

***RET* polymorphisms and the risk of Hirschsprung's disease in a Chinese population**

Cuiping Liu · Lei Jin · Hui Li · Jintu Lou ·
Chunfen Luo · Xuewu Zhou · Ji-Cheng Li

Received: 7 April 2008 / Accepted: 8 June 2008 / Published online: 8 July 2008
© The Japan Society of Human Genetics and Springer 2008

Abstract Hirschsprung's disease (HSCR) is a congenital disorder characterized by intestinal obstructions due to the absence of enteric ganglia along variable lengths of the intestinal tract. *RET* coding mutations have been found in approximately 50% of familial cases, but they only explain a minority of sporadic cases. Here, we report our investigation of a possible role of *RET* non-coding mutations in sporadic HSCR patients. The haplotypes of seven single nucleotide polymorphisms (SNPs), all located in a region 4 kb upstream of the gene through to 23 kb of intron 1, and one SNP in exon 2 were constructed in 125 Han Chinese patients with sporadic HSCR and in 148 Han Chinese controls. Our results indicated that eight SNPs were significantly associated with HSCR ($P < 0.0001$). The C allele of rs2505535 would appear to represent a protecting allele for the Chinese population. One single haplotype composed of these eight markers was present in 59.6% of patients, versus 18.1% of controls. Based on our results, we conclude that non-coding mutations in *RET* have important roles in the development of HSCR. The unknown functional disease variant(s), with a dosage-dependent effect in HSCR, is likely to be located in the 5'-region of the *RET* gene.

Key words Haplotype · Hirschsprung's disease · Linkage disequilibrium · *RET* proto-oncogene · Single nucleotide polymorphism

Introduction

Hirschsprung's disease (HSCR, OMIM 142623) represents the main genetic cause of functional intestinal obstruction, with an incidence of 1/5000 live births. This developmental disorder is a neurocristopathy and is characterized by the absence of the intestinal ganglion cells of the nerve plexuses in variable lengths of the digestive tract. The disease usually presents in infancy, although some patients present with persistent, severe constipation later in life. Symptoms in infants include difficult bowel movements, poor feeding, poor weight gain, and progressive abdominal distention (Chakravarti and Lyonnet 2002).

The length of the aganglionic segment of the large intestine has been used to classify the disease. In short segment HSCR (S-HSCR, 80% of cases), aganglionosis does not extend beyond the upper sigmoid, while in long segment (L-HSCR, 20% of cases), the aganglionosis extends more proximally. There are also rare cases, classified as total colonic aganglionosis (TCA) or total intestinal aganglionosis. Hirschsprung's disease most commonly presents sporadically, although it can be familial (approximately 20% of cases), and it is frequently associated with many other neurocristopathies and with chromosomal abnormalities. Males are affected fourfold more often than females, a difference most prominent in S-HSCR (Amiel et al. 2008).

Hirschsprung's disease is a heterogenic disorder, and a number of genes have been shown to play a role in the disease etiology (Parisi and Kapur 2000; Heanue and Pachnis 2007). Much genetic and functional evidence

C. Liu · L. Jin · J.-C. Li (✉)
Institute of Cell Biology, Zhejiang University,
Hangzhou, Zhejiang 310058, China
e-mail: lijichen@zju.edu.cn

H. Li · J. Lou
Department of Central Laboratory, Children's Hospital,
Zhejiang University School of Medicine, Hangzhou, China

C. Luo · X. Zhou
Department of Pediatric Surgery, Taizhou Hospital,
Taizhou, China

points to the *RET* gene (MIM 164761), located on chromosome 10q11.21, as a major disease-causing locus in HSCR (Edery et al. 1994). *RET* encodes a receptor tyrosine kinase that is expressed in cell lineages derived from the neural crest and which plays a crucial role in the regulation of cell proliferation, migration, differentiation, and survival during embryogenesis and functions as a receptor for growth factors of the glial cell line-derived neurotrophic factor (*GDNF*) family (Iwashita et al. 2001).

Nonetheless, in most studies, mutations in the *RET* coding sequence are identified only in 50% of familial and 10–20% of sporadic cases, which in total is only a small fraction of the HSCR patient population (Angrist et al. 1995; Gabriel et al. 2002). Despite this, several linkage and association studies still point to *RET* as a major HSCR gene (Bolk et al. 2000; Gabriel et al. 2002). These studies have indicated that the *RET* locus is linked to the disease in almost all familial cases, regardless of their mutation status, and is also associated with HSCR in a large proportion of patients with sporadic HSCR, who do not have *RET* coding mutations. These results suggest the existence of common non-coding *RET* variants or common polymorphisms that act in HSCR pathogenesis.

Several recent reports have shown that one haplotype, which starts in the 5'-region of the *RET* locus and spans approximately 27 kb [4 kb encompassing the 5'-untranslated region (UTR) and exon 1, 23 kb encompassing intron 1 and exon 2], is associated with the disease. The associated haplotype has a comparable frequency in all European populations (56–62%), while in Chinese Hong Kong populations, the frequency is substantially higher (85%) (Borrego et al. 2003; Garcia-Barcelo et al. 2003, 2005; Burzynski et al. 2004; Fitze et al. 2003; Pelet et al. 2005).

The associated haplotype was reconstructed with single nucleotide polymorphism (SNP) marker alleles. Some of these SNP alleles have been studied in more detail to determine a possible functional effect, such as rs1800858 located in exon 2 and two promoter SNPs, rs10900296 and rs10900297 (also known as SNP-5 G > A and SNP-1 A > C, respectively) (Fitze et al. 2002). However, none of these have been proven to be a real functional variant or to be in linkage disequilibrium (LD) with other closely located mutation(s).

Recent studies have suggested that the HSCR-associated *RET* mutations may be enhancer mutations, located approximately in the middle of intron 1 (Burzynski et al. 2005; Emison et al. 2005). Our previous study was consistent with this hypothesis. In particular, we further found three SNPs (rs2506005, rs2435357, rs2506004) in the enhancer region that were significantly associated with HSCR in Han Chinese patients (Zhang et al. 2007). However, other non-coding region *RET* variants in Han Chinese HSCR patients are largely unknown.

In the study reported here, we tested for allelic and haplotypic associations of *RET* with HSCR and investigated the possible role of haplotypes formed of eight SNPs in the pathogenesis of sporadic HSCR in a Chinese population.

Materials and methods

Subjects

This study was approved by the Ethics Committee of Zhejiang University, and all subjects gave informed consent for the genetic analyses. Blood samples were obtained from 125 unrelated patients (96 males, 29 females) diagnosed with sporadic HSCR. Their diagnoses were based on the histological examination of either biopsy or surgical resection material for the absence of enteric nerve plexuses. Twenty-six patients were affected with L-HSCR (including four with TCA) and 99 with S-HSCR. One hundred and forty-eight control DNA samples were obtained from a panel of unaffected individuals of Han Chinese backgrounds.

SNP selection and genotyping

We selected eight SNPs in the *RET* gene from the SNP database (<http://www.ncbi.nlm.nih.gov/snp/>). These markers included one SNP (rs741763: C > G, located 4 kb upstream of the transcriptional start site) and six SNPs located in intron 1 [IVS1 + 1813: G > A (rs2435365); IVS1 + 2846: C > A (rs1864410); IVS1 + 3460: C > T (rs2435364); IVS1 + 6000: C > A (rs2435362); IVS1-7593: G > A (rs752975); IVS1-2863: T > C (rs2505535)] and one coding SNP (rs1800858; in exon 2, also known as SNP2) (Fig. 1). Genomic DNA was isolated from peripheral blood leucocytes by standard procedures. Genotyping was performed by the ligase detection reaction, which has

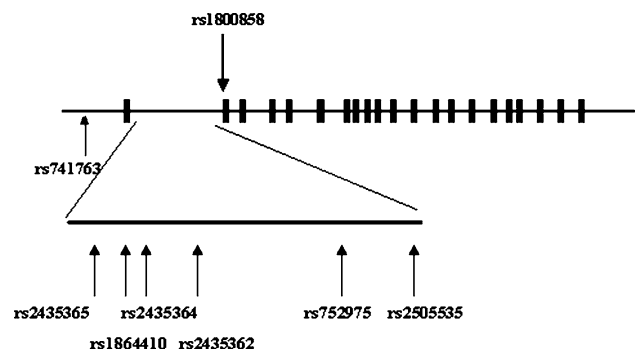


Fig. 1 Schematic representation of the *RET* gene and the locations of polymorphisms analyzed in this study. The *RET* gene structures are given in the upper part of the figure; the intron 1 region is magnified and approximate positions of typed single nucleotide polymorphisms (SNPs) are shown in the lower part

been demonstrated to be a highly specific and sensitive assay for SNP detection (Xiao et al. 2006). Fluorescent dye-labeled PCR products were used with an ABI PRISM 377 DNA Sequencer for genotyping, and data were analyzed using GENEMAPPER software (Applied Biosystems, Foster City, CA).

Statistical analysis

Deviation from Hardy–Weinberg equilibrium (HWE) was examined in controls by the χ^2 test. Based on the logistic regression method, the case–control association of genotypes in five inheritance models (co-dominant, dominant, recessive, over-dominant, log-additive) was tested, and the odds ratios (OR) and 95% confidence intervals (95% CI) were tested using SNPSTATS software (availability: <http://www.bioinfo.iconcologia.net/SNPstats>) (Sole et al. 2006). Subsequent statistical analyses were performed using SHEsis software (<http://www.analysis.bio-x.cn>) (Shi and He 2005). D' and r^2 were calculated to evaluate the magnitude of LD. Haplotype frequencies were estimated using the implementation of the EM algorithm coded into the HAPLO.STATS package (<http://www.mayoresearch.mayo.edu/mayo/research/biostat/schaid.cfm>). The association analysis of haplotypes was similar to that of genotypes with logistic regression, and the results are shown as the OR and 95% CI. The analysis used a two-tailed estimation of significance. Statistical significance was defined as $P < 0.05$.

Results

Nucleotide polymorphism analysis

For all eight SNPs, the genotypic distribution in controls conformed to HWE. Tables 1 and 2 summarize the allele and genotype frequencies of these eight SNPs. P values for Hardy–Weinberg proportions of the SNPs are also shown in Table 1. We found that a large proportion of our patients were homozygous at the markers that showed the highest association with HSCR, whereas these homozygous genotypes were comparatively lower in controls (Table 2). The ORs for the homozygote genotypes ranged from 3.909 to 21.523, whereas those for the heterozygote genotypes ranged from 0.982 to 6.087. Interestingly, only the A/G genotype for rs752975 had a significant difference in distribution between HSCR cases and controls.

To investigate the association of *RET* SNPs with HSCR phenotypes, the allelic distribution of the polymorphisms was analyzed between cases with L-HSCR and cases with S-HSCR. The variant C allele of rs2505535 was significantly under-represented in L-HSCR patients ($P = 0.015$; Table 3). Other SNPs showed no significant difference between S-HSCR and L-HSCR patients.

When logistic regression was used to carry out association analysis after modeling of the SNPs effects as additive, dominant, or recessive, the best inheritance models were obtained according to the smallest AIC (Akaike information

Table 1 The eight SNPs of the *RET* gene investigated in the cases ($n = 125$) and controls ($n = 148$)

dbSNP ID	Position	Allele	Allelic frequency		Statistic		OR (95% CI)
			Patients/cases <i>N</i> (Freq)	Controls <i>N</i> (Freq)	χ^2	<i>P</i> value	
rs741763	–4361C→G	G	217 (0.868)	204 (0.689)	24.548	<0.001	2.966 (1.908–4.410)
	HWE (<i>P</i>)			0.25			
rs2435365	IVS1 + 1813C→T	A	210 (0.84)	195 (0.691)	16.083	<0.001	2.342 (1.536–3.572)
	HWE (<i>P</i>)			1			
rs1864410	IVS1 + 2846G→T	A	190 (0.76)	128 (0.435)	58.63	<0.001	4.107 (2.835–5.950)
	HWE (<i>P</i>)			0.74			
rs2435364	IVS1 + 3460C→T	T	213 (0.852)	190 (0.642)	30.951	<0.001	3.212 (2.105–4.809)
	HWE (<i>P</i>)			0.37			
rs2435362	IVS1 + 6000C→A	A	194 (0.776)	139 (0.476)	51.324	<0.001	3.813 (2.620–5.552)
	HWE (<i>P</i>)			0.74			
rs752975	IVS1-7593G→A	A	208 (0.832)	174 (0.592)	37.267	<0.001	3.415 (2.278–5.121)
	HWE (<i>P</i>)			0.73			
rs2505535	IVS1-2863A→G	C	68 (0.272)	156 (0.542)	40.049	<0.001	0.316 (0.220–0.454)
	HWE (<i>P</i>)			0.74			
rs1800858	c135G(exon2)	A	168 (0.724)	131 (0.455)	17.566	<0.001	2.190 (1.513–3.170)
	HWE (<i>P</i>)			0.87			

HWE, Hardy–Weinberg Equilibrium; SNP, single nucleotide polymorphism; OR, odds ratio; 95% CI, 95% confidence interval; *N*, number of alleles; Freq, frequency

Table 2 Genotype frequencies of polymorphic variants of *RET* in cases ($n = 125$) and controls ($n = 148$)

SNP	Genotype	Patients/cases <i>N</i> (Freq)	Controls <i>N</i> (Freq)	OR (95% CI)
rs741763	C/C	4 (0.032)	11 (0.074)	1
	C/G	25 (0.200)	70 (0.473)	0.982 (0.286–3.367)
	G/G	96 (0.768)	67 (0.453)	3.940 (1.203–12.902)
Totals		125	148	
rs2435365	G/G	4 (0.032)	13 (0.092)	1
	A/G	32 (0.256)	61 (0.433)	1.705 (0.514–5.658)
	A/A	89 (0.712)	67 (0.475)	3.909 (1.223–12.497)
Totals		125	141	
rs1864410	C/C	13 (0.104)	48 (0.327)	1
	A/C	34 (0.272)	70 (0.476)	1.793 (0.858–3.748)
	A/A	78 (0.624)	29 (0.197)	9.600 (4.564–20.192)
Totals		125	147	
rs2435364	C/C	4 (0.032)	16 (0.108)	1
	C/T	29 (0.232)	74 (0.500)	1.568 (0.483–5.085)
	T/T	92 (0.736)	58 (0.392)	6.526 (2.079–20.492)
Totals		125	148	
rs2435362	C/C	11 (0.088)	41 (0.281)	1
	A/C	34 (0.272)	71 (0.486)	1.837 (0.843–4.004)
	A/A	80 (0.640)	34 (0.233)	8.519 (3.925–18.493)
Totals		125	146	
rs752975	G/G	6 (0.048)	74 (0.503)	1
	A/G	30 (0.240)	23 (0.156)	6.087 (5.956–43.450)
	A/A	89 (0.712)	50 (0.348)	21.523 (8.747–52.957)
Totals		125	147	
rs2505535	T/T	70 (0.560)	29 (0.201)	1
	C/T	42 (0.336)	74 (0.514)	7.613 (3.563–16.267)
	C/C	13 (0.104)	41 (0.285)	1.790 (0.863–3.714)
Totals		125	144	
rs1800858	G/G	11 (0.095)	42 (0.292)	1
	A/G	42 (0.362)	73 (0.507)	0.447 (0.183–1.092)
	A/A	63 (0.543)	29 (0.201)	2.056 (1.048–4.036)
Totals		116	144	

N, number of genotypes; Freq, frequency

criterion) value (Table 4). For rs741763, rs2435364, and rs752975, the recessive model was accepted as the best inheritance model; for rs1864410, rs2435362, and rs2505535, the co-dominant model was accepted as the best inheritance model; for rs2435365 and rs1800858, the log-additive model was accepted as the best inheritance model. Interestingly, the C allele of rs2505535 was represented as a protecting allele.

Pairwise linkage disequilibrium (LD)

Pairwise LD between the eight SNPs was calculated for cases and controls. D' and r^2 for all possible pairs of SNPs are shown in Fig. 2a and b, respectively. We found strong LD ($D' > 0.75$) among seven non-coding region SNPs

(rs741763, rs2435365, rs1864410, rs2435364, rs2435362, rs752975, rs2505535).

Haplotype analyses

Based on the single-locus association analyses, we estimated frequencies for the haplotypes consisting of the eight marker loci. Eighteen different haplotypes encompassing the eight SNPs were observed in cases and controls. Among these, five haplotypes showed significant differences between patients and controls (Table 5: haplotype-C, -G, -H, -P, -Q). Haplotype analysis showed that these eight risk-associated alleles were present in the same risk haplotype, haplotype-G. Its frequency was estimated to be 59.6% among patients and 18.1% among controls

Table 3 Allele frequencies (%) of the SNPs in S-HSCR and L-HSCR cases

dbSNP ID	Allele	Frequency (%)		Statistic								
		S-HSCR (99)	L-HSCR (26)	C (148)	χ^2 (S-HSCR vs. L-HSCR)	P value	χ^2 (S-HSCR vs. C)	P value	OR (95% CI)	χ^2 (L-HSCR vs. C)	P value	OR (95% CI)
rs741763	G	0.859	0.094	0.689	0.736	0.391	18.511	0.000	2.738 (1.712–4.379)	10.137	0.010	4.239 (1.633–11.008)
rs2435365	A	0.704	0.808	0.669	0.51	0.475	19.927	0.000	2.772 (1.754–4.381)	3.98	0.048	2.079 (1.001–4.318)
rs1864410	A	0.747	0.808	0.439	0.819	0.366	45.824	0.000	3.78 (2.548–5.607)	24.028	0.000	5.363 (2.593–11.094)
rs2435364	T	0.859	0.827	0.642	0.327	0.567	28.182	0.000	3.387 (2.128–5.393)	6.844	0.090	2.665 (1.251–5.681)
rs2435362	A	0.768	0.808	0.476	0.379	0.538	41.723	0.000	3.632 (2.432–5.423)	19.476	0.000	4.417 (2.233–9.546)
rs752975	A	0.813	0.904	0.416	2.425	0.119	76.739	0.000	6.12 (3.999–9.367)	42.207	0.000	13.221 (5.111–34.203)
rs2505535	C	0.303	0.135	0.272	5.954	0.015	27.061	0.000	0.368 (0.251–0.539)	29.242	0.000	0.132 (0.057–0.302)
rs1800858	A	0.742	0.660	0.724	1.313	0.252	37.368	0.000	3.442 (2.296–5.161)	7.178	0.007	2.326 (1.240–4.366)

S, Short; L, long

($P < 0.01$). Another haplotype, haplotype-H, which differs only from the risk haplotype in the eighth SNP, was present in 7.4% of patients and 22.1% of controls and appeared to be a protective haplotype ($P < 0.01$). We also found two haplotypes, I and H, that were present only in cases and in controls, respectively, confirming the important role of non-coding markers in the development of HSCR.

Discussion

Hirschsprung’s disease is considered to be a genetic disease caused by genetic alterations. To date, at least 11 genes have been associated with sporadic or syndromic forms of HSCR. These ‘HSCR genes’ are generally related to the developmental program of neural crest cells and include the *RET* proto-oncogene, *GDNF*, neurturin (*NTN*), endothelin 3 (*EDN3*), endothelin receptor (*EDNRB*), endothelin-converting enzyme 1 (*ECE1*), transcriptional factors *SOX10* and *PHOX2B*, Smad interacting protein 1 (*SIP1*), *KIAA1279*, and *TITF1* (Parisi and Kapur 2000; Garcia-Barcelo et al. 2007; Heanue and Pachnis 2007).

Among these identified genes, *RET* has been proved to be a major genetic risk factor. Mutations in the other ten genes are rare and found primarily in syndromic HSCR cases. However, for the more common sporadic form of HSCR, *RET* coding mutations have been found in no more than 20% of patients, whereas non-coding mutations in *RET* are suggested to impart susceptibility in other HSCR cases (Brooks et al. 2005; Emison et al. 2005; Griseri et al. 2007).

Since the ‘haplotype’ in the 5’-region of the *RET* locus was identified, many studies have focused on the common variants in this halotype (Fitze et al. 2003; Burzynski et al. 2004; Garcia-Barcelo et al. 2005; Pelet et al. 2005). These findings indicate the possible presence of a mutation(s) in the *RET* promoter or in its regulatory sequences. A number of studies have reported that predisposing haplotypes for HSCR are characterized by the presence of the A variant of SNP2 (Fitze et al. 1999; Garcia-Barcelo et al. 2003), suggesting that the A C A combination of the two promoter and the SNP2 alleles may represent the core haplotype associated with HSCR, acting as a modifying risk allele in the development of the condition.

To refine the mapping of such alleles and characterize their genetic actions, we investigated a sample of 125 sporadic Han Chinese HSCR cases using SNP analysis across the 5’-genomic domain of the *RET* locus, from the 5’-UTR to exon 2. Our results represent the first experimental evidence indicating that all of the seven non-coding SNPs are present in Chinese populations. Moreover, all seven markers are strongly associated with the disease, with significant differences in the frequencies of particular

Table 4 Association analysis of eight SNPs with HSCR using logistic regression

dbSNP ID	Inheritance model	Genotype	Cases	Controls	OR (95% CI)	<i>P</i>	AIC
rs741763	Recessive	C/C-C/G	29 (23.2%)	81 (54.7%)	1	<0.0001	351.7
		G/G	96 (76.8%)	67 (45.3%)	4.00 (2.36–6.78)		
rs2435365	Log-additive	–	–	–	2.33 (1.51–3.58)	1.00E-04	355.7
rs1864410	Co-dominant	C/C	13 (10.4%)	48 (32.6%)	1	<0.0001	325.7
		A/C	34 (27.2%)	70 (47.6%)	1.79 (0.86–3.75)		
		A/A	78 (62.4%)	29 (19.7%)	9.93 (4.71–20.95)		
rs2435364	Recessive	C/C-C/T	33 (26.4%)	90 (60.8%)	1	<0.0001	347.2
		T/T	92 (73.6%)	58 (39.2%)	4.33 (2.58–7.25)		
rs2435362	Co-dominant	C/C	11 (8.8%)	41 (28.1%)	1	<0.0001	330.8
		A/C	34 (27.2%)	71 (48.6%)	1.78 (0.82–3.90)		
		A/A	80 (64%)	34 (23.3%)	8.77 (4.03–19.08)		
rs752975	Recessive	G/G-A/G	36 (28.8%)	97 (66%)	1	<0.0001	340.9
		A/A	89 (71.2%)	50 (34%)	4.80 (2.86–8.04)		
rs2505535	Co-dominant	T/T	70 (56%)	29 (20.1%)	1	<0.0001	337.2
		C/T	42 (33.6%)	74 (51.4%)	0.24 (0.13–0.42)		
		C/C	13 (10.4%)	41 (28.5%)	0.13 (0.06–0.28)		
rs1800858	Log-additive	–	–	–	3.06 (2.07–4.53)	<0.0001	324.7

AIC, Akaike's information criterion

alleles in patients versus controls. Although our results are also similar to those obtained from a Caucasian population (Emison et al. 2005), they additionally indicate that the C allele of rs2505535 may play a protective role in the pathogenesis of HSCR in our Chinese population in comparison to that in the Caucasian population (Emison et al. 2005). It would appear that the T allele of rs2505535 acts as the associated allele for HSCR in our Chinese population.

Since these SNPs were first reported in a Chinese population, it is essential to calculate the inheritance models by logistic regression analysis and investigate the associations of *RET* SNPs with the HSCR phenotype for further study of HSCR in the Chinese population. We found that except for the C allele of rs2505535, there was no significant difference between S-HSCR and L-HSCR patients in terms of SNPs, suggesting that the non-coding region variants have little association with the severity of the disease. One of the more interesting findings of our study is the special role of rs2505535 in the pathogenesis of the Chinese HSCR population, which differs from previous studies. By checking allele frequency data for rs2505535 on HAPMAP, we found that the allele frequencies are rather different between different ethnic groups, with the G allele frequencies being 0.97, 0.8, 0.456 among African–American, Caucasian, and Chinese, respectively. These results imply that even if the SNP is associated with genetic factors involved in the development of HSCR, the latter may be different and variable in different populations. To date there has been little information reported on the relationship between the

non-coding variants and the phenotype of disease. Whether this specific genetic factor is involved in the severe HSCR phenotype in Chinese population requires further investigation.

The vast majority of informative sporadic HSCR cases have been found to be homozygous for the predisposing *RET* 5'-haplotype. A large proportion of our cases proved to have homozygous genotypes, whereas these homozygous genotypes appeared at a comparatively low frequency in the controls. This is especially the case for rs752975, where the A/A genotype revealed a highly increased risk for the development of HSCR (OR > 20), and for rs1864410, where the OR for the A/A genotype was 9.6. This high risk may be related to features of their structures: both are located in conserved regions that contain binding sites for relevant transcription factors. Interestingly, rs1864410 is recognized by ubiquitous transcription factors, such as *SPI-erk1(1)* and *AP2/SPI* (Burzynski et al. 2005). These data suggest that rs1864410 and rs752975 may function as candidate disease-associated variants.

Unexpectedly, rs1800858 was not in LD with the seven other non-coding SNPs in our population. This result deviates from that found in previous studies (Burzynski et al. 2004). There are little data on these seven non-coding SNPs; however, they may account, in part, for the ethnic-dependant prevalence of the disease. Future studies will be required to clarify this point.

Predisposing haplotypes for HSCR have been extensively studied. Burzynski et al. (2004) reported haplotypes

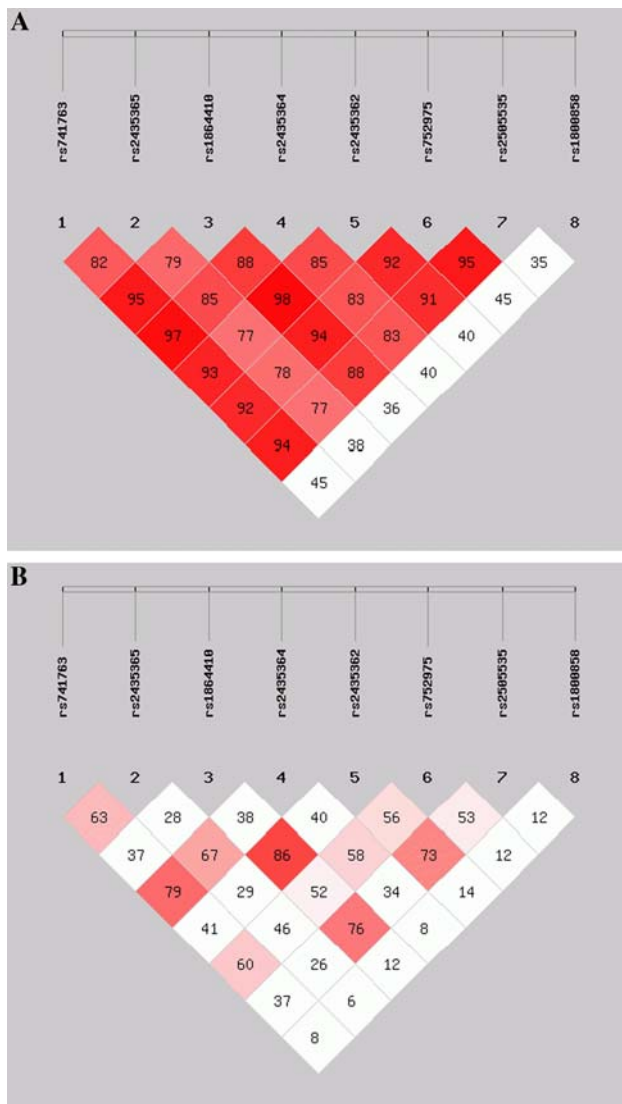


Fig. 2 Linkage disequilibrium (LD) analysis of SNPs in the *RET* region. The white bar above the SNP names represents the relative location of each SNP along the gene. Below the SNPs are haplotype blocks estimated by SHEsis software. **a** The number at the intersection of each pair of SNPs represents the pairwise D' values between two SNPs. **b** The number at the intersection of each pair of SNPs represents the pairwise r^2 values between two SNPs

of 13 DNA markers within and flanking *RET* in a European population. These researcher found six markers (rs741763, SNP-5, SNP-1, rs2435362, rs2565206, and rs1800858) in the 5'-region of *RET* that were strongly associated with the disease. The largest distortions in allele transmission were found at the same markers, suggesting that the 5'-UTR of *RET* may play a key role in HSCR, even when no *RET* mutation was found (Burzynski et al. 2004).

In a subsequent study, these same researchers identified 84 sequence differences by sequencing the haplotype region in a patient and in a control homozygous for the risk

and the non-risk haplotypes, respectively. They found that only one region, located in intron 1, was conserved among all species (including avians) and that only one SNP, rs2506004, was present in this region (Burzynski et al. 2005). However, they neither tested larger samples nor performed a functional analysis. Fortunately, Emison et al. recently genotyped 29 SNPs in 126 families of an American population; these researchers found that 12 of these were localized in the 5'-UTR region and in intron 1 of *RET*. They also showed that two SNPs (rs741763 and rs2505997) in the 5'-UTR region, six SNPs (rs2435365, rs2435364, rs2435362, rs2435357, rs752975, and rs2505535) in intron 1, and one SNP (rs1800858) in the *RET* protein-coding region have the greatest statistical significance and the largest transmission distortions. The highest transmission frequency was observed for one of the alleles from marker rs2435357 (also known as the enhancer mutation), located in intron 1 (Emison et al. 2005). The frequency of the disease-associated haplotypes in the consortium controls corresponded to the frequency of the rs2435357 'T' allele (approx. 24%) in the world population, estimated on the basis of HAPMAP data. A much higher frequency for the longer associated haplotype was observed in Chinese patients (72–76%) (Emison et al. 2005; Zhang et al. 2007).

In our study, we also evaluated the *RET* 'risk haplotypes' present in a Han Chinese population in the 5'-UTR of *RET*. Interestingly, there was one haplotype in the 5'-region of *RET*—G A A T A A T A—that was over-represented in Han Chinese patients with HSCR. If this haplotype is functional, 59.6% of the Han Chinese patients could share the same genetic aetiology. Our results indicate that our identification of the haplotype in the Han Chinese population was in accordance with the 'risk haplotype' reported in the Caucasian population, but with different markers. In addition, our results failed to show a relatively higher frequency in the 'haplotype' reported by Garcia-Barcelo et al. (2005). This difference may be due to the different markers we used.

In conclusion, we observed a very strong association for eight SNPs in the enhancer and 5'-region of *RET*. An increased risk of HSCR was most evident for patients homozygous for the associated alleles. The haplotype consisting of these markers showed similar results, indicating that a strong founder effect is present in our population. The alleles and, consequently, the haplotype, found in our study are similar to those found by others who analyzed the Caucasian population. Our results lead us to conclude that non-coding mutations in *RET* play important roles in the development of HSCR. The unknown functional disease variant(s) with a dosage-dependent effect in HSCR is likely located between the enhancer region and exon 2 of the *RET* gene.

Table 5 Haplotype frequencies of eight SNPs

Haplotype	Allele at marker	rs741763	rs2435365	rs1864410	rs2435364	rs2435362	rs752975	rs2505535	rs1800858	Patients/cases N (Freq)	Controls N (Freq)	P value	ORs (95% CI)
A	C	A	C	C	C	G	C	A	0.00 (0.000)	2.71 (0.010)	0.023	–	
B	C	G	C	C	A	G	C	G	0.00 (0.000)	3.74 (0.014)	0.071	–	
C	C	G	C	C	C	G	C	A	5.57 (0.024)	27.76 (0.103)	0.000	0.211 (0.083–0.536)	
D	C	G	C	C	C	G	C	G	20.26 (0.087)	37.12 (0.137)	0.072	0.592 (0.333–1.053)	
E	G	A	A	C	A	A	T	G	0.00 (0.000)	3.41 (0.013)	0.085	–	
F	G	A	A	T	A	A	C	A	8.27 (0.036)	3.99 (0.015)	0.135	2.446 (0.731–8.190)	
G	G	A	A	T	A	A	T	A	138.17 (0.596)	48.87 (0.181)	0.000	7.214 (4.754–10.946)	
H	G	A	A	T	A	A	T	G	17.22 (0.074)	59.58 (0.221)	0.000	0.276 (0.156–0.490)	
I	G	G	A	T	A	A	T	A	4.09 (0.018)	0.00 (0.000)	0.047	1622.365 (87.787–29982.514)	
J	G	G	A	T	A	A	T	G	0.51 (0.002)	3.07 (0.011)	0.293	0.189 (0.021–1.740)	
K	G	A	C	T	A	A	C	A	2.00 (0.009)	3.04 (0.011)	0.761	0.757 (0.126–4.554)	
L	G	A	C	T	C	A	C	G	11.39 (0.049)	8.56 (0.032)	0.329	1.565 (0.633–3.873)	
M	G	A	C	T	C	A	T	A	0.00 (0.000)	3.27 (0.012)	0.091	–	
N	G	A	C	T	C	G	C	A	2.33 (0.010)	7.41 (0.027)	0.156	0.357 (0.081–1.573)	
O	G	A	C	T	C	G	C	G	5.82 (0.025)	15.51 (0.057)	0.07	0.418 (0.158–1.102)	
P	G	A	C	T	C	A	C	A	1.80 (0.008)	17.14 (0.063)	0.001	0.114 (0.024–0.536)	
Q	G	A	C	T	C	A	T	G	0.00 (0.000)	16.53 (0.064)	0.000	0.000 (0.000–0.001)	

N, number of haplotypes; Freq, frequency

Acknowledgments This work was supported by National Found for Fostering Talents of Basic Science (No.J0730856) and the Medical Science Foundation of Zhejiang Province (No. 2002A027).

References

- Amiel J, Sproat-Emison E, Garcia-Barcelo M, Lantieri F, Burzynski G, Borrego S, Pelet A, Arnold S, Miao X et al (2008) Hirschsprung disease, associated syndromes and genetics: a review. *J Med Gene* 45:1–14
- Angrist M, Bolk S, Thiel B, Puffenberger EG, Hofstra RMW, Buys CHCM, Cass DT, Chakravarti A (1995) Mutation analysis of the RET receptor tyrosine kinase in Hirschsprung disease. *Hum Mol Genet* 4:821–830
- Bolk S, Pelet A, Hofstra RMW, Angrist M, Salomon R, Croaker D, Buys CHCM, Lyonnet S, Chakravarti A (2000) 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* 97:268–273
- Borrego S, Wright F, Fernandez R, Williams N, Lopez-Alonso M, Davuluri R, Antinolo G, Eng CA (2003) 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* 72:88–100
- Brooks AS, Oostra BA, Hofstra RM (2005) Studying the genetics of Hirschsprung's disease: unraveling an oligogenic disorder. *Clin Genet* 67:6–14
- Burzynski GM, Nolte IM, Osinga J, Ceccherini I, Twigt B, Maas S, Brooks A, Verheij J, Plaza Menacho I, Buys CHCM, Hofstra RMW (2004) 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* 12:604–612
- Burzynski GM, Nolte IM, Bronda A, Bos KK, Osinga J, Plaza Menacho I, Twigt B, Maas S (2005) Identifying candidate Hirschsprung disease associated RET variants. *Am J Hum Genet* 76:850–858
- Chakravarti A, Lyonnet S (2002) Hirschsprung disease. In: Scriver CS, Sly WS, Valle D, Beaudet AL (eds) *The metabolic and molecular bases of inherited disease*. McGraw-Hill, New York, pp 6231–6255
- Ederly P, Lyonnet S, Mulligan LM, Pelet A, Dow E, Abel L, Holder S, Nihoul-Fekete C, Ponder BA, Munnich (1994) A mutations of the RET proto-oncogene in Hirschsprung's disease. *Nature* 367:378–380
- Emison ES, McCallion AS, Kashuk CS, Bush RT, Grice E, Lin S, Portnoy ME, Cutler DJ, Green ED, Chakravarti A (2005) A common sex-dependent mutation in a RET enhancer underlies Hirschsprung disease risk. *Nature* 434:857–863
- Fitze G, Schreiber M, Kuhlisch E, Schackert HK, Roesner D (1999) Association of RET proto-oncogene codon 45 polymorphism with Hirschsprung disease. *Am J Hum Genet* 65:1469–1473
- Fitze G, Cramer J, Ziegler A, Schierz M, Schreiber M, Kuhlisch E, Roesner D, Schackert HK (2002) Association between c135G/A genotype and RET proto-oncogene germline mutations and phenotype of Hirschsprung's disease. *Lancet* 359:1200–1205
- Fitze G, Appelt H, Konig I, Gorgens H, Stein U, Walther W, Gossen M, Schreiber M, Ziegler A, Roesner D, Schackert HK (2003) Functional haplotypes of the RET proto-oncogene promoter are associated with Hirschsprung disease (HSCR). *Hum Mol Genet* 12:3207–3214
- Gabriel SB, Salomon R, Pelet A, Angrist M, Amiel J, Fornage M, Attié-Bitach T, Olson JM, Hofstra R, Buys CHCM, Steffann J, Munnich A, Lyonnet S, Chakravarti A (2002) Segregation at three loci explains familial and population risk in Hirschsprung disease. *Nat Genet* 31:89–93
- Garcia-Barcelo MM, Sham MH, Lui VC, Chen BL, Song YQ, Lee WS, Yung SK, Romeo G, Tam PK (2003) Chinese patients with sporadic Hirschsprung's disease are predominantly represented by a single RET haplotype. *J Med Genet* 40:e122
- Garcia-Barcelo M, Ganster RW, Lui VC, Leon TY, So MT, Lau AM, Fu M, Sham MH, Knight J, Zannini MS, Sham PC, Tam PK (2005) TTF-1 and RET promoter SNPs: regulation of RET transcription in Hirschsprung's disease. *Hum Mol Genet* 14:191–204
- Garcia-Barcelo MM, Lau DK, Ngan ES, Leon TY, Liu TT, So MT, Miao XP, Lui VC, Wong KK, Ganster RW, Cass DT, Croaker GD, Tam PK (2007) Evaluation of the thyroid transcription factor-1 gene (TTF1) as a Hirschsprung's disease locus. *Ann Hum Genet* 71:746–754
- Griseri P, Bachetti T, Di Duca M, Ravazzolo R, Ceccheini I (2007) A common variant located in the 3' UTR of the RET gene is associated with protection from Hirschsprung disease. *Hum Mutat* 28:168–176
- Heanue TA, Pachnis V (2007) Enteric nervous system development and Hirschsprung's disease: advances in genetic and stem cell studies. *Nat Rev Neurosci* 8:466–479
- Iwashita T, Kurokawa K, Qiao S, Murakami H, Asai N, Kawai K, Hashimoto M, Watanabe T, Ichihara M, Takahashi M (2001) Functional analysis of RET with Hirschsprung mutations affecting its kinase domain. *Gastroenterology* 121:24–33
- Parisi MA, Kapur RP (2000) Genetics of Hirschsprung disease. *Curr Opin Pediatr* 12:610–617
- Pelet A, de Pontual L, Clement-Ziza M, Salomon R, Mugnier C, Matsuda F, Lathrop M, Munnich A, Feingold J, Lyonnet S, Abel L, Amiel J (2005) Homozygosity for a frequent and weakly penetrant predisposing allele at the RET locus in sporadic Hirschsprung disease. *J Med Genet* 42:e18
- Shi YY, He L (2005) SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci. *Cell Res* 15:97–98
- Sole X, Guino E, Valls J, Iñiesta R, Moreno V (2006) SNPSTATS: a web tool for the analysis of association studies. *Bioinformatics* 22:1928–1929
- Xiao ZX, Xiao JX, Jiang YX, Zhang SY, Yu MH, Zhao JL, Wei DZ, Gao HM (2006) A novel method based on ligase detection reaction for low abundant YIDD mutants detection in hepatitis B virus. *Hepatol Res* 34:150–155
- Zhang XN, Zhou MN, Qiu YQ, Ding SP, Qi M, Li JC (2007) Genetic analysis of RET, EDNRB, and EDN3 genes and three SNPs in MCS + 9.7 in Chinese patients with isolated Hirschsprung disease. *Biochem Genet* 45:523–527