RET genotypes comprising specific haplotypes of polymorphic variants predispose to isolated Hirschsprung disease

Salud Borrego, Agustín Ruiz, María Eugenia Saez, Oliver Gimm, Xin Gao, Manuel López-Alonso, Antonio Hernández, Fred A Wright, Guillermo Antiñolo, Charis Eng

Unidad de Genética Médica y Diagnóstico Prenatal, Hospital Universitario "Virgen del Rocío", Avda Manuel Suirot s/n, 41013 Sevilla, Spain
S Borrego
A Ruiz
M E Saez
G Antiñolo

Clinical Cancer Genetics Program, Comprehensive Cancer Center and Division of Human Genetics, Department of Internal Medicine, Ohio State University, Columbus, OH, USA
O Gimm
C Eng

Human Cancer Genetics Program, Ohio State University Comprehensive Cancer Center, 420 W 12th Avenue, Room 690 MRF, Columbus, OH 43210, USA
O Gimm
X Gao
F Wright
C Eng

Cirugía Infantil, Hospital Universitario "Virgen del Rocío", Sevilla, Spain
M López-Alonso
A Hernández

Cancer Research Campaign Human Cancer Genetics Research Group, University of Cambridge, Cambridge, UK
C Eng

Correspondence to:
Dr Borrego,
gantinolog@sego.es

Abstract

Background—Hirschsprung disease (HSCR), which may be sporadic or familial, occurs in 1:5000 live births and presents with functional intestinal obstruction secondary to aganglionosis of the hindgut. Germline mutations of the *RET* proto-oncogene are believed to account for up to 50% of familial cases and up to 30% of isolated cases in most series. However, these series are highly selected for the most obvious and severe cases and large familial aggregations. Population based studies indicate that germline *RET* mutations account for no more than 3% of isolated HSCR cases. Recently, we and others have noted that specific polymorphic sequence variants, notably A45A (exon 2), are overrepresented in isolated HSCR.

Purpose—In order to determine if it is the variant per se, a combination thereof, or another locus in linkage disequilibrium which predisposes to HSCR, we looked for association of *RET* haplotype(s) and disease in HSCR cases compared to region matched controls.

Methods—Seven loci across *RET* were typed and haplotypes formed for HSCR cases, their unaffected parents, and region matched controls. Haplotype and genotype frequencies and distributions were compared among these groups using the transmission disequilibrium test and standard case-control statistic.

Results—Twelve unique haplotypes, labelled A-L, were obtained. The distributions of haplotypes between cases and controls ($\chi_{11}^2 = 81.4$, p<<0.0001) and between cases and non-transmitted parental haplotypes were significantly different ($\chi_{11}^2 = 53.1$, p<0.0001). Genotypes comprising pairs of haplotypes were formed for cases and controls. There were 38 different genotypes among cases and controls combined. Inspection of the genotypes in these two groups showed that the genotype distribution between cases and controls was distinct ($\chi_{37}^2 = 93.8$, p<<0.0001). For example, BB, BC, BD, and CD, all of which contain at least one allele with the polymorphic A45A, are prominently represented among HSCR cases, together accounting for >35% of the case genotypes, yet these four genotypes were not represented among the population matched normal controls. Conversely, AA, AG, DD, GG, and GJ, none of which contains A45A, are commonly represented in the controls, together accounting for 43% of the control genotypes, and yet they are never seen among the HSCR cases.

Conclusions—Our data suggest that genotypes comprising specific pairs of *RET* haplotypes are associated with predisposition to HSCR either in a simple autosomal recessive manner or in an additive, dose dependent fashion.

(*J Med Genet* 2000;**37**:572–578)

The *RET* proto-oncogene, on 10q11.2, is a major susceptibility gene for Hirschsprung disease (HSCR, MIM 142623), a common disorder occurring in 1 in 5000 live births and characterised by the absence of the intramural ganglia of Meissner and Auerbach in the hindgut, which results in functional intestinal obstruction.[1–4] HSCR most commonly presents as isolated cases although it can be familial and may be inherited as an autosomal dominant or autosomal recessive trait, with reduced penetrance and male predominance.[1 5] Currently, at least seven related genes are believed to play some aetiological role in the pathogenesis of hereditary syndromic and non-syndromic HSCR.

The *RET* proto-oncogene encodes a receptor tyrosine kinase expressed in derivatives of the neural crest and neurectoderm.[6 7] Gain of function germline mutations in the *RET* proto-oncogene are associated with multiple endocrine neoplasia type 2 (MEN 2), an autosomal dominantly inherited cancer syndrome characterised by medullary thyroid carcinoma, phaeochromocytoma, and hyperparathyroidism.[8 9] Interestingly, loss of function germline *RET* mutations have been found in HSCR.[3 4] Depending on the series, up to 50% of familial HSCR cases and anywhere between 10 and 35% of sporadic cases were reported to be accounted for by loss of function germline *RET* mutations.[10–12] However, these series were highly selected, usually for familial cases or severe presentations. The only population based series, however, estimates the frequency of germline *RET* mutation in 69 unselected HSCR cases to be 7% and only 3% of isolated HSCR cases in this population based cohort had germline *RET* mutations.[13] Although several other putative HSCR susceptibility

genes have been proposed, including those that encode glial cell line derived neurotrophic growth factor (*GDNF*), one of the ligands for *RET*,[14–16] endothelin-3[17 18] and endothelin receptor-beta (*EDNRB*),[19–24] only germline heterozygous mutations in *EDNRB* occur in a significant minority of non-syndromic HSCR.[22]

When HSCR and MEN 2 occur together, the great majority are found to have C620R and C618R mutations.[8 25] A single case family segregating both HSCR and MEN 2 was found to harbour a germline *RET* C620S mutation, and at the time of original ascertainment, the only subject to have both HSCR and MEN 2 carried a C620S *RET* mutation and the homozygous sequence polymorphism in exon 2, A45A (c.135G→A).[26] Extending this observation in a population based series of isolated HSCR cases from western Andalucia, Spain, we found that the A45A (c.135G→A) sequence variant and L769L (c.2307T→G) (exon 13) were significantly over-represented (p<0.0006) when compared to region matched, race matched normal controls.[27] In contrast, two other polymorphisms, G691S (c.2071C→A, exon 11) and S904S (c.2712C→G, exon 15), were underrepresented in the HSCR patients compared to controls (p=0.02).[27] Interestingly, similar findings were independently obtained in a series of HSCR cases from different population bases, Germany[28] and the UK.[29] Using the same 64 isolated HSCR patients from our previous population based study[27] and newly accrued normal parents of these cases, we therefore sought to determine whether distinct germline *RET* sequence variant haplotypes could be directly associated with predisposition to HSCR.

Materials and methods

ISOLATED HSCR CASES
The western Andalucia region of Spain is serviced by the University Hospital "Virgen del Rocio" in Seville and is the major referral centre for HSCR, and so all HSCR cases seen at this institution may be considered representative of the population. This study used the first 64 consecutive cases of clinically sporadic HSCR seen at this institution in the first 13.5 months of study, described in detail previously.[27]

CLINICALLY UNAFFECTED PARENTS OF HSCR CASES AND NORMAL CONTROLS
Near the end of the study period, the unaffected parents of the isolated HSCR cases were recontacted to participate in the research protocol (in accordance with local institutional human protection committee rules). Forty nine parental couples and eight single parents agreed to participate.

Normal controls were unselected, unrelated, race matched subjects from Andalucia without a diagnosis of HSCR and who did not attend the medical genetic, paediatric surgical, or gastroenterological clinics.

MUTATION ANALYSIS
Genomic DNA was extracted from peripheral blood leucocytes from HSCR cases, their clinically unaffected parents or relatives, and normal controls using standard techniques.[30] To examine variant status at each of the polymorphic loci within *RET*, the appropriate *RET* amplicon for each of exons 2, 3, 7, 11, 13, 14, and 15 was generated as previously described.[25 27] The presence or absence of each polymorphism within each amplicon was assessed by differential restriction digestion with the appropriate enzymes as described,[27 31 32] according to the manufacturers' recommendations (Roche Life Technologies and Pharmacia Biotech). The products of restriction digestion were fractionated by electrophoresis through 2% agarose or 6% polyacrylamide gels and visualised under UV transillumination after ethidium bromide staining. When the primers were 5' labelled with fluorescent dyes, then the restricted amplicons were subjected to electrophoresis through an Alf-Express Automated DNA Sequencer (Pharmacia Biotech).

STATISTICAL ANALYSIS
Allelic frequencies at all seven *RET* polymorphic loci were determined and haplotypes formed (table 1). The frequencies of each haplotype were compared between the cases and race matched, region matched, normal controls who were unrelated to the HSCR subjects. Comparisons were performed using χ^2 contingency table tests with Yates's correction or (when computationally feasible) Fisher's exact test for tables with small expected cell counts. The criterion for statistical significance was set at α=0.05. Odds ratios with Cornfield 95% confidence intervals were also generated at each haplotype. The asymptotic p values from the transmission disequilibrium tests were supported in each case by exact p values computed via direct evaluation or 10 000 simulations.

Where available, parental haplotypes were examined in the context of the affected children's haplotypes. Transmitted and nontransmitted haplotypes were noted and compared. Statistical analyses and simulations were performed using S-Plus 4.5 (Mathsoft Inc). Stepwise logistic regression was performed using SAS v6.09 (SAS Institute), with entry and retention significance level criteria of 0.05. Although the transmitted haplotypes are matched with non-transmitted haplotypes,

Table 1 *RET haplotypes comprising various combinations of single nucleotide variants*

Haplo	A45A	V125V	A432A	G691S	L769L	S836S	S904S
A							
B	+						
C	+				+		
D			+				
E					+	+	
F	+		+		+		
G				+			+
H		+					
I			+	+			+
J	+			+			+
K	+		+				
L					+		

+, Variant present at locus.
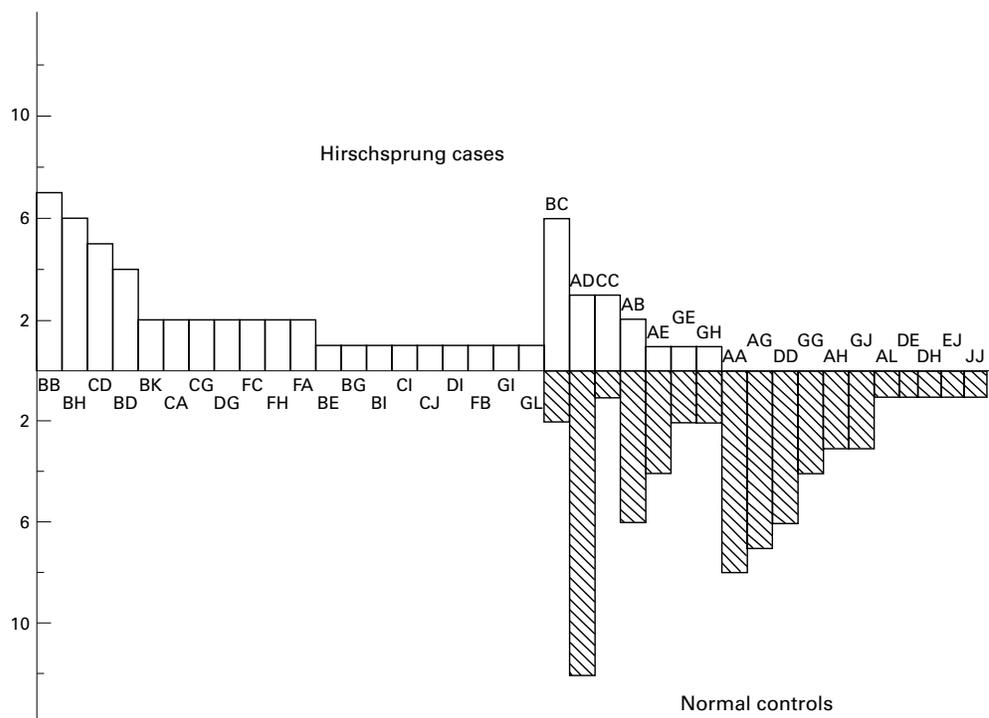Blank, wild type at locus.

Figure 1   Frequency of RET genotypes in HSCR cases (white bars) versus normal controls (striped bars). X axis, genotypes; Y axis, absolute number of cases.

unconditional logistic regression was appropriate (chapter 6 of Breslow and Day[33]) because marginal distributions of the predictors are assumed not to contain information about relative risk.

Results

RET HAPLOTYPE ANALYSIS IN HSCR AND UNRELATED CONTROLS

Haplotypes, comprised of variants across seven polymorphic loci in the coding region of *RET*, could be generated for 62 of the 64 HSCR cases and for 65 of the 104 race matched, region matched controls. Phase was successfully determined for 62 of 64 HSCR cases, 54 parents, and 65 controls. Among the 62 HSCR cases, seven with both parents lost to follow up had their haplotypes inferred from the haplotype compositions of the other subjects. This inference has no effect on the transmission disequilibrium test results, which use only families with informative parents. There were 12 unique haplotypes, labelled A-L, among cases and controls combined (tables 1 and 2A). Since recombination events among loci will be extremely rare, the haplotypes are viewed as individual alleles of a single locus. Apart from the haplotype comprised of the wild type allele at each of the polymorphic loci (haplotype A), there were 11 haplotypes (B-L) made up of various combinations and permutations of sequence variant and wild type sequence at each locus (table 1). The three most common haplotypes represented among controls occurring in more than 15% of chromosomes were A (36.2%), D (A432A only, rest wild type; 20%), and G [G691S (c.2071C→A) and S904S (c.2712C→G; 17.7%)] (table 2A, fig 1). In contrast, the most common haplotypes occurring in 15% or more of HSCR chromosomes were haplotypes B (A45A (c.135G→A) only; 33% of chromosomes) and C (A45A

Table 2A   Comparison of haplotype frequencies among HSCR cases and population normal controls

	HSCR patients		Normal controls		Statistic*	
Haplotype	No chrom	%	No chrom	%	χ_1^2	p value
A	10	8.1	47	36.2	27.2	<0.0001
B	38	30.7	8	6.2	24.0	<0.0001
C	25	20.2	4	3.1	16.7	<0.0001
D	15	12.1	26	20	2.4	0.12
E	3	2.4	8	6.2	1.3	0.25
F	6	4.8	0	0	Fisher	0.01
G	9	7.3	23	17.7	5.4	0.02
H	9	7.3	6	4.6	0.4	0.53
I	5	4.0	7	5.4	0.04	0.83
J	1	0.8	0	0	Fisher	0.49
K	2	1.6	0	0	Fisher	0.24
L	1	0.8	1	0.8	Fisher	1.0
Totals	124	100	130	100		

*The χ_1^2 is performed with Yates's correction. Where noted, the Fisher two tailed exact test was used.

Table 2B   Comparison of haplotype frequencies between HSCR patients, non-transmitted chromosomes of their parents, and control population. (Both homozygous and heterozygous parents are included)

Haplotypes	HSCR patients* (n=62 patients and 124 chromosomes)	Non-transmitted chromosomes (107 chromosomes)	Control population (n=65 unrelated subjects and 130 chromosomes)
A	10	27	47
B	38	12	8
C	25	5	4
D	15	26	26
E	3	5	8
F	7	1	0
G	9	17	23
H	9	2	6
I	4	2	7
J	1	2	0
K	2	0	0
L	1	8	1
Total	124	107	130

Comparison of transmitted *v* non-transmitted: χ_{11}^2 = 53.1, p<0.0001.
Comparison of non-transmitted *v* control: χ_{10}^2 = 19.1, p=0.039.

Table 3 Transmission disequilibrium testing comparing each haplotype to the group of remaining haplotypes

	Haplo											
	A	B	C	D	E	F	G	H	I	J	K	L
χ^2_1	14.2	10.5	10.7	4.5	0.5	4.0	3.9	4.5	1.3	0.3	2.0	5.4
p value	0.0002	0.0012	0.0011	0.0339	0.4795	0.0455	0.0495	0.0339	0.2568	0.5637	0.1573	0.0196

(c.135G→A) and L769L (c.2307T→G) only; 19%). Overall, the distribution of HSCR and control alleles across these 12 haplotypes was significantly different (χ^2_{11} =81.4, p<<0.0001, table 2A, B).

Three haplotypes, B, C, and F, were overrepresented among HSCR cases compared to normal controls (table 2A). Haplotype B, having only the polymorphic A45A (c.135G→A) and the rest of the loci wild type, was found in 38 (30.7%) HSCR chromosomes and only eight (6.2%) control chromosomes (OR=6.74 (2.84-16.53), χ^2_1=24.0, p<<0.0001). Haplotype C, harbouring A45A (c.135G→A) and L769L (c.2307T→G), with the rest of the loci wild type, was represented among 25 (20.2%) HSCR alleles compared to four (3.1%) control alleles (OR=7.95 (2.61-32.25), χ^2_1=16.7, p<<0.0001). Haplotype F, comprising A45A (c.135G→A), A432A, and L769L (c.2307T→G), comprised seven (5.6%) HSCR alleles compared to no control chromosomes (p=0.006, Fisher two tailed exact test).

Two *RET* haplotypes, A and G, appeared under-represented among HSCR cases compared to matched normal controls (table 2A). Haplotype A (all loci wild type) occurred among 10 (8.1%) HSCR chromosomes compared to 47 (36.2%) control chromosomes (OR=0.15 (0.07-0.34), χ^2_1=27.2, p<<0.0001). Haplotype G, with G691S (c.2071C→A) and S904S (c.2712C→G) only, was found in nine (7.3%) HSCR alleles compared to 23 (17.7%) control chromosomes (χ^2_1=5.4, p=0.02).

Other haplotypes, D, E, H, I, and L, showed no difference in distribution between HSCR cases and normal controls (table 2A). However, because haplotypes J and K are rare even among HSCR cases, and never observed in normal controls, no conclusions may be drawn because of small numbers.

Table 4 Transmission disequilibrium test by individual RET polymorphic locus

		Non-transmitted				
Transmitted		Wild type	Variant	TDT	p value	Exact p value
A45A	Wild type	37	9	28.5	<0.0001	5.3×10^{-8}
	Variant	50	11			
V125V	Wild type	99	1	4.5	0.033	0.035
	Variant	7	0			
A432A	Wild type	60	23	0.2	0.647	0.271
	Variant	20	4			
G691S	Wild type	73	19	1.6	0.209	0.075
	Variant	12	3			
L769L	Wild type	63	15	2.5	0.114	0.077
	Variant	25	4			
S836S	Wild type	99	5	0.5	0.480	0.145
	Variant	3	0			
S904S	Wild type	73	19	1.6	0.209	0.075
	Variant	12	3			

ASSOCIATION OF *RET* HAPLOTYPE WITH HSCR BY THE TRANSMISSION DISEQUILIBRIUM TEST

Because the non-transmitted parental haplotypes are available, we have also used these for analyses of association with disease that are highly robust to population stratification.[34 35] Table 2B shows the distribution of haplotypes that are transmitted to HSCR probands, those that are not transmitted from the parents, and those from the control group. The non-transmitted *v* control haplotypes do not appear to differ greatly in distribution, although a comparison of these two groups is significant at the 0.05 level (χ^2_{11}=19.1, p=0.039). A similar test of non-transmitted *v* transmitted haplotypes (chapter 4 of Lange[36]) is highly significant (χ^2_{11}=53.1, p<0.0001). The most significant contributions from individual haplotypes are A (p=0.012), B (p=0.0017), C (p=0.0027), D (p=0.046), H (p=0.034), and L (p=0.0196).

The transmission disequilibrium test[37] tests for association in a manner that acknowledges the matching of observed alleles, or in this instance haplotypes, in the parents of affected subjects. Only heterozygous parents are informative when performing this test. Although originally designed for biallelic markers, we applied multiple allelic/haplotype extensions here.[35] Table 3 presents the results comparing each haplotype to the group of remaining haplotypes. The highest frequencies of nontransmission were observed with haplotypes A, D, and G, and this was more pronounced when compared to the frequency of these haplotypes among the affected offspring. In contrast, haplotypes B, C, F, and H had the lowest rates of parental non-transmission, that is, these haplotypes were the most frequently transmitted from parent to affected offspring. As an overall test of association, Spielman and Ewens[35] propose a summation of the individual contributions. The corresponding statistic is χ^2_{11}=56.7, p<<0.0001. Another overall test uses the Bonferroni corrected p value to the maximum χ^2_1 statistic of the individual haplotypes, which yields p=0.002. Simulation based methods conditioned on the parental genotypes produced p values of 0.0001 and 0.0004, respectively.

ASSOCIATION OF INDIVIDUAL POLYMORPHIC LOCI WITH DISEASE

A series of transmission disequilibrium tests were also performed for each of the seven polymorphic loci (table 4). In these comparisons, only A45A (c.135G→A) and V125V were significant at the 0.05 level. However, G691S (c.2071C→A), L769L (c.2307T→G), and S904S (c.2712C→G) were suggestive with exact p<0.10.

In order to analyse more fully the effect of each locus on the transmission of a haplotype

Table 5 Logistic regression data

Variables	Significance level (p)	Odds ratio	95% CI for odds ratio
Stepwise logistic regression on all seven loci and 21 interaction terms			
A45A	0.0001	7.38	(3.93, 13.87)
V125V	0.006	9.11	(1.82, 45.49)
Regression on each locus			
A45A	0.0001	6.22	(3.36, 11.53)
V125V	0.075	4.11	(0.84, 20.08)
A432A	0.058	0.78	(0.43, 1.45)
G691S	0.081	0.52	(0.25, 1.10)
L769L	0.047	1.89	(1.0, 3.60)
S836S	0.069	0.51	(0.11, 2.23)
S904S	0.081	0.52	(0.25, 1.10)

to the affected proband, we performed stepwise multiple logistic regression (see Methods). Here, the response is defined as the event that a haplotype is transmitted to the affected offspring, and the potential predictors are the allelic states (wild type *v* variant) of the seven loci. With the inclusion of the seven loci and 21 two way interaction terms, only the main effects of A45A (c.135G→A) and V125V remained in the model (table 5). The effect of G691S (c.2071C→A)/S904S (c.2712C→G) is no longer suggestive in models which include A45A (c.135G→A).

HAPLOTYPE PAIRS (GENOTYPE) IN HSCR CASES AND CONTROLS

Genotypes comprising pairs of *RET* haplotypes were generated for cases, their participating parents, and controls. The three most common genotypes among the 65 controls include AD (12 or 18.5%), AG (seven or 10.8%), and AA (eight or 12.3%) (fig 1). Interestingly, only three (4.8%) HSCR cases carried one of these three genotypes and all three were AD. The three most common HSCR genotypes were BB (seven or 11.3%), BC (six or 9.7%), and BH (six or 9.7%). Only two (3.1%) normal controls carried any of these three genotypes and both were BC.

In summary, there were 38 different genotypes among cases and controls combined. Inspection of the genotypes in these two groups showed that the genotype distribution between cases and controls was distinct (χ^2_{37}=93.8, p<<0.0001). For example, BB, BC, BD, and CD are prominently represented among HSCR cases, together accounting for >35% of the case genotypes, yet these four genotypes are not represented among the region matched, race matched normal controls. Conversely, AA, AG, DD, GG, and GJ are commonly represented in the controls, together accounting for 43% of the control genotypes, and yet they are never seen among the HSCR cases.

Discussion

Among the western Andalucian HSCR cases, all *RET* haplotypes harbouring A45A (c.135G→A) appeared to be over-represented among HSCR cases compared to region matched, race matched, normal controls, thus confirming previous single site analyses and our hypotheses.[27] Interestingly, the L769L (c.2307T→G) polymorphism, which was previously found to be over-represented in single locus analysis in HSCR,[27] was very rare by

itself. It almost always occurred with the A45A (c.135G→A) variant, suggesting perhaps that it could be in linkage disequilibrium with the A45A (c.135G→A) allele. Conversely, selected non-A45A (c.135G→A) bearing haplotypes such as A and G were over-represented among controls. Even more powerfully associated with HSCR are the genotypes comprising specific *RET* haplotype pairs, such that our data might suggest an autosomal recessive or dose dependent (additive) low penetrance mechanism for a large proportion of isolated HSCR.

The precise haplotype-HSCR and single locus-HSCR associations compared to unrelated controls differ slightly from those compared to the unaffected parents. In our earlier report,[27] G691S (c.2071C→A)/S904S (c.2712C→G) were reported as conferring an apparent protective effect. Although the current results parallel an earlier report using unrelated control samples, the requirements of the transmission disequilibrium test reduce the effective sample size, so that the effect of G691S (c.2071C→A)/S904S (c.2712C→G) is not statistically significant here at the 0.05 level. Similarly, the stepwise multiple logistic regression shows that the effect of G691S (c.2071C→A)/S904S (c.2712C→G) is no longer suggestive in models which include A45A (c.135G→A). It can be seen by inspection of the haplotypes that variant A45A (c.135G→A) is associated with wild type sequence at codons 691 and 904 such that only one transmitted haplotype comprised variants A45A (c.135G→A) and G691S (c.2071C→A)/S904S (c.2712C→G). Thus, the apparent "HSCR protective" effect of variant G691S (c.2071C→A)/S904S (c.2712C→G) is largely confounded with the (stronger) effect at codon 45. Because of the association of the two loci, it is difficult to explore their precise joint effect on transmission status further. The observations using the transmission disequilibrium test have the advantage of being highly robust to population stratification, and there is evidence (table 2A, B) that the control haplotype distribution differs from that of the non-transmitted haplotypes.

Although this exploratory data set is relatively small, there are strong indications that haplotype pairs interact with each other to modulate HSCR phenotype. Under a simple additive (that is, dose dependent) or autosomal recessive model, therefore, we would expect that the presence of two HSCR associated haplotypes would be overwhelmingly associated with HSCR only and not controls. Conversely, the presence of two control associated haplotypes in a single subject should be strongly associated with normal controls and not with HSCR. In accordance with this model, the BB genotype was observed in seven (11.3%) HSCR cases but no controls. The AA, DD, and GG genotypes combined have been found in a total of 18 (27.7%) controls but no HSCR cases. Extrapolation of these observations would then lead to the hypothesis that heterozygous combinations of HSCR associated haplotypes would also be mainly associated with HSCR; similarly, heterozygous com-

binations of control associated haplotypes would not be associated with HSCR. Since haplotypes B, C, and F, singly, are over-represented in HSCR, we would expect to see the BC, BF, and CF genotypes mainly in HSCR and, indeed, this was observed in the western Andalucia data set where nine (14.5%) HSCR cases carry one of these three genotypes compared to two controls (BC, 3%) (fig 1). Conversely, then, the AD, AG, and DG genotypes should be mainly observed in controls and not HSCR. Indeed, in the western Andalucian data set, 27 (41.5%) controls have one of these three genotypes compared to five (8.1%) HSCR cases.

A HSCR associated haplotype occurring together with a control associated haplotype might give clues as to which, if any, predominates, that is, if a pure autosomal recessive model or, instead, a dose dependent model, holds for the entire cohort. The most obvious genotypes include AB, AC, AF, and BG. Although the F haplotype is relatively uncommon, it would appear that the F haplotype is particularly associated with HSCR and is relatively predominant. All genotypes with at least one F chromosome are only observed in HSCR cases no matter what the opposite chromosome is. The AF genotype, comprising the control associated A haplotype and the HSCR associated F haplotype, has been seen only in two HSCR cases and no controls, with the caveat of small numbers. The AB genotype was observed among six (9.2%) controls and two (3.2%) HSCR cases, perhaps suggesting a shared effect although another explanation is also possible (see below).

Genotypes with at least one E haplotype are worthy of note. The E haplotype contains the S836S variant which has been found to be associated with cases diagnosed with sporadic medullary thyroid carcinoma, in particular those whose tumours harbour somatic *RET* M918T.[32] If this observation is extrapolated, then the E haplotype may be viewed as a low penetrance "gain of function" *RET* allele, and should not be associated with HSCR and perhaps could "protect" against HSCR. The E haplotype is relatively rare in various populations at large[28 31 32] as well as in HSCR and so single site analysis at codon 836 and haplotype analysis (above) among the western Andalucian HSCR were not able to show that S836S is significantly under-represented in this population compared to controls because of small numbers. However, inspection indicates that there are clearly fewer S836S alleles among HSCR compared to controls both in the western Andalucian data set[27] and that from a German population.[28] Most genotype combinations with haplotype E, for example, AE, DE, GE, and EJ, occur among normal controls.

There are genotypes in the western Andalucian cohort of cases and controls which do not appear to "fit" our hypotheses. One of the most obvious examples would be the AD genotype. The A haplotype is significantly over-represented among controls compared to HSCR and the D haplotype is probably over-represented among controls as well. Thus,

the prediction under a simple, additive model would be that the AD genotype is observed only in controls. Although 12 (18.5%) controls carry the AD genotype, three (5%) isolated HSCR cases also have the AD genotype. It is possible that HSCR cases found to have non-HSCR associated genotypes harbour (that is, require) germline high penetrance mutations which result in HSCR. Similarly, it might be predicted under our simple additive model that the BC genotype would only be observed among HSCR cases but not controls. However, while six (9.7%) HSCR cases have this genotype, two (3%) controls also have this genotype. Currently, there is no obvious universal explanation except that other loci are also involved, a postulate that has already been established[38] and recently expanded, at least for familial HSCR.[39] However, it is also possible that the "controls" do have HSCR with decreased expression, perhaps manifesting as constipation and hence these subjects have not sought medical attention.

In summary, our observations show that *RET* genotypes comprising specific haplotypes of *RET* coding sequence variants are associated with isolated HSCR while other distinct haplotype pairs are associated with control status. Our data suggest that genotypes with haplotypes comprising one specific variant, A45A (c.135G→A), are particularly associated with HSCR. Whether it is the *RET* haplotype per se or whether particular haplotype(s) are in linkage disequilibrium with another low penetrance, autosomal recessive, HSCR susceptibility gene is as yet unknown.

A larger case-control series in this population as well as similar studies in other populations are required to confirm our findings. Systematic functional analyses will also need to be performed to help elucidate the precise mechanism of haplotype associated susceptibility.

1 Passarge E. The genetics of Hirschsprung's disease. *N Engl J Med* 1967;**276**:138-41.
2 Lyonnet S, Bolino A, Pelet A, Abel L, Nihoul-Fekété C, Briard ML, Mok-Siu V, Kaariainen H, Martucciello G, Lerone M, Puliti A, Luo Y, Weissenbach J, Devoto M, Munnich A, Romeo G. A gene for Hirschsprung disease maps to the proximal long arm of chromosome 10. *Nat Genet* 1993;**4**:346-50.
3 Edery P, Lyonnet S, Mulligan LM, Pelet A, Dow E, Abel L, Holder S, Nihoul-Fekété C, Ponder BAJ, Munnich A. Mutations of the *RET* proto-oncogene in Hirschsprung's disease. *Nature* 1994;**367**:378-80.
4 Romeo G, Ronchetto P, Luo Y, Barone V, Seri M, Ceccherini I, Pasini B, Bocciardi R, Lerone M, Kääriäinen H, Martucciello G. Point mutations affecting the tyrosine kinase domain of the *RET* proto-oncogene in Hirschsprung's disease. *Nature* 1994;**367**:377-8.
5 Badner JA, Sieber WK, Garver KL, Chakravarti A. A genetic study of Hirschsprung disease. *Am J Hum Genet* 1990;**46**:568-80.
6 Santoro M, Rosato R, Grieco M, Berlingieri MT, Luca-Colucci D'Amato G, de Franciscis V, Fusco A. The *ret* proto-oncogene is consistently expressed in human pheochromocytomas and thyroid medullary carcinomas. *Oncogene* 1990;**5**:1595-8.

7 Nakamura T, Ishizaka Y, Nagao M, Hara M, Ishikawa T. Expression of the *ret* proto-oncogene product in human normal and neoplastic tissues of neural crest origin. *J Pathol* 1994;**172**:255-60.

8 Eng C, Clayton D, Schuffenecker I, Lenoir G, Cote G, Gagel RF, Ploos van Amstel HK, Lips CJM, Nishisho I, Takai S-I, Marsh DJ, Robinson BG, Frank-Raue K, Raue F, Xue F, Noll WW, Romei C, Pacini F, Fink M, Niederle B, Zedenius J, Nordenskjöld M, Komminoth P, Hendy GN, Gharib H, Thibodeau SN, Lacroix A, Frilling A, Ponder BAJ, Mulligan LM. The relationship between specific *RET* proto-oncogene mutations and disease phenotype in multiple endocrine neoplasia type 2: International *RET* Mutation Consortium analysis. *JAMA* 1996;**276**:1575-9.

9 Eng C. *RET* proto-oncogene in the development of human cancer. *J Clin Oncol* 1999;**17**:380-93.

10 Attié 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.

11 Angrist M, Bolk S, Thiel B, Puffenberger EG, Hofstra RM, Buys CHCM, Cass DT, Chakravarti A. Mutation analysis of the RET receptor tyrosine kinase in Hirschsprung disease. *Hum Mol Genet* 1995;**4**:821-30.

12 Eng C, Mulligan LM. Mutations of the *RET* proto-oncogene in multiple endocrine neoplasia type 2, related sporadic tumours and Hirschsprung disease. *Hum Mutat* 1997;**9**:97-109.

13 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.

14 Salomon R, Attié T, Pelet A, Bidaud C, Eng C, Amiel J, Sarnacki S, Goulet O, Ricour C, Nihoul-Fékété C, Munnich A, Lyonnet S. Germline mutations of the RET ligand, GDNF, are not sufficient to cause Hirschsprung disease. *Nat Genet* 1996;**14**:345-7.

15 Angrist M, Bolk S, Halushka M, Lapchak PA, Chakravarti A. Germline mutations in glial cell line-derived neurotrophic factor (*GDNF*) and *RET* in a Hirschsprung disease patient. *Nat Genet* 1996;**14**:341-4.

16 Ivanchuk SM, Myers SM, Eng C, Mulligan LM. *De novo* mutation of GDNF, ligand for RET/GDNFR-α receptor complex in Hirschsprung disease. *Hum Mol Genet* 1996;**5**:2023-6.

17 Edery P, Attié T, Amiel J, Pelet A, Eng C, Hofstra RMW, Martelli H, Bidaud C, Munnich A, Lyonnet S. Mutation of the endothelin-3 gene in Waardenburg-Hirschsprung disease (Shah-Waardenburg syndrome). *Nat Genet* 1996;**12**:442-4.

18 Hofstra RMW, Osinga J, Tan-Sindhunata G, Wu Y, Kamsteeg EJ, Stulp RP, van Ravenswaaij-Arts C, Majoor-Krakauer D, Angrist M, Chakravarti A, Meijers C, Buys CHCM. A homozygous mutation in the endothelin-3 gene associated with a combined Waardenburg type 2 and Hirschsprung phenotype (Shah-Waardenburg syndrome). *Nat Genet* 1996;**12**:445-7.

19 Auricchio A, Casari G, Staiano A, Ballabio A. Endothelin-B receptor mutations in patients with isolated Hirschsprung disease from a non-inbred population. *Hum Mol Genet* 1996;**5**:351-4.

20 Amiel J, Attié T, Jan D, Pelet A, Edery P, Bidaud C, Lacombe D, Tam P, Simeoni J, Flori E, Nihoul-Fékété C, Munnich A, Lyonnet S. Heterozygous endothelin receptor B (*EDNRB*) mutations in isolated Hirschsprung disease. *Hum Mol Genet* 1996;**5**:355-7.

21 Bidaud C, Salomon R, Vancamp G, Pelet A, Attié T, Eng C, Bonduelle M, Amiel J, Nihoul-Fékété C, Willems PJ, Munnich A, Lyonnet S. Endothelin-3 gene mutations in isolated and syndromic Hirschsprung disease. *Eur J Hum Genet* 1997;**5**:247-51.

22 Chakravarti A. Endothelin receptor-mediated signaling in Hirschsprung disease. *Hum Mol Genet* 1996;**5**:303-7.

23 Kusafuka T, Wang Y, Puri P. Novel mutations of the endothelin-B receptor gene in isolated patients with Hirschsprung's disease. *Hum Mol Genet* 1996;**5**:347-9.

24 Puffenberger EG, Hosoda K, Washington SS, Nakao K, de Wit D, Yanagisawa M, Chakravarti A. A missense mutation of the endothelin B receptor gene in multigenic Hirschsprung's disease. *Cell* 1995;**79**:1257-66.

25 Mulligan LM, Eng C, Attié T, Lyonnet S, Marsh DJ, Hyland VJ, Robinson BG, Frilling A, Verellen-Dumoulin C, Safar A, Venter DJ, Munnich A, Ponder BAJ. Diverse phenotypes associated with exon 10 mutations of the *RET* proto-oncogene. *Hum Mol Genet* 1994;**3**:2163-7.

26 Borrego S, Eng C, Sánchez B, Sáez ME, Navarro E, Antiñolo G. Molecular analysis of *RET* and *GDNF* genes in a family with multiple endocrine neoplasia type 2A and Hirschsprung diease. *J Clin Endocrinol Metab* 1998;**83**:3361-64.

27 Borrego S, Saez ME, Ruiz A, Gimm O, Lopez-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.

28 Fitze G, Schreiber M, Kuhlisch E, Schackert HK, Roesner D. Association of the *RET* proto-oncogene codon 45 polymorphism with Hirschsprung disease. *Am J Hum Genet* 1999;**65**:1469-73.

29 Sancandi M, Ceccherini I, Costa M, Fava M, Chen B, Wu Y, Hofstra R, Laurie T, Griffths M, Burge D, Tam PK. Incidence of RET mutations in patients with Hirschsprung's disease. *J Pediatr Surg* 2000;**35**:139-42.

30 Eng C, Smith DP, Mulligan LM, Nagai MA, Healey CS, Ponder MA, Gardner E, Scheumann GFW, Jackson CE, Tunnacliffe A, Ponder BAJ. Point mutation within the tyrosine kinase domain of the *RET* proto-oncogene in multiple endocrine neoplasia type 2B and related sporadic tumours. *Hum Mol Genet* 1994;**3**:237-41.

31 Ceccherini I, Hofstra RMW, Luo Y, Stulp RP, Barone V, Stelwagen T, Bocciardi R, Nijveen H, Bolino A, Seri M, Ronchetto P, Pasini B, Bozzano M, Buys CHCM, Romeo G. DNA polymorphisms and conditions for SSCP analysis of the 20 exons of the *ret* proto-oncogene. *Oncogene* 1994;**9**:3025-9.

32 Gimm O, Neuberg DS, Marsh DJ, Dahia PLM, Hoang-Vu C, Raue F, Hinze R, Dralle H, Eng C. Over-representation of a germline *RET* sequence variant in patients with sporadic medullary thyroid carcinoma and somatic *RET* codon 918 mutation. *Oncogene* 1999;**18**:1369-70.

33 Breslow NE, Day NE. *Statistical methods in cancer research*. Vol 1. *The analysis of case-control studies*. Lyon: IARC Press, 1998.

34 Falk CT, Rubinstein P. Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. *Ann Hum Genet* 1987;**51**:227-33.

35 Spielman RS, Ewens WJ. The TDT and other family-based tests for linkage disequilibrium and association. *Am J Hum Genet* 1996;**59**:983-9.

36 Lange K. *Mathematical and statistical methods for genetic analysis*. Berlin: Springer, 1998.

37 Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 1993;**52**:506-16.

38 Puffenberger EG, Kauffman E, Bolk S, Matise T, Washington S, Angrist M, Weissenbach J, Garver K, Mascari M, Ladda R, Slaugenhaupt S, Chakravarti A. Identity-by-descent and association mapping of a recessive gene for Hirschsprung's disease on chromosome 13q22. *Hum Mol Genet* 1994;**3**:1217-25.

39 Bolk S, Pelet A, Hofstra RMW, Angrist M, Salomon R, Croaker D, Buys CHCM, 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.