finnish populations, respectively, led to the hypothesis that these Y chromosomes might also carry constitutive duplications that would counteract the deleterious effects of the gene dosage reduction. According to the hypothesis that gene dosage differences generally underlie phenotypic variation, we would expect to have seen a higher rate of del/dupl events in the normospermic group examined here. Our results excluded this possibility; on the contrary, they suggested that deletion followed by duplication may negatively affect spermatogenic efficiency. Subjects with deletion/duplications showed lower sperm concentration in both normospermic and infertile groups, although the finding reached significance only in the normospermic group. A possible explanation could be that deletion followed by duplications may indicate a higher propensity to genomic instability which may not be restricted to the *AZFc* region.

The definition of different *DAZ* and *CDY1* copies in our study has two implications. Several studies have indicated that *DAZ1/DAZ2* deletions are restricted to gr/gr carriers with impaired sperm production[14 18 21] and it was therefore proposed that these two copies may be biologically more important than the others. We found no significant phenotypic difference according to the deletion of different *DAZ* and *CDY1* gene copies, implying that these genes are either irrelevant to sperm output (which is unlikely to be true for both) or functionally equivalent. The second implication is that whereas the definition of the Y haplogroups provides information about the Y chromosome structure as a whole, the removal of different *DAZ* and *CDY1* copies may reflect a particular *AZFc* structure on which gr/gr deletion took place, which can be present in several lineages, albeit at different frequencies. Our data indicate that special rearrangements such as inversion polymorphism (which should precede, for example, the removal of *CDY1b* copy in case of sY1291 loss) are probably not relevant for the phenotypic expression of deletions. On the other hand we found significant differences in the haplogroup distribution between carriers with the loss of *CDY1a* and *CDY1b*, indicating that some structural rearrangements in the *AZFc* region may be associated with particular haplogroups (eg, *CDY1b* deletion in haplogroup I).

Table 5 Sperm concentration ($\times 10^6$/ml) of gr/gr deleted and gr/gr deleted–b2/b4 duplicated (single or multiple) men, classified on the basis of their Y haplogroup

	Normal sperm count, mean (SD) (n)			Abnormal sperm count, mean (SD) (n)			
	gr/gr del	gr/gr del–b2/b4 dupl and MD	Total 1	gr/gr del	gr/gr del–b2/b4 dupl and MD	Total 2	Total 1+ Total 2
Y haplogroup							
F*(xK)	65.2 (49.1) (6)	23.0 (1)	59.2 (47.6) (7)	4.3 (4.7) (34)	3.7 (4.5) (8)	4.2 (4.6) (42)	12.1 (26.1) (49)
P	45.8 (27.6) (14)	43.2 (9.5) (2)	45.5 (25.8) (16)	2.3 (3.0) (40)	2.6 (3.6) (5)	2.3 (3.0) (45)	13.7 (23.2) (61)
K*(xP)	94.0 (55.1) (2)	20.5 (1)	69.5 (57.6) (3)	1.8 (3.0) (5)	0.0 (1)	1.5 (2.8) (6)	24.2 (44.6) (9)
E	126.5 (37.5) (2)	34.1 (1)	95.7 (59.6) (3)	1.1 (1.9) (8)	3.8 (4.2) (4)	2.0 (3.0) (12)	20.7 (44.9) (15)
Total	61.4 (42.0) (24)	32.8 (11.8) (5)	—	3.0 (3.8) (87)	3.2 (3.9) (18)	—	—

Del, deletion; dupl, duplication; MD, multiple duplication.
Data are mean (SD) (n) with n shown on the lower line; in the cases where n = 1, there is no SD.

This study has important implications also for clinical practice. Although we did not find an evident correlation between Y haplogroups and the phenotypic expression of gr/gr rearrangements, it is worth noting that geographical differences were found. It is therefore likely that a genetic factor with a relatively mild phenotypic effect may be more penetrant in certain populations than in others and this may contribute to bias in admixed populations. Although we were unable to distinguish between "pathogenic" and "neutral" deletions, gr/gr deletions as a group are considered a significant risk factor for spermatogenic failure.[28 29] This conclusion might encourage some laboratories to screen for this genetic anomaly before using assisted reproductive techniques, as this genetic risk factor will be obligatorily transmitted to the male offspring. In this regard it is important to note that in the populations examined here, the false deletion rate of 5.3% in the sY1291-based test as it is routinely used is a cause for concern. Care should be taken to distinguish amplification failures from true deletions. Suspected deletions should be verified through the use of different PCR conditions or by redesigning the assay into a multiplex format with internal positive control fragments or performing a gene dosage analysis as used in this study. Second, there is at present no justification for subtyping of gr/gr deletions to inform counselling of carriers of European decent. We emphasise, however, that this last conclusion is only relevant to carriers with ancestry restricted to western Europe; in other areas of the world the matter remains open, and in a population with mixed western European/Japanese ancestry, identification of the low-risk haplogroup D gr/gr deletion chromosomes would be positively indicated.

In conclusion, our study has provided evidence that neither the definition of Y haplogroups nor the definition of gr/gr subtypes on the basis of the dosage or the missing *DAZ* and *CDY* copies are able to provide an explanation for the heterogeneous phenotype found in gr/gr deletion carriers. However, our study provides evidence for significant geographical differences in the distribution of deletion subtypes which may affect the outcome of case control association studies in different geographical areas.

Author affiliations: [1] Andrology Unit, Department of Clinical Physiopathology, University of Florence, Viale Pieraccini, Florence, Italy; [2] The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK; [3] Monash Institute of Medical Research and the Australian Research Centre of Excellence in Biotechnology and Development, Monash University, Clayton, Australia; [4] Institute of Reproductive Medicine University of Münster, Münster, Germany; [5] Department of Growth and Reproduction Copenhagen, University Hospital Rigshospitalet, Copenhagen, Denmark; [6] Human Genetics Research Group, Biochemistry and Molecular Genetics Service, Hospital Clínic, Faculty of Medicine, University of Barcelona and IDIBAPS, Barcelona, Spain; [7] Laboratoire de Biologie de la Reproduction, Pôle Mère Enfant, Laboratoire de Génétique et Biologie Moléculaire Section Molécules, Saint-Etienne, France; [8] Andrology-Urology Division, National Health Center, Budapest, Hungary; [9] University of Modena and Reggio Emilia Dept. of Medicine, Endocrinology, Metabolism and Geriatric, Modena, Italy; [10] Institut Clínic of Gynecology, Obstetrics and Neonatology, Hospital Clínic, Barcelona, Spain; [11] Laboratoire d'Histologie Embryologie Cytogénétique Biologie de la Reproduction, France; [12] Division of Endocrinology, Institute of Internal Medicine Polytechnic University of Marche, Ancona, Italy; [13] Deptarment of Surgery, Biology Section, University of Siena, Siena, Italy; [14] Prince Henry's Institute, Monash Medical Centre, Clayton, Australia; [15] Monash IVF Pty Ltd, Richmond, Australia

Acknowledgements: We would like to thank: the Italian Ministry of Education and Research (MIUR) for the grant PRIN 2008–2010 recipient C.K and the University of Florence; the Wellcome Trust (YX and CT-S); Ministerio de Ciencia y Tecnologia BFU2006-03479/BMC (RO). For clinical input and useful discussions, we thank M Maggi, I Noci, D M de Kretser, HWG Baker, E Carlsen and N E Skakkebaek; for technical assistance and useful discussions, we thank S Degl'Innocenti, F Nuti, E Guarducci, A Reilly, D Cram, M Lynch, S Borchert and D Garn; and for carrying out the basic semen analysis, we thank L Gambera and F Serafini, Center for Couple Sterility, Department of Pediatrics, Obstetrics and Reproductive Medicine, University of Siena, Siena, Italy.

Competing interests: None declared.

REFERENCES

1. **Krausz C,** Degl'Innocenti S. Y chromosome and male infertility: update, 2006. *Front Biosci* 2006;**11**:3049–61.
2. **McElreavey K,** Krausz C. Sex Chromosome Genetics '99. Male infertility and the Y chromosome. *Am J Hum Genet* 1999;**64**:928–33.
3. **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.
4. **Noordam MJ,** Repping S. The human Y chromosome: a masculine chromosome. *Curr Opin Genet Dev* 2006;**16**:225–32.
5. **Lin YW,** Hsu LC, Kuo PL, Kuo PL, Huang WJ, Chiang HS, Yeh SD, Hsu TY, Yu YH, Hsiao KN, Cantor RM, Yen PH. Partial duplication at AZFc on the Y chromosome is a risk factor for impaired spermatogenesis in Han Chinese in Taiwan. *Hum Mutat* 2007;**28**:486–94.
6. **Vogt PH.** AZF deletions and Y chromosomal haplogroups: history and update based on sequence. *Hum Reprod Update* 2005;**11**:319–36.
7. **Yen P.** The fragility of fertility. *Nat Genet* 2001;**29**:243–4.
8. **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.
9. **Krausz C,** Giachini C. Genetic risk factors in male infertility. *Arch Androl* 2007;**53**:125–33.
10. **Ravel C,** Chantot-Bastaraud S, El Houate B, Mandelbaum J, Siffroi JP, McElreavey K. GR/GR deletions within the azoospermia factor c region on the Y chromosome might not be associated with spermatogenic failure. *Fertil Steril* 2006;**85**:229–31.
11. **Machev N,** Saut N, Longepied G, Terriou P, Navarro A, Levy N, Guichaoua M, Metzler-Guillemain C, Collignon P, Frances AM, Belougne J, Clemente E, Chiaroni J, Chevillard C, Durand C, Ducourneau A, Pech N, McElreavey K, Mattei MG, Mitchell MJ Sequence family variant loss from the AZFc interval of the human Y chromosome, but not gene copy loss, is strongly associated with male infertility. *J Med Genet* 2004;**41**:814–25.
12. **Carvalho CM,** Zuccherato LW, Bastos-Rodrigues L, Santos FR, Pena SD. No association found between gr/gr deletions and infertility in Brazilian males. *Mol Hum Reprod* 2006;**12**:269–73.
13. **de Llanos M,** Ballesca JL, Gazquez C, Margarit E, Oliva R. High frequency of gr/gr chromosome Y deletions in consecutive oligospermic ICSI candidates. *Hum Reprod* 2005;**20**:216–20.
14. **Giachini C,** Guarducci E, Longepied G, Degl'Innocenti S, Becherini L, Forti G, Mitchell MJ, Krausz C.The gr/gr deletion(s): a new genetic test in male infertility? *J Med Genet* 2005;**42**:497–502.
15. **Hucklenbroich K,** Gromoll J, Heinrich M, Hohoff C, Nieschlag E, Simoni M. Partial deletions in the AZFc region of the Y chromosome occur in men with impaired as well as normal spermatogenesis. *Hum Reprod* 2005;**20**:191–7.
16. **Lynch M,** Cram DS, Reilly A, O'Bryan MK, Baker HW, de Kretser DM, McLachlan RI. The Y chromosome gr/gr subdeletion is associated with male infertility. *Mol Hum Reprod* 2005;**11**:507–12.
17. **Becherini L,** Guarducci E, Degl'Innocenti S, Rotondi M, Forti G, Krausz C. DAZL polymorphisms and susceptibility to spermatogenic failure: an example of remarkable ethnic differences. *Int J Androl* 2004;**27**:375–81.
18. **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.
19. **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.
20. **Repping S,** van Daalen SK, Brown LG, Korver CM, Lange J, Marszalek JD, Pyntikova T, van der Veen F, Skaletsky H, Page DC, Rozen S. High mutation rates have driven extensive structural polymorphism among human Y chromosomes. *Nat Genet* 2006;**38**:463–7.
21. **Ferlin A,** Tessari A, Ganz F, Marchina E, Barlati S, Garolla A, Engl B, Foresta C. Association of partial AZFc region deletions with spermatogenic impairment and male infertility. *J Med Genet* 2005;**42**:209–13.
22. **Karafet TM,** Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res* 2008;**18**:830–8.
23. **Paracchini S,** Arredi B, Chalk R, Tyler-Smith C. Hierarchical high-throughput SNP genotyping of the human Y chromosome using MALDI-TOF mass spectrometry. *Nucleic Acids Res* 2002;**30**:e27.

24. **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, Côrte-Real HB, de Knijff P, Decorte R, Dubrova YE, Evgrafov O, Gilissen A, Glisic S, Gölge 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, Nørby S, Pandya A, Parik J, Patsalis PC, Pereira L, Peterlin B, Pielberg G, Prata MJ, Previderé 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.
25. **Semino O,** Passarino G, Oefner PJ, Lin AA, Arbuzova S, Beckman LE, De Benedictis G, Francalacci P, Kouvatsi A, Limborska S, Marcikiae M, Mika A, Mika B, Primorac D, Santachiara-Benerecetti AS, Cavalli-Sforza LL, Underhill PA. The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. *Science* 2000;**290**:1155–9.
26. **World Health Organization**. WHO Laboratory Manual for the Examination of Human Semen and Sperm-Cervical Mucus Interaction 3rd. *Cambridge University Press, Cambridge, UK* 1999:4–33.
27. **Giachini C,** Nuti F, Marinari E, Forti G, Krausz C. Partial AZFc deletions in infertile men with cryptorchidism. *Hum Reprod* 2007;**22**:2398–403.
28. **Giachini C,** Laface I, Guarducci E, Balercia G, Forti G, Krausz C. Partial AZFc deletions and duplications: clinical correlates in the Italian population. *Hum Genet* 2008;**124**:399–410.
29. **Tuttelmann F,** Rajpert-De Meyts E, Nieschlag E, Simoni M. Gene polymorphisms and male infertility--a meta-analysis and literature review. *Reprod Biomed Online* 2007;**15**:643–58.