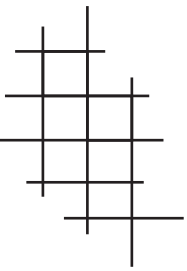
# Genetics dissection and functional studies in Hirschsprung disease

Sribudiani, Yunia

# Genetic Dissection and Functional Studies in Hirschsprung Disease

Yunia Sribudiani



Genetic Dissection and Functional Studies in Hirschsprung Disease
Y. Sribudiani

PhD Thesis – Department of Genetics
University Medical Center Groningen, University of Groningen
Groningen, the Netherlands

This study was financially supported by University of Groningen and
Research Institute GUIDE.

ISBN/EAN: 978-94-6182-039-6

Cover Ilustration: Ella Elviana, www.ellasworks.blogspot.com
Cover & book layout: Briyan B Hendro, www.brianos.daportfolio.com
Printed by: Off Page, Amsterdam, www.offpage.nl

Printing of this thesis was financially supported by: University of Groningen,
GUIDE and Department of Genetics - UMCG.

RIJKSUNIVERSITEIT GRONINGEN

# Genetic Dissection and Functional Studies in Hirschsprung Disease

**Proefschrift**

ter verkrijging van het doctoraat in de
Medische Wetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. E. Sterken,
in het openbaar te verdedigen op
woensdag 21 december 2011
om 14.30 uur

door

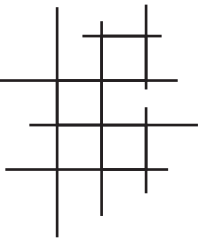
**Yunia Sribudiani**

geboren op 28 juni 1977
te Bandung, Indonesië

| | | |
|---|---|---|
| Promotor | : | Prof. dr. R.M.W. Hofstra |
| Beoordelingscommissie | : | Prof. dr. I. Ceccherini<br>Prof. dr. U.L.M. Eisel<br>Prof. dr. H.W.G.M. Boddeke |

# CONTENTS

# 1 General Introduction



# 1. Hirschsprung Disease: Diagnosis and Treatment

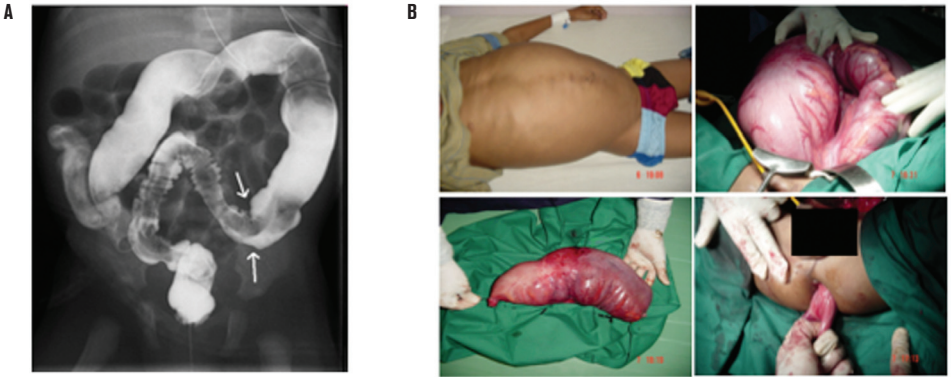
## 1.1 Diagnosis

Hirschsprung (HSCR) disease, also known as congenital megacolon, is characterized by the absence of enteric ganglia (aganglionosis) in the submucosal and myenteric plexuses of the gastrointestinal (GI) tract leading to a life-threatening bowel distention which presents itself usually shortly after birth.[1] HSCR disease was first recognized in 1887 by the Danish pediatrician Harald Hirschsprung. He described two boys who died at age 8 and 11 years with severe constipation. HSCR is clinically characterized by a failure to pass meconium (the first stool) within 48 hours after birth, abdominal distention, vomiting or neonatal enterocoilitis.[2] In particular, enterocolitis can prove life-threatening in HSCR. The mortality rate of HSCR disease due to enterocolitis ranges from 0-30%. The variation is most likely due to differences in diagnosing Hirschsprung-associated enterocolitis (HAEC).[3,4] The gold standard for diagnosing HSCR disease is the histopathology of a full-thickness rectal biopsy in which the neurons and glial cells (ganglia) are absent in both the myenteric and submucosal plexus of the GI tract of HSCR patients.[5] Measurement of the pressures in the rectum (rectal manometry) and X-rays made with a barium enema are methods that can provide extra information to the physician for an accurate diagnosis. The dilated (mega)colon is proximal to the aganglionic region, which narrows towards the distal part and an empty rectum is a common finding.[1] In 75% of cases, the diagnoses is made within six weeks of birth and in 90% of cases, the disease is diagnosed before the age of five years.[6]

The length of the aganglionic region varies and based on this HSCR disease is classified into short-segment HSCR (S-HSCR), long-segment HSCR (L-HSCR) and total colonic/intestinal aganglionosis (TCA/TIA). S-HSCR is the most common type (80% of cases) and the length of the aganglionic region does not extend beyond the upper sigmoid. Fifteen percent of the patients have L-HSCR in which the aganglionic region extends beyond the upper sigmoid. In about 5% of cases, the TCA/TIA is diagnosed. In these patients the complete colon and the terminal part of the ileum show an absence of ganglia and in rare cases even the whole bowel can lack ganglia, which is a lethal condition.[1, 7-9]

HSCR is the most common developmental disorder of the enteric nervous system (ENS) with a prevalence of 1 in 5000 newborns, although variation of the prevalence is observed between ethnic populations.[10] Interestingly, a sex-dependent penetrance is observed in S-HSCR with a male:female ratio of 4:1. With respect to these findings, it is interesting to report that the transmission of disease-causing alleles in this group is also disturbed, as 78% of the commonly shared *RET* allele, the major risk factor for HSCR development, are maternally de-

rived.[11, 12] Whether the male:female ratio is caused by these differences in the transmission of associated alleles is not yet known. HSCR most commonly presents sporadically, although in approximately 20% of cases, several members in a family are affected (familial HSCR). In 70% of cases, HSCR occurs as an isolated trait and in the other 30% HSCR is associated with other congenital malformation or syndromes.




http://radiopaedia.org/images/135545

http://praanadah-hospital.com/procedures/hirschsprung_disease

**Figure 1 -** A) Barium enema X-ray of HSCR patient's bowel, the arrows shows the beginning of the aganglionic region. B) Treatment of HSCR patient with pull-through surgery to remove the aganglionic region and rejoin the healthy ganglionated colon to the anus.

## 1.2 Treatment

Current treatment of HSCR is surgical resection of the aganglionic part of the gastrointestinal (GI) tract. These operations, so called pull-through operations, are mostly based on the procedures developed by Swenson, Duhamel or Soave. In this surgical procedure the affected region of the colon is removed and the healthy ganglionated region of the colon is rejoined with the anus.[13-15] A one-stage procedure (one operation) is possible when the diagnosis of HSCR is made before the colon dilates. Otherwise a primary colostomy is temporarily required before the patient has the second pull-through surgery to remove the aganglionic region. Although the surgical procedure relieves the patient's constipation and has proven life-saving, the long-term outcome remains rather poor, irrespective of the length of the aganglionic region. A retrospective study performed by Catto-Smith and colleagues showed that fecal incontinence is very common and episodic constipations have also been reported.[16]

These clinical findings have led researchers to think about developing alternative treatments for HSCR disease. One method that seems promising is ENS replacement therapy.

Metzger and colleagues showed that enteric nervous system stem cells (ENSSCs) can be isolated from postnatal human gut mucosal biopsy and used to generate neurosphere-like bodies (NLBs). More importantly, they showed that these NLBs have the same stem cell plasticity as those generated from embryonic human gut tissue and, when grafted onto aganglionic chick hindgut, these NLBs have the ability to migrate and differentiate into glial cells and different types of ENS neurons. Whether these grafted cells function correctly and whether they repair the ENS system needs further investigation. These preliminary results show, however, that ENSSCs transplantation might have therapeutic potential and could, in the future, be considered as an additional therapy for HSCR disease.[17]

# 2. Pathogenesis of Hirschsprung Disease: Genetic Background

## 2.1 Genes involved in the development of HSCR disease

The genetic dissection of HSCR has led to the identification of 12 genes and five susceptibility loci implicated in the development of HSCR disease. These 12 genes are expressed either in the enteric neural crest cells (ENCCs) or in other cells in the gut along the path of ENCCs migration during ENS development. The inheritance of HSCR is considered to be complex. The disease can be transmitted as a dominant trait or as a recessive trait, but in the majority of cases it is probably polygenic with differences in sex ratio, with a male predominance in S-HSCR (4:1), incomplete penetrance and variable expression; associations with a large number of syndromes and congenital malformations have been observed.[1] Linkage analyses of multiplex HSCR families revealed that the *RET* (REarrange during Transfection) gene, located at 10q11.2, is the major risk factor for Hirschsprung disease as almost all families showed linkage with *RET*.[11, 18, 19] Coding sequence mutations in *RET* are responsible for a dominant form of HSCR (with incomplete penetrance) and coding and splice site mutations have been identified in up to 50% of familial cases and 15-35% of sporadic cases.[20] The mutations are scattered throughout the *RET* coding sequence, including large and micro-deletions and a variety of point mutations. *RET* mutations associated with HSCR are believed to cause a loss of function (haploinsufficiency).[21-23] However, *RET* mutations on their own might not result in aganglionosis of part of the GI tract, as the penetrance of the *RET* mutations (in general) is 72% in males and 51% in females. Yet *RET* is believed to be the major gene underlying HSCR, as evidenced primarily in families enriched for L-HSCR.[11] Mutations in the rest of the 11 genes *EDNRB* [24], *EDN3* [25, 26], *GDNF* [27, 28], *NTN* [29], *SOX10* [30], *PHOX2B* [31], *ECE1* [32], *KIAA1279/KBP* [33], *ZFHX1B* [34, 35], *TTF-1* [36] and *NRG1* [37] do not exceed 20% of the cases, supporting a genetic

heterogeneity for this disorder. Finally, rare heterozygous germline mutations of *EDNRB* in combination with *RET* germline mutations have also been detected in patients showing the HSCR phenotype.[38] This supports the concept of synergistic heterozygosity for HSCR, i.e. that the disease phenotype could be the result of the cumulative effect of at least two mutations in different genes.

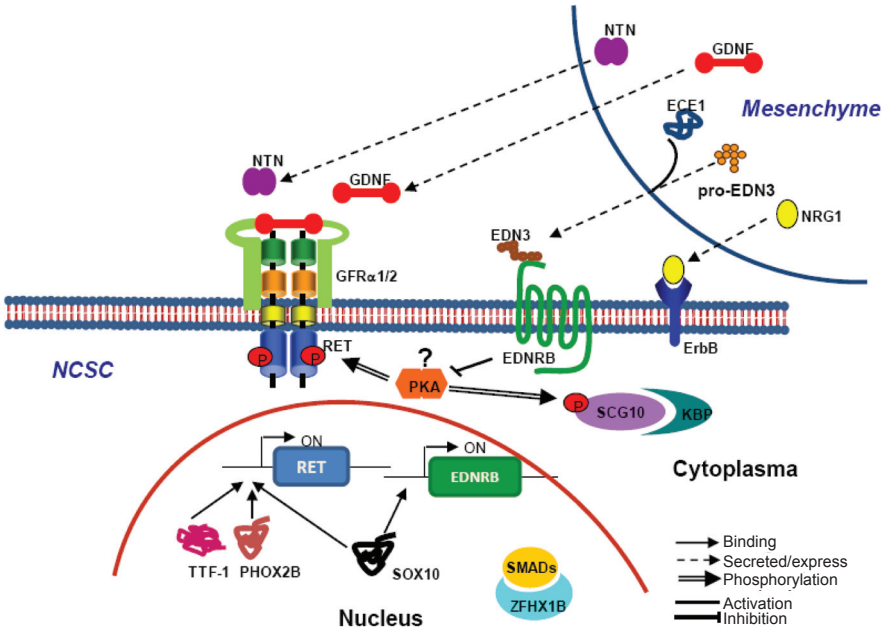



**Figure 2 -** Schematic representation of the encoded proteins known to be associated with HSCR development and showing their possible interactions.

Mutations in most of the 11 genes are found mostly in the syndromic cases where HSCR is associated with other congenital malformation (see Table 1). Mutations in *EDNRB*, *EDN3* and *SOX10* were identified in a patient with Shah-Waardenburg syndrome (WS4), which is characterized by congenital hearing loss, pigmentary abnormalities of the hair, skin and eyes, and HSCR disease.[30, 39] Mutations in *PHOX2B* have been identified in patients with congenital central hypoventilation (CCHS) and HSCR disease. This syndrome is also known as Haddad syndrome. CCHS is a rare disorder characterized by impairment of autonomic control of spontaneous respiration in the absence of other lung or cardiac disease.[31] Mutations in *ZFHX1B* have been identified in patients with Mowat-Wilson syndrome, which is an autosomal dominant disorder characterized by mental retardation, epilepsy, delayed motor development, and HSCR disease.[40] Mutations in *KIAA1279* (now called kinesin-binding protein or

*KBP*) have been identified in patients with Goldberg-Shprintzen (GOSH) syndrome, a rare autosomal recessive disorder characterized by HSCR, microcephaly, mental retardation, and polymicrogria.[33] A mutation in *ECE1* was identified in a single patient with craniofacial and cardiac defects.[32]

Besides mutations in these 11 genes, chromosomal abnormalities are frequently observed and found in 12% of all syndromic HSCR cases. Trisomy chromosome 21 (Down Syndrome (DS)) is the most frequent chromosomal abnormality (>90% cases) associated with HSCR disease. On the other hand, 77% of DS patients have associated gastrointestinal abnormalities, of which 2-15% can be explained by HSCR disease.[41] This might implicate chromosome 21 in the etiology of HSCR, so it is not surprising that DS patients have a 40-fold higher chance of getting HSCR compared to the normal population.[42] The other possible explanation could be that the genetic imbalance disturbs the homeostasis of genes expressed during development that eventually contribute to the disease phenotype.[43] Interestingly, besides these possibilities, a variant in the *RET* gene also seems to be related to the HSCR phenotype in DS patients. It was shown that a common single nucleotide polymorphism (SNP), rs2435357 (C>T), located in the conserved region of intron 1 of the *RET* gene, when homozygous for the T-allele, gives DS patients a 5.3-fold increased risk of developing HSCR compared to DS patients who are homozygous for the wild-type C-allele.[44] Furthermore, the frequencies of the T-allele differ significantly between individuals with DS alone (0.26±0.04), HSCR alone (0.61±0.04) and those with HSCR and DS (0.41±0.04), showing there is association and genetic interaction between RET and the chromosome 21 gene dosage.[42]

**Table 1 -** HSCR associated genes and their clinical features (Modified from Brooks A et al. 2005).

| Gene Symbol | Position | Inheritance | Phenotype |
|---|---|---|---|
| RET | 10q11.2 | Dominant, incomplete penetrance | Non-syndromicMEN2A |
| GDNF | 5p13 | Non-Mendelian | Non-syndromic |
| NTN | 19p13 | Non-Mendelian | Non-syndromic |
| EDNRB | 13q22 | Recessive<br>Dominant (de novo in 80%) | Shah-Waardenburg<br>Non-Syndromic |
| EDN3 | 20q13 | Recessive<br>Dominant, incomplete penetrance | Shah-Waardenburg<br>Non-syndromic |
| PHOX2B | 4p12 | Dominant (de novo in 90%) | Haddad Syndrome (CCHS) |
| SOX10 | 22q13 | Dominant (de novo in 75%) | Shah-Waardenburg |
| ECE1 | 1p36 | Dominant (de novo) | Congenital heart formation |
| ZFHX1B (SIP1) | 2q22 | Dominant (de novo) | Mowat-Wilson |
| KIAA1279 (KBP) | 10q22.1 | Recessive | Goldberg-Shprintzen |
| TTF1 (TITF1) | 14q13 | - | Non-syndromic |
| NRG1 | 8p21 | - | Non-syndromic |

## 2.2 Non-coding RET variants

As mentioned above, *RET* coding mutations have been identified in 50% of the familial cases. However, regardless of the *RET* coding mutation status, almost all familial cases are linked to the RET locus.[11, 45] This suggests that non-coding *RET* mutations must play a major role in the remaining familial cases. This idea was corroborated in association studies on sporadic (simplex) cases with and without *RET* coding mutations, performed on several Caucasian populations and an Asian population. These studies revealed a strong association between a certain haplotype (covering 27 kb in total) and the disease. This haplotype starts 4 kb upstream of the *RET* transcription start site, going all the way to the beginning of exon 2. This haplotype is present in 56%-62% of patients, but in only 20% of the controls tested in the Caucasian population. In the Chinese patient population, the frequency of this same haplotype was 85% but 40% in the controls. This last finding might partially explain the higher incidence of HSCR in Asians compared to Caucasians. Several groups have focused their studies on fine-mapping the associated region and tried to verify the location of the causative variant.[12, 36, 46-48]

### SNP -1(G>A) and -5 (A>C)

Two variants located in the *RET* promoter, SNP rs10900296 (G>A) and rs10900297 (A>C) located at -5 and -1 from the *RET* transcription start site respectively, were selected as probably causative. It was shown that the HSCR associated haplotype (AC) decreased the *RET* promoter activity in luciferase assays.[46] Extended functional studies performed by Garcia-Barcelo and colleagues showed that these two variants and their surrounding sequences were the binding site of TTF-1 (Thyroid Transcription Factor 1). The AC-haplotype reduces the binding affinity of TTF-1, hence reducing *RET* promoter activity.[36] However, this effect proved to be cell-line-dependent and it is therefore difficult to conclude whether these two variants really are the causative ones. Another aspect that makes the interpretation difficult is the fact that these two variants are in a region with strong linkage disequilibrium (LD), which might mask the real functional variant(s).[49]

### SNP rs2435357 (C>T)

A similar study was performed by Emison and colleagues. They genotyped 28 SNPs spanning a region of 175 kb in 126 trios. The region covers *RET* and two other downstream genes, *GALNACT-2* and *RASGEF1A*. They identified HSCR-associated SNPS in *RET* (13 out of 17), *GALNACT-2* (3 out of 7) and *RASGEF1A* (2 out of 4). By multi-species alignment, they identified that one of the associated SNPs in *RET*, rs2435357 (C>T), is located in a highly con-

served region called MSC+9.7 (Multi Species Conserved). This variant is not only conserved among nine different mammalian species, but also showed the largest transmission distortion (τ: 0.8) in transmission disequilibrium test (TDT) analysis. This data led the authors to postulate that SNP rs2435357 is most likely a disease-causing variant. They supported this idea with reporter (luciferase) assays and showed that, indeed, the MCS+9.7 region, containing the disease-associated variant (T-allele), reduced promoter activity when compared to the wild-type sequence (C-allele).[12] Additional functional studies on SNP rs2435357 by the same group proved that SOX10, a transcription factor, binds to a sequence motif containing this specific variant.[50] In addition *in vivo* functional studies on MCS+9.7 were performed by Grice and colleagues. They generated transgenic mice carrying this conserved region coupled to the *LacZ* reporter gene and studied the spatial temporal LacZ expression during embryonic development. They showed that LacZ expression, under the regulation of MSC+9.7, has a similar spatial temporal expression as the *RET* gene, indicating this region contributes to *RET* expression regulation.[51]

### SNP rs2506004 (C>A)

Burzynski and colleagues sequenced the shared common haplotype region from 10 kb upstream of the *RET* gene through intron 1 (a total of 33 kb) in a patient homozygous for the common risk haplotype and in a control individual homozygous for the most common non-risk haplotype. They identified eight variants, located in highly conserved regions, within putative transcription binding sites. Among these variants, SNP rs2506004 (C>A) proved to be conserved not only in vertebrates but also in an avian species (chicken). They hypothesized that the region containing the SNP might be an enhancer for *RET* expression and that this SNP would result in a reduction of *RET* expression.[48] The variant is also located in the MSC+9.7 region, 217 bp downstream of SNP rs2435357 (C>T). These two SNPs are in strong linkage disequilibrium ($r^2$=0.98). Recently we identified NXF/ARNT2 and SIM2/ARNT2 as the transcription factors and repressors, respectively, for the wild-type C-allele of SNP rs2506004. Moreover, transfection of *Nxf/Arnt2* and *Sim2/Arnt2* constructs into a mouse neuroblastoma cell line (Neuro2A) enhanced and reduced the RET endogenous expression, respectively, proving that NXF/ARNT2 and SIM2/ARNT2 can indeed regulate RET expression endogenously (see Chapter 4).[52]

Interestingly, the gene encoding one of the above transcription factors, SIM2, is located in the Down syndrome (DS) critical region on chromosome 21. It has been shown that overexpression of *SIM2* reduces the expression of *Drebrin* (*DBN1*), encoding a protein that plays an important role in synaptic plasticity and hence suggested to play a role in the development of DS.[53,

[54] The overrepresented of *SIM2* in DS patients, due to the trisomy of chromosome 21, might lead to *SIM2* overexpression and, as a consequence, not only down-regulate *DBN1* but also down-regulate *RET* expression. This might then contribute to the development of HSCR and partly explain why DS patients have a 40-fold higher risk of HSCR disease compared to the normal population (Chapter 4).[52] Together these data suggest that these four variants, all identified in the *RET* non-coding sequences, may play important roles in HSCR disease development.

## 2.3 Susceptible HSCR loci

As mentioned above, *RET* coding mutations are the major risk factor for HSCR development. However, a *RET* mutation alone may not be sufficient to cause HSCR as the penetrance of mutant alleles, in general, is incomplete in HSCR patients (72% in males and 51% in females). Non-coding mutations, as described above, also contribution to the phenotype, especially in S-HSCR. Furthermore, mutations in other 11 HSCR genes can only explain a small proportion of the cases. In the majority of sporadic and simplex S-HSCR cases, it is therefore likely that mutations occur in the coding regions of as yet unknown genes.

Several studies have been searching for additional susceptibility loci, with or without them being in conjunction with a *RET* mutation. Bolk and colleagues (2000) performed a linkage analysis in 12 multiplex HSCR families. They split them into two groups: the first consisted of families that were linked to 10q1.2 and all had severe *RET* mutations (six families), while the second group consisted of five families that were also linked to the *RET* region but no clear *RET*-coding sequence mutation had been identified, and one family that was not linked to the *RET* region (and obviously had no *RET* mutation). Bolk and colleagues segregated a new locus at 9q31 in the second group (the six families without *RET* coding mutations). It was hypothesized that 9q31 is probably a modifier locus for development of HSCR disease.[45]

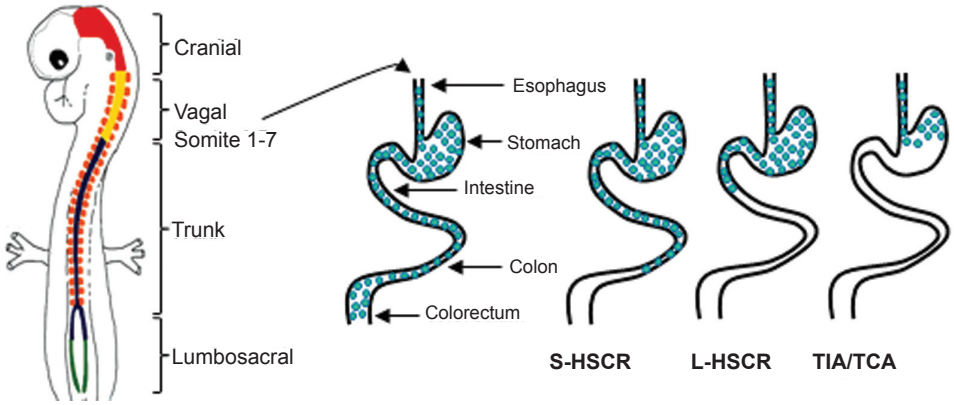
In another study, Gabriel and colleagues (2002) performed a genome-wide scan in 49 nuclear families with S-HSCR. They carried out a sib-pair analysis and found significant allele sharing with markers on 10q11, where *RET* is located ($p = 2.4 \times 10^{-7}$), on 19q12 ($p = 6.6 \times 10^{-4}$) and on 3p21 ($p=1.9 \times 10^{-4}$), with non-parametric LOD scores of 5.02, 3.46 and 3.23, respectively. To validate these data, they checked the families mutation status for *RET*. Surprisingly, they only identified *RET* coding sequence mutations in only 40% of *RET*-linked families, suggesting the importance of non-coding variants. The risk ratios, as determined by the associated risk haplotypes at 10q11, 3p21 and 19q21 were 8.3, 4.2 and 5.0-fold, respectively. Using a multiplicative model, the frequency of individuals being heterozygous for these three loci is close to the observed population incidence. Gabriel et al. therefore hypothesized that all three loci were necessary and possibly even sufficient for the observed occurrence of S-HSCR.[11]

A fourth locus was identified by performing a genome-wide scan on 43 Mennonite trios, all belonging to the same large kindred. They identified three loci, two of which were known loci (13q22.3-q31.1 where the *EDNRB* gene is located and 10q11.21 containing the *RET* gene) and one new locus on 16q23.3. The LOD scores of the loci were 55.60, 5.6 and 3.01, respectively, showing that *EDNRB* is the primary susceptibility factor for the disease phenotype in this kindred.[55]

Studying a large multi-generational Dutch family with an isolated HSCR phenotype resulted in the identification of a new susceptibility locus on 4q31-32 (the fifth locus), with a parametric LOD score of 2.7.[56] The low penetrance of the locus in this family suggests that this mutation is necessary but not sufficient for disease development. In this thesis we describe our search for the disease-causing gene on this locus and for modifiers elsewhere in the genome by exome sequencing performed in this family (see Chapter 3).

# 3. Pathogenesis of Hirschsprung Disease: Developmental background

## 3.1 Enteric nervous system (ENS) development




(Adapted from Dyer, MA, et al., J Clin Pathol 57, 2004)

**Figure 3 -** Schematic representation of the neural tube. The enteric neural crest cells (ENCCs) are formed in the vagal region, in somites 1-7. From there they migrate rosto-caudally to colonize the whole gastrointestinal (GI) tract, later forming the enteric nervous system. Defects in ENCCs migration, proliferation or differentiation could lead to HSCR disease. The different lengths of the aganglionosis region in the GI tract (S-HSCR, L-HSCR and TCA/TIA) define the severity of the disease.

The enteric nervous system (ENS) is a complex system that contains as many neurons as the spinal cord and it functions independently from the central nervous system (CNS). It coordinates normal bowel motility and changes the blood flow and secretion of water and electrolytes. The intrinsic innervation of the gut wall consists of two interconnected ganglionated nerve plexuses: the myenteric plexus, which is located between the longitudinal and circular muscle layers, and the submucous plexus, which is located between the mucosa and the circular muscle layers.[57] Neurons and glial cells of the ENS all derive from neural crest stem cells (NCSCs), which are pluripotent stem cells. They are formed at the time the neural tube closes, at the border of the neural plate and the non-neural ectoderm. They directly migrate towards different structures in the embryo to form distinct (neural crest-derived) tissues. These tissues include the skin's pigment cells, the craniofacial bones and cartilage, smooth muscle cells, peripheral and enteric neurons, and glial cells of the ENS .[58, 59] It is the NCSCs which originate from the vagal region of the hind brain, specifically from somites 1 to 7, that give rise to the vast majority of the parasympathetic neurons and glial cells of the ENS.[60] These precursor cells migrate from the vagal region towards the foregut rostocaudally in the 4$^{th}$ week of gestation in humans, colonize the whole GI tract and reach the end of the hindgut in the 7$^{th}$ week. [57,61, 62] These precursor cells termed enteric neural crest-derived cells (ENCCs) when they enter the foregut.[57] A small part of the ENS precursor cells derive from the sacral region. These cells migrate in the opposite direction from those from the vagal region and move from the end of the hindgut upwards. The sacral ENS precursor cells contribute to ENS development in the gut by colonizing the distal part of the GI tract, however, the real contribution of the sacral neural crest cells in ENS development is still under debate.[63] For instance, a study in chick embryos suggested that sacral neural crest cells differentiate only into glial cells and not into the enteric neurons.[64]

All in all, the colonization of the whole GI tract by ENCCs and their later differentiation into neurons and glial cells is a very complex process. Timing and harmonization of the necessary transcription factors, growth factors, their receptors and the signals they give is essential for each stage of ENS development.[65, 66] Any defect in this process (ENCCs migration, proliferation, differentiation or survival of ENS progenitors) can lead to the development of HSCR.

## 3.2 Signaling pathways and genes involved in ENS development *RET Signaling Pathway*

RET is a transmembrane Tyrosine Kinase Receptor (RTK) that plays an important role in several intracellular signaling cascades that regulate differentiation, proliferation, migration and cell survival. Activation of RET signaling is unique among RTKs as it requires a multiprotein complex instead of the customary direct receptor-ligand interaction. The RET activa-

tion complex comprises a soluble ligand (GDNF, Neurturin, Persephin or Artemin), GFRα1-4 (gylcosyl-phosphatidylinositol membrane-anchored co-receptor of GDNF) and RET itself. A complex containing the ligand and co-receptor heterodimers brings two RET molecules together and triggers auto-phosphorylation of tyrosine residues in the TK-domain of RET (see Figure 2).[67, 68]

RET signaling has been proven critical in ENS development as mice deficient for GDNF, RET and GFRα1 show severe defects in enteric innervations (i.e. aganglionosis).[21, 22, 27] But how RET signaling is involved in ENS development is only partly known. What *is* known is that GDNF, in particular, plays a crucial role, as it serves as a chemo-attractant for the RET-positive vagal ENCCs. GDNF is highly expressed in the stomach ahead of the migrating ENCCs wave-front and its expression is elevated in the caecum when ENCCs migrate towards the distal part of the gut.[66, 69] Besides GDNF, several other ligands can activate the RET signaling pathway, namely Neurturin (NTN), Artemin (ARTN) and Persephin (PSPN) that bind to GFRα2, GFRα3 and GFRα4, respectively. Mouse *Ntn* and *Gfrα2* mutants show ENS defects, revealing their importance in ENS development.[70, 71] Furthermore, HSCR patients carrying *NTN* mutations have also been reported.[29]

Even though the RET signaling pathway is known to play a pivotal role in many processes during ENS development, its downstream effectors, the proteins that transmit signals within the cells, are still largely unknown. In Chapter 5 we describe expression array studies performed on mouse ENCCs to gain more insight into the RET signaling pathways on GDNF stimulation.

### *EDNRB-EDN3 Signaling Pathway*

A second pathway involved in HSCR is the Endothelin Receptor type B (EDNRB) pathway. This belongs to the guanine nucleotide binding protein (G-protein-coupled) receptor family, which is activated by its ligand Endothelin 3 (EDN3). END3-EDNRB signaling is required for the development of melanocytes and enteric neurons.[72] Upstream of EDN3, ECE1 (Endothelin Converting Enzyme), the enzyme that converts the inactive precursor of END3 into the active EDN3 peptide, thus activates the EDNRB-EDN3 signaling pathway (see Figure 2).

In mouse embryonic gut, *Ednrb* is expressed in the migrating ENCCs and *Edn3* is mainly expressed in the midgut and the hindgut mesoderm during the early phases of ENCCs migration, and at high levels in the caecum and the proximal colon when the ENCCs colonize the terminal gut region.[73, 74] This pattern of expression suggests that, as for GDNF-RET signaling, EDN3-EDNRB signaling plays an important role in ENCCs migration. Furthermore, it also suggests that EDN3-EDNRB plays a role in cell proliferation as the *Edn3* mutant mouse has

fewer neural crest stem cells compared to the wild-type mouse.[73] Moreover, cell differentiation is inhibited by EDN3-EDNRB signaling.[75] Together these studies showed that EDN3-EDNRB signaling is important for ENCCs migration and proliferation, and for maintaining ENCCs in their progenitor state during ENS development.

By using a system that allows the regulation of expression of *Ednrb* at different stages of development, Shin and colleagues showed that EDNRB is required during a limited period of ENS development in mouse, namely between embryonic days 10 to 12.5.[72] *Edn3*, *Ednrb* and *Ece1* knockout mice show aganglionosis and pigmentary abnormalities, similar to the phenotypic abnormalities seen in human patients carrying mutations in these genes (Shah-Waardenburg syndrome). Mutations on these genes have been identified in approximately 5% of HSCR cases.[1]

### *RET and EDNRB Interaction*

The RET and the EDNRB signaling pathways were considered independent signaling pathways, however there are several findings that indicate that the two pathways are indirectly connected. A study by Puffenberger and colleagues identified a homozygous *EDNRB* mutation, which was found in association with a specific *RET* haplotype in a large Mennonite kindred with HSCR disease.[76] The data was later corroborated in a genome-wide association study of 43 trios from the same Mennonite population. Joint transmission of a specific *RET* haplotype and the previously identified *EDNRB* mutation with HSCR was found more often than would have been expected by chance.[55] In an Italian patient with HSCR, both a *RET* mutation, inherited from the unaffected mother, and an *EDNRB* mutation, from the unaffected father, were identified.[38] These studies suggest that there might be a genetic interaction between RET and EDNRB.

Studies using animal models support this hypothesis. McCallion and colleagues (2003) intercrossed mice carrying a *Ret* null mutation ($Ret^{-}/Ret^{+}$) and mice carrying an *Ednrb* loss-of-function mutation, $Ednrb^{s}/Ednrb^{s}$ or $Ednrb^{s\text{-}l}/Ednrb^{s}$ (s: piebald, s-l: piebald lethal). Of the offspring, 100% of the males carrying the $Ret^{-}/Ret^{+;}\ Ednrb^{s}/Ednrb^{s}$ genotype were affected, whereas only 70% of the females were affected with variable lengths of aganglionic colon. In mice offspring carrying the $Ret^{-}/Ret^{+;}\ Ednrb^{s\text{-}l}/Ednrb^{s}$ genotype, a genotype that expresses less EDNRB, both males and females were equally affected, showing that this sex difference is gene-dose dependant. This suggests there is epistasis between *EDNRB* and *RET*.[77] Similar studies showed that, whereas $Ret^{51/51}$ and $Edn3^{ls/ls}$ (lethal spotting) mice display colonic aganglionosis, combinations of these mutant alleles lead to almost complete intestinal aganglionosis.[73]

Finally, SOX10, a transcription factor that regulates *RET* expression, also influences *ED-*

*NRB* expression. Mice carrying $Sox10^{-/-}$ or $Sox10^{-/+}$ either show reduced or loss of *EDNRB* expression, respectively.[78]

### *Other Signaling Pathways*

The signaling pathways described above are ones related to known HSCR genes. However, there are many more signaling pathways that play important roles in ENS development, such as the NOTCH, WNT and TGF-β /BMPs signaling pathways. NOTCH signaling is involved in various aspects of neurogenesis. It could decrease cell proliferation (although not solely) and it is known to act as a negative regulator for neuron differentiation.[79, 80]

WNT signaling induces cell proliferation and maintains the pluripotency of stem cells. However, a recent study showed that the result of Wnt signaling can differ between different types of stem cells. In ENCCs, for example, WNT signaling is important for cell lineage decision-making rather than for maintaining the pluripotency of the stem cells, which is a function that WNT signaling has in hematopoietic and CNS stem cells. Deletion of Wnt1 and Wnt3a leads to a reduction of neural crest derivates.[81] The WNT signaling pathway interacts with other signaling pathways, such as the NOTCH and the TGF-β/BMP pathways.

BMP are expressed in submucosal and myenteric plexus and BMP4 has been shown to direct or indirectly influence ENCCs migration. Even though BMP4 alone is not sufficient to promote cell migration, in the absence of BMP4 activity, the number of ENCCs that reach the hindgut is reduced. Furthermore, BMP4 and BMP2 are also required *in vitro* for cell aggregation of isolated enteric neurons, which might indicate an important role for BMPs in ganglion formation during ENS development.[82] Chalazonist and colleagues proposed an interaction between GDNF and BMP4 in ENS. This was based on their data that the combination of these two proteins synergistically enhanced enteric neural development.[83] This study indicates there is possible cross-talk between the RET and BMP signaling pathways in ENS development. Another possibility is that GDNF directly influences BMP signaling independent of RET, since it has been suggested that GDNF could also interact with neural cell adhesion molecule (NCAM), which has been known play a role in a number of developmental process including cell migration, neurite outgrowth, and synaptic plasticity.[84]

## 4. Aims and Outline of this Thesis

The aims of the studies described in this thesis were to: (1) identify new HSCR-causing genes in associated loci reported in previous studies, in particular in 9q31 and 4q31-32.[45,56] (2) per-

form functional assays on a non-coding RET variant (SNP rs2506004) identified by Burzynski et al.[48] to reveal the pathogenic nature of the variant and to determine the molecular mechanism of this SNP with respect to RET expression regulation; and (3) perform gene expression profiling of ENCCs in the presence and absence of GDNF to gain a better insight into the RET signaling pathway in this cell type.

Bolk and colleagues showed that the 9q31 region contains a gene or mutation that acts as a modifier in the development of HSCR disease.[45] This locus was found in families that were linked to the *RET* region but in whom no clear *RET*-coding sequence mutation was identified and in one family that was not linked to the *RET* region. In **Chapter 2,** we focus on fine mapping the 9q31 region by genotyping 301 tagging-SNPs spanning a region of 7 Mb, in 137 Dutch trios (121 patients without a *RET*-coding mutation and 16 patients with *RET*-coding mutations) and in Chinese case-controls (173 patients and 436 controls) study.

In a previous linkage study performed by Brooks and colleagues on a Dutch multi-generational HSCR family, linkage was identified for a region on 4q31-32.[56] In **Chapter 3** we tried to discover the causative gene in this locus using exome sequencing.

Previous studies showed strong association of HSCR with a haplotype containing part of the 5′ region of the RET gene. This haplotype was transmitted to 60-85% of the HSCR patients, but only to 20-40% of the controls. Some of the variants present on this haplotype have been reported and suggested as HSCR-causative variant(s) by functional studies. For some possible causative variants the pathogenic nature has not yet been examined, in particular, SNP rs2506004 (C>A) that was identified by Burzynski et al. In **Chapter 4**, we describe functional studies of this SNP. By *in silico* analysis we found that the C-allele (wild-type) is located in a Central Midline Element (CME) sequence, which is a binding site for NXF/ARNT2 transcription activators and SIM2/ARNT2 repressors, respectively. Subsequently, we determined whether these transcription factors could indeed regulate RET expression and whether the mutant A-allele reduced the enhancer activity of this CME region on RET expression.

It has been shown that the GDNF-RET signaling pathway plays a central role in ENS development. However, the actual downstream effectors of RET signaling, in particular in ENCCs, is still largely unknown. To gain a better insight into the downstream effects of GDNF-RET signaling, we performed expression microarray experiments on mouse embryonic gut and also on ENCCs isolated from mouse embryonic gut (Embryonic day 14.5), untreated and treated with GDNF, the ligand of RET. We analyzed the microarray data not only with single-gene analysis, but also performed pathway analysis using Gene Set Enrichment Analysis (GSEA) software. The results are presented in **Chapter 5**.

The summary, general discussion and future perspective are presented in **Chapter 6**.

# REFERENCES


1. Amiel J, Lyonnet S. Hirschsprung disease, associated syndromes, and genetics: a review. J Med Genet 2001;38:729-739.

2. Hirschsprung H. Stuhltragheit neugeborener in folge von dilatation und hypertrophie des colons. Jahrb Kinderh 1887;27:1-7.

3. Ikeda K, Goto S. Diagnosis and treatment of Hirschsprung's disease in Japan. An analysis of 1628 patients. Ann Surg 1984;199:400-405.

4. Rescorla FJ, Morrison AM, Engles D, et al. Hirschsprung's disease. Evaluation of mortality and long-term function in 260 cases. Arch Surg 1992;127:934-941.

5. Whitehouse FR, Kernohan JW. Myenteric plexus in congenital megacolon; study of 11 cases. Arch Intern Med (Chic) 1948;82:75-111.

6. Devos AS, Blickman JG. Radiological Imaging of the Digestive Tract in Infants and Children. Springer Verlag 2007; ISBN:3540407332.

7. Passarge E. The genetics of Hirschsprung's disease. Evidence for heterogeneous etiology and a study of sixty-three families. N Engl J Med 1967;276:138-143.

8. Garver KL, Law JC, Garver B. Hirschsprung disease: a genetic study. Clin Genet 1985;28:503-508.

9. Moore SW, Rode H, Millar AJ, et al. Familial aspects of Hirschsprung's disease. Eur J Pediatr Surg 1991;1:97-101.

10. Torfs CP, Christianson RE. Anomalies in Down syndrome individuals in a large population-based registry. Am J Med Genet 1998;77:431-438.

11. Gabriel SB, Salomon R, Pelet A, et al. Segregation at three loci explains familial and population risk in Hirschsprung disease. Nat Genet 2002;31:89-93.

12. Emison ES, McCallion AS, Kashuk CS, et al. A common sex-dependent mutation in a RET enhancer underlies Hirschsprung disease risk. Nature 2005;434:857-863.

13. Dasgupta R, Langer JC. Transanal pull-through for Hirschsprung disease. Semin Pediatr Surg 2005;14:64-71.

14. Mattioli G, Pini Prato A, Giunta C, et al. Outcome of primary endorectal pull-through for the treatment of classic Hirschsprung disease. J Laparoendosc Adv Surg Tech A 2008;18:869-874.

15. Antao B, Radhwan T, Samuel M, et al. Short-pouch and low-anastomosis Duhamel procedure results in better fecal control and normal defecation pattern. Dis Colon Rectum 2005;48:1791-1796.

16. Catto-Smith AG, Trajanovska M, Taylor RG. Long-term continence after surgery for Hirschsprung's disease. J Gastroenterol Hepatol 2007;22:2273-2282.

17. Metzger M, Caldwell C, Barlow AJ, et al. Enteric nervous system stem cells derived from human gut mucosa for the treatment of aganglionic gut disorders. Gastroenterology 2009;136:2214-2225.

18. Angrist M, Kauffman E, Slaugenhaupt SA, et al. A gene for Hirschsprung disease (megacolon) in the pericentromeric region of human chromosome 10. Nat Genet 1993;4:351-356.

19. Lyonnet S, Bolino A, Pelet A, et al. A gene for Hirschsprung disease maps to the proxi-

mal long arm of chromosome 10. Nat Genet 1993;4:346-350.

20. Hofstra RM, Wu Y, Stulp RP, Elfferich P, et al. RET and GDNF gene scanning in Hirschsprung patients using two dual denaturing gel systems. Hum Mutat 2000;15: 418-429.

21. Edery P, Lyonnet S, Mulligan LM, et al. Mutations of the RET proto-oncogene in Hirschsprung's disease. Nature 1994;367:378-380.

22. Romeo G, Ronchetto P, Luo Y, et al. Point mutations affecting the tyrosine kinase domain of the RET proto-oncogene in Hirschsprung's disease. Nature 1994;367: 377-378.

23. Pasini B, Borrello MG, Greco A, et al. Loss of function effect of RET mutations causing Hirschsprung disease. Nat Genet 1995;10:35-40.

24. Puffenberger EG, Hosoda K, Washington SS, et al. A missense mutation of the endothelin-B receptor gene in multigenic Hirschsprung's disease. Cell 1994;79: 1257-1266.

25. Edery P, Attie T, Amiel J, et al. Mutation of the endothelin-3 gene in the Waardenburg-Hirschsprung disease (Shah-Waardenburg syndrome). Nat Genet 1996;12:442-444.

26. Hofstra RM, Osinga J, Tan-Sindhunata G, et al. A homozygous mutation in the endothelin-3 gene associated with a combined Waardenburg type 2 and Hirschsprung phenotype (Shah-Waardenburg syndrome). Nat Genet 1996;12:445-447.

27. Angrist M, Bolk S, Halushka M, et al. Germline mutations in glial cell line-derived neurotrophic factor (GDNF) and RET in a Hirschsprung disease patient. Nat Genet 1996;14:341-344.

28. Salomon R, Attie T, Pelet A, et al. Germline mutations of the RET ligand GDNF are not sufficient to cause Hirschsprung disease. Nat Genet 1996;14:345-347.

29. Doray B, Salomon R, Amiel J, et al. Mutation of the RET ligand, neurturin, supports multigenic inheritance in Hirschsprung disease. Hum Mol Genet 1998;7:1449-1452.

30. Pingault V, Bondurand N, Kuhlbrodt K, et al. SOX10 mutations in patients with Waardenburg-Hirschsprung disease. Nat Genet 1998;18:171-173.

31. Amiel J, Laudier B, Attie-Bitach T, et al. Polyalanine expansion and frameshift mutations of the paired-like homeobox gene PHOX2B in congenital central hypoventilation syndrome. Nat Genet 2003;33:459-461.

32. Hofstra RM, Valdenaire O, Arch E, et al. A loss-of-function mutation in the endothelin-converting enzyme 1 (ECE-1) associated with Hirschsprung disease, cardiac defects, and autonomic dysfunction. Am J Hum Genet 1999;64:304-308.

33. Brooks AS, Bertoli-Avella AM, Burzynski GM, et al. Homozygous nonsense mutations in KIAA1279 are associated with malformations of the central and enteric nervous systems. Am J Hum Genet 2005;77:120-126.

34. Cacheux V, Dastot-Le Moal F, Kaariainen H, et al. Loss-of-function mutations in SIP1 Smad interacting protein 1 result in a syndromic Hirschsprung disease. Hum Mol Genet 2001;10:1503-1510.

35. Wakamatsu N, Yamada Y, Yamada K, et al. Mutations in SIP1, encoding Smad interacting protein-1, cause a form of Hirschsprung disease. Nat Genet 2001;27: 369-370.

36. Garcia-Barcelo M, Ganster RW, Lui VC, et al. TTF-1 and RET promoter SNPs: regulation of RET transcription in Hirschsprung's disease.

Hum Mol Genet 2005;14: 191-204.

37. Garcia-Barcelo MM, Tang CS, Ngan ES, et al. Genome-wide association study identifies NRG1 as a susceptibility locus for Hirschsprung's disease. Proc Natl Acad Sci U S A 2009;106:2694-2699.

38. Auricchio A, Griseri P, Carpentieri ML, et al. Double heterozygosity for a RET substitution interfering with splicing and an EDNRB missense mutation in Hirschsprung disease. Am J Hum Genet 1999;64:1216-1221.

39. Pingault V, Bondurand N, Lemort N, et al. A heterozygous endothelin 3 mutation in Waardenburg-Hirschsprung disease: is there a dosage effect of EDN3/EDNRB gene mutations on neurocristopathy phenotypes? J Med Genet 2001;38:205-209.

40. Garavelli L, Donadio A, Zanacca C, et al. Hirschsprung disease, mental retardation, characteristic facial features, and mutation in the gene ZFHX1B (SIP1): confirmation of the Mowat-Wilson syndrome. Am J Med Genet A 2003;116A: 385-388.

41. Moore SW. Down syndrome and the enteric nervous system. Pediatr Surg Int 2008;24:873-883.

42. Arnold S, Pelet A, Amiel J, et al. Interaction between a chromosome 10 RET enhancer and chromosome 21 in the Down syndrome-Hirschsprung disease association. Hum Mutat 2009;30:771-775.

43. Rachidi M, Lopes C, Charron G, et al. Spatial and temporal localization during embryonic and fetal human development of the transcription factor SIM2 in brain regions altered in Down syndrome. Int J Dev Neurosci 2005;23:475-484.

44. de Pontual L, Pelet A, Clement-Ziza M, et al. Epistatic interactions with a common hypomorphic RET allele in syndromic Hirschsprung disease. Hum Mutat 2007;28: 790-796.

45. Bolk S, Pelet A, Hofstra RM, et al. 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 U S A 2000;97:268-273.

46. Fitze G, Appelt H, Konig IR, et al. Functional haplotypes of the RET proto-oncogene promoter are associated with Hirschsprung disease (HSCR). Hum Mol Genet 2003;12:3207-3214.

47. Fernandez RM, Boru G, Pecina A, et al. Ancestral RET haplotype associated with Hirschsprung's disease shows linkage disequilibrium breakpoint at -1249. J Med Genet 2005;42:322-327.

48. Burzynski GM, Nolte IM, Bronda A, et al. Identifying candidate Hirschsprung disease-associated RET variants. Am J Hum Genet 2005;76:850-858.

49. Griseri P, Bachetti T, Puppo F, et al. A common haplotype at the 5' end of the RET proto-oncogene, overrepresented in Hirschsprung patients, is associated with reduced gene expression. Hum Mutat 2005;25:189-195.

50. Emison ES, Garcia-Barcelo M, Grice EA, et al. Differential contributions of rare and common, coding and noncoding Ret mutations to multifactorial Hirschsprung disease liability. Am J Hum Genet 2010;87:60-74.

51. Grice EA, Rochelle ES, Green ED, et al. Evaluation of the RET regulatory landscape reveals the biological relevance of a HSCR-implicated enhancer. Hum Mol Genet 2005;14:3837-3845.

52. Sribudiani Y, Metzger M, Osinga J, et al. Variants in RET associated with Hirschsprung's

disease affect binding of transcription factors and gene expression. Gastroenterology 2011;140:572-582.

53. Ema M, Ikegami S, Hosoya T, Mimura J, et al. Mild impairment of learning and memory in mice overexpressing the mSim2 gene located on chromosome 16: an animal model of Down's syndrome. Hum Mol Genet 1999;8:1409-1415.

54. Meng X, Peng B, Shi J, et al. Effects of overexpression of Sim2 on spatial memory and expression of synapsin I in rat hippocampus. Cell Biol Int 2006;30:841-847.

55. Carrasquillo MM, McCallion AS, Puffenberger EG, et al. Genome-wide association study and mouse model identify interaction between RET and EDNRB pathways in Hirschsprung disease. Nat Genet 2002;32:237-244.

56. Brooks AS, Leegwater PA, Burzynski GM, et al. A novel susceptibility locus for Hirschsprung's disease maps to 4q31.3-q32.3. J Med Genet 2006;43:e35.

57. Heanue TA, Pachnis V. Enteric nervous system development and Hirschsprung's disease: advances in genetic and stem cell studies. Nat Rev Neurosci 2007;8:466-479.

58. Le Douarin NM, Teillet MA. The migration of neural crest cells to the wall of the digestive tract in avian embryo. J Embryol Exp Morphol 1973;30:31-48.

59. Barembaum M, Bronner-Fraser M. Early steps in neural crest specification. Semin Cell Dev Biol 2005;16:642-646.

60. Yntema CL, Hammond WS. The origin of intrinsic ganglia of trunk viscera from vagal neural crest in the chick embryo. J Comp Neurol 1954;101:515-541.

61. Durbec P, Marcos-Gutierrez CV, Kilkenny C, et al. GDNF signalling through the Ret receptor tyrosine kinase. Nature 1996;381:789-793.

62. Kapur RP, Yost C, Palmiter RD. A transgenic model for studying development of the enteric nervous system in normal and aganglionic mice. Development 1992;116:167-175.

63. Burns AJ. Migration of neural crest-derived enteric nervous system precursor cells to and within the gastrointestinal tract. Int J Dev Biol 2005;49:143-150.

64. Pomeranz HD, Rothman TP, Gershon MD. Colonization of the post-umbilical bowel by cells derived from the sacral neural crest: direct tracing of cell migration using an intercalating probe and a replication-deficient retrovirus. Development 1991;111: 647-655.

65. Taraviras S, Marcos-Gutierrez CV, Durbec P, et al. Signalling by the RET receptor tyrosine kinase and its role in the development of the mammalian enteric nervous system. Development 1999;126:2785-2797.

66. Young HM, Newgreen D. Enteric neural crest-derived cells: origin, identification, migration, and differentiation. Anat Rec 2001;262:1-15.

67. Arighi E, Borrello MG, Sariola H. RET tyrosine kinase signaling in development and cancer. Cytokine Growth Factor Rev 2005;16:441-467.

68. Richardson DS, Lai AZ, Mulligan LM. RET ligand-induced internalization and its consequences for downstream signaling. Oncogene 2006;25:3206-3211.

69. Natarajan D, Marcos-Gutierrez C, Pachnis V, et al. Requirement of signalling by receptor tyrosine kinase RET for the directed migration of enteric nervous system progenitor cells during mammalian embryogenesis. Develop-

ment 2002;129:5151-5160.

70. Heuckeroth RO, Enomoto H, Grider JR, et al. Gene targeting reveals a critical role for neurturin in the development and maintenance of enteric, sensory, and parasympathetic neurons. Neuron 1999;22:253-263.

71. Rossi J, Herzig KH, Voikar V, et al. Alimentary tract innervation deficits and dysfunction in mice lacking GDNF family receptor alpha2. J Clin Invest 2003;112: 707-716.

72. Shin MK, Levorse JM, Ingram RS, et al. The temporal requirement for endothelin receptor-B signalling during neural crest development. Nature 1999;402:496-501.

73. Barlow A, de Graaff E, Pachnis V. Enteric nervous system progenitors are coordinately controlled by the G protein-coupled receptor EDNRB and the receptor tyrosine kinase RET. Neuron 2003;40:905-916.

74. Leibl MA, Ota T, Woodward MN, et al. Expression of endothelin 3 by mesenchymal cells of embryonic mouse caecum. Gut 1999;44:246-252.

75. Bondurand N, Natarajan D, Barlow A, et al. Maintenance of mammalian enteric nervous system progenitors by SOX10 and endothelin 3 signalling. Development 2006;133:2075-2086.

76. Puffenberger EG, Kauffman ER, Bolk S, et al. Identity-by-descent and association mapping of a recessive gene for Hirschsprung disease on human chromosome 13q22. Hum Mol Genet 1994;3:1217-1225.

77. McCallion AS, Stames E, Conlon RA, et al. Phenotype variation in two-locus mouse models of Hirschsprung disease: tissue-specific interaction between Ret and Ednrb. Proc Natl Acad Sci U S A 2003;100:1826-1831.

78. Southard-Smith EM, Kos L, Pavan WJ. Sox10 mutation disrupts neural crest development in Dom Hirschsprung mouse model. Nat Genet 1998;18:60-64.

79. Okamura Y, Saga Y. Notch signaling is required for the maintenance of enteric neural crest progenitors. Development 2008;135:3555-3565.

80. Cornell RA, Eisen JS. Notch in the pathway: the roles of Notch signaling in neural crest development. Semin Cell Dev Biol 2005;16:663-672.

81. Kleber M, Sommer L. Wnt signaling and the regulation of stem cell function. Curr Opin Cell Biol 2004;16:681-687.

82. Goldstein AM, Brewer KC, Doyle AM, et al. BMP signaling is necessary for neural crest cell migration and ganglion formation in the enteric nervous system. Mech Dev 2005;122:821-833.

83. Chalazonitis A, D'Autreaux F, Guha U, et al. Bone morphogenetic protein-2 and -4 limit the number of enteric neurons but promote development of a TrkC-expressing neurotrophin-3-dependent subset. J Neurosci 2004;24:4266-4282.

84. Paratcha G, Ledda F, Ibanez CF. The neural cell adhesion molecule NCAM is an alternative signaling receptor for GDNF family ligands. Cell 2003;113:867-879.

# 2

# Fine mapping of the 9q31 Hirschsprung disease locus

CS Tang[1,#], Y Sribudiani[5,#], XP Miao[2,7], AR de Vries[5], G Burzynski[5,6], MT So[2], YY Leon[2], BH Yip[1,2], J Osinga[5], KJWS Hui[2], JBGM Verheij[5], SS Cherny[1], PKH Tam[2,4], PC Sham[1,3,4], RMW Hofstra[5,*], MM Garcia-Barceló[2,4,*]

[1]Department of Psychiatry, [2]Department of Surgery, [3]Genome Research Centre, [4]Centre for Reproduction, Development, and Growth of The University of Hong Kong; Hong Kong, China; [5]University Medical Center Groningen, University of Groningen, Department of Genetics, Groningen, The Netherlands; [6]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, USA; [7]Department of Epidemiology and Biostatistics, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

**#These authors contributed equally to this work**

**Hum Genet, 2010 Jun; 127(6):675-83**

## ABSTRACT

Hirschsprung disease (HSCR) is a congenital disorder characterized by the absence of ganglia along variable lengths of the intestine. The *RET* gene is the major HSCR gene. Reduced penetrance of *RET* mutations and phenotypic variability suggest the involvement of additional modifying genes in the disease. A *RET*-dependent modifier locus was mapped to 9q31 in families bearing no coding sequence (CDS) *RET* mutations. Yet, the 9q31 causative locus is to be identified. To fine-map the 9q31 region, we genotyped 301 tag-SNPs spanning 7 Mb on 137 HSCR Dutch trios. This revealed two HSCR-associated regions that were further investigated in 173 Chinese HSCR patients and 436 controls using the genotype data obtained from a genome-wide association study (GWAS) recently conducted.

Within one of the two identified regions *SVEP1* SNPs were found associated with Dutch HSCR patients in the absence of *RET* mutations. This ratifies the reported linkage to the 9q31 region in HSCR families with no *RET* CDS mutations. However, this finding could not be replicated. In Chinese, HSCR was found associated with *IKBKAP*. In contrast, this association was stronger in patients carrying *RET* CDS mutations with $p$=5.10x10$^{-6}$ [OR=3.32 (1.99, 5.59)] after replication. The HSCR-association found for *IKBKAP* in Chinese suggest population specificity and implies that *RET* mutation carriers may have an additional risk. Our finding is supported by the role of *IKBKAP* in the development of the nervous system.

# INTRODUCTION

Hirschsprung disease (HSCR; aganglionic megacolon) is a congenital disorder characterized by the absence of enteric ganglia along a variable length of hindgut. There is significant ethnic variation in the incidence of the disease, and it is most often found among Asians (2.8 per 10,000 live births.[1 2] Non-familial HSCR has a complex pattern of inheritance and manifests with low, sex-dependent penetrance and variability in the length of the aganglionic segment, according to which patients are classified into short segment (S-HSCR; 80%), long segment (L-HSCR; 15%), and total colonic aganglionosis (TCA 5%). The male:female ratio (M:F) is ≈4:1 among S-HSCR patients and ≈1:1 among L-HSCR patients. HSCR presents mostly sporadically although it can be familial (5-20% of cases) where the recurrence risks to sibs vary from 1.5–33% depending on the gender and the length of the aganglionic segment in the proband, and the gender of the sibling.[3]

The *RET* gene, encoding a tyrosine-kinase receptor, is the major HSCR causing gene and its expression is crucial for the development of the enteric ganglia.[4, 5] Mutations in the coding sequence (CDS) of *RET* account for up to 50% of the familial cases and between 15%-20% of the sporadic cases, indicating that additional HSCR-causing mutations exist.[6] Other HSCR genes identified so far mainly code for protein members of interrelated signalling pathways involved in the development of enteric ganglia: RET, endothelin receptor B (EDNRB), and the transcriptional regulator SOX10. Yet, mutations in genes other than *RET* (*GDNF, GFRA1, NRTN, PHOX2B, NKX2.1, SOX10, NRG1, EDNRB, EDN3, ECE-1, KIAA1279, ZFXH1B, NTRK3, L1CAM*) account for only 7%.[7-15] Failure to identify *RET* CDS mutations in some of the *RET*-linked families suggests that noncoding *RET* mutations, mutations in regulatory regions of *RET* could also contribute significantly to the disease. Indeed, common *RET* single nucleotide polymorphisms (SNPs) are strongly associated with HSCR, with the largest contribution to risk made by a functional SNP (rs2435357) lying in an enhancer-like sequence in intron 1.[16-23] We have recently reported the genetic interaction of *RET* intron 1 rs2435357 SNP with HSCR-associated *NRG1* SNPs whereby the risk to HSCR is increased by 2.3-fold for the *RET* heterozygous genotype.[24] Reduced penetrance of *RET* and other HSCR genes mutations and variable expression of the HSCR phenotype indicates that the disease may result from the combined effect of several genes whereby the outcome would be altered *RET* expression.[25-28]

*RET* modifiers have been mapped to the following chromosomal regions: (i) 3p21; (ii) 19q12 [29]; (iii) 4q31.3-q32.3 [30] and (iv) 9q31. Despite these important findings, the genes in these loci are yet to be identified. The 9q31 locus segregated in families of European descent harbouring no or hypomorphic *RET* mutations, yet the families still showed linkage to 10q12 (*RET* locus). [31] To identify the HSCR locus in the 9q31 candidate region, we have conducted dense genotyping of the region a Dutch patient cohort and evaluated the same region in Chinese.

# MATERIALS AND METHODS

The overall study was approved by the institutional review board of The University of Hong Kong together with the Hospital Authority (IRB: UW 07-292).

## Dutch (Groningen) family-based association study

The Dutch patient cohort consisted of HSCR 140 trios that were genotyped using Illumina GoldenGate platform for 370 SNPs spanning the 7 Mb (from 108.5–115.5 Mb) of the 9q31 HSCR-associated region as described in Bolk *et al*. Most of these 370 were tag SNPs. Quality control based on identity-by-descent (IBD) revealed 3 duplicated trios. SNPs with missing genotypes (>5%) and with minor allele frequency (MAF)<5% and/or violating Hardy-Weinberg equilibrium ($p$<0.001) were excluded, leaving a total of 301 SNPs for association study. PLINK [32] was then used to conduct the TDT analysis in the resulting 137 Dutch trios. Sixteen HSCR probands carried *RET* CDS mutations.

## Chinese case-control analysis

*Subjects*

For analysis of the 9q31 region in the Chinese sample, we revised the genotype data of the candidate 9q31.3 region obtained from a previously GWAS conducted on a total of 181 HSCR Chinese patients and 386 healthy individuals (discovery group) using *Affymetrix* GeneChip®Human Mapping 500K Array as previously described. [24] To that data, we included 92 additional controls recently genotyped (see below).

For the replication of the association found when patients were stratified according to the *RET* mutation status, we included 21 independent HSCR cases bearing *RET* CDS mutations and 71 control individuals from Mainland China (replication group).Characteristics of the patients are depicted in supplementary Tables 1 and 2.

*Revision of the Affymetrix 500K data*

Following the recent recommendations on quality control standards[33] we imposed a more stringent filtering criterion based on heterozygosity to the GWAS described above. Heterozygosity screening aims at the identification of individuals who have more heterozygous genotype calls than the average of the sample set, since the excess of heterozygosity is likely due to cross-contamination of samples. Hence, outliers should be excluded from the analysis. Prior to evaluating the heterozygosity of the samples, we pruned the SNP data set to avoid bias that could have been introduced by large quantities of SNPs in strong LD. In

brief, we pruned the dataset (by removing 1 of a pair of SNPs with $r^2$>0.25 within a window of 200 SNPs and shifts of 50 SNPs forward at each step) using PLINK. In total, 56,000 SNPs in LD with each other remained. The heterozygosity screening excluded outliers with more heterozygous calls than the average (estimated inbreeding coefficient lower than 3 standard deviations from the average).

After this filtering criteria, a total of 173 HSCR cases (31 patients bearing *RET* CDS mutations) and 436 controls were included in the analysis. Restricting by the region analyzed by the Dutch group, 988 SNPs falling within the 9q31 7 Mb candidate region (from 108.5–115.5 Mb) were selected. The mean call rate was 99.09%. After filtering by MAF, genotyping rate and Hardy-Weinberg equilibrium, 981 SNPs were left for analysis.

*Correction for population structure and association analysis*

We used the EIGENSOFT package on the whole genome data pruned for LD as described above to detect population substructure. The Tracy-Widom test (implemented in EIGENSOFT) nominated only the first two axes of variation as significant ($p$<0.05). The first axis of variation corresponded to the ancestral differences between Northern and Southern Chinese as seen in supplementary Figure 1. To correct for these ancestral differences, we applied EIGENSTRAT [34] (also implemented in EIGENSOFT) on the candidate regions based on the two axes of variation nominated by the Tracy-Widom test. The Cochran-Armitage trend test was then used to assess the levels of association of the SNPs with HSCR.

*Replication stage*

EIGENSTRAT could not be applied to correct the replication sample for population stratification as no genealogy information can be inferred when only two SNPs are genotyped. Logistic regression was carried out instead to assess the association in the combined sample (discovery + replication samples). We included a covariate for the sample ancestral origin (northern and southern Chinese ancestry) and a covariate to distinguish GWAS and replication samples, thereby controlling for any possible allele frequency differences between known populations and between stages of association analysis. We also tested the interaction terms between each covariate and the genotype, which were found to not be significant ($p$>0.5 for both, indicating that the effect is not significantly different between subpopulations or stages). Thus, only the two covariates were included in the final model for combined analysis.

# RESULTS

## Family-based association test on Dutch trios identifies two associated regions on 9q31

Through TDT analysis of 301 SNPs on 137 Dutch trios, we identified 17 SNPs (supplementary Table 3) ($p$<0.05) showing asymmetrical transmission from parents. Of these, 10 SNPs could be grouped into 2 moderately associated peaks of 5 SNPs each, **9q31A** (110.7–111.1Mb) and **B** (112.1-112.7Mb).

**The 9q31A region** encompasses a gene-rich spot which contains 4 genes *IKBKAP*, *C9orf6*, *CTNNAL1* and *C9orf5* within a single LD block that spans about 100kb. Among the 5 associated SNPs, the strongest evidence for association with HSCR was found for **rs12351693** ($p$=0.023) and **rs10979637** ($p$=0.023) located within the intronic region of *CTNNAL1* (Table 1 and Figure 1A).

*CTNNAL1* encodes alpha-catulin**,** a protein that modulates the Rho signalling pathway by providing a scaffold for the *ARHGEF1*.[35] The expression of its mouse homologue, *Ctnnal1,* was found to be severely reduced in intestines of *Ret* deficient mice ($Ret^{k-/k-}$) when compared to normal mice ($Ret^{+/+}$) and consequently, the gene has been considered a putative *RET* modifie.[36]

However, the strongest associations lied in 9q31B and mapped upstream the *TXNDC8* gene (p=4.8x$10^{-3}$ for **rs7038415)**, and in the intronic region of ***SVEP1*** ($p$=2.9x$10^{-3}$ for **rs10816998)** (Table 1). Other associated SNPs within the 9q31B region, mapped to the *MUSK* gene. *MUSK* is required for neuromuscular junction formation in vivo and for the survival and development of discrete neuronal subpopulations.[37] The associations described above were confirmed by performing 1 million **permutations (Table 1 and Figure 1).**

*SVEP1 is associated with HSCR in Dutch patients without RET mutations*

As it was reported that the 9q31 HSCR susceptibility locus segregated exclusively in *RET*-linked families depleted of *RET* CDS mutations[31], we stratified the cases according to their *RET* CDS mutation status. We defined *RET* CDS mutations as DNA changes in the coding region or exon/intron boundaries of *RET* that lead to frame-shift, non-sense or missense mutations in the RET protein and that have not been found in at least 400 control chromosomes. PolyPhen [38] was used to predict the effect of the missense mutations on the RET protein. Truncating mutations were considered highly deleterious.

The 301 SNPs in the 9q31 region were checked again for association. The rs10816998 and rs7038415 SNPs within *SVEP1* were strongly associated with HSCR in families without *RET* mutations. After stratification of the patients, the OR attained a significance level of

$p$=5.33x10$^{-5}$ with OR=2.379 (95% CI:1.542, 3.671) for rs10816998 and $p$=7.50x10$^{-5}$ with OR=2.345 (95% CI:1.518, 3.622) for rs7038415. These association values survived the conservative Bonferroni correction for multiple testing (≈300 SNPs). This finding in the Dutch population is in line with the reported linkage of 9q31 with HSCR in patients of European descent bearing no or hypomorphic *RET* mutations.[31]

The *SVEP1* SNPs rs10816998 and rs7038415, were then genotyped in 107 independent Dutch patients without ***RET*** CDS mutations and 183 controls. Unfortunately, the p value of the replication did not reach statistical significance.

**Table 1 -** 9q31A and B SNPs associated with HSCR in Dutch trios samples

| **9q31A** | | Allele | | | | |
|---|---|---|---|---|---|---|
| | Position (bp) | A | B | Overlapping gene | Location | *p*-value[a] |
| rs12351693 | 110751336 | G | A | CTNNAL1 | intron 13 | 0.023 |
| rs10979637 | 110784448 | C | A | CTNNAL1 | intron 6 | 0.023 |
| rs12555920 | 110805558 | T | C | CTNNAL1 | intron 1 | 0.041 |
| rs1333344 | 110816045 | A | G | CTNNAL1 | downsteam | 0.046 |
| | | | | C9orf5 | upstream | |
| rs1044905 | 110821316 | T | C | C9orf5 | 3' UTR | 0.028 |
| **9q31B** | | | | | | |
| rs7038415 | 112157004 | A | C | TXNDC8 | upstream | 0.0048 |
| | | | | SVEP1 | downstream | |
| rs10816998 | 112173562 | G | A | SVEP1 | intron 46 | 0.0029 |
| rs3010815 | 112473999 | T | C | MUSK | intron 1 | 0.046 |
| rs2766998 | 112571406 | T | C | MUSK | intron 9 | 0.029 |
| rs3780529 | 112601072 | C | A | MUSK | intron14 | 0.035 |

A: minor allele; B: major allele; a: p-value for the transmission disequilibrium test (TDT).

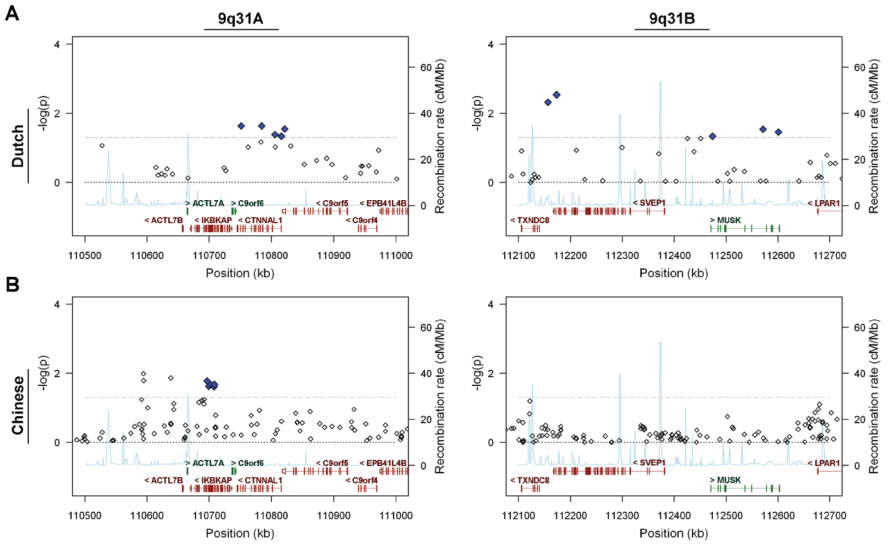



**Figure 1 -** Association analysis on 9q31A (110500-111000kb) (*left panel*) and 9q31B (112000-112500kb) (*right panel*). (A) TDT analysis on 137 Dutch HSCR trios and (B) Case-control study on Chinese using 500k. Associated SNPs listed in Table 1 and 2 are represented as blue diamonds. Recombination rate and genes in the region are also shown (green, "+" strand; red, "-" strand).

## Revision of the population-based GWAS on Chinese identifies IKBKAP as a HSCR candidate gene

We used the GWAS data (above described) to investigate the association observed in the Dutch. After EIGENSTRAT correction for stratification (see material and methods and supplementary material), genotypes of 981 SNPs spanning the entire 9q31 region were tested for association on 173 HSCR patients (31 with *RET* CDS mutations) and 436 controls.

In the Chinese dataset, 6 SNPs within the **9q31A** region showed association with HSCR. No association was found in **9q31B.** The loci displaying the strongest statistical evidence for association were **rs10979596** and **rs10979597** (in intron 25 of the *IKBKAP* gene) which are in complete LD with each other ($r^2$=1) (Figure 1B and Table 2). Four additional SNPs within *IKBKAP* also showed marginally significant association ($p$<0.03) with HSCR. These include the *IKBKAP* I816L missense polymorphism (rs2230793, $p$=0.024) and 3 intronic SNPs (Table 2). Mutations in *IKBKAP* (inhibitor of kappa light polypeptide gene enhancer in B-cells kinase complex-associated protein, involved in transcription elongation) are associated with a neurodevelopmental disease, the Riley-Day syndrome, also known as familial dysautonomia (FD).[39, 40] FD is an autosomal recessive disorder that occurs almost exclusively in persons of Ashkenazi Jewish descent. This neuropathy results from depletion of subsets of autonomic

and sensory neurons.[41] FD patients are born with fewer neurons in their dorsal root and sympathetic ganglia than their peers, and over time, neurons become more scarce.[42] Most importantly, some patients with FD also suffer from gastrointestinal dysfunction shortly after birth and interestingly, the co-occurrence of both FD and HSCR has been reported.[43] Interestingly, the region encompassing our 6 *IKBKAP* associated SNPs includes the region where FD-causing mutations lie. *Ikbkap* is expressed in the rat developing gut and *IKBKAP* in the human colon, supporting a biological role for this gene in the development of enteric nervous system.[42, 44]

**Table 2 -** 9q31A SNPs associated with HSCR in Chinese

| 9q31A | | Allele | | | | Frequencies | | |
|---|---|---|---|---|---|---|---|---|
| | Position (bp) | A | B | IKBKAP | OR (95% CI) | Cases | Controls | *p*-value[a] |
| rs10979596 | 110696824 | G | T | intron 25 | 1.53 (1.10, 2.13) | 0.20 | 0.14 | 0.017 |
| rs10979597 | 110696853 | A | T | intron 25 | 1.53 (1.10, 2.13) | 0.20 | 0.14 | 0.017 |
| rs2230793 | 110699304 | G | T | I816L exon 23 | 1.58 (1.20, 2.09) | 0.32 | 0.23 | 0.024 |
| rs2275630 | 110702228 | G | A | intron 19 | 1.68 (1.12, 2.51) | 0.13 | 0.08 | 0.021 |
| rs10979607 | 110707646 | G | A | intron 14 | 1.49 (1.08, 2.06) | 0.20 | 0.14 | 0.025 |
| rs4369056 | 110708259 | T | C | intron 14 | 1.58 (1.21, 2.08) | 0.33 | 0.24 | 0.021 |

A: minor allele; B: major allele; [a]: corrected trend test computed using EIGENSTRAT.

*The association of IKBKAP is stronger in Chinese HSCR patients with RET mutations*

The Chinese HSCR patients were also stratified according to the *RET* CDS mutation status. We then tested again the 981 SNPs in the 9q31 region for association by independently comparing *RET* CDS mutation carriers (31 Chinese HSCR patients; discovery group) against 436 controls as well as non-carriers (142 HSCR patients) versus the same control set.

Twelve SNPs were found associated with HSCR in *RET* mutations carriers ($p<0.01$). Importantly, 6 of these SNPs mapped to the *IKBKAP* loci in the 9q31A and had been found associated with HSCR before the stratification of the patients according to the mutation status (Figure 2). This data suggested that the initial association detected was mainly driven by this subset of *RET* mutation carriers and that the original marginal association had been masked prior stratification because of the small number of carriers. Among the 12 SNPs segregating with *RET* mutations, four *IKBKAP* SNPs survived the conservative Bonferroni correction for multiple testing (≈1000 SNPs) (Table 3). The frequency of the risk SNPs alleles tended to increase with the severity of the *RET* mutations, yet the small sample does not allow to draw any conclusion (Table 3).

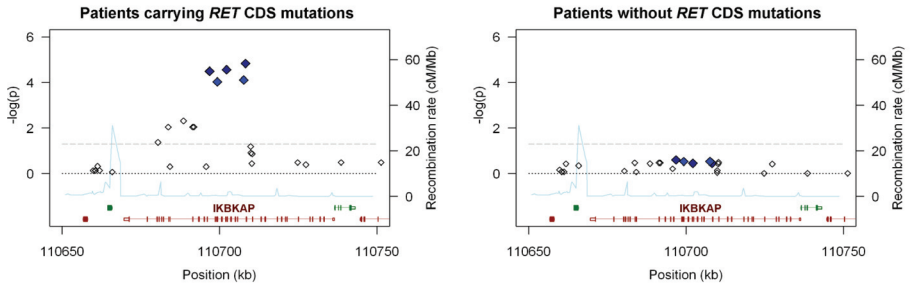



**Figure 2 -** Association analysis on *RET* Chinese CDS mutation carriers (left) and Chinese non-carriers (right). Blue diamonds represent associated SNPs before stratification by mutation status. Dark blue diamonds indicate the 4 SNPs that survived the Bonferroni correction multiple for testing (only 3 dark blue diamonds can be observed as there is overlapping between two SNPs).

**Table 3 -** IKBKAP SNPs associated[a] with HSCR patients with RET CDS mutations

| | Risk allele | MAF of risk allele | | | | | OR (95% CI) | Trend test p-value[d] |
|---|---|---|---|---|---|---|---|---|
| | | RET[b] | Highly deleterious | Damaging | Tolerating | Controls | | |
| rs4369056 | T | 0.52 | 0.75 | 0.50 | 0.47 | 0.24 | 3.44 (2.04, 5.80) | $1.48 \times 10^{-5}$ |
| rs2275630 | G | 0.25 | 0.38 | 0.30 | 0.19 | 0.08 | 3.83 (2.03, 7.23) | $2.79 \times 10^{-5}$ |
| rs10979596[c] | G | 0.34 | 0.63 | 0.35 | 0.32 | 0.14 | 3.21 (1.83, 5.62) | $3.25 \times 10^{-5}$ |
| rs10979597[c] | A | | | | | | | |

[a]: Survived Bonferroni correction for multiple testing; [b]: Frequency in HSCR patients carrying RET CDS mutations; [c]: same frequencies for these two SNPs as $r^2=1$; [d]: corrected trend test computed using EIGENSTRAT; MAF: minor allele frequency; Note that the MAFs increase with the severity of the *RET* mutation.

To validate the association of the *IKBKAP* SNPs with HSCR patients carrying *RET* CDS mutations, we genotyped (by PCR followed by direct sequencing) the 2 perfectly linked SNPs (rs10979596 and rs10979597) in an independent set of 21 HSCR patients (replication group) bearing *RET* mutations and 71 controls collected from mainland China. These SNPs were chosen for replication because they fall in the same amplicon and on the basis of genetic homogeneity between Northern and Southern Chinese. We successfully detected the association of rs10979596 and rs10979597 in the replication sample, with logistic regression association value of $p=9.45 \times 10^{-3}$ and OR=3.38 (95% CI: 1.35, 8.49) (Table 4). Logistic regression was also used for the combined analysis of HSCR samples bearing *RET* mutations (discovery, -31 patients- and replication -21 patients). This yielded a highly significant association of

$p$=2.71x10$^{-6}$ and OR=3.15 (95% CI:1.95, 5.09) for the heterozygous genotype under an additive model. Considering the possible population stratification, we employed the same logistic regression approach conditioning on the strata defined by sample origins (Northern and Southern Chinese), obtaining a significant level of association of $p$=5.75x10$^{-6}$ with OR=3.29 (95% CI:1.97, 5.51) (Table 4). Results of the latter approach were very similar, indicating that samples collected inside and outside Hong Kong (discovery and replication phases) were homogeneous, with little stratification.

**Table 4 -** Summary of the statistics for the *IKBKAP* rs10979596 and rs10979597 markers in *RET* mutation carriers

| Phase | N | | MAF | | Logistic[b] | | | |
|---|---|---|---|---|---|---|---|---|
| | *RET*[a] | Controls | *RET*[a] | Controls | OR (95% CI) | *p* (adjusted) | OR (95% CI) | *p* (non-adjusted) |
| 500K | 31 | 436 | 0.34 | 0.14 | 3.31 (1.77, 6.18) | 1.78 x 10$^{-4}$ | 3.38 (1.86, 6.12) | 6.13 x 10$^{-5}$ |
| Replication | 21 | 71 | 0.29 | 0.11 | NA | NA | 3.38 (1.35, 8.49) | 9.45 x 10$^{-3}$ |
| Combined | 52 | 507 | 0.319 | 0.13 | 3.32 (1.99, 5.59) | 5.10 x 10$^{-6}$ | 3.15 (1.95, 5.09) | 2.71 x 10$^{-6}$ |

NA: all cases and control used in the replication stage had the same ancestral origin of northern Chinese;
[a]: HSCR patients carrying RET CDS mutations
[b]: P-value adjusted for population substructure in GWAS (500K) analysis and for both the substructure and strata defining analysis stages (500K and replication) in combined analysis

## DISCUSSION

This study aims at the identification of a HSCR susceptibility locus in the 9q31 region. As the association of this region was initially reported in families of European descent, we first explored the region in a Dutch cohort. As the evaluation of an association in two populations of different ethnic origins would definitively increase confidence in the finding, we also explored the 9q31 region in the Chinese. We reasoned that this strategy would allow us to determine i) whether the 9q31 locus exerted an effect in patients of European descent other than those in whom the locus was initially identified, ii) assess whether the association of the 9q31 locus with HSCR is population specific, and if not; iii) use population differences in LD to fine-map the gene of interest.

Our data show an initial strong association of *SVEP1* SNPs with Dutch HSCR patients

bearing no *RET* CDS mutations. This finding ratifies the reported linkage to the 9q31 region in HSCR families with no *RET* CDS mutations. The case-control replication of the *SVEP1* association in the Dutch did not reach statistical significance indicating that the initial association was spurious or that there is a population stratification issue in the samples used for replication. This point would need further exploration in a much larger sample. The role of *SVEP1* is not well characterized. Study of its possible implications in the development of the nervous system should follow.

Intriguingly, the analysis of the region in the Chinese population demonstrated the existence of a HSCR-associated locus (*IKBKAP*) that, although within 9q31, is not correlated at all with the *SVEP1* locus identified in the Dutch. The association of *IKBKAP* SNPs with HSCR in Chinese became much more prominent in those patients with *RET* CDS mutations which is at odds with the initial findings described by Bolk and our own findings in the Dutch. Importantly, the *IKBKAP* association with HSCR bearing *RET* CDS mutations was replicated in the same Chinese population. No *IKBKAP* association was detected in the Dutch (supplementary Table 4).

We investigated whether the lack of cross-population replication of the associations of *SVEP1* and *IKBKAP* SNPs with HSCR could be due to population LD and allele frequencies differences, in which case, some true associations may not be replicated, regardless of the sample size of the study. Close examination of the *SVEP1* and *IKBKAP* regions showed no major LD differences between CEU and CHB HapMap populations indicating that lack of replication across populations was not due to differences in LD structure (Supplementary Figures 2 and 3). Lack of replication could also be attributed to a false positive result on the Chinese sample although we think this is unlikely since the initial association found in Chinese was replicated in an independent sample.

In view of these results, we would conclude that the 9q31 locus that appears associated with HSCR in families bearing no *RET* CDS mutations is indeed population specific, as the findings in Chinese point to a different locus with a completely different genetic behavior than that identified in the Dutch. As 9q31 locus was reported to segregate exclusively with *RET*-linked families of Caucasian origin whose members had weak or hypomorphic *RET* CDS mutations, the investigation of *SVEP1* rare variants in these type of patients is warranted.

The association of *IKBKAP* in Chinese implies that *RET* mutation carriers may have an additional risk. As yet, no genotype-phenotype correlation could be found which, after all, is in line with the genetic complexity inherent to HSCR. Had parental DNA been available, we might have been able to provide an argument for the role of *IKBKAP* SNPs on the penetrance of inherited *RET* CDS mutations. SNPs may modulate the penetrance or phenotype of a par-

ticular mutation through epistasis as exemplified by the exclusive action of the *FGFR2* and *MAP3K1* SNPs in *BRCA2* mutation carriers.[45]

## FUNDING

This work was supported by grants from the Hong Kong Research Grants Council HKU 765008M and HKU 775907M and The University of Hong Kong Seed Funding 200709159003 and 200611159152 to MGB and PT respectively. Support was also obtained from the University Grants Committee of Hong Kong (AoE/M-04/04) and from The University of Hong Kong Genomics Strategic Research Theme. SSC and PCS are supported by NIH Grant EY-12562. YS is supported by a grant from the Bernoulle Foundation and GB was supported by a NWO grant (no. 901-04-225).

## ACKNOWLEDGEMENTS

We would like to express our gratitude to all the subjects who participated in the study. This work supported by grants from Hong Kong Research Grants Council HKU 765008M and HKU 775907M and the University of Hong Kong Seed Funding 200709159003 and 200611159152 to MGB and PT, respectively. Support was also obtained from the University Grants Committee of Hong Kong (AoE/M-04/04) and from the University of Hong Kong Genomics Strategic Research Theme. SSC and PCS are supported by NIH Grant EY-12562. YS is supported by grant from the Bernoulle Foundation and GB was supported by NOW grant (no.901-04-225).

# REFERENCES


1. Amiel J, Sproat-Emison E, Garcia-Barcelo M, et al. Hirschsprung disease, associated syndromes and genetics: a review. J Med Genet 2008;45:1-14.

2. Torfs CP, Christianson RE. Anomalies in Down syndrome individuals in a large population-based registry. Am J Med Genet 1998;77:431-438.

3. Badner JA, Sieber WK, Garver KL, et al. A genetic study of Hirschsprung disease. Am J Hum Genet 1990;46:568-580.

4. Edery P, Lyonnet S, Mulligan LM, et al. Mutations of the RET proto-oncogene in Hirschsprung's disease. Nature 1994;367:378-380.

5. Romeo G, Ronchetto P, Luo Y, et al. Point mutations affecting the tyrosine kinase domain of the RET proto-oncogene in Hirschsprung's disease. Nature 1994;367:377-378.

6. Hofstra RM, Wu Y, Stulp RP, et al. RET and GDNF gene scanning in Hirschsprung patients using two dual denaturing gel systems. Hum Mutat 2000;15:418-429.

7. Amiel J, Laudier B, Attie-Bitach T, et al. Polyalanine expansion and frameshift mutations of the paired-like homeobox gene PHOX2B in congenital central hypoventilation syndrome. Nat Genet 2003;33:459-461.

8. Angrist M, Bolk S, Halushka M, et al. Germline mutations in glial cell line-derived neurotrophic factor (GDNF) and RET in a Hirschsprung disease patient. Nat Genet 1996;14:341-344.

9. Brooks AS, Bertoli-Avella AM, Burzynski GM, et al. Homozygous nonsense mutations in KIAA1279 are associated with malformations of the central and enteric nervous systems. Am J Hum Genet 2005;77:120-126.

10. Fernandez RM, Sanchez-Mejias A, Mena MD, et al. A novel point variant in NTRK3, R645C, suggests a role of this gene in the pathogenesis of Hirschsprung disease. Ann Hum Genet 2009;73:19-25.

11. Hofstra RM, Osinga J, Tan-Sindhunata G, et al. A homozygous mutation in the endothelin-3 gene associated with a combined Waardenburg type 2 and Hirschsprung phenotype (Shah-Waardenburg syndrome). Nat Genet 1996;12:445-447.

12. Hofstra RM, Valdenaire O, Arch E, et al. A loss-of-function mutation in the endothelin-converting enzyme 1 (ECE-1) associated with Hirschsprung disease, cardiac defects, and autonomic dysfunction. Am J Hum Genet 1999;64:304-308.

13. Pingault V, Bondurand N, Kuhlbrodt K, et al. SOX10 mutations in patients with Waardenburg-Hirschsprung disease. Nat Genet 1998;18:171-173.

14. Puffenberger EG, Hosoda K, Washington SS, et al. A missense mutation of the endothelin-B receptor gene in multigenic Hirschsprung's disease. Cell 1994;79:1257-1266.

15. Wakamatsu N, Yamada Y, Yamada K, et al. Mutations in SIP1, encoding Smad interacting protein-1, cause a form of Hirschsprung disease. Nat Genet 2001;27:369-370.

16. Burzynski GM, Nolte IM, Osinga J, et al. 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;12:604-612.

17. Burzynski GM, Nolte IM, Bronda A, et al. Identifying candidate Hirschsprung disease-associated RET variants. Am J Hum Genet 2005;76:850-858.

18. Fernandez RM, Boru G, Pecina A, et al. Ancestral RET haplotype associated with Hirschsprung's disease shows linkage disequilibrium breakpoint at -1249. J Med Genet 2005;42:322-327.

19. Fitze G, Cramer J, Ziegler A, et al. Association between c.135 G/A genotypes and RET proto-oncogene germline mutations and phenotype of Hirschsprung's disease. Lancet 2002;359:1200-1205.

20. Fitze G, Appelt H, Konig IR, et al. Functional haplotypes of the RET proto-oncogene promoter are associated with Hirschsprung disease (HSCR). Hum Mol Genet 2003;12:3207-3214.

21. Garcia-Barcelo MM, Sham MH, Lui VC, et al. Chinese patients with sporadic Hirschsprung's disease are predominantly represented by a single RET haplotype. J Med Genet 2003;40:e122.

22. Garcia-Barcelo M, Ganster RW, Lui VC, et al. TTF-1 and RET promoter SNPs: regulation of RET transcription in Hirschsprung's disease. Hum Mol Genet 2005;14:191-204.

23. Sancandi M, Griseri P, Pesce B, et al. 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;40:714-718.

24. Garcia-Barcelo MM, Tang CS, Ngan ES, et al. Genome-wide association study identifies NRG1 as a susceptibility locus for Hirschsprung's disease. Proc Natl Acad Sci U S A 2009;106:2694-2699.

25. Cantrell VA, Owens SE, Chandler RL, et al. Interactions between Sox10 and EdnrB modulate penetrance and severity of aganglionosis in the Sox10Dom mouse model of Hirschsprung disease. Hum Mol Genet 2004;13:2289-2301.

26. Carrasquillo MM, McCallion AS, Puffenberger EG, et al. Genome-wide association study and mouse model identify interaction between RET and EDNRB pathways in Hirschsprung disease. Nat Genet 2002;32:237-244.

27. Lang D, Epstein JA. Sox10 and Pax3 physically interact to mediate activation of a conserved c-RET enhancer. Hum Mol Genet 2003;12:937-945.

28. McCallion AS, Stames E, Conlon RA, et al. Phenotype variation in two-locus mouse models of Hirschsprung disease: tissue-specific interaction between Ret and Ednrb. Proc Natl Acad Sci U S A 2003;100:1826-1831.

29. Gabriel SB, Salomon R, Pelet A, et al. Segregation at three loci explains familial and population risk in Hirschsprung disease. Nat Genet 2002;31:89-93.

30. Brooks AS, Leegwater PA, Burzynski GM, et al. A novel susceptibility locus for Hirschsprung's disease maps to 4q31.3-q32.3. J Med Genet 2006;43:e35.

31. Bolk S, Pelet A, Hofstra RM, et al. 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 U S A 2000;97:268-273.

32. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559-575.

33. Sklar P, Smoller JW, Fan J, et al. Whole-genome association study of bipolar disorder. Mol Psychiatry 2008;13:558-569.

34. Price AL, Patterson NJ, Plenge RM, et al. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006;38:904-909.

35. Park B, Nguyen NT, Dutt P, et al. Association of Lbc Rho guanine nucleotide exchange factor with alpha-catenin-related protein, alpha-catulin/CTNNAL1, supports serum response factor activation. J Biol Chem 2002;277:45361-45370.

36. Heanue TA, Pachnis V. Expression profiling the developing mammalian enteric nervous system identifies marker and candidate Hirschsprung disease genes. Proc Natl Acad Sci U S A 2006;103:6919-6924.

37. Kiba T, Kintaka Y, Suzuki Y, et al. Ventromedial hypothalamic lesions change the expression of neuron-related genes and immune-related genes in rat liver. Neurosci Lett 2009;455:14-16.

38. Sunyaev S, Ramensky V, Koch I, et al. Prediction of deleterious human alleles. Hum Mol Genet 2001;10:591-597.

39. Anderson SL, Coli R, Daly IW, et al. Familial dysautonomia is caused by mutations of the IKAP gene. Am J Hum Genet 2001;68:753-758.

40. Slaugenhaupt SA, Blumenfeld A, Gill SP, et al. Tissue-specific expression of a splicing mutation in the IKBKAP gene causes familial dysautonomia. Am J Hum Genet 2001;68:598-605.

41. Slaugenhaupt SA, Gusella JF. Familial dysautonomia. Curr Opin Genet Dev 2002;12:307-311.

42. Mezey E, Parmalee A, Szalayova I, et al. Of splice and men: what does the distribution of IKAP mRNA in the rat tell us about the pathogenesis of familial dysautonomia? Brain Res 2003;983:209-214.

43. Azizi E, Berlowitz I, Vinograd I, et al. Congenital megacolon associated with familial dysautonomia. Eur J Pediatr 1984;142:68-69.

44. Bar-Shai A, Maayan C, Vromen A, et al. Decreased density of ganglia and neurons in the myenteric plexus of familial dysautonomia patients. J Neurol Sci 2004;220:89-94.

45. Antoniou AC, Beesley J, McGuffog L, et al. Common breast cancer susceptibility alleles and the risk of breast cancer for BRCA1 and BRCA2 mutation carriers: implications for risk prediction. Cancer Res 2010;70:9742-9754.

**Supplementary Table 1 -** Characteristics of the Chinese HSCR patients included in the study

| 500K n=173[a] | S-HSCR n=154 | | L-HSCR n=8 | | TCA n=10 | | Undetermined n=1 | |
|---|---|---|---|---|---|---|---|---|
| | Males | Females | Males | Females | Males | Females | Males | Females |
| | 126 {27} (9) | 28 {3} (3) | 2 | 6 {1} | 7 {2} | 3 | 0 | 1 {1} |
| With RET CDS mutation n=31 | S-HSCR n=28 | | L-HSCR n=1 | | TCA n=2 | | | |
| | Males | Females | Males | Females | Males | Females | | |
| | 21 (2) | 7 {2} (1) | 0 | 1 | 0 | 2 | | |
| Replication n=21 | S-HSCR n=14 | | L-HSCR n=4 | | TCA n=1 | | Undetermined n=2 | |
| | Males | Females | Males | Females | Males | Females | Males | Females |
| | 13 | 1 | 4{1} | 0 | 1 | 0 | 2 | 0 |

[a]number of patients after quality control filtering. { } additional anomalies; ( ) Down syndrome; U: undetermined

**Supplementary Table 2 -** Characteristics of the Dutch HSCR patients included in the study

| n=132 | S-HSCR n=84 | | L-HSCR n=24 | | TCA n=5 | | Undetermined n=19 | |
|---|---|---|---|---|---|---|---|---|
| | Males | Females | Males | Females | Males | Females | Males | Females |
| | 65 | 19 {3} | 19 | 5 | 4 | 1 | 18 | 1 |

{ } additional anomalies; U: undetermined

**Supplementary Table 3**

| SNP | Position (bp) | Allele | | T:U[a] | OR (95% CI) | *p*-value |
|---|---|---|---|---|---|---|
| | | A | B | | | |
| rs10978770 | 109050289 | G | A | 83:54 | 1.54 (1.09, 2.17) | 0.013 |
| rs7038659 | 109095510 | A | G | 15:33 | 0.45 (0.25, 0.84) | 0.0094 |
| rs180533 | 109124082 | T | C | 16:34 | 0.47 (0.26, 0.85) | 0.011 |
| rs12351693 | 110751336 | G | A | 20:8 | 2.50 (1.10, 5.68) | 0.023 |
| rs10979637 | 110784448 | C | A | 20:8 | 2.50 (1.10, 5.68) | 0.023 |
| rs12555920 | 110805558 | T | C | 20:9 | 2.22 (1.01, 4.88) | 0.041 |
| rs1333344 | 110816045 | A | G | 50:72 | 0.69 (0.48, 1.00) | 0.046 |
| rs1044905 | 110821316 | T | C | 48:72 | 0.67 (0.46, 0.96) | 0.028 |
| rs10119758 | 111324700 | A | G | 52:30 | 1.73 (1.11, 2.72) | 0.015 |
| rs9987721 | 111331292 | T | C | 52:30 | 1.73 (1.11, 2.72) | 0.015 |
| rs7038415 | 112157004 | A | C | 76:45 | 1.69 (1.17, 2.44) | 0.0048 |
| rs10816998 | 112173562 | G | A | 78:45 | 1.73 (1.20, 2.50) | 0.0029 |
| rs3010815 | 112473999 | T | C | 55:78 | 0.71 (0.50, 1.00) | 0.046 |
| rs2766998 | 112571406 | T | C | 32:52 | 0.62 (0.40, 0.96) | 0.029 |
| rs3780529 | 112601072 | C | A | 41:24 | 1.71 (1.03, 2.83) | 0.035 |
| rs384408 | 113584922 | A | G | 18:32 | 0.56 (0.32, 1.00) | 0.048 |
| rs1411681 | 114223965 | T | G | 10:21 | 0.48 (0.22, 1.01) | 0.048 |

A: Minor allele; B: Major allele; [a]: 'T and 'U' indicate transmitted and untransmitted minor allele of each SNP

**Supplementary Table 4 -** Replication of the 9 associated SNPs on Dutch cases and controls

| SNP ID | Position | | Minor allele | Major allele | MAF | | *p*-value |
|---|---|---|---|---|---|---|---|
| | | | | | Case | Control (HapMap CEU) | |
| rs7873793 | 110683696 | intron 31 | A | G | 0.20 | 0.19 | 0.84 |
| rs10816758 | 110688553 | intron 29 | G | A | 0.41 | 0.32 | 0.18 |
| rs10118853 | 110691843 | intron 28 | A | C | 0.20 | 0.19 | 0.81 |
| rs10979596 | 110696824 | intron 25 | G | T | 0.048 | 0.075 | 0.39 |
| rs10979597 | 110696853 | intron 25 | A | T | 0.048 | 0.075 | 0.39 |
| rs2230793 | 110699304 | exon 23 | G | T | 0.23 | 0.21 | 0.65 |
| rs2275630 | 110702228 | intron 19 | G | A | 0.00 | 0 | NA |
| rs10979607 | 110707646 | intron 14 | G | A | 0.048 | 0.08 | 0.33 |
| rs4369056 | 110708259 | intron 14 | T | C | 0.095 | 0.15 | 0.24 |

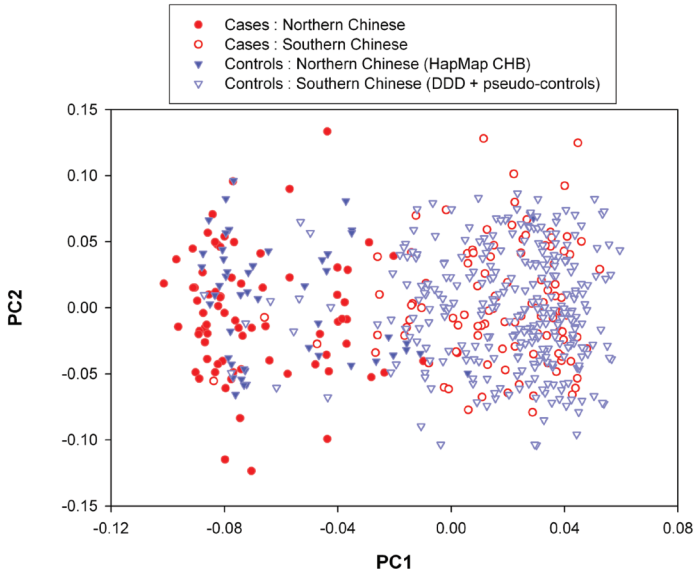



**Figure 1 -** Plot of principal components identified by EIGENSOFT.

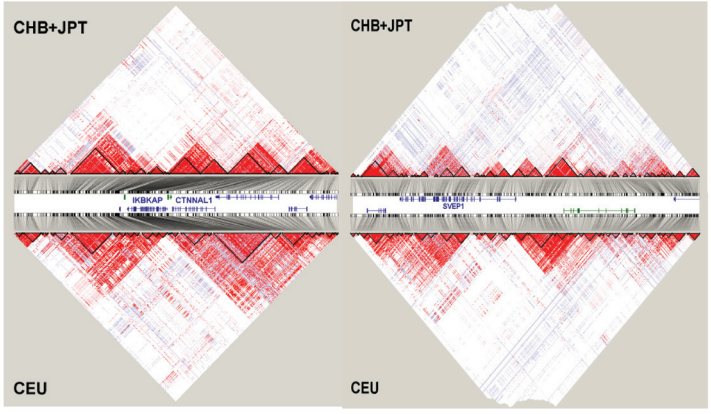



**Figure 2 -** LD structure across the 9q31A (right) and 9q31B (left) regions in CEU and CHB+JPT

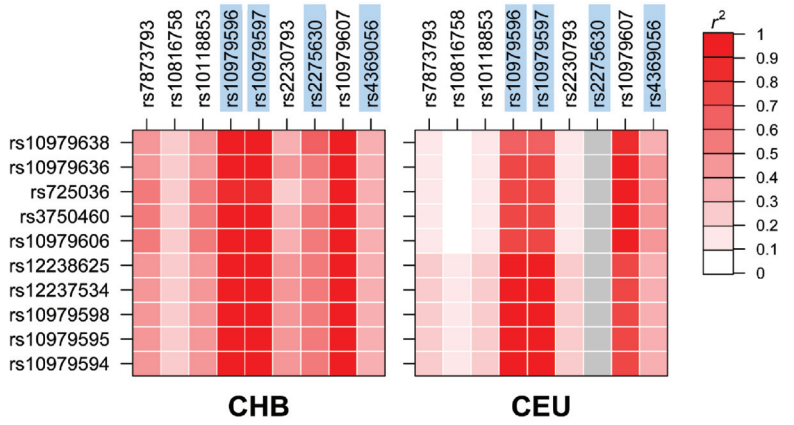



**Figure 3 -** LD structure of *IKBKAP* SNPs in CHB and CEU. The associated SNPs in Chinese which fail to replicate in Dutch are shown along the *x* axis (highlighted in blue: *p*-value passing multiple testing). SNPs in moderate to high LD in CHB (*y*-axis) can also be tagged effectively ($r^2$>0.8) by at least one of these nine associated SNPs (*x*-axis) in CEU, demonstrating the population-specific association. LD information for monomorphic SNP (rs2275630) in CEU is not shown.

# 3 Exome Sequencing in a Family with Hirschsprung Disease

Y. Sribudiani[1], N. Veldman[1], K.C. Rajendra[1], M.M. Alves[1], P. van der Vlies[1], W. Koetsier[2], M. Metzger[3], J. Osinga[1], L. Boven[1], C. van Diemen[2], M.A. Swertz[2], J.B.G. Verheij[1], I.T. Shepherd[4], A.S. Brooks[5], R.M.W. Hofstra[1]

[1] Department of Genetics, University Medical Center Groningen, University of Groningen, P.O. Box 30.001, 9700 RB Groningen, the Netherlands.
[2] Genomics Coordination Center, Department of Genetics, University Medical Center Groningen, University of Groningen, P.O. Box 30.001, 9700 RB Groningen, the Netherlands.
[3] Translational Centre for Regenerative Medicine, University of Leipzig, Philipp-Rosenthal-Strasse 55, 04103 Leipzig, Germany.
[4] Department of Biology, Emory University, Atlanta, USA
[5] Department of Clinical Genetics, Erasmus Medical Center, Westzeedijk 112, 3016 AH Rotterdam, the Netherlands

**manuscript in preparation**

## ABSTRACT

Hirschsprung disease is a congenital disorder characterized by the absence of enteric ganglia in a variable length of the intestinal tract. A linkage study performed on a large, multi-generational, Dutch family with five family members affected with Hirschsprung disease revealed significant linkage to 4q31.3-q32.3. We have sequenced the exomes of two patients from this family, focusing our search for mutations on this linkage region. We assumed an autosomal dominant model of inheritance with low penetrance, suggesting that the mutation in the linkage region is necessary but not sufficient to cause disease development. Exome sequencing revealed only one possible mutation in the linkage region, a single nucleotide change in exon 20 of the *LRB* gene. This variant was validated in all five affected family members. As this variant is located downstream of *MAB21L2*, a gene which plays a role in enteric neural crest cells (ENCCs) migration during ENS development, we hypothesized that this mutation might regulate *MAB21L2* expression in ENS, thereby disturbing ENS development. However, we identified this variant in four individuals out of 220 screened controls and once in 87 Hirschsprung patients. It is therefore unlikely that it is disease-associated. As this was the only variant we identified in the linkage region, we must have missed the causative mutation, which is quite possible since the coverage of the coding sequences, in general, was only approximately 65%.

# INTRODUCTION

Hirschsprung (HSCR) disease is a congenital disorder characterized by the absence of enteric ganglia in a variable length of the intestinal tract. Based on the length of the aganglionic region, HSCR is classified into three types: short-segment (S-HSCR), long-segment (L-HSCR) and total colonic aganglionis (TCA). The estimated incidence of HSCR varies in different populations; it is lowest in Caucasians (1:5000) and highest in Asians (1.4:5000). Mostly, HSCR occurs sporadically and most cases are diagnosed with S-HSCR and only in a minority of cases (5-15%) is more than one patient seen in a family (familial HSCR).[1]

The pattern of inheritance of sporadic cases can be multifactorial, recessive or dominant with low penetrance, while familial cases can be either recessive or autosomal dominant, in both cases with incomplete penetrance[2]. So far, 12 genes and five loci have been implicated in HSCR and *RET* is considered to be the major gene in HSCR etiology. The 12 genes are: *RET* [3], *EDNRB* [4], *EDN3* [5, 6], *GDNF* [7, 8], *NTN* [9], *SOX10* [10, 11], *PHOX2B* [12], *ECE1* [13], *KIAA1279/KBP* [14], *ZFHX1B* [15, 16], *TTF-1/TITF1* [17, 18] and *NRG1* [19]. Besides *RET,* most other genes are involved in what is called syndromic-HSCR, HSCR patients who also suffer from other congenital malformations besides HSCR. Altogether, mutations in these 12 genes explain no more that ~20% of the total heritability, suggesting that other disease associated genes must exist.

In a previous linkage study, Brooks and colleagues identified a linkage region on 4q31.3-q32.3 in a multi-generational, Dutch HSCR family with five affected members.[20] The pattern of inheritance in this family seems to be autosomal dominant with low penetrance, indicating that the mutation in this region is necessary but not sufficient to cause the disease. It therefore seemed likely that in addition to the chromosome 4 mutation, other modifier genes or variants contribute to the development of the disease. This 4q linkage region contains at least 57 genes. From a functional point of view, Brooks *et al.* picked the *MAB21L2* gene as the most promising candidate, however no coding mutation was found in this gene.

We have now sequenced all exons (the exome) of two affected family members from different branches of the family in an attempt to identify the disease-associated mutation in the linkage region. Furthermore, we looked for possibly other modifier mutations elsewhere in the coding regions in the genome.

# MATERIALS AND METHODS

## Patients and DNA Samples

DNA from the five affected individuals, from 220 normal controls and from 87 HSCR patients was isolated from peripheral blood leucocytes using a standard protocol. For exome sequencing, we selected two affected members from different branches of the family (V-1 and V-4, see Figure-1). Validation of some identified variants by Sanger sequencing was performed in these two individuals and on three more affected family members (IV-3, V-2 and V-3).

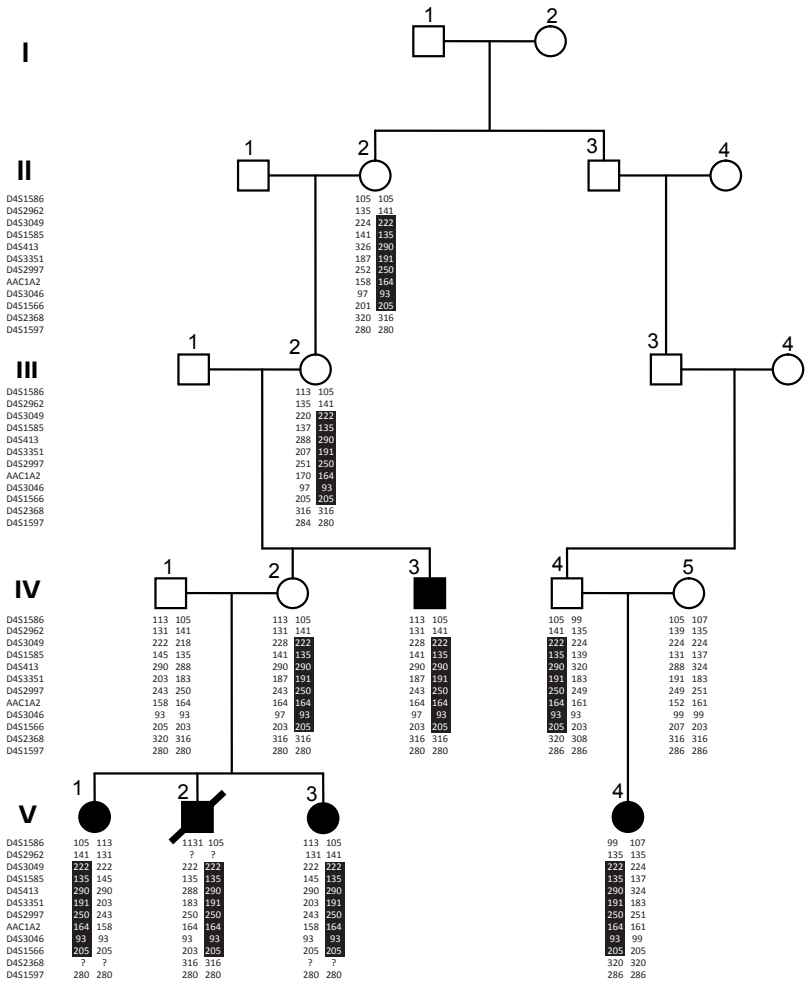



**Figure 1 -** Pedigree structure and haplotypes. Patients are represented as black symbols. The segregating haplotype, 4q31-q32, is depicted. The minimum critical region at chromosome 4 spans 11.7 cM between D4S3049 and D4S1566 (adapted from Brooks et al. 2006).

## Exome Sequencing and Pipeline Data Analysis

Five microgram of DNA from each individual was used for exome sequencing. For exome capturing and enrichment, we used the hybrid capture SureSelect Human All Exon kit (Agilent Technologies, Amsterdam, the Netherlands) which targets most human exons (approximately 38 Mb). The captured fragments were sequenced on a Genome Analyzer II (Illumina, San Diego, USA). The alignment of the sequencing results, the SNP calling and data filtering was performed using a pipeline set up by the Genomics Coordination Centre (GCC), hosted by the Department of Genetics, UMCG, Groningen. The pipeline is based on the genome analysis toolkit (GATK) version 1.0.4418 developed at the Broad institute (Boston, USA). dbSNP v1.29 is incorporated into this pipeline for filtering of known SNPs. The SNPs that passed the threshold of all the criteria of the pipeline and that were unknown in dbSNP v1.29 were further filtered by checking whether they were known in a later version of dbSNP (dbSNP v1.31) or mentioned in the 1000 Genomes database. SNPs were further checked for their presence in both the sequenced patients and in an unrelated sample that was run at the same time on the Genome Analyzer II (Illumina). The list of variants of this unrelated sample we call "in-house controls". This last step was incorporated into the whole filtering process to reduce the number of false-positives introduced during the enrichment and sequencing processes. As the pattern of inheritance is autosomal dominant, we focused on searching for heterozygous mutations.

## Deletion-Insertion Polymorphism (DIP) Analysis

Deletion-insertion polymorphism (DIP) analysis was performed using a program called CLC bio (CLCbio, Muehltal, Germany). It was performed only for chromosome 4. Again, as the pattern of inheritance seems to be autosomal dominant, we focused on searching for heterozygous DIPs and we set the minimum variant frequency at 30%, with a window length of 11 bp, a maximum number of gaps and mismatches of 2, and a coverage of at least 10.

## Validation of Candidate Variants (PCR and Sequencing)

PCR reactions were performed in a final volume of 15 μl, containing 100 ng of DNA, 1.5 μl of 10X PCR buffer (Invitrogen, Leek, the Netherlands), 15mM $MgCl_2$, 0.15 μl of 25mM dNTPs, 1 unit of Taq Polymerase (Invitrogen) and 1 μl of 20 pmol/μl of each primer. The following primers were used for *LRBA* 5'CCACATAACTTAAGGTTGATTC3' and 5'GATATAAGGAGATGTGGCTG3'; for *ARMC9* 5' TGGCCAGGTTTGTCTTGC3' and 5' GGCTTCCTAAAGGCCTCTTG3'; for *CD27* 5' CTGGGATTACAGGCATGAG3' and 5' GCCTGAGACAGTTCTTACG3'. The PCR conditions were as follows: pre-heating 1 min at 95°C, followed by 35 cycles of 45 seconds at 95°C, 45 seconds at 60°C and 45 seconds at 72°C. The

final elongation step was 10 min at 72°C. PCR purification was performed by using EXOSAP IT according to the manufacture's manual (GE Healthcare, Eindhoven, the Netherlands). For Sanger sequencing we used 5 μl of purified PCR product.

# RESULTS

## Exome Sequencing

A total of 2.9-3.0 GB of paired-end sequence data was generated per individual, with a mean read length of 101 bps. About 84.4% of bases were mapped to the target exons and, on average, 64.8% of the exome was covered at least 10 times. A summary of the exome sequencing performance is shown in Table 1.

13,403 and 14,637 variants per individual passed the pipeline's thresholds for SNP calling. After filtering with dbSNP build 129, version 131 and the 1000 Genome database, using the SeattleSeq annotation, 501 and 544 novel variants were identified per individual. To identify potentially pathogenic variants, we focused only on missense, splice-site, and nonsense mutations also insertions and deletions (indels). An overview of our SNP calling results is shown in Table 2. DIP analysis on chromosome 4 did not identify any heterozygous deletion or insertion.

**Table 1 -** Summary of the statistics of exome sequencing analysis of two HSCR patients.

| | Patient V-1 | Patient V-4 |
|---|---|---|
| Raw data yield (bp) | 2.974.801.379 | 3.023.364.300 |
| Target region (bp | 38.925.539 | 38.925.539 |
| Base-pairs mapped to genome | 2.325.231.999 | 2.336.642.676 |
| Base-pairs mapped to target region | 1.303.155.126 | 1.382.373.365 |
| Base-pairs mapped to target and flanking regions | 2.036.529.862 | 2.065.636.143 |
| Mean read length (bp) | 101 | 101 |
| Mean depth of target region | 37.34 | 38.71 |
| Coverage of target region | 33.47 | 35.51 |
| Fraction of targets covered ≥ 2X (%) | 80.14 | 83.29 |
| Fraction of targets covered ≥ 10X (%) | 62.85 | 66.70 |
| Fraction of targets covered ≥ 20X (%) | 50.39 | 53.70 |
| Fraction of targets covered ≥ 30X (%) | 40.29 | 42.82 |
| Capture specificity (%) | 43.80 | 45.72 |

**Table 2 -** Overview of the SNP calling and filtering steps in two HSCR patients.

| | Patient V-1 | Patient V-4 |
|---|---|---|
| **Variants identified (pipeline)** | **13,403** | **14,637** |
| **Unknown SNPs (in dbSNP129)** | **743** | **859** |
| **Unknown SNP in dbSNP131 and the 1000 genome database** | **501** | **544** |
| *coding-synonymous | 115 | 102 |
| *coding-not Mod 3 | 1 | 1 |
| *intergenic | 24 | 20 |
| *intron | 149 | 172 |
| *missense | 196 | 210 |
| *nonsense | 6 | 11 |
| *near-gene | 1 | None |
| *splice-sites (SS) | 1 | 3 |
| *utr-3 | 4 | 19 |
| *utr-5 | 4 | 6 |
| **Missense+SS+nonsense variants** | **203** | **224** |
| Missense+SS+nonsense variants present in both patients | 10 | 10 |
| Variants - probably damaging (Polyphen) | 3 | 3 |
| Variants - benign (Polyphen) | 6 | 6 |
| Variants - unknown (Polyphen) | 2 | 2 |
| No. of genes in overlapped variants | 9 | 9 |
| Private missense + SS + nonsense variants | 193 | 214 |

## Selection of Possible Disease-Causing SNPs

As the two individuals we sequenced are related they should have mutations in common. Ten new variants proved to be present in both individuals (genome-wide) (Table 3). Three of them, however, were also identified in the "in-house controls" and were therefore not considered as possible disease-associated candidate variants. Three of the seven remaining SNPs that were still possibly disease-causing were predicted as probably damaging in Polyphen-2 (Polymorphism Phenotyping v2). One of the three variants was in the linkage region, in exon 20 of *LRBA* (LPS Responsive beige-like Protein Kinase A anchor). A second possibly damaging variant was located in exon 19 of *ARMC9* (chromosome 2) and the third was located in exon 1 of the *OR1F1* gene (chromosome 16).

## Validate of Candidate Variants

To confirm a role for these missense mutations in the disease process, we screened all five affected individuals for the presence of these variants The variant in exon 20 of the *LRBA* gene was confirmed in all five affected family members, while the variant in *ARMC9* was confirmed in three of the five (V-1, V-3 and V-4). The variant on *CDC27* was not confirmed in any of the five individuals, showing that an allele-balance (AB) threshold less than 0.6 is crucial to defining the real variant and notifying a false-positive one. For this reason we did not validate the remaining four candidate variants as all of them had an AB value of more than 0.6. The variant in exon 20 of *LRBA* is a missense mutation that changes an Asparginine to a Serine (N > S). This amino acid (N) is conserved in humans, mouse and zebrafish (Figure 2A).

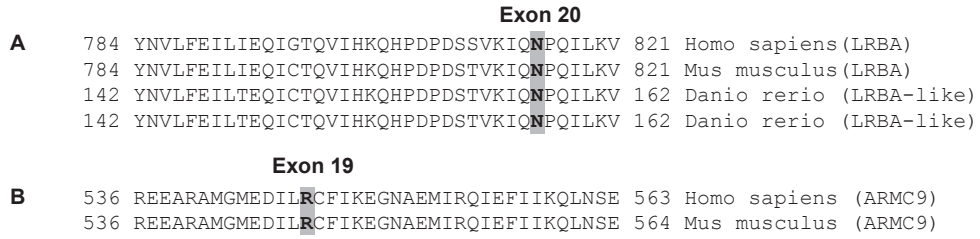



**Figure 2 -** A) Comparative alignment of part of the protein seqeunce of LRBA (using human, mouse and zebrafish data). The amino acid encoded by the *LRBA* variant (N > S) is colored in grey. B) Comparative alignment of part of the ARMC9 amino acid sequence of human and mouse. The amino acid encoded by the *ARMC9* variant (R > C) is colored in red. Zebrafish has a predicted ortholog of ARMC9, however the precise protein sequence is not yet available.

**Table 3 -** Summary of candidate variants.

| Chr | Position | Ref | Variant | Mutation | Amino Acid Changes | Polyphen-2 | Gene Symbol | Case V-1 | | Case V- 4 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | C | AB | C | AB |
| 2 | 232156084 | C | T | missense | ARG > CYS | PD | ARMC9 | 20 | 0.4 | 21 | 0.48 |
| 4 | 151791682 | T | C | missense | ASN > SER | PD | LRBA | 38 | 0.42 | 36 | 0.45 |
| 7 | 142460335 | A | G | missense | LYS > GLU | Benign | PRSS1 | 30 | 0.7 | 45 | 0.64 |
| 7 | 142460369 | G | A | missense | SER > ASN | Benign | PRSS1 | 27 | 0.77 | 50 | 0.72 |
| 10 | 135438967 | C | T | missense | ARG > GLN | PD | FRG2B * | 15 | 0.6 | 37 | 0.54 |
| 11 | 123814439 | A | G | missense | ILE >THR | Benign | OR6T1 | 31 | 0.7 | 32 | 0.34 |
| 16 | 3254293 | G | A | missense | GLY > GLU | PD | OR1F1 | 11 | 0.64 | 13 | 0.54 |
| 17 | 16068396 | G | A | missense | SER > LEU | Benign | NCOR1* | 42 | 0.67 | 30 | 0.67 |
| 17 | 44144993 | C | G | missense | ARG > PRO | Benign | KIAA1267 * | 40 | 0 | 38 | 0 |
| 17 | 45234303 | G | C | missense | ALA > GLY | Benign | CDC27 | 55 | 0.6 | 51 | 0.69 |

* Present in “in-house controls”

Chr: Chromosome, Ref: Reference, PD: Probably damaging, C: Coverage, AB: Allele Balance = Ref.seq/(Ref.Seq+Variant.Seq)

### *LRBA* Variant in Controls and HSCR Patients

We checked for the prevalence of the *LRBA* variant in a control population and in a HSCR population. The HSCR patients that were screened do not carry a *RET* coding sequence mutation. The *LRBA* variant was identified heterozygously in four of the 220 healthy individual (4/440 chromosomes: 0.9%) and heterozygously in one of the 87 HSCR patient (1/174 chromosomes: 0.57%).

### *ARMC9* Variant in Controls

We checked for the prevalence of *ARMC9* variant in a control population and we identified the variant heterozygously in six healthy individual (6/296 chromosomes: 2%).

## DISCUSSION

We identified a possible disease-associated variant in exon 20 of the *LRBA* gene in two cases from a large, multi-generational HSCR family on which we performed exome sequencing. We confirmed the presence of this variant in all five affected family members by Sanger sequencing. However, this variant proved to be present in the control population (0.9%, 4/440 chromosomes), and as the incidence rate of this variant is lower in our HSCR population (0.5%, 1/174 chromosomes) compared to the control population, it is unlikely that the variant is a strong HSCR-predisposing mutation or even a modifier gene for HSCR. A second variant was identified in *ARMC9.* The variant was identified in three of the five affected members (V-1, V-3 and V-4). However, the variant was also present in the control population (0.57%, 6/298 chromosomes) which again makes it unlikely that this variant is a disease-associated.

As the *LRBA* variant was the only variant we found on chromosome 4 that was present in both the exome-sequenced cases, we assume that we have probably missed the causative variant that all the cases in this family should have in common on chromosome 4. This was likely due to technical issues. One of the criteria we used in SNP calling was a minimum coverage of 10. The statistical data presented in Table 1 show that only 62.8% and 66.70% of the target sequences were covered with a minimum number of 10 reads. In fact, we found that when we lowered the coverage to 2, we still only had a coverage of 80.14% and 83.29%, respectively. In short, this means that by taking a coverage of 10, we do miss mutations in almost 40% of the target sequences (Table 1). Furthermore, the enrichment kit used in this experiment covers 81.33% of all human exons. So, in addition to the missed 40% of target sequences, we also missed about 19% of all exomic data. These figures led us to conclude that we might well have

missed the real disease-causing mutation.

Although we did not find any common mutation on chromosome 4, we wondered whether we could still select the best possible candidate gene associated with HSCR based on function in order to perform additional analysis. This led us to *MAB21L2*, which was mentioned earlier. *MAB21L2* is expressed in the neural crest derivatives, including the progenitor cells of the neurons and the glial of the enteric nervous system.[21] MAB21L2 is a downstream target of the TGF-β/SMADs signaling pathway, which is important for ENS development. Knock-down of the MAB21L2 ortholog in zebrafish showed that subsequent ENCCs migration along the intestine is perturbed and a reduction of cell differentiation was seen, all indicating that normal MAB21L2 function is required for normal ENCCs migration during ENS development (I. Sheperd, personal communication).

As mentioned earlier, *MAB21L2* has already been sequenced and no coding mutation was found.[20] We therefore decided to extend our sequencing efforts to the promoter region, and in fact we sequenced 16 Kb upstream of *MAB21L2* in all five affected family members. However, again, no mutation was identified (data not shown). One could argue that our analysis more or less excluded *MAB21L2* as the best candidate HSCR gene. However, this might not be completely true as will become clear from the following reasoning.

*MAB21L2* is located within the *LRBA* gene, shown in Figure 3. The *MAB21L2* is transcribed from the anti-sense strand of *LRBA* gene. This unique nested gene pair (*LRBA* and *MAB21L2)* is conserved in mouse, chick, frog and zebrafish.

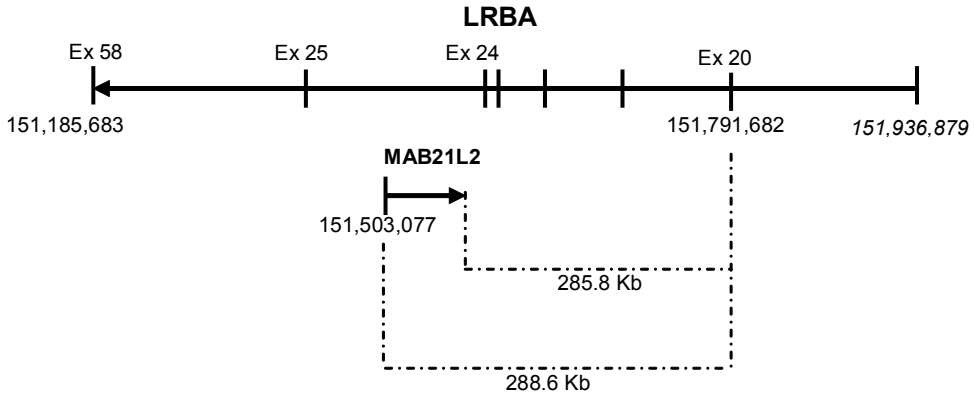



**Figure 3 -** Schematic overview of the *LRBA* and the *MAB21L2* genes and their position in the genome. The variation in exon 20 of *LRBA* is located 288.6 Kb downstream of the start site of *MAB21L2* gene.

Moreover Tsang and colleagues identified five paralogous, non-coding elements (pNEs) that are scattered throughout *LRBA*. They are located both upstream and downstream of *MAB21L2* and could regulate the expression of a reporter gene in a tissue-specific manner during mouse embryonic development, comparable to the expression patterns of MAB21L2.[22] Mutations in the LRBA locus might also have a regulatory effect on the gene that is transcribed from the anti-sense strand, namely *MAB21l2*. As we screened only part of the *LRBA/MAB21L2* locus, we cannot exclude that there are mutations in the locus that have an influence on *MAB21L2* expression.

We therefore propose to re-sequence the affected family members, but with a higher coverage and for a larger targeted sequence, including the entire LRBA genomic region (thereby including the multiple species conserved region that are involved in expression regulation of *MAB21L2*)

## REFERENCES


1. Passarge E. The genetics of Hirschsprung's disease. Evidence for heterogeneous etiology and a study of sixty-three families. N Engl J Med 1967;276:138-143.

2. Badner JA, Sieber WK, Garver KL, et al. A genetic study of Hirschsprung disease. Am J Hum Genet 1990;46:568-580.

3. Romeo G, Ronchetto P, Luo Y, et al. Point mutations affecting the tyrosine kinase domain of the RET proto-oncogene in Hirschsprung's disease. Nature 1994;367: 377-378.

4. Puffenberger EG, Hosoda K, Washington SS, et al. A missense mutation of the endothelin-B receptor gene in multigenic Hirschsprung's disease. Cell 1994;79: 1257-1266.

5. Edery P, Attie T, Amiel J, et al. Mutation of the endothelin-3 gene in the Waardenburg-Hirschsprung disease (Shah-Waardenburg syndrome). Nat Genet 1996;12:442-444.

6. Hofstra RM, Osinga J, Tan-Sindhunata G, et al. 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: 445-447.

7. Angrist M, Bolk S, Halushka M, et al. Germline mutations in glial cell line-derived neurotrophic factor (GDNF) and RET in a Hirschsprung disease patient. Nat Genet 1996;14:341-344.

8. Salomon R, Attie T, Pelet A, et al. Germline mutations of the RET ligand GDNF are not sufficient to cause Hirschsprung disease. Nat Genet 1996;14:345-347.

9. Doray B, Salomon R, Amiel J, et al. Mutation of the RET ligand, neurturin, supports multigenic inheritance in Hirschsprung disease. Hum Mol Genet 1998;7:1449-1452.

10. Pingault V, Bondurand N, Kuhlbrodt K, et al. SOX10 mutations in patients with Waardenburg-Hirschsprung disease. Nat Genet

1998;18:171-173.

11. Paratore C, Eichenberger C, Suter U, Sox10 haploinsufficiency affects maintenance of progenitor cells in a mouse model of Hirschsprung disease. Hum Mol Genet 2002; 11:3075-3085.

12. Amiel J, Laudier B, Attie-Bitach T, et al. Polyalanine expansion and frameshift mutations of the paired-like homeobox gene PHOX2B in congenital central hypoventilation syndrome. Nat Genet 2003;33:459-461.

13. Hofstra RM, Valdenaire O, Arch E, et al. A loss-of-function mutation in the endothelin-converting enzyme 1 (ECE-1) associated with Hirschsprung disease, cardiac defects, and autonomic dysfunction. Am J Hum Genet 1999;64:304-308.

14. Brooks AS, Bertoli-Avella AM, Burzynski GM, et al. Homozygous nonsense mutations in KIAA1279 are associated with malformations of the central and enteric nervous systems. Am J Hum Genet 2005;77:120-126.

15. Cacheux V, Dastot-Le Moal F, et al. Loss-of-function mutations in SIP1 Smad interacting protein 1 result in a syndromic Hirschsprung disease. Hum Mol Genet 2001;10:1503-1510.

16. Wakamatsu N, Yamada Y, Yamada K, et al. Mutations in SIP1, encoding Smad interacting protein-1, cause a form of Hirschsprung disease. Nat Genet 2001;27: 369-370.

17. Garcia-Barcelo M, Ganster RW, Lui VC, et al. TTF-1 and RET promoter SNPs: regulation of RET transcription in Hirschsprung's disease. Hum Mol Genet 2005;14: 191-204.

18. Garcia-Barcelo MM, Lau DK, Ngan ES, et al. Evaluation of the thyroid transcription factor-1 gene (TITF1) as a Hirschsprung's disease locus. Ann Hum Genet 2007; 71:746-754.

19. Garcia-Barcelo MM, Tang CS, Ngan ES, et al. Genome-wide association study identifies NRG1 as a susceptibility locus for Hirschsprung's disease. Proc Natl Acad Sci U S A 2009;106:2694-2699.

20. Brooks AS, Leegwater PA, Burzynski GM, et al. A novel susceptibility locus for Hirschsprung's disease maps to 4q31.3-q32.3. J Med Genet 2006;43:e35.

21. Mariani M, Baldessari D, Francisconi S, et al. Two murine and human homologs of mab-21, a cell fate determination gene involved in Caenorhabditis elegans neural development. Hum Mol Genet 1999;8:2397-2406.

22. Tsang WH, Shek KF, Lee TY, et al. An evolutionarily conserved nested gene pair - Mab21 and Lrba/Nbea in metazoan. Genomics 2009;94:177-187.

# 4 Variants in *RET* Associated With Hirschsprung Disease Affect Binding of Transcription Factors and Gene Expression

Yunia Sribudiani[1], Marco Metzger[2,3], Jan Osinga[1], Amanda Rey[1], Alan.J.Burns[2], Nikhil Thapar[2], Robert.M.W.Hofstra[1,*]

[1] Department of Genetics, University Medical Center Groningen, University of Groningen, P.O. Box 30.001, 9700 RB Groningen, the Netherlands.
[2] Gastroenterology and Neural Development Units, Institute of Child Health, University College London, London, United Kingdom.
[3] Translational Centre for Regenerative Medicine, University of Leipzig, Philipp-Rosenthal-Strasse 55, 04103 Leipzig, Germany.

**Gastroenterology, 2011 Feb; 140(2):572-582.e2**

## ABSTRACT

**Background & Aims:** Two non-coding *RET* variations, the T allele of SNP rs2435357 (Enh1:C>T) and the A allele of SNP rs2506004 (Enh2:C>A), are strongly associated with Hirschsprung disease (HSCR) susceptibility. These SNPs are in strong linkage disequilibrium and located in an enhancer element in intron 1 of the *RET* gene. For Enh1 it has been shown that the disease-associated T allele results in reduced expression in Luciferase experiments, via disturbed SOX10 binding, compared to the non-disease associated C allele. We aimed to determine whether Enh2-A is also a functional variant.

**Methods:** We evaluated the function of Enh2-A by Luciferase assays of constructs containing both alleles, separately or in combinations. We performed *in silico* analysis to identify a possible transcription activator or repressor that binds to Enh2-C (non-disease associated variant).

**Results:** *Luciferase* assays showed that not only the Enh1-T allele but also the Enh2-A allele decreased Luciferase expression. In silico analysis identified the sequences of Enh2-C and its surroundings sequences (-A**C**GTG-) as a potential binding site for the NXF/ARNT2 and SIM2/ARNT2 heterodimers. The binding affinity of NXF/ARNT2 to Enh2-C was confirmed by Electrophoresis Mobility Shift Assays and Supershift. Transfections of NXF/ARNT2 or SIM2/ARNT2 into neuroblastoma cell lines increased and decreased *RET* expression, respectively.

**Conclusions:** Our results show that more than one SNP on an associated haplotype can influence disease development. Our data strongly implicate the involvement of NXF, ARNT2 and SIM2 in *RET* expression regulation, thereby contributing to HSCR development.


**Abbreviation**: CME, central nervous system midline enhancer; ENS, enteric nervous system; HSCR, Hirschsprung disease; MCS, multi-species conserved sequences; NLBs, neurosphere-like bodies; NPE, nuclear protein extract; NCSCs, neural crest stem cells; SNP, single nucleotide polymorphism


**Keywords**: Hirschsprung disease, RET expression, Down Syndrome

# INTRODUCTION

Loss-of-function mutations in the *RET* gene are associated with Hirschsprung's disease (HSCR), a congenital disorder characterized by the absence of ganglia in the distal part of the gut. Genetic dissection of HSCR has led not only to the identification of mutations in *RET*, but also to coding mutations in 10 other genes. Of these eleven genes, *RET* is believed to be the major factor in the etiology of HSCR. Coding sequence mutations in *RET* have been found in up to 50% of familial and 15 to 35% of sporadic cases of HSCR.[1] However, genome-wide linkage analysis on multigenerational HSCR families and sibpairs studies on smaller HSCR families showed that in almost all of the families, *RET* seems to play a role[2, 3]. This led to the hypothesis that non-coding RET mutations likely play a role in these familial cases. It further led us to hypothesize that non-coding variants may also play a role in the sporadic cases.

Association studies carried out on several European and Asian patient populations, mostly without *RET* mutations, indeed showed a very strong association with a specific *RET* haplotype (Figure 1)[1, 4-15]. In fact, regardless of the population, the same haplotype, spaning ~27 Kb (4 Kb of the 5' UTR and exon 1; 23 Kb of intron 1 and exon 2) is present in most HSCR patients. It is now believed that the mutation(s) on this haplotype are probably present in intron 1, within an enhancer region[13, 14]. In both zebra fish and mice it was shown that the reporter gene expression patterns driven by this enhancer element (called MCS+9.7) are similar to *RET*, both in a tissue and in a time-wise manner[16, 17]. There are two SNPs in this enhancer sequence: rs2435357 (Enh1:C>T) (also called RET+3)[13] and rs2506004 (Enh2:C>A) (also called IVS+9494)[14], of which the Enh1-T and the Enh2-A alleles both are strongly associated with HSCR[13, 14]. Emison proposed that Enh1-T was most likely the disease-associated allele based on conservation and functional arguments[14 18]. However, the Enh2-A allele, an allele in strong linkage disequilibrium with Enh1-T (r = 0.969), was also present in the constructs used for the functional interpretation of Enh1-T[14]. Thus we cannot totally exclude that this Enh2-A allele also contributes to disease development. We therefore investigated the role of this SNP in *RET* expression regulation.

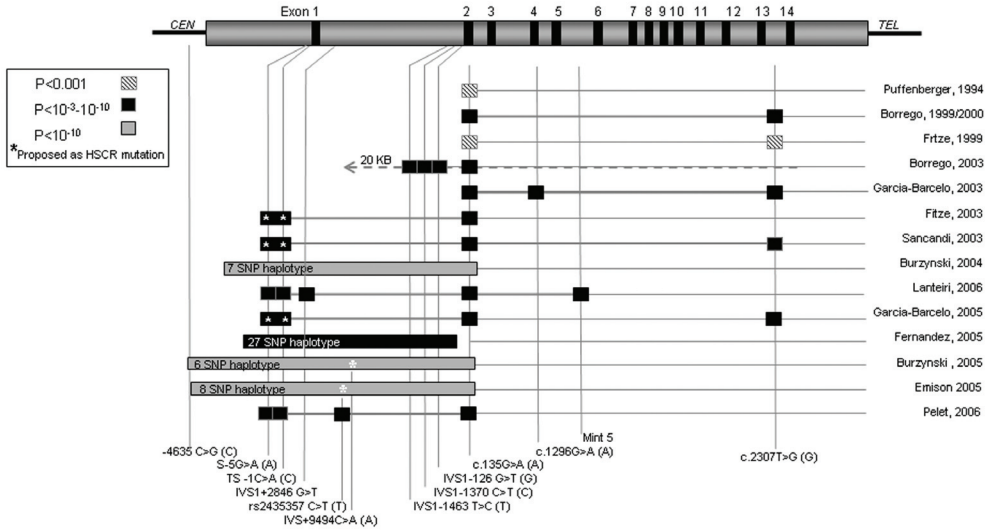



**Figure 1 -** An overview of the reported Hirschsprung (HSCR) disease-associated SNPs

# MATERIALS AND METHODS

## Construction of Vectors

The enhancer elements, containing one or both HSCR-associated SNPs, were PCR amplified using patients DNA as template, and cloned into KpnI/XhoI sites of a pGL3-*Luciferase* construct already containing the RET promoter or the SV40 promoter (Promega, Madison, WI, USA). Constructs containing a combination of non-disease- and disease-associated variants for both SNPs were generated in RETprom+Enh1-C/Enh2-C by site-directed mutagenesis (QuickChange XL Site-Directed Mutagenesis kit, Stratagene, Texas, USA). For primers see Supplementary Table 2. The sequences of the inserts were checked by direct sequencing.

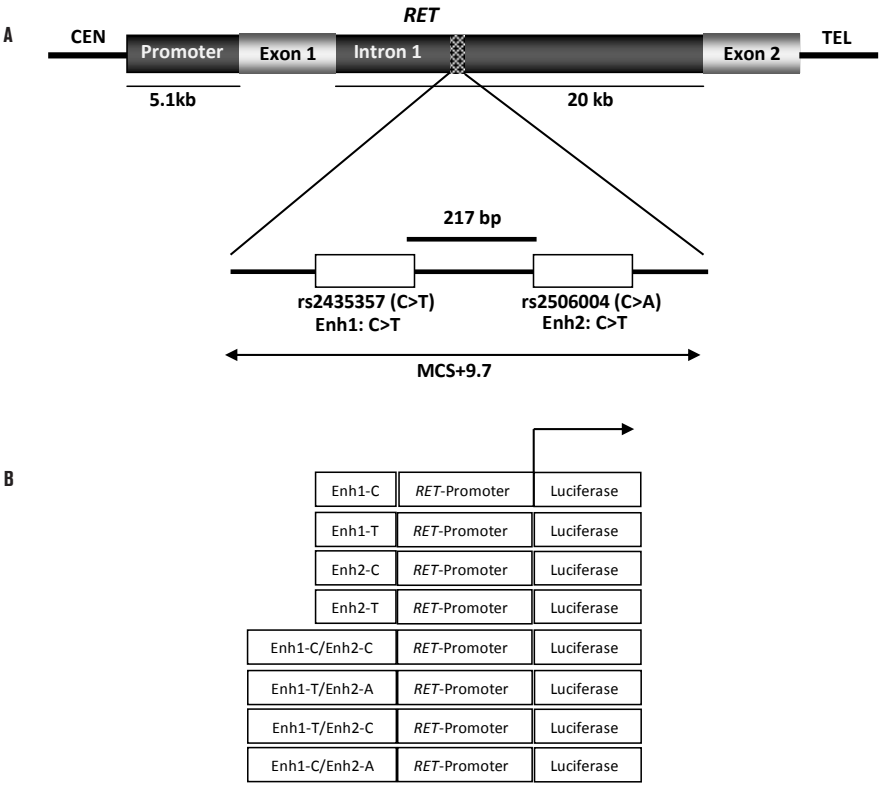



**Figure 2 -** Schematic overview of the *RET* gene and the constructs used in this study. (A) SNP rs2435357 (Enh1:C>T) and SNP rs2506004 (Enh2:C>A) are both located in a conserved region of intron 1, called MCS+9.7, separated by 217 bp. (B) RET and SV40 promoter constructs coupled to *Luciferase* and enhancer sequences used in this study.

## Luciferase Assays

Transient transfections were performed using a mouse neuroblastoma cell line (Neuro-2A) which was cultured according to the ATCC protocol. $0.5 \times 10^6$ Neuro2A cells were plated in 35 mm dishes for 48 hours prior to transfection. Cells were co-transfected with 100 ng of pRL-SV40-*Renilla Luciferase* (Promega, Madison, USA) as an internal control and 1 mg of the reporter constructs containing different enhancer elements (Figure 2B) using 4 ml of Lipofectamine™2000 (Invitrogen, Carlsbad, CA, USA) according to the manufacture's protocol. Luciferase activity was analyzed with a Dual-Luciferase-Reporter Assay (Promega) 24 hours after transfection. The ratios of Firefly Luciferase compared to Renilla Luciferase (normalized-Luciferase unit) were used to determine the activity of each promoter and enhancer element.

The results presented in the graphs are fold changes of normalized-Luciferase expression of each construct to the control (pGL3-RET or pGL3-SV40 promoter). Activation or repression of Luciferase expression by NXF/ARNT2 and SIM2/ARNT2, respectively were performed by co-transfection the *NXF*/*ARNT2* (1mg/500ng) or *SIM2*/*ARNT2* (1mg/500ng) with 1 mg of reporter construct containing different enhancer elements. In all cases the amount of the transfected DNA was the same, if necessary this was achieved by adding pCDNA3. The assays were performed in three, independent, triplicate experiments. For the dosages effect experiment, three different dosages of *NXF*/*ARNT2* or *SIM2*/*ARNT2* (100/50 ng, 500/250ng and 1 mg/ 500 ng) were used.

## Statistical Analysis

Statistical analysis of the Luciferase assay results was performed using the one-way ANOVA test. The results represent the mean ± S.E.M. Statistical analysis was performed by GraphPad Prism 4. P-values of $< 0.05$ were considered significant.

## Protein Analysis

Western blotting was performed to check the level of RET expression in the Neuro2A cell line, after *NXF*/*ARNT2* or *SIM2*/*ARNT2* transfections. Transient transfection was performed as described above. For the mock control, cells were only transfected with 1.5 mg pcDNA3. After 24 hours cells were harvested and total cell lysates were isolated using CellLytic™ (Sigma Aldrich, Saint Louis, USA) according to the manufacture's protocol. Protein concentrations were determined using the BCA protein assay (Pierce Biotechnologies, Rockford, IL, USA) and measured on a NanoDrop® ND-1000 (Thermo Scientific, Waltham, MA, USA). 30 mg of cell lysate was used for Western blotting. Primary antibodies anti-RET, and anti-b-Actin were applied and visualized with the respective IgG-HRP secondary antibodies (see supplementary Table 3).

## Neural Crest Stem Cells (NCSCs) Isolation and Culture

C57BL/6 mouse embryos E11.5, E12.5, E13.5, E14.5 and E15.5, were sacrificed and the gut (jejunum to rectum) was used for the isolation of NCSCs. The gut was mechanically dissected and tissues were incubated in collagenaseXI /dispaseII enzyme solution (collagenase XI [750U/ml], Sigma Aldrich, Saint Louis, USA; dispase II [250 mg/ml], Roche, IN, USA) in PBS, incubated at 37°C for up to 5 minutes. Digested tissue was triturated until homogenous cell suspensions were created. Cell suspensions were washed and seeded onto 35 mm Petri dishes coated with

fibronection 2 mg/cm$^2$ (Sigma Aldrich) in 2 ml of DMEM-F12 medium (PAA) supplemented by N2 supplement 1:100, B27 supplement 1:50 (Invitrogen, Carlsbad, CA, USA), Penicillin (100 IU/ml)/ streptomycin (100 µg/ml) (PAA), 20 ng/ml FGF, 20 ng/ml EGF (Peprotech EC, London, UK) . Half of the medium was refreshed every two days and fresh EGF and FGF were added to the medium. The neural crest-derived cells were kept in culture as neurosphere-like bodies (NLBs) for 14 days.

## RT-PCR

Total RNA was isolated from NLBs cultures using the RNeasy mini kit (Qiagen) according to the manufacture's protocol. Two micrograms of RNA were transcribed into cDNA with the Ready-To-Go First Strand Synthesis kit (GE-Health care) according to the manufacture's instruction. PCRs were performed by using 1.25 ml of cDNA for 35 cycles using the primers given in supplementary Table 4.

## Immunohistochemistry

Immunostaining was performed on cryosections (12 mm) of frozen NLBs. NLBs were centrifuged at 180*g* in a microtube and then frozen in OCT embedding compound (Tissue-Tek; Sakura Finetek, Tokyo, Japan). Cryosections of NLBs were fixed in 4% PFA for 10 minutes at room temperature and rinsed three times in PBS before application of blocking solution (PBS containing 10% donkey serum, 0.1% Triton X-100) for 30 minutes at room temperature. Primary antibodies anti-RET, NXF, SIM2 and ARNT2 were applied and visualized with the respective fluorochrome-conjugated secondary antibodies (see supplementary Table 5). Cell nuclei were stained with 4'-6-diamidino-2-phenylindole [DAPI] solution (Invitrogen; 100 ng/ml) and coverslipped with Kaiser's gelatine (Merck, Darmstadt, Germany). The images were captured on a Nikon inverted fluorescence microscope (Nikon, Dusseldorf, Germany).

## EMSA and Supershift Assays

Nuclear protein extracted from Neuro2A cells transfected with *NXF*/*ARNT2*, *SIM2*/*ARNT2*, *SIM2*, *NXF* and *ARNT2* constructs were isolated using a Nuclear Protein extraction kit (Sigma Aldrich, Saint Louis, USA). Complementary oligonucleotides (30 bp) containing either Enh2-C or Enh2-A were designed (see Supplementary Table 6). Sense and antisense oligonucleotide were 3'end labeled (Biotin3'end-DNA labeling kit, Pierce Biotechnologies, Rockford, IL, USA) according to the manufacture's protocol. Binding experiments were performed using an EMSA kit (Pierce) according to the manufacture's protocol using 20 mg of nuclear protein

and 1 ml of 25 fmol of biotin labeled oligonucleotides. In the competitive reactions a 1000 times excess of unlabeled oligonucleotides was added. In the Supershift experiments, 1 ml rabbit polyclonal NXF antibody (Abcam, Cambridge, UK) was mixed with 10 mg of total lysate of a NXF and ARNT2 positive cell line (SK-N-SH). Gelshifts were visualized using the LightShift® Chemiluminescent EMSA kit (Pierce).

# RESULTS

## Mutations in either Ehn1 or Ehn2 reduce reporter gene expression

Constructs expressing Luciferase, driven by the *RET* or *SV40* promoter coupled to the *RET* enhancer sequence containing SNP rs245357 (Enh1:C>T) or SNP rs2506004 (Enh2:C > A), were transfected into Neuro2A cells. The inserts containing Enh1, Enh2 and Enh1/Enh2 differ in length; 300bp, 400bp and 500 bp respectively. These differences might have an effect on binding affinity. Therefore, we compared the expression of the mutant constructs to their WT counterparts. Constructs with the *SV40* promoter showed that both the disease-associated alleles (Enh1-T/Enh2-A) separately or in combination, decreased the Luciferase expression approximately two-fold (Figure 3A). Constructs in which the *RET* enhancer sequence was coupled to the *RET* promoter, however, showed that only when the two disease variants were both present, were able to decrease Luciferase expression (Figure 3B). In these RET promoter experiments, we did not test whether one or both mutant alleles were necessary for the reduction of Luciferase expression. Therefore, by using site-directed mutagenesis, we made constructs that contain a combination of one non-disease-associated variant and its counterpart disease-associated variant (Enh1-C/Enh2-A or Enh1-T/Enh2-C). We could show that a mutation in either variant decreased the Luciferase expression (Figure 3C).

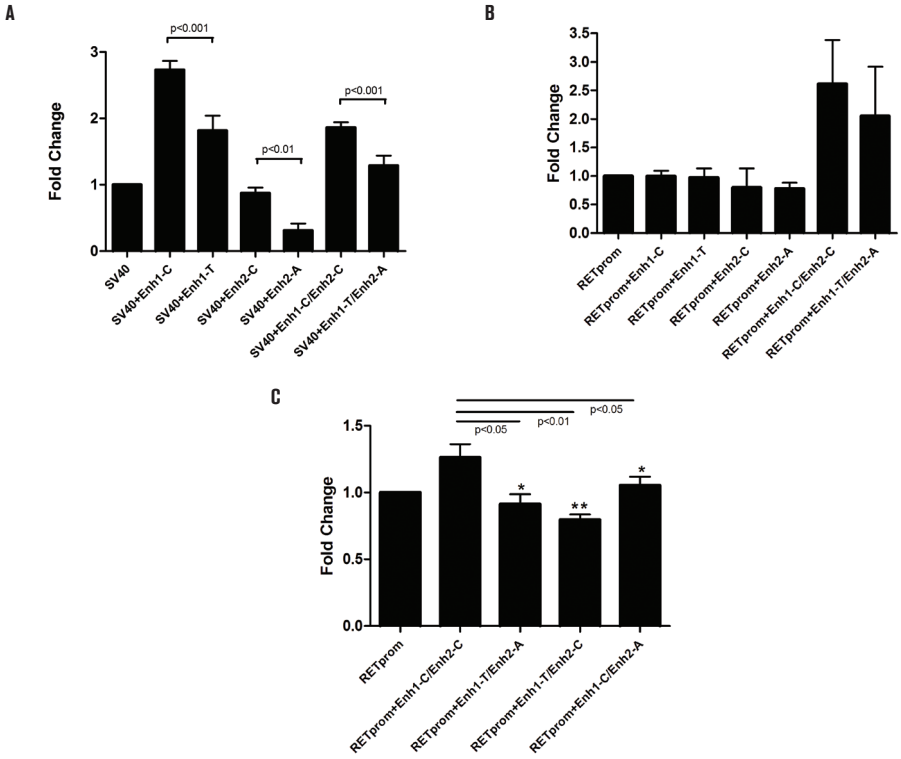



**Figure 3 -** Panel (A) shows under SV40 promoter regulation, both the Enh1-T and the Enh2-A separately or in combination decreased the Luciferase expression approximately 2-fold. (B) Under *RET* promoter regulation, only when both diseases associated SNPs are present they can decrease Luciferase expression. (C) Constructs with the *RET* promoter coupled to the enhancer sequence containing combination of a non-disease associated variant and its disease-associated counterpart (Enh1-T/Enh2-C and Enh1-C/Enh2-A),, showed that both SNPs could significantly decrease Luciferase expression. The bars indicate the mean ± S.E.M of the fold changes of Luciferase expression for each construct. $p < 0.05$ was considered significant.

## In silico analysis of the MCS+9.7 sequence

*MatInspector* (www.genomatix.de) analysis of the MCS+9.7 sequences showed that the Enh2-C allele with its surrounding sequences -A**C**GTG- (also known as the Central Nervous System Midline Element (CME)) is a core binding site of the heterodimers NXF/ARNT2 and SIM2/ARNT2. These heterodimers have the same core binding sequence, but have opposite effects. NXF/ARNT2 is an expression activator, whereas SIM2/ARNT2 represses gene transcription[19].

## Reduced NXF/ARNT2 binding to Enh2-A compared to Enh2-C

To check whether the NXF/ARNT2 heterodimer could indeed bind to the Enh2-C sequence, we performed Electro Mobility Shift Assay (EMSA). Oligonucleotides containing the Enh2-C allele or the Enh2-A allele were mixed with the nuclear protein extract (NPE) of Neuro2A cells that were transiently transfected with either the *NXF* or *ARNT2* constructs or with both constructs together. Protein binding was determined to be sequence specific if the excess of unlabeled oligonucleotides could replace the labeled oligonucleotides (Cp-Sp) while the excess of unrelated-unlabeled oligosnucleotides could not replace it (Ns-SP). NPE of Neuro2A transfected with all these constructs could bind to both oligonucleotides containing Enh2-C or Enh2-A (Figure 4A and B: lane 1, 3 and 4). Only one band of the NPE of Neuro2A transfected with NXF/ARNT2 (Figure 4A lane 1, white arrow) could be completely deleted by an excess of unlabeled oligonucleotides and stay the same when an excess unrelated-unlabeled was used in the competition assay (Figure 4A lane 5 and 6). Furthermore, this specific band has less or no binding to the oligonucleotides containing Enh2-A (Figure 4B lane 1, white arrow). To confirm whether this protein complex bound to Enh2-C is NXF/ARNT2, Supershift assays were performed using an anti-NXF antibody and a total lysate of a SK-N-SH, a human neuroblastoma cell line positive for NXF/ARNT2. Total lysates of SK-N-SH had the same binding pattern as the NPE of Neuro2A transfected by NXF/ARNT2 (Figure 4C and D, lane 1 arrow white and black). When the antibody anti-NXF was applied, the upper band was shifted and the lower band was blocked (Figure 4C and D, lane 4) showing that these two bands contain the NXF protein. Notably the polyclonal ARNT2 antibody was not able to either block or supershift the upper band (data was not shown). This could be because the epitope of ARNT2 is hidden and is not recognized by the antibody. Based on these results, we conclude that the upper band (white arrow) is the NXF/ARNT2 heterodimer. Based on EMSA results (Figure 4A lane 3) and Supershift results (Figure 4C lane 4, black arrow) we assume that the lower band most likely is NXF alone. Similar experiment was performed using NPE of Neuro2A cells transiently transfected with SIM2/ARNT2, SIM2 and ARNT2 (Figure 4E and F). It was shown that SIM2/ARNT2 and SIM2-alone gave a specific band for both Enh2-C and Enh2-A suggesting that only SIM2 binds to the enhancer sequence, independent of the mutation status (Figure 4E and F, lane 1 and 3).

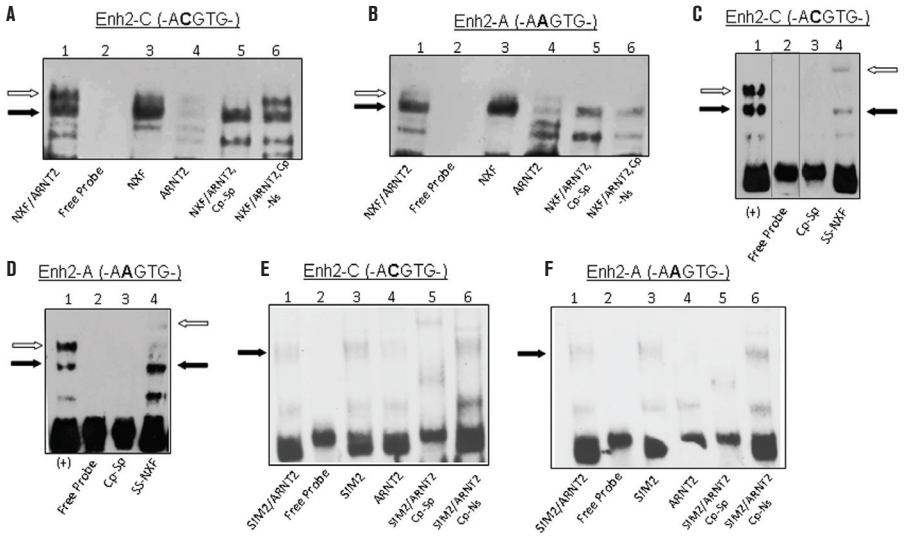



**Figure 4 -** Panel (A) and (B) show EMSA using labeled oligonucleotides containing Enh2-C and Enh2-A with nuclear protein extract (NPE) of Neuro2A cells transfected with *NXF/ARNT2*, *NXF* or *ARNT2* constructs: an extra band (white arrow) appeared when oligonucleotides containing Enh2-C allele were used in combination with NPE containing both *NXF* and *ARNT2*. In a competition assay (Cp-Sp) in lane 5, the extra band as seen in lane 1 disappeared, while excess of unrelated, unlabeled oligonucleotide (Cp-NS), could not compete the specific labeled oligonucleotide (lane 6). (C) and (D) EMSA-Supershift (SS) assays using Enh2-C and Enh2-A labeled oligonucleotide with total cell lysates of the SK-N-SH cell line and anti-NXF antibody showed that the highest band was shifted upwards and the second lower band became weaker (was blocked) (lane 4). Based on Fig. 4C, the highest band (white arrow), which is shifted by the NXF antibody, is the NXF/ARNT2 complex, and the second band (black arrow), for which the binding is blocked in lane 4, most likely is NXF. (E) and (F) EMSA using labeled oligonucleotides containing Enh2-C and Enh2-A with NPE of Neuro2A transfected with SIM2/ARNT2, SIM2 or ARNT2 contruct. SIM2 has the same binding affinity to the enhancer sequences containing Enh2-C or Enh2-A (lane 1 and 3), the specificity of this binding is confirmed in competition specific and non-specific assays (lane 5 and 6).

## *RET*, *NXF, SIM2* and *ARNT2* are expressed in NCSCs isolate from mouse embryonic gut

Neurosphere-like bodies (NLBs), propagated in vitro from the gut of mouse embryos E11.5-E15.5 (Figure 5A) were checked for expression of *RET*, *NXF*, *ARNT2* and *SIM2* by RT-PCR. The result showed that *RET*, *SIM2*, *NXF* and *ARNT2* are expressed in all stages analyzed. The RET expression increase in time, while NXF and SIM2 expression reach the highest level on E12.5 to 14.5 and decrease on E15.5. The neural progenitor status of RET positive cells within NLBs cultures was demonstrated by co-immunostaining with the neural pro-

genitor marker p75 (Figure 5B). In addition, the majority of SIM2 and ARNT2-positive cells co-expressed RET (Figure 5C and 5D). The only available commercial NXF antibody revealed no satisfying results on mouse tissue and it therefore could not be used for immunostaining (data not shown).

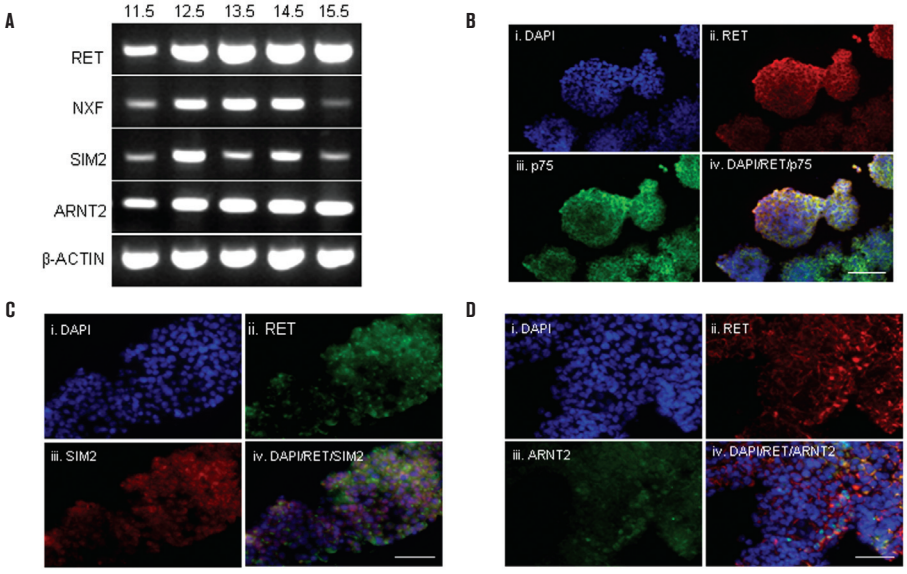



**Figure 5 -** Panel (A): By using RT-PCR method it is showed that *RET*, *NXF, SIM2* and *ARNT2* are expressed in NCSCs isolated from mouse embryonic gut: E11.5, 12.5, 13.5, 14.5 and 15.5. (B) Progenitor status of RET positive cells in neurosphere-like body (NLBs) cultures by co-immunostaining with the neural crest marker p75: i) DAPI staining, ii) RET expression, iii) p75 expression and iv) co-expression of RET and p75. C) SIM2 co-expression with RET in NLBs cultures: i) DAPI staining, ii) RET expression, iii) SIM2 expression and iv) co-expression of RET and SIM2. D) ARNT2 co-expression with RET in NLBs cultures: i) DAPI staining, ii) RET expression, iii) ARNT2 expression and iv) co-expression of RET and ARNT2. Scale bars: 100 mm in B (iv), 50 mm in C (iv) and D (iv).

### NXF/ARNT2 and SIM2/ARNT2 increases and decrease *RET* expression respectively via the *RET* promoter and the enh2 sequences.

Not only the *RET* enhancer sequence (MCS +9.7) but also the *RET* promoter (starting 5.1 kb upstream of the *RET* start site) contains six CME (Central Nervous System Midline Element) variants. NXF/ARNT2 and SIM2/ARNT2 might regulate *RET* expression via binding to the enhancer region containing the Enh2-C allele and the *RET* promoter. To prove this, we performed co-transfections of *NXF/ARNT2* or *SIM2/ARNT2* with the *RET* promoter (RETprom) constructs with or without the enhancer sequences containing Enh2-C or Enh2-A, into Neuro2A, a cell line not endogenously expressing NXF, SIM2 and ARNT2. Figure 6A shows that co-transfection of *NXF/ARNT2* with the *RET* promoter construct increases the Luciferase expression 1.25-fold, while co-transfection with RETprom+Enh2-C allele, the Luciferase levels go up 1.75-fold. The RETprom+Enh2-A allele results in a Luciferase expression level comparable to that of the RETprom construct without the enhancer sequence. This shows that NXF/ARNT2 can regulate *RET* expression via both the promoter and the enhancer sequences. On the contrary, the co-transfection with *SIM2/ARNT2* decreased the Luciferase expression level two-fold for the RETprom construct (without enhancer). Similar levels of Luciferase expression were obtained for the constructs containing the *RET* promoter coupled to the enhancer (for both the Enh2-C and the Ehn2-A allele). These data suggest that the binding of SIM2/ARNT2 to either of Enh2-C or Enh2-A is ablating the enhancer effect. Therefore the expression levels of the RETprom with Enh2-C or Ehn2-A are equal to those of the RETprom alone, in the present of SIM2/ARNT2 expression. Furthermore, by increasing the amount of constructs transfected, we could induce an increase of up- or down-regulation of gene expression (Figure 6C and D). This shows that expression regulated by NXF/ARNT2 and SIM2/ARNT2, via the *RET* promoter and the enhancer sequence, depends on dosage (Figure 5C).

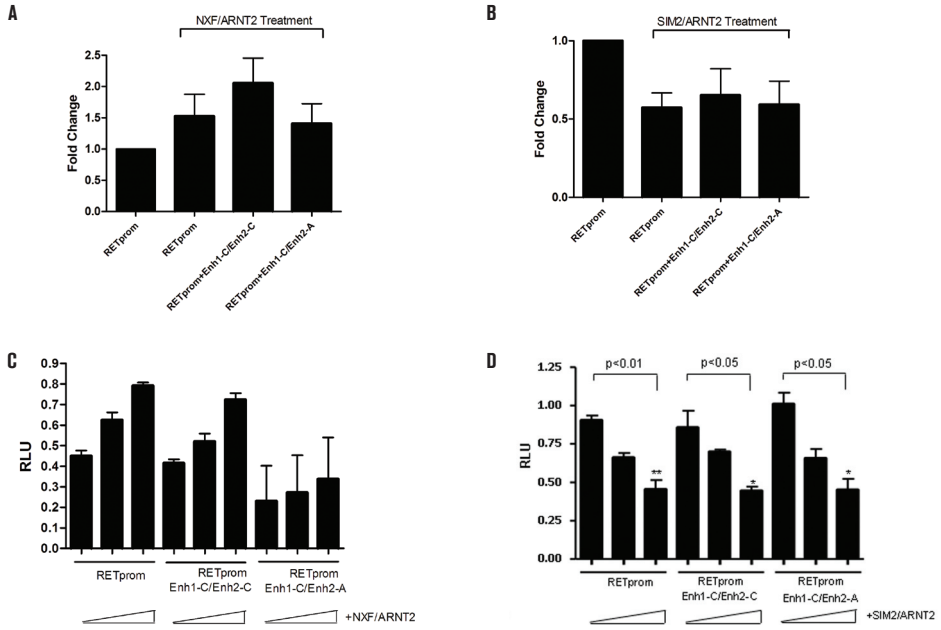



**Figure 6 -** Panel (A) shows the effect of *NXF/ARNT2* to Luciferase expression of constructs containing the *RET* promoter or the *RET* promoter and enhancer containing Enh2-C or Enh2-A. (B) Effect of SIM2/*ARNT2* to Luciferase expression of constructs containing the *RET* promoter or the *RET* promoter and enhancer containing Enh2-C and Enh2-A variants. (C) and (D) a possible (not significant) dosage effect of *NXF/ARNT2* and *SIM2/ARNT2* on Luciferase expression driven by the *RET* promoter and enhancer regulation. Three dosages of *NXF/ARNT2* and *SIM2/ARNT2* constructs were used. The bars indicate the mean ± S.E.M of fold changes of Luciferase expression for each construct after co-transfection of *NXF/ARNT2* or *SIM2/ARNT2*. In figures C and D the bars indicate the mean ± S.E.M of normalized-Luficerase expression for each construct after co-transfection with three different dosages of *NXF/ARNT* or *SIM2/ARNT2*. $p < 0.05$ was considered significant.

## Level of endogenous RET protein expression is up- and down-regulated after *NXF/ARNT2* & *SIM2/ARNT2* transfection, respectively

To check the effect of NXF/ARNT2 and SIM2/ARNT2 on endogenous RET expression, we transfected three different dosages of the *NXF/ARNT2* or *SIM2/ARNT2* constructs into Neuro2A cells. The level of RET endogenous expression was checked by Westernblot. The transfection of the highest dosage of *NXF/ARNT2* in the cell line enhanced the RET expression level 2.25 times, whereas the transfection of the highest dosage of *SIM2/ARNT2* constructs reduced the RET expression level 2.6 times (Figure 7).

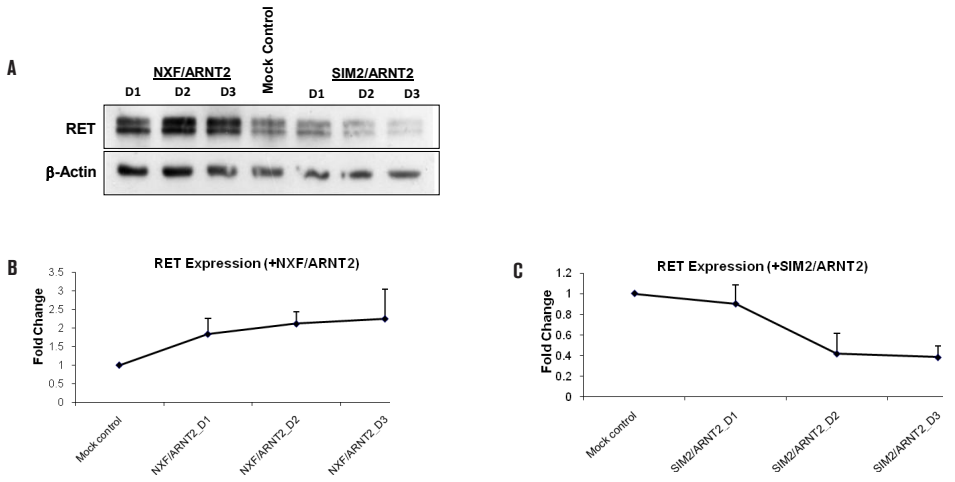



**Figure 7 -** Panel A shows Western blotting of total cell lysate of Neuro2A cells transfected with three different dosages of *NXF/ARNT2* or *SIM2/ARNT2*: D1 (100/50 ng), D2 (500/250 ng) and D3 (1000/500 ng) using anti-RET and anti-b-Actin primary antibodies. (B) Densiometric analysis of the Western blots. The highest dosage of *NXF/ARNT2* enhanced RET expression 2.25-fold compared to the RET expression in the mock controls. (C) Densiometric analysis of the Western blots. The highest dosage of *SIM2/ARNT2* decreases RET expression 2.6-fold compared to the RET expression in the mock controls. The bars indicate the mean ± S.E.M of fold changes for RET endogenous expression after transfection of *NXF/ARNT2* or *SIM2/ARNT2*. $p < 0.05$ was considered significant.

# DISCUSSION

Recent studies showed that the T allele of SNP rs2435357 (Enh1-T), a SNP present in an enhancer element (MCS+9.7) in intron 1 of the *RET* gene, contributes to HSCR development. It was shown in reporter assays that the construct containing the non-disease associated allele, Enh1-C, contrasts to that with the disease-associated allele, Enh1-T, by enhancing luciferase expression 6-fold[14]. Furthermore, *in vitro* assays showed that the effect Enh1-T has on *RET* expression is caused by a disruption of a SOX10 binding site within the enhancer element[18]. However, in the sequence used in all the experiments, a second SNP is also present, namely rs2506004[14]. The A allele of this SNP is also strongly associated with HSCR,[13] what is not surprising as they are only 217 base pairs apart. We investigated whether the disease-associated allele, Enh2-A, also played a role in regulating *RET* expression.

In order to prove whether Enh1-T and Enh2-A, separately or in combination, function in

gene expression regulation, we generated Luciferase-reporter constructs containing the two alleles of only Enh1 or Enh2 and constructs containing alleles of the two SNPs, under the regulation of the *SV40* or the *RET* promoter (RETprom). Luciferase assays showed that the disease-associated allele of each SNP, separately or in combination, under *SV40* regulation, could reduce Luciferase expression (Figure 3A). However, under *RET* promoter regulation, only the construct containing both non-disease associated SNPs together could yield the enhancer effect, while the disease-associated SNPs, together, decreased Luciferase expression (Figure 3B). It is likely that the constructs used to show the effect for the SNPs independently were lacking sequences necessary for the binding of *RET*-specific transcription factors. This experiment still left the possibility that only one of the SNPs is the functional variant. To define the function of each SNP, we generated constructs containing the non-disease associated allele of one SNP with the disease-associated allele of the other SNP, under *RET* promoter regulation. Luciferase assays of these constructs showed that Enh1-T or Enh2-A in the existence of its non-disease associated counterpart could significantly reduce Luciferase expression. The effect of the Enh1-T allele is bigger than for the Enh2-A allele, but the combination of both SNPs did not double the effect (Figure 3C). This experiment showed that disease-associated alleles of both of these SNPs can contribute to the development of HSCR.

To elucidate the effect of Enh2-A on the functioning of the enhancer, and to identify a protein candidate that could bind to this SNP, we analyzed the enhancer sequences containing the non-disease and disease-associated variants by *MatInspector* software (www.genomatix.de). The Enh2-C (non-disease) allele and its surrounding sequence proved to be present in the -A**C**GTG- sequence, also known as the Central Nervous System Midline Element (CME).[20] This CME sequence element and three other CME variants (-**G**CGTG-, -**T**CGTG- and –**C**CGTG-) are binding site of NXF/ARNT2 and the SIM2/ARNT2 heterodimers.[19].

NXF is a basic-Helix-Loop-helix/Per-Arnt-Sim (bHLH/PAS) protein, which is mainly expressed in neuronal tissues. It was shown in a previous study that NXF can form a heterodimer with ARNT2. Moreover, it was proven that NXF/ARNT2 competes with the SIM2/ARNT2 heterodimer in binding to the CME sequences.[19] These two heterodimers regulate gene expression for the *Drebrin* gene as transcription activator and repressor, respectively, via binding to the CME sequences in the *Drebrin*'s promoter. We therefore hypothesized that both NXF/ARNT2 and SIM2/ARNT2 are also involved in regulating RET expression via the binding of these two heterodimers to the CME in the Enh2 sequence. Moreover, as we detected six CME variants in the promoter of RET, we cannot exclude that these sites also play a role in the expression regulation. By EMSAs and a Supershift assay, using a NXF antibody, we showed

that the NXF/ARNT2 heterodimer could indeed bind to the Enh2-C (non-disease associated) allele and that the heterodimer has less binding affinity to Enh2-A, the disease-associated allele (Figure 4A, 4B, 4C and 4D, lane 1). EMSAs of SIM2/ARNT2 showed that also SIM2 could bind to Enh2 but that it had the same binding affinity to the wild type and the mutant Enh2 sequences.

A first step to prove our hypothesis was to show that these proteins are expressed in the (migrating) neural crest stem cells (NCSC), the progenitors of the cells that are lacking in the hindgut of HSCR patients. Therefore, we checked by RT-PCR and immunohistochemical staining the presence of NXF, SIM2 and ARNT2 in NCSC isolated from embryonic guts (E11.5 until E15.5). By day 15.5 the migration of the NCSCs and the formation of the enteric nervous system is complete in the mouse. We show that RET expression is up-regulated during this period, while SIM2 and NXF were expressed in all stages analyzed with the highest expression at E12.5 to E14.5. Furthermore, immunostaining of neural progenitor cells isolated from mouse embryonic gut day 13.5 showed that both SIM2 and ARTN2 are expressed (Figure 5C and D).

To check whether NXF/ARNT2 and SIM2/ARNT2 can indeed activate and inhibit the expression of RET, via the *RET* promoter and/or via the enhancer region containing the Enh2-C allele, we co-transfected *NXF*/*ARNT2* or *SIM2*/*ARNT2* constructs with the *RET* promoter-*Luciferase* constructs, with or without the enhancer sequences containing the Enh2-C or Enh2-A alleles, into Neuro2A cell lines. The results show that NXF/ARNT2 could enhance the Luciferase expression via the *RET* promoter and via the enhancer sequence containing Enh2-C. The disease-associated allele, Enh2-A, reduces the Luciferase expression to the level of the RET promoter only (Figure 6A). In contrast, co-transfection of *SIM2*/*ARNT2* resulted in a decrease of Luciferase expression.

Furthermore, the level of Luciferase expression proved to depend on the dosage of NXF/ARNT2 and SIM2/ARNT2, as by increasing concentrations of NXF/ARNT2 or SIM2/ARNT2 a further up- or down-regulation of Luciferase expression was measured, respectively (Figure 6C and D). To further strengthen these data, we also performed a comparable dosage-effect experiment in mouse Neuro2A cells, checking endogenous RET expression. The results confirmed the patterns of gene expression found in our initial Luciferase assays. The highest dosage of *NXF*/*ARNT2* increased RET expression levels by 2.25-fold, while the highest dosage of *SIM2*/*ARNT2* reduced RET expression levels by 2.6-fold (Figure 7).

SIM2 is located in the critical region for Down Syndrome on chromosome 21 and it is expressed in the fetal brain, particularly in regions crucial for cognitive processes. Moreover,

it was shown that overexpression of SIM2 in rat and mouse hippocampus impairs the ability of learning and memory by inhibiting synaptic plasticity.[21, 22] Interestingly, in Down syndrome patients who have three copies of the SIM2 gene, the structure of the brain regions expressing SIM2 are altered.[23] What this overexpression of SIM2 leads to is only partly known, although it was shown that overexpression of this gene leads to reduced expression of Drebrin, a protein playing an important role in synaptic plasticity.[18] Its expression is known to be decreased in the brains of Down syndrome patients.[24] We have now shown that SIM2 overexpression results in a down-regulation of RET. This finding might partly explain why we see a 40-fold higher risk for HSCR disease in Down syndrome patients than in the general population of newborns.

In this study we have proved that not one but at least two SNPs on the HSCR-associated haplotype can contribute to the disease phenotype. However it is still unclear when exactly these SNPs act in the disease development. They might deregulate proliferation during the first stages of embryogenesis, or impair migration of the NCSCs, or play a role during proliferation when the cells reach their final destination in the gut. We cannot exclude the possibility that the different disease-associated SNPs play a role at different stages of development, or are involved in different processes. Previous studies showed evidence for the involvement of at least two additional SNPs on the same haplotype. These SNPs are located at -5 and -1 from the start site of the RET gene (Figure 1). It was also shown that the TTF-1 protein binds to this region and that TTF-1 can activate the RET promoter[10, 12]. Our new data and previously published data hold an important message for the study of complex diseases, namely that more than one SNP on an associated haplotype might influence disease development. So far, almost all studies have stopped when a SNP was found with a proven functional effect. We have shown that even when such a SNP has been found, there may be more SNPs involved in the disease development. This means that polygenic diseases are even more complex than originally thought.

## Acknowledgements

We thank Jackie Senior for editing the manuscript. We thank Prof.Dr. Isabella Ceccherini (Laboratorio di Genetica Molecolare, Instituto G Gaslini, Genoa, Italy) for providing us with the pGL3-RETpromoter plasmid, Dr. Sabourin (Department of Cellular and Molecular Medicine, University of Ottawa, Canada) for providing the pCDNA3-NXF (rat) and pCDNA1-ARNT2 (rat) expression constructs, and Dr. Cheng Xu (Department of Embryology, Carnegie Institution of Washington, Baltimore, USA) for providing the pCDNA3-SIM2 (mouse) expressing construct.

# REFERENCES


1. Azmiel J, Sproat-Emison E, Garcia-Barcelo M, et al. Hirschsprung disease, associated syndromes and genetics: a review. J Med Genet 2008;45:1-14.

2. Bolk S, Pelet A, Hofstra RM, et al.. 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 U S A 2000;97:268-273.

3. Gabriel SB, Salomon R, Pelet A, et al. Segregation at three loci explains familial and population risk in Hirschsprung disease. Nat Genet 2002;31:89-93.

4. Angrist M, Kauffman E, Slaugenhaupt SA, et al. A gene for Hirschsprung disease (megacolon) in the pericentromeric region of human chromosome 10. Nat Genet 1993;4:351-356.

5. Borrego S, Saez ME, Ruiz A, et al. 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;36:771-774.

6. Borrego S, Ruiz A, Saez ME, et al. RET genotypes comprising specific haplotypes of polymorphic variants predispose to isolated Hirschsprung disease. J Med Genet 2000;37:572-578.

7. Borrego S, Wright FA, Fernandez RM, et al. 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.

8. Garcia-Barcelo MM, Sham MH, Lui VC, Chen BL, et al. Chinese patients with sporadic Hirschsprung's disease are predominantly represented by a single RET haplotype. J Med Genet 2003;40:e122.

9. Fitze G, Appelt H, Konig IR, et al. Functional haplotypes of the RET proto-oncogene promoter are associated with Hirschsprung disease (HSCR). Hum Mol Genet 2003;12:3207-3214.

10. Sancandi M, Griseri P, Pesce B, et al. 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;40:714-718.

11. Lantieri F, Griseri P, Puppo F, et al. Haplotypes of the human RET proto-oncogene associated with Hirschsprung disease in the Italian population derive from a single ancestral combination of alleles. Ann Hum Genet 2006;70:12-26.

12. Garcia-Barcelo M, Ganster RW, Lui VC, et al. TTF-1 and RET promoter SNPs: regulation of RET transcription in Hirschsprung's disease. Hum Mol Genet 2005;14:191-204.

13. Emison ES, McCallion AS, Kashuk CS, et al. A common sex-dependent mutation in a RET enhancer underlies Hirschsprung disease risk. Nature 2005;434:857-863.

14. Burzynski GM, Nolte IM, Bronda A, et al. Identifying candidate Hirschsprung disease-associated RET variants. Am J Hum Genet 2005;76:850-858.

15. Pelet A, de Pontual L, Clement-Ziza M, et al. Homozygosity for a frequent and weakly penetrant predisposing allele at the RET locus in sporadic Hirschsprung disease. J Med Genet 2005;42:e18.

16. Grice EA, Rochelle ES, Green ED, et al. Evaluation of the RET regulatory landscape reveals the biological relevance of a HSCR-implicated enhancer. Hum Mol Genet 2005;14:3837-3845.

17. Fisher S, Grice EA, Vinton RM, et al. Conservation of RET regulatory function from human to zebrafish without sequence similarity. Science 2006;312:276-279.

18. Emison ES, Garcia-Barcelo M, Grice EA, et al. Differential contributions of rare and common, coding and noncoding Ret mutations to multifactorial Hirschsprung disease liability. Am J Hum Genet 2010;87:60-74.

19. Ooe N, Saito K, Mikami N, et al. Identification of a novel basic helix-loop-helix-PAS factor, NXF, reveals a Sim2 competitive, positive regulatory role in dendritic-cytoskeleton modulator drebrin gene expression. Mol Cell Biol 2004;24:608-616.

20. Wharton KA, Jr., Franks RG, Kasai Y, et al. Control of CNS midline transcription by asymmetric E-box-like elements: similarity to xenobiotic responsive regulation. Development 1994;120:3563-3569.

21. Meng X, Peng B, Shi J, et al. Effects of overexpression of Sim2 on spatial memory and expression of synapsin I in rat hippocampus. Cell Biol Int 2006;30:841-847.

22. Ema M, Ikegami S, Hosoya T, et al. Mild impairment of learning and memory in mice overexpressing the mSim2 gene located on chromosome 16: an animal model of Down's syndrome. Hum Mol Genet 1999;8:1409-1415.

23. Rachidi M, Lopes C, Charron G, et al. Spatial and temporal localization during embryonic and fetal human development of the transcription factor SIM2 in brain regions altered in Down syndrome. Int J Dev Neurosci 2005;23:475-484.

24. Shim KS, Lubec G. Drebrin, a dendritic spine protein, is manifold decreased in brains of patients with Alzheimer's disease and Down syndrome. Neurosci Lett 2002;324:209-212.

# SUPPLEMENTARY DATA

**Supplementary Table 1 -** Primers used for PCR, length of product and construct name

| Enhancer containing HSCR-associated variant | Primers | Product length (bp) | Constructs |
|---|---|---|---|
| Enh1-C and Enh1-T | F : 5'-TAGTGGTACCTTGCACATGGACATGATCTG -3'<br>R: 5'-TAGTCTCGAGTGCTTCCGGCTTCTATGCAA -3' | 300 | RETprom+Enh1-C<br>And<br>RETprom+Enh1-T |
| Enh2-C and Enh2-A | F : 5'-TAGTGGTACCGCCAGTGAGGCTGGTGATTA-3'<br>R: 5'-TAGTCTCGAGGATCCCTGCACGCACGAACA-3' | 490 | RETprom+Enh2-C<br>And<br>RETprom+Enh2-A |
| Enh1-C/Enh2-C<br>and<br>Enh1-T/Enh2-A | F: 5'-GTTGGGTACCGCCAGGGCCAGTGAAC-3'<br>R: 5'-TGGGCTCGAGGGCAGCCCAAGGCTAC-3' | 400 | RETprom+Enh1-C/<br>Enh2-C<br>And<br>RETprom+Enh1-T/<br>Enh2-A |

**Supplementary Table 2 -** Primers used for site-directed-mutagenesis

| Construct used for template | Primers | Constructs |
|---|---|---|
| Enh1-C/Enh2-C | F: 5'-CAGTGACCCTTACATGGTCATCCACAGGCCACTTG -3'<br>R: 5'-CAAGTGGCCTGTGGATGACCATGTAAGGGTCACTG -3' | RETprom+Enh1-T/Enh2-C |
| Enh1-C/Enh2-C | F: 5'-CTGGTCTGGGTATGGAAGTGTTGGGTGGCCATCG-3'<br>R: 5'-CGATGGCCACCCACACTTCCATACCCAGACCGAG -3' | RETprom+-Enh1-C/Enh2-A |

**Supplementary Table 3 -** Antibodies used in Westernblot

| | Antibodies | Host | Dilution | Supplier |
|---|---|---|---|---|
| **Primary** | RET<br>β-Actin | Goat<br>mouse | 1 : 500<br>1 : 10.000 | R&D system, MN, USA<br>MP Biomedical, Irvine, CA, USA |
| **Secondary** | Anti-goat IgG-HRP<br>Anti-mouse IgG-HRP | Rat<br>Goat | 1 : 3000<br>1 : 5000 | MP Biomedical, Irvine, CA, USA<br>Bio-Rad |

**Supplementary Table 4 -** Primers used for RT-PCR amplification and product length

| Gene | Primers | Product length | Tm= °C |
|---|---|---|---|
| RET | F : 5'- CCACTGGGCCTCTATTTCTC -3'<br>R: 5'- TGCCCTCCTTCCGCTTAAAC -3' | 799 | 50 |
| NXF | F : 5'- TTGACCCTGCTGACCATCTC -3'<br>R: 5'- GTCTGGGAAGGTAGCACTGCTG -3' | 505 | 59.3 |
| SIM2 | F: 5'- TCACCTGCTCCAAGAGTACGA -3'<br>R: 5'- TGGCTGACATCCAGTTTGTC -3' | 765 | 50 |
| ARNT2 | F: 5'-ACCAGCAGGAATGACCATAC -3'<br>R : 5'-CTTGGCCTTTCAGCTTAACC-3' | 317 | 50 |
| β-ACTIN | F : 5'-ATATCGCTGCGCTGGTCGTC -3'<br>R : 5'-AGGATGGCGTGAGGGAGAGC -3' | 517 | 50 |

**Supplementary Table 5 -** Antibodies used in Immunohistochemistry

| | Antibodies | Host | Dilution | Supplier |
|---|---|---|---|---|
| **Primary** | RET | Goat | 1 : 50 | R&D system, MN, USA |
| | SIM2a | Goat | 1 : 200 | MP Biomedical, Irvine, CA, USA |
| | ARNT2 | Rabbit | 1 : 50 | Santa Cruz |
| | $p75^{NTR}$ | Rabbit | 1 : 250 | Promega, Mannheim, Germany |
| **Secondary** | Anti-Goat IgG Alexa 488<br>Anti-Rabbit IgG Allexa 488 | Donkey | 1 : 500 | Invitrogen, Carlsbad, CA, USA |
| | Anti-Goat-IgG CY3<br>Anti-Rabbit IgG CY3 | Donkey | 1 : 500 | Invitrogen, Carlsbad, CA, USA |

**Supplementary Table 6 -** Oligonucleotides containing Enh-2C and Enh-2A used for EMSA

| Oligonucleotides containing | Sequences |
|---|---|
| **Enh2-C** | Sense : 5'- GGTCTGGGTATGGACGTGTGGGTGGTGGCC -3'<br>Anti-sense : 5'-GGCCACCACCCACACGTCCATACCCAGACC-3' |
| **Enh2-A** | Sense : 5'- GGTCTGGGTATGGAAGTGTGGGTGGTGGCC -3'<br>Anti-sense : 5'- GGCCACCACCCACACTTCCATACCCAGACC -3' |

# 5 RET Mediated Gene Expression in ENS Precursors shows Involvement of RET in Innate Immune Response

Y. Sribudiani[1], R.S.N. Fehrmann[1*], M. Metzger[2*], D. Schriemer[1*], E. Binder[3], W. Douwenga[4], E.M. Wesseling[5], U.L.M. Eisel[4], A.J. Burns[3], B.J.L. Eggen[5], N. Thapar[3], R.M.W. Hofstra[1]

[1] Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands.
[2] Translational Centre for Regenerative Medicine, University of Leipzig, Leipzig, Germany.
[3] Gastroenterology and Neural Development Units, University College London Institute of Child Health, London, United Kingdom.
[4] Department of Molecular Neurobiology, University of Groningen, Groningen, the Netherlands.
[5] Department of Neuroscience, Section of Medical Physiology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands.
[*] Equal contribution

***Submitted***

## ABSTRACT

Signaling of the RET receptor is crucial for the migration, proliferation, differentiation and survival of neural crest cells (ENCCs) that form the enteric nervous system (ENS). Disturbances in RET signaling are associated with ENS defects, as is seen in patients with Hirschsprung disease.

To gain new insight into the pathways involved in or triggered by RET in ENCCs, we performed gene expression profiling with RNA isolated from GDNF- (the RET ligand) stimulated and non-stimulated mouse ENCCs. Microarray data on ENCCs was compared with microarray data generated from mouse embryonic gut E14.5 and analyzed by gene set enrichment analysis (GSEA) and single-gene analysis methods. Pathways that were up-regulated upon GDNF stimulation included those for ATP- and biosynthesis while those significantly down-regulated upon GDNF stimulation were several known development-dependent pathways, such as the Notch, Tgf-β/Bmp and Wnt signaling pathways. Moreover, we saw that the Toll-like receptor (Tlrs) signaling pathway, important for apoptosis and the innate immune response, was down-regulated upon GDNF stimulation. This is interesting as Hirschsprung patients often suffer from inflammation of the colon and small intestine (enterocolitis). By qPCR we were able to confirm down-regulation of *Tlr2*, *Tlr4* and *Tlr5* and downstream effectors of the Tlrs, such as *Tnfα* and *Il-6*, pointing toward the involvement of Ret signaling in the immuno-modulatory capacity of ENCCs.

In conclusion, GDNF stimulation promotes proliferation of ENCCs rather than differentiation. RET seems to be involved in regulating the innate immune response of ENCCs via the TLRs pathway which is possibly linking RET to enterocolitis.

# INTRODUCTION

The *RET* proto-oncogene encodes a transmembrane glycoprotein belonging to the receptor tyrosine kinase (RTKs) family. During vertebrate embryogenesis, *RET* is expressed in motor and catecholaminergic neurons of the central nervous system (CNS), in the developing excretory system and in all lineages of the peripheral nervous system (PNS), including neural crest cells (ENCCs). ENCCs originating from the vagal region of the hindbrain contribute to the vast majority of parasympathetic neurons and glial cells of the enteric nervous system (ENS).[1] In mouse, vagal neural crest cells migrate in a rostocaudal direction to the foregut at E9.5 and reach the hindgut at E15.5, while sacral crest cells migrate in the opposite direction at E13.5.[1-3] In humans, ENCCs start to invade the foregut at week 4 and reach the terminal hindgut at week 7 of gestation.[4] This process of proliferation and migration of ENCCs into the whole gastrointestinal (GI) tract, followed by differentiation into ENS neurons and glial cells, is very complex. The timing and harmonious action of transcription factors, growth factor, their receptors and the signaling effectors is essential for each stage of development of the ENS.[5, 6] The RET, Retinoic Acid (RA), WNT, NOTCH, BMP and FGF signaling pathways are some of the pathways that are known to play important roles in ENCCs development in the gut.

Defects in proper migration, proliferation or differentiation of ENS progenitors can lead to the development of Hirschsprung disease (HSCR), the most common ENS-related disorder. It is characterized by the absence of neurons and glial cells in the myenteric and submucosal plexus of the enteric nervous system and is mostly limited to the most distal part of the colon. The inheritance of HSCR is complex: the disease can be transmitted as a dominant trait, or as a recessive trait, but in the majority of cases it is believed to be polygenic. There are clear differences in the sex ratio of patients, with a male predominance in short-Hirschsprung disease (S-HSCR) (4:1). The disease has an incomplete penetrance, variable expression and is associated with a large number of syndromes and congenital malformations.[7] Over the last 15 years, at least 12 genes that play an important role in the development of HSCR disease have been identified. The *RET* gene is considered to be the major disease risk factor as coding sequence mutations are identified in up to 50% of familial cases and in 15-35% of sporadic HSCR cases.[8] In fact, the RET signaling pathway is believed to be one of the most important signaling cascades in regulating proper migration, proliferation, differentiation and survival of ENCCs during ENS development.[1, 9] This is corroborated by several animal studies as mice deficient for the Ret receptor but also for the Ret co-receptor Gfr$\alpha$1, or the Ret ligand Gdnf, all show severe defects in enteric innervation.[10-13]

Little is known about the signaling cascades triggered by RET in ENCCs or the actual

downstream effectors of RET signaling in these cells. We therefore analyzed the RET network and its downstream effectors in ENCCs in detail. ENCCs were isolated from E14.5 mouse embryonic gut and cultured in the presence or absence of GDNF followed by expression profiling. We also performed gene expression profiling of mouse gut (E14.5). This study has identified genes and signaling pathways that are trigged by the RET signaling pathway in ENCCs and that can further contribute to our understanding of RET signaling pathway and its role in HSCR development.

## MATERIALS AND METHODS

### *Animals*

Transgenic mice carrying Wnt1-Cre-recombinase and a LoxP/R26-YFP site were obtained from the Jackson Laboratory, California, USA, and maintained on a C57BL/6 background. The age of the embryos was set at E0.5 on the day when the vaginal plug was seen. Mouse embryos carrying the transgenes produce green fluorescent ENCCs, and are easily identified by the green cranial face and green gut under a fluorescence microscope (Figure 1B).

### *ENCCs isolation, in vitro culturing and GDNF stimulation*

Pregnant female transgenic mice were sacrificed at stage E14.5. The embryos were removed and transferred into cold PBS solution. Under the microscope, only the fluorescent embryos indicating transgene expression were selected. We then carefully dissected the gut (foregut to hindgut) and transferred these into cold PBS. The embryonic gut from a single embryo was dissected and treated with 1 ml of collagenase/dispase enzyme solution (Sigma Aldrich, Saint Louis, USA] as described before.[14] The resulting cell suspension was diluted in 1 ml of DMEM/F-12 medium supplemented by 1% N2 (Invitrogen, Carlsbad, CA, USA), 2% B27 (Invitrogen), 20ng/ml FGF (Peprotech EC, London, UK), 20 ng/ml EGF (Peprotech EC) and 1% Penicillin-Streptomycin (Invitrogen). The cell suspension was filtered using sterile cell strainers with 40 μm nylon mesh (Becton Dickinson, Oxford, UK). Single YFP-ENCCs were sorted via a MoFlow cell sorter (Beckman Coulter, London, UK) and directly seeded into a 24-wells plate coated with fibronectin (2 μg/cm², Sigma Aldrich) filled with 0.5 ml of culture medium. YFP-ENCCs cells from one gut were divided into two wells of 24-wells plate. One well was filled with normal medium and one with medium containing GDNF (50ng/ml). The plates were incubated at 37oC for 14 days in a humidified atmosphere with 5% CO2. Half of the medium was changed every two days and fresh GDNF (only for GDNF treated cells),

FGF and EGF were added. Small neurosphere-like bodies (NBLs) appeared after 3-5 days and were kept in culture for 14 days without splitting.

### *RNA isolation and gene expression profiling*

Cells were washed twice with 1 ml of RNAse free 1X PBS. RNA isolation was performed using the RNeasy Mini Kit (Qiagen, Crawley, UK) according to the manufacture's instruction. The RNA yield, purity and integrity were determined using the Agilent 2100 BioAnalyzer with 6000 Nano Chips (Agilent Technologies, Amsterdam, the Netherlands). Custom DNA microarray experiments and RNA quality controls were performed by ServiceXS (Leiden, the Netherlands) using the GEneChip Mouse Genome 430 2.0 arrays of Affymetrix. These array experiments were performed on RNA isolated from nine independent ENCCs cultures, untreated and treated with GDNF (in total 18 samples) and eight arrays were hybridized with RNA isolated from the whole gut of embryonic mice, C57BL/6 wild-type E14.5.

### *Data analysis*

Student's *t*-tests, with Bonferroni multiple testing corrections ($p< 0.05$) were used for the single gene analysis. Gene Set Enrichment Analysis (GSEA) was performed using GSEA 2.0 (Broad Institute, Cambridge, MA, USA).[15, 16] In our analysis, we assembled genes sets from well-accepted and widely used curated pathway databases, such as the KEGG and TRANSFAC.[17, 18] We used a nominal p-value < 0.05 and a false discovery rate (FDR) < 0.10 to define sets of genes that were significantly differentially expressed between two conditions (with or without GDNF).

### *ENCCs GDNF activation and immunomodulatory capacity*

ENCCs (3–4x$10^4$cells/well) were seeded into 24-well plates. After one night of culturing, GDNF 50 ng/ml of medium (Peprotech EC) was added and incubated for 48 hours, untreated cells served as control. Supernatants were collected and secreted protein levels of Tnfα and Il-6 were measured by mouse Tnfα and Il-6 ELISAs LEGEND MAX™ (Biolegend, San Diego, USA). ELISAs were performed on three independent experiments in triplicate (n = 9).The cells from which we used the supernatant for the ELISAs were also used to isolate RNA for a qRT-PCR experiment (see below). For this, cells were washed with 1X PBS and RNA was isolated using the GeneJet™ RNA purification kit (Fermentas, St.Leon-Rot, Germany) and cDNA was synthesized by RevertAid™ H Minus First Strand cDNA Synthesis (Fermentas).

### *qRT-PCR*

RNA levels of *Tlr2, Tlr4, Tlr5, Tnfα, Il-6, Il1-β, p50 (Nf-κb component)* and *Ifnγ* were quantified by real-time PCR using the 7900HT Fast Real-Time PCR System (Applied Biosystems, Nieuwerkerk a/d IJssel, the Netherlands). Primers were designed to cross exon-exon boundaries for either the forward or the reverse primer. The length of the amplicons designed were around 51-209 bp and the annealing temperature was always 60ºC. The primers we used are listed in Supplementary Table 1. qRT-PCRs were performed on three independent experiments in triplicate (n=9) with iTaq™SYBR®Green Supermix ROX (Bio-Rad, Veenendaal, the Netherlands). For data normalization we used the expression data of the house-keeping gene GAPDH (glyceraldehyde-3-phosphate dehydrogenase). Data were analyzed using the method described by Livak.[19]

### *Immunohistochemistry*

Immunostaining was performed on cryosections (12 μm) of mouse embryonic gut E14.5. Cryosections of mouse embryonic gut were either fixed in 4% PFA (Sox10), ice-cold acetone (Ret) or 70% ethanol/glycine (Gfap) for 10 minutes at respective temperatures. When acetone was used for fixation, sections were dried for a further 4 hours. All slides were rinsed three times in PBS before applying a blocking solution (PBS containing 10% donkey serum, 0.1% Triton X-100) for 30 minutes at room temperature. For Sox10 staining, sections were additionally pretreated by boiling in citrate buffer (10mM, pH 6.0; Sigma) for 10 minutes. Primary antibodies, anti-mouse Ret, Sox10, Gfap, Tlr2, Tlr4 and Tlr5 were applied and visualized with the respective fluorochrome-conjugated secondary antibodies (see Supplementary Table 2). Cell nuclei were stained with 4'-6-diamidino-2-phenylindole [DAPI] solution (Invitrogen; 100 ng/ml) and cover-slipped with Kaiser's gelatine (Merck, Darmstadt, Germany). The images were captured on a Nikon inverted fluorescence microscope (Nikon, Dusseldorf, Germany).

### *Statistical analysis of qRT-PCR*

Statistical analysis of the qRT-PCR results was performed using the Student's *t*-test. The results represent the means ± standard deviation (SD). Statistical analysis was performed by GraphPad Prism 4. P values $<0.05$ and $<0.005$ (one-tailed distribution) were considered significant.

# RESULTS

## *ENCCs isolation*

The applied cell culture protocol resulted in almost pure ENCCs cultures as confirmed by fluorescence microscopy (see Figure 1C). After treatment with GDNF, we noticed that neurosphere-like bodies (NLBs) were larger in size (diameter: 100 μm) than the untreated control culture (diameter: 30-50 μm) (Figure 1C).

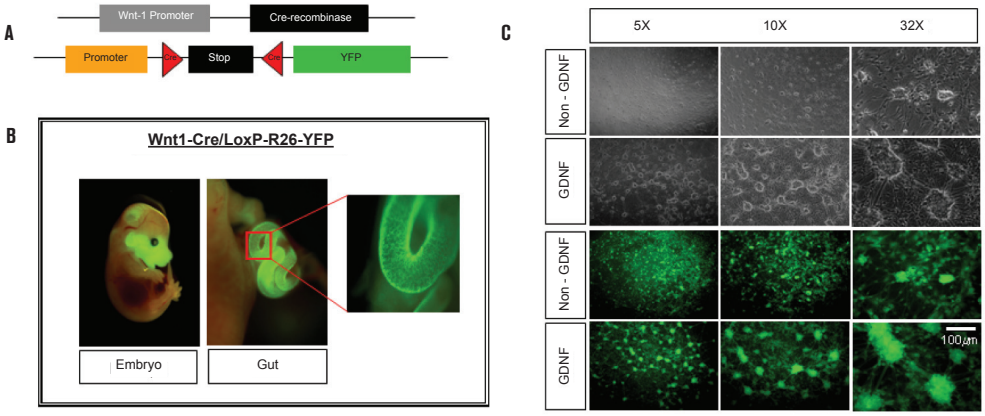



**Figure 1 –** A) Schematic overview of the *Wnt1-Cre-recombinase* and *Loxp-YFP* cassettes. B) Wnt1-Cre/Loxp-YFP embryonic mouse E14.5 showing YFP-positive cells in the neural crest-derived craniofacial structures, and in the gut. C) ENCCs after 14 days of culture, with and without GDNF treatment (5X, 10X and 32X magnification).

## Expression array analysis

## *Single Gene Analysis (Student's t-test, Bonferroni p<0.05)*

Microarray data of gene expression profiling obtained from mouse embryonic gut, ENCCs untreated and treated with GDNF were used for single gene analysis using the Student's *t*-test with Bonferroni multiple correction testing $p<0.05$. A schematic overview of the experiments and of the single gene analysis is given in Figure 2. The data collected from single gene analysis can be divided into five groups. The first group consists of genes that are significantly differentially expressed in ENCCs treated with GDNF versus those untreated. From the 428 genes in this list (see Supplementary Table 3), the top 25 are presented in Table 1. Genes that up-regulated (398 genes) were those important for the GTP metabolic process, for GTPase activity, cell-cell signaling, synaptic vesicle transport, neuron projection and genes important for neurogenesis, such as *Spock2*, *Vgf*, *Cend1* and *Dner* (Delta/Notch-like EGF-related Receptor).

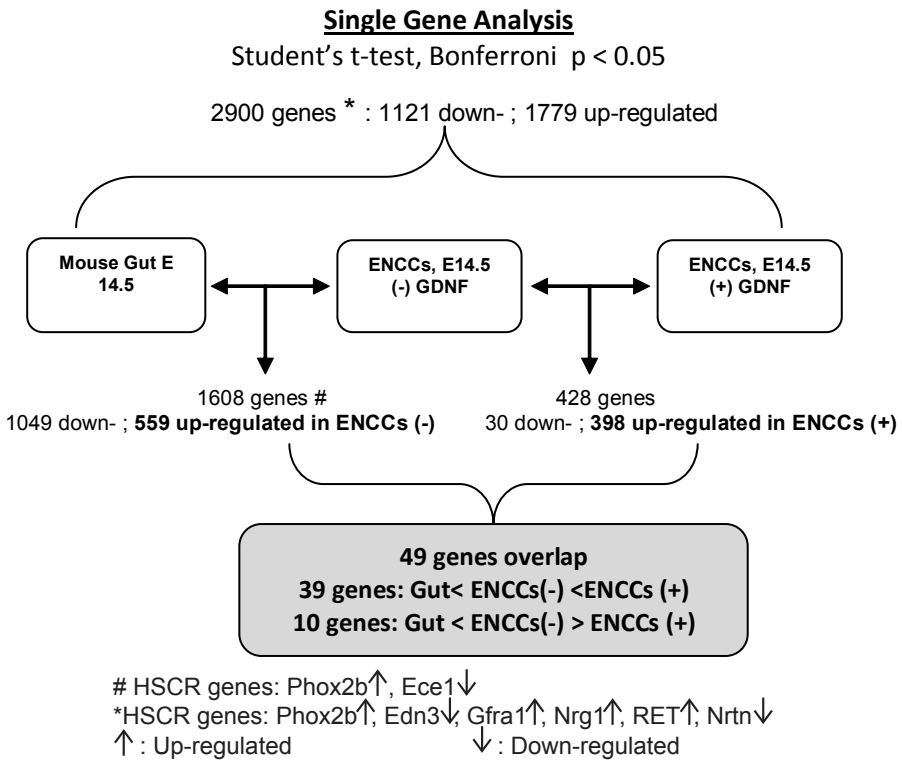



**Figure 2 –** Schematic overview of the experiments performed and the outcome of the single gene analysis

The second group consists of genes that are expressed significantly different in ENCCs compared to the gut. We identified 1608 genes belong to this group, among these were two known HSCR genes *Phox2b* and *Ece1* that were up-regulated and down-regulated in ENCCs compared to gut, respectively (see Table 2). There were 559 genes that up-regulated in ENCCs compared to gut, the top 25 genes are presented in Table 3. Among these genes are *Sox2* and *Gfrα2.*

The third group consists of genes that are up-regulated in ENCCs compared to the gut and further up-regulated in ENCCs upon GDNF treatment. There are 39 genes belonging to this group (see Supplementary Table 4). Among these genes are *Scg10 (Stmn2)* that has recently been proposed as a candidate gene for HSCR [20, 21] and *Snca* (α-synuclein) a gene known to be involved in Parkinson's disease.[22]

The fourth group consist of genes that are up-regulated in ENCCs compared to the gut, but down-regulated in ENCCs upon GDNF treatment. Only 10 such genes belong to this group (see Supplementary Table 5), including *Rora* (Retinoic acid receptor-related orphan receptor alpha), a transcription factor important for cell survival and differentiation of Purkinje cells.[23]

The fifth group consists of the genes that were expressed significantly different in ENCCs treated with GDNF compared to the gut. There were 2900 genes belong to this group. Among these were the known HSCR genes: *Ret, Phox2b, Gfrα1, Nrtn* and *Nrg1* were all up-regulated in ENCCs whereas *Edn3* was down-regulated in ENCCs (see Table 2).

**Table 1 -** Single Gene Analysis: Top 25 genes that were expressed significantly higher in ENCCs treated with GDNF compare to untreated ENCCs (Bonferroni : $p < 0.05$)

| No. | Gene Symbol | Gene Name | Human Chromosome |
|---|---|---|---|
| 1 | *Trnp1* | TMF1-regulated nuclear protein 1 | 1p36.11 |
| 2 | *Spock2* | sparc/osteonectin, cwcv and kazal-like domains proteoglycan (testican) 2 | 10q25.3 |
| 3 | *Tmem59l* | transmembrane protein 59-like | 19p12 |
| 4 | *Aplp1* | amyloid beta (A4) precursor-like protein 1 | 19q13.12 |
| 5 | *Mblac2* | metallo-beta-lactamase domain containing 2 | 5q14.3 |
| 6 | *Vgf* | VGF nerve growth factor inducible | 7q22 |
| 7 | *Dner* | delta/notch-like EGF repeat containing | 2q36.3 |
| 8 | *Rlbp1l2 ( CLVS2)* | clavesin 2 | 6q22.31 |
| 9 | *2900062L11Rik* | - | - |
| 10 | *Rnasel* | ribonuclease L (2',5'-oligoisoadenylate synthetase-dependent) | 1q25 |
| 11 | *Slc25a22* | solute carrier family 25 (mitochondrial carrier: glutamate), member 22 | 11p15.5 |
| 12 | *A730017C20Rik (no human ortholog* | - | - |
| 13 | *Rnf157* | ring finger protein 157 | 17q25.3 |
| 14 | *Resp18* | regulated endocrine-specific protein 18 | 2q35 |
| 15 | *B3galnt1* | beta-1,3-N-acetylgalactosaminyltransferase 1 (globoside blood group) | 3q25 |
| 16 | *Zdhhc22* | zinc finger, DHHC-type containing 22 | 14q24.3 |
| 17 | *Cyp4x1* | cytochrome P450, family 4, subfamily X, polypeptide 1 | 1p33 |
| 18 | *Stard7* | StAR-related lipid transfer (START) domain containing 7 | 2p11.1 |
| 19 | *Wnk3* | WNK lysine deficient protein kinase 3 | xp11.22 |
| 20 | *6530401D17Rik (human orth: SPCS3)* | signal peptidase complex subunit 3 homolog (S. cerevisiae) | 4q34.2 |
| 21 | *A730017C20Rik* | - | - |
| 22 | *Pmvk* | phosphomevalonate kinase | 1q21.3 |
| 23 | *Samd10* | sterile alpha motif domain containing 10 | 20q13.33 |
| 24 | *Cend1* | cell cycle exit and neuronal differentiation 1 | 11p15.5 |
| 25 | *Cmas* | cytidine monophosphate N-acetylneuraminic acid synthetase | 12p12.1 |

**Table 2 -** Observed expression differences of HSCR-associated genes

| HSCR Gene | ENCCs/Gut | p-Value | ENCCs(+GDNF)/Gut | p-Value |
|---|---|---|---|---|
| *Ret* | - | - | Up-regulated | 0.00025 |
| *Phox2b* | Up-regulated | 0.03796 | Up-regulated | 0.00375 |
| *Edn3* | - | - | Down-regulated | 0.00106 |
| *Gfra1* | - | - | Up-regulated | 0.00125 |
| *Ece1* | Down-regulated | 3.14E-08 | - | - |
| *Nrg1* | - | - | Up-regulated | 2.02E-12 |
| *Nrtn* | - | - | Down- | 0.00789 |

**Table 3 -** Single Gene Analysis: Top 25 genes that were expressed significantly higher in ENCCs compared to the gut (Bonferroni : $p < 0.05$)

| No. | Gene Symbol | Gene Name | Human Chromosome |
|---|---|---|---|
| 1 | *Ptpn21* | protein tyrosine phosphatase, non-receptor type 21 | 14q31 |
| 2 | *Mdm2* | transformed mouse 3T3 cell double minute 2 | 12q13-q14 |
| 3 | *Cdkn1a* | cyclin-dependent kinase inhibitor 1A | 6p21.1 |
| 4 | *Tcfap2a* | transcription factor AP-2 alpha (activating enhancer binding protein 2 alpha | 6p24.3 |
| 5 | *(C4orf34)* | - | - |
| 6 | *Atf6* | activating transcription factor 6 | 1q22-q23 |
| 7 | *Afap1* | actin filament associated protein 1 | 4p16 |
| 8 | *Trim11* | tripartite motif-containing 11 | 1q42.13 |
| 9 | *BC031353* | KIAA1370 | 15q21.2-q21.3 |
| 10 | *Rora* | Retinoic acid receptor-related orphan receptor alpha | 15q22.2 |
| 11 | *Pcdh7* | protocadherin 7 | 4p15 |
| 12 | *Tm9sf3* | transmembrane 9 superfamily member 3 | 10q24.2 |
| 13 | *1110059G02Rik* | - | - |
| 14 | *Cpne2* | copine II | 16q13 |
| 15 | *Cep170* | centrosomal protein 170kDa | 1q44 |
| 16 | *Tmem111* | tr ansmembrane protein 111 | 3p25.3 |
| 17 | *Nek7* | NIMA (never in mitosis gene a)-related kinase 7 | 1q31.3 |
| 18 | *Sox2* | SRY (sex determining region Y)-box 2 | 3q26.33 |
| 19 | *Bnip2* | BCL2/adenovirus E1B 19kDa interacting protein 2 | 15q21.3 |
| 20 | *Casp12* | caspase 12 (gene/pseudogene) | 11q22.3 |
| 21 | *Ndst1* | N-deacetylase/N-sulfotransferase (heparan glucosaminyl) 1 | 5q33.1 |
| 22 | *App* | Amyloid beta (A4) precursor protein | 21q21.2 |
| 23 | *Gfra2* | GDNF family receptor alpha 2 | 8p21.3 |
| 24 | *Lphn3* | latrophilin 3 | 4q13.1 |
| 25 | *1110032E23Rik* | - | - |

We identified differential expression of a number of transcription factors that are known to play an important role in embryonic and ENS development, such as transcription factors belonging to the Sox (SRY-related HMG-box), Fox (Forkhead) and Hox (homeobox) transcription factors family. We also identified differential expression of *Dlx1* and *Dlx2* which are required for coordinating and programing neurite maturation and migration[24] and Etv1 that is known to play an important role in differentiation of dopaminergic neurons in *C.elegans* and mouse.[25] The summary of transcription factor expression differences is presented in Table 4.

**Table 4 -** Transcription factors that up- and down-regulated in ENCCs (+GDNF)/ENCCs, ENCCs/Gut and ENCCs(+GDNF)/Gut

| Gene | ENCCs(+GDNF) /ENCCs | p-Value | ENCCs/ Gut | p-Value | ENCCs(+GDNF) /Gut | p-Value |
|---|---|---|---|---|---|---|
| *Sox2* | - | - | Up- | 3.30E-07 | Up- | 0.00280 |
| *Sox4* | - | - | Up- | 0.0134 | - | - |
| *Sox6* | - | - | Up- | 0.00247 | - | - |
| *Sox17* | - | - | Down- | 0.02739 | - | - |
| *Sox18* | - | - | Down- | 0.02430 | Down- | 0.00163 |
| *Sox21* | - | - | - | - | Down- | 0.00085 |
| *Phox2b* | - | - | Up- | 0.03796 | Up- | 0.00375 |
| *Foxk2* | Up- | 0.00514 | - | - | - | - |
| *Foxf1a* | - | - | - | - | Down- | 0.04777 |
| *Foxp2* | - | - | - | - | Down- | 0.04632 |
| *Foxd3* | - | - | Up- | 0.00092 | - | - |
| *Dlx1* | - | - | - | - | Up- | 5.87E-05 |
| *Dlx2* | - | - | Up- | 0.01264 | - | - |
| *Hoxb4* | - | - | - | - | Up- | 0.00205 |
| *Hoxb5* | - | - | - | - | Up- | 6.21E-05 |
| *Hoxa9* | - | - | - | - | Down- | 0.013288 |
| *Hoxc8* | - | - | - | - | Down- | 0.007206 |
| *Etv1* | - | - | - | - | Up- | 4.13E-07 |
| *Etv5* | - | - | - | - | Up- | 3.35E-05 |
| *Ets2* | - | - | - | - | Down- | 0.041972 |
| *Cdx1* | - | - | Down- | 0.009328 | Down- | 0.003229 |
| *Cdx2* | - | - | Down- | 0.003811 | Down- | 0.003811 |

## Pathways Analysis using the Gene Set Enrichment Analysis (GSEA) method

The pathway analysis we performed showed that some pathways previously implicated in ENS as well as in embryonic development, such as Tgfβ/Bmp, Notch and Wnt were down-regulated in ENCCs on GDNF stimulation (Table 5). The Notch signaling pathway for instance is known to be involved in various aspects of neurogenesis, although it is also involved in cell proliferation.[26, 27] Within the Notch pathway we saw down-regulation of *Notch1*, *Notch2*, *Jag1*, and the Notch co-activators *Maml1* and *Maml2*. With respect to the Wnt signaling pathway, it is known that Wnt signaling in ENCCs is important for cell lineage decision-making rather than maintaining cells in their progenitor state.[28] *Wnt* (2, 5B, 9A and 10B), and Wnt receptor *Fzd* (2 and 6) are some of the genes in the Wnt signaling pathway that were down-regulated by GDNF. The Tgf-β/Bmp signaling pathway plays a role in many cellular processes including cell proliferation, apoptosis, differentiation and migration. *Tgf-β*, *Tgfβr1*, *Smads*, *Bmps*, *Bmpr1*, *Rock1* are some of the Tgf-β/Bmp signaling pathway genes that are down-regulated by GDNF. Rock signaling is involved in regulating neural precursor cell (NPC) migration. Knock down of *Rock1* and *Rock2* in mouse promotes NPC migration to the olfactory bulb.[29] With respect to Bmps, Bmp2 and Bmp44, these proteins are known to be involved in neuronal differentiation of post-migratory enteric neural crest.[30, 31] Bmp2 is also known as a stimulator of Mash1 expression, which is an important protein for neurogenesis.[32] Rock1, together with the Semaphorins and the Plexin protein family (a Semaphorin co-receptor) are all part of the axon guidance signaling pathway. We found that this axon signaling pathway is down-regulated during GDNF stimulation, suggesting that neuronal differentiation of ENCCs is suppressed upon RET activation.

Adherens junctions, tight junctions (cell-cell adhesion) and focal adhesion (cell-matrix adhesion) play an important role in cell migration. In this study, genes that belong to these signaling pathways were down-regulated on GDNF treatment. Down-regulation could lead to loss of, or less tight, cell-cell and cell-matrix contact, a crucial process for cell migration and proliferation. *Tjp1* (tight junction 1) and *Snai2* (Slug) are genes that were down-regulated and these are members of the tight and adherens junction signaling pathways, respectively. Slug is known to play a role in ENCCs migration and homozygous deletion leads to the development of Waardenburg syndrome type 2D, an auditory-pigmentary syndrome.[33]

We also observed that cell cycle signaling pathways are down-regulated on GDNF stimulation. Histone deacetylases (*Hdac1* and *Hdac2*), *Mdm2*, *Cdkn1a* (*p21*), *Cdkn2b* (*p15*), *Rb1* are all on the list of cell cycle signaling genes that are down-regulated on GDNF treatment. HDACs have a role in cell growth arrest, differentiation and cell death, while Mdm2 inhibits the G0/G1-S phase transition of human diploid cells, and over-expression of MDM2 would induce

growth arrest in normal cells. Mdm2, p21, p15 and Rb1 are negative regulators of the cell cycle. Down-regulation of the expression of these genes will most likely reduce their growth-inhibitory function and induce cell proliferation as a consequence. Although GDNF treatment results in a down-regulation of many genes involved in the cell cycle and cell growth, there are also some down-regulated genes that play a role in cell division, such as *Ccnd* (*cyclin D*) and *Cdks* (*cyclin-dependent kinase*). The harmonization of negative and positive regulators of the cell cycle process eventually results in cell division or cell cycle arrest. In this context it is interesting to note that the ENCCs cultures with and without GDNF (Figure 1C) show a clear difference, namely that GDNF-treated ENCCs have bigger Neurosphere-Like Bodies (NLBs) than those not treated with GDNF, pointing towards the higher proliferative state of these treated cells

On the other hand, signaling pathways that play pivotal roles in apoptosis, such as the p53 and Toll-like receptors (Tlrs) pathways are down-regulated on GDNF treatment. The down-regulation of genes involved in cell cycle arrest and apoptosis upon Ret activation could point toward a model in which ENCCs proliferate and survive better during ENS development.

Finally the JAK-STAT signaling pathway is known to be activated by RET.[34] Our microarray results, to our surprise, showed that this pathway is down-regulated on GDNF treatment. However, most of the genes that are down-regulated in this signaling pathway are ones important for cytokines and their receptors' expression. Among these are *Il-15, Il-12b, Il-6r, Il-28R, Ifnαρ1,* and *Ifnγr1*. Together with the Tlrs signaling pathways, it seems that GDNF activation reduces the immunomodulatory capacity of ENCCs.

**Table 5 -** Pathways analysis (ENCCs_GDNFneg Vs ENCCs_GDNFpos)

| Database | Pathway | P-Value | FDR | #Genes |
|---|---|---|---|---|
| **KEGG** | Up-regulated in GDNF (+) | | | |
| | Citrate Cycle | 0.01 | 0.06 | 12 |
| | Oxidative phosphorylation | < 0.001 | 0.025 | 39 |
| | Peroxisome | 0.001 | 0.040 | 19 |
| | Steroid Biosynthesis | 0.004 | 0.039 | 6 |
| | Valin-Leucin-Isoleucin | 0.006 | 0.070 | 11 |
| | PPAR | 0.002 | 0.049 | 15 |
| | Butanoate metabolism | 0.009 | 0.029 | 16 |
| | Propanoate metabolism | 0.005 | 0.079 | 12 |
| | Parkinson's disease | <0.001 | 0.043 | 35 |
| | Down-regulated in GDNF (+) | | | |
| | ECM Receptor | 0.005 | 0.075 | 21 |
| | Adherens Junction | < 0.001 | < 0.001 | 28 |
| | Axon Guidance | < 0.001 | 0.052 | 44 |
| | Cell cycle | < 0.001 | 0.056 | 53 |
| | Chronic Myeloid | < 0.001 | 0.054 | 38 |
| | Focal Adhesion | < 0.001 | 0.026 | 68 |
| | JAK-STAT | 0.001 | 0.084 | 56 |
| | Lysosome | 0.003 | 0.089 | 24 |
| | N-Glycan Biosynthesis | 0.01 | 0.076 | 17 |
| | Notch Signaling | < 0.001 | 0.001 | 19 |
| | P53 Signaling | < 0.001 | < 0.001 | 32 |
| | Ribosome | < 0.001 | < 0.001 | 58 |
| | Spliceosome | < 0.001 | 0.007 | 45 |
| | TGF-β Signaling | < 0.001 | < 0.001 | 30 |
| | Tight Junction | 0.005 | 0.08 | 41 |
| | TOLL-Like Receptor | 0.009 | 0.08 | 40 |
| | Wnt Signaling | 0.003 | 0.07 | 42 |
| | Pathaways in Cancer | < 0.001 | 0.057 | 95 |
| | Splicesome | < 0.001 | 0.076 | 45 |
| **TRANSFAC** | Up-regulated in GDNF (+) | | | |
| | SF1 | < 0.001 | 0.005 | 41 |
| | ATF3 | 0.002 | 0.017 | 34 |
| | TCF11 (MAFG) | 0.006 | 0.082 | 41 |
| | AHR-ARNT | 0.007 | 0.016 | 5 |
| | EGR3 | 0.02 | 0.04 | 11 |
| | LXR | 0.015 | 0.033 | 10 |
| | NFE2 | < 0.001 | 0.015 | 50 |
| | NRSF (a.k.a REST) | < 0.001 | < 0.001 | 23 |
| | PAX4 | 0.01 | 0.10 | 30 |
| | RFX1 | < 0.001 | 0.017 | 35 |
| | Down-regulated in GDNF (+) | | | |
| | CDX2 | < 0.001 | 0.008 | 63 |
| | CEBP | < 0.001 | 0.002 | 69 |
| | PPARg | < 0.001 | 0.004 | 20 |
| | SOX5 | < 0.001 | 0.008 | 70 |
| | TEF1 | < 0.001 | 0.008 | 66 |
| | CIZ1 | < 0.001 | < 0.001 | 60 |
| | ETS | < 0.001 | 0.007 | 70 |
| | FOXO1 | < 0.001 | 0.006 | 77 |
| | FREAC2 | < 0.001 | 0.007 | 70 |
| | HP1 | < 0.001 | 0.007 | 61 |
| | IPF1 | < 0.001 | 0.008 | 83 |
| | NFAT | < 0.001 | 0.008 | 88 |
| | NFYC | < 0.001 | 0.005 | 55 |
| | PAX5 | < 0.001 | 0.004 | 41 |

FDR : False Discovery Rate ; # Number of genes in leading edge subset of GSEA results

## GDNF stimulation significantly reduces Tlr2, Tlr4, Tlr5, Tnfα and Il-6 mRNA levels in ENCCs

Pathway analysis showed that the Tlr signaling pathway is down-regulated in ENCCs on GDNF stimulation (see Table 5). In particular, among the genes listed in the Tlr pathway, Tlr2 and Tlr5 were notably down-regulated after GDNF treatment. To confirm these data we performed qRT-PCR on *Tlr2* and *Tlr5* in ENCCs treated and untreated with GDNF. As it is known from previous studies that Tlr4 is also expressed in the myenteric plexus[35, 36], we included *Tlr4* in this experiment as a positive control. As activation of Tlr signaling is known to activate Nf-κb (p65-p50 complex) and initiate the innate immune response by inducing expression of pro-inflammatory cytokines, we also measured the mRNA level of several pro-inflammatory cytokines such as Tnfα, Ifnγ, Il-6 and Il-1β.

d significantly reduced in ENCCs on GDNF stimulation compared to the untreated cells. The results are presented in Figure 3A-B. Tnfα expression levels in the ENCCs were not detectable by ELISA. Il-6, which was detectable by ELISA, showed no significant difference in the level of expression between the control and the GDNF-treated cells.

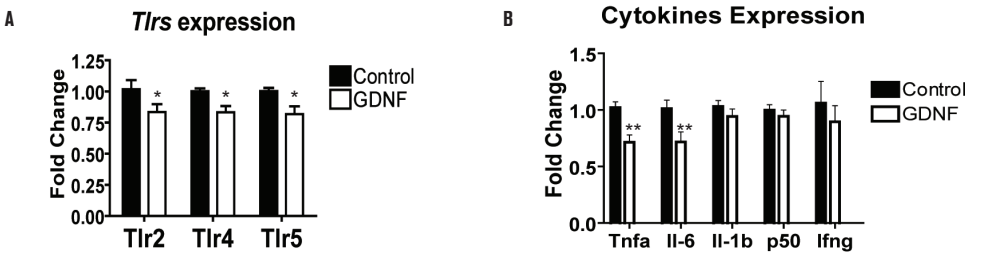
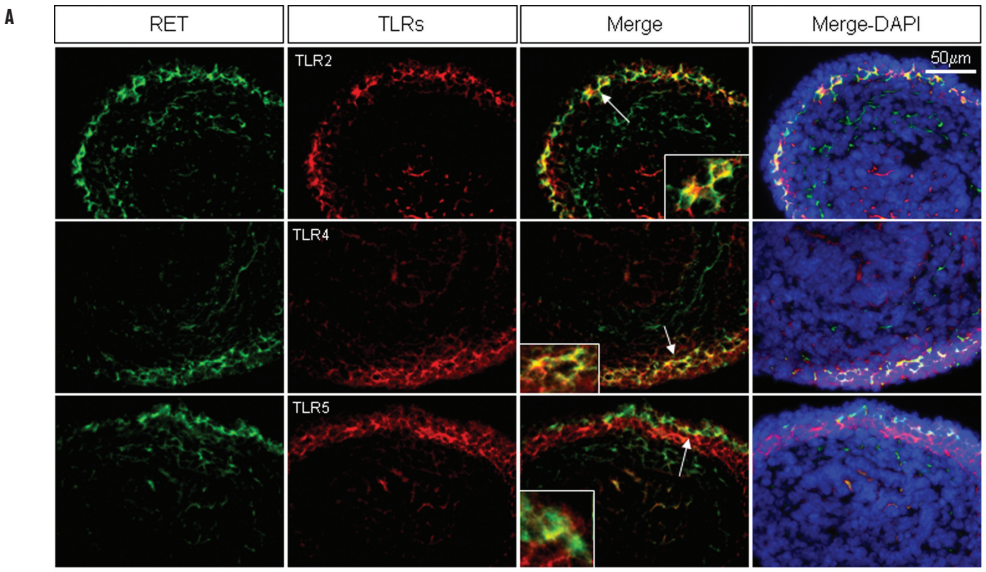
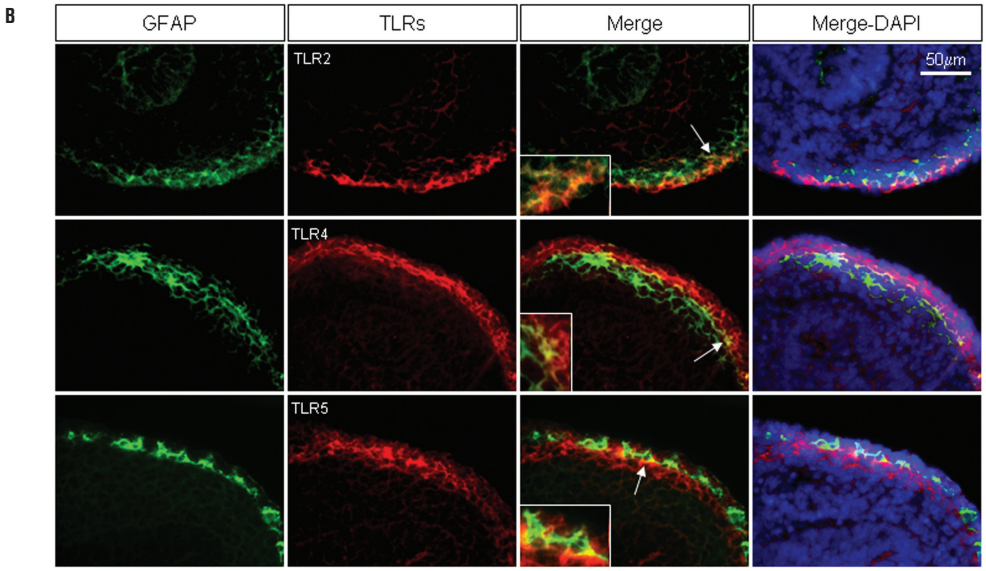



**Figure 3 -** GDNF significantly reduces the mRNA level of: A) *TLR2*, *TLR4* and *TLR5* in ENCCs, B) *TNFα* and *Il-6* in ENCCs. The bars indicate the means ± SD of the fold changes of the mRNA level for each gene between control and GDNF treated cells. *p < 0.05 and **p < 0.005 were considered significant (one-tailed distribution).

## Tlr2, 4 and 5 are expressed in enteric glial, neurons and ENS progenitor cells of the embryonic myenteric plexus

Tlr2, Tlr4 and Tlr5 are expressed in the myenteric plexus of mouse embryonic colon E14.5. In particular, the Tlrs could be demonstrated in the glial cells, as shown by co-staining with Gfap, and in neurons as shown by co-staining with Ret. Moreover they were already detected in the ENS progenitors as shown by co-staining with Sox10 (Figure 4A-C). In all the stainings, Tlrs did not overlap completely with Ret, Gfap or Sox10. It seems that parts of the smooth muscle layer were also positively-stained for Tlrs, especially for Tlr5. This is in line with previous reports describing similar tissue expression of Tlr4.[36]


A
RET
TLRs
Merge
Merge-DAPI
TLR2
TLR4
TLR5
50μm


B
GFAP
TLRs
Merge
Merge-DAPI
TLR2
TLR4
TLR5
50μm

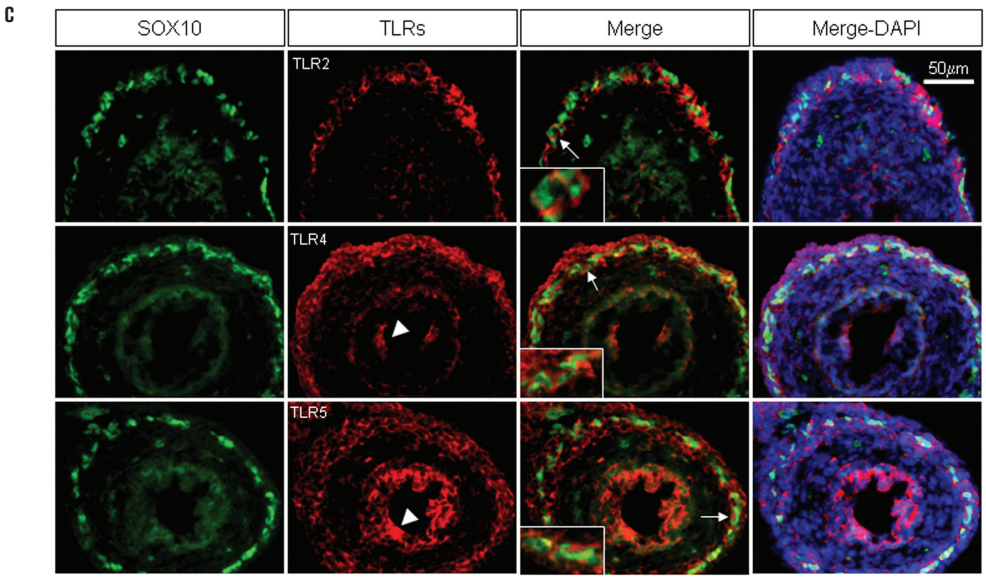



**Figure 4 -** Tlr2, Tlr4 and Tlr5 are expressed in the myenteric plexus of embryonic mouse gut E14.5. A) Co-staining of Tlrs with Ret, B) Co-staining of Tlrs with the glial marker (Gfap), and C) Co-staining of Tlrs with ENS progenitor marker Sox10. The merged images of each Tlr with Ret, Gfap and Sox10 staining are shown in the third column of each panel. The magnifications of co-stained Tlrs are shown in the respective insets. Tlr4 and Tlr5 are also expressed on colon epithelial cells as shown by arrowheads in Figure C.

## DISCUSSION

Several genes have been implicated in HSCR development, including *RET*, *GDNF*, *EDNRB*, *EDN3*, *SOX10*, *NTN*, *ECE1*, *ZFXH1B*, *KBP*, *TTF1* and *NRG1*. The majority of these genes can be divided into three groups: the ones involved in the RET pathways (*RET*, *GDNF*, *NTN*), the ones involved in EDNRB pathways (*EDNRB*, *EDN3*, *ECE-1*), and transcription factors that can affect both RET and/or EDNRB pathways (*SOX10*, *ZFXH1B*, *PHOX2B*). Increasing evidence has shown that the proteins encoded by these genes are interconnected. A genome-wide association study, using 43 trios from a genetically isolated Mennonite population, reported statistically significant joint transmission of RET and EDNRB alleles.[37] Furthermore, mice homozygous for the recessive hypomorphic allele of *Ednrb* (EdnrbS or Piebald) and heterozygous for a *Ret* null mutation ($Ret^{+/-}$) showed high frequencies of aganglionosis. These findings are suggestive of epistasis between EDNRB and RET and this idea was further supported by *Ret*/*Ednrb* mouse crosses.[38] Similar studies showed that, whereas $Ret^{51/51}$ and $Edn3^{ls/ls}$ (lethal

spotting) mice display colonic aganglionosis, combinations of these mutant alleles lead to almost complete intestinal aganglionosis.[39]

Furthermore, the transcription factor SOX10 is involved in the regulation of both *RET* and *EDNRB* expression.[40-42] PHOX2B is also a transcription factor that is expressed in several classes of differentiating neurons of both the peripheral and central nervous systems. Mice with a homozygous disruption of *Phox2b* show lack of enteric ganglia and no Ret expression which suggests that Phox2b might play a regulatory role in *Ret* expression.[43] All these data indicate a central role for RET in the development of both HSCR and ENS. It was these findings that encouraged us to undertake this study. By determining the *RET*-associated gene networks in ENS progenitors we hoped to gain insight not only into ENS development but also a better understanding of how HSCR might develop, eventually leading to HSCR candidate genes.

To gain such insight we performed gene expression profiling studies on RNA isolated from mouse ENCCs treated and untreated with GDNF, the ligand for the RET receptor, and compared these data to each other and to the expression profile from the whole mouse embryonic gut (all E14.5). The data were used to perform pathway analysis using GSEA and for a single gene analysis by a Student's *t*-test.

## Single Gene Analysis

With our single gene analysis we identified a large number of genes that were differentially expressed in ENCCs upon GDNF treatment. Moreover, the gene expression profiles from the mouse gut helped us to identified ENCC-specific genes. The number of genes we identified, as differentially expressed in these experiments was very high, even after multiple testing correction. The reason we did identify so many significantly differentially expressed genes is mainly due to the large number of replicates we analysed.

The data generated gives new insights in which genes are important for ENS, development, hence important for HSCR development. By comparing genes profile of ENCCs to that of mouse embryonic gut, we identified two genes that directly can be linked to RET namely, Phox2b which acts as a transcription factor for Ret, [43, 44]and Gfr$\alpha$2 that is a known co-receptor of Ret.45 Moreover, we also identified Sox2 which in a recent study proved to be an excellent ENS progenitor cell marker.[46] Genes belonging to this group, i.e genes expressed in ENCCs but not in the gut, might therefore be considered candidate ENS progenitor markers. Among the down-regulated genes, the known HSCR gene, *Ece1,* was identified. Furthermore, by comparing genes profile of ENCCs treated with GDNF to that of the mouse embryonic gut, we identified six of 12 known-HSCR genes (see Table 2).

We, also identified genes that were up-regulated in ENCCs compared to those in the gut and further up-regulated or down-regulated upon GDNF treatment (see Supplementary Table 4 and 5). It suggests that these genes are ENCCs specific and are downstream effectors of Ret. Snca *(α-synuclein)* was among the genes that were up-regulated (higher in NCC when compared to the gut and further up-regulated upon GDNF treatment). This suggests that defects in RET signaling will most likely reduce the expression of α-synuclein. Interestingly, in Parkinson's patients, the presence of Lewy bodies (aggregation of α-synuclein) in their gut is associated with chronic constipation. These findings suggest that α-synuclein might play an important role in ENS functioning.[47, 48]

Several comparable expression profiling studies, all aiming at a better understanding of ENS and HSCR development and to find new candidate HSCR genes, have been performed. Different study designs have been used: Heanue and colleagues compared gene expression of mouse embryonic gut from wild type (WT) mice with those of *Ret* knock-out mice ($Ret^{-/-}$)[49]; Vohra and colleagues performed a study with a design comparable to Heanue et al. However, they analysed mouse intestines of two different developmental stages (E14 and P0)[50]; Iwashita and colleagues compared gene expression profiles of neural crest stem cells (NCSCs) with the profiles made of total mouse embryos while Ngan and colleagues compared gene expression profiles of ENCCs isolated from mice WT with those from ENCCs isolated from *Ret* mutant mice (E11.5). These ENCCs were treated and untreated with GDNF at two different time-points (8 and 16 hours) before RNA isolation.[51, 52] Each study has its own advantages and disadvantages. The use of knock-out mice (Heanue) clearly has an advantage as also the influence of other endonegous RET ligands can be measured. However using gut instead of ENCCs might be considered a drawback as the ENS progenitor cells are less than 5% of the total cell population in the mouse embryonic gut. The use of NCSCs without culturing as described by Iwashita has an advantage over cultured ENCCs (Ngan et al. and our study) as it omits the potential of *in vitro* influences on gene expression. However to study the RET-GDNF genes network in ENCCs (or NCSCs) without GDNF treatment might not cover completely the dynamic changes of RET-dependent genes expression. Furthermore, the use of expression profiles from gut (our study) compared to those from total embryos (Iwashita) for instance might have some advantages as differences in expression of genes in non-ENCCs of the gut might be identified as well. The study by Ngan et al. was very similar to our study, although they did not compare the ENCCs profiles with the expression profile of the gut or total embryo. Because of this it is difficult to identify ENCCs specific genes. In our study we could detect six of the 12 known-HSCR genes by comparing the gene expression profiles of ENCCs (+GDNF) to the expression profiles of the gut, whereas ENCCs treated versus un-

treated did not detect any of these 12 known-HSCR genes. This shows that this comparison is very informative. The strength of our study is the use almost pure ENCCs isolated from mice carrying Wnt1-Cre-recombinase and a LoxP/R26-YFP, and the fact that we used large numbers of replicates. The last resulted in, even after Bonferroni multiple testing corrections ($p< 0.05$), large numbers of significantly down or up regulated genes.

When comparing our results with those generated by the other studies we saw that 70% of genes (33/47) that were thoroughly examined by Heannue et al. (expression in the ENS by *in situ* hybridization) were present in our data set. Among these are *Tgfβ2, Mab21l1, Sncg, Etv1, L1cam* and *Fgf13.* From the 38 genes thoroughly examined by Vohra et al (expression in the ENS by *in situ* hybridization) 52% of genes (20/38) proved to be present in our gene set. Among these genes are *Elavl4, Ncam2, Hoxb5, Ndrg4, Stmn3* and *Dlx2.* The comparison with the study of Iwashita et al. we saw 27% of genes (29/107 annotated genes) were present in our data. Among these are *Ret, Gfrα1, Gfrα2, Emb* and *Syngr1.* The comparison with the study of Ngan et al., we saw 58.6% of genes (17/29) were present in our data. Among these are *Mapk8ip2, Tesc, Kcnd2* and *Vip.*

These comparisons show considerable overlap between the studies. We did not find major differences between these 4 studies and our study. The differences found, different gene sets, most likely are the result of the different study designs used.

## Pathway analysis

GSE analysis based on the data of the KEGG database showed that upon GDNF stimulation, several sets of genes were significantly up- or down-regulated in the ENCCs (results shown in Table 5, see also supplementary Figure 1A). The Pathways that were up-regulated were those which are important for energy supply (citrate cycle and oxidative phosphorylation), for steroid and valin-leucin-isoleucin biosynthesis and lipid metabolism (peroxisome and peroxisome proliferators-activated receptors/ PPARs). Signaling pathways that are considered important for ENS development such as the Notch, Tgfβ/Bmp and Wnt signaling were all significantly down-regulated upon GDNF stimulation. Sets of genes important for cell-cell and cell-matrix contacts (e.g. adherens junctions, tight junction and focal adhesion formation), cell cycle arrest, apoptosis (p53, Toll-like receptor signaling) and innate immune system signaling (Toll-like receptor signaling) were also down-regulated. We summarize the possible biological function of the signaling pathways mentioned above in the supplementary Figure 1B.

Pathway analysis based on the TRANSFAC database showed similar results to those collected from the KEGG database. Genes under the regulation of transcription factors that are

important for determining cell differentiation, such as Sox5 and Foxo1, are down-regulated by GDNF, while genes under the regulation of Egr3, a transcription factor that is important for cell growth and cell migration, are up-regulated on GDNF treatment. The same goes for Rest, a transcription repressor which is highly expressed in the neuronal progenitor cells and which is known as a negative master regulator of neurogenesis.[53, 54]

## TLRs signaling pathway

TLRs recognize viral and bacterial infections and induce an innate immune response by activating pro-inflammatory cytokines such as TNF$\alpha$. Furthermore, several studies have shown that TLRs can also induce apoptosis through the activation of Caspases (as reviewed by B. Salaun).[55]

Recently it was shown that TLR4 is expressed in neurons of the myenteric plexus of the human illieum.[36] Tlr3 and Tlr7 are expressed in the ENS and in the dorsal root ganglia (DRG).[35] Our microarray data and our immunostainings show that Tlr2, Tlr4 and Tlr5 are expressed in ENCCs and in the ENS (the progenitors, neurons and glial cells, see Figure 4A-C). The expression of Tlrs in the ENS precursor cells could indicate that these cells play a role in the immunomodulatory mechanism against viral and bacterial infection. This might seem illogical as in the normal situation the ENS is not directly in contact with the luminal contents. However, bacteria and viruses can cross the epithelial barrier and reach the deeper layers of cells, such as the ENS, and thereby induce an immune response in these cells.[56, 57] Moreover, we show that all three Tlrs were down-regulated upon GDNF stimulation, implying that these particular *Tlrs* genes are down-stream effectors of Ret signaling. Based on these data, we hypothesize that RET-GDNF activation will reduce the expression of TLRs and, consequently, will also reduce the expression of cytokines. Therefore, defects in RET signaling, as often seen in HSCR patients, will result in higher expression levels of TLRs, which subsequently lead to an overexpression of cytokines. This condition might well contribute to activating the innate immune response system, which might manifest itself as HSCR-associated enterocolitis. This idea is corroborated by an animal study which demonstrated that mRNA levels of *Tlr2* and *TlrR4* are increased in the intestinal epithelial cells (IEC) of enterocolitis rat.[58] Furthermore, it has also been shown that TLR4 expression is elevated in the intestinal mucosa of humans and that TLR4 plays a critical role in the development of necrotizing entercolitis .[59]

In summary, there seems to be a general trend in our data that genes encoding proteins that stimulate cell differentiation or that are negative regulators of cell proliferation, and genes that inhibit cell-cell adhesion and apoptosis were generally down-regulated in neural crest cells upon treatment with GDNF. On the other hand, genes that encode proteins that have the

opposite effect, such as genes for an early growth response, those involved in the inhibition of cell differentiation, and those that are important for energy supply and biosynthesis, were up-regulated. Furthermore, in the ENCCs we saw a connection between Ret signaling and the innate immune response, as stimulation of Ret resulted in down-regulation of several Tlrs and their pro-inflammatory cytokines. These data provide new insight into the gene networks activated upon RET stimulation by GDNF in ENCCs. They also will serve as a catalog of genes possibly involved in HSCR. This might will prove extremely helpful once we start sequencing exomes or even genomes of HSCR patients and revealing enormous numbers of DNA variants, most of which will not in fact play a role in HSCR.


## Acknowledgements

We thank Jackie Senior for editing the manuscript. This study was supported by grant from Dutch Digestive Foundation (MLDS project WO 09-62).


# REFERENCES


1. Durbec PL, Larsson-Blomberg LB, Schuchardt A, et al. Common origin and developmental dependence on c-ret of subsets of enteric and sympathetic neuroblasts. Development 1996;122:349-358.

2. Druckenbrod NR, Epstein ML. The pattern of neural crest advance in the cecum and colon. Dev Biol 2005;287:125-133.

3. Kapur RP, Yost C, Palmiter RD. A transgenic model for studying development of the enteric nervous system in normal and aganglionic mice. Development 1992;116:167-175.

4. Wallace AS, Burns AJ. Development of the enteric nervous system, smooth muscle and interstitial cells of Cajal in the human gastrointestinal tract. Cell Tissue Res 2005;319:367-382.

5. Taraviras S, Pachnis V. Development of the mammalian enteric nervous system. Curr Opin Genet Dev 1999;9:321-327.

6. Young HM, Newgreen D. Enteric neural crest-derived cells: origin, identification, migration, and differentiation. Anat Rec 2001;262:1-15.

7. Amiel J, Lyonnet S. Hirschsprung disease, associated syndromes, and genetics: a review. J Med Genet 2001;38:729-739.

8. Hofstra RM, Wu Y, Stulp RP, et al. RET and GDNF gene scanning in Hirschsprung patients using two dual denaturing gel systems. Hum Mutat 2000;15:418-429.

9. Taraviras S, Marcos-Gutierrez CV, Durbec P, et al. Signalling by the RET receptor tyrosine kinase and its role in the development of the mammalian enteric nervous system. Development 1999;126:2785-2797.

10. Schuchardt A, D'Agati V, Larsson-Blomberg L, Costantini F, et al. Defects in the kidney

and enteric nervous system of mice lacking the tyrosine kinase receptor Ret. Nature 1994;367:380-383.

11. Sanchez MP, Silos-Santiago I, Frisen J, et al. Renal agenesis and the absence of enteric neurons in mice lacking GDNF. Nature 1996;382:70-73.

12. Pichel JG, Shen L, Sheng HZ, et al. Defects in enteric innervation and kidney development in mice lacking GDNF. Nature 1996;382:73-76.

13. Enomoto H, Araki T, Jackman A, et al. GFR alpha1-deficient mice have deficits in the enteric nervous system and kidneys. Neuron 1998;21:317-324.

14. Alves MM, Osinga J, Verheij JB, et al. Mutations in SCG10 are not involved in Hirschsprung disease. PLoS One 2010;5:e15144.

15. Mootha VK, Lindgren CM, Eriksson KF, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 2003;34:267-273.

16. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545-15550.

17. Dahlquist KD, Salomonis N, Vranizan K, et al. GenMAPP, a new tool for viewing and analyzing microarray data on biological pathways. Nat Genet 2002;31:19-20.

18. Kanehisa M, Goto S, Hattori M, et al. From genomics to chemical genomics: new developments in KEGG. Nucleic Acids Res 2006;34:D354-D357.

19. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Method 2001;25:402-408.

20. Alves MM, Burzynski G, Delalande JM, et al. KBP interacts with SCG10, linking Goldberg-Shprintzen syndrome to microtubule dynamics and neuronal differentiation. Hum Mol Genet 2010;19:3642-3651.

21. Korbel JO, Tirosh-Wagner T, et al. The genetic architecture of Down syndrome phenotypes revealed by high-resolution analysis of human segmental trisomies. Proc Natl Acad Sci U S A 2009;106:12031-12036.

22. Spillantini MG, Crowther RA, Jakes R, et al. alpha-Synuclein in filamentous inclusions of Lewy bodies from Parkinson‘s disease and dementia with lewy bodies. Proc Natl Acad Sci U S A 1998;95:6469-6473.

23. Boukhtouche F, Doulazmi M, Frederic F, et al. RORalpha, a pivotal nuclear receptor for Purkinje neuron survival and differentiation: from development to ageing. Cerebellum 2006;5:97-104.

24. Cobos I, Borello U, Rubenstein JL. Dlx transcription factors promote migration through repression of axon and dendrite growth. Neuron 2007;54:873-888.

25. Flames N, Hobert O. Gene regulatory logic of dopamine neuron differentiation. Nature 2009;458:885-889.

26. Cornell RA, Eisen JS. Notch in the pathway: the roles of Notch signaling in neural crest development. Semin Cell Dev Biol 2005;16:663-672.

27. Okamura Y, Saga Y. Notch signaling is required for the maintenance of enteric neural crest progenitors. Development 2008;135:3555-3565.

28. Kleber M, Sommer L. Wnt signaling and the regulation of stem cell function. Curr Opin Cell Biol 2004;16:681-687.

29. Leong SY, Faux CH, Turbic A, et al. The rho kinase pathway regulates mouse adult neural precursor cell migration. Stem Cells 2011;29:332-343.

30. Lo L, Sommer L, Anderson DJ. MASH1 maintains competence for BMP2-induced neuronal differentiation in post-migratory neural crest cells. Curr Biol 1997;7:440-450.

31. Pisano JM, Colon-Hastings F, Birren SJ. Postmigratory enteric and sympathetic neural precursors share common, developmentally regulated, responses to BMP2. Dev Biol 2000;227:1-11.

32. Shah NM, Groves AK, Anderson DJ. Alternative neural crest cell fates are instructively promoted by TGFbeta superfamily members. Cell 1996;85:331-343.

33. Sanchez-Martin M, Rodriguez-Garcia A, Perez-Losada J, et al. SLUG (SNAI2) deletions in patients with Waardenburg disease. Hum Mol Genet 2002;11:3231-3236.

34. Plaza Menacho I, Koster R, van der Sloot AM, et al. RET-familial medullary thyroid carcinoma mutants Y791F and S891A activate a Src/JAK/STAT3 pathway, independent of glial cell line-derived neurotrophic factor. Cancer Res 2005;65:1729-1737.

35. Barajon I, Serrao G, Arnaboldi F, et al. Toll-like receptors 3, 4, and 7 are expressed in the enteric nervous system and dorsal root ganglia. J Histochem Cytochem 2009;57:1013-1023.

36. Rumio C, Besusso D, Arnaboldi F, et al. Activation of smooth muscle and myenteric plexus cells of jejunum via Toll-like receptor 4. J Cell Physiol 2006;208:47-54.

37. Carrasquillo MM, McCallion AS, Puffenberger EG, et al. Genome-wide association study and mouse model identify interaction between RET and EDNRB pathways in Hirschsprung disease. Nat Genet 2002;32:237-244.

38. McCallion AS, Stames E, Conlon RA, et al. Phenotype variation in two-locus mouse models of Hirschsprung disease: tissue-specific interaction between Ret and Ednrb. Proc Natl Acad Sci U S A 2003;100:1826-1831.

39. Barlow A, de Graaff E, Pachnis V. Enteric nervous system progenitors are coordinately controlled by the G protein-coupled receptor EDNRB and the receptor tyrosine kinase RET. Neuron 2003;40:905-916.

40. Lang D, Epstein JA. Sox10 and Pax3 physically interact to mediate activation of a conserved c-RET enhancer. Hum Mol Genet 2003;12:937-945.

41. Zhu L, Lee HO, Jordan CS, et al. Spatiotemporal regulation of endothelin receptor-B by SOX10 in neural crest-derived enteric neuron precursors. Nat Genet 2004;36:732-737.

42. Lang D, Chen F, Milewski R, et al. Pax3 is required for enteric ganglia formation and functions with Sox10 to modulate expression of c-ret. J Clin Invest 2000;106:963-971.

43. Pattyn A, Morin X, Cremer H, et al. The homeobox gene Phox2b is essential for the development of autonomic neural crest derivatives. Nature 1999;399:366-370.

44. Leon TY, Ngan ES, Poon HC, et al. Transcriptional regulation of RET by Nkx2-1, Phox2b, Sox10, and Pax3. J Pediatr Surg 2009;44:1904-1912.

45. Jing S, Yu Y, Fang M, et al. GFRalpha-2 and GFRalpha-3 are two new receptors for ligands of the GDNF family. J Biol Chem

1997;272:33111-33117.

46. Heanue TA, Pachnis V. Prospective Identification and Isolation of Enteric Nervous System Progenitors Using SOX2. Stem Cells 2010.

47. Kaye J, Gage H, Kimber A, et al. Excess burden of constipation in Parkinson's disease: a pilot study. Mov Disord 2006;21:1270-1273.

48. Lebouvier T, Neunlist M, Bruley des Varannes S, et al. Colonic biopsies to assess the neuropathology of Parkinson's disease and its relationship with symptoms. PLoS One 2010;5:e12728.

49. Heanue TA, Pachnis V. Expression profiling the developing mammalian enteric nervous system identifies marker and candidate Hirschsprung disease genes. Proc Natl Acad Sci U S A 2006;103:6919-6924.

50. Vohra BP, Tsuji K, Nagashimada M, et al. Differential gene expression and functional analysis implicate novel mechanisms in enteric nervous system precursor migration and neuritogenesis. Dev Biol 2006;298:259-271.

51. Iwashita T, Kruger GM, Pardal R, et al. Hirschsprung disease is linked to defects in neural crest stem cell function. Science 2003;301:972-976.

52. Ngan ES, Lee KY, Sit FY, et al. Prokineticin-1 modulates proliferation and differentiation of enteric neural crest cells. Biochim Biophys Acta 2007;1773:536-545.

53. Schoenherr CJ, Anderson DJ. The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. Science 1995;267:1360-1363.

54. Chong JA, Tapia-Ramirez J, Kim S, et al. REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons. Cell 1995;80:949-957.

55. Salaun B, Romero P, Lebecque S. Toll-like receptors' two-edged sword: when immunity meets apoptosis. Eur J Immunol 2007;37:3311-3318.

56. Bomsel M, Alfsen A. Entry of viruses through the epithelial barrier: pathogenic trickery. Nat Rev Mol Cell Biol 2003;4:57-68.

57. Niedergang F, Kweon MN. New trends in antigen uptake in the gut mucosa. Trends Microbiol 2005;13:485-490.

58. Mandat Schultz A, Bonnard A, Barreau F, et al. Expression of TLR-2, TLR-4, NOD2 and pNF-kappaB in a neonatal rat model of necrotizing enterocolitis. PLoS One 2007;2:e1102.

59. Leaphart CL, Cavallo J, Gribar SC, et al. A critical role for TLR4 in the pathogenesis of necrotizing enterocolitis by modulating intestinal injury and repair. J Immunol 2007;179:4808-4820.

60. Schmidt K, Glaser G, Wernig A, Wegner M, Rosorius O. Sox8 is a specific marker for muscle satellite cells and inhibits myogenesis. J Biol Chem 2003;278:29769-29775.

# SUPPLEMENTARY DATA

**Supplementary Table 1 -** Primers used for qPCR amplification and product length

| Gene | Primers | Product length |
|---|---|---|
| TLR2 | F: 5' CTCAGAGGATGCTACGAGCTC 3'<br>R: 5' GAATAGAGGTGAAAGACCTGGAGC 3' | 147 |
| TLR4 | F: 5' GAATCCCTGCATAGAGGTAGTTCC 3'<br>R: 5' TGATCCATGCATTGGTAGGTAATATTA 3' | 51 |
| TLR5 | F: 5' TCAGCAGGATCATGGCATGTC 3'<br>R: 5' TGAAGATCACACCTATGAGCAAGTCA 3' | 51 |
| TNFa | F: 5' GACGTGGAACTGGCAGAAGA 3'<br>R: 5' GCCACAAGCAGGAATGAGAA 3' | 101 |
| Il-6 | F: 5' CCTCTCTGCAAGAGACTTCCATCCA 3'<br>R: 5' GGCCGTGGTTGTCACCAGCA 3' | 64 |
| Il-1b | F: 5' GGCAGGCAGTATCACTCATT 3'<br>R: 5' AAGGTGCTCATGTCCTCATC 3' | 94 |
| P50 | F: 5' CATGGTGGTTGGCTTTGCA 3'<br>R: 5' ACACGCCTCTGTCATCCGTG 3' | 91 |
| IFNg | F: 5' ACTGGCAAAAGGATGGTGAC 3'<br>R: 5' GCTGATGGCCTGATTGTCTT 3' | 98 |
| Gapdh | F: 5' CATCAAGAAGGTGGTGAAGC 3'<br>R: 5' ACCACCCTGTTGCTGTAG 3' | 209 |

**Supplementary Table 2 -** Antibodies used in immunohistochemistry

| Gene | Antibodies | Host | Dilution | Supplier |
|---|---|---|---|---|
| **Primary** | Ret | Goat | 1:500 | R&D system, Minneapolis, USA |
| | Sox10 | Mouse | 1:2 | Kindly provided by Prof.M.Wegner60 |
| | Gfap | Mouse | 1:100 | Dako, Hamburg, Germany |
| | Tlr2 | Rabbit | 1:100 | Sigma Aldrich, Saint Louis, USA |
| | Tlr4 | Rabbit | 1:100 | Sigma Aldrich, Saint Louis, USA |
| | Tlr5 | Rabbit | 1:100 | Sigma Aldrich, Saint Louis, USA |
| **Secondary** | Anti-goat Alexa 488 | Donkey | 1:500 | Invitrogen, Karlsruhe, Germany |
| | Anti-mouse Alexa 488 | Donkey | 1:500 | Invitrogen, Karlsruhe, Germany |
| | Anti-mouse Alexa 488 | Donkey | 1:500 | Jackson Immunoresearch, Hamburg, Germany |

**Supplementary Table 3 -** Single Gene Analysis: Genes were expressed significantly different in ENCCs treated with GDNF compare to untreated ENCCs (Bonferroni : $p < 0.05$)

| Rank | Affymetrix Name | Gene Symbol | T | Bonferronni |
|---|---|---|---|---|
| **Up-regulated** | | | | |
| 1 | 1453008_at | Trnp1 | -49.711785 | 1.34E-06 |
| 2 | 1435026_at | Spock2 | -43.762722 | 3.70E-06 |
| 3 | 1452825_at | Tmem59l | -42.657373 | 4.54E-06 |
| 4 | 1435857_s_at | Aplp1 | -40.88871 | 6.35E-06 |
| 5 | 1437261_at | Mblac2 | -35.864703 | 1.80E-05 |
| 6 | 1436094_at | Vgf | -33.761245 | 2.92E-05 |
| 7 | 1456379_x_at | Dner | -33.271202 | 3.28E-05 |
| 8 | 1437292_at | Rlbp1l2 | -31.037683 | 5.69E-05 |
| 9 | 1428333_at | 2900062L11Rik | -29.839986 | 7.78E-05 |
| 10 | 1426604_at | Rnasel | -29.703856 | 8.07E-05 |
| 11 | 1452653_at | Slc25a22 | -29.609467 | 8.27E-05 |
| 12 | 1437528_x_at | A730017C20Rik | -29.389713 | 8.78E-05 |
| 13 | 1434427_a_at | Rnf157 | -27.830259 | 0.000135269 |
| 14 | 1417988_at | Resp18 | -26.224421 | 0.000216607 |
| 15 | 1418736_at | B3galnt1 | -26.088397 | 0.000225711 |
| 16 | 1459723_at | Zdhhc22 | -25.644417 | 0.000258556 |
| 17 | 1441662_at | Cyp4x1 | -25.349314 | 0.000283353 |
| 18 | 1449628_s_at | Stard7 | -25.219314 | 0.000295116 |
| 19 | 1443924_at | Wnk3 | -25.062297 | 0.000310062 |
| 20 | 1451679_at | 6530401D17Rik | -24.824377 | 0.000334356 |
| 21 | 1427893_a_at | Pmvk | -24.621329 | 0.00035679 |
| 22 | 1435756_at | Samd10 | -24.571621 | 0.000362536 |
| 23 | 1421349_x_at | Cend1 | -24.564025 | 0.000363423 |
| 24 | 1426662_at | Cmas | -24.553795 | 0.000364622 |
| 25 | 1417428_at | Gng3 | -24.541655 | 0.000366051 |
| 26 | 1455260_at | Lcorl | -24.130619 | 0.000418331 |
| 27 | 1423280_at | Stmn2 | -24.044308 | 0.000430345 |
| 28 | 1418671_at | Capn5 | -23.886055 | 0.000453394 |
| 29 | 1418400_at | Larp6 | -23.766463 | 0.000471734 |
| 30 | 1460262_a_at | 1700037H04Rik | -23.762846 | 0.000472301 |
| 31 | 1455888_at | Lingo2 | -23.645762 | 0.000491093 |
| 32 | 1427482_a_at | Car8 | -23.417563 | 0.000530181 |
| 33 | 1455792_x_at | Ndn | -23.39489 | 0.000534252 |
| 34 | 1422711_a_at | Pnck | -23.389166 | 0.000535285 |
| 35 | 1450863_a_at | Dclk1 | -23.246707 | 0.000561739 |
| 36 | 1434754_at | Garnl4 | -23.232605 | 0.000564437 |
| 37 | 1454687_at | Lrfn5 | -23.111085 | 0.000588295 |
| 38 | 1436134_at | Scn2b | -23.073987 | 0.000595802 |
| 39 | 1440132_s_at | Prkar1b | -22.863039 | 0.000640594 |
| 40 | 1424624_at | 2900011O08Rik | -22.81546 | 0.000651211 |
| 41 | 1460239_at | Tspan13 | -22.703032 | 0.000677096 |
| 42 | 1438370_x_at | Dos | -22.614644 | 0.000698258 |
| 43 | 1435105_at | Rnf208 | -22.604722 | 0.000700679 |

| | | | | |
|---|---|---|---|---|
| 44 | 1433876_at | Lrrc24 | -22.526199 | 0.00072018 |
| 45 | 1423352_at | Crispld1 | -22.47565 | 0.000733055 |
| 46 | 1436087_at | LOC100047231 | -22.311623 | 0.000776651 |
| 47 | 1422605_at | Ppp1r1a | -22.177671 | 0.000814419 |
| 48 | 1460292_a_at | Smarca1 | -22.159762 | 0.000819623 |
| 49 | 1424443_at | Tm6sf1 | -22.156432 | 0.000820595 |
| 50 | 1448312_at | Pcsk2 | -21.990273 | 0.000870767 |
| 51 | 1455652_at | Kif3a | -21.95774 | 0.000880989 |
| 52 | 1433443_a_at | Hmgcs1 | -21.884633 | 0.000904454 |
| 53 | 1416965_at | Pcsk1n | -21.881448 | 0.000905492 |
| 54 | 1455080_at | Ppp1r16b | -21.802173 | 0.000931768 |
| 55 | 1450773_at | Kcnd2 | -21.669177 | 0.000977796 |
| 56 | 1436450_at | D11Bwg0517e | -21.570151 | 0.001013727 |
| 57 | 1452240_at | Brunol4 | -21.456076 | 0.001056969 |
| 58 | 1416935_at | Trpv2 | -21.285768 | 0.001125436 |
| 59 | 1447725_at | C030034E14Rik | -21.204243 | 0.001159963 |
| 60 | 1447738_s_at | Ankrd13d | -21.110093 | 0.001201326 |
| 61 | 1434961_at | Asb1 | -21.097184 | 0.001207125 |
| 62 | 1450659_at | Rgs7 | -21.007074 | 0.001248493 |
| 63 | 1424482_at | Arhgef7 | -20.930141 | 0.001285075 |
| 64 | 1437079_at | Slc18a2 | -20.809908 | 0.001344682 |
| 65 | 1436179_a_at | Dnajc5 | -20.385681 | 0.001581199 |
| 66 | 1451092_a_at | Rangap1 | -20.242769 | 0.001671131 |
| 67 | 1434657_at | Gls | -20.023423 | 0.001820554 |
| 68 | 1420499_at | Gch1 | -19.889409 | 0.001919206 |
| 69 | 1423282_at | Pitpna | -19.852766 | 0.001947215 |
| 70 | 1459881_at | Fbll1 | -19.848639 | 0.001950398 |
| 71 | 1428370_at | 1500011B03Rik | -19.796438 | 0.001991169 |
| 72 | 1418829_a_at | Eno2 | -19.701832 | 0.002067524 |
| 73 | 1454972_at | Atcay | -19.57182 | 0.002177875 |
| 74 | 1441197_at | 9530059O14Rik | -19.554945 | 0.002192676 |
| 75 | 1442180_at | Dleu7 | -19.524284 | 0.00221986 |
| 76 | 1448832_a_at | Cplx1 | -19.437734 | 0.002298659 |
| 77 | 1457755_at | Gng8 | -19.31441 | 0.002416422 |
| 78 | 1424017_a_at | Hint1 | -19.163534 | 0.002569807 |
| 79 | 1427229_at | Hmgcr | -19.154594 | 0.002579232 |
| 80 | 1432464_a_at | 2310057J16Rik | -19.053404 | 0.002688674 |
| 81 | 1441905_x_at | Snrpn /// Snurf | -19.049584 | 0.002692906 |
| 82 | 1418209_a_at | Pfn2 | -18.993873 | 0.00275549 |
| 83 | 1426615_s_at | Ndrg4 | -18.990207 | 0.002759665 |
| 84 | 1451008_at | St8sia3 | -18.922381 | 0.002838206 |
| 85 | 1423317_at | 3110001D03Rik | -18.83419 | 0.002944109 |
| 86 | 1428765_at | Meg3 | -18.818445 | 0.002963479 |
| 87 | 1451322_at | Cmbl | -18.758269 | 0.003038838 |
| 88 | 1457092_at | Fam19a1 | -18.679781 | 0.003140383 |
| 89 | 1448595_a_at | Bex1 | -18.617769 | 0.003223309 |
| 90 | 1460223_a_at | Epb4.9 | -18.575406 | 0.003281371 |
| 91 | 1421152_a_at | Gnao1 | -18.494733 | 0.003395217 |
| 92 | 1421493_a_at | Rgs20 | -18.48937 | 0.003402941 |
| 93 | 1433791_at | Rab9b | -18.465285 | 0.003437872 |
| 94 | 1455014_at | Hint3 | -18.252221 | 0.003765047 |

| | | | | |
|---|---|---|---|---|
| 95 | 1453078_at | 2610002M06Rik | -18.167393 | 0.003904918 |
| 96 | 1456392_at | Negr1 | -18.164024 | 0.003910593 |
| 97 | 1429022_at | Adcyap1r1 | -18.118995 | 0.003987329 |
| 98 | 1441944_s_at | Gpr135 | -18.100384 | 0.004019536 |
| 99 | 1448660_at | Arhgdig | -18.070941 | 0.004071088 |
| 100 | 1436552_at | Jakmip2 | -18.063323 | 0.004084547 |
| 101 | 1418062_at | Eef1a2 | -18.011533 | 0.004177382 |
| 102 | 1435641_at | Mgat4a | -17.985843 | 0.004224308 |
| 103 | 1451499_at | Cadps2 | -17.976442 | 0.004241627 |
| 104 | 1424308_at | Slc24a3 | -17.974323 | 0.004245543 |
| 105 | 1428089_at | Slitrk1 | -17.954827 | 0.004281753 |
| 106 | 1451394_at | Dpp6 | -17.943084 | 0.004303732 |
| 107 | 1436605_at | Tkt | -17.929896 | 0.004328565 |
| 108 | 1437234_x_at | Prmt2 | -17.917984 | 0.004351133 |
| 109 | 1437618_x_at | Gpr85 | -17.803552 | 0.004574799 |
| 110 | 1431403_a_at | Mtap7d2 | -17.72493 | 0.004735946 |
| 111 | 1434155_a_at | 2310061I04Rik | -17.693038 | 0.00480312 |
| 112 | 1415671_at | Atp6v0d1 | -17.665455 | 0.004862083 |
| 113 | 1450888_at | Napb | -17.632404 | 0.004933805 |
| 114 | 1420554_a_at | Rac3 | -17.560163 | 0.005094746 |
| 115 | 1431086_s_at | Pcmt1 | -17.539057 | 0.005142872 |
| 116 | 1428354_at | Foxk2 | -17.538998 | 0.005143007 |
| 117 | 1435630_s_at | Acat2 | -17.510206 | 0.005209486 |
| 118 | 1455313_at | Ablim2 | -17.498944 | 0.00523575 |
| 119 | 1434346_at | Rufy2 | -17.491789 | 0.005252513 |
| 120 | 1450693_at | Rgs17 | -17.436449 | 0.005384226 |
| 121 | 1417997_at | Ngb | -17.427255 | 0.005406465 |
| 122 | 1423724_at | Zwint | -17.418391 | 0.005428004 |
| 123 | 1438672_at | Parvb | -17.404948 | 0.005460853 |
| 124 | 1435014_at | Rab39b | -17.374322 | 0.005536528 |
| 125 | 1438619_x_at | Zdhhc14 | -17.338304 | 0.005627033 |
| 126 | 1422941_at | Wnt16 | -17.293055 | 0.005743098 |
| 127 | 1416392_a_at | Atp6v0c | -17.225169 | 0.005922297 |
| 128 | 1440270_at | Fgf12 | -17.217381 | 0.005943254 |
| 129 | 1442884_at | Hgf | -17.201465 | 0.00598634 |
| 130 | 1425893_a_at | Fhit | -17.112659 | 0.006233277 |
| 131 | 1449983_a_at | Nqo2 | -17.103442 | 0.006259555 |
| 132 | 1449377_at | Nicn1 | -17.072009 | 0.006350107 |
| 133 | 1436237_at | Ttc9 | -17.059961 | 0.006385204 |
| 134 | 1451503_at | Nol3 | -17.025724 | 0.006486134 |
| 135 | 1455258_at | Kcnc2 | -16.969797 | 0.006654878 |
| 136 | 1428550_at | Ydjc | -16.941611 | 0.006741787 |
| 137 | 1433987_at | Hpcal4 | -16.937396 | 0.006754889 |
| 138 | 1456108_x_at | Rnf112 | -16.917326 | 0.006817684 |
| 139 | 1449491_at | Card10 | -16.881509 | 0.006931384 |
| 140 | 1452332_at | Ccdc85a | -16.846635 | 0.007044136 |
| 141 | 1443327_at | D130043K22Rik | -16.840347 | 0.007064684 |
| 142 | 1434170_at | Decaf12l1 | -16.833805 | 0.007086133 |
| 143 | 1438054_x_at | Ppp2r2c | -16.825827 | 0.007112391 |
| 144 | 1426249_at | Adrbk1 | -16.819502 | 0.007133283 |
| 145 | 1438667_at | 5730410E15Rik | -16.816992 | 0.007141596 |

| | | | | |
|---|---|---|---|---|
| 146 | 1452711_at | Atl1 | -16.798857 | 0.007201962 |
| 147 | 1417609_at | Ube2a | -16.792496 | 0.007223273 |
| 148 | 1418178_at | LOC100047943 | -16.758641 | 0.00733789 |
| 149 | 1455796_x_at | Olfm1 | -16.736336 | 0.007414517 |
| 150 | 1455509_at | Nacad | -16.717595 | 0.007479597 |
| 151 | 1460570_at | Pgbd5 | -16.691408 | 0.007571608 |
| 152 | 1426956_a_at | Trp53bp1 | -16.678548 | 0.007617257 |
| 153 | 1422852_at | Cib2 | -16.675134 | 0.007629427 |
| 154 | 1452800_a_at | Apoo | -16.673311 | 0.007635937 |
| 155 | 1442786_s_at | Rufy3 | -16.627761 | 0.007800581 |
| 156 | 1422072_a_at | Gstm6 | -16.609159 | 0.007868961 |
| 157 | 1422977_at | Gp1bb | -16.586468 | 0.007953284 |
| 158 | 1435113_x_at | Stmn3 | -16.580414 | 0.007975952 |
| 159 | 1420880_a_at | Ywhab | -16.576786 | 0.007989571 |
| 160 | 1415993_at | Sqle | -16.490009 | 0.008323233 |
| 161 | 1451895_a_at | Dhcr24 | -16.453165 | 0.008469606 |
| 162 | 1435605_at | Actr3b | -16.433418 | 0.008549245 |
| 163 | 1416098_at | Syngr3 | -16.413526 | 0.008630315 |
| 164 | 1426712_at | Slc6a15 | -16.379991 | 0.008768951 |
| 165 | 1428996_at | Pdp2 | -16.323407 | 0.00900857 |
| 166 | 1435917_at | Ociad2 | -16.31635 | 0.009038965 |
| 167 | 1428578_s_at | Ppfia4 | -16.305943 | 0.009084001 |
| 168 | 1433787_at | Nell1 | -16.276958 | 0.009210769 |
| 169 | 1459890_s_at | 1110008P14Rik | -16.227167 | 0.009433186 |
| 170 | 1455321_at | Ddhd1 | -16.208564 | 0.009517823 |
| 171 | 1420545_a_at | Chn1 | -16.197733 | 0.009567496 |
| 172 | 1437018_at | Pnma2 | -16.190394 | 0.009601318 |
| 173 | 1459980_x_at | Rab3a | -16.187444 | 0.009614952 |
| 174 | 1417672_at | Slc4a10 | -16.177399 | 0.009661535 |
| 175 | 1450675_at | Smap2 | -16.166271 | 0.009713438 |
| 176 | 1452915_at | Prkar2a | -16.141985 | 0.009827805 |
| 177 | 1435200_at | Ncrna00086 | -16.131233 | 0.009878915 |
| 178 | 1454787_at | Zdhhc9 | -16.095105 | 0.010052862 |
| 179 | 1426255_at | Nefl | -16.074527 | 0.010153473 |
| 180 | 1423418_at | Fdps | -16.052983 | 0.010260016 |
| 181 | 1418085_at | Prkcz | -16.000829 | 0.010523169 |
| 182 | 1448991_a_at | Ina /// LOC100047943 | -15.915841 | 0.010968373 |
| 183 | 1419580_at | Dlg4 | -15.895684 | 0.011077038 |
| 184 | 1425760_a_at | Pitpnm1 | -15.797923 | 0.011621484 |
| 185 | 1417788_at | Sncg | -15.797841 | 0.011621952 |
| 186 | 1435083_at | Ctxn1 | -15.788938 | 0.011673013 |
| 187 | 1437211_x_at | Elovl5 | -15.773696 | 0.011761007 |
| 188 | 1437147_at | Gabrg2 | -15.765563 | 0.011808265 |
| 189 | 1426345_at | Prepl | -15.718717 | 0.012084646 |
| 190 | 1460740_at | LOC100046457 | -15.699916 | 0.012197603 |
| 191 | 1451747_a_at | Atg12 | -15.694463 | 0.012230591 |
| 192 | 1417471_s_at | D1Ertd622e | -15.692671 | 0.012241453 |
| 193 | 1417799_at | Atp6v1g2 | -15.682707 | 0.012302034 |
| 194 | 1428664_at | Vip | -15.662437 | 0.012426332 |
| 195 | 1431076_at | Add2 | -15.647016 | 0.012521837 |
| 196 | 1448663_s_at | Mvd | -15.623501 | 0.012669052 |

| | | | | |
|---|---|---|---|---|
| 197 | 1427045_at | Synpo | -15.611553 | 0.012744603 |
| 198 | 1433714_at | Sult4a1 | -15.595042 | 0.012849827 |
| 199 | 1434027_at | Rcan3 | -15.578334 | 0.012957308 |
| 200 | 1455269_a_at | Coro1a | -15.536606 | 0.013230161 |
| 201 | 1433566_at | Rasl10b | -15.530133 | 0.01327306 |
| 202 | 1418361_at | Gas8 | -15.521722 | 0.013329038 |
| 203 | 1450961_a_at | Tceal6 | -15.513995 | 0.013380698 |
| 204 | 1441317_x_at | Jakmip1 | -15.508048 | 0.013420608 |
| 205 | 1453127_at | Ppm1j | -15.499416 | 0.013478782 |
| 206 | 1433947_at | Rab37 | -15.485122 | 0.013575728 |
| 207 | 1457587_at | Kcnq5 | -15.465515 | 0.013709985 |
| 208 | 1454782_at | Bai3 | -15.45587 | 0.013776571 |
| 209 | 1418474_at | Fam158a | -15.449029 | 0.013824022 |
| 210 | 1423328_at | Gdap1 | -15.443272 | 0.013864094 |
| 211 | 1434969_at | Brunol5 | -15.397101 | 0.014190232 |
| 212 | 1422431_at | Magee1 | -15.392193 | 0.014225406 |
| 213 | 1430642_at | 2900001G08Rik | -15.387067 | 0.01426224 |
| 214 | 1417840_at | 1500031L02Rik | -15.375277 | 0.014347378 |
| 215 | 1434951_at | Armc8 | -15.360494 | 0.014454934 |
| 216 | 1436148_at | Tnr | -15.34392 | 0.014576593 |
| 217 | 1426236_a_at | Glul | -15.343505 | 0.01457965 |
| 218 | 1437185_s_at | Tmsb10 | -15.327694 | 0.01469681 |
| 219 | 1424182_at | Acat1 | -15.307308 | 0.01484943 |
| 220 | 1420013_s_at | Lss | -15.287392 | 0.015000248 |
| 221 | 1416023_at | Fabp3 | -15.281706 | 0.015043629 |
| 222 | 1420955_at | Vsnl1 | -15.278605 | 0.015067338 |
| 223 | 1417524_at | LOC100044812 | -15.272629 | 0.015113157 |
| 224 | 1427356_at | Fam89a | -15.226715 | 0.015470435 |
| 225 | 1448211_at | Atp6v0e2 | -15.22373 | 0.01549399 |
| 226 | 1424069_at | Napg | -15.222628 | 0.015502694 |
| 227 | 1460576_at | Exoc6 | -15.20585 | 0.01563592 |
| 228 | 1417676_a_at | Ptpro | -15.205287 | 0.015640412 |
| 229 | 1450466_at | Cdk5r2 | -15.190918 | 0.015755567 |
| 230 | 1450853_at | Tle4 | -15.179773 | 0.015845535 |
| 231 | 1419551_s_at | Stk39 | -15.158714 | 0.016017118 |
| 232 | 1434455_at | Fbxo44 | -15.110964 | 0.016413938 |
| 233 | 1442023_at | A530030E21Rik | -15.076419 | 0.016707893 |
| 234 | 1437855_at | Mtap4 | -15.067419 | 0.016785435 |
| 235 | 1454959_s_at | Gnai1 | -15.065485 | 0.016802159 |
| 236 | 1424529_s_at | Cgref1 | -15.056176 | 0.016882887 |
| 237 | 1458413_at | Fbxw8 | -15.055287 | 0.016890616 |
| 238 | 1433737_at | LOC677213 /// Uhmk1 | -15.042434 | 0.017002847 |
| 239 | 1436795_at | 9630058J23Rik | -15.026237 | 0.017145479 |
| 240 | 1449259_at | Rab3d | -15.011606 | 0.017275464 |
| 241 | 1431255_at | Calr3 | -14.991252 | 0.017458146 |
| 242 | 1439770_at | 6430598A04Rik | -14.978284 | 0.017575656 |
| 243 | 1437912_at | Iqsec3 | -14.968554 | 0.017664417 |
| 244 | 1417002_at | 0610012G03Rik | -14.955263 | 0.017786472 |
| 245 | 1422692_at | Sub1 | -14.924587 | 0.018071797 |
| 246 | 1449298_a_at | Pde1a | -14.916057 | 0.018152048 |
| 247 | 1427280_at | Scn2a1 | -14.8936 | 0.018365234 |

| | | | | |
|---|---|---|---|---|
| 248 | 1418453_a_at | Atp1b1 | -14.886328 | 0.018434873 |
| 249 | 1425248_a_at | Tyro3 | -14.85919 | 0.018697362 |
| 250 | 1423906_at | Hsbp1 | -14.845171 | 0.018834599 |
| 251 | 1426094_at | Rhbdl1 | -14.823518 | 0.019048801 |
| 252 | 1439031_at | Jph4 | -14.760349 | 0.019689448 |
| 253 | 1453003_at | Sorl1 | -14.756716 | 0.01972702 |
| 254 | 1420872_at | Gucy1b3 | -14.745665 | 0.019841815 |
| 255 | 1451257_at | Acsl6 | -14.731314 | 0.019991989 |
| 256 | 1424455_at | Gprasp1 | -14.730476 | 0.020000802 |
| 257 | 1417022_at | Slc7a3 | -14.707008 | 0.020249271 |
| 258 | 1451935_a_at | Spint2 | -14.699556 | 0.020328891 |
| 259 | 1431826_a_at | Brsk2 | -14.690931 | 0.02042148 |
| 260 | 1452779_at | 3110006E14Rik | -14.679031 | 0.020550015 |
| 261 | 1434270_at | Cbx6-Nptxr /// Nptxr | -14.66256 | 0.020729408 |
| 262 | 1427004_at | Fbxo2 | -14.658023 | 0.020779128 |
| 263 | 1449229_a_at | Cdkl2 | -14.642309 | 0.020952374 |
| 264 | 1428638_at | Efhc2 | -14.636313 | 0.021018912 |
| 265 | 1440531_at | Rbm11 | -14.624637 | 0.021149153 |
| 266 | 1449494_at | Rab3c | -14.586689 | 0.021578746 |
| 267 | 1435028_at | Wdr7 | -14.580738 | 0.021646998 |
| 268 | 1428387_at | Acsl3 | -14.57717 | 0.021688032 |
| 269 | 1456210_at | 5430407P10Rik | -14.573942 | 0.021725238 |
| 270 | 1434641_x_at | Sez6l2 | -14.556641 | 0.02192584 |
| 271 | 1424386_at | Reep2 | -14.514886 | 0.022418612 |
| 272 | 1424145_at | Prr3 | -14.511926 | 0.022454011 |
| 273 | 1423569_at | Gatm | -14.501683 | 0.022577001 |
| 274 | 1417683_at | Diablo | -14.494192 | 0.022667423 |
| 275 | 1456708_at | 2610028E06Rik | -14.490687 | 0.022709871 |
| 276 | 1451040_at | Dtd1 | -14.481389 | 0.022822909 |
| 277 | 1424764_at | Sez6l | -14.472233 | 0.022934832 |
| 278 | 1434640_at | Nrsn2 | -14.468432 | 0.022981479 |
| 279 | 1426272_at | Lmbr1 | -14.447456 | 0.023240822 |
| 280 | 1460397_at | Armc6 | -14.435543 | 0.023389564 |
| 281 | 1418046_at | Nap1l2 | -14.434213 | 0.023406234 |
| 282 | 1420824_at | Sema4d | -14.428484 | 0.023478205 |
| 283 | 1433607_at | Cbln4 | -14.392068 | 0.023941506 |
| 284 | 1455829_at | Usp14 | -14.382448 | 0.024065607 |
| 285 | 1436099_at | AI836003 | -14.377229 | 0.024133227 |
| 286 | 1424124_at | Mospd2 | -14.372671 | 0.024192468 |
| 287 | 1423642_at | Tubb2c | -14.364696 | 0.024296502 |
| 288 | 1455481_at | Ids | -14.341279 | 0.024604888 |
| 289 | 1450647_at | Hps3 | -14.302782 | 0.025121437 |
| 290 | 1426167_a_at | Camk4 | -14.280286 | 0.025428918 |
| 291 | 1417374_at | Tuba4a | -14.245118 | 0.025918074 |
| 292 | 1439971_at | 6330439K17Rik | -14.23563 | 0.026051842 |
| 293 | 1436403_at | Fam171a2 | -14.231944 | 0.026104025 |
| 294 | 1419062_at | Epb4.1l3 | -14.186862 | 0.026751752 |
| 295 | 1455369_at | Apba1 | -14.181556 | 0.026829163 |
| 296 | 1423055_at | Nsg1 | -14.173006 | 0.026954431 |
| 297 | 1436624_at | Dnm3 | -14.150986 | 0.027280081 |
| 298 | 1435123_at | Efr3b | -14.145269 | 0.027365344 |

| | | | | |
|---|---|---|---|---|
| 299 | 1455636_at | Lsamp | -14.144254 | 0.027380527 |
| 300 | 1428405_at | Hcfc1r1 | -14.139908 | 0.02744558 |
| 301 | 1428209_at | Bex4 | -14.121669 | 0.027720511 |
| 302 | 1424935_at | Gdap1l1 | -14.107253 | 0.027940012 |
| 303 | 1418263_at | Ddx25 | -14.105193 | 0.027971533 |
| 304 | 1448530_at | Gmpr | -14.103178 | 0.028002399 |
| 305 | 1451783_a_at | Kifap3 | -14.09768 | 0.02808683 |
| 306 | 1437744_at | Slitrk4 | -14.092544 | 0.02816596 |
| 307 | 1423285_at | Coch | -14.090176 | 0.028202538 |
| 308 | 1418493_a_at | Snca | -14.075296 | 0.028433546 |
| 309 | 1448193_at | 5730403B10Rik | -14.031732 | 0.02912215 |
| 310 | 1417626_at | Pde4dip | -14.027722 | 0.029186469 |
| 311 | 1435389_at | Reps2 | -14.006277 | 0.029533148 |
| 312 | 1416102_at | Ywhaz | -13.949724 | 0.030469732 |
| 313 | 1455021_at | Gabbr1 | -13.938683 | 0.030656444 |
| 314 | 1426517_at | Gnaz | -13.932864 | 0.03075536 |
| 315 | 1436674_at | Rundc3a | -13.929905 | 0.030805795 |
| 316 | 1457045_at | Galnt13 | -13.929549 | 0.030811866 |
| 317 | 1450053_at | Kif2a | -13.90716 | 0.031196596 |
| 318 | 1419493_a_at | Tpd52 | -13.900943 | 0.031304376 |
| 319 | 1417559_at | Sfxn1 | -13.900464 | 0.031312704 |
| 320 | 1449286_at | Ntng1 | -13.870132 | 0.031844757 |
| 321 | 1416936_at | Aatk | -13.864694 | 0.031941213 |
| 322 | 1422533_at | Cyp51 | -13.857524 | 0.032068874 |
| 323 | 1422592_at | Ctnnd2 | -13.847363 | 0.032250802 |
| 324 | 1437075_at | Frmd3 | -13.791405 | 0.033273573 |
| 325 | 1422583_at | Rab3b | -13.786418 | 0.033366474 |
| 326 | 1449423_at | Mast1 | -13.786152 | 0.033371438 |
| 327 | 1437472_at | Unc13a | -13.761468 | 0.033835643 |
| 328 | 1452451_at | mCG_21548 | -13.755737 | 0.033944448 |
| 329 | 1424988_at | Mylip | -13.751982 | 0.034015968 |
| 330 | 1434676_at | Mtmr9 | -13.747122 | 0.034108754 |
| 331 | 1418164_at | Stx2 | -13.72877 | 0.034461743 |
| 332 | 1416253_at | Cdkn2d /// Gm4694 | -13.72401 | 0.034553967 |
| 333 | 1417154_at | Slc25a14 | -13.721546 | 0.034601811 |
| 334 | 1428434_at | Zcchc12 | -13.719647 | 0.034638726 |
| 335 | 1420505_a_at | Stxbp1 | -13.716675 | 0.034696601 |
| 336 | 1426068_at | Slc7a4 | -13.699416 | 0.035034879 |
| 337 | 1438164_x_at | Flot2 | -13.691769 | 0.035185919 |
| 338 | 1429723_at | 6330409N04Rik | -13.648523 | 0.036054028 |
| 339 | 1429013_at | Mtap7d2 | -13.630973 | 0.036413156 |
| 340 | 1424691_at | 5930434B04Rik | -13.62781 | 0.036478301 |
| 341 | 1444085_at | Pdss2 | -13.626078 | 0.036514028 |
| 342 | 1428525_at | 4930488B01Rik | -13.624238 | 0.036552022 |
| 343 | 1434398_at | Nkrf | -13.615835 | 0.03672614 |
| 344 | 1454800_at | Morn2 | -13.60874 | 0.036873865 |
| 345 | 1448995_at | Pf4 | -13.580105 | 0.037476873 |
| 346 | 1451490_at | Lyplal1 | -13.566281 | 0.037771939 |
| 347 | 1455087_at | D7Ertd715e | -13.565406 | 0.037790709 |
| 348 | 1451122_at | Idi1 | -13.562816 | 0.037846303 |
| 349 | 1416256_a_at | Tubb5 | -13.537593 | 0.038392615 |

| | | | | |
|---|---|---|---|---|
| 350 | 1421396_at | Pcsk1 | -13.529433 | 0.038571235 |
| 351 | 1438176_x_at | Snap47 | -13.508043 | 0.039043896 |
| 352 | 1431046_at | Ppfia3 | -13.50241 | 0.039169451 |
| 353 | 1429266_at | Tmhs | -13.502068 | 0.039177096 |
| 354 | 1418746_at | Pnkd | -13.490861 | 0.039428275 |
| 355 | 1430034_at | Cct4 | -13.469681 | 0.039907937 |
| 356 | 1434083_a_at | Elmod1 | -13.440734 | 0.040574116 |
| 357 | 1416022_at | Fabp5 | -13.436336 | 0.040676421 |
| 358 | 1424072_at | 2010107G23Rik | -13.435663 | 0.040692096 |
| 359 | 1418444_a_at | Gde1 | -13.422497 | 0.041000217 |
| 360 | 1452445_at | Slc41a2 | -13.411501 | 0.04125955 |
| 361 | 1449381_a_at | Pacsin1 | -13.410627 | 0.041280234 |
| 362 | 1427347_s_at | Tubb2a | -13.378011 | 0.042060757 |
| 363 | 1420048_at | C78859 | -13.361665 | 0.042458149 |
| 364 | 1438396_at | Ocrl | -13.352984 | 0.042670904 |
| 365 | 1440928_at | D630037F22Rik | -13.352895 | 0.042673085 |
| 366 | 1418588_at | Nrsn1 | -13.342067 | 0.042940144 |
| 367 | 1458622_at | Ntrk2 | -13.332992 | 0.043165406 |
| 368 | 1435968_at | Ptprn2 | -13.330754 | 0.043221154 |
| 369 | 1448812_at | Hpcal1 | -13.325271 | 0.043358091 |
| 370 | 1456633_at | Trpm3 | -13.313853 | 0.043644829 |
| 371 | 1452473_at | Prr15 | -13.309305 | 0.043759635 |
| 372 | 1456523_at | C77713 | -13.293546 | 0.044160059 |
| 373 | 1434535_at | Krt222 | -13.285202 | 0.044373736 |
| 374 | 1415845_at | Syt4 | -13.27695 | 0.044586181 |
| 375 | 1435783_at | Fam169a | -13.272256 | 0.044707546 |
| 376 | 1455882_x_at | Vwc2 | -13.272005 | 0.044714039 |
| 377 | 1420679_a_at | Aig1 | -13.254862 | 0.045160473 |
| 378 | 1438305_at | Rims1 | -13.221414 | 0.046045986 |
| 379 | 1419546_at | Atp6v1c1 | -13.218557 | 0.046122538 |
| 380 | 1430780_a_at | Pmm1 | -13.215341 | 0.046208845 |
| 381 | 1434315_at | Nipal3 | -13.21243 | 0.046287131 |
| 382 | 1428049_a_at | Nudt16l1 | -13.211124 | 0.046322306 |
| 383 | 1426448_at | Pja1 | -13.197898 | 0.046680205 |
| 384 | 1439459_x_at | Acly | -13.17696 | 0.047253141 |
| 385 | 1448600_s_at | Vav3 | -13.173794 | 0.047340438 |
| 386 | 1426887_at | Nudt11 | -13.157439 | 0.04779438 |
| 387 | 1424029_at | Tspyl4 | -13.157206 | 0.047800892 |
| 388 | 1435852_at | Spred3 | -13.140179 | 0.048278733 |
| 389 | 1455144_s_at | AU040829 | -13.129505 | 0.048581007 |
| 390 | 1434105_at | Epm2aip1 | -13.105586 | 0.049266118 |
| 391 | 1423221_at | Tubb4 | -13.099685 | 0.049436807 |
| 392 | 1453065_at | Aldh5a1 | -13.090774 | 0.049695791 |
| 393 | 1453120_at | Tmx4 | -13.086395 | 0.049823632 |
| 394 | 1451033_a_at | Trpc4 | -13.076864 | 0.050103114 |
| 395 | 1452323_at | Spryd3 | -13.076565 | 0.05011191 |
| 396 | 1436448_a_at | Ptgs1 | -13.066511 | 0.050408738 |
| 397 | 1433596_at | Dnajc6 | -13.065176 | 0.050448271 |
| 398 | 1455148_at | Tmem130 | -13.064905 | 0.050456309 |
| 399 | 1451229_at | Hdac11 | -13.058079 | 0.050659175 |

| Down-Regulated | | | | |
|---|---|---|---|---|
| 1 | 1426048_s_at | Tcfap2a | 18.8011969 | 0.002984862 |
| 2 | 1426812_a_at | Fam129b | 18.6740158 | 0.00314799 |
| 3 | 1450994_at | Rock1 | 17.4197496 | 0.005424697 |
| 4 | 1442063_at | Adamtsl1 | 16.7292229 | 0.007439144 |
| 5 | 1436840_x_at | Rpl35 | 16.4723071 | 0.008393202 |
| 6 | 1456135_s_at | Pxn | 16.4592421 | 0.008445266 |
| 7 | 1423771_at | Prkcdbp | 16.0717801 | 0.010166986 |
| 8 | 1430978_at | Rps25 | 16.029682 | 0.010376663 |
| 9 | 1423425_at | Plbd2 | 15.4008967 | 0.014163098 |
| 10 | 1449007_at | Btg3 /// Gm7334 | 15.3044709 | 0.014870813 |
| 11 | 1428113_at | Tmtc4 | 15.2629261 | 0.015187883 |
| 12 | 1424704_at | Runx2 | 15.0035771 | 0.017347267 |
| 13 | 1460036_at | Ap1s2 | 14.9437791 | 0.017892688 |
| 14 | 1449484_at | Stc2 | 14.723464 | 0.020074682 |
| 15 | 1439439_x_at | Eef1d | 14.7222485 | 0.02008752 |
| 16 | 1423536_at | Strn3 | 14.7105706 | 0.020211329 |
| 17 | 1448785_at | Runx1t1 | 14.6854947 | 0.020480086 |
| 18 | 1436325_at | Rora | 14.5638037 | 0.021842541 |
| 19 | 1435972_at | Cast | 14.5157617 | 0.022408155 |
| 20 | 1418135_at | Aff1 | 14.3268176 | 0.02479752 |
| 21 | 1448347_a_at | Caprin1 | 14.3194122 | 0.024896819 |
| 22 | 1449099_at | Lrba | 14.2400722 | 0.025989119 |
| 23 | 1421043_s_at | Arhgef2 | 14.1050282 | 0.027974052 |
| 24 | 1460644_at | Bckdk | 13.9108203 | 0.031133331 |
| 25 | 1423753_at | Bambi | 13.8244339 | 0.032665566 |
| 26 | 1460211_a_at | Kdelr1 | 13.7376391 | 0.034290641 |
| 27 | 1435777_at | Itpripl2 | 13.6238068 | 0.036560938 |
| 28 | 1416133_at | Efr3a | 13.619167 | 0.036656982 |
| 29 | 1418564_s_at | Serbp1 | 13.1277282 | 0.048631537 |

T (-) : Up-regulated in ENCCs with GDNF

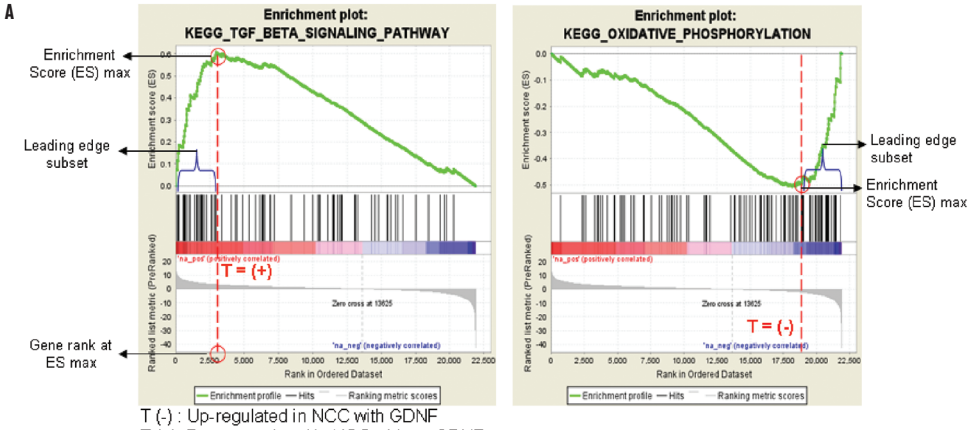
T (+): Down-regulated in ENCCs without GDNF

**Supplementary Table 4 -** Single Gene Analysis: Genes with higher expression in ENCCs than in the gut and further up-regulated in ENCCs (+ GDNF) (Bonferroni : $p < 0.05$)

| No | Gene Symbol | Gene Name | Human Chromosome |
|---|---|---|---|
| 1 | Aplp1 | amyloid beta (A4) precursor-like protein 1 | 19q13.12 |
| 2 | Napg | N-ethylmaleimide-sensitive factor attachment protein, gamma | 18p11.22 |
| 3 | Mtap4 | microtubule-associated protein 4 | 3p21.31 |
| 4 | Dpp10 | Dipeptidyl-peptidase 10 | 2q14.1 |
| 5 | Brsk2 | BR serine/threonine kinase 2 | 11p15.5 |
| 6 | Sez6l2 | Seizure related 6 homolog (mouse)-like 2 | 16p11.2 |
| 7 | hcfc1r1 | Host cell factor C1 regulator 1 (XPO1 dependent) | 16p13.3 |
| 8 | Snca | Synuclein, alpha | 4q22.1 |
| 9 | Gabbr1 | gamma-aminobutyric acid (GABA) B receptor, 1 | 6p22.1 |
| 10 | Galnt13 | UDP-N-acetyl-alpha-D-galactosamine:polypeptide | 2q23.3 |
| 11 | Cend1 | cell cycle exit and neuronal differentiation 1 | 11p15.5 |
| 12 | Stmn2/Scg10 | stathmin-like 2 | 8q21.13 |
| 13 | Pcsk1n | proprotein convertase subtilisin/kexin type 1 inhibitor | Xp11.23 |
| 14 | Brunol4 | Bruno-like protein 4 | 18q12.2 |
| 15 | Napb | N-ethylmaleimide-sensitive factor attachment protein, beta | 20p11.21 |
| 16 | Kcnq5 | potassium voltage-gated channel, KQT-like subfamily, member 5 | 6q13 |
| 17 | Gdap1 | ganglioside-induced differentiation-associated protein 1 | 8q21.11 |
| 18 | Asb1 | ankyrin repeat and SOCS box-containing 1 | 2q37.3 |
| 19 | Arhgef7 | Rho guanine nucleotide exchange factor (GEF) 7 | 13q34 |
| 20 | Dnajc5 | DnaJ (Hsp40) homolog, subfamily C, member 5 | 20q13.33 |
| 21 | Ndrg4 | NDRG family member 4 | 16q21 |
| 22 | Negr1 | neuronal growth regulator 1 | 1p31.1 |
| 23 | Rufy2 | RUN and FYVE domain containing 2 | 10q21.3 |
| 24 | Rgs17 | Regulator of G-protein signaling 17 | 6q25.2 |
| 25 | Rab39b | RAB39B, member RAS oncogene family | Xq28 |
| 26 | Kcnc2 | Potassium voltage gated channel, Shaw-related subfamily, member 2 | 12q21.1 |
| 27 | Ppp2r2c | protein phosphatase 2, regulatory subunit B (PR 52), gamma isoform | 4p16.1 |
| 28 | Atl1 | atlastin GTPase 1 | 14q22.1 |
| 29 | slc6a15 | solute carrier family 6 (neurotransmitter transporter), member 15 | 12q21.31 |
| 30 | Frmd3 | FERM domain containing 3 | 9q21.32 |
| 31 | mtmr9 | myotubularin related protein 9 | 8p23.1 |
| 32 | stxbp1 | syntaxin binding protein 1 | 9q34.11 |
| 33 | Pcsk1 | proprotein convertase subtilisin/kexin type 1 | 5q15 |
| 34 | Pmm1 | brain glucose-1,6-bisphosphatase | 22q13.2 |
| 35 | 6330439K17Rik /C20orf12 | RIKEN cDNA 6330439K17 gene/ chromosome 20 open reading frame 12 | 20p11.23 |
| 36 | Atp6v0e2 | ATPase, H+ transporting, lysosomal V0 subunit E2 | 7q36.1 |
| 37 | Cdkl2 | cyclin-dependent kinase-like 2 | 4q21.21 |
| 38 | Slc41a2 | solute carrier family 41, member 2 | 12q24.11 |
| 39 | Tmx4 | thioredoxin-related transmembrane protein 4 | t20p12 |

**Supplementary Table 5 -** Genes with higher expression in ENCCs than in the gut and down-regulated in ENCCs (+ GDNF) (Bonferroni : $p < 0.05$)

| No | Gene | Gene Name | Human Chromosome |
|---|---|---|---|
| 1 | Aff1 | AF4/FMR2 family, member 1 | 4q21.3 |
| 2 | Cast | Calpastatin | 5q15 |
| 3 | Tmtc4 | Transmembrane and tetratricopeptide repeat containing 4 | 13q32.3 |
| 4 | Btg3 | B-cell translocation gene 3 | 21q21.1 |
| 5 | Prkcdbp | protein kinase C, delta binding protein | 11p15.4 |
| 6 | Adamtsl1 | ADAMTS-like 1 | 9p22.2 |
| 7 | Fam129b | family with sequence similarity 129, member B | 9q34.11 |
| 8 | Tcfap2a | transcription factor AP-2 alpha | 6p24.3 |
| 9 | Rora | RAR-related orphan receptor A | 15q22.2 |
| 10 | arhgef2 | Rho/Rac guanine nucleotide exchange factor (GEF) 2 | 1q22 |



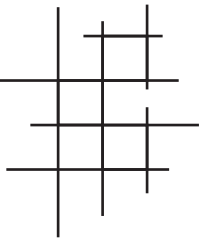

T (-) : Up-regulated in NCC with GDNF
T (+): Down-regulated in NCC without GDNF

B

| | | Biosynthesis | Differentation | Migration | Proliferation | Apoptosis | Miscellaneous |
|---|---|---|---|---|---|---|---|
| **Up-regulated** by GDNF in ENCCs | Parkinson's disease | | | | | | ■ |
| | Citrate Cycle | ■ | | | | | |
| | Oxidative phosphorylation | ■ | | | | | |
| | Peroxisome | ■ | | | | | |
| | Steroid Biosynthesis | ■ | | | | | |
| | Valin-Leucin-Isoleucin | ■ | | | | | |
| | PPAR | ■ | | | | | |
| | Butanoate metabolism | ■ | | | | | |
| | Propanoate metabolism | ■ | | | | | |
| **Down-regulated** by GDNF in ENCCs | N-Glycan Biosynthesis | ■ | | | | | |
| | Notch | | ■ | | | | |
| | Wnt | | ■ | | | | |
| | TGFβ/BMP | | ■ | | | | |
| | Axon guidance | | ■ | | | | |
| | Adherens junction | | | ■ | ■ | | |
| | Tight junction | | | ■ | ■ | | |
| | Focal adhesion | | | ■ | ■ | | |
| | Cell cycle | | | | ■ | | |
| | p53 | | | | ■ | ■ | |
| | TLRs | | | | | ■ | |
| | Pathaways in Cancer | | | | | | ■ |
| | Splicesome | | | | | | ■ |
| | Ribosome | | | | | | ■ |
| | Chronic Myeloid | | | | | | ■ |

**Supplementary Figure 1 -** A) An example of GSEA enrichment profile of the down-regulated (Tfb-β signaling pathway) and the up-regulated (Oxidative Phosphorylation signaling pathway) in ENCCs treated with GDNF. B) Summary of GSEA enriched signaling pathways as extracted from the KEGG database (when comparing ENCCs treated with GDNF to untreated cells). The possible biological function in ENCCs during ENS development is given.

# 6 General Discussion and Future Prespective



HSCR is a congenital disease characterized by the absence of ganglia in the myenteric and submucosal plexus of the gastrointestinal (GI) tract.[1] So far, 12 genes and five loci have been found associated with disease development. Of these genes *RET* has proved to be the major risk factor for HSCR, as 50% of the familial cases and 15-35% of the sporadic cases have *RET* coding sequence (CDS) mutations.[2] Mutations in the other 11 genes (*EDNRB*[3], *EDN3*[4, 5], *GDNF*[6, 7], *NTN*[8], *SOX10*[9], *PHOX2B*[10], *ECE1*[11], *KIAA1279/KBP*[12], *ZFHX1B*[13, 14], *NRG1*[15] and *TTF1*[16]) occur in no more than 20% of the remaining cases, most of which are syndromic-HSCR cases. The aims of the studies described in this thesis were first: to identify new HSCR-genes in associated loci that have been reported in previous studies, in particular in 9q31 and 4q31.1-q32.2.[17, 18] Second, we investigated whether the non-coding *RET* SNP rs2506004 (C>A), which we had previously identified as a candidate disease-associated mutation, was indeed pathogenic.[19] For this, we performed functional assays and we revealed the possible molecular mechanism by which the mutation pinpointed by this SNP contributes to disease development. As RET seems to be the central factor for HSCR- and enteric nervous system (ENS) development, we studied the RET networks in detail by performing gene expression profiling studies of mouse embryonic enteric neural crest cells (ENCCs), the progenitor cells of the ENS, untreated or treated with GDNF, the ligand of RET.

## 1. Identification of disease-causing genes in HSCR-associated loci

Mutations in the coding regions of the 12 genes that have been identified so far only explain about 15-35% of the sporadic HSCR cases and slightly more than 50% of the familial HSCR cases. Furthermore, reduced penetrance of the coding sequence (CDS) mutations of the 12 genes and the phenotypic variability suggests the involvement of other modifier genes or factors in the disease development. Bolk and colleagues identified a *RET*-dependent modifier locus on 9q31 in families without *RET* CDS mutations but who are linked to the *RET* locus (with the exception of one family). In Chapter 2 we tried to fine map this locus and to identify the disease-causing gene by genotyping 301 tag-SNPs spanning 7 Mb of the 9q31 region. As the 9q31 locus was identified in patients of European descent, we performed this experiment on 137 HSCR Dutch trios. We grouped the probands based on their *RET* CSD mutations status (121 probands without the *RET* CDS mutation and 16 probands with it). We identified the SNPs rs10816998 and rs7038415, located within *SVEP1,* as strongly associated with HSCR in the patients who do not carry a *RET* CDS mutation. This result is in line with the finding of Bolk and colleagues.[17] However, we could not replicate this result in the independent

case-control study (again using Dutch patients and controls). This might be due to population stratification issues, but it might also mean that this locus, in the end, does not contain a common disease variant for the Dutch HSCR patients. If such a common mutation does exist, the sample size we used was simply not large enough to detect significant association. We also replicated our study in 173 Chinese patients and 436 controls and we identified associated SNPs in different locations than those initially found in the Dutch population. These SNPs are located in the intronic and exonic regions of *IKBKAP.* This result was replicated in a second independent patient cohort, also from China. In contrast to the previous finding by Bolk, this association was stronger in patients with *RET* CDS mutations. We checked whether there were major linkage disequilibrium (LD) differences between the CEU and CHB HapMap populations in this locus. This was not the case, so we concluded that the differences we found were not due to differences in LD structure but due to the different ethnic origins of the patient cohorts. In view of these results, we concluded that in the Chinese population, the 9q31 locus, which was found associated with HSCR in families bearing *RET* CDS mutations, might at least contain a common disease-associated variant in the Chinese patient population. As this Chinese locus was not identified in the Dutch patient population, suggesting population specificity. The unknown variant in the Chinese patient population is likely present in the *IKBKAP* gene.

Brooks and colleagues identified a linkage region on 4q31.1-q32.2 in a multi-generational Dutch family affected by HSCR disease.[18] This region contains 57 genes. Using an exome sequencing approach on two out of five affected family members (**Chapter 3**), we tried to identify the disease-causing gene in this locus. We found only one possible disease-causing variant in the linkage region, in exon 20 of the *LRBA* gene. This variant was present in all five affected family members. However, we looked for the presence of this variant in a control population and unfortunately found it was present in heterozygously 4 of the 220 controls (0.9%, 4 out of 440 chromosomes). This makes it hard to conclude that it is the disease-associated mutation. As our screen was genome-wide, we were also able to look for disease modifiers not present on chromosome 4. A possible modifier present in three out of the five affected family members was found on chromosome 2, in exon 19 of *ARMC9.* Again however, we found this variant in 2% of the control population (6 out of 289 chromosomes). These results showed that these variants are unlikely to be HSCR-predisposing mutations.

Why did we not find any clear mutation on chromosome 4 although linkage was identified? Did we miss the mutation? It seems we did. The enrichment method we used only covered around 81.33% of all human exons, meaning that we lost 18.67% data to begin with. Furthermore, only 62.8% and 66.70% of the data for the two patients had a minimum coverage

of 10, the criteria we set for SNP calling. In total, we were unable to look at coding mutations in more than 40% of the exome. This we think is probably the best explanation for not finding the real causative-variant in the linkage region.

This finding also directly points to the difficulties we, as researchers, will meet when performing exome sequencing. Of course, we might find a mutation (and would be happy), but if unsuccessful, the question remains of what should we do next? Wait for better enrichment kits or develop locus-specific exome enrichment kits? This seems like a good idea and might solve most of the problems. However, we should not totally exclude the possibility that the mutation is present in an intronic or regulatory sequence. For the *LRBA* gene, we speculate in Chapter 3 that there may indeed be non-coding mutations that influence not the *LRBA* gene but the *MAB21L2* gene, which is present within the *LRBA* locus (for an extended discussion on this, see Chapter 3). If so, the enrichment kits might just not cover the mutation. We could of course go back to specific genes in the linkage area and focus on those, and think of making enrichment kits for these candidates (both coding and non-coding). But, as shown in Chapter 3, we did go back to the best candidate gene, *MAB21L2,* and sequence the entire genome region, including the regulatory sequences, but we still did not find any candidate disease-causing mutation. So focusing on a candidate gene is not always the answer to the problem. In the case of *MAB21L2,* we might extend the sequence region to the entire *LRBA* gene (see above) but even this might not prove successful. These difficulties are also nicely illustrated by a previous study we performed on a family with Goldberg-Shprintzen syndrome.[12] None of the candidate genes in the linkage area proved to be mutated, and in fact it was the last gene on a list of 35 that finally proved to be mutated. In this specific family, the unbiased approach of exome sequencing might have been helpful.

Another problem that became clear in this study is that our search for modifiers will be extremely difficult. Screening exomes results in hundreds of variants, while screening large numbers of patients and controls will be necessary to prove any possible involvement. Furthermore, the variants we found were not present in the larger databases such as dbSNP and 1000 Genomes, but by screening population-specific controls, we confirmed their presence in our population. This last finding shows that we really need to build population-specific databases.

Clearly, exome sequencing is a great tool that will help us in an unbiased way to find mutations, but equally clearly, this approach will not always be successful and the results obtained should be examined with great care.

## 2. Functional studies on HSCR-associated variant SNP rs2506004 (C>A)

As mentioned above, mutations in coding sequences of the known HSCR genes only explain some of the HSCR cases, hence it is predicted that there are more genes involved in the disease development. However, besides other genes, non-coding mutations in the known genes might also contribute to disease development. Indeed, in the past six years, substantial evidence was presented that *RET* non-coding sequence mutations play an important role in disease development, especially in sporadic S-HSCR cases. At least three *RET* non-coding variants have been identified and proven to be functional variants, as the mutant allele of these variants decreased reporter gene expression in luciferase assays. Two out of the three variants are located in the *RET* promoter region (rs10900296, G>A; rs10900297, A>C) and one is located in the multi-species conserved region (MCS+9.7) in intron 1 (rs2435357, C>T) of *RET*. We identified another associated-variant (SNP rs2506004, C>A) which is located 217 bp downstream of SNP rs2435357, also in MCS+9.7.[19] We considered this variant a good candidate since it is highly associated with HSCR disease and conserved not only in mammals but also in avian species (chicken). Interestingly, when Emison and colleagues performed luciferase assays for the region containing SNP rs2435357 (named Enh1 from now onward), they also included SNP rs2506004 (named Enh2 from now onward). In their *in vitro* studies they showed that the T- (disease-associated) allele of Enh1 reduced the luciferase expression and disturbed the binding site of SOX10, the transcription factor of *RET*. However, it has not yet been proven that Enh2, which was present in the construct used in these studies, did not contribute to the disease development. We therefore decided to perform functional assays for Enh2 (**Chapter 4**). Luciferase assays showed that both of the disease-associated alleles (Enh1-T and Enh2-A) reduced the luciferase expression independently under regulation of the *RET*-promoter.

Using *in silico* analysis we identified Enh2-C, the non-disease allele, and its surrounding sequence to be present in the -A**C**GTG- sequence, also known as the Central Nervous System Midline Element (CME), is a binding site of Nxf/Arnt2 and the Sim2/Arnt2 heterodimers. By electrophoresis mobility shift assay (EMSA) we showed that Nnxf/Arnt2 has less binding affinity to CME containing the Enh2-T (disease associated) allele compared to CME containing the Enh2-C (non-disease) allele, while Sim2/Arnt2 has the same binding affinity to both disease and non-disease associated alleles. Interestingly, we identified six CME regions in the *RET* promoter. This led us to hypothesize that NXF/ARNT2 and SIM2/ARNT2 regulated *RET* expression not only by binding to Enh2 but also by binding to the *RET* promoter. We proved this hypothesis by co-transfection of constructs containing *Nxf*/*Arnt2* or *Sim2*/*Arnt2* with the *RET* promoter coupled to *luciferase* into a mouse neuroblastoma cell line (Neuro2A). The re-

sults showed that Nxf/Arnt2 and Sim2/Arnt2 increase and decrease the luciferase expression, respectively. Furthermore, the transfection of *Nxf/Arnt2* or *Sim2/Arnt2* into mouse neuroblastoma cells line (Neuro2A) showed increased and decreased endogenous RET expression. This strengthens the idea that NXF/ARNT2 and SIM2/ARNT2 are transcription regulators for RET expression. It is interesting to mention that *SIM2* is located in the critical region of Down's syndrome (DS) in chromosome 21. It has been hypothesized that overexpression of SIM2 in DS patients reduces the expression of Drebrin (DBN1), a protein that is important for synaptic plasticity. In this study we showed that overexpression of Sim2 leads to down-regulation of Ret expression. This could explain why DS patients have a 40-fold higher risk of having HSCR disease. In conclusion, in this study we showed that Enh1 and Enh2 are both functional variants and the disease-associated alleles of both variants decrease *luciferase* expression. We also showed that Nxf/Arnt2 and Sim2/Arnt2 are expressed in the enteric nervous system (ENS) progenitor cells and these heterodimers are a transcription activator and repressor, respectively, for *RET* expression.

Besides the fact that we could confirm the probable involvement of the SNP we had previously identified, we had to conclude that multiple SNPs present on the same disease-associated haplotype may play a role in the disease, likely at different stages of development, or are involved in different processes. As mentioned above, previous studies showed evidence for the involvement of at least two additional SNPs on the same haplotype. These SNPs are located at -5 and -1 from the start site of the *RET* gene (Figure 1 in Chapter 4). It was also shown that the TTF-1 protein binds to this region and that TTF-1 can activate the *RET* promoter[16, 20, 21]. Therefore, previously published data and our new data hold an important message for the study of complex diseases, namely that more than one SNP on an associated haplotype might influence disease development. So far, almost all studies have stopped if they found a SNP with a proven functional effect. We have shown that even when one such SNP has been found, there may still be more involved in the disease development. This means that polygenic diseases are even more complex than originally thought.

As we have identified new transcription regulators for *Ret*, there is the question of whether variants/mutations in those transcription regulators might also contribute to disease development. So far, association to 11q13 (where NXF is located) or 15q25.1 (where ARNT2 is located) has not been reported. Rare coding mutations have not yet been looked for, but whether they exist in these transcription factors and cause or contribute to HSCR can, however, be doubted. These transcription factors are involved in many developmental processes and if any mutation were present, it is most likely it would result in a complex clinical phenotype. It is more likely that the sequences in HSCR genes to which these transcription fac-

tors bind (the CME regions) might contain SNPs that disturb proper binding, as shown for SNP rs2506004 in MCS+9.7. We have already performed one study to test this hypothesis. We checked the six CME regions within the *RET* promoter for mutations in 107 HSCR patients (patients who do not carry *RET* CDS) but did not find any mutation (unpublished data).

## 3. RET mediated gene expression in ENS precursors

The 12 HSCR genes that have been identified so far can be roughly classified into three groups: those involved in the RET pathways (*RET*, *GDNF*, *NTN*), those involved in EDNRB pathways (*EDNRB*, *EDN3*, *ECE-1*), and transcription factors that can affect both RET and/or EDNRB pathways (*SOX10*, *ZFXH1B*, *PHOX2B*). These three protein groups/pathways, however, are not totally independent. For instance, a genome-wide association study using 43 trios from a genetically isolated Mennonite population reported statistically significant joint transmission of RET and EDNRB alleles.[22, 23] Furthermore, mice homozygous for the recessive hypomorphic allele of *Ednrb* (EdnrbS or Piebald) and heterozygous for a *Ret* null mutation ($Ret^{+/-}$) showed high frequencies of aganglionosis. These findings suggest epistasis between EDNRB and RET, and this idea was further supported by *Ret*/*Ednrb* mouse crosses.[24] Similar studies showed that, whereas $Ret^{51/51}$ and $Edn3^{ls/ls}$ (lethal spotting) mice display colonic aganglionosis, combinations of these mutant alleles lead to almost complete intestinal aganglionosis.[25] Moreover, the transcription factor SOX10 is involved in regulating both *RET* and *EDNRB* expression.[26-28] PHOX2B is also a transcription factor that is expressed in several classes of differentiating neurons of both the peripheral and central nervous systems. Mice with a homozygous disruption of *Phox2b* show no enteric ganglia and no *Ret* expression, which suggests that Phox2b might play a regulatory role in *Ret* expression.[29]

These data together indicate that all the known HSCR genes seem to be connected, physically or indirectly, and they point to a central role for RET in the development of both HSCR and the enteric nervous system (ENS) development. Therefore, understanding the protein networks in which the HSCR associated proteins are involved, in particular the RET associated pathways, might help us not only in understanding ENS development, but also in identifying the missing HSCR genes.

However, these pathways are still largely unknown, in particular the signaling pathways regulated by RET in ENCCs. To gain more insight into the RET protein network, we performed gene expression profiling studies on RNA isolated from mouse ENCCs, untreated or treated with GDNF, the ligand for the RET receptor (**Chapter 5**). We compared these data to

each other and also to the expression profile from the whole mouse embryonic gut (all E14.5). The data were used to perform pathway analysis using GSEA and for single gene analysis.

With single gene analysis we identified genes that were differentially expressed in ENCCs with GDNF compared to those in untreated ENCCs. Moreover, by having gene expression profiling from the mouse gut, we also identified genes that were differentially expressed in ENCCs compared to those in the gut. This group of genes could be considered as the ENS marker candidate resource, as Sox2 that was recently identified as a maker of ENS progenitor cells by Heanue et al. was listed in this group.[30] Two HSCR-associated genes, *Phox2b* and *Ece1*, were up- and down-regulated respectively in ENCCs compared to those in the gut. By comparing the gene expression profiling of ENCCs treated with GDNF to those in the gut, we were able to identify more HSCR-associated genes, including *Ret* itself, *Phox2b, Edn3, Nrg1, Gfrα1* and *Nrtn.* This led to the question whether these genes, including *RET*, are down-effectors of the RET signaling pathway and hence RET signaling would have a positive feedback loop system. As it was known from a previous study that PHOX2B is a transcription activator for RET expression, the up-regulation of PHOX2B upon GDNF stimulation, as shown in this study, would be a possible mechanism of the RET positive feedback loop system. We also identified *Snca* (α-Synuclein), a gene known to be involved in Parkinson's disease, as being up-regulated in ENCCs and further up-regulated upon GDNF stimulation. Our data suggest that defects in RET signaling would most likely reduce the expression of α-synuclein. Interestingly, in Parkinson's patients, the presence of Lewy bodies (aggregation of α-synuclein) in their gut is associated with chronic constipation. These findings suggest that α-synuclein might play an important role in ENS functioning. Recently, it has been proposed that detecting Lewy bodies in the gut could be a method for the early detection of Parkinson's disease.

Despite many interesting new findings in single gene analysis, it is still rather difficult to gain a bigger picture and conclude what actually happens when we stimulate RET with GDNF in ENCCs. RET is believed to be important for ENCCs proliferation, differentiation, migration and survival during ENS development, but it has not yet been reported in detail which signaling pathways triggered by RET play these roles. To get more insight into the RET signaling pathway in ENCCs, we therefore decided to analyze our data by pathway analysis, using the gene set enrichment analysis (GSEA) method.[31] The results showed that some pathways previously implicated in ENS, as well as in embryonic development, such as Tgfβ/Bmp, Notch and Wnt were down-regulated in ENCCs upon GDNF stimulation. These pathways are known to be important for cell differentiation and neurogenesis. It has been reported earlier that GDNF stimulation inhibits cell apoptosis of ENCCs.[32] In this study, we showed that the signaling pathways important for inducing cell apoptosis, e.g. p53 and TOLL-like

receptors (TLRs), were down-regulated, and the set of genes that induce cell cycle arrest in the cell cycle pathway (p53, mdm2, p21) were also down-regulated. Our study showed that to induce cell proliferation, instead of activating cell division genes, the RET signaling pathway represses the negative regulators of cell cycle genes. GDNF is known to be a chemo-attractant for ENCCs and to stimulate their migration and we showed that GDNF down-regulated adherens and tight junction pathways in ENCCs. As a consequence the cell-cell contact might loosen and thereby stimulate cell migration. This process is a possible mechanism for how GDNF encourages ENCCs to migrate during ENS development. Altogether, there is a general trend in our data that genes encoding proteins that stimulate cell differentiation, negative regulators of cell proliferation, and those important for cell-cell adhesion and apoptosis are down-regulated in ENCCs upon treatment with GDNF. On the other hand, genes that encode proteins that have the opposite effect, such as genes for early growth response, those involved in inhibiting cell differentiation, and those important for energy supply and biosynthesis, are up-regulated.

It is known that the TLRs signaling pathway can induce cell apoptosis and and innate immune response by stimulating pro-inflammatory cytokine expression. Down-regulation of TLRs expression upon GDNF stimulation in ENCCs might repress the expression of pro-inflammatory cytokines. Furthermore, to our surprise, the JAK-STAT pathway that is known to be activated by RET-GDNF was down-regulated in our data set.[33] We went through the genes listed in the leading subset of genes from GSEA analysis of the JAK-STAT signaling pathway and found many of them are genes encoding cytokines and cytokine receptors. Based on data of these two pathways, it seems that RET is involved in regulating the innate immune response and GDNF activation down-regulates this response in ENCCs. We confirmed Tlrs expression in mouse gut (E14.5) by immunostaining and showed that Tlr2, Tlr4 and Tlr5 are indeed expressed in neuron, glial and ENS progenitor cells in the myenteric plexus. We also confirmed by qRT-PCR that GDNF down-regulates the expression of *Tlr2, Tlr4* and *Tlr5* and some pro-inflammatory cytokines, such as *Tnfα* and ***Il-6***. These data showed that ENCCs have an immunomodulatory capacity and that the RET signaling pathway is involved in regulating the innate immune response. Defects in RET signaling, as often seen in HSCR patients, will induce TLRs expression and subsequently pro-inflammatory cytokine expression, and activate the innate immune response. Overexpression of pro-inflammatory cytokines in the gut might lead to HSCR-associated enterocolitis. However, the level of Tnf*α* expression in ENCCs was very low so that we could not detect protein levels by ELISA. On the other hand, for the Il-6, even though we were able to detect its protein level by ELISA, we were not able to detect the changes in protein level in ENCCs, untreated or treated with GDNF. We think that even

though ENCCs may play a role in the immune response, their response is not as fast as that of immune cells and that it therefore takes much longer for them to produce pro-inflammatory cytokines. The immune response is a very complex process, where the cross-talk between cytokines produced by different cells is needed to modulate the response. The *in vitro* study that we performed might not be a perfect model for studying such a complex process. A better and more straightforward method to prove the involvement of RET-GDNF signaling in the immune response would be an *in vivo* study using transgenic mouse bearing $RET^{(+/-)}$ and $RET^{(-/-)}$, and wild-type mouse, to sensitize them with bacteria and virus infections in their gut and compare their sensitivity in developing enterocolitis. We could also perform quantitative immunostaining for TLRs and RET in myenteric plexus and check for its association with enterocolitis development in these mice. This study could be combined with *in vitro* studies on ENCCs isolated from the transgenic and wild-type mice mentioned above. We could then check the expression levels of the pro-inflammatory cytokines upon stimulation with various pathogens recognized by TLRs (e.g. LPS, LPA, Flagellin).

## 4. Overall conclusion

In this thesis we describe both genetic and functional studies performed to identify and characterize HSCR-associated mutations and genes. Our work clearly shows the many difficulties encountered when working on a complex disease.

Identification of a gene in a susceptibility locus proved difficult. Both our attempts on chromosomes 4 and 9 were largely unsuccessful (**Chapters 2 and 3**). Larger patient cohorts for the association and exome sequence studies and better enrichment kits for the exome sequencing are necessary to be more successful.

Moreover, we should not forget that non-coding sequence variants can also play an important role in complex diseases (**Chapters 4**). This should certainly be taken into consideration if exome sequencing proves unsuccessful (**Chapter 3**). Importantly, we showed that multiple SNPs on a haplotype can be involved in disease development.

Finally, we gained insight into the signaling properties of RET in ENCCs (**Chapter 5**). Not only does this study offer new insight into ENS-related processes and a long list of candidate HSCR genes, it also made us hypothesize that RET could be a factor in the occurrence of enterocolitis in HSCR patients.

# REFERENCES


1. Amiel J, Lyonnet S. Hirschsprung disease, associated syndromes, and genetics: a review. J Med Genet 2001;38:729-739.

2. Hofstra RM, Wu Y, Stulp RP, Elfferich P, et al. RET and GDNF gene scanning in Hirschsprung patients using two dual denaturing gel systems. Hum Mutat 2000;15:418-429.

3. Puffenberger EG, Hosoda K, Washington SS, et al. A missense mutation of the endothelin-B receptor gene in multigenic Hirschsprung's disease. Cell 1994;79:1257-1266.

4. Edery P, Attie T, Amiel J, et al. Mutation of the endothelin-3 gene in the Waardenburg-Hirschsprung disease (Shah-Waardenburg syndrome). Nat Genet 1996;12:442-444.

5. Hofstra RM, Osinga J, Tan-Sindhunata G, et al. A homozygous mutation in the endothelin-3 gene associated with a combined Waardenburg type 2 and Hirschsprung phenotype (Shah-Waardenburg syndrome). Nat Genet 1996;12:445-447.

6. Angrist M, Bolk S, Halushka M, et al. Germline mutations in glial cell line-derived neurotrophic factor (GDNF) and RET in a Hirschsprung disease patient. Nat Genet 1996;14:341-344.

7. Salomon R, Attie T, Pelet A, et al. Germline mutations of the RET ligand GDNF are not sufficient to cause Hirschsprung disease. Nat Genet 1996;14:345-347.

8. Doray B, Salomon R, Amiel J, et al. Mutation of the RET ligand, neurturin, supports multigenic inheritance in Hirschsprung disease. Hum Mol Genet 1998;7:1449-1452.

9. Pingault V, Bondurand N, Kuhlbrodt K, et al. SOX10 mutations in patients with Waardenburg-Hirschsprung disease. Nat Genet 1998;18:171-173.

10. Amiel J, Laudier B, Attie-Bitach T, et al. Polyalanine expansion and frameshift mutations of the paired-like homeobox gene PHOX2B in congenital central hypoventilation syndrome. Nat Genet 2003;33:459-461.

11. Hofstra RM, Valdenaire O, Arch E, et al. A loss-of-function mutation in the endothelin-converting enzyme 1 (ECE-1) associated with Hirschsprung disease, cardiac defects, and autonomic dysfunction. Am J Hum Genet 1999;64:304-308.

12. Brooks AS, Bertoli-Avella AM, Burzynski GM, et al. Homozygous nonsense mutations in KIAA1279 are associated with malformations of the central and enteric nervous systems. Am J Hum Genet 2005;77:120-126.

13. Cacheux V, Dastot-Le Moal F, Kaariainen H, et al. Loss-of-function mutations in SIP1 Smad interacting protein 1 result in a syndromic Hirschsprung disease. Hum Mol Genet 2001;10:1503-1510.

14. Wakamatsu N, Yamada Y, Yamada K, et al. Mutations in SIP1, encoding Smad interacting protein-1, cause a form of Hirschsprung disease. Nat Genet 2001;27:369-370.

15. Garcia-Barcelo MM, Tang CS, Ngan ES, et al. Genome-wide association study identifies NRG1 as a susceptibility locus for Hirschsprung's disease. Proc Natl Acad Sci U S A 2009;106:2694-2699.

16. Garcia-Barcelo M, Ganster RW, Lui VC, et al. TTF-1 and RET promoter SNPs: regulation of RET transcription in Hirschsprung's disease. Hum Mol Genet 2005;14:191-204.

17. Bolk S, Pelet A, Hofstra RM, et al. 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 U S A 2000;97:268-273.

18. Brooks AS, Leegwater PA, Burzynski GM, et al. A novel susceptibility locus for Hirschsprung's disease maps to 4q31.3-q32.3. J Med Genet 2006;43:e35.

19. Burzynski GM, Nolte IM, Bronda A, et al. Identifying candidate Hirschsprung disease-associated RET variants. Am J Hum Genet 2005;76:850-858.

20. Fitze G, Schierz M, Kuhlisch E, et al. ovel intronic polymorphisms in the RET proto-oncogene and their association with Hirschsprung disease. Hum Mutat 2003;22:177.

21. Sancandi M, Griseri P, Pesce B, et al. ingle nucleotide polymorphic alleles in the 5' region of the RET proto-oncogene define a risk haplotype in Hirschsprung's disease. J Med Genet 2003;40:714-718.

22. Carrasquillo MM, McCallion AS, Puffenberger EG, et al. Genome-wide association study and mouse model identify interaction between RET and EDNRB pathways in Hirschsprung disease. Nat Genet 2002;32:237-244.

23. Carrasquillo MM, McCallion AS, Puffenberger EG, et al. Genome-wide association study and mouse model identify interaction between RET and EDNRB pathways in Hirschsprung disease. Nat Genet 2002;32:237-244.

24. McCallion AS, Stames E, Conlon RA, et al. Phenotype variation in two-locus mouse models of Hirschsprung disease: tissue-specific interaction between Ret and Ednrb. Proc Natl Acad Sci U S A 2003;100:1826-1831.

25. Barlow A, de Graaff E, Pachnis V. Enteric nervous system progenitors are coordinately controlled by the G protein-coupled receptor EDNRB and the receptor tyrosine kinase RET. Neuron 2003;40:905-916.

26. Lang D, Epstein JA. Sox10 and Pax3 physically interact to mediate activation of a conserved c-RET enhancer. Hum Mol Genet 2003;12:937-945.

27. Zhu L, Lee HO, Jordan CS, et al. Spatiotemporal regulation of endothelin receptor-B by SOX10 in neural crest-derived enteric neuron precursors. Nat Genet 2004;36:732-737.

28. Lang D, Chen F, Milewski R, et al. Pax3 is required for enteric ganglia formation and functions with Sox10 to modulate expression of c-ret. J Clin Invest 2000;106:963-971.

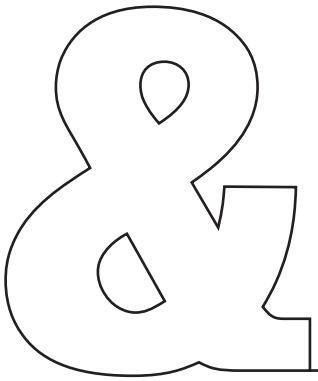
29. Pattyn A, Morin X, Cremer H, et al. The homeobox gene Phox2b is essential for the development of autonomic neural crest derivatives. Nature 1999;399:366-370.

30. Heanue TA, Pachnis V. Prospective Identification and Isolation of Enteric Nervous System Progenitors Using SOX2. Stem Cells 2010.

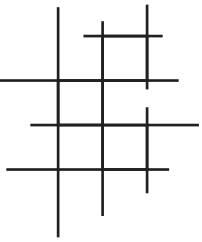
31. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545-15550.

32. Bordeaux MC, Forcet C, Granger L, et al. The RET proto-oncogene induces apoptosis: a novel mechanism for Hirschsprung disease. EMBO J 2000;19:4056-4063.

33. Plaza Menacho I, Koster R, van der Sloot AM, et al. RET-familial medullary thyroid carcinoma mutants Y791F and S891A activate a Src/JAK/STAT3 pathway, independent of glial cell line-derived neurotrophic factor. Cancer Res 2005;65:1729-1737.



Summary
Samenvatting
Ringkasan
List of Abbreviation
Acknowledgments



## Summary

In **Chapter 1,** we review the genetics, the developmental background, the clinical characteristics, and the diagnosis and treatment of Hirschsprung disease (HSCR). HSCR is a congenital disease characterized by the absence of ganglia in the myenteric and submucosal plexus of the gastrointestinal (GI) tract. The incidence of HSCR is 1 in 5000 live births and most HSCR patients present as sporadic cases, although in about 5-15% of the cases the disease is familial. In about 70% of the cases, HSCR occurs as an isolated trait, while the rest are associated with other diseases or syndromes. So far at least 12 genes and five loci (9q31, 4q31-q32, 3p21, 19q21 and 16q23) have been found associated with HSCR development. *RET* is a major genetic risk factor for HSCR as *RET* coding sequence (CDS) mutations are identified in 50% of the familial cases and in 15-35% of the sporadic cases. Mutations in 11 other genes (*EDNRB*, *EDN3*, *GDNF*, *NTN*, *SOX10*, *PHOX2B*, *ECE1*, *KIAA1279*/*KBP*, *ZFHX1B*, *NRG1* and *TTF1*) occur in around 20% of the cases, most of which prove to be syndromic. This leads to our hypothesis that there are probably several other disease-causing genes.

Bolk and colleagues identified linkage to 9q31 in five large HSCR families, all of which also showed linkage to *RET* (10q11.2), but none of them had a clear *RET* coding mutation and in one family that was not linked to *RET* and obviously did not have a *RET* coding mutation. In **Chapter 2** of this thesis, we describe the fine mapping of this locus, to identify the disease-contributing gene. We genotyped 301 tag-SNPs spanning 7 Mb on 9q31 region in 137 HSCR Dutch trios, of whom 121 probands did not carry a *RET* CDS mutation while 16 probands did. We identified two HSCR-associated SNPs in the gene *SVEP1* (rs10816998 and rs7038415) in patients not having a *RET* CDS mutation confirming Bolk's data. However, this result was not replicated in 107 independent HSCR Dutch patients without a *RET* CDS mutation. In a Chinese HSCR population, HSCR was found associated with SNPs in *IKBKAP* but not in *SVEP1.* This association was stronger in patients with *RET* CDS mutations than those without *RET* CDS mutations. The data was confirmed in an independent Chinese HSCR cohort. We concluded that association with chromosome 9 is population-specific and that in Chinese HSCR patients the gene likely involved is *IKBKAP*. The data, as found in the Chinese patients, further implies that Chinese *RET* mutation carriers may have an additional risk through this associated *IKBKAP* variant.

In a previous linkage study performed by Brooks and colleagues on a Dutch multi-generational HSCR family, linkage was identified for a region on 4q31-q32. In **Chapter 3** we tried to identify the disease-causing mutation in this linkage region. All of this was done by exome sequencing. We identified only one variant in the linkage region, in exon 20 of the *LRBA* gene.

The *LRBA* gene harbors *MAB21L2* and MAB21L2 plays a role in ENCCs migration during ENS development. It was therefore postulated that this variant might influence not *LRBA* but *MAB21L2*. However, although this variant was present in all five affected family members, it was also identified in 0.9% of healthy controls (4/440). It is thus unlikely that this variant is the disease-causing variant. But as this was the only variant we identified in the linkage region, we assume that we must have missed the causative mutation, possibly because the enrichment method and criteria we used was only able to analyze 65% of the target region properly. We will therefore repeat this experiment with a better enrichment method which covers all the exons in the genome.

Previous studies showed that, regardless of the *RET* CDS mutation status, almost all familial cases of HSCR are linked to the *RET* locus. Furthermore, association studies performed by several groups in Caucasian and Asian populations revealed a strongly associated RET haplotype. This haplotype covers 27 kb, starting 4 kb upstream of *RET* through intron 1 to the beginning of exon 2. Some of the variants present on this haplotype have been reported or suggested as HSCR-causative variants. However, for some associated variants, in particular SNP rs2506004 (C>A), which is present in an enhancer region for *RET*, the pathogenic nature has not yet been examined. In **Chapter 4**, we describe functional studies performed on this SNP (rs2506004, C>A). By *in silico* analysis we found that the C-allele (wild type) is located in a Central Midline Element (CME) sequence element, which is a binding site for Nxf/Arnt2 and Sim2/Arnt2, transcription activators and repressors, respectively. Luciferase assays showed that the enhancer element containing the A- (disease-associated) allele, when coupled to the *RET* promoter, reduced Luciferase expression compared to that measured for the same construct having the C- (non-disease associated) allele, or to the construct having the *RET* promoter only. We checked the binding affinity of Nxf/Arnt2 and Sim2/Arnt2 heterodimers to the oligos containing the C-allele or A-allele by EMSA and a supershift assay. We confirmed that Nxf/Arnt2 binds to the C-allele much more strongly than to the A-allele, while Sim2/Arnt2 has an equal binding capacity to the A- and C-alleles. The effects of Nxf/Arnt2 and Sim2/Arnt2 on the expression of Ret were further confirmed by co-transfection of constructs containing *Nxf/Arnt2* or *Sim2/Arnt2* into a mouse neuroblastoma (Neuro2A) cells line. We checked the endogenous Ret expression on protein level by Western blot and could show that Nxf/Arnt2 and Sim2/Arnt2 increased and decreased the endogenous expression of Ret, respectively. SIM2 is located in the Down syndrome's (DS) critical region in chromosome 21, and interestingly, 77% of DS patients have gastrointestinal (GI) abnormality, in which 2-15% of the cases have HSCR disease. As we proved that Sim2 is a transcription repressor of Ret, this might partly explain the high incidence of HSCR disease in DS patients.

All the published data indicate that the known HSCR genes identified so far seem to be connected, physically or indirectly. Moreover, they point to a central role for RET in the development of both HSCR and the enteric nervous system (ENS). Therefore, understanding the protein networks in which the HSCR-associated proteins are involved, in particular the RET-associated pathways, might help us to not only in understanding better ENS development, but also to identify candidate HSCR genes. However, these pathways are still largely unknown, in particular the signaling pathways that are regulated by RET in neural crest cells (ENCCs). To gain more insight into the RET protein network, we performed gene expression profiling studies on RNA isolated from mouse ENCCs, untreated or treated with GDNF, the ligand for the RET receptor (**Chapter 5**). We compared these data to each other and also to expression profiles from the whole mouse embryonic gut (all E14.5). The data were used to perform pathway analysis using gene set enrichment analysis (GSEA) and for a single gene analysis. The results of the pathway analysis showed a general trend in ENCCs treated with GDNF, towards down-regulation of genes encoding proteins that stimulate cell differentiation, negative regulators of cell proliferation, and proteins important for cell-cell adhesion and apoptosis. On the other hand, genes that encoded proteins that have the opposite effect, such as genes for early growth response, those involved in the inhibition of cell differentiation, and genes that are important for energy supply and biosynthesis, were up-regulated.

The Tlrs signaling pathway which plays a role in cell apoptosis and the innate immune response was also down-regulated by GDNF stimulation. By immunostaining we confirmed that Tlr2, Tlr4 and Tlr5 were expressed in the neuron, glial and ENS progenitor cells in the myenteric plexus of mouse gut E14.5. We also confirmed by qPCR that down-regulation of Tlrs in ENCCs by GDNF stimulation reduced the mRNA level of the pro-inflammatory cytokines Tnf$\alpha$ and Il-6 significantly. This result showed, for the first time, the connection between RET signaling and the innate immune response of ENCCs in the gut. This suggests that the inactivation of RET (e.g. as in RET CDS mutations in HSCR patients) could lead to the activation of the innate immune response (i.e hypersensitivity of immune response), which might manifest itself as HSCR-associated enterocolitis.

Together, our studies led to the identification of a candidate HSCR gene in the 9q31 locus, specifically for Chinese HSCR patients. We revealed a possible mechanism for SNP rs2506004 in regulating the expression of *RET* and hence revealed how the disease-associated allele contributes to disease development. Furthermore, we studied the RET signaling pathway in ENCCs in more detail. Our studies have led to more insight into HSCR- and ENS development, the genes and pathways involved, and how HSCR might develop.

# Samenvatting

In **hoofdstuk 1** bespreken we de genetica, de embryologische ontwikkeling, de klinische kenmerken en de diagnose en behandeling van de ziekte van Hirschsprung (HSCR). HSCR is een aangeboren ziekte die wordt gekenmerkt door de afwezigheid van ganglia in de myenteric en submucosale plexi van het maag-darmkanaal. De incidentie van HSCR is 1 op de 5000 levend geborenen. De meeste HSCR patiënten zijn sporadisch (zijn de enige in de familie) maar familiaire HSCR komt wel degelijk voor. We zien dit in ongeveer 5-15% van alle gevallen. In 70% van alle gevallen hebben de patiënten alleen HSCR, terwijl de overige 30% van de patiënten naast HSCR ook een andere ziekte of een syndroom heeft. Tot nu toe zijn 12 genen en vier chromosomale gebieden (dit noemen we loci) gevonden die geassocieerd zijn met de ontwikkeling van HSCR. *RET* is de belangrijkste erfelijke risicofactor voor HSCR, omdat in 50% van de familiaire gevallen een *RET* mutatie wordt gevonden. Dit getal is lager in de niet-familiaire gevallen (15-35%). Mutaties in de 11 andere genen (*EDNRB*, *EDN3*, *GDNF*, *NTN*, *SOX10*, *PHOX2B*, *ECE1*, *KIAA1279*/*KBP*, *ZFHX1B*, *NRG1* en *TTF1*) verklaren ongeveer 20% van de rest van de gevallen. De meeste patiënten met mutaties in de 11 andere genen blijken een syndroom te hebben waar HSCR een onderdeel van is. Omdat een groot deel van de patiënten geen mutatie heeft in de 12 genen is onze hypothese dat er waarschijnlijk nog verscheidene andere ziekte veroorzakende genen moeten zijn.

Bolk en collega's identificeerden een koppeling van de ziekte aan een stukje van chromosoom 9 (9q31). Ze vonden dit in vijf grote HSCR families. Vier van deze families bleken echter ook te koppelen met het *RET* gen (10q11.2), maar in geen van deze families was een duidelijke *RET* mutatie gevonden. De vijfde familie, die niet gekoppeld kon worden aan het *RET* gen, had logischerwijs ook geen *RET* mutatie. In **hoofdstuk 2** van dit proefschrift beschrijven we de zoektocht naar de genmutatie in dit gebied op chromosoom 9. We genotypeerden 301 tag-SNPs (Single Nucleotide Polymorphism (SNP)) verspreid over het gekoppelde gebied in 137 Nederlandse HCSR patiënten en hun ouders (121 hadden geen *RET* mutatie, 16 hadden dat wel). We identificeerden 2 HSCR geassocieerde SNPs in een gen genaamd *SVEP1* (rs10816998 en rs7038415). We vonden de associatie in patiënten die geen *RET* mutatie hebben. Dit bevestigde de data van Bolk. Echter, dit resultaat herhaalde zich niet in 107 onafhankelijke Nederlandse HSCR patiënten zonder een *RET* mutatie. In een Chinese HSCR populatie, bleek ook een associatie aanwezig te zijn maar met een ander gen, genaamd *IKBKAP*. Deze associatie was sterker in de groep patiënten met een *RET* mutatie dan in de groep zonder *RET* mutatie. De Chinese data kon wel worden bevestigd in een onafhankelijke groep Chinese HSCR patiënten. We concludeerden daarom dat de associatie met chromosoom 9 populatie specifiek

is en dat in de Chinese HSCR patiënten waarschijnlijk het gen *IKBKAP* betrokken is bij de ontwikkeling van de ziekte.

In een studie, beschreven door Alice Brooks en haar collega's werd in een grote Nederlandse HSCR familie koppeling gevonden tussen de ziekte en een regio op chromosoom 4 (4q31-q32). In **hoofdstuk 3** hebben we een poging gedaan het gen en de genmutatie te vinden die de ziekte in deze familie veroorzaakt. Om dit te doen hebben we alle genen in het genoom geanalyseerd ('Exoom sequencing'). We vonden één variant in het gekoppelde gebied, in exon 20 van het *LRBA* gen. In het *LRBA* gen ligt ook een ander gen, te weten *MAB21L2.* Van dit gen is bekend dat het een rol speelt in de migratie van NCCs en dus betrokken is bij de ontwikkeling van het zenuwstelsel van de darm. Mogelijk zou een variant in *LRBA* een invloed kunnen hebben op *MAB21L2* en dus indirect HSCR kunnen veroorzaken. Toen we echter controles gingen analyseren op aanwezigheid van de variant bleek deze ook in de normale populatie voor te komen (4/440), wat het zeer onwaarschijnlijk maakt dat de variant de ziekte kan veroorzaken. Waarschijnlijk hebben we dus de mutatie gemist wat mogelijk komt doordat we niet alle genen in het koppelingsgebied helemaal hebben kunnen bekijken. We hebben kunnen berekenen dat we slechts 65% van alle genen hebben bekeken, dus mogelijk zit er een mutatie in het missende deel. Daarom gaan we dit experiment herhalen.

Eerdere studies lieten zien dat, ongeacht het hebben van een mutatie in het *RET* gen, bijna alle familiaire gevallen aan *RET* zijn gekoppeld. Dit blijkt echter niet alleen het geval te zijn voor de familiaire gevallen, het is ook zo voor de sporadische patiënten. Associatie studies, uitgevoerd op patiënten groepen uit Europa en Azië, lieten sterke associatie zien met één specifiek *RET* haplotype. Dit haplotype spreidt zich uit over een gebied van 27.000 bp, startende 4000 bp upstream van de *RET* startsite tot en met het einde van exon 2. Een aantal aanwezige varianten in dit gebied is gemeld of gesuggereerd als HSCR-veroorzakende varianten. Echter, niet voor alle geassocieerde varianten, in het bijzonder SNP rs2506004 (C>A) is de pathogenicitiet ook daadwerkelijk onderzocht. In **hoofdstuk 4** beschrijven we functionele studies die voor deze SNP (rs2506004, C>A) zijn uitgevoerd. Met behulp van *in silico* analyse hebben we gevonden dat het C-allel (wild type) gelokaliseerd is in een Central Midline Element (CME) sequentie element, dat een bindingsplaats is voor de eiwitcomplexen Nxf/Arnt2 en Sim2/Arnt2, respectievelijk transcriptie activatoren en repressors. Luciferase assays lieten zien dat de sequentie met het ziekte geassocieerde allel, wanneer zij gekoppeld is aan de *RET* promotor, verminderde expresse liet zien als we die vergeleken met het niet ziekte geassocieerde allel. We controleerden de bindingsaffiniteit van de Nxf/Arnt2 en Sim2/Arnt2 heterodimeren aan DNA oligos met het C-allel of het A-allel met behulp van EMSA (elektroforetische mobiliteit shift assay) en een supershift assay. We bevestigen dat Nxf/Arnt2 veel sterker

bindt aan het C-allel dan aan het A-allel, terwijl Sim2/Arnt2 een gelijke bindingscapaciteit heeft voor de A- en C-allelen. De effecten van Nxf/Arnt2 en Sim2/Arnt2 op de expressie van *RET* werden verder bevestigd door co-transfecties met constructen die Nxf/Arnt2 of Sim2/Arnt2 tot expressie brachten in een muize neuroblastoma (Neuro2A) cellijn. We controleerden de endogene RET expressie op eiwit niveau met behulp van Western blotting en konden laten zien dat Nxf/Arnt2 en Sim2/Arnt2, respectievelijk de endogene expressie van RET verhogen en lieten afnemen. We hebben hiermee aangetoond dat de SNP daadwerkelijk bij de ziekte is betrokken. Sim2 bevindt zich in het kritische gebied voor Down syndroom op chromosoom 21. 77% van de patiënten met het syndroom van Down hebben maagdarm (GI) afwijkingen, waarvan 2-15% van de gevallen HSCR hebben. Omdat Sim2 een transcriptie repressor voor *RET* is, zou overexpressie van SIM2 veroorzaakt door een extra kopie van het gen, de hoge incidentie van HSCR in Down syndroom patiënten mogelijk kunnen verklaren.

Alle gepubliceerde data geven aan dat de tot nu toe geïdentificeerde HSCR genen direct of indirect met elkaar verbonden zijn. Bovendien wijst alle data op een centrale rol voor RET in de ontwikkeling van zowel HSCR als het zenuwstelsel van de darm. Het ontrafelen van het RET geassocieerde eiwitnetwerk zou ons daarom niet alleen inzicht geven in hoe het zenuwstelsel in de darm zich ontwikkelt, maar ook zou het helpen in het identificeren van kandidaat HSCR genen. Deze kennis, en dan met name de eiwitnetwerken rond RET in de voorloper celen van het darm zenuwstelsel enteric neurale-lijst-cellen (ENCCs), ontbreekt helaas nog. Om hier inzicht in te krijgen hebben we genexpressie studies uitgevoerd op NCCS geïsoleerde uit de darm van de muis. In deze cellen hebben we RET geactiveerd of juist niet door ze te behandelen met GDNF, de activator (ligand ) voor de RET receptor (**hoofdstuk 5**). De genexpressie data van de ENCCs hebben we met elkaar vergeleken (wel of niet gestimuleerd met GDNF) en vergeleken met het expressieprofiel van de gehele (embryonale) muizendarm. De data hebben we geanalyseerd met behulp van een programma dat kijkt naar signaalroutes en gennetwerken die verrijkt zijn in de dataset waarin de onderzoeker geïnteresseerd is ('gene set enrichment analysis, GSEA') en we hebben de data geanalyseerd per gen (zogenaamde 'single gene analysis'). Uit de resultaten die gegenereerd waren met behulp van het GSEA programma bleek dat, over het algemeen, een activatie van RET resulteert in een lagere expressie van genen die coderen voor eiwitten die betrokken zijn bij celdifferentiatie, eiwitten die celdeling negatief beïnvloeden en eiwitten die belangrijk zijn voor cel-cel interacties en apoptose. De expressie gaat echter juist omhoog voor genen die coderen voor eiwitten die celdifferentiatie tegengaan en eiwitten die een rol spelen in de energie bevoorrading en biosynthese.

Een eiwitnetwerk dat ook als geheel een lagere expressie liet zien was de Toll like recep-

tor (Tlr) signaalroute, een route betrokken bij apoptose en bij de aangeboren (aspecifieke) immuunrespons. We hebben met behulp van immuno-kleuringen kunnen bevestigen dat Tlr2, Tlr4 en TLr5 inderdaad tot expressie komen in de neuronen, de glia en de zenuwcel voorlopercellen (ENCCs) in de darmwand van de muis. We konden ook laten zien dat GDNF stimulatie van RET leidt tot een lagere expressie van de eiwitten die door Tlrs worden geactiveerd, Tnfalpha, and IL-6. We laten hiermee voor het eerst zien dat RET activiteit een effect heeft op de de aangeboren immuunrespons van de ENCCs die zich in de darm bevinden. HSCR patiënten hebben vaak RET mutaties die resulteren in een verlaging van de activiteit van RET en dit zal daarom zorgen voor een hogere immuunrespons. HSCR patiënten hebben vaak last van chronische ontstekingen aan de darm (HSCR-associated enterocolitis). Onze bevindingen zouden dit mogelijk kunnen verklaren.

Concluderend kunnen we zeggen dat onze studies hebben geresulteerd in de identificatie van een HSCR gen in het 9q31 gebied, specifiek voor Chinese HSCR patiënten. Verder hebben we kunnen aantonen dat SNP rs2506004 betrokken is bij het ontstaan van HSCR en hebben we het mechanisme achter de mutatie opgehelderd. Ook hebben we ons inzicht vergroot met betrekking tot het RET eiwit netwerk in de ENCC. Onze resultaten geven meer duidelijkheid over het ontstaan van het zenuwstelsel van de darm, over de genen betrokken bij HSCR en leveren een lijst van goede kandidaatgenen voor HSCR op.

# RINGKASAN

**Bab 1** mengulas genetika, biologi perkembangan, karakteristik klinis, diagnosis dan pengobatan penyakit Hirschsprung (HSCR). HSCR adalah penyakit bawaan yang ditandai oleh ketidakadaanya ganglia dalam pleksus submukosa dan *myenteric* di saluran pencernaan. Insiden penyakit HSCR adalah 1 dalam setiap 5000 kelahiran, yang mana mayoritas merupakan kasus sporadis dan sekitar 5-15% merupakan kasus turunan. Sekitar 70% kasus merupakan kasus HSCR terisolasi sedangkan sisanya merupakan kasus sindromik, yaitu HSCR yang disertai sindrom lainnya. Sampai saat ini, 12 gen dan lima lokus (9q31, 4q31-P32, 3p21, 19q21 dan 16q23) telah teridentifikasi berasosiasi dengan perkembangan penyakit HSCR. Gen *RET* adalah faktor genetis utama penyakit HSCR dan mutasi pada gen *RET* (CDS) teramati pada sekitar 50% kasus turunan dan pada 15-35% kasus sporadis. Mutasi pada 11 gen lainnya (*EDNRB, EDN3, GDNF, NTN, SOX10, PHOX2B, ECE1, KIAA1279/KBP, ZFHX1B, NRG1 dan TTF1*) hanya terjadi pada sekitar 20% kasus, yang sebagian besar ditemukan pada kasus sindromik. Hal Ini mengarah pada hipotesis bahwa terdapat gen-gen lain yang berkontribusi kepada perkembangan penyakit HSCR.

Bolk dan kolega mengidentifikasi lokus 9q31 berasosiasi dengan HSCR pada enam keluarga. Lima diantaranya menunjukkan asosiasi dengan lokus 10q11.2, tempat gen *RET* berada. **Bab 2** memaparkan pemetaan lokus 9q31 dalam mengidentifikasi gen yang berkontribusi terhadap perkembangan penyakit HSCR. Kami memeriksa genotip 301 tag-SNP (*single nucleotide polymorphism*) yang mencakup 7 Mb (juta pasang basa) pada lokus 9q31 dari sampel DNA 137 penderita penyakit HSCR beretnis Belanda beserta DNA orang tuanya (data trios). Sebanyak 121 diantara penderita tersebut tidak memiliki mutasi pada gen *RET*, sementara 16 lainya memiliki mutasi pada gen *RET*. Kami mengidentifikasi dua SNP pada gen *SVEP1* (rs10816998 dan rs7038415) berasosiasi dengan HSCR pada pasien yang tidak memiliki mutasi gen *RET*. Akan tetapi, hasil ini tidak dapat dikonfirmasi ulang pada *cohort* pasien lain dari etnis yang sama (107 pasien etnis Belanda yang tidak memiliki mutasi *RET*). Percobaan serupa diakukan pula dengan sampel DNA yang diperoleh dari penderita penyakit HSCR beretnis Cina. Pada penelitian yang menggunakan pendekatan *case-control study* ini, teridentifikasi beberapa SNP yang terletak pada gen *IKBKAP* berasosiasi dengan HSCR. Asosiasi ini lebih kuat pada pasien dengan mutasi *RET* dibandingkan pasien tanpa mutasi *RET*. Data ini dapat dikonfirmasi ulang pada *cohort* HSCR pasien etnis Cina dan dapat disimpulkan bahwa asosiasi kromosom 9 dengan HSCR adalah populasi spesifik dan pada pasien HSCR etnis Cina terdapat kemungkinan keterlibatan gen *IKBKAP*. Data yang diperoleh menunjukkan bahwa pembawa mutasi *RET* pada etnis Cina dapat memiliki tambahan resiko menderita

penyakit HSCR jika memiliki varian gen *IKBKAP*.

Melalui metoda *Linkage analysis*, Brooks dan kolega mengidentifikasi locus 4q31-32 berasosiasi dengan HSCR pada keluarga multi-generasi etnis Belanda. **Bab 3** memaparkan penelitian guna mengidentifikasi gen pada lokus 4q31-32 yang bertanggung jawab atas perkembangan penyakit HSCR pada keluarga multi-generasi tersebut. Variasi pada exon 20 gen *LRBA* dapat diidentifikasi dengan menggunakan teknik *exome sequencing*. Pada untai DNA *anti-sense* gen *LRBA* terdapat gen *MAB21L2* yang berperanan penting pada proses migrasi ENCCs saat perkembangan ENS. Penelitian telah menunjukkan bahwa beberapa daerah gen *LRBA* adalah pengatur ekspresi gen *MAB21L2*. Hal ini menimbulkan spekulasi bahwa varian yang ditemukan pada exon 20 gen *LRBA* adalah bagian dari daerah pengatur ekspresi *MAB21L2*. Varian gen *LRBA* ini terbukti dimiliki oleh seluruh penderita HSCR pada keluarga tersebut di atas. Namun demikian, varian ini juga ditemukan pada 4 dari 440 kromosom kontrol yang diperiksa (0.9% dari populasi). Oleh karena kemungkinan varian tersebut adalah varian yang bertanggung jawab terhadap perkembangan penyakit HSCR adalah kecil. Akan tetapi, karena varian pada exon 20 gen *LRBA* adalah satu-satunya varian yang teridentifikasi pada lokus 4q31-32, maka besar kemungkinan varian penyebab HSCR yang sesungguhnya belum teridentifikasi. Hal ini bisa disebabkan karena metoda pengayaan (*enrichment method*) dan analisis *exome sequencing* yang digunakan hanya mampu menganalisis sekitar 65% exon genom manusia. Oleh karena itu, penelitian selanjutnya akan menggunakan *exome sequencing* dengan metoda pengayaan yang mencakup seluruh exon genome manusia.

Penelitian sebelumnya menunjukkan bahwa, terlepas dari status mutasi gen *RET*, hampir seluruh kasus HSCR turunan berasosiasi dengan lokus *RET*. *Association study* yang telah dilakukan oleh beberapa kelompok peneliti atas populasi Kaukasia dan Asia mengidentifikasi adanya *general RET haplotype* yang berasosiasi dengan penyakit HSCR. *Haplotype* ini mencakup 27 kb (ribu pasang basa), mulai dari 4 kb di bagian hulu gen *RET*, intron 1 dan sampai ke bagian awal ekson 2. Beberapa varian yang diidentifikasi pada *haplotype* ini telah dilaporkan dan diusulkan sebagai varian yang terlibat dalam perkembangan penyakit HSCR. Namun pengujian fungsi beberapa varian seperti halnya SNP rs2506004 (C> A) yang terletak di intron 1 gen *RET* belum pernah dilakukan. **Bab 4** memaparkan hasil uji fungsi SNP rs2506004 (C> A). Melalui analisis *in silico* teridentifikasi bahwa C-alel (*wild type*) SNP rs2506004 terletak pada elemen *Central Nervous System Midline Enhancer* (CME). CME merupakan urutan DNA tempat menempelnya Nxf/Arnt2 dan Sim2/Arnt2 yang masing-masing merupakan aktivator dan repressor transkripsi. *Luciferase assay* menunjukkan penurunan ekspresi gen *Luciferase* jika elemen *enhancer* yang memiliki A-alel (alel yang berasosiasi dengan HSCR) digabungkan dengan *promoter RET*; dibandingkan dengan elemen *enhancer* yang

memiliki C-alel atau dibandingkan dengan *promoter RET* saja (kontrol). Afinitas Nxf/Arnt2 dan Sim2/Arnt2 *heterodimer* terhadap *oligonucleiotide* pembawa C-alel atau A-alel kami uji dengan menggunakan metoda EMSA (*Electrophoretic Mobility Shift Assay*) dan *supershift assays*. Hasil pengujian menunjukkan Nxf/Amt2 memiliki afinitas yang lebih tinggi terhadap oligonukleotida pembawa C-alel, sementara Sim2/Amt2 memiliki afinitas yang tidak berbeda terhadap A-alel maupun C-alel. Efek Nxf/Arnt2 dan Sim2/Arnt2 terhadap ekspresi gen *RET* dibuktikan lebih lanjut dengan pemeriksaan level ekspresi protein *RET endogenous* pada sel Neuro2A saat ditransfeksi dengan plasmid pengkode *Nxf/Arnt2* atau *Sim2/Arnt2* melalui metoda *Western blot*. Hasilnya menunjukan bahwa Nxf/Arnt2 dan Sim2/Arnt2 masing-masing meningkatkan dan menurunkan ekspresi RET protein pada sel Neuro2A. Salah satu hal yang menarik adalah, gen *SIM2* terletak pada "*critical region of Down Syndrome* (DS)" pada kromosom 21, yang mana 77% pasien DS mengalami gangguan saluran pencernaan dan 2-15% kasus tersebut adalah HSCR. Penelitian ini telah menunjukkan bahwa SIM2 adalah repressor transkripsi *RET*. Hal ini dapat menjelaskan tingginya insiden penyakit HSCR pada pasien DS.

Sejauh ini, data menunjukan gen-gen penyebab penyakit HSCR saling terkait, baik secara langsung maupun tidak, dengan gen *RET*. Data-data tersebut mengarah pada peran sentral gen *RET* pada perkembangan sistem syaraf *enteric* (ENS) dan penyakit HSCR. Oleh karena itu, dengan mempelajari jaringan protein *RET* pada *enteric neural crest cells* (ENCCs) yang merupakan *progenitor* sel neuron dan glial pada ENS dapat membantu memahami proses perkembangan ENS dan mengidentifikasi kandidat gen yang terlibat pada perkembangan penyakit HSCR. Untuk mempelajari jaringan protein *RET* pada ENCCs kami mengisolasi RNA ENCCs mencit yang distimulasi dan tidak distimulasi dengan GDNF. Dengan menggunakan *microarray*, RNA tersebut digunakan untuk memperoleh data profil ekspresi gen (**Bab 5**). Hasil *microarra*y juga dibandingkan dengan profil ekpresi gen dari usus mencit. Hasil analisis dengan menggunakan analisis gen tunggal dan analisis jaringan protein (GSEA) menunjukkan bahwa pada sel yang diberi perlakuan GDNF terdapat kecenderungan sebagai berikut, adanya represi ekspresi gen-gen yang menstimulasi diferensiasi sel, *negative regulator* sel proliferasi dan gen yang penting untuk kontak antar sel serta *apoptosis*. Di sisi lain, tardapat peningkatan ekspresi gen-gen untuk respon awal pertumbuhan, yaitu yang terlibat dalam penghambatan diferensiasi sel, dan gen yang penting dalam pengaturan pasokan energi dan biosintesis.

Selain represi di atas, stimulasi GDNF pun merepresi alur sinyal *Toll-like receptors* (TLRs) yang berperan dalam *apoptosis* dan renspon imun bawaan. Kami menunjukkan bahwa Tlr2, Tlr4 dan Tlr5 diekspresikan pada neuron, glial dan progenitor sistem syaraf *enteric* saluran

pencernaan mencit dengan mengunakan tenik *immunostaining.* qRT-PCR mengkonfirmasi bahwa GDNF dapat menurunkan secara signifikan ekspresi gen *TLRs* dan juga beberapa *cytokines* pencetus reaksi inflamasi seperti IL-6 dan TNF$\alpha$. Untuk pertamakalinya data yang diperoleh menunjukan bahwa gen *RET* terlibat dalam pengaturan respon imun bawaan. Hal ini mengarahkan pada hipotesa bahwa inaktivasi gen *RET* (seperti halnya mutasi RET CDS pada penderita HSCR) akan mengakibatkan aktivasi reaksi imunitas bawaan *(hypersensitivitas)* yang bisa mengarah pada perkembangan penyakit *enterocolitis.*

Hasil penelitian yang telah kami lakukan mengarah pada identifikasi kandidat gen HSCR pada lokus 9q31 penderita HSCR etnis Cina. Kami juga berhasil mengungkapkan mekanisme keterlibatan SNP rs2506004 (C>A) pada pengaturan ekspresi gen *RET* sehingga dapat menjelaskan bagaimana A-alel pada SNP tersebut berkontribusi terhadap perkembangan penyakit HSCR. Selain itu, kami mempelajari jaringan protein *RET* pada ENCCs dengan lebih terperinci. Secara umum hasil penelitian kami memberikan informasi mengenai gen-gen dan jaringan protein yang terlibat pada perkembangan ENS dan bagaimana mekanisme perkembangan penyakit HSCR.

# List of Abbreviations

| | |
|---|---|
| ARMC9 | Armadillo repeat containing 9 |
| ARNT2 | Aryl-hydrocarbon receptor translocator 2 |
| ARTN | Artemin |
| BMPs | Bone morphogenetic proteins |
| bp | Base pair |
| BRCA2 | Breast cancer 2, early onset |
| CCHS | Congenital central hypoventilation |
| CDS | Coding sequences |
| CME | Central midline element |
| CNS | Central nervous system |
| Cp-Ns | Competitive non-specific |
| Cp-Sp | Competitive specific |
| CTNNAL1 | Catenin (adherin-associated protein), alpha-like 1 |
| DAPI | 4',6-diamidino-2-phenylindole |
| dbSNP | Database single nucleotide polymorphism |
| DIP | Deletion-insertion polymorphism |
| DMEM | Dulbecco's Modified Eagle Medium |
| DNA | Deoxyribonucleic acid |
| DS | Down's syndrome |
| E14.5 | Embryonic day 14.5 |
| ECE1 | Endothelin converting enzyme |
| EDN3 | Endothelin 3 |
| EDNRB | Endothelin receptor type B |
| EGF | Epidermal growth factor |
| ELISA | Enzyme-linked immunosorbent assay |
| EMEM | Eagle's minimal essential medium |
| EMSA | Electrophoretic mobility shift assay |
| ENCCs | Enteric neural crest cells |
| Enh | Enhancer |
| ENS | Enteric nervous system |
| FD | Familial dysautonomia |
| FDR | False discovery rate |
| FGF | Fibroblast growth factor |
| FGFR2 | Fibroblast growth factor receptor 2 |
| FOX | Forkhead box |
| GALNACT-2 | Chondroitin sulfate N-acetylgalactosaminyltransferase 2 |
| GAPDH | Glyceraldehyde-3-phosphate dehydrogenase |
| GATK | Genome analysis toolkit |
| GDNF | Glial cell derived neurotrophic factor |
| GFAP | Glial fibrillary acidic protein |
| GFRa1 | Glycosyl-phosphatidylinositol-anchored co-receptor of GDNF-1 |
| GI | Gastrointestinal |
| GOSH | Goldberg-Shprintzen |
| GSEA | Gene set enrichment analysis |
| GTP | Guanosine-5'- triphosphate |
| GWAS | Genome association study |
| HAEC | Hirschsprung-associated enterocolitis |
| HOX | Homebox |
| HSCR | Hirschsprung |
| IEC | Intestinal epithelial cells |
| IFNg | Interferon gamma |
| IKBKAP | Inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase complex-associated protein |
| IL-6 | Interleukin 6 |

| | |
|---|---|
| JAK | Janus tyrosine kinase |
| KBP | Kinesin binding protein |
| KBP | Kilo base |
| KEGG | Kyoto encyclopedia of genes and genomes |
| LD | Linkage disequilibrium |
| L-HSCR | Long-segment Hirschsprung |
| LOD | Logarithm of the odds |
| LPS | Lipopolysaccharide |
| LRBA | LPS-responsive vesicle trafficking, beach and achor containing |
| MAB21L2 | Mab-21(*C.elegans*)-like 2 |
| MAF | Minor allele frequency |
| MAP3K1 | Mitogen-activated protein kinase kinase kinase 1 |
| Mb | Mega base pairs |
| MCS | Multiple-species conserved sequences |
| MUSK | Muscle, skeleta, receptor tyrosine kinase |
| NCAM | Neural cell adhesion molecule |
| NCSCs | Neural crest stem cells |
| NEC | Necrotizing enterocolitis |
| NLBs | Neurosphere-like bodies |
| NPE | Nuclear protein extract |
| NRG1 | Neuregulin 1 |
| NTN | Neurturin |
| NXF (NPAS4) | Neuronal PAS domain protein 4 |
| OR | Odd ratio |
| P53 | Tumor protein 53 |
| PBS | Phosphate buffer saline |
| PCR | Polymerase chain reaction |
| PFA | Paraformaldehyde |
| PHOX2B | Paired-like homeobox 2B |
| pNEs | Paralogous, non-coding elements |
| PPARs | Peroxisome proliferators-activated receptors |
| PSPN | Persephine |
| qRT-PCR | Quantitative reverse transcription-polymerase chain reaction |
| RASGEF1A | RasGEF domain family, member 1A |
| RET | Rearrange during transfection |
| RETprom | RET promoter |
| RNA | Ribonucleic acid |
| RTK | Tyrosine kinase receptor |
| SD | Standdard deviation |
| S-HSCR | Short-segment Hirschsprung |
| SIM2 | Single-minded homolog 2 |
| SNP | Single nucleotide polymorphism |
| SOX10 | SRY (sex determining region Y)-box 10 |
| STAT | Signal transducer and activator of transcription |
| SV40 | Simian vacuolating virus 40 |
| SVEP1 | Sushi, von Willebrand factor A, EGF and pentraxin domain |
| TCA | Total colonic aganglionosis |
| TDT | Transmission disequilibrium test |
| TGFb | Transforming growth factor, beta 1 |
| TIA | Total intestinal aganglionosis |
| TLRs | TOLL-like receptors |
| TRANSFAC | Transcription factor binding predictions |
| TTF1 | Thyroid transcription factor 1 |
| WNT | Wingless-type MMTV integration site family, member 1 |
| WS4 | Shah-Waardenburg syndrome type 4 |
| YFP | Yellow fluorescent protein |
| ZFHX1B | Zinc finger homebobox protein 1B |

## Acknowledgments

This last five years have been a wonderful and interesting journey for me, I have learned many new things and met a lot of new people. Finally, I can finish what I started and feel grateful that I have had so many people helping and encouraging me to get through this PhD period. It is now time for me to express my gratitude to all the people with whom I worked together or who contributed their help and friendship during my PhD program.

First of all I would like to thank my supervisor Prof. Hofstra. Dear Robert, It was at the end of 2004 when I first heard your name from my boss Dr. Joseph Hilgers. He suggested that I go to your lab and do a research project on HNPCC, as this was the project I was working on back then in Indonesia. But it was not until the end of 2005 that I sent you an email and made an appointment to talk about an internship for my MSc degree. Instead of working on the HNPCC project, I asked you to give me another project so I could learn something different. You gave me a project on Hirschsprung (HSCR) disease, a disease that I had never heard of (despite finding out later that my older sister, whom I never met, had passed away because of this disease). Exciting project! I had a lot to learn since my background was not in genetics, but I enjoyed working on this project very much together with Greg Burzynski. This 9-month project was extended for another 5 years with me as your PhD student. You've always been supportive of all my ideas and gave me so much freedom to explore all the things I wanted to learn and do. I learned to be more independent, to trust my own judgment and to be more creative than ever. I am grateful to have met you and I sincerely thank you for all the opportunities, for your trust, your guidance and all the support you have given to me. Thank you for always being there and for being someone that I can talk to in difficult times, this really means a lot to me.

I would like to thank the members of my reading committee, Prof. Isabella Cechirinni, Prof. U.L.M. Eisel and Prof. H.W.G.M. Boddeke, for their time and remarks on my thesis.

I would like to thank the head of Genetics Department (past and present). Dear Prof. Buys, dear Charles, thank you for welcoming me for the very first time when I came to work in this department, thank you for all the good words you always had for me. Dear Prof. Wijmenga, dear Cisca, thank you for all your input and critical comments on my presentations during the Tuesday meetings and for always encouraging the PhD students to broaden their knowledge and learn about other things than only our own research subject.

I would like to thank Jackie Senior for all her time and effort in editing my thesis. Despite the fact that sometimes I wasn't on time handing in my manuscripts, you were always willing to help me. Thank you also for always giving me feedback during my presentations and help-

ing me to improve my English.

I would like to thank Intan Taufiq, Wangsa Tirya Ismaya and Neng Fisheri who helped me by editing the summary in bahasa Indonesia and also to Helga and Bote for helping me with the Dutch summary. Special thanks to Ella Elviana for the cover illustration and to Briyan Hendro for the cover and book layout.

I am grateful to all my collaborators for their contributions to the work presented in this thesis, especially Dr. Marcia Barcelo and Clara, our collaborators in Hong Kong. It was good to work with you on the chromosome 9 paper, thank you for everything and it is always nice to meet and to share experiences at HSCR consortium meetings. Dr. N. Thapar and Dr. A.J. Burn, dear Nikhil and Alan, thank you for welcoming me in your lab and allowing me to work there for five months. It was good experience and I learned a lot during my stay in London. Dr. M. Metzger, dear Marco, thank you for teaching me how to isolate ENCCs from mouse embryos and how to culture them. I enjoyed working with you, Nikhil and Alan on the RET non-coding variants and expression papers and thank you for helping with my experiments and answering questions.

Dear Rudolf, it was nice to work with you on the expression paper. Thank you for helping me analyze the microarray data. The GSEA data turned out to be better than I expected and have led to an interesting hypothesis. Thank you and good luck with your training in oncology.

Dear Wanda, it's nice to know you and it's always been good to talk to you. Thank you for trying your best to get mouse embryos for my experiments, you were really my savior! Dear Evelyn, thank you so much for teaching me the qRT-PCR, all the tips and tricks, from designing the primers to analyzing the data. Dear Lydia and Hans from the Department of Pathology, thank you for helping with the Luciferase assay and LacZ assays experiments.

Dear Alice Brooks, Morris, Pieter, Cleo and Wil, Ian Sheperd, it was great to work with you all on the exome paper and thank you for all your help. Hopefully, next time, we will have more luck and really find the "true" disease-causing variant in this linkage region. Dear Pieter and Cleo, thank you for the CLC-Bio tutorial, you've been a great help. Thanks also to Martijn and Freerk for helping with the pipeline data over the last few months.

Thank you to my Italian friends Isa, Marta and Ivana, for sharing your scientific results on RET regulating the immune response. Thank you also to all the members of the HSCR consortium, it was always great to meet you all and share our latest scientific results.

To my partners in crime in the HSCR/RET group: Greg Burzynski, Maria, Christine, Hans Verbeek, Thomas, Duco and Rajendra. It's been nice to meet you all, share experiences, share bench work and share an office, thank you for everything and for being such good col-

leagues. Dear Greg, it was nice to work with you on my MSc project, and although you are now far away, you've always been willing to help me during my PhD period. Thanks a lot Greg! Dear Maria, we started nearly everything together, and although we come from two different cultures, I am glad we get along so well together and can share our worries, dreams and going through the good and bad times of PhD life. Thank you for everything and I wish you all the best in the future. Dear Christine, life is not the same without you here, I'm less talkative! I really used to talk a lot to you and about almost everything. It doesn't matter which continent you're on, you always have time to chat with me and show concern, not only for my work but for me as your friend. Thank you also to your family who always invited me to share Christmas dinner. It was a good and warm feeling to be part of your family. Good luck with your MD/PhD program and in the future as a medical doctor, I'm sure you will become a great one. Dear Thomas, it was nice to have you here in the department even though it was only for a short period, thank you for your friendship. Dear Duco, it was great to work with you on the expression paper. Dear Rajendra thank you for helping me on the exome paper, both of you were a great help in the last part of my PhD work. I wish you both good luck with the rest of your own PhDs. Dear Amanda and Nina, my students, thank you for all your help in my research project, I wish you all the best with your future plans!

Dear Jan (Osinga), you were the first person in this department who helped me with all my experiments and taught me some lab techniques. It's always nice to work with you, a lot of laughing, sharing a lot of stories and, most of all, you were the person who persistently talked to me in Dutch. :-D It helped improve my Dutch, too bad that we are no longer office mates because my Dutch is disappearing. Thank you also for being one of my *paranimfen*. I wish you good health and happiness always.

I would like to thank to Joke Verheij for helping and provide me with all the information I needed related to HSCR patients. I thank Bart Eggen for always being willing to spend time discussing my research. It's been a great help and you always have an answer to all my questions and gave me so many ideas on how to solve the problems. Dear Conny, Peter, Jorieke, Nicole, Siobhan, thank for your comments and suggestions and sharing your experiences during the Friday meetings.

Dear Lude and Gerard, thank you for helping me with the eQTL data. It was interesting to discuss things with you both together, I felt like I was in the middle of nowhere but "in a good way", since it stimulated my brain to work hard, to the limit. :-D Dear Gerard, your two years in Indonesia meant you could more or less understand my point of view as an Indonesian living in the Netherlands and so that I could talk to you about everything related to cultures, religion and even politics. :-D However, when it comes to statistics, I almost always give

up… This is really not my 'thing' but thank you for helping me. Dear Vinod and Sebo, thank you for your suggestions and helping me check the HSCR genes network using GeneGo.

Dear Ellen and Dineke, thank you for always being willing to answer all my questions. Dear Ellen, I'm so grateful to have you here in this department, you were always enthusiastic about my work and gave me a lot of ideas on how to answer my research question. The Tuesday meetings were much more interesting and fun when you were there with all your questions and comments. Dear Dineke, thank you for being so kind and discussing in detail so many things related to my experiments, like my EMSA results. Thank you too for tricks and tips on all the cloning stuff.

Dear Mentje, Ria, Bote, Joke and Helena, our secretaries, thank you very much for all your help (how will I survive without you all? :-D). I would also like to thank Mathieu for helping me solve the problems with the *SPOED bestellingen*. Dear Yvonne, Jan Jongbloed, Krista Bos, Krista Koi, Bea, Bart Mol, Josh and all the people in the DNA diagnostic lab, thank you for always being willing to help me with DNA samples and all the things related to patient data. Dear Marina, thank you for taking care of the finances, and Edwin, thank you for solving all the computer-related problems. Dear Karen, Jantine, Michiel, Ludolf and all the people in the research lab, thank you for everything, it's been nice working with you all and sharing experiences in the lab.

Dear Helga, Renee, Bahram and Ana Ferreira, the HNPCC people, thanks for all the nice lunches, BBQs in the summer, and dinners. We always found good topics to talk and laugh about. Dear Ana, it was nice to talk to you and share stories; to my surprise we almost always had the same opinion about a lot of things.

To all my officemates and ex-officemates: Eva, Karin, Suzanne, Noortje, Mitja, Rodrigo, Omid, Celine, Olga, Senapati and Asia (or shall I call you Jan Manhattan? :D), I enjoyed sharing this room with you all, despite sometimes feeling the room was too small for all of us. :-D I sincerely apologize if sometimes I sang too loudly and disturbed you, from now on you will be free of my noise! :-D

To all my friends and colleagues in the Department of Genetics, past and present, Jianghua, Tjako, Renee, Annemieke, Mats, Gerben, Anna Posafalvi, Mahdi, Justyna, Jihane, Gosia, Agata, Isis, Paul, Barbara and Javier  thank you for all the great times and I wish you all the best.

Home is where the heart is, Groningen is like my second home, and part of the reason is because I have made such great friends here. I would like to thank Asal Hamarneh, Ramune Nedveckyte and Adriana Mantos, for their friendship and support especially for the first two years I lived in Groningen. Thank you for always being there for me. I would like

to thank all my WB housemates, my very best friend mbak Ponky, Mbak Diana, Dini, mbak Puri, Mira, mbak Ike, mas Itob, mas Wangsa, Pak Agus, Ican, Adit, Indra, Senaz, Insanu, Bu Heni and Mas Amal, how dull my life would be in Groningen without you all, thank you for the great times we spent together in Westerbadstraat, sharing stories, worries, dreams, hopes and of course delicious Indonesian cooking, these are unforgettable memories. To my dear housemates in Bezettingslaan, Neng Fisheri, Mbak Ari and Wiwin, thank you for being such wonderful housemates, we started our PhD programs almost at the same time and now we come to the end. I am grateful for having you all to share the good and bad times during these last four years. Good luck with your own theses. :-D Dear mbak Puri and mbak Ari, to have you both like I have my older sisters, felt safe and warm, no matter what problems I had, you always had a shoulder for me to cry on. Dear Neng and Mira, I often wondered how it feels to have little sisters, now I have at least two :-D, and I feel so grateful and blessed to know you. Dear Neng, thank you for being willing to be one of my *paranimfen*, it means a lot to me.

I would also like to thank the families of pak Elfami Yaman and Uni Yulia, and of Guntur and Intan for welcoming me into your home, and of Mas Eko and mbak Inne, Mas Ismail and mbak Agnes, Kang Nandang and teh Nisa. It's been nice to meet you all, share many good moments and trips to places in Europe. E-travel will not be forgotten and hopefully we will keep in contact no matter where we are. Thank you also to the family of Pak Tri and mbak Indah, Wak Aisyah and Om Meno, Mas Chalid and mbak Nona, Om Basuki and Tante Pantja, Mas Yayok and Mbak Lia, Mbak Tina, Mbak Poppy, Mas Donny and Nike, Kang Intan and the Enci, Bayu and Indres, Mas Adi and mbak Sista for the wonderful times we had here in Groningen.

Dear Dini, Dowty, Helmy and Hafidz, how grateful I am to have best friends like you. Thank you for being good listeners, good companions, you were always there during the hardest times of my stay in Netherlands. Although now almost all of us live in different countries, I hope we can keep our friendship forever.

Dear Robby, Lia Atwa, Rachma, Eryth, Iqbal, Astri, Faizah, Alya, Nizar, Irfan, Iging, Desti, Arramel and Puti how wonderful life in Groningen is because of all of you, thank you for everything, for all the wonderful gatherings we had. I wish you all the best for the future and good luck with the remaining PhD work. :-D Dear Mbak Roga, bu Mariana and Okta, sharing stories of PhD life with you made me feel like I was not alone, we went through similar good and bad times. Mbak Roga and bu Maria, good luck with your theses, and Okta, congratulations on being the first of us to reach the finish line, I am happy for you and good luck with your next journey! I would also like to thank to Teguh, kang Hendi, Mas Titah, Seno, Kadek, Wahyu, Tita, Cicin, Wisnu, Yuli, Mutia, mbak Yuni, Kenzie and all friends in deGromiest

(Indonesian Muslims' association in Groningen) and PPIG (Indonesian students' association in Groningen) for our great friendship during my stay in Groningen.

My special thanks also to Dr. Debbie Retnoningrum. Dear bu Debbie, I could not imagine how my life would have been if I hadn't met you. You are the one who introduced me to how fascinating research life is. Thank you very much for the unconditional support you've given me since I was in the BSc program up to now, the end of my PhD. Dr. Joseph Hilgers, dear Jo, my first boss, who became more like a father to me, how I wish you were here with me at the end of my PhD work, but I am sure you are watching from above and will be proud of me. Dear Jo, indeed you were right, this world is not as big as I thought. I am happy that I listened to you and dared to make one of the biggest decisions in my life, to go abroad and experience the differences. Although you will not read this, I sincerely want to thank you for all your love and support.

To my beloved family, Kepada Ibu dan bapak tercinta, tidak ada kata-kata lain yang bisa Uyung ungkapan selain terimakasih yang tidak terhingga atas semua kasih sayang dan pengorbanan Ibu dan bapak selama ini. Walaupun bapak sudah tidak bersama kita lagi, semoga beliau bisa ikut merasakan kebahagian Uyung hari ini. Semoga apa yang Uyung capai selama ini bisa membahagiakan dan membuat ibu bangga. Terimakasih untuk selalu mendukung semua keputusan yang Uyung pilih dan percaya sepenuhnya bahwa Uyung bisa menyelesaikan semua dengan baik. Terimakasih untuk kakak-kakaku, untuk teh Dewi, teh Lilis, teh Kinkin, Aa Dodo, Aa Budi atas semua dukungannya selama ini. Terimakasih untuk keluarga Wa Maman Tistajaya atas dukungan dan juga bantuannya merawat ibu selama ini. Terimakasih untuk de An-an yang sudah bersedia untuk datang dan menemani bi Uyung di hari penting ini. How I miss you all and wish that you were here with me today. Semoga Allah selalu melimpahkan kasih sayang-Nya dan selalu memberkahi kita sekeluarga.