

Ancestral *RET* haplotype associated with Hirschsprung's disease shows linkage disequilibrium breakpoint at –1249

R M Fernandez, G Boru, A Peciña, K Jones, M López-Alonso, G Antiñolo, S Borrego, C Eng

J Med Genet 2005;42:322–327. doi: 10.1136/jmg.2004.023960

The *RET* proto-oncogene, located at 10q11.2, encodes receptor tyrosine kinase expressed during neural crest development and is involved in different human neurocristopathies such as the multiple endocrine neoplasia type 2 (MEN 2) syndromes and Hirschsprung's disease (HSCR).¹ HSCR (OMIM 142623), or aganglionic megacolon, is a common developmental disorder characterised by the absence of enteric neurones in variable lengths of the distal gastrointestinal tract, resulting in functional intestinal obstruction. Although *RET* is considered to be the major gene involved in HSCR, only a subset of HSCR patients can be attributable to traditional germline *RET* mutations (50% of familial HSCR and 10–15% of sporadic cases in selected series,^{2,3} or 30% of familial and 3% of sporadic cases in a population based study⁴). In an even smaller subset of HSCR patients (5–10%), germline variants are present in other genes, generally related to the developmental programme of neural crest cells, such as the glial cell line derived neurotrophic factor (*GDNF*), neurturin (*NTN*), endothelin 3 (*EDN3*), endothelin receptor β (*EDNRB*), endothelin converting enzyme 1 (*ECE1*), transcriptional factors *SOX10* and *PHOX2B*, and Smad interacting protein 1 (*SIP1*).^{5–7} In addition, *RET* modifying gene loci at 9q31, 3p21, 10q11, and 19q12 have been described associated with an HSCR phenotype.^{8,9} Because traditional germline mutations accounted for such a small subset of HSCR, we sought to determine whether there are more common susceptibility factors which predispose to a majority of HSCR cases. The silent variant A45A was described by our group in the only member from a MEN 2A/HSCR family co-segregating both phenotypes.¹⁰ When this variant in codon 45 (exon 2) and the *RET* haplotypes carrying it were found to be highly associated with a large subset of non-familial HSCR,^{11,12} it prompted us to propose that this was the associated variant itself, or that A45A was in linkage disequilibrium with the putative low penetrance susceptibility locus that would account for the majority of HSCR cases. Subsequently, we have shown that A45A anchors ancestral haplotypes in linkage disequilibrium, with the putative common founding susceptibility locus estimated to be 22 to 50 kb upstream.¹³ These observations have been confirmed in other populations,^{14–16} suggesting that this founder locus dates to the Stone Age.¹³

In this study we have undertaken systematic screening of the region upstream of the anchoring A45A single nucleotide polymorphism (SNP)—comprising *RET* intron 1, exon 1, and promoter—in 117 population based HSCR cases to find this putative founding locus.

METHODS

Patients and controls

In this study we included a series of 117 cases affected by clinically sporadic HSCR (20% female, 80% male), and their unaffected parents when available. More specifically, the triads (composed of the HSCR patient and both parents) were complete for 104 of the cases, while for 12 patients only one

Key points

- Several studies have shown that there are specific haplotypes of the *RET* proto-oncogene associated with the sporadic forms of Hirschsprung's disease (HSCR). Based on such studies, linkage disequilibrium (LD) mapping estimates predicted the presence of an HSCR locus located at 22–50 kb upstream of codon 45 (exon 2) of *RET*. The aim of this study was to identify a founding locus responsible for the majority of HSCR cases in a particular series.
- After a systematic mutational screening upstream of exon 2 of *RET*, the variants found were genotyped in 117 HSCR cases and 100 controls by fluorescence resonance energy transfer (FRET) or direct sequencing. Haplotype and genotype distributions were compared among these groups using standard case-control statistical analyses.
- A specific *RET* haplotype was identified, clearly linked to HSCR phenotype. The LD of the analysed variants was maintained along the whole region studied up to position –1249. The ancestral haplotype associated with HSCR was characterised by the presence of specific single nucleotide polymorphisms (at –200 and –196) in proximity to the transcriptional start site. Functional modelling using luciferase expression assays showed a significantly depressed activity for the HSCR linked haplotype at –200/–196 in comparison with other combinations associated with controls.
- These results seem to disprove the existence of an HSCR causing mutation as conceived in the traditional sense, but strengthen the concept of a specific combination of markers conferring susceptibility to the disease in a low penetrance fashion. It is conceivable that such an "HSCR haplotype", together with other events occurring in other genes, might give rise to the disease. This would be in agreement with a polygenic model for the disease.

of the parents was available. There was only one patient for whom we could not collect DNA from either parent. The characteristics of our cohort of HSCR patients have been described in detail previously.^{11–13,17} In addition, we also included a group of 100 normal controls comprising

Abbreviations: FRET, fluorescence resonance energy transfer; HSCR, Hirschsprung's disease; LD, linkage disequilibrium; SNP, single nucleotide polymorphism

Table 1 Allelic frequencies for the *RET* polymorphic loci in patients with Hirschsprung's disease (HSCR) and controls

<i>RET</i> sequence variants	Allelic frequency		χ^2 test with Yates correction
	HSCR	Controls	
-1249C→T	Polymorphic T	0.34	$\chi^2 = 3.32, p = 0.08$
	Wild type C	0.66	
-200A→G	Polymorphic G	0.39	$\chi^2 = 86.32, p < 0.00000001$
	Wild type A	0.61	
-196C→A	Polymorphic A	0.15	$\chi^2 = 43.7, p < 0.00000001$
	Wild type C	0.85	
IVS1+53G→A	Polymorphic A	0.15	$\chi^2 = 43.7, p < 0.00000001$
	Wild type G	0.85	
IVS1+620A→G	Polymorphic G	0.15	$\chi^2 = 43.7, p < 0.00000001$
	Wild type A	0.85	
IVS1+779C→T	Polymorphic T	0.15	$\chi^2 = 43.7, p < 0.00000001$
	Wild type C	0.85	
IVS1+2820C→T	Polymorphic T	0.15	$\chi^2 = 43.7, p < 0.00000001$
	Wild type C	0.85	
IVS1+2834C→T	Polymorphic T	0.12	$\chi^2 = 18.2, p = 0.000194$
	Wild type C	0.88	
IVS1+2846T→G	Polymorphic G	0.39	$\chi^2 = 75.2, p < 0.00000001$
	Wild type T	0.61	
IVS1+3105T→C	Polymorphic C	0.39	$\chi^2 = 75.2, p < 0.00000001$
	Wild type T	0.61	
IVS1+3991del12	deleted	0.86	$\chi^2 = 35.8, p < 0.00000001$
	Wild type	0.14	
IVS1+6136T→G	Polymorphic G	0.15	$\chi^2 = 43.7, p < 0.00000001$
	Wild type T	0.85	
IVS1+6174A→G	Polymorphic G	0.15	$\chi^2 = 43.7, p < 0.00000001$
	Wild type A	0.85	
IVS1+6383-84CC→GG	Polymorphic GG	0.15	$\chi^2 = 43.7, p < 0.00000001$
	Wild type CC	0.85	
IVS1+6411A→G	Polymorphic G	0.15	$\chi^2 = 43.7, p < 0.00000001$
	Wild type A	0.85	
IVS1+7236A→G	Polymorphic G	0.39	$\chi^2 = 75.2, p < 0.00000001$
	Wild type A	0.61	
IVS1+7436G→A	Polymorphic A	0.39	$\chi^2 = 75.2, p < 0.00000001$
	Wild type G	0.61	
IVS1+8406del16	deleted	0.86	$\chi^2 = 35.8, p < 0.00000001$
	Wild type	0.14	
IVS1+9202C→T	Polymorphic T	0.39	$\chi^2 = 75.2, p < 0.00000001$
	Wild type C	0.61	
IVS1+9277T→C	Polymorphic C	0.39	$\chi^2 = 75.2, p < 0.00000001$
	Wild type T	0.61	
IVS1+9494A→C	Polymorphic C	0.39	$\chi^2 = 75.2, p < 0.00000001$
	Wild type A	0.61	
IVS1+11260A→T	Polymorphic T	0.39	$\chi^2 = 75.2, p < 0.00000001$
	Wild type A	0.61	
IVS1+11369C→T	Polymorphic T	0.17	$\chi^2 = 35.8, p < 0.00000001$
	Wild type C	0.83	
IVS1+11637C→T	Polymorphic T	0.17	$\chi^2 = 35.8, p < 0.00000001$
	Wild type C	0.83	

Statistical analysis of the allelic distribution in both groups.

unselected, unrelated, race, age, and sex matched individuals from Spain.

Informed consent was obtained from all the individuals studied, in accordance with the respective committees for protection of human subjects.

Identification, characterisation, and genotyping of *RET* sequence variants

Genomic DNA was extracted according to standard protocols.¹⁸ Mutational screening of the whole intron 1 and promoter of *RET* was undertaken by direct sequencing of overlapping fragments obtained by polymerase chain reaction (PCR), using an Applied Biosystems 3730 DNA sequencer, and aligning our results with the sequence provided in Genbank accession number NT 033985.

Large scale genotyping of each of the variants was done either by fluorescence resonance energy transfer (FRET) technology or by direct sequencing of the PCR products. The FRET internal probes were designed according to the sequences of interest and purchased to TIB Molbiol.

Conditions for both FRET and sequencing, and the sequences of primers and probes, are available on request.

Statistical analysis

Allelic and genotypic frequencies of the *RET* polymorphisms analysed were calculated and then compared between patients and controls. Comparisons were made using χ^2 analysis with Yates' correction, with statistical significance set at $p < 0.05$.

Haplotypes were constructed comprising 23 of the 42 novel variants (from -200A→C to IVS1+11637C→T) as well as the other three SNPs located in the 3' end of the intron which had been analysed previously by our group.¹³ Generation of these extended haplotypes was possible and accurate because of the availability of the complete triad for 89% of the HSCR patients, which allowed us to reconstruct and compare the transmitted versus non-transmitted alleles. Based on the observation of complete linkage disequilibrium of an important part of the studied variants, and the existence of only a few combinations of these, it was plausible, and reliable, to infer the haplotypes for controls.

The haplotype distribution was compared between cases and controls. In addition, parental haplotypes were examined in the context of the affected children's haplotypes, so that transmitted and non-transmitted haplotypes were noted and their frequencies compared.

Promoter constructs and assays

Plasmid construction for the *RET* minimal promoter region

The 5' UTR region of the *RET* genomic DNA that comprises the variants -200A→G and -196C→A was PCR amplified using a forward primer (5' GGGCTAGCCCGCACTGAGCTC CTACAC-3') and a reverse primer (5' CCCTCGAGCTGGAGG GACTGCGGCTA-3'). These primers span the minimum promoter region (-313 to -24 from translational start site) and both have a restriction site (*NheI* and *XhoI* respectively) at the 5' end.

DNA samples with GA, GC, and AC (-200, -196) polymorphic variants were PCR amplified. The PCR products and the pGL3-basic luciferase report vector (Promega) were digested with *NheI* and *XhoI*. The fragments were cloned upstream of the *XhoI* site and transformed into *E coli* TOPO 10 competent cells following the manufacturer's instruction (Invitrogen). Transformants with the desired fragment were screened both by restriction digest and by direct sequencing using Big-Dye Terminators v 20 cycle sequencing kit (Applied Biosystems), and analysed on an ABI Prism 3700 DNA analyser (Applied Biosystems).

Transient transfection and dual-luciferase assay

The dual-luciferase reporter assay system (Promega) was used to assess the activity of the *RET* promoter variants (indicated above). A human MTC cell line TT was seeded at approximately 6×10^5 cells in a six-well tissue culture plate with RPMI supplemented with 10% fetal calf serum. Cultures were incubated for 18–24 hours at 37°C in a CO₂ humidified incubator.

When the cells were 40–60% confluent, they were washed and supplemented with serum-free media and transiently

transfected and co-transfected with constructs and pRL-TK (*Renilla* plasmid as internal control). Transfection was carried out using DMRIE transfection agent following the manufacturer's instructions (Life Technologies). Transfected cells were harvested 48 hours post-transfection and were lysed to carry out the dual-luciferase reporter assay according to the manufacturer's instructions (Promega).

RESULTS

Novel SNPs and other types of polymorphic loci were sought upstream of IVS1-1463 by direct sequencing of the 23 kb intron 1, exon 1, and the region 12 kb upstream of the translational start site. This revealed 42 novel SNPs and two insertion/deletion polymorphisms (available on request and posted to the *JMG* website). Once identified, all the variants were tested in a small group of patients and controls and the preliminary results showed a strong association with HSCR. Because our fundamental aim was the identification of the putative functional low penetrance susceptibility locus for HSCR guided by our previous statistical calculation,¹³ we only genotyped the variants located nearest to the 5' end of *RET* in the entire series of patients and controls. Following this strategy, we observed a lack of association of specific alleles with disease state at -1249C/T suggesting that linkage disequilibrium had broken ($p = 0.08$). Thus we subsequently concentrated on the region between -1249 and IVS1+11637,

	0	1a	1b*	1c	1d*	1e*	2a	2b*	2c	2d*	2e*	3*	4*
-200A→G	-	+	+	+	+	+	+	+	+	+	+	+	+
-196C→A	-	+	-	-	-	+	-	-	-	+	+	-	+
IVS1+53G→A	-	-	-	-	-	-	-	-	-	-	-	-	-
IVS1+620A→C	-	-	-	-	-	-	-	-	-	-	-	-	-
IVS1+779C→T	-	+	+	-	-	+	-	-	-	+	+	-	+
IVS1+2820C→T	-	-	-	-	-	-	-	-	-	-	-	-	-
IVS1+2834C→T	-	-	-	-	+	-	+	-	-	-	-	-	-
IVS1+2846T→G	-	+	+	+	+	+	+	+	+	+	+	+	+
IVS1+3105T→C	-	+	+	+	+	+	+	+	+	+	+	+	+
IVS1+3991del12	+	-	+	+	+	-	+	+	+	-	-	+	-
IVS1+6136T→G	-	-	-	-	-	-	-	-	-	-	-	-	-
IVS1+6174A→G	-	+	+	-	-	-	-	-	-	+	+	-	+
IVS1+6383-84CC→GG	-	-	-	-	-	-	-	-	-	-	-	-	-
IVS1+6411A→G	-	-	-	-	-	-	-	-	-	-	-	-	-
IVS1+7236A→G	-	+	+	+	+	+	+	+	+	+	+	+	+
IVS1+7436G→A	-	+	+	+	+	+	+	+	+	+	+	+	+
IVS1+8406del16	+	-	+	+	+	-	+	+	+	-	-	+	-
IVS1+9202C→T	-	-	-	-	-	-	-	-	-	-	-	-	-
IVS1+9277T→C	-	+	+	+	+	+	+	+	+	+	+	+	+
IVS1+9494A→C	-	+	+	+	+	+	+	+	+	+	+	+	+
IVS1+11260A→T	-	-	-	-	-	-	-	-	-	-	-	-	-
IVS1+11369C→T	-	+	+	+	+	+	+	+	+	+	+	+	+
IVS1+11637C→T	-	+	-	-	+	+	-	+	-	+	-	+	+
IVS1-1463T→C	-	-	-	-	-	-	-	-	-	-	-	-	-
IVS1-1370C→T	-	-	-	-	-	-	-	-	-	-	-	-	-
IVS1-126G→T	-	-	-	-	-	-	-	-	-	-	-	-	-

The haplotypes marked with an asterisk have a very low frequency.
+/-, presence/absence of the polymorphic variant.

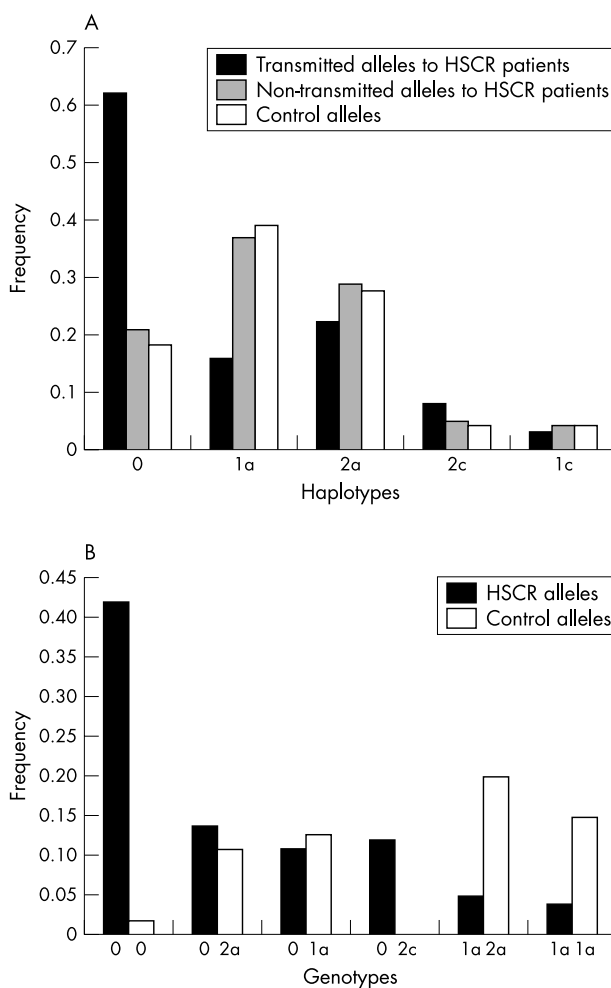


Figure 1 (A) Distribution of the extended haplotypes in transmitted and non-transmitted alleles within the Hirschsprung's disease (HSCR) families and in controls. (B) Distribution of the pairs of haplotypes in HSCR and controls.

which was found to contain 22 novel SNPs and two insertion/deletion polymorphisms (table 1). We observed that the distribution of each variant was significantly different in HSCR cases from controls, showing an under-representation of all SNPs and an over-representation of the deletions (table 1, $p < 0.0001$).

Because 89% of parental pairs were available for the HSCR cases, haplotypes comprising the analysed variants and those at the 3' end of the intron (IVS1-1463T→C, IVS1-1370C→T and IVS1-126G→T)¹³ could be constructed for 92 of the cases and inferred for the rest. Although we found 13 different haplotypes among the HSCR cases and their parents, some of these were very rare (table 2). In fact, each of the 13 extended haplotypes could be defined as different subtypes within the five fundamental intronic haplotypes previously published (0-4).¹³ Of note, haplotype 0 (previously defined as "HSCR haplotype")¹³ is always linked to a unique combination of alleles (table 2), and represents the most common haplotype among our cases (62%) compared with controls (17%).

As parents of control subjects were not available, the haplotype construction for this group was inferred from the haplotypes of the HSCR cases, as described elsewhere.^{12, 13} Using this method, haplotypes could be inferred for 82 of the 100 controls analysed. As shown in fig 1A, the difference in haplotype distribution observed in the cohort of patients and in controls was highly significant ($\chi^2 = 76.96$,

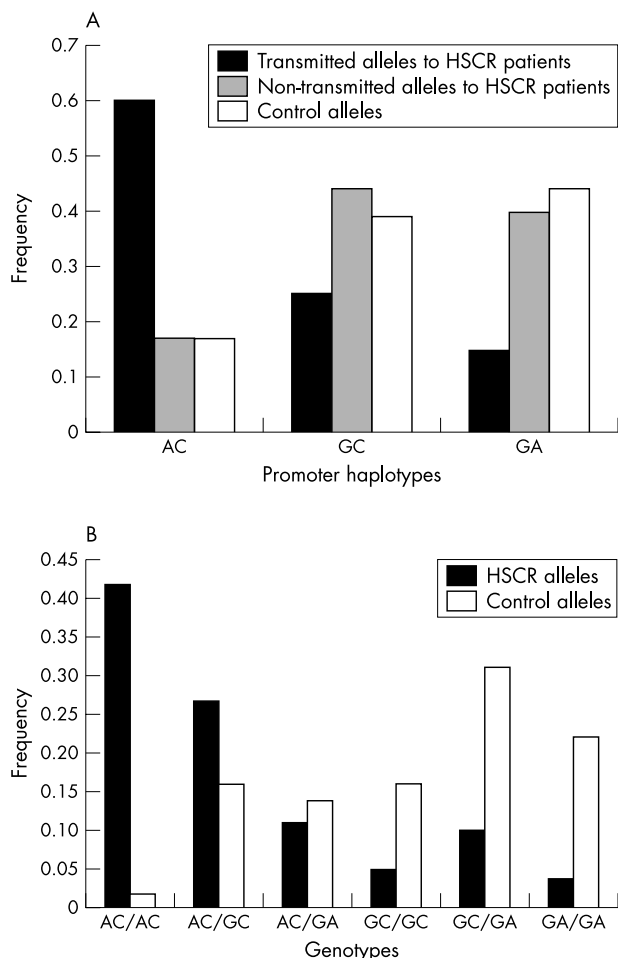


Figure 2 (A) Distribution of haplotypes comprising the *RET* promoter variants in transmitted and non-transmitted alleles within the Hirschsprung's disease (HSCR) families, and in controls. (B) Distribution of pairs of haplotypes comprising the *RET* promoter variants in HSCR patients and controls.

$p < 0.00000001$). In addition, we also calculated the haplotype frequencies among the non-transmitted alleles of the HSCR triads, and again observed significant differences when comparing transmitted alleles with HSCR patients ν non-transmitted alleles ($\chi^2 = 66.47$, $p < 0.00000001$; fig 1A). Conversely, the frequencies for each haplotype and the distribution are similar in the non-transmitted and the control alleles ($\chi^2 = 0.88$, $p = 0.93$).

We also compared the distribution of the pairs of haplotypes (genotypes) between patients and controls (fig 1B). Genotype 00 (42%) predominated in HSCR cases, while in the control group the most common genotype was 1a2a (21%), followed by 1a1a (15%). Perhaps the most interesting finding was that the "HSCR associated" genotype 00 was only present in 2% of controls.

Because only two variants, -200A→G and -196C→A, lie within the promoter region and are in proximity to the transcriptional start site (at -195), we calculated the frequencies of the short promoter specific haplotypes comprising only the loci in this region in cases and controls. Such *RET* promoter haplotypes could easily be assigned for all the individuals tested because of the genotyping technique employed (see fig 4). As shown in fig 2A, the short promoter haplotype distribution was significantly different in cases and controls, as was the frequency of particular transmitted alleles to HSCR patients ν non-transmitted alleles ($\chi^2 = 83.64$, $p < 0.00000001$), or ν control alleles ($\chi^2 = 84.02$, $p < 0.00000001$). It is important to note that haplotype AC is in complete linkage disequilibrium with all the risk haplotypes described so far¹¹⁻¹³ (table 2). In contrast, the frequencies for haplotypes GC (-200G -196C) and GA (-200G -196A) are considerably lower among HSCR chromosomes (25% and 15%) than among non-transmitted (44% and 40%) or control chromosomes (39% and 44%). Similar conclusions can be drawn from the comparison of the distribution of pairs of haplotypes (genotypes), where the differences observed are even more pronounced (fig 2B).

Luciferase reporter assay of *RET* promoter function

Luciferase activity was assayed in TT cells with constructs containing the three naturally existing promoter haplotypes at -200 and -196, GA, GC, and AC (see table 2). GC (or + - in table 2) had the highest expression (fig 3). AC (or - -), which is over-represented in HSCR, had the lowest expression (fig 4). Interestingly, GA (+ +) had an expression level less than that of GC but much more than AC, bringing its expression level much closer to that of GC than that of AC (fig 3).

DISCUSSION

It is currently well established that the *RET* proto-oncogene is the major gene implicated in the pathogenesis of Hirschsprung's disease. Besides traditional germline loss of function mutations causing HSCR, several common *RET*

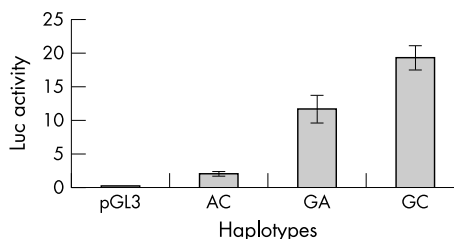


Figure 3 Luciferase assay showing decreased expression from the Hirschsprung's disease (HSCR) associated -200 -196 haplotype AC (corresponding to - - in table 2) compared with the control associated GA (+ +) haplotype.

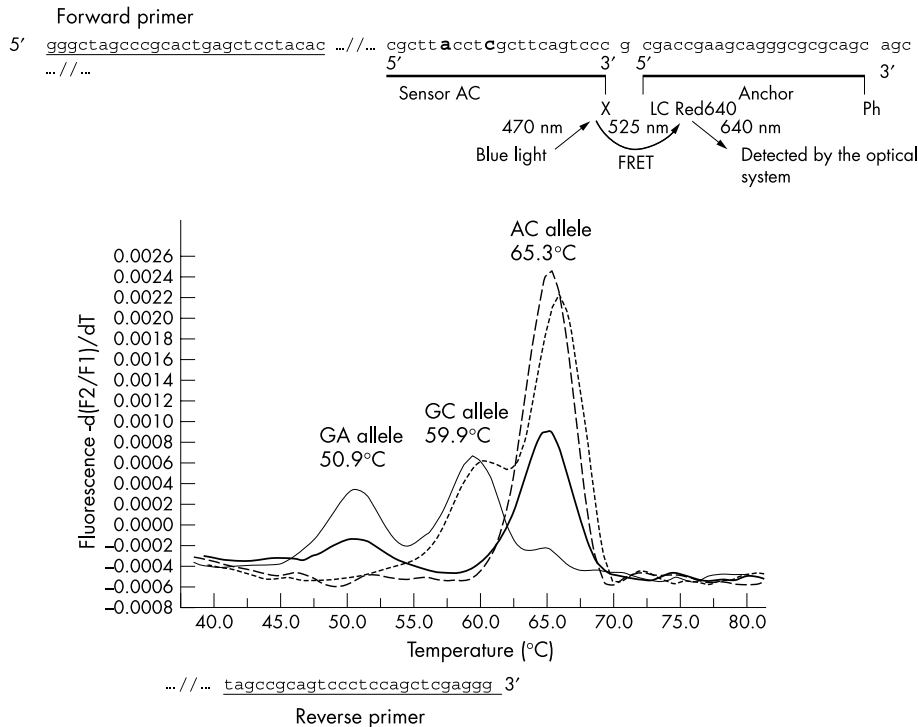


Figure 4 Fluorescence resonance energy transfer (FRET) pattern for the simultaneous detection of the *RET* variants $-200A \rightarrow G$ and $-196C \rightarrow A$.

proto-oncogene polymorphisms, present with different frequencies in different populations, have been associated with a variable risk of developing HSCR. In this sense, the HSCR associated variant firstly described was the silent polymorphism A45A, for which an association with the disease has been demonstrated in several cohorts of patients with different ethnic backgrounds, such as the Spanish, German, Italian, or Chinese.^{11 12 19–21}

Several hypotheses have been put forward to explain the underlying cause of such an association, including a direct action of the variant or the existence of linkage disequilibrium with some functional, still unknown allele. While no functional assays have yet confirmed the first hypothesis, the second has gained major relevance because of the results of subsequent studies. More specifically, we observed that the c.135A allele, irrespective of the 3' haplotype, was in linkage disequilibrium with a group of markers within the 3' end of the first intron (haplotype 0: IVS1–1463T, IVS1–1370C, IVS1–126G).¹³ These results strongly suggested the existence of a low penetrance locus of susceptibility for HSCR located 5' upstream all these markers.

Recently, the analysis of the promoter region of the *RET* proto-oncogene in three different populations of HSCR patients has revealed the existence of two new SNPs located at -5 and -1 from the transcription start site ($-200A \rightarrow G$ and $-196C \rightarrow A$) in linkage disequilibrium with A45A.^{14–16} Moreover, there is a strong association with HSCR in such populations, supporting the postulate previously mentioned. In particular, the “ACA” haplotype was shown to be associated with HSCR and to present a significantly lower activity in dual-luciferase expression assays in vitro, compared with those haplotypes identified in the majority of normal controls.¹⁴ In agreement with these data, it has been suggested that the “ACA” combination may represent a core haplotype associated with HSCR, acting as a modifying risk allele in the development of the condition.¹⁴

Based on our previous results, our group has carried out a systematic mutational screening along the 5' region of *RET*

comprising the whole intron 1, exon 1, and the region 12 kb upstream of the translational start site, in order to identify the putative founding locus accounting for the majority of sporadic HSCR cases.¹³ Although no germline mutations were identified in our patients, we found several novel polymorphisms that are in linkage disequilibrium with the so called haplotype 0¹³ and A45A, and strongly associated with HSCR. This association was supported not only by the results from comparative studies between HSCR cases and normal controls, but also from the TDT analysis in the HSCR triads. The linkage disequilibrium (LD) was maintained along the entire analysed region until position -1249 , where the lack of association of specific alleles with the disease suggested that LD is broken. The relevance of our results resides in the finding of a putative LD breakpoint which would allow us to delimit the 5' end of the “HSCR linked” *RET* region at a distance of approximately 24 kb with respect to A45A, in agreement with what it had been previously calculated by LD mapping estimates.¹³ This finding would be in contrast to the results of a previous study, in which the presence of a variant linked to A45A was described, located at approximately -4 kb from the translational start site.¹⁶ It would be plausible if the differences obtained in both studies reflected the distinct genetic background of the two populations examined. Future analyses of the region in different populations may help to elucidate this.

In conclusion, our findings seem to rule out the existence of an HSCR causing mutation, as conceived in the traditional sense, but strengthen the view that there is a specific combination of markers conferring susceptibility to the disease in a low penetrance fashion. The data derived from our functional in vitro studies, in agreement with other previously reported results,¹⁴ suggest that the haplotype 0 (table 2) may result in a lower level of expression of the *RET* gene. Thus homozygous individuals for this haplotype (62% of our cases and only 2% of controls) would present with considerably lower *RET* levels than those with other haplotype pairs. As the pathogenesis of HSCR is generally

ascribed to loss of function or haploinsufficiency of the *RET* gene, the decrease in *RET* transcription rate associated with this haplotype could have a functional effect linked to *RET* messenger RNA expression. It is conceivable that such a "HSCR haplotype", together with other events occurring in other genes, might give rise to the disease, which would be concordance with a polygenic model of the disease. However, the precise mechanism of actuation of this *RET* haplotype is not completely clear. Thus further studies are warranted to elucidate the functional relevance of this haplotype in the pathogenesis of Hirschsprung's disease.

ACKNOWLEDGEMENTS

This study was partially funded by grants 1R01HD39058 from the National Institutes of Health, Bethesda, Maryland, USA (to GA, SB, and CE); Red de Centros 03/05 from the Instituto de Salud Carlos III, Spain (to RMF, GA, and SB); and CAA 56/03 from the Consejería de Salud de la Junta de Andalucía, Spain (to RMF, AP, GA, and SB). RMF is a doctoral fellow of the Fundación Ramon Areces. AP is supported by Instituto de Salud Carlos III contract CP03/00045. CE is a recipient of the Doris Duke distinguished clinical scientist award.

Authors' affiliations

R M Fernandez, A Peciña, G Antiñolo, S Borrego, Unidad Clínica de Genética y Reproducción, Hospitales Universitarios Virgen del Rocío, Seville, Spain

G Boru, K Jones, C Eng, Human Cancer Genetics Program, Comprehensive Cancer Center, The Ohio State University, Columbus, Ohio, USA

M López-Alonso, Servicio de Cirugía Infantil, Hospitales Universitarios Virgen del Rocío, Seville, Spain

Competing interests: none declared

Correspondence to: Dr Salud Borrego, Unidad Clínica de Genética y Reproducción, Hospitales Universitarios Virgen del Rocío, Avda. Manuel Siurot s/n, 41013, Seville, Spain; salud.borrego.sspa@juntadeandalucia.es

Revised version received 2 August 2004

Accepted for publication 11 August 2004

REFERENCES

- Eng C. *RET* proto-oncogene in the development of human cancer. *J Clin Oncol* 1999;17:380–93.
- Attie T, Pelet A, Edery P, Eng C, Mulligan LM, Amiel J, Boutrand L, Beldjord C, Nihoul-Fékété C, Munnich A, Ponder BAJ, Lyonnet S. Diversity of *RET* proto-oncogene mutations in familial and sporadic Hirschsprung disease. *Hum Mol Genet* 1995;4:1381–6.
- Angrist M, Bolk S, Thiel B, Puffenberger, Hofstra RM, Buys CH, Cass DT, Chakravarti A. Mutation analysis of the *RET* receptor tyrosine kinase in Hirschsprung disease. *Hum Mol Genet* 1995;4:821–30.
- Svensson PJ, Molander JL, Eng C, Anvret M, Nordenskjöld A. Low frequency of *RET* mutations in Hirschsprung disease in Sweden. *Clin Genet* 1998;54:39–44.
- Chakravarti A, Lyonnet S. Hirschsprung Disease. In: Scriver CS, et al. *The metabolic and molecular bases of inherited disease*. New York: McGraw-Hill, 2002:6231–55.

- Wakamatsu N, Yamada Y, Yamada K, Ono T, Nomura N, Taniguchi H, Kitoh H, Mutoh N, Yamanaka T, Mushiaki K, Kato K, Sonta S, Nagaya M. Mutations in *SIP1*, encoding Smad interacting protein-1, cause a form of Hirschsprung disease. *Nat Genet* 2001;27:369–70.
- Amiel J, Laudier B, Attie-Bitach T, Trang H, de Pontual L, Gener B, Trochet D, Etchevers H, Ray P, Simonneau M, Vekemans M, Munnich A, Gaultier C, Lyonnet S. Polyalanine expansion and frameshift mutations of the paired-like homeobox gene *PHOX2B* in congenital central hypoventilation syndrome. *Nat Genet* 2003;33:459–61.
- Bolk S, Pelet A, Hofstra RM, Angrist M, Salomon R, Croaker D, Buys CH, Lyonnet S, Chakravarti A. A human model for multigenic inheritance: phenotypic expression in Hirschsprung disease requires both the *RET* gene and a new 9q31 locus. *Proc Natl Acad Sci USA* 2000;97:268–73.
- Gabriel SB, Salomon R, Pelet A, Angrist M, Amiel J, Fornage M, Attie-Bitach T, Olson JM, Hofstra R, Buys C, Steffann J, Munnich A, Lyonnet S, Chakravarti A. Segregation at three loci explains familial and population risk in Hirschsprung disease. *Nat Genet* 2002;31:89–93.
- Borrego S, Eng C, Sánchez B, Sáez ME, Navarro E, Antiñolo G. Molecular analysis of the *ret* and *GDNF* genes in a family with multiple endocrine neoplasia type 2A and Hirschsprung disease. *J Clin Endocrinol Metab* 1998;83:3361–4.
- Borrego S, Sáez ME, Ruiz A, Gimm O, López-Alonso M, Antiñolo G, Eng C. Specific polymorphisms in the *RET* proto-oncogene are over-represented in patients with Hirschsprung disease and may represent loci modifying phenotypic expression. *J Med Genet* 1999;36:771–4.
- Borrego S, Ruiz A, Sáez ME, Gimm O, Gao X, Lopez-Alonso M, Hernandez A, Wright FA, Antinolo G, Eng C. *RET* genotypes comprising specific haplotypes of polymorphic variants predispose to isolated Hirschsprung disease. *J Med Genet* 2000;37:572–8.
- Borrego S, Wright FA, Fernandez RM, Williams N, Lopez-Alonso M, Davuluri R, Antinolo G, Eng C. A founding locus within the *RET* proto-oncogene may account for a large proportion of apparently sporadic Hirschsprung disease and a subset of cases of sporadic medullary thyroid carcinoma. *Am J Hum Genet* 2003;72:88–100.
- Fitze G, Appelt H, König IR, Gorgens H, Stein U, Walther W, Gossen M, Schreiber M, Ziegler A, Roesner D, Schackert HK. Functional haplotypes of the *RET* proto-oncogene promoter are associated with Hirschsprung disease (HSCR). *Hum Mol Genet* 2003;12:3207–14.
- Sancandi M, Griseri P, Pesce B, Patrone G, Puppo F, Lerone M, Martucciello G, Romeo G, Ravazzolo R, Devoto M, Ceccherini I. Single nucleotide polymorphic alleles in the 5' region of the *RET* proto-oncogene define a risk haplotype in Hirschsprung's disease. *J Med Genet* 2003;40:714–18.
- Burzynski GM, Nolte IM, Osinga J, Ceccherini I, Twigg B, Maas S, Brooks A, Verheij J, Plaza Menacho I, Buys CHCM, Hofstra RMW. Localizing a putative mutation as the major contributor to the development of sporadic Hirschsprung disease to the *RET* genomic sequence between the promoter region and exon 2. *Eur J Hum Genet* 2004;12:604–12.
- Borrego S, Fernandez RM, Dziema H, Niess A, Lopez-Alonso M, Antinolo G, Eng C. Investigation of germline *GFRA4* mutations and evaluation of the involvement of *GFRA1*, *GFRA2*, *GFRA3*, and *GFRA4* sequence variants in Hirschsprung disease. *J Med Genet* 2003;40:e18.
- In: Dracopoli NH, Haines JL, Korf BR, Moir DT, Morton CC, Seidman CE, Seidman JG, Smith DR, eds. *Current protocols in human genetics*. New York: John Wiley and Sons, 1994.
- Fitze G, Schreiber M, Kuhlisch E, Schackert HK, Roesner D. Association of *RET* protooncogene codon 45 polymorphism with Hirschsprung disease. *Am J Hum Genet* 1999;65:1469–73.
- Griseri P, Pesce B, Patrone G, Osinga J, Puppo F, Sancandi M, Hofstra R, Romeo G, Ravazzolo R, Devoto M, Ceccherini I. A rare haplotype of the *RET* proto-oncogene is a risk-modifying allele in Hirschsprung disease. *Am J Hum Genet* 2002;71:969–74.
- Garcia-Barcelo MM, Sham MH, Lui VC, Chen BL, Song YQ, Lee WS, Yung SK, Romeo G, Tam PK. Chinese patients with sporadic Hirschsprung's disease are predominantly represented by a single *RET* haplotype. *J Med Genet* 2003;40:e122.

The gr/gr deletion(s): a new genetic test in male infertility?

C Giachini, E Guarducci, G Longepied, S Degl'Innocenti, L Becherini, G Forti, M J Mitchell, C Krausz

J Med Genet 2005;42:497–502. doi: 10.1136/jmg.2004.028191

Y chromosome microdeletions are the most frequent genetic cause of severe oligozoospermia (<5 million spermatozoa/ml) and azoospermia (absence of spermatozoa in the ejaculate).¹ Microdeletions associated with infertility occur in specific regions of the long arm of the Y chromosome, called azoospermia factor (AZF) regions.^{2–3} In 1996, three types of AZF deletion (AZFa, AZFb, and AZFc) were described by Vogt *et al*; however, after the complete physical map and sequence of the AZFb and AZFc regions was produced,⁴ it became evident that the AZFb and AZFc intervals partially overlap.⁵ The Y chromosome is extremely rich in repetitive sequences, organised in amplicons.⁶ Ampliconic sequences are characterised by sequence pairs showing nearly complete (>99.9%) nucleotide identity, organised in massive palindromes. These repeated sequences may undergo genetic exchange through gene conversion—that is, non-reciprocal transfer of sequence information occurring between duplicated sequences within the chromosome, a process that could account for the >99.9% nucleotide identity between the arms of a palindrome. Although this mechanism may serve to preserve Y chromosome genes from the gradual accumulation of deleterious mutations and thus prolong their genetic fitness,⁶ this peculiar organisation also provides the structural basis for deletions and rearrangements.

The classical AZFc deletion, which removes 3.5 Mb between the b2/b4 amplicons, is the most frequent type of deletion. Taking into consideration the Y chromosome structure and the suggested deletion mechanism, a number of other possible partial deletions have been proposed in both the AZFb and AZFc regions.^{7–8} The frequency and the pathological significance of these partial deletions is not yet clear, although recently a partial deletion termed gr/gr has been described specifically in infertile men with varying degrees of spermatogenic failure.^{9–10} This deletion removes half the AZFc gene content, including two copies of the major AZFc candidate gene called *DAZ*.¹¹ Another deletion, named b2/b3¹² or u3-gr/gr¹³ or g1/g3,¹⁴ which removes a similar quantity of AZFc genes, seems to have no effect on fertility status in association with a certain Y chromosome background commonly present in northern Eurasian populations (Y haplogroup N).^{12–13} A similar conclusion can also be drawn for the gr/gr deletion found in association with Hgr D2b, which is present in 20% of Japanese men. One of the possible explanations for the heterogeneous phenotype observed in association with complete and partial AZFc deletions is that polymorphisms or mutations are present in the autosomal homologue of *DAZ*, *DAZL*.¹⁵ Indeed, the finding that the human *DAZ* transgene is able to partially rescue the spermatogenic failure of mice homozygous for a null allele of *Dazl*¹⁶ suggests a possible interplay between *DAZL* and *DAZ* in humans. In the case of partial AZFc deletions, we can also speculate that the type and number of missing gene copies or other unknown Y chromosome related factors (for example duplications or beneficial mutations in other parts of the Y chromosome) may also influence the phenotype.

Key points

- Deletion of the AZFc region of the Y chromosome is the most frequent molecular genetic cause of oligo/azoospermia. Within this region, partial deletions have been recently described, including the gr/gr deletion.
- A direct aetiopathogenic role has been suggested for the gr/gr deletion as it was absent in normozoospermic but present in oligospermic men.
- We studied a group of normospermic (n=189) and oligo/azoospermic (n=150) men using a sequence tagged site (STS) +/- method. To gain insight into the molecular basis of the heterogeneous phenotype observed in men with the deletion we: (a) defined the type of *DAZ* and *CDY1* genes deleted and the Y haplogroup, and (b) carried out a mutational screen of *DAZL*, the autosomal homologue of *DAZ*.
- We found that: (a) partial AZFc deletions are not specific for spermatogenic failure, and (b) gr/gr deletion can be considered as a risk factor because its frequency was significantly higher in the oligo/azoospermic group (5.3%) than in controls (0.5%) p<0.012.
- Gene specific analysis revealed three distinct deletion patterns, indicating that further combined studies based on gene copy and haplogroup analysis are likely to provide a means for the distinction between pathogenic and neutral (or compensated by other Y factors) types of deletion.

The aims of this study were: (a) to establish the clinical significance of partial AZFc deletions (that is, if any of them are specific for spermatogenic failure or can be considered a risk factor); and (b) to gain insight through a combined approach (quantitative and qualitative analysis of *DAZ* and *CDY*, definition of Y chromosome haplogroup, and mutational analysis of *DAZL* in men with partial AZFc deletion) into the molecular basis of the heterogeneous phenotype described in men bearing partial AZFc deletions.

MATERIALS AND METHODS

Subjects

The study population consisted of 150 infertile patients (89 idiopathic and 61 with mild abnormal andrological findings: varicocele, monolateral cryptorchidism, recurrent infections) seeking complete andrological investigation for couple infertility at the Andrology Unit of the University Hospital Careggi in Florence, Italy, and 189 normospermic controls of

Abbreviations: AZF, azoospermia factor; SFV, sequence family variant; SNV, single nucleotide variant; STS, sequence tagged site

Italian origin. Cytogenetic analysis and Y chromosome microdeletion screening revealed 46,XY karyotype and the absence of AZF microdeletions in all included patients. Samples were collected using approved protocols and the informed consent of all individuals was obtained.

Screening for partial AZFc deletions

We tested eight STSs, originally described by Repping *et al.*⁹: sY1291, sY1191, sY1161, sY1206, sY1201, sY142, sY1258, and sY1197. In addition, we tested our samples with a complementary set of primers, o1084/1085 and o1276/1277.¹³ We identified gr/gr deletion by the following STS results: sY1291 and o1084/85 negative; sY1161, sY1191, sY1206, sY1201, and o1276/1277 all positive (an example of a PCR duplex result is shown in fig 1). The b2/b3 or u3-gr/gr deletion is characterised by the absence of the STS sY1191-o1276/1277 and the presence of the rest of the STS.

Qualitative and quantitative analysis for *DAZ* copy and *CDYa/CDYb* loss was performed according to Machev *et al.*¹³ For *DAZ*, we chose the sequence family variant (SFV) at STS sY587 in intron 10, which distinguishes *DAZ1/2* from *DAZ3/4*.

For *CDY1* we used a C/A SFV situated 7750 bp 5' of the *CDY1* translation start codon (CDY7750). In order to quantify copy number we used a PCR based method described by Machev *et al.*¹⁶ The PCR products were separated on an automatic sequencer (LI-COR Gene ReadIR 4200), and quantification was performed by ONE-Dscan.

Y chromosome haplotyping

Y chromosome haplotyping was performed as previously published for the YAP, M9, SRY4064, and 92R7 polymorphisms.¹⁷ The Tat polymorphism¹⁸ was assayed by PCR digestion with primers o1489: atgtatatagctctgtagg and o1490: gtaagcataattgagaagggtgcc (annealing 54°C). The 12f2 assay was performed as a duplex with *SRY* primers 3'SRY15 and 3'SRY16 and the 12f2 primers 12f2D and 12f2F.¹⁷ These primer pairs were used at a concentration of 300 nmol/l for each *SRY* primer and 600 nmol/l for each 12f2 primer, and the annealing temperature used was 52°C. Polymorphisms were visualised by restriction enzyme digest for M9 (*HinfI*), SRY4064 (*BsrBI*), Tat (*HpyCH4IV*), and 92R7 (*HindIII*).

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, and K (xN3, P), N3, and P, and one paragroup Y*(xD, E, J, K).

Statistical analysis

Statistical analysis was performed using the statistical package SPSS for Windows (version 11; SPSS, Chicago, IL, USA). We tested the significance of the observed difference in the incidence of gr/gr deletion between our infertile and control populations using Fisher's exact test. Our null hypothesis was that incidence is the same in infertile and control populations. All variables were checked for normal distribution by Kolmogorov-Smirnov one sample test. For comparisons of means between groups of different genotypes, Student's *t* test for independent samples, when normal distribution was observed, was applied. Logarithmic transformation of data was performed in order to normalise the distribution when the presence of log normal distribution was checked. Finally, in case of non-normalised distribution, the non-parametric Mann-Whitney *U* test was applied to achieve the same objective. A *p* value <0.05 was considered statistically significant for each test.

RESULTS

Frequency and type of partial AZFc deletions in patients and normospermic men

Based on the STS +/- analysis^{9,16} it was possible to distinguish between different types of partial deletions (fig 1). We found two types of partial AZFc deletions: the gr/gr (8/339) and the b2/b3 (3/339) deletion, whereas b1/b3 was absent in our study population.

The gr/gr and b2/b3 deletions were found in both the infertile and normospermic groups, although at different frequencies. The frequency of gr/gr deletions was significantly higher in the patient (8/150; 5.3%) than in the normospermic group (1/189; 0.5%) (*p*<0.012; odds ratio (OR) 10, 16; 95% confidence interval (CI) 1.28 to 80.3). In contrast, the frequency of the less common b2/b3 (u3-gr/gr) deletion was not different between the two groups (1.3% v 0.5%; *p* = 0.41).

DAZ and CDY1 gene copy definition

In order to further characterise the deletions, we defined the type of missing *DAZ* (*DAZ1/DAZ2/DAZ3/DAZ4*) and *CDY1* (*CDY1a/CDY1b*) gene copies. In the gr/gr deletion group, we found three types of deletion pattern: *DAZ1/DAZ2+CDY1a*, *DAZ3/DAZ4+CDY1a*, and *DAZ3/DAZ4+CDY1b* (fig 2). In the b2/b3 deletions group, we found two types of deletion pattern: *DAZ3/DAZ4+CDY1a* and *DAZ3/DAZ4+CDY1b*. The presence of different gene copy deletions patterns (five combinations found in our study) indicates that each type of partial deletion may be further divided into subtypes, and although the number of gene copies removed is the same, the missing copy type is different.

Interestingly, *DAZ1/DAZ2* and *CDY1a* were deleted only in the patient group, whereas *DAZ3/DAZ4* and *CDY1b* were deleted in both groups. However, the deletion of *CDY1a* was not consistently associated with the absence of *DAZ1/DAZ2* copies and thus can be deleted with either *DAZ* gene pair. The absence of the *DAZ* and *CDY* copies was further confirmed by a densitometric analysis, according to Machev *et al.*¹⁶ In all the gr/gr and b2/b3 deletion cases, we found a reduction of gene dosage for *DAZ* and *CDY1*. We have therefore no case of b2/b4 duplication among the men with deletions.

Genotype-phenotype association

Both types of deletion were associated with a wide range of sperm count, from azoospermia to normozoospermia (table 1). Although the mean values of the three principal sperm parameters were lower in patients with gr/gr deletion than in patients without, the difference was not significant: 2.1 v 4.5 million spermatozoa/ml for mean sperm concentration, 12.8% v 20.8% for sperm motility, and 11.6% v 15.3% for morphology. As *CDY1a* deletion was a specific feature of the patient group, we calculated the frequency of gr/gr deletions with missing *CDY1a* in both patients and controls; the difference between the two groups was highly significant (*p*<0.003). It is therefore possible that only specific patterns of partial deletions are deleterious for spermatogenesis.

We also sequenced the entire coding sequence of *DAZL* to test whether allelic forms of *DAZL* might underlie the phenotypic variation associated with the partial deletions. We found no new mutations. Only one single nucleotide variant (SNV) was found, at nucleotide position 260.¹⁹ Two patients were heterozygous for this SNV; however, its incidence in normospermic and infertile men is similar¹⁹ and thus it is unlikely to have any significant effect on the phenotype.

Genotype-Y chromosome haplogroup and phenotype relationship

Recent studies reported that b2/b3 (u3-gr/gr or g1/g3) deletions primarily arise in one family of closely related

Table 1 Phenotype of patients and controls bearing gr/gr and *b2/b3 deletions. The entire coding region of *DAZL* was sequenced, and no mutation was found except the single nucleotide polymorphism A→G transition in exon 2 (SNP 260). Patient A46 had undergone surgery at the age of 2 years for unilateral cryptorchidism

Code	Aetiology	DAZL (exons from 1 to 11)	Deleted gene copies		Semen parameters		
			DAZ	CDY	Number (n/ml*10 ⁶)	Motility (A+B %)	Morphology (%)
A49*	Idiopathic	No mutations	3/4	1b	0.38	10	13
A170	Idiopathic	No mutations	3/4	1a	0.9	20	16
A202	Idiopathic	No mutations	1/2	1a	10	30	13
A239	Idiopathic	No mutations	1/2	1a	4,2	14	16
A286	Idiopathic	SNP 260	3/4	1b	0.01	–	–
A 353*	Idiopathic	No mutations	3/4	1a	10	15	22
A 46	Cryptorchidism sx	No mutations	1/2	1a	0.01	–	–
A 184	Varicocele sx	No mutations	1/2	1a	0.6	0	2
A 234	Varicocele sx	SNP 260	3/4	1a	0.7	3	8
CS 64*	–	No mutations	3/4	1b	100	78	40
CS111	–	No mutations	3/4	1b	153	63	30

In contrast to the situation for the complete AZFc deletion, genotype–phenotype correlations for the partial AZFc deletion seem to be more complicated. Using an STS +/- approach^{9,10} we found two types of partial AZFc deletion, gr/gr and b2/b3 (u3-gr/gr or g1/g3) in our study population. Both deletion types were present in the normospermic group,

indicating that these deletions, in contrast to b2/b4 deletions, are not specific for spermatogenic failure. The gr/gr deletion seems to be more frequent than b2/b3 (or u3-gr/gr) in our Italian cohort (73% of the deletions were gr/gr). The similar frequency of the b2/b3 deletion in patients (1.3%) and controls (0.5%) does not support a pathogenic role for this

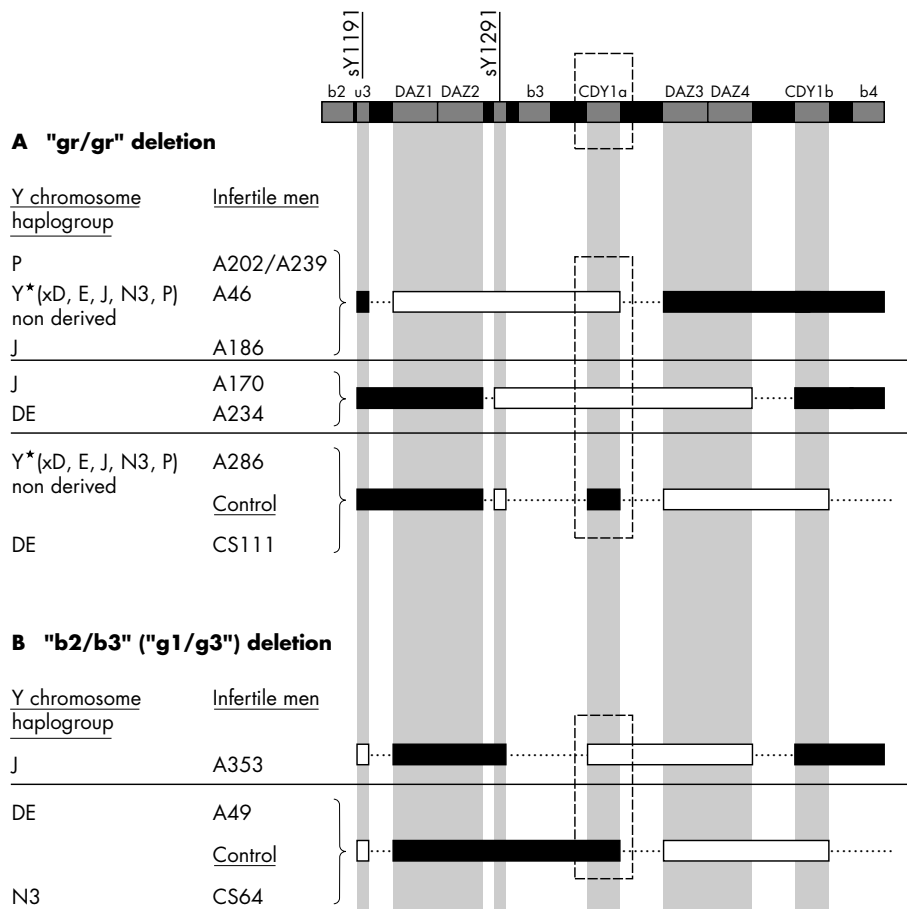


Figure 2 Schematic representation of the five distinct deletion patterns (based on the type of *DAZ* and *CDY* copies deleted) found in the 11 subjects with partial AZFc deletion. The gr/gr and b2/b3 (g1/g3) deletions were defined using a two step analysis based on STS +/- described by Repping *et al*⁹ and Machev *et al*.¹⁶ The orientation of the amplicons/sequences is not indicated because a number of possible inversion events may take place, leading to different deletion patterns. Open and filled boxes indicate the absence or presence of a given marker or gene, respectively.

deletion; however, it has been found at high frequency in the general population of northern Europe in association with hgrN, indicating that hgrN may contain compensating Y linked factors.^{12, 13} Our normospermic man with b2/b3 (or u3-gr/gr) deletion was hgrN3, whereas the two oligospermic men are hgrJ and hgrDE. Although the number is low, we can speculate that b2/b3 (u3-gr/gr; g1/g3) deletions are pathogenic only in association with certain haplogroups.

The frequency of gr/gr deletions was significantly higher in the infertile group with respect to controls, suggesting a possible deleterious effect of this polymorphism on spermatogenic efficiency. Similarly to the Repping study,⁹ the significant OR (10.1) indicates that gr/gr deletions can be considered as new risk factors for oligozoospermia. Although the mean sperm parameters (concentration, motility, and morphology) were not significantly different in the infertile group between men with gr/gr deletions and those without, there was a clear trend toward lower values in the former.

However, the most intriguing phenomenon observed in association with this type of deletion is the extreme heterogeneity of the phenotype, ranging from azoospermia to normozoospermia. In order to gain more insight into this phenomenon, we performed mutational analysis of *DAZL*, an autosomal homologue of *DAZ*. *DAZ* was copied to the Y chromosome relatively recently, in the old world primate lineage, and is 90% identical to its autosomal ancestor *DAZL*.²²

We found no new mutations in the entire coding region of *DAZL*, and the polymorphic Thr12-Ala change (T12A) does not seem to have any modulating effect, probably because of its relatively high frequency in the normospermic group. However, as our mutational analysis was focused exclusively on the coding regions, we cannot exclude promoter variations that might affect the level of expression of *DAZL*.

Interestingly, deletion of the *CDY1a* copy was found only in the patient group, providing an even more significant difference ($p < 0.003$) between the infertile and normospermic group. A similar phenomenon has also been observed in another study population from Mediterranean France.¹³

The different combinations of loci found deleted in cases of partial AZFc deletion indicate that a number of possible inversion events must have preceded these deletions.^{8, 9, 13} Clearly, the type of missing gene copies and the Y chromosome structure (hgr) on which the deletion arises are of fundamental importance for the understanding of a potential cause-effect relationship. Our data seem to support the hypothesis that *DAZ1/DAZ2* copies are functionally more important than *DAZ3/DAZ4*,²³ as the former were missing only in the patient group ($p = 0.037$); however, a recent study found no difference in *DAZ1/DAZ2* deletion frequency between infertile men and fertile controls (with unknown sperm count), which appears to contradict this observation.¹³ On the other hand, our observation that deletion of *CDY1* is strongly associated with infertility is consistent with the findings of this same study.¹⁶ This now requires further confirmation.

Although it is clear that the STS +/- analysis alone does not provide information about the type of missing gene copies and the hgr, we propose a flow chart (fig 1) in order to carry out partial AZFc deletion analysis in a cost effective and relatively simple manner. At this point, it is difficult to decide whether routine screening for partial AZFc deletions will be worthwhile. Clearly, gr/gr deletions, especially those resulting in the loss of *CDY1a*, can be considered a risk factor. However, the ideal situation would be to define a genetic profile specific for spermatogenic failure, although this may not exist.

Future studies in larger, well selected groups of subjects (patients without interfering abnormal andrological findings and normospermic controls) should focus on the combined

definition of the type and copy number of the AZFc genes deleted in men with partial deletions and the haplogroup of the Y chromosome. This will probably provide a means of distinguishing between pathogenic and neutral (or compensated by other Y factors) deletion types. If we consider that, together with the gr/gr deletion, other Y related factors (protective or negative) will be transmitted to the male offspring, we can eventually propose screening for partial AZFc deletion prior to assisted reproductive techniques. However, until a clear definition of pathogenic and non-pathogenic deletions is established, the prediction of the testicular phenotype of the offspring will remain rather vague.

ACKNOWLEDGEMENTS

We are grateful to Dr M Rotondi for his valuable contribution to the statistical analysis of our data. We thank the clinical staff of the Andrology Unit and Reproductive Physiology for patient's recruitment: Professor M Maggi, Dr G Corona, Dr M Magini, Dr L Petrone, Dr A Cilotti, Professor E Coccia, Professor I Noci. The study was supported by a grant from the University of Florence.

Authors' affiliations

C Giachini, E Guarducci, S Degl'Innocenti, L Becherini, G Forti, C Krausz, Andrology Unit, Department of Clinical Physiopathology, Center for Research, Transfer and Higher Education on Chronic, Inflammatory, Degenerative and Neoplastic Disorders, University of Florence, Viale Pieraccini, 6, Florence 50139, Italy
G Longepied, M J Mitchell, Inserm U.491, Faculté de médecine, 13385 Marseille, France; Institut de Médecine de la Reproduction, 13008 Marseille, France

Competing interests: none declared

The first two authors have contributed equally to the study and they should be regarded as joint first authors.

Correspondence to: Dr C Krausz, Andrology Unit, Viale Pieraccini, 6, Florence, 50139, Italy; c.krausz@dfc.unifi.it

Received 13 October 2004

Revised 29 November 2004

Accepted 1 December 2004

REFERENCES

- 1 **Krausz C, Forti G, McElreavey K.** The Y chromosome and male fertility and infertility. *Int J Androl* 2003;**26**:70–5.
- 2 **Tiepolo L, Zuffardi O.** Localization of factors controlling spermatogenesis in the non fluorescent portion of the human Y chromosome long arm. *Hum Genet* 1976;**34**:119–24.
- 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 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 **Repping S, Skaletsky H, Lange J, Silber S, Van der Veen F, Oates RD, Page DC, Rozen S.** Recombination between palindromes P5 and P1 on the human Y chromosome causes massive deletions and spermatogenic failure. *Am J Hum Genet* 2002;**71**:906–22.
- 6 **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.
- 7 **Yen P.** The fragility of fertility. *Nat Genet* 2001;**29**:243–4.
- 8 **Vogt PH.** Genomic heterogeneity and instability of the AZF locus on the human Y chromosome. *Mol Cell Endocrinol* 2004;**224**:1–9.
- 9 **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.

- 10 **Repping S**, Korver CM, Oates RD, Silber S, Van der Veen F, Page DC, Rozen S. Are sequence family variants useful for identifying deletions in the human Y chromosome? *Am J Hum Genet* 2004;**75**:514–17.
- 11 **Reijo RA**, Lee TY, Salo P, Alagappan R, Brown LG, Rosenberg M, Rozen S, Jaffe T, Straus D, Hovatta O, De la Chapelle 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.
- 12 **Repping S**, Van Daalen SK, Korver CM, Brown LG, Marszalek JD, Gianotten J, Oates RD, Silber S, Van der Veen F, Page DC, Rozen S. A family of human Y chromosomes has dispersed throughout northern Eurasia despite a 1.8-Mb deletion in the azoospermia factor c region. *Genomics* 2004;**83**:1046–52.
- 13 **Machev N**, Saut N, Longepied G, Terriou P, Navarro A, Levy N, Guichaoua M, Metzler-Guillemain C, Collignon P, Frances A-M, Beloungue J, Clemente E, Chiaroni J, Chevillard C, Durand C, Ducourneau A, Pech N, McElreavey K, Mattei M-G, 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.
- 14 **Fernandes S**, Paracchini S, Meyer LH, Florida 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* 2004;**74**:180–7.
- 15 **Yen PH**. Putative biological functions of the DAZ family. *Int J Androl* 2004;**27**:125–9.
- 16 **Slee R**, Grimes B, Speed RM, Taggart M, Maguire SM, Ross A, McGill NI, Saunders PT, Cooke HJ. A human DAZ transgene confers partial rescue of the mouse Dazl null phenotype. *Proc Natl Acad Sci U S A* 1999;**96**:8040–5.
- 17 **Rosser ZH**, Zerjal T, Hurler 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, Villes 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.
- 18 **Zerjal T**, Dashnyam B, Pandya A, Kayser M, Roewer L, Santos FR, Schiefenhoel W, Fretwell N, Jobling MA, Harihara S, Shimizu K, Semjiddmaa 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.
- 19 **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.
- 20 **Simoni M**, Bakker E, Krausz C. EEA/EMQN best practice guidelines for molecular diagnosis of Y-chromosomal microdeletions. State of the Art 2004. *Int J Androl* 2004;**27**:240–9.
- 21 **Krausz C**, Quintana-Murci L, McElreavey K. Prognostic value of Y deletion analysis: what is the clinical prognostic value of Y chromosome microdeletion analysis? *Hum Reprod* 2000;**15**:1431–4.
- 22 **Saxena R**, Brown LG, Hawkins T, Alagappan RK, Skaletsky H, Reeve MP, Reijo R, Rozen S, Dinulos MB, Distechi CM, Page DC. The DAZ gene cluster on the human Y chromosome arose from an autosomal gene that was transposed, repeatedly amplified and pruned. *Nat Genet* 1996;**14**:292–9.
- 23 **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.

CORRECTION

doi: 10.1136/jmg.2004.23960corr1

The Letter to JMG titled, Ancestral RET haplotype associated with Hirschsprung's disease shows linkage disequilibrium at –1249 (*J Med Genet* 2005;**42**:322–27) should have been published as a short report and the following abstract was omitted:

Background: Hirschsprung disease (HSCR) is a complex disorder with traditional germline mutations in *RET* in up to 30% of familial cases and in 3% of sporadic cases in a population-based series. We have previously demonstrated that an ancestral haplotype at the 5' end of *RET* (haplotype 0) was strongly associated with a large subset of isolated HSCR cases and that a putative low

penetrance susceptibility locus was encompassed within this ancestral haplotype, anchored by exon 2 SNP A45A.

Objective: To determine the 5' extent of the HSCR-associated ancestral haplotype by defining the linkage disequilibrium breakpoint in search for the low penetrance susceptibility locus.

Methods: Systematic screening of the region upstream of the anchoring A45A SNP, comprising *RET* intron 1, exon 1, and promoter in 117 population-based HSCR cases and 100 controls. Dual luciferase assay to determine differential activities between SNP combinations near a transcription start site.

Results: New SNPs were found which formed upstream haplotypes, anchored by A45A, in linkage disequilibrium with HSCR ($\chi^2 = 76.96$, $p < 0.00000001$). Linkage disequilibrium appeared to break at the –1249C/T SNP. Further, the HSCR-associated

genotype (00) was found in >60% of HSCR but only 2% of controls. Because only 2 variants, –200A>G and –196C>A, lie within the promoter region and are in proximity to the transcriptional start site (at –195), we modelled these combinations into constructs for luciferase reporter assay. The HSCR-associated SNP combination showed the lowest activity and the control-associated combination, the highest.

Conclusions: Our observations seem to discard the existence of a HSCR-causing mutation as it is conceived in the traditional sense, but strengthen the idea of a specific combination of variants conferring susceptibility to the disease in a low penetrance fashion. The data derived from our functional "in vitro" studies would suggest that the HSCR-associated haplotype 0 may result in a lower level of expression of the *RET* gene.

The journal apologises for this error.