**RIJKSUNIVERSITEIT GRONINGEN**

# Hirschsprung disease – genetics and development

**Proefschrift**

ter verkrijging van het doctoraat in de
Medische Wetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. F. Zwarts,
in het openbaar te verdedigen op
woensdag 25 oktober 2006
om 16.15 uur

door


**Grzegorz Maciej Burzyński**

geboren op 30 mei 1977
te Wrocław, Polen

Promotores: Prof. Dr. R.M.W. Hofstra

Prof. Dr. C.H.C.M. Buys

Beoordelingscommissie: Prof. Dr. G. de Haan

Prof. Dr. I. Ceccherini

Prof. Dr. A.J.M. van Oosterhout

This study was supported by the **Nederlandse Organisatie voor Wetenschappelijk Onderzoek**

The printing costs of this thesis were covered by:
**University of Groningen**
**Nederlandse Organisatie voor Wetenschappelijk Onderzoek**
**J.E. Jurriaanse Stichting**
**Dr. I.R. van de Laar Stichting**
**MRC-Holland**
**Ingeny International BV**


Printed by Cyfrus, Wrocław, Poland
Drawings by Natalia Barbara Burzyńska

## CONTENTS

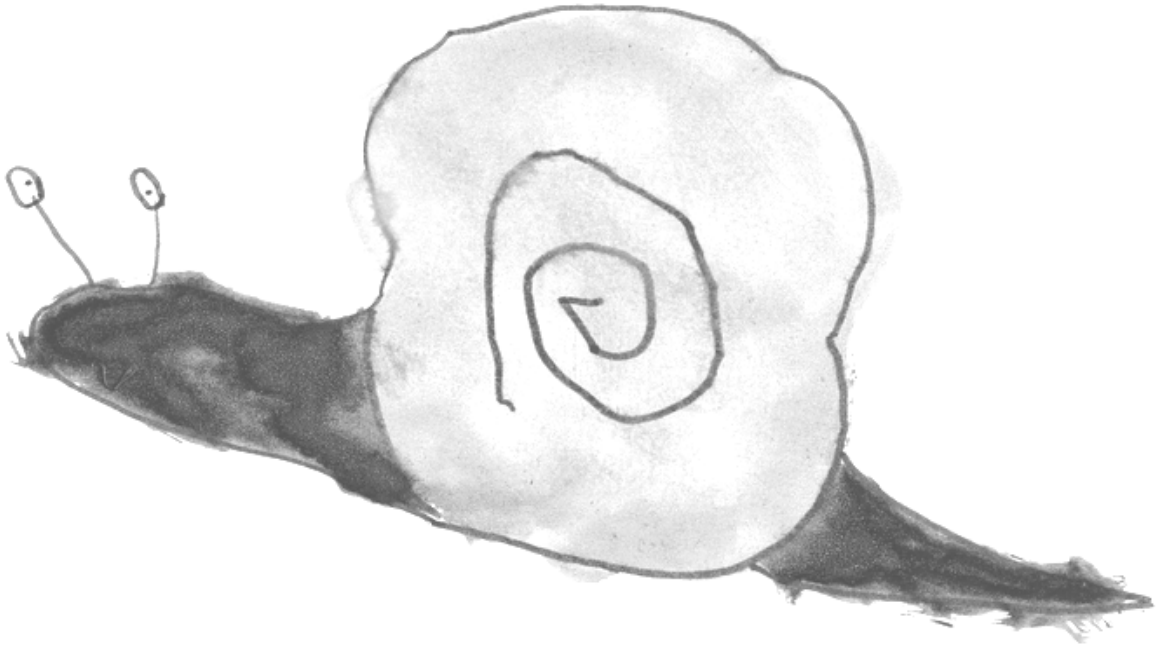
## Scope

Hirschsprung disease (HSCR) is a complex genetic disorder that is characterized by the lack of the Enteric Nervous System (ENS) in the myenteric and submucosal plexuses of the gastrointestinal tract. HSCR is a neurocristopathy and can result from aberrant migration, proliferation or differentiation of vagal neural crest cells. The length of the affected intestine classifies HSCR. The most common form of the disease (~80% of patients) is called short-segment HSCR (S-HSCR), where aganglionosis does not extend beyond the sigmoid colon. Approximately 20% of the patients is diagnosed with long-segment HSCR (L-HSCR), in which aganglionosis extents proximal to the sigmoid. The population incidence is generally assumed to be 1/5000 live births, though it varies among patients of distinct ethnic origin. A sex bias is also present in this disorder as approximately three times more males are affected. The pattern of inheritance for L-HSCR is autosomal dominant with low penetrance whereas the inheritance of the much more common S-HSCR is believed to be multifactorial. The major gene in HSCR is *RET*. Nevertheless, mutations in the coding sequence of *RET* are hardly being found in the main classes of patients.

The work included in this thesis had two main goals. Firstly, to elucidate further the role of *RET* in those cases where no coding sequence mutations were found. We focused on the most common group of patients, namely those with S-HSCR who lack a disease history in the family. We found, that the *RET* gene is involved in nearly all HSCR cases and that the majority of mutations are not lying within the coding regions, but in the regulatory sequences of the *RET* locus.

Secondly, baring in mind that HSCR is genetically complex disease, we attempted to find new genes and loci involved in HSCR development and more generally genes encoding proteins,

which contribute to ENS development. In a collaborative effort with the Rotterdam group a new HSCR susceptibility locus was identified in a multigenerational HSCR family. Further, ENS precursors cells, isolated from mouse embryos' intestines were studied and expression profiling was performed. Genes encoding proteins belonging to the RET pathway or genes encoding proteins affected by RET signaling in these cells, were identified. By this approach we hope to get a better insight in ENS development, and HSCR susceptibility.



# Hirschsprung disease – genetics and development

## Abstract

Hirschsprung disease (HSCR) is a developmental disorder characterized by the absence of enteric neurons in myenteric and submucosal plexuses in the distal parts of gastrointestinal tract. Any failure occurring on the level of migration, differentiation, proliferation or interaction with surrounding tissue can lead to a HSCR phenotype. Since the precursors of the enteric neurons are a sub-population of cells derived from the neural crest that migrate in rostro-caudal direction in the gut HSCR is considered a neurocristopathy.

The approximate incidence of HSCR is 1/5000 for the world population. In 90% of the cases HSCR occurs in a single patient in a family (simplex) and in 70% HSCR occurs as an isolated disease trait. In the remaining 30% cases HSCR is accompanied by other congenital anomalies, which in part are due to chance, can be a consequence of a chromosomal abnormality associated with HSCR or are due to a mutation(s) giving rise to a more complex (syndromic) phenotype.

HSCR is considered a genetic disease, it is caused by genetic alterations. Until now mutations in 10 genes and allele sharing with 4 loci have been identified. as a risk factors of developing HSCR. Of the genes identified *RET* proved to be the major genetic risk factor. In the common simplex cases, however, no more than 15% of patients possess *RET* coding mutations, but most patients do show haplotype sharing at the *RET* locus. Mutations in the other 9 genes are rare and mostly found in the syndromic HSCR cases. Judging from the low, sex-dependent penetrance of HSCR, its variable expression and multiple susceptibility loci discovered in linkage and association studies, isolated HSCR should be considered as an oligogenic disorder. We review

the most recent findings in HSCR genetics as well as the developmental conditions of the disorder, underlining the importance of all processes that precursors of enteric neurons need to undergo before successfully colonizing the gut. We discuss the signaling pathways involved and their spatial and temporal cross talks that enable the development of enteric innervation. This review links results from genetic studies and developmental knowledge and considers research strategies that take into account their interplay.

## 1. Genetic background

Hirschsprung disease (HSCR [MIM 142623]) is a congenital disorder characterized by intestinal obstruction due to an absence of enteric ganglia along variable lengths of the intestinal tract. The length of the aganglionic segment of the large intestine has been used to classify the disease. In short segment HSCR (S-HSCR, 80% cases) aganglionosis does not extend beyond the upper sigmoid while in long segment (L-HSCR, 20% cases), the aganglionosis extends more proximally. There are also rare cases classified as total colonic aganglionosis (TCA) and total intestinal aganglionosis (TIA), both lethal condition involving aganglionosis of the whole bowel (Passarge,et al., 1967; Holschneider et al., 1982; Garver et al., 1985; Chakravarti and Lyonnet et al., 2001).

Most common (90%) are single cases in a family (simplex cases) most of which have S-HSCR. The inheritance pattern of simplex HSCR is likely either multifactorial or recessive with low penetrance (Badner et al., 1990). L-HSCR is mainly encountered in familial cases, mostly showing an, autosomal dominant pattern of inheritance with incomplete penetrance and variation in expression. The estimated incidence is 1/5000 births, but varies in different ethnic groups, being lowest in Caucasians and highest in Asians (Torfs et al., 1998). Males are four times often affected than females, a difference most prominent in S-HSCR. Interestingly, a difference is also observed between paternal and maternal transmission of associated alleles (Holschneider et al., 1982; Gabriel 2002; Emison et al. et al., 2005). In a study by Gabriel et al. (2002), 78% of shared *RET* alleles were maternally derived. The reason for the difference in parental transmission of associated alleles is not known nor if this phenomenon could explain the sex bias.

## 1.1 Genes involved in HSCR

To date, ten genes have been implicated in HSCR (Table 1): *RET* (encoding a tyrosine kinase receptor) (Edery et al., 1994a; Romeo et al., 1994), *GDNF* – encoding the ligand of *RET* (Angrist et al., 1996; Salomon et al., 1996) and *NTN*, encoding also a possible *RET* ligand (Doray et al., 1998); genes from the endothelin pathway – *EDNRB* (Puffenberger et al., 1994), *EDN3* (Edery et al., 1996; Hofstra et al., 1996), *ECE1* (Hofstra et al., 1999), *ZFHX1B* (Cacheux et al. et al., 2001; Wakamatsu et al. et al., 2001) and *PHOX2B* (Amiel et al. et al., 2003) – two genes encoding transcription factors; and common to both the *RET* and endothelin pathways the gene encoding the transcription factor *SOX10* (Pingault et al., 1998; Paratore et al. et al., 2002). Recently, we identified homozygous nonsense mutations in *KIAA1279* (Brooks et al. et al., 2005).

With the exception of *RET*, mutations in these genes, are rare and show mostly reduced penetrance. By contrast, *RET* mutations are comparatively frequent and *RET* is therefore considered to be the major gene in HSCR etiology. For multiplex non-syndromic cases, mutation rates in RET vary in the literature between 11% and 53%, and for simplex non-syndromic cases between 9% and 35% (Angrist et al., 1993; Attie et al., 1995; Seri et al., 1997; Hofstra 2000). It is generally assumed that the mutation rates in the coding sequence of *RET* or in regions relevant for splicing are about 50% in the familial cases and 15% in simplex patients. Mutations in the other genes together do not exceed 10% and approximately half of these are mutations in the *EDNRB* receptor gene (Amiel et al., 1996; Kusafuka et al., 1996; Auricchio et al., 1996; reviewed by Verheij and Hofstra, 2004). Furthermore, most of these genes are involved in syndromic forms of HSCR, where patients show other malformations in addition to colonic aganglionosis. For example, deafness and pigmentary abnormalities as seen in Shah-Waardenburg syndrome are

caused by mutations in *EDNRB*, *EDN3* and *SOX10*. Central hypoventilation, neuroblastoma and autonomic dysfunction as seen in CCHS (also named Haddad or Ondine-Hirschsprung syndromes) are caused by mutations in *PHOX2B* and mental retardation, microcephaly, epilepsy and corpus callosum agenesis as found in Mowat-Wilson syndrome are caused by *ZFHX1B* mutations. The protein encoded by *KIAA1279* is responsible for the syndromic forms of HSCR called Goldberg-Shprintzen syndrome (Brooks et al. et al., 2005). *KIAA1279* seems to play a pivotal role in both central nervous system (CNS) and enteric nervous system (ENS) development, as almost all patients are diagnosed with microcephaly, mental retardation and polymicrogyria, all representing CNS defects, in combination with HSCR an ENS defect (Brooks et al. et al., 2005). Syndromes associated with HSCR have been extensively reviewed by Amiel and Lyonnet (2001).

**Table 1. Characteristics of genes involved in HSCR and their mouse models**

| Gene | Chrom. location | OMIM | Associated syndromes with HSCR | Enteric phenotypes of mouse models |
|---|---|---|---|---|
| ***RET*** | 10q11.2 | 164761 | - | KO - aganglionosis of small and large intestine; RET51 – colon aganglionosis |
| ***GDNF*** | 5p13 | 600837 | - | KO - aganglionosis of small and large intestine |
| ***NTN*** | 19p13 | 602018 | - | - |
| ***EDNRB*** | 13q22 | 131244 | Shah-Waardenburg, ABCD | KO - aganglionosis of distal colon; $s^l$ – recessive phenotype resembling KO |
| ***EDN3*** | 20q13 | 131242 | Shah-Waardenburg | KO - aganglionosis of distal colon; $l^s$ - recessive phenotype resembling KO |
| ***SOX10*** | 22q13 | 602229 | Shah-Waardenburg | Dom – aganglionosis of distal colon |
| ***ECE-1*** | 1p36 | 600423 | - | KO - aganglionosis of distal colon |
| ***ZFHX1B*** | 2q22 | 605802 | Mowat-Wilson | - |
| ***PHOX2B*** | 4p12 | 603851 | Ondine–Hirschsprung | KO – aganglionosis of the whole GI tract |
| ***KIAA1279*** | 10q22.1 | 609367 | Goldberg-Shprintzen | - |

$s^l$ - piebald lethal; $l^s$ – lethal spotting; *RET*51 – monoisoformic animals, lacking *RET*9 form; *Dom* – dominant megacolon; KO – knock out

## 1.2 Susceptibility loci

As mentioned, coding sequence (CDS) mutations in the discovered genes explain only a minority of HSCR cases and are largely involved in rarer classes of patients, namely in the more severe types of HSCR (L-HSCR, TCA), and in (familial) syndromes cases. In particular in the non-syndromic simplex cases, CDS mutations are not commonly found in any of the genes known to be involved in HSCR. It is therefore not surprising that several susceptibility loci have been mapped in linkage and association studies, though the genes causing this linkage still need to be identified (Bolk et al. et al., 2000; Gabriel et al. et al., 2002). In the study by Bolk et al. (2002) performed on multigenerational families, a modifier locus at 9q31 was found to segregate with multiplex HSCR. However, this locus was only associated when no obvious CDS mutation could be identified in *RET*. From the 12 families studied, 11 were linked to the *RET* locus, although 6 of these did not have a *RET* CDS mutation. It were these last 6 families that showed linkage at 9q31 and led to the conclusion that there is a modifier on chromosome 9, necessary when no strong CDS mutation is present at the *RET* locus. The lack of CDS mutations, but linkage to the *RET* locus, suggest the existence of weaker, regulatory *RET* mutations, possibly diminishing *RET* expression (Bolk et al., 2000). Other susceptibility loci were found in a genome-wide scan performed by Gabriel et al. (2002) on siblings with predominantly S-HSCR. It was shown that 65 out of 67 sibpairs showed haplotype sharing at the *RET* locus, although only 40% had *RET* CDS mutations, resembling the findings in the multigenerational families study (Bolk et al., 2000) and confirming a central role of *RET* in HSCR. In the latter study, loci conferring HSCR risk in addition to *RET* were located at 3p21 and 19q12. Allele sharing in patients was most significant ($p=2.4 \times 10^{-7}$) at 10q11 (RET). The other two loci showed comparable linkage (non-parametric lod score, Z=3.46 at 3p21 and Z=3.43 at 19q12). From the models of multigenic inheritance that were taken into consideration, a multiplicative model of

penetrance resembled best the observed population incidence. This means that all three loci are necessary and as the authors claim, sufficient for the observed occurrence of S-HSCR. Moreover, as the *RET* locus is segregating in almost all S-HSCR nuclear families, it should be considered as the major factor. The two other loci likely act as modifiers of *RET*. Additionally, when sibpairs with severe CDS *RET* mutations were excluded and the data were reanalyzed, nonrandom sharing (p=0.027) at 9q31 was confirmed. The effect of this locus seems to be more prominent in L-HSCR and in multigenerational families (Bolk et al., 2002).

In an other study by the same group (Carrasquillo et al., 2002), a genome-wide scan was performed on 43 Mennonite trios. The highest lod scores were obtained for a locus at 13q22.3-q31.1 (lod score = 55.60), where the *EDNRB* gene is located, which in this kindred represents the primary susceptibility factor. A second locus was at 10q11.21 where the *RET* gene is located (lod score = 5.60). A new locus was discovered on 16q23.3 (lod score = 3.01). No association was found for any of the susceptibility loci that were found in the aforementioned study, indicating a distinct genetic background of this Mennonite kindred. Recently, an other possible new chromosomal locus implicated in HSCR was discovered in a multigeneration Dutch family with isolated HSCR (Brooks et al., 2006) with a parametric multipoint lod score of 2.7 at 4q31-q32. Low penetrance of HSCR in this family might indicate that the locus on chromosome 4 is an important but not sufficient prerequisite for expression of the disease phenotype.

### 1.3 Regulatory mutations at the *RET* locus

As mentioned, CDS mutations in the *RET* gene are present in no more than 15% of the simplex cases. Nevertheless, it was shown in linkage studies performed on sibpairs and multigenerational HSCR families that almost all of the families were linked to the *RET* locus regardless the mutation status (Bolk 2000, Gabriel 2002). Following association studies, carried

out on several European and Asian patients' populations, mostly on simplex cases, confirmed the major role of *RET* also in the non familial cases, in HSCR development. In the majority of HSCR patients studied, association of alleles at the *RET* locus was found, despite the fact that most of these patients did not carry a CDS mutation (Fitze 2003, Sancandi 2003, Burzynski 2004, Fernandez 2005, Garcia-Barcelo 2005, Griseri 2005, Pelet 2005The aforementioned studies also tried to narrow down the region where these unknown non-coding RET mutations are located in order to eventually identify the functional variant(s). These studies showed that one frequent haplotype is associated with the disease. Regardless of the population this haplotype is present in most HSCR patients. It starts in the 5'region of the *RET* locus and spans approximately 27 Kb (4 Kb of the 5' UTR and exon 1; 23 Kb of intron 1 and exon 2). The associated haplotype has a comparable frequency in all European populations (56%-62%) (Borrego 2003, Fitze 2003, Sancandi 2003, Burzynski 2004, Pelet 2005), with the exception of Chinese population in which the frequency is higher (85%), in agreement with higher incidence of HSCR in Asians and the presence op this haplotype in controls (Garcia-Barcelo et al., 2005, Torfs et al., 1998). These findings indicate the possible presence of a mutation(s) in the *RET* promoter or in its regulatory sequences.

The associated haplotype was reconstructed with Single Nucleotide Polymorphism (SNP) marker alleles. Some of these SNP alleles were studied in more detail to determine a possible functional effect. It has been already speculated by Fitze et al. (2002) that variant rs1800858, present in exon 2, can cause the formation of a novel cryptic acceptor site, which would result in aberrant *RET* splicing. However this possibility was ruled out by Griseri et al. (2005) who did not detect any aberrant *RET* splice products in patients with different genotypes present at this SNP. Furthermore, an exon trapping experiment did not detect any splicing aberrations either (Griseri et al., 2005).

Attention was also paid to two promoter SNPs rs10900296 and rs10900297 (also named SNP-5 G>A and SNP-1 A>C, respectively), located just upstream of the *RET* transcription start site. DNA fragments containing the HSCR associated 'AC' haplotype were shown to decrease *RET* promoter activity in luciferase reporter assays performed in some of the studies (Fitze et al., 2003, Fernandez et al., 2005). Furthermore, in an elegant study by Garcia-Barcelo et al. (2005), the transcription factor *TTF-1* (Thyroid Transcription Factor 1) was shown to bind to the region where the two SNPs are located. Presence of the mutant 'AC' alleles significantly decreased the *TTF-1* binding and reduced the transcriptional activity of the *RET* promoter. It was additionally demonstrated that *TTF-1* transcription factor expression coincides with *RET* expression in the developing gut (Garcia-Barcelo et al., 2005). Nevertheless, reduced activity of the promoter, caused by SNPs rs10900296 and rs10900297 was also shown to differ significantly between two different neuroblastoma cell lines (Griseri 2005). Therefore, it is difficult to draw a final conclusion on whether these SNPs are real functional variants or whether they are in Linkage Disequilibrium (LD) with other closely located mutation(s).

Two other studies have found evidence that the HSCR associated *RET* mutations might be an enhancer mutation located approximately in the middle of intron 1 of the *RET* gene (Burzynski 2005, Emison 2005). Both studies not only employed association analyses but also applied comparative genomics, which was probably the key in finding the variants. We sequenced the haplotype region in a patient and a control homozygous for the risk and the non-risk haplotypes, to find all possible sequence variations. Eighty-four sequence differences were identified (Burzynski 2005). For these 84 identified variants we determined whether they were present in putative transcription factor binding sites, in other words whether these variants might cause the loss of such sites. Furthermore, the 5' region was also aligned with the corresponding regions from three mammalian and one avian species, in order to look for evolutionarily conserved

regions, which likely contain relevant regulatory sequences. Interestingly, only one region located in intron 1 was conserved among all the species (including avians) and only one SNP rs2506004 was present in this region. This SNP was highly associated with HSCR and caused loss of the site for the enhancer called ETV4. ETV4 belongs to the Ets family of transcription factors that have important functions in cell development, differentiation, proliferation, apoptosis and tissue remodeling. Most of its target genes are downstream of the RAS/MAP kinase signaling pathways (Oikawa et al., 2004, Davidson et al., 2004), a pathway also involved in RET signal transduction (Hayashi et al., 2000). However, as we did not perform any functional studies on this region, it still needs to be elucidated whether this SNP is a causative mutation in HSCR. In the study by Emison et al. (2005), similar but more extensive approaches were used and additionally, functional assays were employed. Approximately 350 Kb of genomic regions encompassing *RET* and two other genes located 3' to *RET* were compared in 12 vertebrate species. Additionally, 29 SNPs were genotyped in 126 trios and 12 from these were localized in the 5' *RET* UTR region and in intron 1. The highest transmission frequency was observed for one of the alleles from the marker rs2435357, located in intron 1, only 200 bp upstream of the SNP rs2506004 (Burzynski et al., 2005). Both SNPs are located in the same interval, conserved among vertebrate species. Luciferase assays performed by Emison and colleagues on amplicons containing SNP rs2435357, but also rs2506004 and rs2506005, showed reduced SV40 promoter activity when the HSCR-associated allele was present at the rs2435357 locus. It was hypothesized that the occurrence of the mutant allele interrupts binding of retinoic acid to two predicted RARE (Retinoic Acid Response Elements) elements, located within 4 nt on each side of the rs2435357 SNP, though this still needs to be proven experimentally. Retinoic acid has already been documented as a regulator of *RET* expression in cardiac in renal development (Batourina et al., 2001; Shoba et al., 2002a; 2002b) and therefore it is likely that it has also a

function in ENS development. Another, recently published study characterized regulatory elements present at the *RET* locus. This study focused on the intron 1 region, where SNPs rs2506005, rs2435357 and rs2506004 are located (Grice et al., 2005). Of 18 evolutionary conserved regions tested in luciferase assays, 7 were capable of driving transcription significantly better than the promoter alone. Five of them acted in a cell-dependent manner and were active in a mouse neuroblastoma cell lines but not in HeLa cells. As neuroblastomas are neural crest-derived tumors, RET-positive neuroblastomas are expected to be a more appropriate model in studies on HSCR than e.g. HeLa cells. In addition to the enhancing properties of the tested evolutionary conserved regions, most of them (12) also actively repressed promoter activity in HeLa cells. This indicates that these conserved regulatory elements exert a dual function of activation and repression of transcription, depending on the tissue type. Furthermore, Grice et al. (2005) constructed *LacZ* reporter mice in which *Lac*Z was expressed under the regulation of the human MCS+9.7 *RET* intron 1 region, in the context of the *hsp68* promoter. The MCS+9.7 is containing the rs2506005, rs2435357 and rs2506004 SNPs,. This evolutionary conserved region was chosen because it was able to drive luciferase expression *in vitro* at the highest levels and because it contained SNP rs2435357, indicated in the study of Emison et al. (2005) as the causative variant in HSCR. Interestingly, the MCS+9.7 region was able to drive *LacZ* expression to such an extent that it resembled the temporal and spatial expression of *RET* during embryonic development, indicating the relevance of MCS+9.7 in *RET* expression regulation. The transcription factor(s), which likely binds to the evolutionary conserved regions studied by Grice et al. (2005) remain unidentified. Furthermore, the authors were able to show specific binding of some proteins to the earlier mentioned 7 evolutionary conserved regions, though it was not conclusive as specific binding was observed to most of the designed oligonucleotides. Either these EMSA experiments were not able to reproduce the real *in vivo* interactions or indeed the

cooperation of several factors may be responsible for this regulation mechanism. In addition to the already mentioned RARE elements, located on each side of the rs2435357 SNP, the binding site for the SRF protein was predicted to overlap with the SNP locus. The role of these potential sequence sites/variants needs further investigation.

### 1.4 Properties of coding and non-coding mutations in HSCR

Although it is yet unclear which non CDS (RET) variants are functional, their existence and significance in the disease etiology is beyond doubt. Weak (regulatory) *RET* mutation(s) appear to play a major role in HSCR susceptibility, since the 5'*RET*-associated haplotype is present in the major class of patients – i.e. sporadic male cases with S-HSCR. It is this group of patients that hardly possess CDS mutations. On the contrary, CDS mutations occur more often in the rarer patients' classes, namely: in females, in familial cases and in the patients with more severe phenotypes (L-HSCR, TCA) (Emison et al., 2005; Burzynski et al., 2005). It is now becoming clear that CDS and regulatory mutations can have different properties..

In 2004 an International Hirschsprung Disease Consortium (IHDC) was founded to better characterize different mutational types present in HSCR. Groups from the US, Hong-Kong, Italy, Spain, France and Netherlands pooled their sources to increase the power of the analyses and to be able to answer questions with respect to the world-wide HSCR population. Nearly 900 HSCR patients (most of them together with their parents) were genotyped in a unified assay for 14 SNPs encompassing the *RET* locus. Furthermore, phenotypic, demographic and family data were collected from all the groups. First of all, it was reconfirmed for the patients from all HSCR population studied that the same marker alleles from the 5'*RET* locus region showed the highest transmission frequencies. The highest transmission frequency was observed for the mutant 'T' allele from SNP rs2435357 (also called the enhancer mutation). From haplotypes of all

genotyped markers as inferred from the consortium trios, it appeared that this allele was Identical By Descent (IBD) for all patients. Two common disease-associated haplotypes were present in over half of the patients and in approximately 24% of the untransmitted chromosomes, which makes this haplotype also quite common in parents. The frequency of the disease-associated haplotypes in the consortium controls corresponds to the frequency of the rs2435357 'T' allele (~24%) in the world population, estimated on the base of HapMap data (Emison et al., 2005). Both the associated haplotypes carried the 'T' allele and they differ only in the alleles from the 3'SNPs, meaning that, in fact there is only one HSCR-associated haplotype in the 5' *RET* locus region, where the enhancer mutation is located. A much higher frequency for the longer associated haplotype was observed in Chinese patients (76%). It was also the most common haplotype in the untransmitted chromosomes (42%) in the Chinese population. This corresponds with the higher incidence of HSCR in Asians. A reason for the elevated frequency of the enhancer mutation and in general the *RET* locus disease-associated haplotype in the Chinese population, is yet unknown. It was hypothesized that individuals heterozygous for the haplotype might have or had a selective advantage (Emison et al., 2005). However, neither the source of this evolutionary pressure nor whether it is still present or acted only in the past is known. What may support this hypothesis is again data on this SNP from HapMap, and the possibility to compare the distribution of this SNP over 8200 ENCODE loci, from which only approximately 0.5% show a similar or more extreme distribution pattern (Emison et al., 2005, ENCODE project et al., 2004). The SNP rs2435357 mutant allele is not equally distributed among the world population. It is hardly present in the black African population (5%), which may explain that Africans have the lowest incidence of HSCR. In Europeans the frequency of the T allele is estimated to be approximately 25%, in Asians over 40%.

This distribution illustrates a probable introduction of the mutation during the early phases of *Homo sapiens* migration from Africa. In Asians its frequency rose to 40% within a relatively short time period. Indicating the power of evolutionary selective pressure for the 5'*RET* haplotype and the regulatory mutation(s) (Emison et al., 2005).

In the IHDC study the properties of different mutational types (CDS and regulatory mutations) were compared. As apparent from previous studies, CDS mutations are enriched in specific classes of patients, but in most simplex HSCR cases no CDS mutations are found. Different patient classes can be distinguished on the base of several important factors that define HSCR risk. We mainly include here: gender - as a sex bias is present in HSCR patients; familiality – as most patients are sporadic; length of aganglionosis, characterizing the severity of the phenotype and the majority of the patients possess the milder form – S-HSCR, where only the sigmoid part of the colon is affected. Therefore the influence of different types of mutations was compared among the just mentioned categories in order to distinguish the liabilities of CDS and the enhancer mutation to determine how these mutations influence these different risk categories. Clearly in the general population, CDS mutations are much less frequent (less than 1%) than regulatory mutation(s) (~24%) and CDS mutations are generally associated with more severe and less common forms of HSCR. This is to be expected, as these mutations have clear and strong effects on the proper functioning of the RET protein. Hence, the longer the affected part of the colon, the more CDS mutations are observed, with the highest percentage in TCA forms. Predominance of CDS mutations is also visible in multiplex families, where more affected members are present in different generations. The type of HSCR, which occurs in multiplex families, is also predominantly L-HSCR.

An interesting phenomenon to come out from this study relates to the gender of the patients in relation to the type of mutations and the severity of the HSCR phenotype. A significantly higher

percentage of female patients than male patients have CDS mutations. Similarly, more female patients are being diagnosed with L-HSCR and TCA forms. By contrast, the proposed enhancer mutation(s) are more common in male patients diagnosed with mostly S-HSCR. The penetrance of the 5' *RET* haplotype is also different among the risk categories, being highest in males with S-HSCR and in simplex families. On the contrary, an inverse correlation with the recurrence risk of the disease (lower penetrance) is observed for the relatives of female patients in more sever forms of HSCR (L-HSCR and TCA) and for patients from multiplex families. This can be explained by the penetrance and frequency of CDS mutations, which are very rare and show a reduced penetrance (Chakravarti and Lyonnet 2001) still significantly higher than the penetrance of the most common regulatory variants (Emison et al., 2005). Hence, we lean towards the proposed model of a common genetic disease, caused predominantly by very low-penetrant variant(s) commonly present in the population and remaining there, probably due to the positive selection for another phenotype (Emison 2005). This common variant(s) is largely responsible for the S-HSCR forms and affects mainly males

### 1.5 Additional genetic variants

We hypothesize that in addition to the most important enhancer mutation(s) present in the 5'region of the *RET* locus, a substantial number of similar regulatory mutations present at other genomic loci is jointly causing HSCR. It is not known how many genes can participate in this process and one has to realize that the modifier genes involved probably vary significantly between different HSCR populations, depending on their genetic background. This potential variation in modifier genes may provide a possible explanation for the sex bias present in HSCR patients. Despite the genetic heterogeneity of HSCR, we assume that most of the genes involved are encoding proteins that are interconnected in their ability to modulate RET signaling.

Understanding the spatial and temporal cross talk of growth factors, receptors, second-messengers and transcription factors involved in the enteric neural crest specification, migration, proliferation and their eventual differentiation into mature neurons and glia cells of the ENS should bring us closer to a final genetic dissection of the HSCR complexity.

# 2. Developmental background

## 2.1 Embryological origin of the enteric nervous system

### *2.1.1 The vagal neural crest*

The enteric nervous system (ENS), the intrinsic innervation of the gastrointestinal tract, is comprised of neurons and glial cells that are arranged in interconnecting ganglia located between the smooth muscle layers of the gut wall. The ENS functions to control gut processes including motility reflexes, secretion and absorption, blood flow and immune and endocrine functions (Furness et al., 2005). All neurons and glial cells of the ENS are derived from the neural crest as a transient population of cells that give rise to a wide variety of cell types throughout the developing embryo (Le Douarin and Kalcheim, 1999). The precise axial origin of the neural crest-derived precursor cells that form the ENS was originally determined after establishment of the quail-chick interspecies grafting technique by Le Douarin in the late 1960s. Using this grafting procedure, Le Douarin and colleagues described the vagal region of the neural crest, adjacent to somites 1-7, as the major contributor of precursor cells to the ENS, with a minor contribution originating in the lumbosacral neural crest, caudal to the 28$^{th}$ pair of somites, which provides a smaller number of cells to the hindgut (Le Douarin and Teillet, 1973; Le Douarin and Teillet,1974) (Fig. 1). Subsequent analysis of the spatiotemporal colonisation of the gut by neural crest cells (NCSC) has been mainly performed in the chick embryo, due to it's many advantages for use as an experimental model organism. Important observations and experimental results concerning ENS development have, however, also been obtained in zebrafish, birds, mice, rats, horses and humans. Following the pioneering work of Le Douarin and colleagues, many groups have confirmed and extended their findings, using such techniques as regional ablations of the

vagal neural tube (Peters van der Sanden et al., 1993), quail-chick grafting experiments (Burns et al., 2000; Burns and Le Douarin, 1998; Burns and Le Douarin, 2001), injection of the fluorescent lineage tracer DiI (Pomeranz et al., 1991), and cell labelling by injection of a virus vector containing the marker gene *lacZ* into the somites adjacent to the vagal neural tube (Epstein et al., 1994), (reviewed in Burns et al., 2005; Newgreen et al., 2006). Determination of the neural crest contribution to the ENS in mammalian embryos has proved more difficult, mainly because of the technical problems of performing focal, long term labelling of NCSC in embryos that normally develop *in utero*. However, researchers have used markers such as DiI, in conjunction with embryo culture, to trace the fate of vagal NCSC in the mouse (Durbec et al., 1996). Such studies confirmed that the majority of the murine ENS is derived from a population of NCSC originating at the vagal level, adjacent to somites 1-5, while the foregut ENS receives a contribution from a population of cells adjacent to somites 6-7 (Durbec et al., 1996). It also appears that the vagal region of the neural crest contributes to the ENS in zebrafish (Shepherd et al., 2004) and in humans (Wallace and Burns et al., 2005), again highlighting the conservation of developmental processes in ENS formation (also see below).

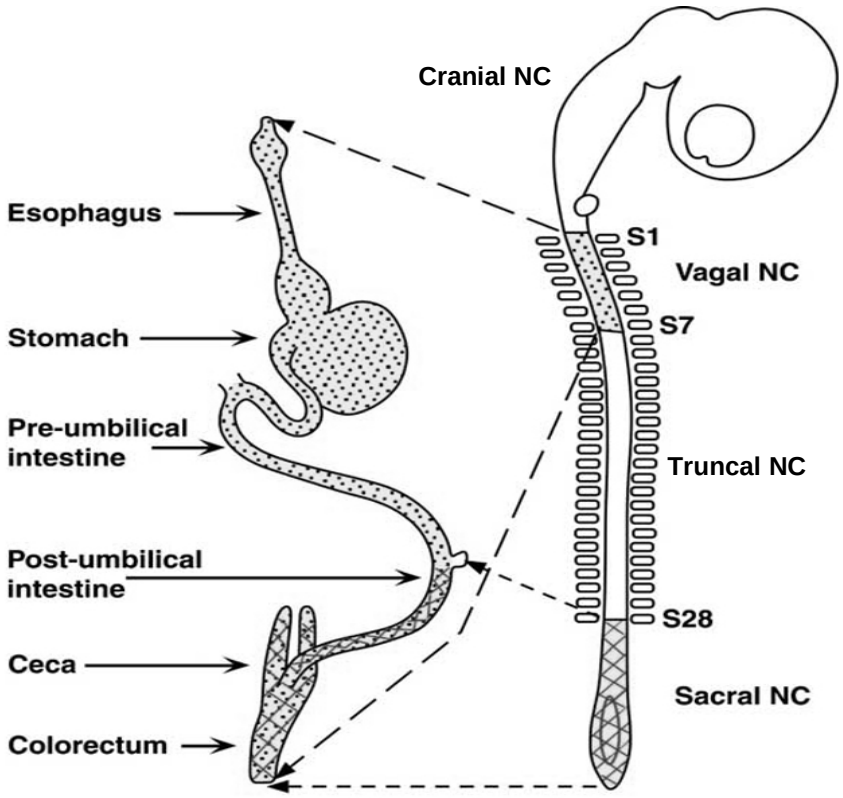



**Figure 1. Schematic diagram of the chicken neuroaxis and adjacent somites on the right and of the GI tract on the left. Vagal crest cells (somites 1-7) migrate rostro-caudally and colonize the whole intestine, while sacral neural crest cells (posterior to somite 28) migrate in the opposite direction, not further than the umbilicus; figure modified from Le Douarin and Kalcheim (1999).**

*2.1.2 The lumbosacral neural crest*

Although the initial findings of Le Douarin and colleagues suggested that the lumbosacral neural crest contributes cells to the chick hindgut (Le Douarin and Teillet, 1973; Le Douarin and Teillet, 1974), the fate of this caudal NCSC population was controversial for a number of years (see Burns and Le Douarin, 2001). Supporting evidence for a contribution of sacral NCSC to intrinsic enteric ganglia was initially obtained using injection of either DiI or replication-deficient retrovirus into the sacral region of chick embryos (Pomeranz et al., 1991), or DiI into chick and mouse embryos (Serbedzija et al., 1991). In these cell labelling studies, sacral NCSC appeared to migrate into the hindgut, early in development, before the arrival of vagal NCSC. However, the fate of labelled sacral NCSC was not confirmed in these studies. The question of sacral cell fate was subsequently addressed using quail-chick interspecies grafting to selectively label sacral

NCSC, and antibody double-labelling to identify quail cells and neuronal and glial phenotypes within the ENS. Burns and colleagues (Burns et al., 2000; Burns and Le Douarin, 1998) showed that sacral NCSC initially form the nerve of Remak (a ganglionated nerve unique to avians that extends along the mesenteric border of the hindgut and midgut), then migrate into the gut, after it has been colonised by vagal NCSC, along nerve fibres derived from the nerve of Remak. Once inside the gut, sacral cells contribute neurons and glial cells to the myenteric and submucosal plexi of the post-umbilical gut, but they are mainly restricted to the terminal region of the colon (reviewed in Burns et al., 2005; Burns and Le Douarin, 2001). Supporting evidence for a later arrival of sacral NCSC into the mouse hindgut has been obtained using *Wnt1-lacZ* transgene expression as an early marker of murine NCSC (Kapur et al., 2000). In these studies, although Wnt1-lacZ-positive sacral NCSC initially colonise pelvic ganglia in the mesenchyme surrounding the hindgut, these sacral NCSC were not found in the hindgut prior to the arrival of vagal NCSC. More recent studies using DbetaH-nlacZ mice (Anderson et al., 2006), or Wnt1-Cre/Rosa-floxed-YFP mice (Druckenbrod and Epstein et al., 2005) where NCSC are genetically marked, suggest that sacral cells colonise the hindgut during a tight developmental window after E13 but before E15.5 (Anderson et al., 2006). However, the fact that the *Wnt1-lacZ* and *DβH-nlacZ* transgenes do not discriminate between vagal and sacral NCSC, highlights the problem of fate mapping sacral NCSC in mice. Clearly, until a sacral NCSC-specific marker becomes available, the precise contribution of sacral cells to the mammalian ENS will be unresolved. Other important questions remaining to be addressed include whether sacral NCSC make specific contributions to particular subsets of enteric neurons and whether RET signalling is necessary for sacral NCSC development.

Whereas the axial origin of the various cell populations of the neural crest that give rise to the ENS in vertebrates is now well established, the precise mechanism that control the process of migration, proliferation and differentiation of enteric precursors along the whole length of the GI tracts is however much less understood. Obviously, these processes need a clear coordinated and harmonious action of growth factors, their receptors, signaling effectors and transcription factors (Taraviras and Pachnis, 1999, Young et al., 2001). Studies by numerous groups have implicated critical roles for the *WNT, NOTCH, BMP, FGF* and retinoic acid signaling pathways in the initial induction and delamination of the neural crest (La Bonne and Bronner-Fraser, 1998; Cornell and Eisen, 2005; Villanueva et al., 2002; Garcia-Castro et al., 2002). In the following section we will review some of the molecules and factors that have been implicated in ENS development, with an accent on genes and proteins previously implicated in HSCR development and on genes differentially expressed in NCSCs upon treatment with Gdnf from regions previously indicated as harboring HSCR susceptibility loci.

## 2.2 The involvement of HSCR loci in ENS development

### *2.2.1 Main signal transduction routes*

As mentioned, mutants in genes encoding members of two main signaling pathways have been found implicated in HSCR development, the *RET* signaling pathway and to a lesser extent the *EDNRB* signaling pathway. For both signaling routes pivotal roles in ENS development have also been inferred from developmental and cell biological studies. (Durbec et al., 1996; Taraviras et al., 1999; Shin et al., 1999).

### *2.2.2 Spatio-temporal induction of RET and EDNRB routes*

GDNF, the ligand of RET serves as a chemoattractant for the migrating vagal cells that express RET towards and within the gut. High levels of GDNF expression were observed in the stomach mesenchyme ahead of the migrating neural crest stem cells (NCSCs) and later in the cecum region (Natarajan et al. 2002, Young et al., 2001). In the GI tract regions posterior to the cecum, no GDNF expression was detected ahead of NCSCs cells was detected. Instead, GDNF was expressed alongside the streams of migrating cells and most likely exert proliferative and survival effects, rather than having a migratory effect (Natarajan et al., 2002).

Et-3, the ligand of EDNRB is expressed in mouse embryos at E10.0 in the midgut and hindgut and in this stage NCSCs colonize the stomach and proximal small intestine. At E11.0 *Et-3* mRNA is detected in cecum and proximal colon and remains there till the end of NCSCs migration. It is documented that ET-3 inhibits NCSC's differentiation (Hearn et al., 1998; Wu et al., 1999) and also stimulates, synergistically with GDNF, the proliferation of ENS progenitors. Interestingly, the endothelin pathway seems to have the inhibitory effect on GDNF-induced cell migration (Barlow et al., 2003).

### *2.2.3 Interactions between RET and EDNRB signaling pathways*

As said, mutations implicated in HSCR are mostly being found in genes encoding proteins belonging to RET and EDNRB receptors signaling pathways (Chakravarti and Lyonnet, 2001). Moreover, mice models with null mutations in either in *Ret*, *GFRalpha1* or *Gdnf* have complete intestinal aganglionosis (Enomoto et al., 1998, Moore et al., 1996, Schuchardt et al., 1994) and mice homozygous for null alleles of *Ednrb* or *Et-3* have aganglionosis of the distal colon (Baynash et al., 1994, Hosoda et al., 1994). These findings confirm the necessity of *RET* signaling along the whole gut and the importance of the endothelin pathway at least at later stages

of neural crest cells migration. Currently, there is growing evidence of the existence of a direct interaction between these two key signal transduction routes, on a genetic as well as on a biochemical level (McCallion et al., 2003, Carrasquillo et al., 2002, Barlow et al., 2003). Carrasquillo and others showed in a Mennonite population that specific alleles of the *RET* and *EDNRB* loci are non-randomly associated with HSCR and much more often jointly transmitted to patients than expected. To test the functional connection between the *RET* and *EDNRB* pathways, McCallion and colleagues have generated intercrossed mice from already existing strains ($Ret^{-}/Ret^{+}$, $Ednrb^{s}/Ednrb^{s}$ and $Ednrb^{s\text{-}l}/Ednrb^{s}$) which resulted in a compound heterozygous offspring for null *Ret* allele (truncating deletion at Lys-748) and loss of function *Ednrb* (*s* – piebald, $s^{l}$ – piebald lethal) alleles (McCallion et al., 2003). The authors showed that particular combination of *Ret* and *Ednrb* genotypes in mice ($Ret^{-}/Ret^{+}$; $Ednrb^{s}/Ednrb^{s}$) explain much better the HSCR transmission and phenotype observed in humans than independent mutants do, which would indicate interactions between both pathways. 100% of the male offspring with this specific genotype was affected, similarly all females manifested variable length of aganglionic colon, though 30% had reduced clinical expression of HSCR, explaining to some extent the observed sex bias in humans. Furthermore, the authors proved that the sex differences observed were dose-dependent as mice with further reduced *Ednrb* expression ($Ret^{-}/Ret^{+}$; $Ednrb^{s\text{-}l}/Ednrb^{s}$) did not present any sex differences and were equally affected. Lastly, it was demonstrated that mice carrying compound genotypes do not show a stronger effect on melanocyte or renal development compared to mice carrying isolated genotypes, indicating a dose-dependant specificity of those genes interactions in ENS development only (McCallion et al., 2003). Work of Barlow et al. (2003) brings closer the spatial and temporal regulation of neural crest migration by the RET and EDNRB routes. Heterozygous mice strains expressing the $Ret^{51}$ and the $Et\text{-}3^{ls}$ alleles were examine and intercrossed. $Ret^{51}$ allele expressing mice do not express the shorter $Ret^{9}$ isoform.

Mice expressing only the *Ret*$^{51}$ allele, manifest aganglionosis of the distal gut, resembling the HSCR phenotype and proved therefore a good model to study the relevance of *Ret* signaling in ENS development in a dose-dependent fashion (de Graaf et al., 2001). *Et-3*$^{ls}$ is a null allele of the *Ednrb* ligand and in a homozygous form it is causing distal intestinal aganglionosis (Hosoda et al., 1994). When compared to Ret expression, Et-3 was expressed ahead of the front of migrating NCSC's. Ednrb expression except the NCSCs was also detected in the mesenchyme of proximal small intestine but absent in cecum and colon. Homozygous embryos for *Ret*$^{51}$ or *Et-3*$^{ls}$ had aganglionosis of the distal colon but also demonstrated transient defects of neurons formation in the small intestine. Double homozygotes for *Ret*$^{51}$ and *Et-3*$^{ls}$ in 70% of the cases developed total intestinal aganglionosis, which indicated strong interaction between these two loci. Moreover, expression data and single homozygotes phenotypes (delay and reduced number of NCSC's in small intestine at E10-E12.5) suggested the involvement of EDNRB pathway in ENS development also in the prececal gut. Furthermore, it was shown by Barlow et al. that *ET-3*, besides the already documented inhibition of NCSCs differentiation (Hearn et al., 1998, Wu et al., 1999), also stimulates synergistically with RET, the proliferation of ENS progenitors. This effect was however visible only upon stimulation of RET signaling. Interestingly, the inhibitory effect of the endothelin pathway on *Gdnf*-induced cell migration was also revealed. It was hypothesized that it plays a role in setting up the boundaries and the pace of NCSC's migration. cAMP dependent Protein Kinase A (PKA) was pinpointed as a likely candidate that mediates interactions between the RET and the EDNRB pathways. ET-3 is thought to be able to reduce the activity of PKA, as a specific PKA inhibitor mimicked the effect of *ET3* (proliferation and inhibition of migration) on NCSCs (Barlow et al., 2003). Possibly *PKA* action is mediated via *RET* serine 696. Phosphorylation of this residue by *PKA* promotes lamellipodia formation in neuroectodermal cells, and by this means stimulates cell migration (Fukuda et al., 2002).

### *2.2.4 Transcription factors*

*SOX10* belongs to the high mobility group (HMG) gene family and for a number of neural crest derived tissues has a critical function in neural crest cell fate specification. It's function in neural crest development has been evolutionarily conserved as in mouse and zebrafish *sox10* null mutants have comparable phenotypes. Neural crest cells start to express SOX10 shortly after delamination from the dorsal neural tube (Southard-Smith et al., 1998). Loss of function mutations (in *Dom* mice) cause defects in Schwann cell and melanocyte differentiation while mice with homozygous mutants exhibit embryonic lethality with multiple defects in neural crest cell migration and differentiation including the ENS (Herbarth et al., 1998; Southard-Smith et al., 1998; Potterf et al., 2001; Sonneberg-Rietmacher et al., 2001; Dutton et al., 2001). SOX10 has been shown to promote survival of NC precursors, stimulate differentiation of these cells into melanocytes and glia cells (reviewed in Mollaaghababa and Pavan, 2003) and at the same time it maintains the multipotency of NCSCs and inhibits neuronal differentiation (Kim J et al., 2003).

The importance of this factor is also reflected in failure of expression of the neurogenic transcription factors MASH1 and PHOX2B (paired-like homeobox 2b) when *SOX10* is absent (Kim J et al., 2003, Elworthy S et al., 2005). *MASH1* and *PHOX2B* are both factors required for the NC to differentiate into autonomic neurons (Guillemot et al., 1993, Pattyn et al., 1999). In mice lacking the homeodomain transcription factor *Phox2b,* all autonomic ganglia, including ENS, fail to form properly (Pattyn et al., 1999) (Fig. 2). Heterozygous mutations in *PHOX2B* are also documented in humans and can cause HSCR, central hypoventilation syndrome (CCHS) and tumours of the sympathetic nervous system like neuroblastoma.

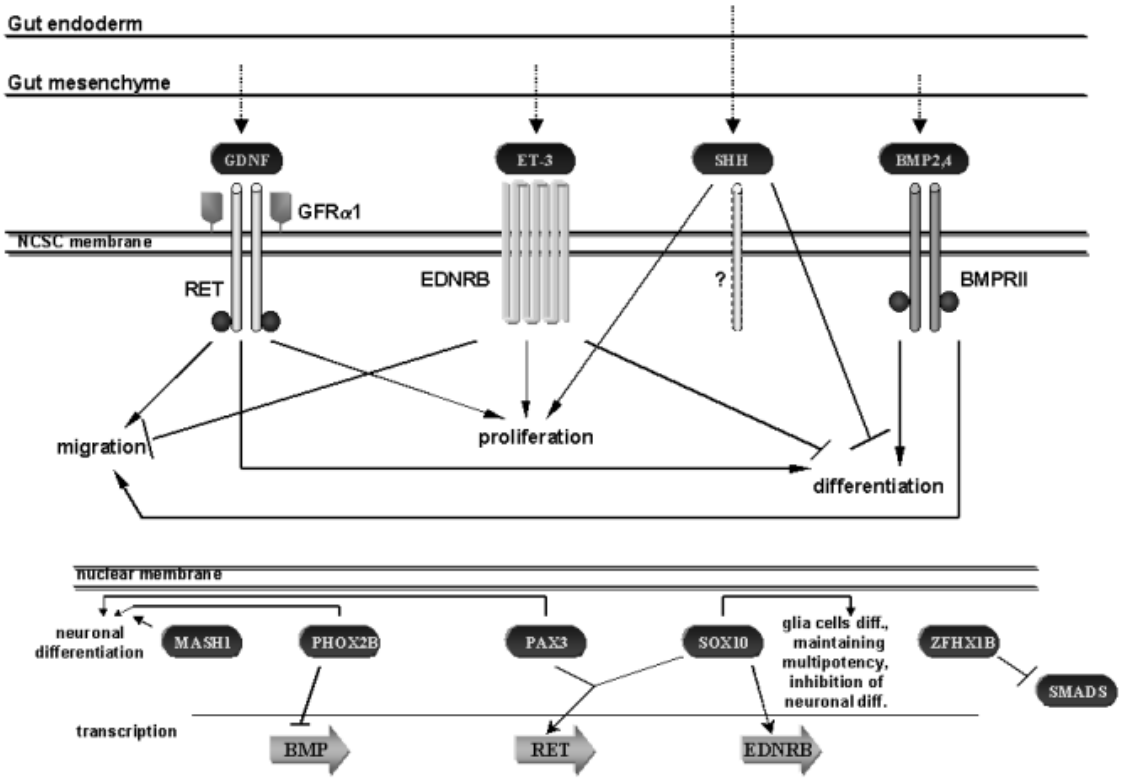



**Figure 2. Schematic representation of the most important signaling pathways acting on NCSC during ENS development; the top of the diagram shows the ligands of the receptors and their origin; GDNF, ET3 and BMPs are derived from mesenchyme whereas SHH from endoderm; receptors in the membrane are RET, EDNRB and BMPRII; the ligand of SHH is yet unknown, it is also suspected that it can act indirectly on ENS precursors; RET and BMP pathways stimulate migration and differentiation of ENS progenitors; SHH, EDNRB and RET signaling promote proliferation whereas EDNRB can inhibit the migration and together with SHH, the differentiation; In the 'nuclear compartment' transcription factors known to regulate these signaling routes are depicted; MASH1, PHOX2B and PAX3 stimulate the differentiation; PHOX2B was shown to inhibit the expression of BMP genes. PAX3 and SOX10 synergistically enhance the transcription of the RET gene, whereas SOX10 positively regulate the transcription of EDNRB gene, proliferation of NCSCs and inhibits the differentiation; ZFHXB was shown to inhibit the transcription of SMADS, involved in the regulation of BMP signaling.**

The aforementioned disorders, depending probably on the genetic background of the patients can overlap with each other with different combinations (Amiel et al., 2003, Trochet et al., 2004, 2005). Only recently, it was shown that the *RET* gene can act as a modifier locus in CCHS families as it caused the co-occurrence of the enteric phenotype in one branch of the family, whereas in other branch, neuroblastoma was diagnosed (De Pontual et al., 2006).

In addition to SOX10's functions in neural crest specification and differentiation it also plays also a significant role in ENS formation by regulating the transcription of the *EDNRB* receptor gene (Kruger et al., 2003; Zhu et al., 2004). Both SOX10 and EDNRB are necessary for enteric neuron and melanocyte formation. However while SOX10 has been shown to directly activate *EDNRB* transcription in ENS development, there appears to be no evidence for a direct interaction of the gene products in melanocyte lineage differentiation (Mollaaghababa et al., 2006). Besides stimulating *EDNRB* expression, it was also shown that it can enhance *RET* gene transcription synergistically with PAX3 (Lang et al., 2003). Thus, SOX10 has been show to influence both key signal transduction pathways known to be critical for normal ENS development and when perturbed, linked to Hirschsprung disease (Fig. 2).

*2.2.5 Other signaling pathways*

Routes that have been discussed so far are almost certainly not the only ones involved in guiding the migration and differentiation of ENS progenitors. It has been demonstrated that proteins belonging to the BMP (Bone Morphogenetic Proteins) family (BMP2, BMP4, their receptors and antagonists like noggin and gremlin) are expressed in both crest-derived and non-crest cells in the fetal gut (Chalazonitis et al., 2004) and modulate the responsiveness of NCSCs to GDNF mediated migration and influence the ganglia formation (Goldstein et al., 2005). Overexpression of *BMP* antagonists result in hypoganglionosis and smaller ganglia size. Moreover, the *GDNF* dependent migration of NCSCs is significantly reduced when *BMP4* signaling is blocked, indicating synergistic effects of these two factors on ENS progenitors migration (Goldstein et al., 2005). The involvement of *BMPs* in the ENS formation is supported by the identification of mutations in the transcription factor *ZFHX1B* in some syndromic HSCR cases, which is a downstream target of the BMP siganling (Cacheux et al., 2001 Nagaya et al.,

2002). While BMP's signaling affects migration of NCSCs, it does not influence the differentiation of these cells triggered by *GDNF* (Goldstein et al., 2005). A potential factor that may modulate this aspect of ENS formation and which is often co-expressed with BMP4 is sonic hedgehog (SHH) a member of the Hedgehog family of signaling molecules. During ENS and GI tract formation, SHH is expressed along the gut in the endoderm (Fu et al., 2005, Bitgood and McMahon., 1995, Stolow and Shi, 1995). SHH displays an opposite function to GDNF as it promotes proliferation of NCSCs, but inhibits differentiation of these cells (Fu et al., 2005). Interconnection of GDNF/ET-3/BMP/SHH pathways still needs clarification but it is very likely that endoderm and mesenchyme derived factors act together in the control of NCSCs migration, proliferation and differentiation of ENS precursors in the intestine (Fig. 2).

### *2.2.6 RET mediated gene expression of ENS precursors*

Signalling of the RET receptor triggers one of the most important signal transduction route, crucial for the migration, proliferation and differentiation of the neural crest stem cells (NCSC's), which are responsible for the formation of the enteric nervous system (ENS). To identify genes involved in/or triggered by the RET pathway, we isolated NCSC's from E.12-E.14 mouse embryo intestines and cultured them in selection media with or without GDNF, the ligand of RET and we performed gene expression profiling. Several, important signaling pathways for ENS were found to be significantly, differentially regulated by Gdnf stimulation in the NCSCs isolated from mouse embryo intestines (Appendix 4). Genes belonging to the Notch and Wnt routes were found to be down-regulated upon Gdnf stimulation. Furthermore, Gdnf seemed to negatively regulate routes promoting differentiation of the ENS progenitors like the retinoic acid (RA) pathway and semaphorins signaling. Conversely, genes and pathways that have been previously shown to stimulate proliferation and/or migration of NCSC's were found to be up-

regulated in cells treated with Gdnf. Among them the FGF pathway, forkhead transcription factors and the BMP signaling route. From the previously identified genes in HSCR, only the Ret co-receptor Gfrα1 was found to be down-regulated by Gdnf. Furthermore, three genes were found – *Svep1, Klf4* and *Arhgef3*, which have human homologues that map to the previously described susceptibility loci on 3p21 and 9q31, which has been found in linkage studies performed on HSCR patients (Tab. 1) (Bolk et al., 2000; Gabriel et al., 2002). *Svep1* and *Klf4* are genes from the chromosome 9 region that were significantly down and up-regulated in cells treated with Gdnf, respectively. Svep1 is a newly described cell surface CAM protein, probably involved in initial cell adhesion process (Shur et al., 2006). Klf4 (Kruppel like factor 4) is a transcription factor, highly expressed in the (Zhao et al., 2004). *Arhgef3*, which was down-regulated, encodes Rho guanine nucleotide exchange factor 3 and may form a complex with G proteins and stimulate Rho-dependent signaling resulting in migration, adhesion or axon growth (Bloch-Gallego et al., 2005). These genes due to their localization and Ret related expression, might be considered as a good candidates genes in Hirschsprung disease.

## 3. Conclusions

From the overwhelming genetic developmental and biochemical data, it is clear that RET plays a central role in the development of ENS and HSCR. One of the main regulators of RET function is the G-protein coupled *EDNRB* receptor transduction route, influencing the proliferation and migration of NCSCs (Shin et al., 1999; Barlow et al., 2003). While these two signaling pathways clearly play a central role, there is increasing evidence of involvement of the BMP and Hedgehog pathways in modulating many of the aspects of ENS NCSCs development (Goldstein et al., 2005; Fu et al., 2005). It is only a question of time before it will appear that other signaling pathways are implicated in the development of the ENS

It has been shown that RET and EDNRB are interconnected and harmoniously guide the migration, proliferation and differentiation of neural crest stem cells. As RET seems to be the central factor, the other routes likely act as modifiers of RET signaling. Therefore, we postulate that in prospective studies *RET*-dependent expression profiling, preferably in the neural crest stem cells should be the starting point. In parallel, association studies should be performed using dense marker scans in already discovered HSCR susceptibility regions or genome wide, in conjunction with comparative genomics and functional approaches with reporter constructs and knock out animal models.

## 4. References


Amiel J, Attie T, Jan D, Pelet A, Edery P, Bidaud C, Lacombe D, Tam P, Simeoni J, Flori E, Nihoul-Fekete C, Munnich A, Lyonnet S.
Heterozygous endothelin receptor B (EDNRB) mutations in isolated Hirschsprung disease. Hum Mol Genet 1996 Mar;5(3):355-7

Amiel J, Laudier B, Attie-Bitach T, Trang H, de Pontual L, Gener B, Trochet D, Etchevers H, Ray P, Simonneau M, Vekemans M, Munnich A, Gaultier C, Lyonnet S. Polyalanine expansion and frameshift mutations of the paired-like homeobox gene PHOX2B in congenital central hypoventilation syndrome.
Nat Genet 2003;33(4):459-61.

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

Anderson RB, Stewart AL, Young HM.
Phenotypes of neural-crest-derived cells in vagal and sacral pathways.
Cell Tissue Res 2006 Jan;323(1):11-25.

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

Angrist M, Kauffman E, Slaugenhaupt SA, Matise TC, Puffenberger EG, Washington SS, Lipson A, Cass DT, Reyna T, Weeks DE, et al.
A gene for Hirschsprung disease (megacolon) in the pericentromeric region of human chromosome10.
Nat Genet 1993 Aug;4(4):351-6.

Attie T, Pelet A, Edery P, Eng C, Mulligan LM, Amiel J, Boutrand L, Beldjord C, Nihoul-Fekete C, Munnich A, et al.
Diversity of RET proto-oncogene mutations in familial and sporadic Hirschsprung disease. Hum Mol Genet 1995 Aug;4(8):1381-6

Auricchio A, Casari G, Staiano A, Ballabio A.
Endothelin-B receptor mutations in patients with isolated Hirschsprung disease from a non-inbred population.
Hum Mol Genet 1996 Mar;5(3):351-4.

Badner J, Sieber W, Garver K, Chakravarti A (1990) A genetic study of Hirschsprung disease. Am J Hum Genet 46:568-580

Barembaum M, Bronner-Fraser M.
Early steps in neural crest specification.
Semin Cell Dev Biol 2005 Dec;16(6):642-6

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 Dec 4;40(5):905-16

Batourina E, Tsai S, Lambert S, Sprenkle P, Viana R, Dutta S, Hensle T, Wang F, Niederreither K, McMahon AP, Carroll TJ, Mendelsohn CL.
Apoptosis induced by vitamin A signaling is crucial for connecting the ureters to the bladder. Nat Genet 2005 Oct;37(10):1082-9.

Baynash AG, Hosoda K, Giaid A, Richardson JA, Emoto N, Hammer RE, Yanagisawa M.
Interaction of endothelin-3 with endothelin-B receptor is essential for development of epidermal melanocytes and enteric neurons.
Cell 1994 Dec 30;79(7):1277-85.

Bitgood MJ, McMahon AP.
Hedgehog and Bmp genes are coexpressed at many diverse sites of cell-cell interaction in the mouse embryo.
Dev Biol 1995 Nov;172(1):126-38.

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;97:268-73.

Borrego S, Wright F, Fernandez R, Williams N, Lopez-Alonso M, Davuluri R, Antinolo G,Eng C
A founding locus within the RET proto-oncogene may account for a large proportion of apparently sporadic Hirschsprung disease and a subset of cases of sporadic medullary thyroid carcinoma.
Am J Hum Genet 2003 72:88-100

Brooks AS, Bertoli-Avella AM, Burzynski GM, Breedveld GJ, Osinga J, Boven LG, Hurst JA, Mancini GM, Lequin MH, de Coo RF, Matera I, de Graaff E, Meijers C, Willems PJ, Tibboel D, Oostra BA, Hofstra RM.
Homozygous nonsense mutations in KIAA1279 are associated with malformations of the central and enteric nervous systems.
Am J Hum Genet 2005 Jul;77(1):120-6.

Brooks AS, Oostra BA, Hofstra RM.
Studying the genetics of Hirschsprung's disease: unraveling an oligogenic disorder. Clin Genet 2005 Jan;67(1):6-14

Burns, A. J.
Migration of neural crest-derived enteric nervous system precursor cells to and within the gastrointestinal tract. Int J Dev Biol 2005 49, 143-50.

Burns, A. J., and Le Douarin, N. M.
Enteric nervous system development: analysis of the selective developmental potentialities of vagal and sacral neural crest cells using quail-chick chimeras.
Anat Rec 2001 262, 16-28.

Burns AJ, Champeval D, Le Douarin NM.
Sacral neural crest cells colonise aganglionic hindgut in vivo but fail to compensate for lack of enteric ganglia.
Dev Biol 2000 Mar 1;219(1):30-43

Burns AJ, Douarin NM.
The sacral neural crest contributes neurons and glia to the post-umbilical gut: spatiotemporal analysis of the development of the enteric nervous system.
Development 1998 Nov;125(21):4335-47

Burzynski GM, Nolte IM, Osinga J, Ceccherini I, Twigt B, Maas S, Brooks A, Verheij J, Plaza Menacho I, Buys CHCM, Hofstra RMW
Localizing a putative mutation as the major contributor to the development of sporadic Hirschsprung disease to the RET genomic sequence between the promoter region and exon 2. Eur J Hum Genet 2004 12:604-612

Cacheux V, Dastot-Le Moal F, Kaariainen H, Bondurand N, Rintala R, Boissier B, Wilson M, Mowat D, Goossens M.
Loss-of-function mutations in SIP1 Smad interacting protein 1 result in a syndromic Hirschsprung disease.
Hum Mol Genet 2001 Jul 1;10(14):1503-10.

Carrasquilo M, McCallion A, Puffenberg E, Kashuk C, Nouri N, Chakravarti A.
Genome-wide association study and mouse model identify interaction between RET and EDNRB pathways in Hirschsprung disease.
Nat Genet 2002;32:237-244.

Chakravarti, A and Lyonnet S.
Hirschsprung's disease. In The Metabolic and Molecular Bases of Inherited Diseases (ed. C. R. Scriver, A. L.
Beaudet, D. Valle, W. S. Sly, B. Childs, K. Kinzler and B. Vogelstein),
2001 pp. 6231-6255. New York: McGraw-Hill.

Chalazonitis A, D'Autreaux F, Guha U, Pham TD, Faure C, Chen JJ, Roman D, Kan L, Rothman TP, Kessler JA, Gershon MD.
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 Apr 28;24(17):4266-82.

Cornell RA, Eisen JS.
Notch in the pathway: the roles of Notch signaling in neural crest development. Semin Cell Dev Biol 2005 Dec;16(6):663-72

Creuzet S, Couly G, Le Douarin NM.
Patterning the neural crest derivatives during development of the vertebrate head: insights from avian studies.
J Anat 2005 Nov;207(5):447-59

Davidson B, Goldberg I, Tell L, Vigdorchik S, Baekelandt M, Berner A, Kristensen GB, Reich R, Kopolovic J.
The clinical role of the PEA3 transcription factor in ovarian and breast carcinoma in effusions.
Clin Exp Metastasis 2004;21(3):191-9

Doray B, Salomon R, Amiel J, Pelet A, Touraine R, Billaud M, Attie T, Bachy B, Munnich A, Lyonnet S.
Mutation of the RET ligand, neurturin, supports multigenic inheritance in Hirschsprung disease.
Hum Mol Genet 1998;7:1449-1452.

Druckenbrod, N. R., and Epstein, M. L.
The pattern of neural crest advance in the cecum and colon.
Dev Biol 2005 287, 125-33.

Durbec, P. L., Larsson-Blomberg, L. B., Schuchardt, A., Costantini, F., and Pachnis, V.
Common origin and developmental dependence on c-ret of subsets of enteric and sympathetic neuroblasts.
Development 1996 122, 349-58.

Durbec P, Marcos-Gutierrez CV, Kilkenny C, Grigoriou M, Wartiowaara K, Suvanto P, Smith D, Ponder B, Costantini F, Saarma M, et al.
GDNF signalling through the Ret receptor tyrosine kinase.
Nature 1996 Jun 27;381(6585):789-93

Dutton K, Pauliny A, Lopes SS, Elworthy S, Carney TJ, Rauch J, Geisler R, Haffter P, Kelsh RN.
Zebrafish colourless encodes sox10 and specifies non-ectomesenchymal neural crest fates.
Development. 2001 Nov;128(21):4113-25.

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.

Edery P, Lyonnet S, Mulligan LM, Pelet A, Dow E, Abel L, Holder S, Nihoul-Fekete C, Ponder B, Munnich A.
Mutations of the RET proto-oncogene in Hirschsprung's disease.
Nature 1994a Jan 27;367(6461):378-80

Edery P, Pelet A, Mulligan LM, Abel L, Attie T, Dow E, Bonneau D, David A, Flintoff W, Jan D, et al. Long segment and short segment familial Hirschsprung's disease: variable clinical expression at the RET locus.
J Med Genet 1994b Aug;31(8):602-6

Elworthy S, Pinto JP, Pettifer A, Cancela ML, Kelsh RN.
Phox2b function in the enteric nervous system is conserved in zebrafish and is sox10-dependent.
Mech Dev 2005 May;122(5):659-69

Emison ES, McCallion AS, Kashuk CS, Bush RT, Grice E, Lin S, Portnoy ME, Cutler DJ, Green ED, Chakravarti A.
A common sex-dependent mutation in a RET enhancer underlies Hirschsprung disease risk.
Nature 2005 Apr 14;434(7035):857-63.

ENCODE Project Consortium.
The ENCODE (ENCyclopedia Of DNA Elements) Project.
Science 2004 Oct 22;306(5696):636-40

Enomoto H, Araki T, Jackman A, Heuckeroth RO, Snider WD, Johnson EM Jr, Milbrandt J.
GFR alpha1-deficient mice have deficits in the enteric nervous system and kidneys.
Neuron 1998 Aug;21(2):317-24.

Epstein, M. L., Mikawa, T., Brown, A. M., and McFarlin, D. R.
Mapping the origin of the avian enteric nervous system with a retroviral marker.
Dev Dyn 1994 201, 236-44.

Fernandez RM, Boru G, Pecina A, Jones K, Lopez-Alonso M, Antinolo G, Borrego S, Eng C.
Ancestral RET haplotype associated with Hirschsprung's disease shows linkage disequilibrium breakpoint at -1249.
J Med Genet 2005 Apr;42(4):322-7.

Fitze G, Appelt H, Konig I, Gorgens H, Stein U, Walther W, Gossen M, Schreiber M, Ziegler A, Roesner D, Schackert HK
Functional haplotypes of the RET proto-oncogene promoter are associated with Hirschsprung disease (HSCR).
Hum Mol Genet 2003 12:3207-3214

Fitze G, Cramer J, Ziegler A, Schierz M, Schreiber M, Kuhlisch E, Roesner D, Schackert HK.
Association between c135G/A genotype and RET proto-oncogene germline mutations and phenotype of Hirschsprung's disease.
Lancet 2002 Apr 6;359(9313):1200-5.

Fu M, Lui VC, Sham MH, Pachnis V, Tam PK.
Sonic hedgehog regulates the proliferation, differentiation, and migration of enteric neural crest cells in gut.
J Cell Biol 2004 Aug 30;166(5):673-84.

Fukuda T, Kiuchi K, Takahashi M.
Novel mechanism of regulation of Rac activity and lamellipodia formation by RET tyrosine kinase.
J Biol Chem 2002 May 24;277(21):19114-21.

Furness, J. B.
"The Enteric Nervous System." 2005, Blackwell Publishing, Oxford.

Gabriel SB, Salomon R, Pelet A, Angrist M, Amiel J, Fornage M, Attie-Bitach T, Olson JM, Hofstra R, Buys C, Steffann J, Munnich A, Lyonnet S, Chakravarti A
Segregation at three loci explains familial and population risk in Hirschsprung disease.
Nat Genet 2002 31:89-93.

Garcia-Barcelo M, Ganster RW, Lui VC, Leon TY, So MT, Lau AM, Fu M, Sham MH, Knight J, Zannini MS, Sham PC, Tam PK.
TTF-1 and RET promoter SNPs: regulation of RET transcription in Hirschsprung's disease. Hum Mol Genet 2005 Jan 15;14(2):191-204.

Garcia-Castro MI, Marcelle C, Bronner-Fraser M.
Ectodermal Wnt function as a neural crest inducer.
Science 2002 Aug 2;297(5582):848-51.

Garver KL, Law JC, Garver B.
Hirschsprung disease: a genetic study.
Clin Genet 1985 Dec;28(6):503-8.

Gershon, M. D., Kirchgessner, A. L. and Wade, P. R.
Functional anatomy of the enteric nervous system. In Physiology of the Gastrointestinal Tract, Vol. 1 (ed. L. R.Johnson), 1994 pp. 381-422. New York: Raven Press.

Goldstein AM, Brewer KC, Doyle AM, Nagy N, Roberts DJ.
BMP signaling is necessary for neural crest cell migration and ganglion formation in the enteric nervous system.
Mech Dev 2005 Jun;122(6):821-33.

de Graaff, E., Srinivas, S., Kilkenny, C., D'Agati, V., Mankoo, B. S., Costantini, F. and Pachnis, V. Differential activities of the RET tyrosine kinase receptor isoforms during mammalian embryogenesis.
Genes Dev 2001 15,2433 -2444.

Graham A, Begbie J, McGonnell I.
Significance of the cranial neural crest.
Dev Dyn 2004 Jan;229(1):5-13

Grice EA, Rochelle ES, Green ED, Chakravarti A, McCallion AS.
Evaluation of the RET regulatory landscape reveals the biological relevance of a HSCR-implicated enhancer.
Hum Mol Genet 2005 Dec 15;14(24):3837-45

Griseri P, Bachetti T, Puppo F, Lantieri F, Ravazzolo R, Devoto M, Ceccherini I.
A common haplotype at the 5’ end of the RET proto-oncogene underlies genetic susceptibility to HSCR development through reduced gene expression. Hum Mut 2005 25:189-195

Guillemot F, Joyner AL.
Dynamic expression of the murine Achaete-Scute homologue Mash-1 in the developing nervous system.
Mech Dev 1993 Aug;42(3):171-85

Hayashi, H., Ichihara, M., Iwashita, T., Murakami, H., Shimono, Y., Kawai, K., Kurokawa, K., Murakumo, Y., Imai, T., Funahashi, H., Nakao, A. and Takahashi, M.
Characterization of intracellular signals via tyrosine 1062 in RET activated by glial cell line-derived neurotrophic factor.
Oncogene 2000 19,4469 -4475

Hearn CJ, Murphy M, Newgreen D.
GDNF and ET-3 differentially modulate the numbers of avian enteric neural crest cells and enteric neurons in vitro.
Dev Biol 1998 May 1;197(1):93-105

Herbarth B, Pingault V, Bondurand N, Kuhlbrodt K, Hermans-Borgmeyer I, Puliti A, Lemort N, Goossens M, Wegner M.
Mutation of the Sry-related Sox10 gene in Dominant megacolon, a mouse model for human Hirschsprung disease.
Proc Natl Acad Sci U S A 1998 Apr 28;95(9):5161-5.

Hofstra RM, Osinga J, Tan-Sindhunata G, Wu Y, Kamsteeg EJ, Stulp RP, van Ravenswaaij-Arts C, Majoor-Krakauer D, Angrist M, Chakravarti A, Meijers C, Buys CH.
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.

Hofstra RM, Valdenaire O, Arch E, Osinga J, Kroes H, Loffler BM, Hamosh A, Meijers C, Buys CH. 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.

Hofstra RM, Wu Y, Stulp RP, Elfferich P, Osinga J, Maas SM, Siderius L, Brooks AS, vd Ende JJ, Heydendael VM, Severijnen RS, Bax KM, Meijers C, Buys CH.
RET and GDNF gene scanning in Hirschsprung patients using two dual denaturing gel systems.
Hum Mutat 2000;15(5):418-29.

Holschneider AM
Hirschsprung's congenital megacolon. The concept of physiopathology and therapy
Med Welt 1982 Feb 12;33(6):210-3

Hosoda K, Hammer RE, Richardson JA, Baynash AG, Cheung JC, Giaid A, Yanagisawa M.
Targeted and natural (piebald-lethal) mutations of endothelin-B receptor gene produce megacolon associated with spotted coat color in mice.
Cell 1994 Dec 30;79(7):1267-76.
Kapur RP

Colonization of the murine hindgut by sacral crest-derived neural precursors: experimental support for an evolutionarily conserved model.
Dev Biol 2000 Nov 1;227(1):146-55

Kapur RP, Yost C, Palmiter RD.
A transgenic model for studying development of the enteric nervous system in normal and aganglionic mice.
Development 1992 Sep;116(1):167-75

Kim J, Lo L, Dormand E, Anderson DJ.
SOX10 maintains multipotency and inhibits neuronal differentiation of neural crest stem cells. Neuron 2003 Apr 10;38(1):17-31

Kruger GM, Mosher JT, Tsai YH, Yeager KJ, Iwashita T, Gariepy CE, Morrison SJ.
Temporally distinct requirements for endothelin receptor B in the generation and migration of gut neural crest stem cells.
Neuron 2003 Dec 4;40(5):917-29.

Kusafuka T, Wang Y, Puri P.
Novel mutations of the endothelin-B receptor gene in isolated patients with Hirschsprung's disease.
Hum Mol Genet 1996 Mar;5(3):347-9.

LaBonne C, Bronner-Fraser M.
Induction and patterning of the neural crest, a stem cell-like precursor population.
J Neurobiol 1998 Aug;36(2):175-89

Lang D, Epstein JA.
Sox10 and Pax3 physically interact to mediate activation of a conserved c-RET enhancer. Hum Mol Genet 2003 Apr 15;12(8):937-45

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 Aug;30(1):31-48

Le Douarin, N. M., and Teillet, M. A.
Experimental analysis of the migration and differentiation of neuroblasts of the autonomic nervous system and of neurectodermal mesenchymal derivatives, using a biological cell marking technique.
Dev Biol 1974 41, 162-84.

Le Douarin, N. M., and Kalcheim, C.
"The Neural Crest." 1999, Cambridge University Press, Cambridge.

McCallion AS, Stames E, Conlon RA, Chakravarti A.
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 Feb 18;100(4):1826-31

Mollaaghababa RM, Hou L, Baxter LL, Loftus SK, Southard-Smith EM, Incao A, Cheng J, Pavan WJ.
Genetic evidence does not support direct regulation of EDNRB by SOX10 in migratory neural crest and the melanocyte lineage.
Mech Dev 2006 Jan 10

Mollaaghababa R, Pavan WJ.
The importance of having your SOX on: role of SOX10 in the development of neural crest-derived melanocytes and glia.
Oncogene 2003 May 19;22(20):3024-34

Moore, M. W., Klein, R. D., Farinas, I., Sauer, H., Armanini, M., Phillips, H., Reichardt, L. F., Ryan, A. M., Carver-Moore, K. and Rosenthal, A.
Renal and neuronal abnormalities in mice lacking GDNF.
Nature 1996 382, 76-79

Natarajan D, Marcos-Gutierrez C, Pachnis V, de Graaff E.
Requirement of signalling by receptor tyrosine kinase RET for the directed migration of enteric nervous system progenitor cells during mammalian embryogenesis.
Development 2002 Nov;129(22):5151-60.

Nagaya M, Kato J, Niimi N, Tanaka S, Wakamatsu N.
Clinical features of a form of Hirschsprung's disease caused by a novel genetic abnormality.
J Pediatr Surg 2002 Aug;37(8):1117-22

Newgreen, D., Young, H. M., and Burns, A. J.
The Development of the enteric nervous system in relation to Hirschsprung's disease. In "Embryos, Genes and Birth Defects" (P. Ferretti, A. J. Copp, C. Tickle, and G. Moore, Eds.), pp. 263-300. 2006, John Wiley and Sons, Chichester.

Oikawa T.
ETS transcription factors: possible targets for cancer therapy.
Cancer Sci 2004 95:626-633

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

Pattyn A, Morin X, Cremer H, Goridis C, Brunet JF.
The homeobox gene Phox2b is essential for the development of autonomic neural crest derivatives.
Nature 1999 May 27;399(6734):366-70.

Passarge E
The genetics of Hirschsprung's disease. Evidence for heterogeneous etiology and a study of sixty-three families.
N Engl J Med 1967 Jan 19;276(3):138-43.

Pelet A, de Pontual L, Clement-Ziza M, Salomon R, Mugnier C, Matsuda F, Lathrop M, Munnich A, Feingold J, Lyonnet S, Abel L, Amiel J.
Homozygosity for a frequent and weakly penetrant predisposing allele at the RET locus in sporadic Hirschsprung disease.
J Med Genet 2005 Mar;42(3):e18

Peters-van der Sanden, M. J., Kirby, M. L., Gittenberger-de Groot, A., Tibboel, D., Mulder, M. P., and Meijers, C.
Ablation of various regions within the avian vagal neural crest has differential effects on ganglion formation in the fore-, mid- and hindgut.
Dev Dyn 1993 196, 183-94.

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

Pomeranz, H. D., Rothman, T. P., and Gershon, M. D.
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, 647-55.

de Pontual L, Pelet A, Trochet D, Jaubert F, Espinosa-Parrilla Y, Munnich A, Brunet JF, Goridis C, Feingold J, Lyonnet S, Amiel J.
Mutations of the ret gene in isolated and syndromic Hirschsprung disease in human disclose major and modifier alleles at a single locus.
J Med Genet 2006 Jan 27

Potterf SB, Mollaaghababa R, Hou L, Southard-Smith EM, Hornyak TJ, Arnheiter H, Pavan WJ.
Analysis of SOX10 function in neural crest-derived melanocyte development: SOX10-dependent transcriptional control of dopachrome tautomerase.
Dev Biol 2001 Sep 15;237(2):245-57

Puffenberger E, Hosoda K, Washington S, Nakao K, de Wit D, Yanagisawa M, Chakravart A. A missense mutation of the endothelin-B receptor gene in multigenic Hirschsprung's disease. Cell 1994;79:1257-1266.

Romeo G, Ronchetto P, Luo Y, Barone V, Seri M, Ceccherini I, Pasini B, Bocciardi R, Lerone M, Kaariainen H, et al. Point mutations affecting the tyrosine kinase domain of the RET proto-oncogene in Hirschsprung's disease.
Nature 1994 Jan 27;367(6461):377-8.

Sancandi M, Griseri P, Pesce B, Patrone G, Puppo F, Lerone M, Martucciello G, Romeo G, Ravazzolo R, Devoto M, Ceccherini I
Single nucleotide polymorphic alleles in the 5' region of the RET proto-oncogene define a risk haplotype in HSCR.
J Med Genet 2003 40:714-718

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;14:345-7.

Schuchardt, A., D'Agati, V., Larsson-Blomberg, L., Costantini, F. and Pachnis, V. Defects in the kidney and enteric nervous system of mice lacking the tyrosine kinase receptor Ret.
Nature 1994 367,380 -383

Serbedzija, G. N., Burgan, S., Fraser, S. E., and Bronner-Fraser, M.
Vital dye labelling demonstrates a sacral neural crest contribution to the enteric nervous system of chick and mouse embryos.
Development 1991 111, 857-66.

Seri M, Yin L, Barone V, Bolino A, Celli I, Bocciardi R, Pasini B, Ceccherini I, Lerone M, Kristoffersson U, Larsson LT, Casasa JM, Cass DT, Abramowicz MJ, Vanderwinden JM, Kravcenkiene I, Baric I, Silengo M, Martucciello G, Romeo G.
Frequency of RET mutations in long- and short-segment Hirschsprung disease.
Hum Mutat 1997;9(3):243-9.

Shepherd, I. T., Pietsch, J., Elworthy, S., Kelsh, R. N., and Raible, D. W.
Roles for GFRalpha1 receptors in zebrafish enteric nervous system development.
Development (2004 131, 241-9.

Shin MK, Levorse JM, Ingram RS, Tilghman SM.
The temporal requirement for endothelin receptor-B signalling during neural crest development.
Nature 1999 Dec 2;402(6761):496-501

Shoba T, Dheen ST, Tay SS.
Retinoic acid influences the expression of the neuronal regulatory genes Mash-1 and c-ret in the developing rat heart.
Neurosci Lett 2002 Feb 1;318(3):129-32.

Shoba T, Dheen ST, Tay SS.
Retinoic acid influences Phox2 expression of cardiac ganglionic cells in the developing rat heart.
Neurosci Lett 2002 Mar 15;321(1-2):41-4

Sonnenberg-Riethmacher E, Miehe M, Stolt CC, Goerich DE, Wegner M, Riethmacher D.
Development and degeneration of dorsal root ganglia in the absence of the HMG-domain transcription factor Sox10.
Mech Dev 2001 Dec;109(2):253-65.

Southard-Smith EM, Kos L, Pavan WJ.
Sox10 mutation disrupts neural crest development in Dom Hirschsprung mouse model. Nat Genet 1998 Jan;18(1):60-4

Stoller JZ, Epstein JA.
Cardiac neural crest.
Semin Cell Dev Biol 2005 Dec;16(6):704-15

Stolow MA, Shi YB.
Xenopus sonic hedgehog as a potential morphogen during embryogenesis and thyroid hormone-dependent metamorphosis.
Nucleic Acids Res 1995 Jul 11;23(13):2555-62

Taraviras S, Marcos-Gutierrez CV, Durbec P, Jani H, Grigoriou M, Sukumaran M, Wang LC, Hynes M, Raisman G, Pachnis V.
Signalling by the RET receptor tyrosine kinase and its role in the development of the mammalian enteric nervous system.
Development 1999 Jun;126(12):2785-97

Taraviras S, Pachnis V.
Development of the mammalian enteric nervous system.
Curr Opin Genet Dev 1999 Jun;9(3):321-7

Torfs CP, Christianson RE.
Anomalies in Down syndrome individuals in a large population-based registry.
Am J Med Genet 1998 Jun 5;77(5):431-8.

Trochet D, Bourdeaut F, Janoueix-Lerosey I, Deville A, de Pontual L, Schleiermacher G, Coze C, Philip N, Frebourg T, Munnich A, Lyonnet S, Delattre O, Amiel J.
Germline mutations of the paired-like homeobox 2B (PHOX2B) gene in neuroblastoma. Am J Hum Genet 2004 Apr;74(4):761-4

Trochet D, O'Brien LM, Gozal D, Trang H, Nordenskjold A, Laudier B, Svensson PJ, Uhrig S, Cole T, Niemann S, Munnich A, Gaultier C, Lyonnet S, Amiel J.
PHOX2B genotype allows for prediction of tumor risk in congenital central hypoventilation syndrome.
Am J Hum Genet 2005 Mar;76(3):421-6

Verheij J and Hofstra RM: Shah-Waardenburg' in "Molecular basis of inborn errors of development" edited by CJ Epstein, RP Erickson and A Wynshaw-Boris. 2004 Pp. 440-446

Villanueva S, Glavic A, Ruiz P, Mayor R.
Posteriorization by FGF, Wnt, and retinoic acid is required for neural crest induction. Dev Biol 2002 Jan 15;241(2):289-301

Wakamatsu N, Yamada Y, Yamada K, Ono T, Nomura N, Taniguchi H, Kitoh H, Mutoh N, Yamanaka T, Mushiake K, Kato K, Sonta S, Nagaya M.

Mutations in SIP1, encoding Smad interacting protein-1, cause a form of Hirschsprung disease.
Nat Genet 2001;27:369-70.

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 Mar;319(3):367-82

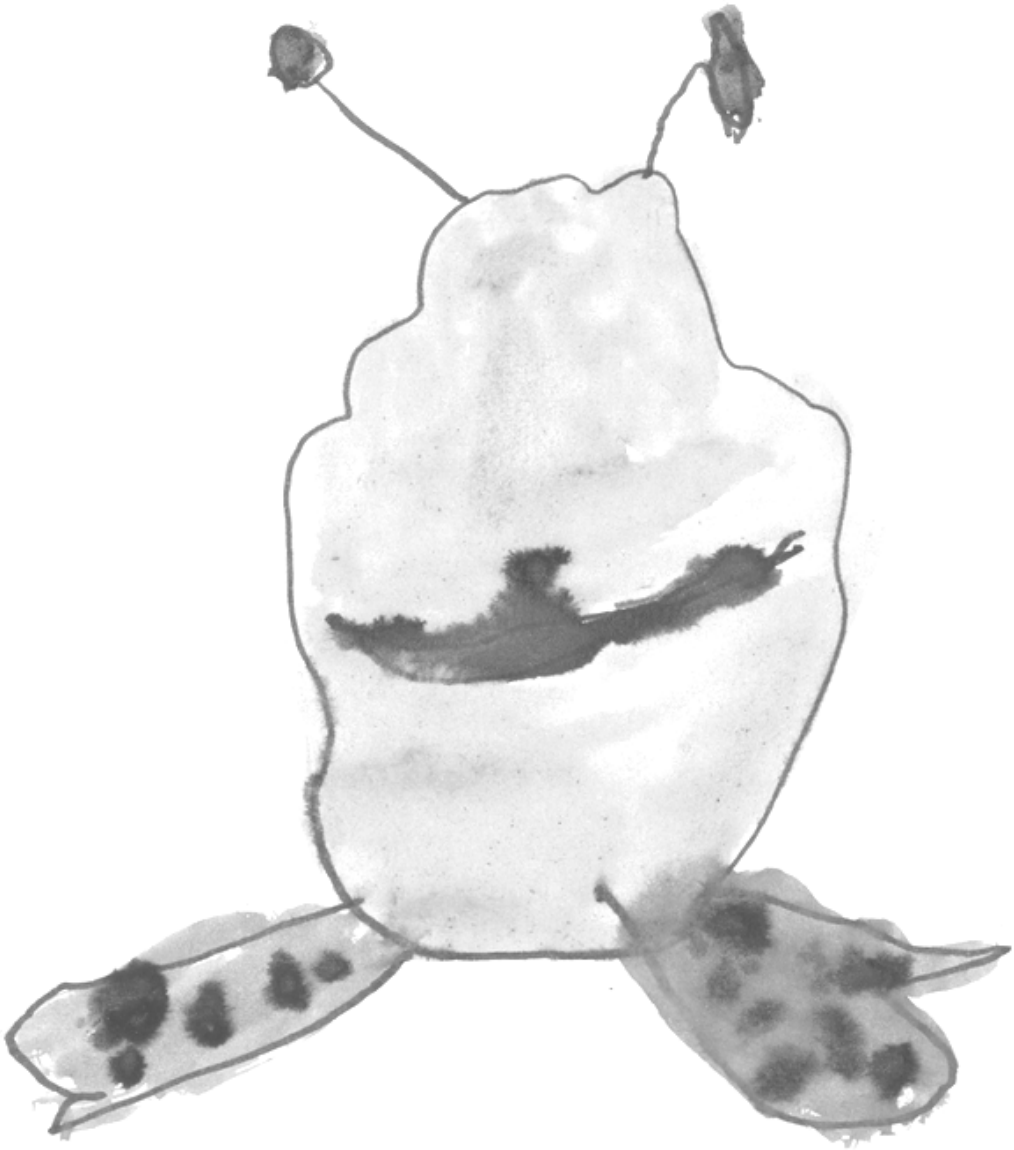
Wu JJ, Chen JX, Rothman TP, Gershon MD.
Inhibition of in vitro enteric neuronal development by endothelin-3: mediation by endothelin B receptors.
Development 1999 Mar;126(6):1161-73.

Yntema CL, Hammond WS.
The origin of intrinsic ganglia of trunk viscera from vagal neural crest in the chick embryo.
J Comp Neurol 1954 Oct;101(2):515-41

Young HM, Hearn CJ, Ciampoli D, Southwell BR, Brunet JF, Newgreen DF.
A single rostrocaudal colonization of the rodent intestine by enteric neuron precursors is revealed by the expression of Phox2b, Ret, and p75 and by explants grown under the kidney capsule or in organ culture.
Dev Biol 1998 Oct 1;202(1):67-84.

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

Zhu L, Lee HO, Jordan CS, Cantrell VA, Southard-Smith EM, Shin MK.
Spatiotemporal regulation of endothelin receptor-B by SOX10 in neural crest-derived enteric neuron precursors.
Nat Genet 2004 Jul;36(7):732-7



# Appendix 1

## Localizing 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

Grzegorz M. Burzynski[1], Ilja M. Nolte[2], Jan Osinga[1], Isabella Ceccherini[3], Bas Twigt[1], Saskia Maas [4], Alice Brooks[5], Joke Verheij [6], Ivan Plaza Menacho[1], Charles H.C.M Buys[1] & Robert M.W. Hofstra[1]

1. Department of Genetics, University Medical Center Groningen, Groningen, The Netherlands

2. Department of Medical Biology, University Medical Center Groningen, Groningen, The Netherlands

3. Laboratorio di Genetica Molecolare, Instituto G. Gaslini, Genova, Italy

4. Department of Clinical Genetics, Academic Medical Centre, Amsterdam, The Netherlands

5. Department of Clinical Genetics, Erasmus Medical Center Rotterdam, The Netherlands

***European Journal of Human Genetics* 2004 Aug;12(8):604-12.**

## Abstract

Hirschsprung disease (HSCR), a congenital disorder characterised by intestinal obstruction due to absence of enteric ganglia along variable lengths of the intestinal tract, occurs both in familial and sporadic cases. *RET* mutations have been found in approximately 50% of the families, but explains only a minority of sporadic cases. This study aims at investigating a possible role of *RET* in sporadic HSCR patients. Haplotypes of 13 DNA markers, within and flanking *RET* have been determined for 117 sporadic HSCR patients and their parents. Strong association was observed for six markers in the 5' region of *RET*. The largest distortions in allele transmission were found at the same markers. One single haplotype composed of these six markers was present in 55.6% of patients vs. 16.2% of controls. Odds ratios revealed a highly increased risk of homozygotes for this haplotype to develop HSCR (OR>20). These results allowed us to conclude that *RET* plays a crucial role in HSCR even when no gene mutations are found. An unknown functional disease variant(s) with a dosage-dependent effect in HSCR is likely located between the promoter region and exon 2 of *RET*.

## Introduction

Hirschsprung disease (HSCR [MIM 142623]) is characterized by the absence of intrinsic ganglion cells in the myenteric and submucosal plexuses of the gastrointestinal tract resulting in a sometimes life threatening intestinal obstruction. HSCR is a disorder with a complex pattern of inheritance. Differences in sex ratio with a male predominance (3:1 to 5:1), incomplete penetrance, variable expression, and association with a large number of syndromes have been observed.[1-3] Genetic analyses of the disorder made clear that HSCR is a heterogenic disease.[3] To date, mutations in nine genes have been found that contribute to the disease: *RET*,[4,5] *GDNF*,[6,7] *NTN*,[8] *EDNRB*,[9] *EDN3*,[10,11] *ECE1*[12] *SOX10*,[13,14] *SIP1*[15,16] and *PHOX2B*.[17] Of all of these genes, *RET* [MIM 164761] is thought to be the major gene, since almost all familial cases are linked to 10q11.2 the region in which the *RET* gene is located.[18,19] However, mutations in *RET* are detected in about 50% of all familial cases.[3,20] and fail to explain 70-90% of the more commonly observed sporadic HSCR cases .[3,20] All other genes do explain no more than about 5-10% of the HSCR cases.[21]

In this study, we aim at investigating a possible role of *RET* in HSCR patients for whom no mutation had been found in the *RET* coding sequence. Therefore, we typed nine SNPs and four microsatellite markers covering the entire genomic region of *RET* (Figure 1) in a group of 117 Dutch sporadic HSCR patients and their 231 parents. Of these, 64 were selected on being investigated for mutations in *RET* and the fact that no mutations could be identified. In 53 patients, *RET* was not screened. We performed single and multilocus association and TDT analysis in order to determine in our patients a possible inherited ancestral haplotype coupled to a putative ancestral disease mutation(s). The two different subgroups of patients, i.e. those screened and found negative for *RET* mutations and those not screened for mutations, were also analyzed separately to see if combining these patients was a valid procedure.

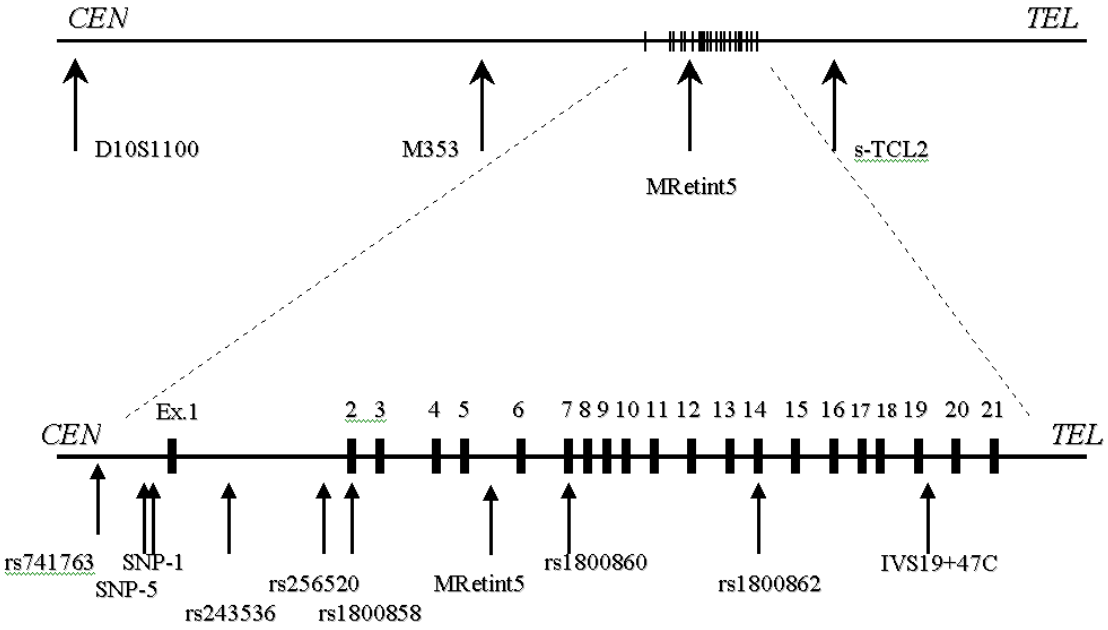



**Figure 1: Schematic representation of the *RET* gene and flanking regions. In the upper part of the figure the positions of the microsatellite markers used in this study are given. In the lower part the *RET* gene structure is magnified and approximate positions of typed SNP's are shown.**

## Materials and Methods

*Studied patient population*

DNA samples for this study were obtained from sporadic HSCR patients and their parents with written informed consent. These patients reside throughout the Netherlands and had been referred to clinical geneticists. Almost all our families (n=113) consist of two parents and the affected child (trios). Of three patients, only one parent was available for the study and one patient participated with his child and spouse. Of these, 64 were screened by means of DGGE and sequence analysis[20] and found negative for *RET* mutations. These 64 were selected from an initial group of 75 sporadic HSCR patients from whom 11 proved to carry a *RET* mutation. Fifty-three did not undergo the screening. As controls, we used the non-transmitted haplotypes of unaffected parents, like does the Haplotype Relative Risk method.[22,23] The two non-transmitted haplotypes

within each trio form pseudo-control genotypes. Under the hypothesis of random mating in our study population, these pseudo-control genotypes are expected to be similar to genotypes of population controls. Therefore, we are allowed to use these genotypes to provide estimates of the number of heterozygotes and homozygotes and their odds ratios in the population.

*Genotyping*

In all patients and their relatives, we genotyped four microsatellite markers and nine single nucleotide polymorphisms (SNPs) in a 400 kb region flanking and containing the *RET* locus (Figure 1). These markers included two SNPs located in the promoter region of the gene (SNP-5; G>A, 202 bp upstream of the start codon, and SNP-1; C>A, 198 upstream of the start codon),[24] and SNP rs741763; G>C located 4 kb upstream of the *RET* gene transcription start site. Two SNPs were located in intron 1 (IVS+6000A>C /rs2435362 and IVS1-126G>T /rs2565206 (http://www.ncbi.nlm.nih.gov/SNP/) and the following SNPs were from exon 2 (c135G>A / A45A, rs1800858)[25], exon 7 (c1296G>A/rs1800860)[25] exon 14 (c2508C>T /S836S, rs1800862)[25] and intron 19 (IVS19+47C>T)[26]. Of the four microsatellite markers, two were located upstream of *RET*, namely *D10S1100*[27] (300 kb upstream) and *M353*[28] (80 kb upstream), one located within intron 5, *MRETint5*[28],and one is located 38 kb downstream the *RET* gene, *s-TCL2*[28] (Figure 1). Genotyping of the microsatellites markers (*D10S1100*, *M353*, *MRETint5*, *s-TCL2*) was performed on the ABI 377 DNA sequencer machine (PE Biosystems). Data were analyzed by using Applied Biosystem software. The SNPs: rs2565206, rs1800858, rs1800860, rs1800862 and IVS19+47C>T were genotyped by restriction enzymes digestion. Additionally, two SNPs in *RET* promoter region, SNP-1 and SNP-5 were typed using the pyrosequencing method (Isogen Bioscience, Maarssen, The Netherlands). Table 1 contains characteristics of all genotyped markers.

**Table 1. Characteristics of the typed markers around and within the RET locus.**

| Name | Position with reference to RET gene and possible alleles | Genomic position according to NCBI (bp) | Heterozygo-sity | Primer sequences | Restriction enzyme used |
|---|---|---|---|---|---|
| **D10S1100** | -300 kb to ATG, CA repeat | 43 047 427 – 43 047 448 | 0.668 | F: 5'-TTGGGCATACTAGGCAAACAGA-3' R: 5'-ACAAGTCAGAGATGGCTTTACC-3' | - |
| **M353** | -80 kb to ATG, CA repeat | 43 267 787 – 43 267 811 | 0.591 | F: 5'-TTTAAGACTGGAACCTCTCC-3' R: 5'-CTTCCAAAATCAATACAGGC-3' | - |
| **rs741763** | -4 kb to ATG G>C | 43 343 542 | 0.490 | F: 5'-GCCTTCCCATGCTACTG-3' R: 5'-CTTCCCTCGCCCTCCTTTCT-3' | AluI |
| **SNP-5** | -202 bp to ATG G>A | 43 347 703 | 0.726 | Sequenced by Isogen Bioscience BV Maarssen, The Netherlands | - |
| **SNP-1** | -198 bp to ATG C>A | 43 347 707 | 0.342 | Sequenced by Isogen Bioscience BV Maarssen, The Netherlands | - |
| **rs2435362** | IVSI+6000A>C | 43 353 975 | 0.488 | F: 5'-*CACATTGCTCTGCCATCGTA*-3' R: 5'-*GGCATTCCTGTCACCAAGTA*-3' | PstI |
| **rs2565206** | IVSI-126G>T | 43 370 977 | 0.379 | F: 5'-*ATGACTTTCCTGTAAAGTGC*-3' R: 5'-*GGAGTTTTTCATCTCTGTTC*-3' | NlaIII |
| **rs1800858** | A45A, c135G>A | 43 371 164 | 0.438 | F: 5'-AGCCTTATTCTCACCATCCC-3' R: 5'-CAGTGTCAGCGGCTGTGATA-3' | EagI |
| **MRetint5** | IVSV+930, CA repeat | 43 378 204 – 43 378 245 | 0.345 | F: 5'-CTGTAACTCCACCACCTTGG-3' R: 5'-GCTCTAGTGTGTCCTTCCCA-3' | - |
| **rs1800860** | A432A c1296G>A | 43 381 883 | 0.780 | F: 5'-GCCATTACAGGCCGGTCCAGC-3' R: 5'-GAGGCCCAGGCTCCAGAAGC-3' | BsmI |
| **rs1800862** | S836S, c2508C>T | 43 390 290 | 0.27 | F: 5'-GAAGACCCAAGCTGCCTGAC-3' R: 5'-GGGCTGGGTGCAGAGCCATATG-3' | AluI |
| **rs2075912** | IVS19+47C>T | 43 397 413 | 0.117 | F: 5'-TGGAGTGACCGGCCATCTCT-3' R: 5'-AAGCATCACAGAGAGGAAGG-3' | MnlI |
| **s-TCL2** | +38 kb downstream RET, CA repeat | 43 439 311 – 43 439 352 | 0.309 | F: 5'-CCAGACTCTCAAAGACCAGG-3' R: 5'-CATAATACTGGCCTATAG-3' | - |

Heterozygosity of the markers is calculated in the control genotypes, which are the pseudo-genotypes comprised of the alleles not transmitted from the parents to their affected child.

*Statistical analyses*

As a quality control on the scoring of the genotypes, all markers were tested for Hardy-Weinberg equilibrium (HWE) before analysis. For this test, we only used DNA specimens from control individuals. If resulting p-values were smaller than 0.05, this was regarded as a sign of low quality of the genotyping data and the corresponding marker was discarded from further analyses.

Linkage disequilibrium (LD) was evaluated in order to investigate the identity-by-descent status of similar haplotypes. If LD is strong, the probability that two haplotypes are not coincidental but identical by descent is high. The strength of LD was measured separately for case and the control chromosomes using a randomization test, which permutes alleles over the haplotypes, determining the significance of D'.

For single locus allelic association analyses, the frequencies of alleles and genotypes, respectively, were compared between patients and controls using a chi-square test. When this test showed a significant result, a Z-test was performed to determine which allele was responsible for the difference. This test assumes that patients and controls are independent samples from the population and that the numbers of specific allele genotypes follow a binomial distribution that can be approximated by a normal distribution. Under the assumption of no assortative mating in the population, transmitted and non-transmitted haplotypes meet these criteria.

Transmission distortion of each allele versus all other alleles together is tested and combined in a multi-allelic transmission/distortion test (TDT; Spielman and Ewens, 1996). This test evaluates whether, among heterozygous parents, one or more alleles are transmitted to their affected child more often than the 50% expectation.

The program TDTPHASE from the software package UNPHASED is used for the estimation of haplotype transmissions and determining the corresponding haplotype risks as a result of over-transmission[29]

Odds ratios (ORs) and 95% confidence intervals (CIs) were calculated without correction for external variables such as age at diagnosis or type of the disease.[30]

All p-values and CIs were corrected for multiple testing for independent tests, i.e. for 13 markers and two groups of patients (all patients and those with unknown or without a *RET* mutation), using a Bonferroni correction. Because of LD between markers in the region, this correction is overstated and the results are, therefore, conservative.

## Results

*Hardy-Weinberg equilibrium (HWE) analysis*

None of the markers showed a significant deviation from HWE in the sample of non-transmitted pseudo-genotypes. Hence, further analyses were performed on all 13 marker loci.

*Linkage disequilibrium*

Strong LD is observed for all markers pairs both for the sample of case chromosomes and for the sample of control chromosomes (data not shown).

*Association analyses*

Table 2 shows that six markers: rs741763, SNP-5, SNP-1, rs2435362, rs2565206 and rs1800858 were strongly associated with the disease, revealing significant differences in frequencies of particular alleles in patients versus controls: for rs741763, the C allele is present in 81.5% of the patients chromosomes and 58.3% of the control chromosomes ($p=2.0\cdot10^{-7}$) and SNP-5, the A allele is present in 66.7% of the patients chromosomes versus 22.5% of the control chromosomes ($p=6.1\cdot10^{-24}$); for SNP-1 the C allele was present in 81.2% of the patients chromosomes versus 58.0% in the control chromosomes ($p=1.1\cdot10^{-6}$); for rs2435362 the A allele was seen in 67.6% of

the patients chromosomes versus 26.2% of the control chromosomes ($p=1.4\cdot10^{-22}$); for rs2565206 allele T was found on 84.7% of the patients chromosomes versus 67.1% of the control chromosomes ($p=1.2\cdot10^{-4}$) and for rs1800858 the A allele was observed on 67.1% of the patients chromosomes versus 23.1% of the control chromosomes ($p=2.2\cdot10^{-24}$).

**Table 2: Frequencies of associated alleles in patients, subgroups and controls with, in brackets, p-values corrected for multiple testing using a Bonferroni correction for two independent subgroups and all independent alleles with an expected number of at least five copies.**

| Marker name | Allele | Controls | All patients | Without mutation | Not tested |
|---|---|---|---|---|---|
| **D10S1100** | 194 | 59.6 | 72.0 (n.s.) | 67.5 (n.s.) | 77.4 (0.018) |
| **M353** | 166 | 28.1 | 32.6 (n.s.) | 27.0 (n.s.) | 38.8 (n.s.) |
| **rs741763** | C | 58.3 | 81.5 ($2.0\cdot10^{-7}$) | 79.7 ($3.4\cdot10^{-4}$) | 83.7 ($3.1\cdot10^{-6}$) |
| **SNP-5** | A | 22.5 | 66.7 ($6.1\cdot10^{-24}$) | 63.7 ($1.7\cdot10^{-12}$) | 69.6 ($1.6\cdot10^{-17}$) |
| **SNP-1** | C | 58.0 | 81.2 ($1.1\cdot10^{-6}$) | 81.0 ($2.1\cdot10^{-4}$) | 81.4 ($1.1\cdot10^{-4}$) |
| **rs2435362** | A | 26.2 | 67.6 ($1.4\cdot10^{-22}$) | 64.9 ($1.5\cdot10^{-12}$) | 70.6 ($3.1\cdot10^{-16}$) |
| **rs2565206** | T | 67.1 | 84.7 ($1.2\cdot10^{-4}$) | 83.9 (0.0058) | 85.7 (0.0020) |
| **rs1800858** | A | 23.1 | 67.1 ($2.2\cdot10^{-24}$) | 64.5 ($1.6\cdot10^{-13}$) | 70.0 ($3.1\cdot10^{-17}$) |
| **MRETInt5** | 243 | 30.3 | 48.5 (0.0030) | 46.4 (n.s.) | 51.1 (0.023) |
| **rs1800860** | G | 84.3 | 85.7 (n.s.) | 90.0 (n.s.) | 83.3 (n.s.) |
| **rs1800862** | C | 93.6 | 99.1 (0.024) | 98.4 (n.s.) | 100.0 ($5.5\cdot10^{-4}$) |
| **IVS19+47C>T** | T | 19.2 | 31.4 (n.s.) | 30.9 (n.s.) | 32.0 (n.s.) |
| **s-TLC2** | 171 | 14.7 | 31.6 ($6.0\cdot10^{-4}$) | 30.7 (0.045) | 32.7 (0.027) |

n.s.=not significant

Separate association analyses were performed on the groups of patients that were known to be *RET* mutation negative and those that were not screened for the mutations in order to exclude possible variations in allelic associations. Because of the fewer number of cases, these analyses were, of course, less powerful than on the whole patients group. Nevertheless, significant associations were found for the same alleles at the same marker loci for both groups. Hence, we assumed that the group of patients not tested for the *RET* mutations was similar to the group of patients negative for a *RET* mutation and we could consider our patients' population as a whole in further analyses. Furthermore, we observed that a large proportion of our patients is homozygous for the associated alleles at the six marker loci that showed the highest association with HSCR, whereas these homozygous genotypes are hardly present in controls (Table 3a). This particularly is the case for SNP-5 allele A homozygotes – 52.0% (53/102) among patients versus 6.4% (8/125) among controls; rs2435362 allele A homozygotes – 53.7% (58/108) versus 10.0% (13/130) and rs1800858 allele A homozygotes – 54.3% (57/105) versus 7.1% (9/126). The ORs for the homozygote genotypes at these three SNPs ranged from 16.95 to 26.50, whereas the ORs for the heterozygote genotypes ranged from 2.48-2.93 (Table 3a).

**Table 3: Odd's ratios and 95% confidence intervals for the six most significantly associated loci, (a) separately and (b) combined in a haplotype. The confidence intervals are corrected for multiple testing (13 loci x 2 subgroups).**

**(a)**

| name | genotype | controls | cases | OR | 95% CI |
|---|---|---|---|---|---|
| | | | | | |
| **rs741763** | G,G | 23 | 9 | 1 | - |
| | C,G | 68 | 23 | 0.86 | 0.21-3.60 |
| | C,C | 45 | 79 | 4.49 | 1.17-17.25 |
| **totals** | | 136 | 111 | | |
| **SNP-5** | G,G | 76 | 19 | 1 | - |
| | A,G | 41 | 30 | 2.93 | 0.99-8.68 |
| | A,A | 8 | 53 | 26.50 | 6.42-109.3 |
| **totals** | | 125 | 102 | | |
| **SNP-1** | A,A | 22 | 10 | 1 | - |
| | C,A | 62 | 18 | 0.64 | 0.15-2.70 |
| | C,C | 42 | 73 | 3.82 | 1.02-14.37 |
| **totals** | | 126 | 101 | | |
| **rs2435362** | C,C | 76 | 20 | 1 | |
| | A,C | 41 | 30 | 2.78 | 0.95-8.15 |
| | A,A | 13 | 58 | 16.95 | 4.97-57.85 |
| **totals** | | 130 | 108 | | |
| **rs2565206** | G,G | 15 | 3 | 1 | - |
| | G,T | 57 | 28 | 2.46 | 0.31-19.73 |
| | T,T | 62 | 80 | 6.45 | 0.85-48.93 |
| **totals** | | 134 | 111 | | |
| **rs1800858** | G,G | 77 | 21 | 1 | - |
| | G,A | 40 | 27 | 2.48 | 0.84-7.32 |
| | A,A | 9 | 57 | 23.22 | 6.04-89.25 |
| **totals** | | 126 | 105 | | |
| **(b)** | | | | | |
| **name** | **genotype** | **controls** | **cases** | **OR** | **95% CI** |
| **Six loci haplotype** | Non-carriers | 80 | 28 | 1 | - |
| | Heterozygotes | 22 | 17 | 2.21 | 0.66-7.39 |
| | Homozygotes | 6 | 46 | 21.90 | 4.86-98.69 |

*Transmission/Disequilibrium Test (TDT)*

For the TDT, 113 case-parents triads were considered. Table 4 shows the results of the multi-allelic TDT, described in detail for the specific associated alleles. Large differences are observed in transmission of alleles at marker loci on the 5' region of the *RET* gene. For rs741763, allele C is transmitted 70 times versus 20 non-transmitted ($p=9.3\cdot10^{-6}$); for SNP-5, allele A is transmitted 97 times versus 14 non-transmitted ($p=2.3\cdot10^{-13}$); for SNP-1, allele C is transmitted 63 versus 17 non-transmitted ($p=1.8\cdot10^{-5}$); for rs2435362, allele A is 100 transmitted versus 14 non-transmitted ($p=5.3\cdot10^{-14}$); for rs2565206 allele T is transmitted 48 times versus 18 non transmitted (p=0.015); for rs1800858 allele A is transmitted 102 times versus 18 non-transmitted ($p=1.2\cdot10^{-12}$). Allele transmissions in the two subgroups (*RET* negative versus not *RET* screened) did not significantly differ from each other. These findings confirm the association results that we observed.

*Haplotype analyses*

Based on the results of the single locus association analyses, we estimated frequencies of transmissions for the haplotypes consisting of the six strongest associated marker loci (Table 5). Out of 23 haplotypes found, the haplotype most frequent among patients was the one consisting of the alleles that gave the highest single locus associations (rs741763 – allele C; SNP-5 - allele A; SNP-1 – allele C; rs2435362 – allele A; rs2565206– allele T; rs1800858 – allele A): it was transmitted in 55.6% and in 16.2% it was not. The corresponding risk of over-transmission of this haplotype was estimated to be 10.31. Odds ratios (ORs) (Table 3b) for the haplotype composed of the six highest associated alleles reveal that the risk of developing HSCR is highly increased, in particular, for homozygotes for the associated haplotype (OR=21.90; 95% CI 4.86-98.69) as compared to the risk of carriers of one associated haplotype (OR=2.21; 95% CI 0.66-7.39).

**Table 4: Transmissions of associated alleles in the entire sample and subgroups with p-values corrected for multiple testing for two independent (sub)groups and all independent alleles with a frequency of at least five copies using a Bonferroni correction.**

| Marker name | Allele | All patients | Without mutation | Not tested |
|---|---|---|---|---|
| **D10S1100** | 194 | 60 tr, 35 ntr (n.s.) | 39 tr, 20 ntr (n.s.) | 21 tr, 15 ntr (n.s.) |
| **M353** | 166 | 51 tr, 33 ntr (n.s.) | 23 tr, 14 ntr (n.s.) | 28 tr, 19 ntr (n.s.) |
| **rs741763** | C | 70 tr, 20 ntr ($9.3 \cdot 10^{-6}$) | 36 tr, 12 ntr (0.036) | 34 tr, 8 ntr (0.0041) |
| **SNP-5** | A | 97 tr, 14 ntr ($2.3 \cdot 10^{-13}$) | 48 tr, 3 ntr ($2.0 \cdot 10^{-8}$) | 49 tr, 11 ntr ($6.3 \cdot 10^{-5}$) |
| **SNP-1** | C | 63 tr, 17 ntr ($1.8 \cdot 10^{-5}$) | 31 tr, 7 ntr (0.0067) | 32 tr, 10 ntr (0.046) |
| **rs2435362** | A | 100 tr, 14 ntr ($5.3 \cdot 10^{-14}$) | 50 tr, 4 ntr ($2.6 \cdot 10^{-8}$) | 50 tr, 10 ntr ($1.6 \cdot 10^{-5}$) |
| **rs2565206** | T | 48 tr, 18 ntr (0.015) | 30 tr, 7 ntr (0.011) | 18 tr, 11 ntr (n.s.) |
| **rs1800858** | A | 102 tr, 18 ntr ($1.2 \cdot 10^{-12}$) | 50 tr, 5 ntr ($8.8 \cdot 10^{-8}$) | 52 tr, 13 ntr ($8.9 \cdot 10^{-5}$) |
| **MRETint5** | 243 | 64 tr, 30 ntr (0.030) | 35 tr, 15 ntr (n.s.) | 29 tr, 15 ntr (n.s.) |
| **rs1800860** | G | 18 tr, 19 ntr (n.s.) | 4 tr, 5 ntr (n.s.) | 14 tr, 14 ntr (n.s.) |
| **rs1800862** | C | 13 tr, 2 ntr (n.s.) | 10 tr, 2 ntr (n.s.) | 3 tr, 0 ntr (n.s.) |
| **IVS19+47C>T** | T | 51 tr, 22 ntr (0.046) | 24 tr, 9 ntr (n.s.) | 27 tr, 13 ntr (n.s.) |
| **s-TLC2** | 171 | 60 tr, 22 ntr (0.0018) | 33 tr, 10 ntr (0.030) | 27 tr, 12 ntr (0.030) |

n.s. = not significant; tr = transmissed; ntr = non-transmitted

**Table 5: Frequencies and relative risks (RR) of transmission of hapotypes consisting of the six most associated loci as estimated using the software package UNPHASED[22]. The risk haplotype is depicted in italic.**

| | Haplotype | freq-T | freq-NT | OR |
|---|---|---|---|---|
| **1.**[a] | CGCAGG | 0.004484 | 0.01345 | 1 |
| **2.** | CGCATG | 0.003612 | 0.005357 | 2.023 |
| **3.** | CGCCGG | 0.1611 | 0.2869 | 1.685 |
| **4.** | CGCCGA | 0.001933 | 0.002867 | 2.023 |
| **5.** | CGCCTG | 0.008969 | 0.0449 | 0.5994 |
| **6.** | CGCCTA | 0.001806 | 0.002678 | 2.023 |
| **7.** | CGACGG | 0.003612 | 0.005357 | 2.023 |
| **8.** | CGACTG | 0.01345 | 0.02242 | 1.8 |
| **9.** | CACAGG | 0.001806 | 0.002678 | 2.023 |
| **10.** | CACAGA | 0.001806 | 0.002679 | 2.023 |
| **11.** | CACATG | 0.01377 | 0.01763 | 2.342 |
| **12.** | CACATA | 0.5557 | 0.1617 | 10.31 |
| **13.** | CACCGG | 0.001806 | 0.002678 | 2.023 |
| **14.** | CACCTG | 0.001806 | 0.002678 | 2.023 |
| **15.** | CACCTA | 0.02242 | 0 | -3 |
| **16.** | GGCCGG | 0.001807 | 0.00268 | 2.023 |
| **17.** | GGCCTG | 0.001806 | 0.002678 | 2.023 |
| **18.** | GGAATG | 0.003612 | 0.005357 | 2.023 |
| **19.** | GGAATA | 0.001806 | 0.002678 | 2.023 |
| **20.** | GGACGG | 0.008969 | 0.02696 | 0.9981 |
| **21.** | GGACTG | 0.1704 | 0.359 | 1.424 |
| **22.** | GGACTA | 0.004487 | 0.01314 | 1.024 |
| **23.** | GACATA | 0.008969 | 0.01345 | 2 |

[a]haplotype 1 is the haplotype consisting of all wildtype alleles.

freq-T-frequency of transmitted haplotypes; freq-NT-frequency of non-transmitted haplotypes; OR-odds ratios; LR = 91.84, df=10, $p=2.3 \cdot 10^{-15}$.

## Discussion

In 70-90% of the sporadic HSCR patients, no mutations in any of the nine known HSCR susceptibility genes are identified. This raises the question whether, despite the fact that no pathogenic mutations can be detected in the majority of sporadic patients, the known HSCR susceptibility genes are involved in the development of the disease, or whether yet unknown genes are responsible. A possible involvement of *RET*, the major gene in HSCR, might be concluded from several studies. Bolk et al.[18] showed that 11 of 12 families investigated were linked to *RET*. In six of them however no clear *RET* mutation could be identified in the coding and flanking intronic sequences, as checked by SSCP, DGGE and sequence analysis. It suggests an involvement of RET although not through a clear *RET* mutation. Several association and haplotype studies also support this hypothesis.[26,31,32] Conserved haplotypes could be constructed using the alleles identified for different markers in and around the *RET* gene.[24,33,34,35] Carrasquilo et al.[33] used the highest number of SNPs in and around the *RET* locus. They found that patients shared common haplotypes with markers in both the proximal and distal part of the gene without having a clear pathogenic mutation. This study was carried out on a Mennonite kindred, an isolated population in which HSCR occurs in 1 in 500 live births. Therefore, it is likely that their findings apply to this population only and that the associated haplotype will be in LD with mutations that are most likely different from those that are present in other populations. In the recent study, Fitze et al.[35] showed that the haplotype ACA comprising alleles from SNPs –5, –1 and rs1800858 (c135G>A) is over-represented in their patients population (66.9% of 80 cases). Moreover, of 58 HSCR patients, all non-mutation carriers, 62.1% proved homozygous for a two-locus (SNP-5, SNP-1) haplotype. Sancandi et al.[24] performed a haplotype analysis on a smaller group of HSCR patients (46 patients and 50 population matched control individuals). They

genotyped two SNPs in the promoter region lying close together (SNP-1 and SNP-5) and SNPs in exon 2 (A45A) and exon 13 (L769L). As relatives of the patients were not available for screening real haplotypes could not be reconstructed. They estimated frequencies of haplotypes consisting of the markers in the promoter region and those in exons 2 and 13. Two of the haplotypes, differing only in the allele from last marker from exon 13, appeared to be much more frequent among HSCR patients. (62% of the patients chromosomes versus 22% of control chromosomes). Assuming that these SNPs are not causative variants themselves, Sancandi et al.[24] localised the unknown variant upstream of exon 2, although based on their results the region between exon 2 and 13 cannot be excluded. Borrego et al.[34] described a haplotype analysis on 103 HSCR patients and their parents. They genotyped three SNPs at the end of intron one (IVS1-1463T>C; IVS1-1370C>T; IVS-126G>T), and seven in exons 2, 3, 7, 11, 13, 14, and 15. Significant haplotype frequency differences were found for the three SNPs in intron 1. From the whole pool of genotyped SNPs, they reconstructed one haplotype that was transmitted 9 times and never non-transmitted. However, when they reconstructed a haplotype consisting solely of the three markers located at the end of intron 1, a haplotype spanning a region of 1.2 kb was observed that was far more common in HSCR patients (59.2%) than in controls (18.5%). Furthermore, they suspected that, based on extrapolation of the strength of LD at the observed SNPs, LD would become stronger upstream to the SNPs from intron 1 and that the susceptibility variant should lie upstream but probably still in the intron 1.

All these studies suggest that the *RET* gene is involved in sporadic HSCR and that an ancestral mutation is likely to be located in the 5' end of the gene. We typed 13 markers, including four microsatellites and nine SNPs within and flanking the gene, in order to better define the region in which this ancestral mutation might be located.

In correspondence with previous studies, significant associations with HSCR were found for SNPs in the 5' region of the *RET* gene. These were even stronger than those published in previous studies. Six successive SNPs, namely: rs741763, SNP-5, SNP-1, rs2435362, rs2565206 and rs1800858 (Table 1 and 2), were found to be strongly associated with HSCR in our population with the highest frequency increase of 23.1% among controls to 67.1% among patients. Associated alleles at marker loci genotyped also in other studies[24,34,35] were found to be the same (SNP-5, allele A; SNP-1, allele C; rs2565206, allele T; rs1800858, allele A). All of the alleles of genotyped markers also showed a significant transmission distortion (Table 3). Furthermore, a large proportion of our patients proved to be homozygous for the alleles at these six marker loci, whereas these homozygous genotypes had a very low frequency in controls. We were able to reconstruct haplotypes consisting of these six SNPs and observed one that was very commonly transmitted to the patients (55.6%) and significantly under-represented among controls (16.2%). Homozygosity for this ancestral haplotype was observed in 50.5% of our patients, versus 5.5% of our controls. It appears that homozygosity for this, most likely European ancestral haplotype, in sporadic HSCR cases gave a much higher increased risk of developing HSCR than heterozygosity, i.e. the disease appears dosage-dependent with respect to a mutation in this ancestral haplotype.

The six strongest associated SNPs are located in the interval spanning from the promoter region and the beginning of the exon 2 (27 kb). The region of association may not extend far 5' to the transcriptional start site of the *RET* gene, since the two microsatellites markers (*M353*, *D10S1100*) located upstream of the *RET* locus do not segregate significantly with HSCR. Nevertheless, we cannot exclude that the unknown mutation is lying upstream of the promoter region as microsatellite *M353* is located 80 kb upstream of *RET*. Alleles of marker *s-TCL2*, which lies downstream of the gene were associated when consider the patients as a whole.

However, when separating the analyses on both group of patients (screened for *RET* and not tested for *RET*), no association between HSCR and an allele of the marker, as well as TDT was found.

We found very strong associations for three SNPs, and in particular large differences in numbers of homozygotes between patients and controls giving ORs, ranging from 16.95 to 26.50 (see table 3). This might suggest that one or more of these SNPs are themselves causally involved in HSCR. For some of the SNPs included this has indeed been proposed. For instance the polymorphism in exon 2 might interfere with correct splicing[34] but experimental proof has not been presented. Polymorphisms in the introns have been suggested to disrupt the binding sites of regulatory proteins and thereby change gene transcription.[34] Calculations indicated four possible binding sites for regulatory proteins. Again, it has not been proven experimentally. Fitze et al.[35] performed functional study on basal *RET* promoter sequences, carrying different haplotypes at loci –5 and –1. Fitze's group has found that expression of a reporter gene in NMB and Vi-856 cell lines is reduced under the *RET* promoter carrying the 'AC' haplotype. The expression was even lower with the 'AA' haplotype, however this haplotype was not present in patients. They came to the conclusions that –5A variant can alter *RET* promoter activity and modulate the HSCR phenotype. It should however be noted that expression of a gene in different cell lines might give contradicting results. Preliminary results of the Ceccherini group indeed show such differences (paper in prep. Ceccherini et al). Sequence analysis of the promoter region up to intron 1 should reveal all possible SNPs which, when typed on the entire sample, might provide sufficient statistical information to identify the true causal mutation(s).

In conclusion, we observed a very strong association for several SNPs in the promoter and in the 5' region of the *RET* gene. An increased risk of HSCR was most evident for patients homozygous for the associated alleles. The haplotype consisting of these markers showed similar

results, indicating that a strong founder effect is present in our population. The alleles and consequently the halpotype found in our study is similar to that found by others who analysed other European patient populations. The ancestral haplotype might therefore be very old making it likely that most of the European HSCR patients share the same disease associated variant(s). We, therefore, expect that these results will allow us to eventually identify the mutation, which is obviously playing a major role in the susceptibility to HSCR.

## Acknowledgements

Supported by NWO grant 901-04-210 and NWO grant 901-04-225

## References

1. Holschneider A. Hirschsprung's congenital megacolon. The concept of physiopathology and therapy. Med Welt 1982;**33**:210-213.

2. Badner J, Sieber W, Garver K, Chakravarti A. Am J Hum Genet 1990;**46**:568-580.

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

4. Edery P, Lyonnet S, Mulligan L *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. Angrist M, Bolk S, Halushka M, Lapchak PA, Chakravarti A. Germline mutations in glial cell line-derived neurotrophic factor (GDNF) and *RET* in a Hirschsprung disease patient.Nat Genet 1996;**14**:341-344.

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. Puffenberger E, Hosoda K, Washington S, Nakao K, de Wit D, Yanagisawa M, Chakravart A. A missense mutation of the endothelin-B receptor gene in multigenic Hirschsprung's disease. Cell 1994;**79**:1257-1266.

10. 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.

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

15. 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.

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, Sham M, Lui V, Chen B, Ott J, Tam P. Association study of PHOX2B as a candidate gene for Hirschsprung's disease. Gut. 2003;**52**:563-567.

18. 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 USA 2000;**97**:268-273.

19. 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

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

21. Chakravarti A. Endothelin receptor-mediated signaling in hirschsprung disease. Hum Mol Genet 1996;**5**:303-307.

22. Falk CT, Rubinstein P. Haplotype relative risks: an easy reliable way to construct to construct a proper control sample for risk calculations. Ann Hum Genet 1987;**51**:227-233.

23. Terwilliger JD, Ott J. A haplotype-based ‘haplotype relative risk’ approach to detect allelic associations. Hum Hered 1992;**42**:337-346.

24. 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 disease. J Med Genet 2003;**40**:714-718.

25. Ceccherini I, Hofstra RM, Luo Y *et al*. DNA polymorphisms and conditions for SSCP analysis of the 20 exons of the ret proto-oncogene. Oncogene 1994;**9**:3025-3029.

26. Griseri P, Pesce B, Patrone G *et al.* A rare haplotype of the *RET* proto-oncogene is a risk modifying allele in Hirschsprung disease. Am J Hum Genet 2002;**71**:969-974

27. Carlson KM. Three dinucleotide repeat polymorphisms closely linked to the RET protooncogene D10S1098, D10S1099 and D10S1100. Hum Mol Genet 1994;**3**:1207.

28. Pasini B, Hofstra RM, Yin L, Bocciardi R *et al.* The physical map of the human RET proto-oncogene. Oncogene 1995;**11**:1737-1743.

29. Breslow N, Day N. The analysis of case-control studies. Statistical methods in cancer research I, p. 134. 1980 Lyon, France: IARC Sci. Publ. No 32.

30 Woolf B. On estimating the relationship between blood group and disease. Ann Hum Genet 1995;**19**:251-253

31. Fitze G, Schreiber M, Kuhlisch E, Schackert H. Association of *RET* protooncogene codon 45 polymorphism with Hirschsprung disease. Am J Hum Genet 1999;**65**:1469-1473.

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

33. Carrasquilo M, McCallion A, Puffenberg E, Kashuk C, Nouri N, Chakravarti A. Genome-wide association study and mouse model identify interaction between *RET* and EDNRB pathways in Hirschsprung disease. Nat Genet 2002;**32**:237-244.

34. Borrego S, Wright F, Fernandez R *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.

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

# Appendix 2

## Identifying candidate Hirschsprung disease associated RET variants

Grzegorz M. Burzynski[1], Ilja M. Nolte[2], Agnes Bronda[1], Krista K. Bos[1], Jan Osinga[1], Ivan Plaza Menacho[1], Bas Twigt[1], Saskia Maas[3], Alice S. Brooks[4], Joke B.G.M. Verheij[1], Charles H.C.M Buys[1] & Robert M.W. Hofstra[1]

1. Department of Genetics, University Medical Center Groningen, Groningen, The Netherlands
2. Department of Medical Biology, University Medical Center Groningen, Groningen, The Netherlands
3. Department of Clinical Genetics, Academic Medical Centre, Amsterdam, The Netherlands
4. Department of Clinical Genetics, Erasmus Medical Center Rotterdam, The Netherlands

***American Journal of Human Genetics*** **2005 May;76(5):850-8**

## Abstract

Patients with sporadic Hirschsprung disease (HSCR) show an increased allele sharing at markers in the 5' region of the *RET* locus, indicating the presence of a common ancestral *RET* mutation. In a previous study, we found a haplotype of six SNPs that was transmitted to 55.6% of our patients, whereas it was present in only 16.2% of the controls we used. Among the patients with that haplotype, 90.8% had it on both chromosomes, which gave a much higher increased risk of developing HSCR than when the haplotype occurred heterozygously. In order to more precisely define the HSCR-associated region and to identify candidate disease-associated variant(s), we sequenced the shared common haplotype region from 10 kb upstream the *RET* gene through intron 1 (in total 33 kb) in a patient homozygous for the common risk haplotype and in a control individual homozygous for the most common non-risk haplotype. A comparison of these sequences revealed eighty-six sequence differences. Eight of these eighty-six variations proved to be in regions highly conserved among different vertebrates and within putative transcription factor binding sites. We therefore, considered these as candidate disease-associated variants. Subsequent genotyping of these eight variants revealed a strong disease association for six of the eight markers. These six markers also showed the largest distortions in allele transmission. Interspecies comparison showed that only one of the six variations was located in a region also conserved in a non-mammalian species, making it the most likely candidate HSCR associated variation.

Hirschsprung disease (HSCR [MIM 142623]) is caused by the absence of intrinsic ganglion cells in the myenteric and submucosal plexuses of the gastrointestinal tract, resulting in severe intestinal obstructions. HSCR is a heterogenic disorder, since a number of genes have been shown to play a role in the disease etiology (Holschneider 1982; Badner et al. 1990; Amiel and Lyonnet 2001). The *RET* gene [MIM 164761] is a main factor in Hirschsprung development. *RET* encodes a transmembrane tyrosine kinase receptor, a protein responsible for triggering a number of downstream signaling pathways involved in crucial processes, such as neural crest cell differentiation, migration and proliferation (Manie et al. 2001). Mutations in *RET* have been found in up to 50% of familial HSCR cases. For the more common sporadic form of HSCR, *RET* coding mutations have been found in not more than 20% of the patients (Hofstra et al. 2000; Amiel and Lyonnet 2001). Several studies have shown, however, that the *RET* locus is linked to the disease in almost all familial cases, regardless of their mutation status (Bolk et al. 2000; Gabriel et al. 2002) and is also associated with HSCR in a large proportion of the sporadic HSCR patients that do not have *RET* coding mutations. Furthermore, recently published articles revealed that similar haplotypes are found in the 5' region of *RET* locus in patients from several European Hirschsprung populations (Borrego et al. 2003; Fitze et al. 2003; Sancandi et al. 2003; Burzynski et al. 2004) indicating the segregation of the same ancestral mutation(s) in the whole European HSCR population. The identified region of association comprises intron 1 of *RET* gene and depending on the markers genotyped by the different groups, different lengths of upstream genomic sequence (Borrego et al. 2003; Fitze et al. 2003; Sancandi et al. 2003; Burzynski et al. 2004). In a previous study, we found that a haplotype consisting of six SNPs, spanning 27 kb (from 4 kb upstream of the gene through the whole intron 1) was transmitted from the parents to our sporadic patients in 55.6% of cases and not transmitted in only 16.2% (Burzynski et al.

2004). Moreover, for individuals homozygous for the risk haplotype the odds ratio to develop HSCR was 21.9, whereas for the heterozygotes it was 2.2, suggesting dose dependence of the unknown variants (Burzynski et al. 2004). Three groups tried to correlate specific haplotypes of two SNPs (which were also part of our risk haplotype) located in the basal *RET* promoter (-5 and -1 bp from the transcription start site) with reduced *RET* expression by performing functional assays in human cell lines (Fitze et al. 2003; Griseri et al. 2005; Garcia-Barcelo et al. 2005). However, results were not consistent, so that these SNPs cannot unambiguously be considered as being pathogenic.

In this study, we attempt to further investigate the region of interest by sequencing 33 kb, from 10 kb upstream of the gene through 23 kb intron 1 and exon 2, in a patient homozygous for the common risk haplotype and in a healthy control individual homozygous for the most common non-risk haplotype. In order to sequence this 33 kb region (genomic position [NCBI]: chromosome 10: 42,846,516 –42,880,012 bp), 49 primer sets were designed. The average size of the amplification products was 750 bp and between amplicons there was an overlap of about 50 bp. Primer sequences can be found on our website. Samples were sequenced by BaseClear, Leiden, The Netherlands. By restricting the comparison for economic reasons to a single patient and a single control individual, one runs the risk of missing SNP(s) in the comparison. However, since the haplotypes are quite common, one might expect that a SNP responsible for the difference will have a very high probability to be found.

Sequence comparison between the two sequences (the sequence of the common HSCR haplotype and the most common control haplotype) revealed 86 different sequence variants (Table 1). In intron 1 a deletion of 12 bp was found (IVS+3991), as where two insertions (IVS1+5274insCT and IVS1-11084insG) and two 2-base substitutions (IVS1+6383-84GG>CC and IVS1-6567CG>GC). The remaining 81 variations are all single nucleotide polymorphisms.

**Table 1. Sequence variations in the common risk haplotype as compared to the most common non-risk haplotype.**

| | Position and type of variant | dbSNP accession no | | Position and type of variant | dbSNP accession no |
|---|---|---|---|---|---|
| 1. | -9557G>C | - | 44. | IVS1+6757A>G | rs1897000 |
| 2. | -9341C>T | rs3026719 | 45. | IVS1+6813C>T | rs1896999 |
| 3. | -9256C>G | rs2505989 | 46. | IVS1+7236G>A | rs2435359 |
| 4. | -7359A>C | rs2505990 | 47. | IVS1+7436A>G | rs2435358 |
| 5. | -6347G>A | rs2505991 | 48. | IVS1+7445A>C | rs2506008 |
| 6. | -6093T>C | - | 49. | IVS1+7994C>G | rs2506007 |
| 7. | -5669G>C | rs2505992 | 50. | IVS1+9494C>A | rs2506004 |
| 8. | -4636G>C | rs732610 | 51. | IVS1+10371G>A | rs2435356 |
| 9. | -4361C>G[a] | rs741763 | 52. | IVS1+10813A>G | - |
| 10. | -4201G>A | rs2082107 | 53. | IVS1+11369T>C | rs2506021 |
| 11. | -4146C>G | - | 54. | IVS1+11476C>T | rs2435342 |
| 12. | -3820C>T | rs2505994 | 55. | IVS1-11490T>C | rs2506022 |
| 13. | -3054G>A | rs2505995 | 56. | IVS1-11084insG | - |
| 14. | -2897T>C | rs2505996 | 57. | IVS1-10285G>T | rs2435343 |
| 15. | -2352T>C | rs2505997 | 58. | IVS1-10032C>T | rs2435344 |
| 16. | -1782G>A | rs2505998 | 59. | IVS1-10017C>T | rs2435345 |
| 17. | -1260T>C | rs2505999 | 60. | IVS1-9651T>G | - |
| 18. | -719C>T | rs2435366 | 61. | IVS1-8047C>T | rs2506023 |
| 19. | -200G>A (SNP-5)[b] | rs10900296 | 62. | IVS1-7593G>A | rs752975 |
| 20. | -196A>C (SNP-1)[b] | rs10900297 | 63. | IVS1-7525C>T | rs752977 |
| 21. | IVS1+620C>A | rs3123712 | 64. | IVS1-7477T>A | rs752978 |
| 22. | IVS1+779T>C | rs2506010 | 65. | IVS1-7298G>A | - |
| 23. | IVS1+1813C>T | rs2435365 | 66. | IVS1-6988G>A | rs2506024 |
| 24. | IVS1+2157C>T | rs2506011 | 67. | IVS1-6567C>G | - |
| 25. | IVS1+2820T>C | rs1864411 | 68 | IVS1-6566G>C | - |
| 26. | IVS1+2846G>T | rs1864410 | 69. | IVS1-6393A>G | - |
| 27. | IVS1+3104C>T | rs1864408 | 70. | IVS1-6257G>C | rs3128726 |
| 28. | IVS1+3442A>G | rs2506013 | 71. | IVS1-5791C>T | rs2505541 |
| 29. | IVS1+3460C>T | rs2435364 | 72. | IVS1-5666C>T | rs2505540 |
| 30. | IVS1+3470G>A | - | 73. | IVS1-5127C>G | - |
| 31. | IVS1+3788C>G | rs2506014 | 74. | IVS1-5107T>C | - |
| 32. | IVS1+3991del12 | - | 75. | IVS1-5038C>T | rs2505539 |
| 33. | IVS1+4247C>G | rs2506015 | 76. | IVS1-5005A>G | rs2435346 |
| 34. | IVS1+5094T>A | rs2506016 | 77. | IVS1-4925G>A | - |
| 35. | IVS1+5274insCT | - | 78. | IVS1-4503G>A | rs2505538 |
| 36. | IVS1+6000C>A[c] | rs2435362 | 79. | IVS1-4073C>A | - |
| 37. | IVS1+6136G>T | rs2506019 | 80. | IVS1-3748C>T | rs2505537 |
| 38. | IVS1+6174G>A | rs2435361 | 81. | IVS1-3702T>A | rs2505536 |
| 39. | IVS1+6294T>C | rs2506020 | 82. | IVS1-2863G>A | rs2505535 |
| 40. | IVS1+6383G>C | - | 83. | IVS1-891A>T | - |
| 41 | IVS1+6384G>C | - | 84. | IVS1-888G>C | - |
| 42. | IVS1+6411G>A | - | 85. | IVS1-881C>T | - |
| 43. | IVS1+6751T>C | rs1897001 | 86. | c135G>A (exon 2)[d] | rs1800858 |

Variants no 9, 19, 20, 36 and 84 are SNPs from our previously characterized 6-loci ancestral haplotype. Positions of variations localized upstream the gene refer to *RET* ATG. Superscripts indicate SNPs previously reported in HSCR studies; a: Burzynski et al. (2004); b: Sancandi et al. (2003); c: Burzynski et al. (2004); d: Ceccherini et al. (1994)

All the variations were found in a homozygous state as well in the patient as in the control's DNA. Table 1 includes as indicated five of the six SNPs that formed our previously identified haplotype. The sixth marker IVS1-126G>T is not included since patient DNA and the control DNA had the same allele at this locus. Twenty variants were located in the region upstream of the gene, 65 in intron 1 and one is located in exon 2 (Table 1). The sequence variation frequency in the upstream region is 1/536 bp, for the intronic region it is 1/353 bp. This might indicate a tighter selection mechanism on the upstream sequence as a consequence of the possible presence of more important regulatory regions. On the other hand, the common risk haplotype observed in our patients also spans the whole of intron 1, supporting the hypothesis that this intronic region might also be involved in the regulation of *RET* expression.

As we aim at delimiting the HSCR-associated region by further genotyping we selected from these 84 sequence variations the most likely candidate disease-associated variants. Variants were first selected on the basis of their location in regions conserved among vertebrate species. To identify evolutionarily conserved domains within this 33 kb region, we compared the human genomic sequence with the genomic DNA sequences of chimpanzee (*Pan troglodytes*), rat (*Rattus norvegicus*), mouse (*Mus musculus*) and chicken (*Gallus gallus*) (downloaded from the UCSC database). Sequence comparisons between these species were performed using the freely available MUltiple sequence Local AligNment and conservation visualization tool – MULAN and the software package Clone Manager Professional Suite (Needleman-Wunsch method). Figure 1 shows the level of evolutionary conservation between the chimpanzee, mouse, rat, chicken and human sequence as a reference. Rat and mouse genomes show a comparable level of conservation with the human sequence. There are 11 regions with a similarity higher than 70%. The distal part of the human intron 1 shows no homology with the mouse and rat counterparts, with the exception of the short region just downstream of exon 2 and of exon 2 itself.

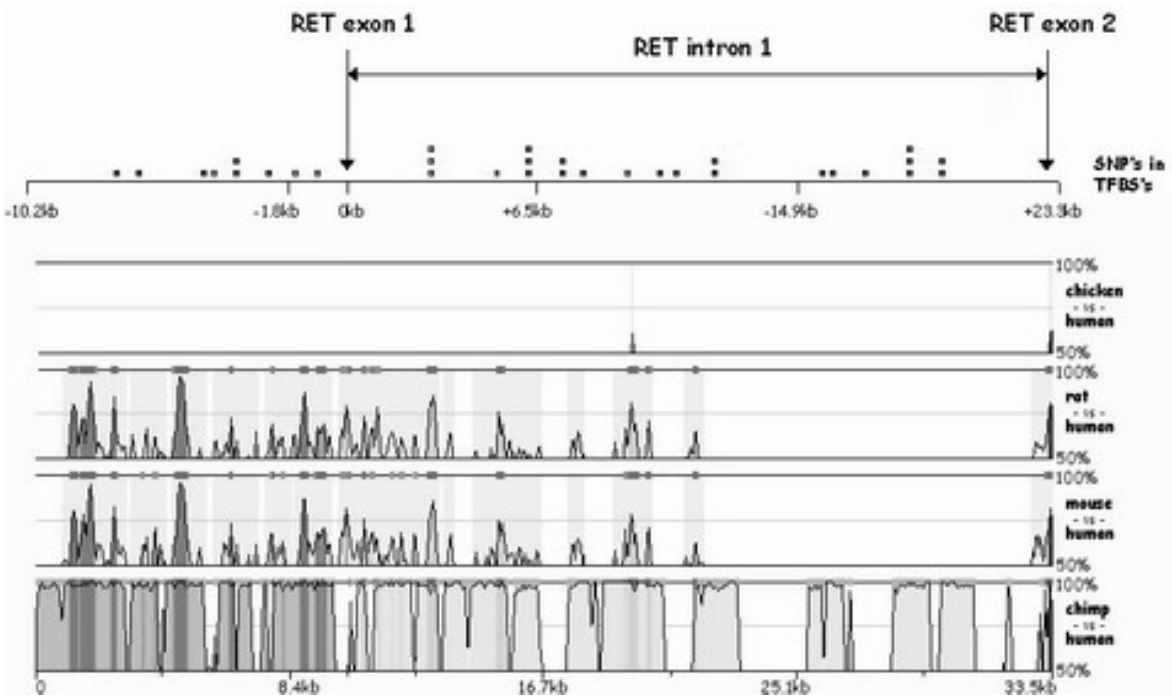



**Figure 1. MULAN plots displaying the percentages of similarity of the 5' region of the *RET* locus in the chicken, rat, mouse and rat genomic sequences, compared to human sequence as a reference. The horizontal gray line indicates 70% identity. The most upper panel of the figure depicts the positions of the 31 SNPs that are located within putative transcription factor binding sites. Those indicated in light gray fall in one of the evolutionarily conserved regions. The upper scale refers to *RET* ATG (0 kb). The lower scale represents the length of the whole region.**

Furthermore, it has to be noticed that the rodents' intron 1 is only 12 kb long, thus roughly half as long as the human intron 1. Moreover, in the mouse genome, homology to the human intron 1 is found in a region 30 kb upstream of the gene and also in intron 18 of the mouse *Ret* gene (data not shown). This could mean that this region has been translocated either in rodents or in primates, since the chimpanzee's intron 1 is highly similar to the human one.

A region that is almost identical between rodents and mammals (99.5%) is located between the positions -5.7 kb and -5.2 kb with reference to *RET* ATG (Fig. 1). Between the human and chicken sequences there is hardly any significant similarity. This might imply that conserved transcriptional regulation of *RET* is restricted to mammals. However, similarity occurs and was found in the region corresponding to the human sequence between positions +9.3 kb and

+9.7 kb, where strong homology with chimpanzee, rat and mouse sequences is observed as well (Fig. 1). This is the only region conserved among all species studied. The second region of moderate conservation in chicken is located at the distal part of the sequence in the beginning of exon 2.

The sequence of the chimpanzee is almost identical to the human one, except for a number of repeats, point mutations or small deletions/insertions. A puzzling observation is that exon 1 is missing from chimpanzee genome database. Instead, there is a repeated sequence in this position. This is probably a sequencing error, as the chimpanzee genome database is still not assembled well enough and many gaps exist between contigs.

The occurrence of sequence homology between different species in non-coding regions (including promoter and intron regions) might suggest that these conserved regions contain binding sites for relevant transcription factors. We, therefore, checked the sequenced genomic region for presence of specific consensus binding elements for mammalian transcription factors and we determined whether the identified DNA variants were present in these putative transcription factor consensus binding elements. For this we used the UCSC Genome Browser and the program GENEQUEST from the DNASTAR software package. We aimed at finding such DNA elements that were lost in the patient DNA, as that might result in a diminished expression and by that explain the HSCR phenotype. We detected 31 base changes leading to loss of such a binding site (Table 2). Among the binding sites that were found, many are recognized by ubiquitous transcription factors such, as *SP1*, *AP2* or *NF*. We did find, however, also consensus binding elements for more specifically acting factors such as *Msx1*, which play a role in neural tube development (Table 2). In the top of Figure 1, the approximate positions are given of those SNPs that cause loss of the putative transcription factor binding sites. Eight of them fall in the previously mentioned evolutionarily conserved regions (-4201G>A, -1260T>C, -719C>T,

IVS1+2820T>C, IVS1+2846G>T, IVS1+3104C>T, IVS1+5094T>A and IVS1+9494C>A) (Fig.1, Table 2).

**Table 2. Sequence variations in the common risk haplotype with loss of a hypothetical consensus binding element**

| | Position | dbSNP accession no | Transcription factor |
|---|---|---|---|
| **1.** | -7359A>C | rs2505990 | CSS(2) |
| **2.** | -6347G>A | rs2505991 | AP2/SP1, SP1erk |
| **3.** | -4201G>A | rs2082107 | SP1-YB1, ApoE-B1 |
| **4.** | -3820C>T | rs2505994 | USF-APP, IgHC.4, NF-mu-E1-CS, GT-2B-RS |
| **5.** | -3054G>A | rs2505995 | HC5 |
| **6.** | -2897T>C | rs2505996 | NF-E1, NF-E1.6 |
| **7.** | -1260T>C | rs2505999 | MEF3_B |
| **8.** | -719C>T | rs2435366 | SP1erk, AP2/SP1, SP1-TPI(4), SP1-hsp70, hsp70.2, SP1-IE-3.3,4.5 |
| **9.** | IVS1+2820T>C | rs1864411 | Uteroglobin HS-2.4-C, AP2-TBXAS1 |
| **10.** | IVS1+2846G>T | rs1864410 | AP2/SP1, SP1-erk1(1), JCV |
| **11.** | IVS1+3104C>T | rs1864408 | NFI-CS3, ELP/SF1/FT2F1, SF1-CYP21 |
| **12.** | IVS1+5094T>A | rs2506016 | BPV-E2.CS2, E2-RS1 |
| **13.** | IVS1+6294T>C | rs2506020 | Ets-1 |
| **14.** | IVS1+6383-84GG>CC | - | SP1 complement factor, SP1-CS4, HC3 |
| **15.** | IVS1+6411G>A | - | AP2 CS4 |
| **16.** | IVS1+6751T>C | rs1897001 | PEA3, IE1.2 |
| **17.** | IVS1+6813C>T | rs1896999 | Site-I(2) |
| **18.** | IVS1+7445A>C | rs2506008 | HNF-5 |
| **19.** | IVS1+9494C>A | rs2506004 | E1AF |
| **20.** | IVS1+10813A>G | - | Msx1 |
| **22.** | IVS1-10032C>T | rs2435344 | E2A-CAMLG, E2A, USFAPP |
| **23.** | IVS1-10017C>T | rs2435345 | SP1-TPI(4), AP-2-beta |
| **24.** | IVS1-7593G>A | rs752975 | GR-MT-IIA |
| **25.** | IVS1-7298G>A | - | AP2-CS5 |
| **26.** | IVS1-5791C>T | rs2505541 | AP2-CS5 |
| **27.** | IVS1-5127C>G | - | TFIID-EIIa |
| **28.** | IVS1-5038C>T | rs2505539 | Uteroglobin HS-2.4-C, T antigen |
| **29.** | IVS-5005A>G | rs2435346 | SP1YB1, MyoD-MCK right site, site F (4,5), AP2-APP |
| **30.** | IVS1-3748C>T | rs2505537 | AP2-beta |
| **31.** | IVS1-3702T>A | rs2505536 | GR-uteroglobin(1,2), GR/Prconnexin43 |

Previously we reported 23 different haplotypes (Burzynski et al. 2004), of which four were homozygously present in our patients or controls, two risk and two non-risk haplotypes. We typed individuals homozygous for these four haplotypes for seven SNPs by direct sequencing, and for one , IVS1+9494C>A, by digestion (HpyCH4IV) of amplified PCR products. For four of the eight markers, including the two most proximal SNPs, SNPs -4201 G>A and -1260 T>C, the

same alleles were identified on both the risk and non-risk haplotypes (Fig. 2). Four SNPs ( SNPs -719C>T, IVS1+2846, IVS1+3104 and IVS1+9494) showed perfect segregation with the disease, meaning that one allele was present on a risk haplotype and the other on a non-risk haplotype. The candidate region for the causal mutation in HSCR is therefore likely bordered by SNP-1260C>A and SNP IVS1-126G>T. As SNPs -4201 G>A and -1260 T>C are not part of the candidate region, only SNPs -719C>T, IVS1+2820T>C, IVS1+2846G>T, IVS1+3104C>T, IVS1+5094T>A and IVS1+9494C>A were subsequently genotyped in all our HSCR families.

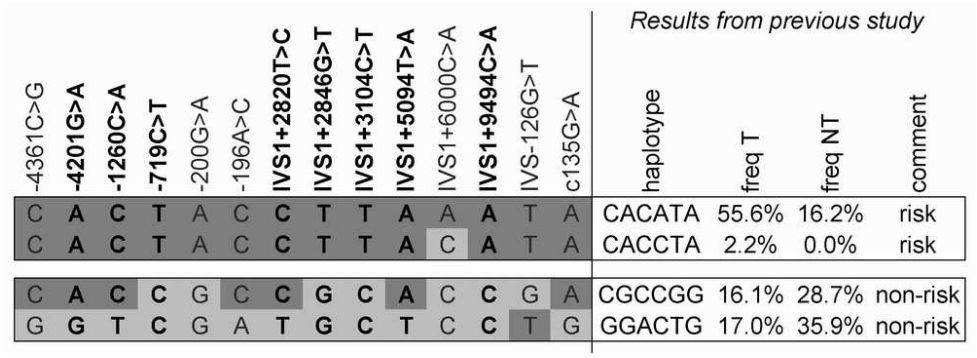



**Figure 2. Haplotypes consisting of 14 SNPs, six from our previous risk haplotype and eight newly selected ones, of four homozygous individuals. Individuals homozygous for one of the haplotypes identified in our previous study (the corresponding SNPs are indicated in gray) were typed for eight new SNPs (indicated in bold). They were also homozygous for all eight new SNPs. Risk alleles are put in red boxes, the non-risk alleles in green ones.**

DNA samples used for genotyping were obtained upon written informed consent from sporadic HSCR patients and their parents residing in the Netherlands and referred to clinical genetic centers. Almost all our families (113 out of 117) consist of two parents and the affected child (trios). In three cases only one parent was available for the study and one patient participated with his child and spouse. Of all patients, 64 were screened by means of DGGE and sequence analysis (Hofstra et al. 2000) and found negative for *RET* mutations. These 64 were selected from an initial group of 72 sporadic HSCR patients from whom 8 appeared to carry a *RET* mutation. Fifty-three of the patients included in the study have not been screened for a RET mutation, which therefore may be present in a few of them. As controls for the transmitted

haplotypes we used the non-transmitted haplotypes of unaffected parents, as does the Haplotype Relative Risk method (Falk and Rubinstein 1987; Terwilliger and Ott 1992). The two non-transmitted haplotypes within each trio form pseudo-control genotypes. Under the hypothesis of random mating in our study population, these pseudo-control genotypes are expected to be similar to genotypes of population controls and can be used to provide estimates of the number of heterozygotes and homozygotes in the population and of the odds ratios of their risk to develop HSCR. The results of the TDT and association tests on the six newly selected SNPs and the six SNPs from our previous study are shown in Figure 3. All SNPs were strongly associated. For six of them (-200G>A [SNP-5], IVS1+2820T>C, IVS1+2846G>T, IVS1+3104C>T, IVS1+5094T>A and IVS1+9494C>A), the association was stronger than for the other six. Corresponding allele frequencies were all in the range of 66.7%-68.6% among patients and 21.8%-25.4% among the pseudo-controls. Haplotype analysis showed that these six risk alleles are present on the same risk haplotype CTACCTTAAATA (Table 3). Its frequency is estimated to be 53.3% among patients and 15.9% among controls (Tab. 3). Another haplotype CCACCTTAAATA, only different from the risk haplotype at the second

SNP, which also contains the risk alleles of our previous six-loci risk haplotype, was present in 6.01% among patients and 1.5% among controls. As this haplotype also confers an increased risk, it does not contribute to further restricting our candidate region. The question remains which of the identified variants is contributing to the disease development. Several candidates have already been put forward. SNP -200G>A (SNP-5) and

–196A>C (SNP-1) that were subjects of functional studies (Fitze et al. 2003; Griseri et al. 2005; Garcia-Barcelo et al. 2005) and are part of the risk haplotype identified in our previous study (Burzynski et al. 2004), were not present among the eight variants mentioned.

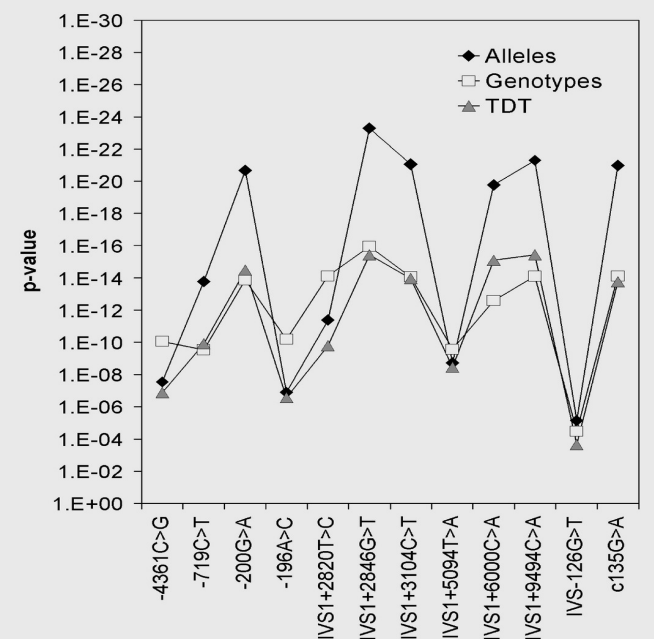



**Figure 3. Association, TDT and genotype results for the 5' *RET* locus SNPs. Results presented regard both markers characterized in the previous study (Burzynski et al. 2004) and 6 new loci. For the single locus allelic and genotypic association analyses, the frequencies of the alleles and the genotypes, respectively, were compared between patients and pseudo-controls using a chi-square test. Transmission distortion of each allele versus all other alleles together was tested and combined in a multi-allelic transmission distortion test (TDT) (Spielman and Ewens, 1996). All P-values and CIs were corrected for multiple testing for independent tests, using a Bonferroni correction.**

**Table 3: Haplotype frequencies over 12 SNPs, estimated by means of an Expectation-Maximization (EM) algorithm implemented in our own software. Newly selected SNPs are in bold.**

| **-4361C>G** | **-719C>T** | **SNP-5** | **SNP-1** | **IVS1+2820T>C** | **IVS1+2846G>T** | **IVS1+3104C>T** | **IVS1+5094T>A** | **IVS1+6000C>A** | **IVS1+9494C>A** | **IVS-126G>T** | **c135G>A** | **Patients (n=230)** | **Controls (n=290)** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| C | **C** | G | C | **C** | **G** | **C** | **A** | C | **C** | G | G | 11.04% | 20.88% |
| C | **C** | G | C | **C** | **G** | **C** | **A** | C | **C** | G | G | 0.00% | 3.43% |
| G | **C** | G | A | **T** | **G** | **C** | **T** | C | **C** | T | G | 11.05% | 24.41% |
| G | **T** | G | A | **T** | **G** | **C** | **T** | C | **C** | T | G | 0.91% | 4.64% |
| C | **C** | A | C | **C** | **T** | **T** | **A** | A | **A** | T | A | 6.05% | 1.46% |
| C | **T** | A | C | **C** | **T** | **T** | **A** | A | **A** | T | A | 53.31% | 15.94% |

These two candidate variations did not fulfill our current criteria, which are based on interspecies sequence conservation and changes in transcription factor binding sites. Nevertheless, Fitze et al. (2003) claimed that SNP –200G>A in particular caused reduced *RET* promoter activity. Griseri et al. (2005) found rather contradicting results, as they observed strong

cell line dependency and no consistent results in promoter activity assays. These two SNPs are localized in the close proximity of a transcription start site. They are part of the *RET* basal promoter, conserved in human, mouse and rat. Because of these contradictory results and of our findings presented here, it is likely that they are in linkage disequlibrium with mutation(s) lying close by. The closest polymorphism that we found upstream of SNP-5 is SNP-719C>T, one of our eight selected polymorphisms. It causes loss of several putative sites recognized by *SP1* and *AP2* transcription factors. It is likely that this SNP is still a part of the basal promoter, and thus a good candidate for future promoter activity studies. Three other SNPs (IVS1+2820T>C, IVS1+2846G>T, IVS1+3104C>T) are lying close together and are located in the proximal part of intron 1 in a region of strong sequence conservation between species (Fig. 1, peak at position -5.7 kb --5.2 kb). The most distal SNP of the eight, - IVS1+9494C>A is lying in an interesting region, since all species including chicken show homology with the human genome at this position (Fig. 1, peak at position +9.3 kb- +9.7 kb). The occurrence of evolutionarily conserved regions located in an intronic sequence supports the idea that intron 1 is involved in transcriptional regulation. Downstream of the last SNP IVS1+9494C>A, there is only one more short region of moderate homology in the human and rodents sequences (Fig. 1, peak at position +11.5 kb - +11.6 kb). Intron 1 in the mouse and the rat genomes is only 12 kb long, whereas in humans it is 23 kb. Likely, it is either a translocation or a deletion of this sequence in rodents or an insertion in primate genomes, since in the mouse we found regions of partial homology to the missing sequence at 30 kb upstream of *Ret* and also in intron 18 (data not shown).

Based on the statistical analyses, no single SNP could be identified as being the causal one (Fig. 3). As the region around SNP IVS1+9494 is the best conserved region, not only in mammals, but even in chicken, this DNA variation is a good candidate for disease susceptibility, although we cannot fully exclude that a nearby DNA variation(s) which we might have missed,

would actually be responsible or act in combination with proposed here candidate. Marker IVS1+9494C>A cause loss of a binding site of the *ETV4* enhancer (ets variant gene 4 *E1A* enhancer binding protein, *E1AF*, DNA binding sequence: 'TGGACGT'). *ETV4* is a member of the *Ets* transcription factors family and is a downstream nuclear target of *Ras-MAP* kinase signaling. Its disturbed functioning was implicated in malignant tumors, such as ovarian carcinoma and breast cancer (Davidson et al. 2004). Transcription factors with an evolutionarily conserved *Ets* domain, play in general important roles in cell development, cell differentiation, cell proliferation, apoptosis and tissue remodeling (Oikawa T. 2004), which suggests that *ETV4* can influence the HSCR phenotype. Functional studies on *ETV4* and its interaction with *RET* regulatory sequences should follow to determine the importance of this enhancer in *RET* transcription.

## Acknowledgements

Supported by NWO grants 901-04-210 and 901-04-225

## Electronic-database information

National Center for Biotechnology Information (NCBI), http://www.ncbi.nlm.nih.gov/

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/

Website of Department of Medical Genetics, Groningen, The Netherlands,

http://www.rug.nl/med/faculteit/disciplinegroepen/medischegenetica/hereditarydiseases/hirschprungDisease

MUltiple sequence Local AligNment and conservation visualization tool, http://mulan.dcode.org/

UCSC Genome Bioinformatics, http://genome.ucsc.edu/

## References

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

Badner J, Sieber W, Garver K, Chakravarti A (1990) A genetic study of Hirschsprung disease. Am J Hum Genet 46:568-580

Bolk S, Pelet A, Hofstra RMW, Angrist M, Salomon R, Croaker D, Buys CHCM, Lyonnet S, Chakravarti A (2000) A human model for multigenic inheritance: phenotypic expression in Hirschsprung disease requires both the *RET* gene and a new 9q31 locus. Proc Natl Acad Sci USA 97:268-273

Borrego S, Wright F, Fernandez R, Williams N, Lopez-Alonso M, Davuluri R, Antinolo G, Eng C (2003) 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 72:88-100

Burzynski GM, Nolte IM, Osinga J, Ceccherini I, Twigt B, Maas S, Brooks A, Verheij J, Plaza Menacho I, Buys CHCM, Hofstra RMW (2004) Localizing a putative mutation as the major contributor to the development of sporadic Hirschsprung disease to the RET genomic sequence between the promoter region and exon 2. Eur J Hum Genet 12:604-612

Ceccherini I, Hofstra RM, Luo Y, Stulp RP, Barone V, Stelwagen T, Bocciardi R, Nijveen H, Bolino A, Seri M
(1994) DNA polymorphisms and conditions for SSCP analysis of the 20 exons of the ret proto-oncogene.
Oncogene 9:3025-3029

Davidson B, Goldberg I, Tell L, Vigdorchik S, Baekelandt M, Berner A, Kristensen GB, Reich R, Kopolovic J (2004) The clinical role of the PEA3 transcription factor in ovarian and breast carcinoma in effusions. Clin Exp Metastasis 21:191-199

Falk CT, Rubinstein P (1987) Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. Ann Hum Genet 51:227-233

Fitze G, Appelt H, Konig I, Gorgens H, Stein U, Walther W, Gossen M, Schreiber M, Ziegler A, Roesner D, Schackert HK (2003) Functional haplotypes of the RET proto-oncogene promoter are associated with Hirschsprung disease (HSCR). Hum Mol Genet 12:3207-3214

Gabriel SB, Salomon R, Pelet A, Angrist M, Amiel J, Fornage M, Attie-Bitach T, Olson JM, Hofstra R, Buys C, Steffann J, Munnich A, Lyonnet S, Chakravarti A (2002) Segregation at three loci explains familial and population risk in Hirschsprung disease. Nat Genet 31:89-93

Garcia-Barcelo M, Ganster RW, Lui VC, Leon TY, So MT, Lau AM, Fu M, Sham MH, Knight J, Zannini MS, Sham PC, Tam PK (2005) TTF-1 and RET promoter SNPs: regulation of RET transcription in Hirschsprung's disease. Hum Mol Genet 14:191-204

Griseri P, Bachetti T, Puppo F, Lantieri F, Ravazzolo R, Devoto M, Ceccherini I (2005) A common haplotype at the 5' end of the *RET* proto-oncogene underlies genetic susceptibility to HSCR development through reduced gene expression. Hum Mut 2005 25:189-195

Hofstra RMW, Wu Y, Stulp RP Elfferich P, Osinga J, Maas SM, Siderius L, Brooks AS, vd Ende JJ, Heydendael VM, Severijnen RS, Bax KM, Meijers C, Buys CH (2000) RET and GDNF gene scanning in Hirschsprung patients using two dual denaturing gel systems. Hum Mut 15:418-429

Holschneider A (1982) Hirschsprung's congenital megacolon. The concept of physiopathology and therapy. Med Welt 33:210-213

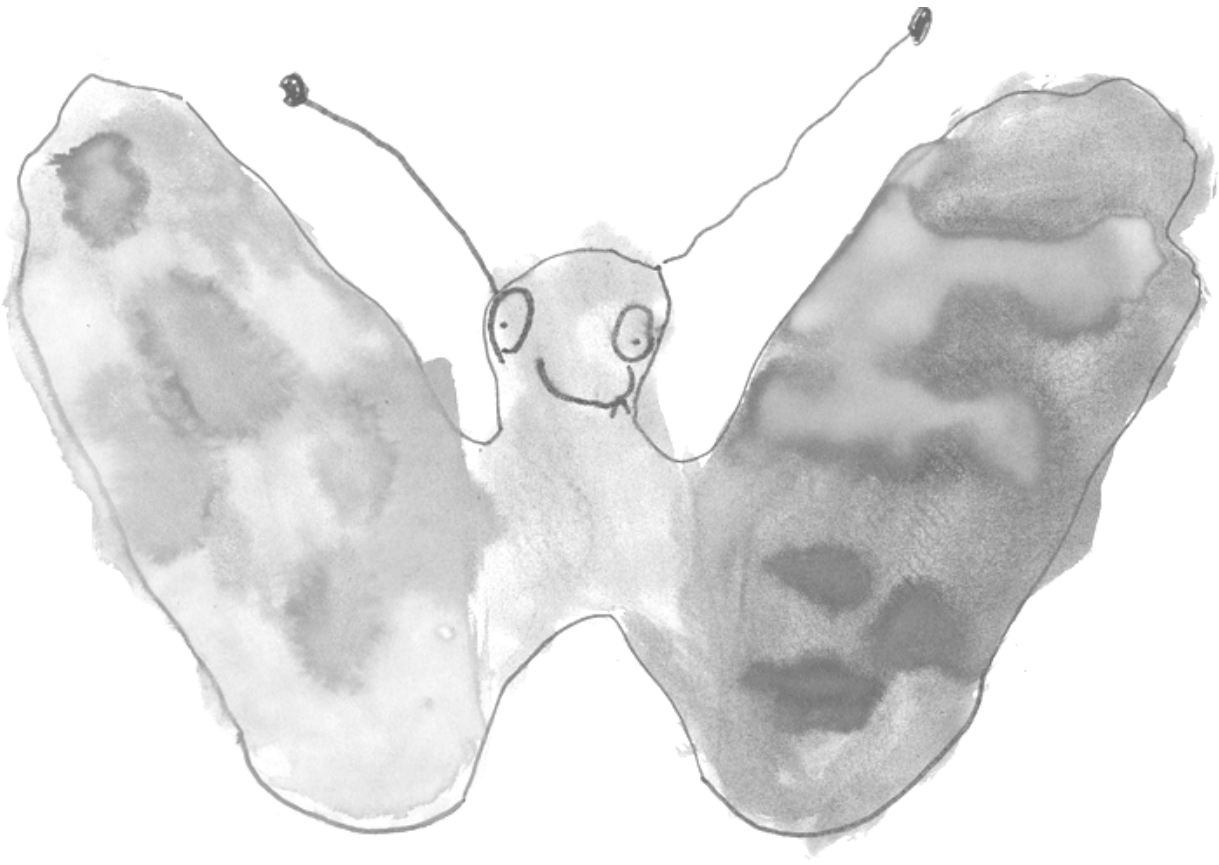
Manie S, Santoro M, Fusco A, Billaud M (2001) The RET receptor: function in development and dysfunction in congenital malformation. Trends Genet 17:580-589

Oikawa T (2004) ETS transcription factors: possible targets for cancer therapy. Cancer Sci 95:626-633

Sancandi M, Griseri P, Pesce B, Patrone G, Puppo F, Lerone M, Martucciello G, Romeo G, Ravazzolo R, Devoto M, Ceccherini I (2003) Single nucleotide polymorphic alleles in the 5' region of the *RET* proto-oncogene define a risk haplotype in Hirschsprung disease. J Med Genet 40:714-718

Spielman RS, Ewens WJ (1996) The TDT and other family-based tests for linkage disequilibrium and association. Am J Hum Genet 59:983-989

Terwilliger JD, Ott J (1992) A haplotype-based 'haplotype relative risk' approach to detect allelic associations. Hum Hered 42:337-346



# Appendix 3

## A novel susceptibility locus for Hirschsprung disease maps to 4q31.3-q32.3

Alice S. Brooks[1*], Peter A. Leegwater[1,2*], Grzegorz M. Burzynski[3], Patrick J. Willems[4], Bianca de Graaf[1], Irene van Langen[5], Peter Heutink[1,6], Ben A.Oostra[1], Robert M.W. Hofstra[3], Aida M. Bertoli-Avella[1]

* Both authors contributed equally to this work

1. Department of Clinical Genetics, Erasmus MC, Rotterdam, the Netherlands
2. Present address: Department of Clinical Sciences of Companion Animals, University of Utrecht, the Netherlands
3. Department of Genetics, University Medical Center Groningen, Groningen, The Netherlands
4. Synergene, Antwerp, Belgium
5. Department of Medical Genetics, Academic Medical Center, Amsterdam, The Netherlands
6. Present address: Section Medical Genomics, Department of Human Genetics and Department of Biological Psychology, VU University Medical Center, Amsterdam, The Netherlands

***Journal of Medical Genetics* 2006 Jul;43(7):e35.**

## Abstract

We report on a multigenerational family with isolated Hirschsprung disease (HSCR). Five patients were affected by either short-segment or long-segment HSCR. The family consists of two main branches with 4 patients (3 sibs and 1 maternal uncle) and 1 patient, respectively. Analysis of the *RET* gene, the major gene involved in HSCR susceptibility, revealed neither linkage nor mutations. A genome-wide linkage analysis was performed, revealing suggestive linkage to a region on 4q31-q32 with a maximum parametric multipoint LOD score of 2.7. Furthermore, non-parametric analysis of the genome-wide scan data revealed a NPL score of 2.54 (p= 0,003) for the same region on chromosome 4q (D4S413-D4S3351). The minimum linkage interval spans a region of 11.7 cM (12.2 Mb). No genes within this chromosomal interval have previously been implicated in HSCR. Considering the low penetrant-disease in this family, the 4q locus may be necessary but not sufficient to cause HSCR in the absence of modifying loci elsewhere in the genome; we hypothesize that the penetrance of this gene might be *RET*-dependent. Our results are suggesting the existence of a new susceptibility locus for HSCR at 4q31.3-q32.3.

## Introduction

Hirschsprung disease (HSCR, OMIM 143623) is a congenital disorder characterized by the absence of enteric neurons, which are neural-crest derived, in the digestive tract. Delayed passage of meconium is the cardinal symptom in neonates with HSCR. If untreated, bowel hypomotility leads to severe constipation often associated with obstruction, gross distention of the bowel and vomiting. The prevalence of HSCR is approximately 1 in every 5000 liveborns. In the majority of cases (75%-80%) the aganglionosis typically involves the rectum and the sigmoid (short segment Hirschsprung disease: SS-HSCR). In 20-25% of the patients the aganglionosis extends proximally of the rectosigmoid, and the disease is called long segment HSCR (LS-HSCR).[1] HSCR mostly presents as an isolated congenital malformation (non-syndromic HSCR), but can be found in association with other congenital abnormalities (syndromic HSCR).[2]

So far mutations in 10 genes (*RET*, *GDNF*, *EDNRB*, *EDN3*, *ECE1*, *SOX10*, *ZFHX1B*, *NTN*, *PMX2b*, and *KIAA1279*) have been implicated in HSCR. [3-15]

The *RET* gene located at 10q11.2 is the major susceptibility locus in HSCR: 15%-35% of the sporadic patients have inactivating mutations in the coding sequence of *RET* [16-19], whereas linkage analysis showed that all, but one, autosomal dominant families are linked to *RET*. [20] High-penetrance mutations in the coding sequence of the *RET* gene, however, are found in only 50% of the *RET* linked families. *RET* linked multigenerational families without a coding sequence mutation, might have a mutation in the non-coding sequence of the *RET* gene, including alterations in intronic and promoter sequences or harbor (frequent) variants that do change the function of the RET protein slightly.[20] Furthermore, similar haplotypes are found in the 5' region of the *RET* locus in HSCR populations patients from all over the world indicating the segregation of the same ancestral variant(s). [21-25] Evidence is accumulating that specific

(common) non-coding low-penetrance variants just before the gene and within intron 1 of *RET* are associated with HSCR susceptibility.[21, 24, 26-29]

Single mutations leading to either isolated or syndromic HSCR have been found in the above-mentioned 10 genes, although evidence is building up that HSCR is a multigenic congenital malformation in the majority of HSCR patients. A “multiplicative model” has been suggested which assumes that additional loci are involved apart from the *RET* locus, and their individual effects can be multiplied.[30]

Furthermore, 4 HSCR susceptibility loci (9q31, 3p21, 19q12, and 16q23) have been identified, harboring unidentified genes.[20, 30, 31] Linkage at 9q31 was reported in 5 families that also showed linkage with *RET*, however no causative *RET* mutation could be identified. A sixth family that was *RET*-unlinked, however, was linked to the locus at 9q31.[20] Furthermore susceptibility loci at 3p21 and 19q12 have been identified in affected nuclear families, suggesting that these 2 loci probably function as *RET*- dependent modifiers. Non-random allele sharing was also found at 9q31 in those nuclear families in which no *RET* mutation was identified, confirming the segregation of the 9q31 locus in multiplex families.[20, 30] In an inbred Mennonite population, *RET* not only interacts with *EDNRB*, but also with an unknown gene on chromosome 16q23. This locus is probably only of importance in this genetic isolated population [31] in which HSCR can be associated with symptoms also found in Waardenburg-Shah syndrome.[7] Conversely, no linkage with the other susceptibility loci at 3p21, 9q31 and 19q12 was found in this Mennonite kindred.

Clearly, HSCR is a heterogeneous congenital malformation. It is estimated that only 30% of cases can be attributed to mutations of the known genes.[32] Thus, a considerable number of additional genes involved in enteric nervous system development will be identified in the

future.[27] Here, we describe a five-generation family with non-syndromic HSCR and show evidence suggesting the presence of a new susceptibility locus on chromosome 4.

## Methods

### Patients

The family we describe here is a five-generation pedigree of native Dutch origin with 5 cases presenting HSCR. The segregation pattern is compatible with an autosomal dominant mode of inheritance with incomplete penetrance. One branch of the family consists of a sibship with 3 affected children and an affected uncle. Two sisters (V-1 and V-2) have SS-HSCR. Their brother (V-3) was diagnosed with total intestinal aganglionosis, both large and small intestines were aganglionic. Given the poor prognosis, a joint medical and parental decision was made for conservative care of HSCR; he died at the age of 1 month. Delayed passage of meconium led to the suspected diagnosis of HSCR in all 3 children. Suction and/or full thickness biopsies were consistent with this diagnosis. Their maternal uncle (IV-3) was operated during childhood because of SS-HSCR. The other branch of the family contains one affected female with SS-HSCR (V-4). Her paternal grandfather (III-3) is a cousin of the maternal grandmother (III-2) of the 3 sibs (Fig. 1). Congenital malformations indicative of syndromic HSCR were lacking in all 5 patients. Brainstem evoked response audiometry, which we performed because of his expected early death, showed no abnormalities in patient V-3. Chronic severe constipation was not reported in II-2, III-2, IV-4, although IV-2 as well as her father III-1 suffered from severe constipation in childhood. Informed consent was given by the parents and the adult patient (IV-3). Genomic DNA was isolated from peripheral blood obtained from II-2, III-2, IV-1, IV-2, IV-3, IV-4, IV-5, V-1, V-2, V-3 and V-4 using standard protocols.[33]

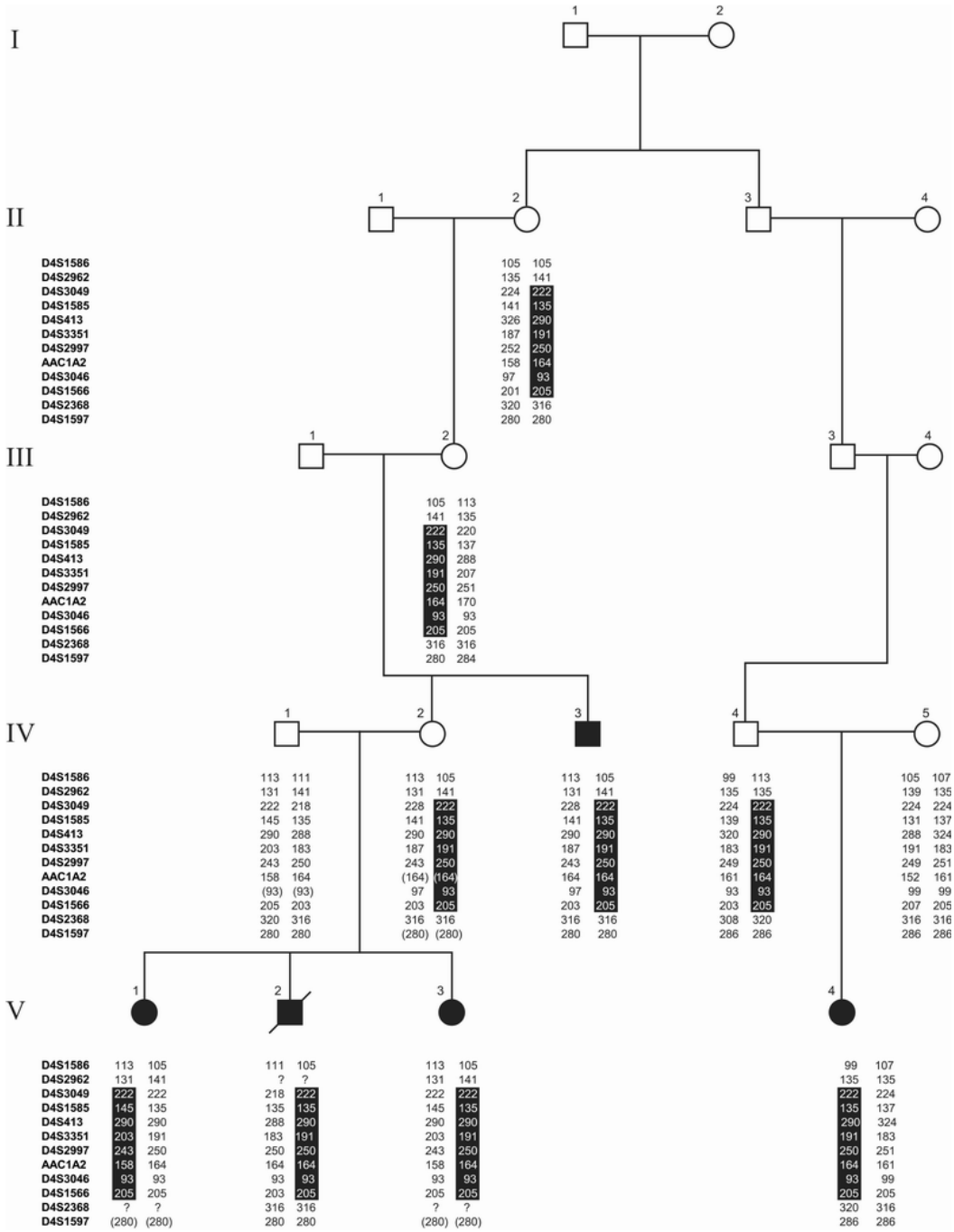



**Figure 1 Pedigree structure and haplotypes. Patients are represented as blackened symbols. The segregating 4q31-q32 haplotype is depicted. The minimum critical region at chromosome 4 spans 11.7 cM between D4S3049 and D4S1566.**

**Analysis of markers encompassing the *RET* locus**

The following markers D10S141 (- 1 Mb to *RET*), *RET*int5, D10S1099 (1.4 Mb downstream to *RET*) and 5 SNPs (rs741763, rs2435362, rs2565206, rs2506004) from the 5'*RET* region were analyzed as reported by us elsewhere [25, 26]and rs2435357 reported by others.[27]

**Mutational analysis of *RET***

Mutation analysis of the 21 exons of *RET* was performed in proband V-1 and V-4 as described by us elsewhere. [19]

**Genome wide linkage analysis**

We performed a systematic genome scan using the ABI Prism MD-10 set (Applied Biosystems) consisting of 382 markers (STRPs), with an average spacing of 10 cM. Additional markers for further characterization of candidate regions were selected from the sex-average Marshfield genetic map or newly designed. Genomic DNA (20 ng) was used as a template in 7.5 μl PCR reactions, with 5 pmol of oligonucleotides, 0.3 units Amplitaq Gold polymerase in Gold buffer (Applied Biosystems) and 2.5 mM MgCl2. The thermal cycling consisted of an initial incubation at 95 $^{o}$C for 5 minutes, followed by 10 cycles of 95 $^{o}$C for 30 seconds, 55 $^{o}$C for 15 seconds and 72 $^{o}$C for 30 seconds and 25 cycles of 92 $^{o}$C for 30 seconds, 55 $^{o}$C for 15 seconds and 72 $^{o}$C for 30 seconds. PCR products were pooled and loaded on an ABI3100 automated sequencer (Applied Biosystems). Data were analyzed using GeneMapper software (v 2.0). Mega2 [34] was used to process the genetic data into the appropriate format and perform data validation checks. Simulation analysis to estimate the probability of detecting genetic linkage given the pedigree structure (statistical power) was performed with the SLINK program.[35] Due to uncertainties

related to the correct genetic model in this pedigree, parametric and non-parametric linkage analyses were performed using SimWalk2 (version 2.9).[36]

For the parametric analysis we specified an autosomal dominant mode of inheritance, a mutant allele frequency of 0.01% with a penetrance of 40% and equal marker allele frequencies. Pedigree's Location Scores were calculated; these location scores are directly comparable to multipoint LOD scores. For the non-parametric analysis, the max-tree statistic (the largest number of affecteds inheriting an allele from one founder-allele) is reported. This statistic was designed for traits best modeled by dominant inheritance and was formerly known as STAT B. The NPL_ALL (STAT E) statistics, a measure of whether a few founder-alleles are overly presented in affecteds (suiTable for an additive model) is reported as well. A large value of the statistic indicates a high degree of identity by descent (IBD) allele sharing among the patients, usually a result above 2 can be considered significant. Empirical p-values (10,000 simulations) are also reported. Namely, this p-value is the probability of obtaining a value for that statistic equal to, or more extreme than, the observed value, if the trait were unlinked to the markers.

## Results

### Linkage analysis to the *RET* locus and sequence analysis of the *RET* gene

The family we investigated is a five-generation pedigree of native Dutch origin with 5 cases presenting HSCR (Fig. 1). Since the majority of the multigenerational families with HSCR show linkage to the *RET* gene [20], we investigated the *RET* locus by haplotype analysis and mutational analysis. We observed that the 5 patients (IV-3, V-1, V-2, V-3 and V-4) did not share the same haplotype at the *RET* locus (Table 1), excluding linkage with the *RET* locus in this family. We also performed sequence analysis of the entire coding region of the *RET* gene. Direct sequencing revealed no mutations in patients V-1 (branch 1) and V-4 (branch 2).

**Table 1. *RET* haplotypes**

| | II-2 | III-2 | IV-1 | IV-2 | IV-3 | IV-4 | IV-5 | V-1 | V-2 | V-3 | V-4 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **D10S141** | 3 1 | 3 2 | 2 2 | 4 3 | 4 2 | 3 4 | 2 2 | 2 4 | 2 4 | 2 4 | 3 2 |
| **rs741763** | C G | C G | G G | C C | C G | C C | G G | G C | G C | G C | C G |
| **rs2435362** | C C | C C | A A | C C | C C | C C | C A | A C | A C | A C | C C |
| **rs2435357** | C C | C C | T T | C C | C C | C C | C T | T C | T C | T C | C C |
| **rs2506004** | C C | C C | A A | C C | C C | C C | C A | A C | A C | A C | C C |
| **rs2565206** | G T | G T | G G | G G | G T | G G | T G | G G | G G | G G | G T |
| **Retint5** | 3 2 | 3 2 | 1 3 | 2 3 | 2 2 | 3 1 | 2 1 | 1 2 | 1 2 | 3 2 | 3 2 |
| **D10S1099** | 5 2 | 5 5 | 5 4 | 5 5 | 5 5 | 2 3 | 5 1 | 5 5 | 5 5 | 4 5 | 2 5 |

### Genome search

Simulation analysis (SLINK) yielded an average LOD score of 1.54 and a maximum of 2.15. We performed a genome wide search using 382 STRPs. Results from the parametric linkage analysis excluded most of the genome (data not shown). Only two genomic regions displayed LOD scores above 1, a region on chromosome 17 (D17S798 mLOD score=1.15) and on chromosome 4 (mLOD=1.84). We tested additional markers and performed haplotype analysis. The region on

chromosome 17 was rapidly excluded since one patient (IV-3) was not sharing the haplotype observed in patients from branch 1.

The highest multipoint LOD score (mLOD=1.84) was obtained for chromosome 4 between markers D4S424 and D4S413. This region fully segregated with the disease phenotype (Fig. 1). When we saturated the chromosome 4 region with additional markers, a maximum mLOD of 2.7 (Fig. 2) was reached between markers D4S1585 and D4S3351 which is consistent with suggestive linkage according to the Lander-Kruglyak guidelines for significance thresholds.[37]

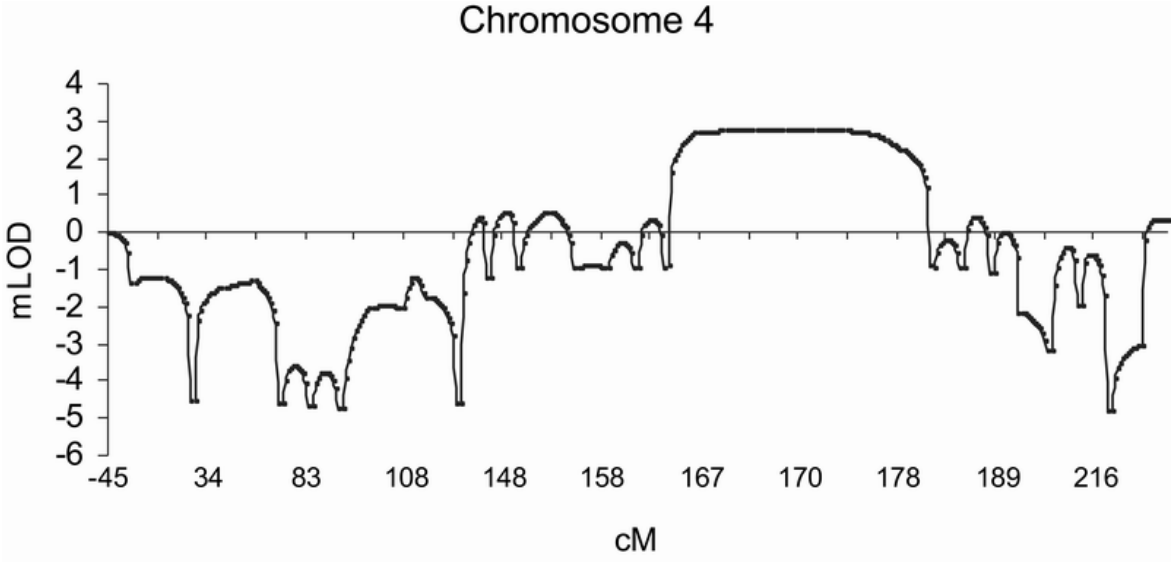



**Fig. 2 Multipoint LOD-score analysis with several markers on chromosome 4. The X- axis represents the chromosomal position of the markers. The Y- axis indicates the multipoint LOD score, showing a peak LOD score of 2.7.**

We also performed non-parametric analysis for our genome wide scan data. Only for this region on chromosome 4q, we found non-parametric scores (NPL) above 1. The maximum NPL was found for marker D4S413 (158.0 Mb, NCBI build 35.1) located between markers D4S1585 and D4S3351 (NPL=2.55, p=0.003), under both dominant and additive mode of inheritance (Table 2).

Recombination events can be identified in individual IV-4 and show that marker D4S2962 at the centromeric site, and D4S2368 at the telomeric site limits the critical region. The maximum critical region between D4S2962 and D4S2368 spans approximately 16.4 cM (19.7 Mb according to NCBI D4S1566 (11.7 cM or 12.2 Mb).

**Table 2. Results from the non-parametric analysis on chromosome 4q after fine mapping.physical the minimum shared region extend from marker D4S3049 until map, build 35.1),**

| Marker Name | Genetic Position cM* | Physical Position Mb† | Max-tree‡ | NPL_All§ | Empirical p-value |
|---|---|---|---|---|---|
| D4S1575 | 132.1 | 135.1 | 0.35 | 0.38 | 0.418 |
| **D4S1597** | 169.4 | 170.2 | 0.74 | 0.79 | 0.163 |
| **D4S1644** | 143.3 | 142.1 | 0.80 | 0.83 | 0.148 |
| **D4S424** | 144.6 | 142.6 | 0.82 | 0.84 | 0.143 |
| **D4S1625** | 146.0 | 143.9 | 0.84 | 0.85 | 0.141 |
| **D4S1586** | 147.1 | 147.1 | 0.87 | 0.86 | 0.136 |
| **EDNRA** | -- | 148.8 | 0.86 | 0.86 | 0.137 |
| **D4S2962** | 153.0 | 150.7 | 0.87 | 0.86 | 0.137 |
| **D4S3049** | 155.2 | 155.1 | 1.60 | 1.60 | 0.025 |
| **D4S1585** | 158.0 | 157.9 | 2.42 | 2.42 | 0.004 |
| **D4S413** | 158.0 | 158.7 | 2.54 | 2.55 | 0.003 |
| **D4S3351** | 158.7 | 159.9 | 2.54 | 2.54 | 0.003 |
| **D4S2997** | 158.7 | 160.0 | 2.52 | 2.51 | 0.003 |
| **AAC1A2** | 158.0 | 160.0 | 2.51 | 2.51 | 0.003 |
| **D4S3046** | 162.5 | 163.7 | 1.55 | 1.57 | 0.030 |
| **D4S1566** | 166.9 | 167.3 | 0.99 | 0.99 | 0.102 |
| **D4S2368** | 167.6 | 169.1 | 0.83 | 0.84 | 0.144 |
| **D4S2431** | 176.2 | 175.2 | 0.51 | 0.63 | 0.234 |
| **D4S415** | 181.4 | 179.1 | 0.10 | 0.11 | 0.776 |

* According to Marshfield sex average genetic map
† According to NCBI physical map, build 35.1
‡Max-tree is the allele-sharing statistic for traits best modeled as dominant inheritance.
§NPL-All is the statistics for traits following an additive inheritance.

**5' *RET* common risk haplotype analysis**

To test whether non-coding low-penetrance variants just before or within intron 1 of *RET* were associated with HSCR susceptibility in a part of this family, we typed 5 SNPs all being part of an ancestral haplotype.[26] Spouse IV-1 (married-in individual) was homozygote for the GATAG haplotype that is part of the core risk SNP haplotype detected in European, European-American and Asian-American patients with sporadic HSCR.[26, 27] The 3 affected sibs (V-1, V-2 and V-3) were thus carriers of the core risk haplotype. However, the ther 2 patients (IV-3 and V-4) were not carrying the 5' *RET* common risk haplotype (Table 1).

## Discussion

We identified a five-generation family with 5 patients affected with HSCR. The patients were connected to a common ancestor within 3-4 generations. Two branches with HSCR patients within the family were identified. The inheritance pattern in the family is compatible with an autosomal dominant mode of inheritance with reduced penetrance, although an oligogenic model with the contribution of 2 or more loci could not be discarded.

Since *RET* is the major gene involved in HSCR susceptibility in multigenerational families [20], linkage to the *RET* locus and mutations in the coding sequence of *RET* were first excluded. Subsequently, we performed a genome-wide scan to map the disease gene(s) in this family. Model-free or non-parametric linkage analysis methods are more robust than parametric or model-dependent analysis, when the mode of inheritance or the genetic model is uncertain, such as the pedigree with isolated HSCR reported here. Consequently, we performed both parametric and non-parametric linkage analysis. Both methods highlighted the same region on chromosome 4q, no other known HSCR susceptibility loci showed positive LOD scores. When adjacent markers for chromosome 4 were tested, the evidence of linkage became stronger and we could observe a common haplotype extending at least 11.7 cM that was inherited by all the affected individuals from their common ancestor. Our results are strongly suggesting the existence of a novel HSCR susceptibility locus on chromosome 4q.

Clearly, the chromosome 4q locus has incomplete penetrance. Is this 4q locus solely leading to HSCR or alternatively, is the phenotype only expressed in the presence of other susceptibility loci? Modifier loci either can increase susceptibility and severity of the phenotype or can act protectively to confer resistance to the disease in the face of a predisposing mutation.[38] Since HSCR is a well known complex, oligogenic disorder

[20, 30, 31], we consider the possibility of another gene that could act in synergy with the chromosome 4q locus. However, we could not find evidence pointing to other genomic regions where the patients showed an excess of allele sharing, including the known HSCR susceptibility loci at chromosome 3, 9, 16 or 19. Subsequently, we focused again our attention at the *RET* gene.

Can variants within the *RET* gene explain the difference in penetrance observed in both branches? We investigated whether all patients shared a haplotype similar to the common risk haplotype defined by SNPs located in the 5' region of the *RET* locus reported in Dutch HSCR patients.[25] We identified non risk haplotypes in patients IV-3 and V-4. However, the spouse IV-1and the 3 affected children V-1, V-2, and V-3 were homozygote and heterozygotes, respectively, for the GATAG haplotype, which contains the core risk haplotype detected in European, European-American and Asian-American patients with sporadic HSCR.[26, 27] The third and fourth SNPs (RET3+, rs2435357 and IVS1+9494, rs2506004) are particularly interesting as disease-associated *RET* variants, because of the homology and evolutionary conservation between rodents and primates and the differences in allele/genotype frequencies among patients and controls.[26] Recent data show that RET3+ might lie within, and might compromise the activity of an enhancer-like sequence in *RET* intron 1.[27] These findings make us hypothesize that the 5' *RET* risk haplotype in combination with the identified chromosome 4 locus is involved in the high disease penetrance and severity observed in the sibship with 3 affected children.

A recent genome-wide scan identified novel modifier loci of aganglionosis in a *SOX10* Dom mouse model of HSCR on mouse chromosomes 3, 5, 8, 11 and 14. [39] The identified region on chromosome 14 is consistent with the position *EDNRB*, a known *SOX10* modifier. [40] Interestingly, the chromosome 4 locus we reported here is close to the syntenic region on mouse chromosome 8.

We looked for candidate genes in the minimum 12.2 Mb-linked region at 4q31.3-32.3 (between markers D4S3049 and D4S1566) in the human genome sequence. This region contains at least 57 genes, in accordance with the National Center for Biotechnology Information (NCBI) build 35.1 of the human genome and the Ensemble Genome Browser. The maximum 20-Mb linked region between D4S2962 and D4S1597 encompasses 93 genes, including several interesting functional candidate genes that are proposed to be involved in neural crest development or neuronal development.

Unfortunately, the maximum genetic interval contains far too many candidates to begin functional evaluation of each gene individually; the best positional and functional candidate we could identify is *Mab21L2* (named after male abnormal 21 in C. Elegans). *Mab21L2* is expressed in the central nervous system and neural crest in midgestation embryogenesis in mice.[41] Furthermore, *Mab21L2* is linked to the TGFbeta signaling pathway to which also *ZFHX1B* belongs. [12] [42] *ZFHX1B*, the gene involved in Mowat-Wilson syndrome, a syndromic form of HSCR, is a transcription factor that functions as a TGF-beta receptor substrate. For these reasons we sequenced the complete coding region of the *Mab21L2* gene for mutations; however we did not identify a sequence variant in *Mab21L2*. Besides *Mab21L2*, many other candidate genes are located in the region. Several proteins encoding neuropeptide Y receptors (NPY2R, NPY1R and NPY5R) are located in the minimum 12.2 Mb-linked region. In mammals NPY, the ligand, is mainly found in cells derived from the neural crest, and is widely distributed in the central and peripheral nervous system.[43] Furthermore, the gene for the secreted frizzled related protein 2 (*SFRP2*) is located in this region.[44] Wnt-frizzled signaling is involved in neural crest formation.[45]

The 4q locus may be necessary but not sufficient to cause HSCR in the absence of modifying loci elsewhere in the genome. Linkage analysis in additional multigenerational HSCR

families and association studies in sporadic patients with high-density marker sets covering the entire interval will be necessary to confirm this suggestive linkage to chromosome 4q. Eventually, identification of the causative gene defect will specify the susceptibility to HSCR conferred by this novel locus at 4q31.3-q32.3.

## Electronic database information

Accession numbers and URLs for data in this article are as follows:

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim (for MIM 142623).

## Acknowledgment

We are grateful to the family who made this study possible. This work was supported in part by The Termeulen foundation (grant to ASB) and the Nederlandse organisatie voor Wetenschappelijk Onderzoek (grants 901-04-210 and 901-04-225 to RMWH).

## References


1. Puri P. Hirschsprung's disease. In: Holschneider A, ed. *Hirschsprung's disease,* 2 ed. Amsterdam: Harwood Academic Publishers, 2000;129-135.

2. Amiel J, Lyonnet S. Hirschsprung disease, associated syndromes, and genetics: a review. *J Med Genet* 2001;**38**(11):729-39.

3. Romeo G, Ronchetto P, Luo Y, Barone V, Seri M, Ceccherini I, Pasini B, Bocciardi R, Lerone M, Kaariainen H, et al. Point mutations affecting the tyrosine kinase domain of the RET proto-oncogene in Hirschsprung's disease. *Nature* 1994;**367**(6461):377-8.

4. Edery P, Lyonnet S, Mulligan LM, Pelet A, Dow E, Abel L, Holder S, Nihoul-Fekete C, Ponder BA, Munnich A. Mutations of the RET proto-oncogene in Hirschsprung's disease. *Nature* 1994;**367**(6461):378-80.

5. Angrist M, Bolk S, Halushka M, Lapchak PA, Chakravarti A. Germline mutations in glial cell line-derived neurotrophic factor (GDNF) and RET in a Hirschsprung disease patient. *Nat Genet* 1996;**14**(3):341-4.

6. Salomon R, Attie T, Pelet A, Bidaud C, Eng C, Amiel J, Sarnacki S, Goulet O, Ricour C, Nihoul-Fekete C, Munnich A, Lyonnet S. Germline mutations of the RET ligand GDNF are not sufficient to cause Hirschsprung disease. *Nat Genet* 1996;**14**(3):345-7.

7. Puffenberger EG, Hosoda K, Washington SS, Nakao K, deWit D, Yanagisawa M, Chakravart A. A missense mutation of the endothelin-B receptor gene in multigenic Hirschsprung's disease. *Cell* 1994;**79**(7):1257-66.

8. Hofstra RM, Osinga J, Tan-Sindhunata G, Wu Y, Kamsteeg EJ, Stulp RP, van Ravenswaaij-Arts C, Majoor-Krakauer D, Angrist M, Chakravarti A, Meijers C, Buys CH. 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**(4):445-7.

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

10. Hofstra RM, Valdenaire O, Arch E, Osinga J, Kroes H, Loffler BM, Hamosh A, Meijers C, Buys CH. 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**(1):304-8.

11. Pingault V, Bondurand N, Kuhlbrodt K, Goerich DE, Prehu MO, Puliti A, Herbarth B, Hermans-Borgmeyer I, Legius E, Matthijs G, Amiel J, Lyonnet S, Ceccherini I, Romeo G, Smith JC, Read AP, Wegner M, Goossens M. SOX10 mutations in patients with Waardenburg-Hirschsprung disease. *Nat Genet* 1998;**18**(2):171-3.

12. Wakamatsu N, Yamada Y, Yamada K, Ono T, Nomura N, Taniguchi H, Kitoh H, Mutoh N, Yamanaka T, Mushiake K, Kato K, Sonta S, Nagaya M. Mutations in SIP1, encoding Smad interacting protein-1, cause a form of Hirschsprung disease. *Nat Genet* 2001;**27**(4):369-70.

13. Doray B, Salomon R, Amiel J, Pelet A, Touraine R, Billaud M, Attie T, Bachy B, Munnich A, Lyonnet S. Mutation of the RET ligand, neurturin, supports multigenic inheritance in Hirschsprung disease. *Hum Mol Genet* 1998;**7**(9):1449-52.

14. Amiel J, Laudier B, Attie-Bitach T, Trang H, de Pontual L, Gener B, Trochet D, Etchevers H, Ray P, Simonneau M, Vekemans M, Munnich A, Gaultier C, Lyonnet S. Polyalanine expansion and frameshift mutations of the paired-like homeobox gene PHOX2B in congenital central hypoventilation syndrome. *Nat Genet* 2003;**33**(4):459-61.

15. Brooks AS, Bertoli-Avella AM, Burzynski GM, Breedveld GJ, Osinga J, Boven LG, Hurst JA, Mancini GM, Lequin MH, de Coo RF, Matera I, de Graaff E, Meijers C, Willems PJ, Tibboel D, Oostra BA, Hofstra RM. Homozygous Nonsense Mutations in KIAA1279 Are Associated with Malformations of the Central and Enteric Nervous Systems. *Am J Hum Genet* 2005;**77**(1):120-6.

16. Angrist M, Bolk S, Thiel B, Puffenberger EG, Hofstra RM, Buys CH, Cass DT, Chakravarti A. Mutation analysis of the RET receptor tyrosine kinase in Hirschsprung disease. *Hum Mol Genet* 1995;**4**(5):821-30.

17. Attie T, Pelet A, Edery P, Eng C, Mulligan LM, Amiel J, Boutrand L, Beldjord C, Nihoul-Fekete C, Munnich A, et al. Diversity of RET proto-oncogene mutations in familial and sporadic Hirschsprung disease. *Hum Mol Genet* 1995;**4**(8):1381-6.

18. Seri M, Yin L, Barone V, Bolino A, Celli I, Bocciardi R, Pasini B, Ceccherini I, Lerone M, Kristoffersson U, Larsson LT, Casasa JM, Cass DT, Abramowicz MJ, Vanderwinden JM, Kravcenkiene I, Baric I, Silengo M, Martucciello G, Romeo G. Frequency of RET mutations in long- and short-segment Hirschsprung disease. *Hum Mutat* 1997;**9**(3):243-9.

19. Hofstra RM, Wu Y, Stulp RP, Elfferich P, Osinga J, Maas SM, Siderius L, Brooks AS, vd Ende JJ, Heydendael VM, Severijnen RS, Bax KM, Meijers C, Buys CH. RET and GDNF gene scanning in Hirschsprung patients using two dual denaturing gel systems. *Hum Mutat* 2000;**15**(5):418-29.

20. Bolk S, Pelet A, Hofstra RM, Angrist M, Salomon R, Croaker D, Buys CH, Lyonnet S, Chakravarti A. A human model for multigenic inheritance: phenotypic expression in Hirschsprung disease requires both the RET gene and a new 9q31 locus. *Proc Natl Acad Sci U S A* 2000;**97**(1):268-73.

21. Borrego S, Wright FA, Fernandez RM, Williams N, Lopez-Alonso M, Davuluri R, Antinolo G, Eng C. A founding locus within the RET proto-oncogene may account for a

large proportion of apparently sporadic Hirschsprung disease and a subset of cases of sporadic medullary thyroid carcinoma. *Am J Hum Genet* 2003;**72**(1):88-100.

22. Fitze G, Schierz M, Kuhlisch E, Schreiber M, Ziegler A, Roesner D, Schackert HK. Novel intronic polymorphisms in the RET proto-oncogene and their association with Hirschsprung disease. *Hum Mutat* 2003;**22**(2):177.

23. Sancandi M, Griseri P, Pesce B, Patrone G, Puppo F, Lerone M, Martucciello G, Romeo G, Ravazzolo R, Devoto M, Ceccherini I. Single nucleotide polymorphic alleles in the 5' region of the RET proto-oncogene define a risk haplotype in Hirschsprung's disease. *J Med Genet* 2003;**40**(9):714-8.

24. Griseri P, Bachetti T, Puppo F, Lantieri F, Ravazzolo R, Devoto M, Ceccherini I. 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**(2):189-95.

25. Burzynski GM, Nolte IM, Osinga J, Ceccherini I, Twigt B, Maas S, Brooks A, Verheij J, Plaza Menacho I, Buys CH, Hofstra RM. 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**(8):604-12.

26. Burzynski GM, Nolte IM, Bronda A, Bos KK, Osinga J, Plaza Menacho I, Twigt B, Maas S, Brooks AS, Verheij JB, Buys CH, Hofstra RM. Identifying candidate Hirschsprung disease-associated RET variants. *Am J Hum Genet* 2005;**76**(5):850-8.

27. Emison ES, McCallion AS, Kashuk CS, Bush RT, Grice E, Lin S, Portnoy ME, Cutler DJ, Green ED, Chakravarti A. A common sex-dependent mutation in a RET enhancer underlies Hirschsprung disease risk. *Nature* 2005;**434**(7035):857-63.

28. Pelet A, de Pontual L, Clement-Ziza M, Salomon R, Mugnier C, Matsuda F, Lathrop M, Munnich A, Feingold J, Lyonnet S, Abel L, Amiel J. Homozygosity for a frequent and weakly penetrant predisposing allele at the RET locus in sporadic Hirschsprung disease. *J Med Genet* 2005;**42**(3):e18.

29. Fernandez RM, Boru G, Pecina A, Jones K, Lopez-Alonso M, Antinolo G, Borrego S, Eng C. Ancestral RET haplotype associated with Hirschsprung's disease shows linkage disequilibrium breakpoint at -1249. *J Med Genet* 2005;**42**(4):322-7.

30. Gabriel SB, Salomon R, Pelet A, Angrist M, Amiel J, Fornage M, Attie-Bitach T, Olson JM, Hofstra R, Buys C, Steffann J, Munnich A, Lyonnet S, Chakravarti A. Segregation at three loci explains familial and population risk in Hirschsprung disease. *Nat Genet* 2002;**31**(1):89-93.

31. Carrasquillo MM, McCallion AS, Puffenberger EG, Kashuk CS, Nouri N, Chakravarti A. Genome-wide association study and mouse model identify interaction between RET and EDNRB pathways in Hirschsprung disease. *Nat Genet* 2002;**32**(2):237-44.

32. Chakravarti A. Hirschsprung disease. In: Scriver E, ed. *The metabolic and molecular bases of inherited disease.,* 8 ed. New York: McGraw-Hill, 2001;6231-6255.

33. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988;**16**(3):1215.

34. Mukhopadhyay N, Almasy L, Schroeder M, Mulvihill WP, Weeks DE. Mega2: data-handling for facilitating genetic linkage and association analyses. *Bioinformatics* 2005;**21**(10):2556-7.

35. Weeks DE, Ott J. Risk calculations under heterogeneity. *Am J Hum Genet* 1989;**45**(5):819-21.

36. Sobel E, Lange K. Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. *Am J Hum Genet* 1996;**58**(6):1323-37.

37. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995;**11**(3):241-7.

38. Nadeau JH. Modifier genes and protective alleles in humans and mice. *Curr Opin Genet Dev* 2003;**13**(3):290-5.

39. Owens SE, Broman KW, Wiltshire T, Elmore JB, Bradley KM, Smith JR, Southard-Smith EM. Genome-wide linkage identifies novel modifier loci of aganglionosis in the Sox10Dom model of Hirschsprung disease. *Hum Mol Genet* 2005;**14**(11):1549-58.

40. Cantrell VA, Owens SE, Chandler RL, Airey DC, Bradley KM, Smith JR, Southard-Smith EM. Interactions between Sox10 and EdnrB modulate penetrance and severity of aganglionosis in the Sox10Dom mouse model of Hirschsprung disease. *Hum Mol Genet* 2004;**13**(19):2289-301.

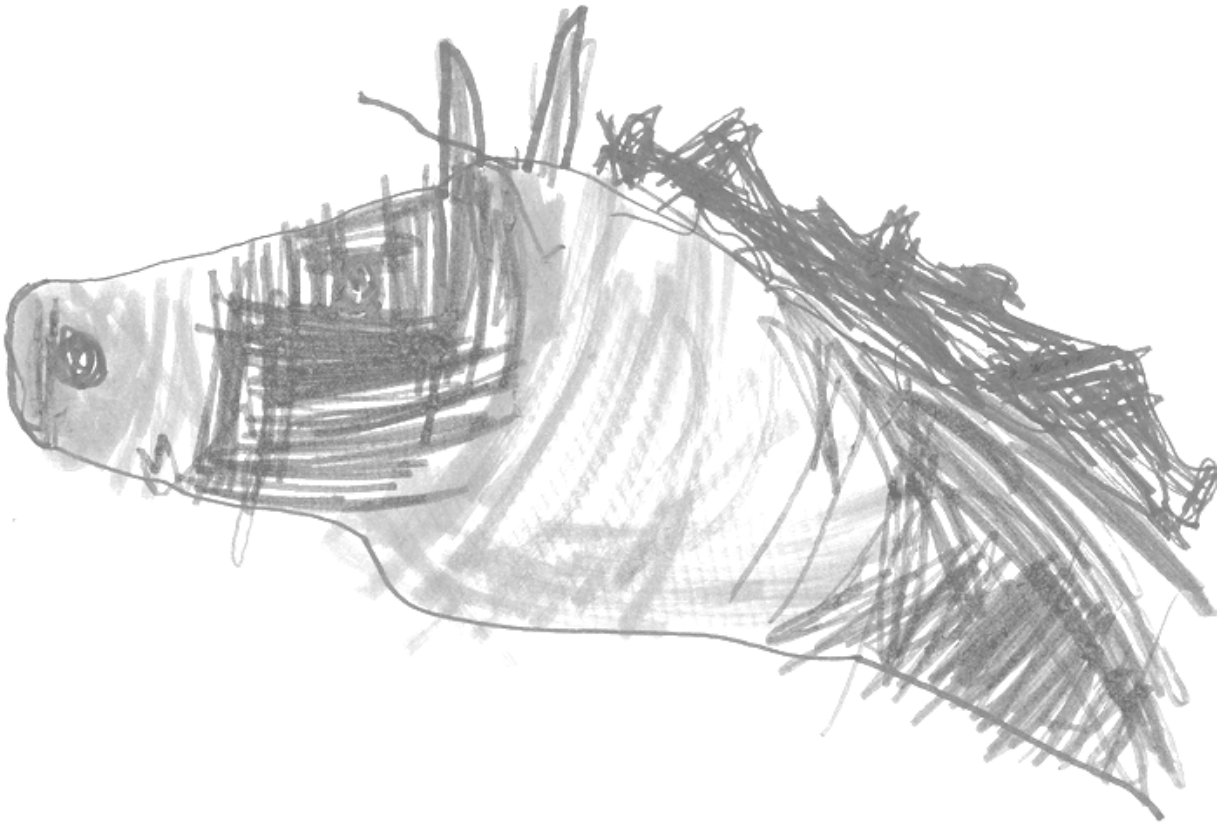
41. Mariani M, Baldessari D, Francisconi S, Viggiano L, Rocchi M, Zappavigna V, Malgaretti N, Consalez GG. Two murine and human homologs of mab-21, a cell fate determination gene involved in Caenorhabditis elegans neural development. *Hum Mol Genet* 1999;**8**(13):2397-406.

42. Amiel J, Espinosa-Parrilla Y, Steffann J, Gosset P, Pelet A, Prieur M, Boute O, Choiset A, Lacombe D, Philip N, Le Merrer M, Tanaka H, Till M, Touraine R, Toutain A, Vekemans M, Munnich A, Lyonnet S. Large-scale deletions and SMADIP1 truncating mutations in syndromic Hirschsprung disease with involvement of midline structures. *Am J Hum Genet* 2001;**69**(6):1370-7.

43. Thorsell A, Heilig M. Diverse functions of neuropeptide Y revealed using genetically modified animals. *Neuropeptides* 2002;**36**(2-3):182-93.

44. Chang JT, Esumi N, Moore K, Li Y, Zhang S, Chew C, Goodman B, Rattner A, Moody S, Stetten G, Campochiaro PA, Zack DJ. Cloning and characterization of a secreted frizzled-related protein that is expressed by the retinal pigment epithelium. *Hum Mol Genet* 1999;**8**(4):575-83.

45. Wu J, Saint-Jeannet JP, Klein PS. Wnt-frizzled signaling in neural crest formation. *Trends Neurosci* 2003;**26**(1):40-5.



# Appendix 4

## RET mediated gene expression of ENS precursors

Grzegorz M. Burzynski[1], Anita Wiersema[2], Bert Dontje[2], Frans Gerbens[1], Gerald de Haan[2], Robert M. W. Hofstra[1]

1. Department of Genetics, University Medical Center Groningen, Groningen, The Netherlands
2. Department of Stem Cell Biology, University Medical Center Groningen, Groningen, The Netherlands

**Manuscript in preparation**

## Abstract

Signalling of the RET receptor triggers one of the most important signal transduction route, crucial for the migration, proliferation and differentiation of the neural crest stem cells (NCSC's), which are responsible for the formation of the enteric nervous system (ENS). We isolated NCSC's from E.12-E.14 mouse embryo intestines and cultured them in selection media with or without GDNF, the ligand of RET. To identify genes involved in/or triggered by the RET pathway, we performed gene expression profiling using Agilent Whole Mouse Genome Arrays. Not only would this give information on GDNF-induced RET signalling in NCSC's, it might also be helpful in identifying genes involved in the development of Hirschsprung disease (HSCR). Several, important signalling pathways for ENS and, more generally, embryonic development were found to be significantly, differentially regulated by Gdnf stimulation. Genes belonging to the Notch and Wnt routes were found to be downregulated upon Gdnf stimulation. Furthermore, Gdnf seemed to negatively regulate routes promoting differentiation of the ENS progenitors like the retinoic acid (RA) pathway and semaphorins signalling. Conversely, genes and pathways that have been previously shown to stimulate proliferation and/or migration of NCSC's were found to be upregulated in cells treated with Gdnf. Among them the FGF pathway, forkhead transcription factors and the BMP signalling route. From the previously identified genes in HSCR, only the Ret co-receptor Gfr$\alpha$1 was found to be down-regulated by Gdnf. We also identified 3 genes, differentially regulated upon Gdnf stimulation, which are located in regions previously shown to harbour HSCR susceptibility genes (3p21 and 9q31). In sum, we provide evidence that Gdnf stimulation support the maintenance of pluripotency, proliferation and migration of the NCSC's. Further, we provide a profile of genes' expression in the ENS precursors, regulated by the RET signalling and identified possible candidate HSCR predisposing genes.

## Introduction

The neural crest (NC) is formed at the time of neural tube closure, at the border of the neural plate and non-neural ectoderm. Shortly after the neural tube closure is complete, neural crest cells undergo an epithelial to mesenchymal transition, they delaminate from the neuroepithelium and migrate along various pathways in the embryo to form distinct neural crest tissue derivatives (Barembaum and Fraser, 2005). A predominant region for ENS formation is the vagal region, corresponding to somites 1-7, giving rise to the vast majority of parasympathetic neurons and glia cells of the ENS. Vagal crest cells upon delamination (in mouse E.8.5-9.0) migrate ventrally to the region posterior to the branchial arches, close to the dorsal aorta. Subsequently cells enter the foregut mesenchyme, migrate caudally and eventually colonize the whole length of the GI tract (Durbec et al., 1996, Kapur et al., 1992). In mouse, vagal crest cells enter the foregut at E.9.0-E9.5, and by E11.5 the whole small intestine is colonized. The entire GI tract is colonized by E13.5 (Durbec et al., 1996; Kapur et al., 1992; Young et al., 1998). Respectively in humans, ENS precursors invade the foregut at embryonic week 4 and reach the terminal hindgut at week 7 (Wallace and Burns, 2005). ENS precursors form the myenteric and submucosal plexuses, of which the myenteric plexus arises earlier than the submucosal plexus. Ganglia develop gradually from the foregut to the hindgut region, what resembles the vagal cells migration wave. Submucosal ganglia are formed later by cells which migrate transversely from the myenteric region towards the gut lumen (Gershon et al., 1994).

This process of proliferation and migration of NCSC's from the dorsal neural tube towards the foregut and later colonization of the whole GI tract, followed by differentiation into ENS neurons and glia cells is a very complex procedure. Timing and harmonious action of growth factors, their receptors, signalling effectors and transcription factors is essential for each stage of ENS

development (Taraviras and Pachnis, 1999, Young et al., 2001). Important during the first steps in the induction and delamination of neural crest is signalling of the WNT, NOTCH, BMP, FGF and Retinoic Acid (RA) signalling routes (La Bonne and Bronner-Fraser, 1998; Cornell and Eisen et al., 2005; Villanueva et al., 2002; Garcia-Castro et al., 2002). Subsequently, neural crest cells acquire distinct fate, depending on the region of origin and upon migration to different embryonic regions where surrounding tissue express appropriate instructive signals.

The RET protein with GFRα's as co-receptors, and GDNF as its ligand are the major components of the most important signalling cascade in ENS development (Durbec et al., 1996; Taraviras et al., 1999). GDNF serves as a chemoattractant for the migrating vagal cells that express RET towards and within the gut. High levels of GDNF expression are observed in the stomach mesenchyme ahead of the migrating NCSC's and later in the cecum region assuring proper course and successful colonization of the small intestine (Natarajan et al. 2002, Young et al., 2001). However, in the GI tract regions posterior to the cecum, no GDNF expression ahead of NCSC's cells was detected. Instead, GDNF is expressed alongside the streams of the migrating cells and most likely exert proliferative and survival rather than having a migratory effect (Natarajan et al., 2002).

Defects in proper migration, proliferation or differentiation of the ENS progenitors can lead in humans to the most common disorder affecting enteric neurons – Hirschsprung disease (HSCR) [MIM 142623]. It is a congenital disorder characterized by intestinal obstruction due to an absence of enteric ganglia among variable lengths of the intestinal tract. HSCR is genetically a complex, both heterogenic and non-Mendelian disease, and to date 10 genes have been shown to be involved in the disease etiology (see first chapter). Nevertheless, the RET gene is a major disease risk factor as the majority of the mutations are being found either in the RET coding region or in its regulatory sequences (Chakravarti and Lyonnet, 2001; Emison

et al., 2005). Other genes and susceptibility loci, which were found to be involved in HSCR, are believed to belong to the RET signalling cascades or are believed to interact with the RET route.

As mentioned, RET signalling is extremely important in NCSC migration proliferation and differentiation, nevertheless the network that transmits RET signals within the cell and more importantly the actual downstream effectors are largely unknown. Similarly, understanding the interconnections between different signalling cascades is poorly understood. In this study we tried to get more insight into the RET network and downstream effectors by the isolation of migratory vagal neural crest cells from mice embryo intestines. We have compared gene expression profiles of cells where RET signalling was triggered by Gdnf to those where RET signalling was inactive. By this approach we hoped to identify genes belonging to or interfering with RET signalling pathway, also those important in HSCR. We make an overview of the genes, whose expression in ENS progenitors was significantly influenced by Gdnf stimulation.

## Material and Methods

*Isolation of NCSC's and culture conditions*

Anesthetized pregnant female C57BL/6 mice were sacrificed by cervical dislocation; embryos at E11-14 were removed and transferred to PBS solution. Mouse embryonic guts were dissected after washing in a 0,1% trypsin EDTA solution, followed by gentle mechanical dissociation using a 26G needle. Cell suspensions were then cultured in 6 well plates coated with fibronectin. Spheres appeared after 5-6 days of culture (Fig. 1a). They were isolated, mechanically dissociated and passaged to T25 culture flasks. Cells were grown in SFEM-medium, which was supplemented with 1% penicillin-streptomycin, 1% N2-supplement, 2% B27 supplement, 15% Chicken Embryo Extract (CEE), 20ng/ml FGF-2 20ng/ml IGF I and 20ng/ml EGF in a humidified atmosphere with 5% $CO_2$ incubator at 37°C. Culture medium was changed twice a week, spheres were mechanically dissociated and passaged.

*Immunocytochemistry*

Immunocytochemistry was performed on differentiated and proliferating cells. Cells were washed in PBS and fixed with 3,7% paraformaldehyde, washed once more with PBS, and incubated with 7,5% BSA (Invitrogen) for 30 minutes. Cells were then incubated with primary antibodies for 2 hours at 37°C or overnight at 4°C. Primary antibodies that were used were directed against Nestin (1:200, Chemicon), Nerve Growth Factor (p75) Receptor (1:300, ATSBIO), Sox10 (1:200, Sigma), Mash1 (1:300, CeMines), Ret H-300 (1:200, Santa Cruz), Ret (P)Y1062 (1:400, Santa Cruz), Glial Fibrillary Acidic Protein (GFAP, Invitrogen). Cells were washed and incubated with

appropriate fluorescently labeled secondary antibodies (Jackson Labs, West Grove, PA, USA) for 1 h at room temperature, followed by a DAPI staining (2mg/ml, Sigma) to visualize the nuclei. As a control for primary antibody staining tissue without primary antibody incubation was used. As a control for secondary antibody staining tissue without secondary antibody incubation was performed.

*GDNF stimulation and RNA isolation*

After 2 passages the addition of FGF-2, IGF-I and EGF growth factors was stopped. Instead, half of the culture was stimulated with Gdnf (50ng/ml). RNA was isolated after 48h using Qiagen RNeasy Mini Kit, accordingly to the protocol provided by the manufacturer. The RNA yield and purity were determined spectrophotometrically by measuring the absorbance at 260 nm and 280 nm. RNA integrity was checked on 1% agarose gels.

*RNA amplification, and gene expression profiling*

Linear amplification of mRNA was performed essentially according to a protocol of the Dutch Cancer Institute (www.nki.l.nl/nkidep/pa/microarray/protocols.html). Briefly, amplification started with first strand cDNA synthesis from 2 μg of total RNA, using Superscript II RT-polymerase (GIBCO) and a specific oligo(dT) primer containing a 17 bp T7 polymerase recognition site. After second strand synthesis, double-stranded cDNA was purified and the yield was determined spectrophotometrically. Transcription was performed with the T7 Megascript kit (Ambion) as described by the manufacturer and aminoallyl-UTP (Ambion) was incorporated as described by ‘t Hoen et al. (2003).

Custom DNA microarray experiments and RNA quality control were performed by ServiceXS, Netherlands using *in-situ* synthesized oligonucleotide Mouse Whole Genome Array provided by

Agilent Technologies. Data analysis was performed by using BRB ArrayTools developed by Dr. Richard Simon and Amy Peng Lam.

## Results

*Properties of the cells*

NCSC's were stained with a range of antibodies to confirm neural crest origin of the isolated cells, their ability to differentiate into the expected descendant cell types and importantly, their level of Ret expression and Gdnf responsiveness (Fig. 1). Immunostainings were carried out at several different stages of the cultures (day 2, 5 and 10). The protein that primarily served as a biomarker for neural crest origin was Nerve Growth Factor (p75) Receptor, which is expressed by all neural crest cell populations. All cells in cultures were expressing the p75 receptor (Fig. 1b). Additionally, all cells do express Sox10, Mash1 and Ret verifying the neural crest origin of the cells (Fig 1c-f). Sox10 is expressed in neural crest already at the very early stages of cell delamination from the neural tube and almost 100% of cultured cells were Sox10 positive (Fig. 1c). Mash1 is a transcription factor that is often used as an early neuronal marker and is expressed by neural crest stem cells as well as by differentiating neurons. Similarly to Sox10, Mash1 expression was nearly 100% in each stage of the culture (Fig. 1d). Not all of the cells were positive for Ret, especially at the very early stages of the cultures (day 1-2) Ret expression was weak (Fig. 1e). Nevertheless, later immunostainings, at days 5 or 10, revealed Ret expression in virtually all cells (Fig 1f).

Differentiating capacities of cultured NCSC's were tested by Nestin and GFAP stainings (Fig. 1 g-j). Cells cultured in the complete medium with all the growth factors present (FGF-2, IGF-I and EGF) remained largely undifferentiated (~80%) even after 15 days of culturing. However when

growth factors addition was stopped, the differentiation was increasing. Moreover, the addition of Gdnf seemed to arrest cell differentiation towards the glia cells, based on visual observations. The majority of cells (~80%) in complete medium differentiated towards neuronal fate (Fig. 1g, h), and around 20% was expressing glia cells marker GFAP (Fig. 1i, j). Still, cultured NCSC's showed bipotent differentiation, which confirmed their proper origin. Further, double immuno-stainings with Ret and either Nestin or GFAP showed that differentiating or differentiated cells remained Ret positive (Fig. 1h, j).

Lastly, immunocytochemistry with an antibody raised against phosphorylated Ret tyrosine 1062 were carried out to check Ret protein responsiveness to Gdnf stimulation (Fig. 1k, l). Cells from cultures stimulated and unstimulated with Gdnf were fixed and stained after 48h. Cells stimulated with Gdnf showed a high rate of Ret Y1062 phosphorylation (~70% of the cells), whereas phosphorylated Ret was present only in around 30% of unstimulated cells (Fig. 1k, l).

*Gene expression profiling*

Gene expression profiling revealed a great number of genes differentially regulated depending on Gdnf stimulation. We set a 3-fold expression difference as a threshold of significance. In this condition the number of downregulated genes (664) is approximately double to the number of upregulated genes (378). Lists of differentially regulated genes, their accession numbers, a short description and the expression fold differences can be found in the supplement, in tables 1 and 2.

## Discussion

Many genes encoding proteins belonging to signalling pathways previously implicated in ENS development as well as proteins known to be involved in embryonic development, in general, such as the Notch and Wnt signalling pathways, were found to be differentially regulated in the NCSCs upon Gdnf stimulation (Tab. 1).

The Notch1 and Notch signalling pathway components Hey1 and Hes1 were both significantly down-regulated in cells stimulated with Gdnf, though Notch ligand - Delta (Dlk1) was highly expressed. Hey1 and Hes1 both are downstream signalling targets of Notch molecules and belong to basic helix-loop-helix (bHLH) transcription factors implicated in cell fate decision and boundary formation, and proven crucial in neural crest region formation. (Cornell et al., 2005). This downregulation might indeed be Gdnf dependent as Gdnf signalling acts as a chemoattractor for neural crest cells and cause their directional migration towards the gut. It might down-regulate early transduction cascades in NCSC's and initiate the differentiation process of neural crest cells into ENS cells. On the other hand, over-expression of the Notch ligand – Dlk1 can possibly be explained by the fact that Delta proteins are membrane-bound ligands expressed usually by neighboring cells, which triggering Notch signalling in targeted cells (Artavanis-Tsakonas et al., 1999). It can be hypothesized that migrating ENS precursors exert an influence on surrounding tissue and their companions, and promote specific secondary fates, including glial differentiation and neural fate inhibition (Cornell et al., 2005).

With respect to the Wnt pathway, Wnt2, Fzd1 (frizzled, Wnt receptor) and the Wnt antagonists Frzb (frizzled related) and Sfrp1 (secreted frizzled related) were all down-regulated. In contrast, the Tle1, groucho co-repressor of Tcf/Lef factors, which are downstream targets of Wnt signalling, was significantly up-regulated. Wnt signalling, similarly as Notch, is very important in

the early stages of neural crest development and plays a role in the induction, delamination and early differentiation of the neural crest (Schmidt et al., 2005). The Ret pathway possibly silence the early signalling routes and command cells to move towards the gut and to differentiate into appropriate cell lineages. Additionally, it has been previously shown that the expression of Sox10 transcription factor is controlled by Wnt proteins (Honore et al., 2003). Sox10 was shown to promote survival of NC precursors, stimulate differentiation of these cells into melanocytes and glia cells (Mollaaghababa and Pavan et al., 2003) and inhibits neuronal differentiation (Kim J et al., 2003). These facts can also explain the down-regulation of Wnt pathway by Gdnf.

We did not identify any differential expression of Bmp molecules or their receptors. However, Bmp antagonists like gremlin and Bmper were strongly downregulated in NCSC's treated with Gdnf (Tab. 1). This might indicate that Gdnf allows for moderate Bmp expression and signalling. This would be consistent with Wnts down-regulation as usually Bmps exert opposing effects on neural crest cells when compared to Wnt molecules (Kleber et al., 2005). Furthermore, it is documented that Bmp proteins (e.g. Bmp4) are able to promote migration of ENS progenitors within the gut, synergistically with Gdnf (Goldstein et al., 2005)

We also observed a down-regulation of the Retinoic Acid (RA) receptor -RARβ and Crabp1 gene, which encodes a cellular RA binding protein. It is know that RA can promote differentiation of neural crest cells and Gdnf appears to prevent this (Hansford and Marshall, 2003). RA has also shown to have an opposing effect on FGF signalling (Mao and Lee, 2005; Diez del Corral and Storey, 2004) and in agreement with this we observed up-regulation of Fgf2, Fgf7 and the receptor Fgfrl1. In addition, genes encoding proteins Hmga1 and Hmga2, which are known regulators of active chromatin, mainly expressed in embryonic tissues, were up-regulated (Tab. 1). These genes have been previously shown to abrogate RA signalling in neuroblastoma cell lines and prevent cells differentiation and growth inhibition (Giannini et al., 1999).

Furthermore, several genes belonging to the forkhead box family of transcription factors (Foxc1, Foxc2, Foxf2 and Foxg1), which participate in specification of the different mesodermal subpopulations that arise during gastrulation (Kaestner et al., 2000), were highly up-regulated (Tab. 1). Foxg1 promotes proliferation and inhibits differentiation. In the study by Martynoga et al. (2005) Foxg1 was shown to control telencephalic precursor neurogenesis proliferation via regulation of Fgf signalling and differentiation via regulation of Bmp signalling. Moreover, it can interact with Tle1, which was also found to be up-regulated in this data set, and synergistically repress transcription of the target genes (Marcal et al., 2005).

Foxc1 and Foxc2 have been shown to be involved in kidney development and to affect expression of Gdnf in the intermediate mesoderm (Kume et al., 2000).

The most interesting gene we found in this group is Foxf2. This gene has previously been shown to be expressed in neural crest cells and can regulate the proliferation and survival of the gut epithelium together with Foxf1 (Ormestad M et al., 2005). Foxf2, similarly as Foxf1 is expressed in the splanchnic mesoderm, but in contrast Foxf2 is also expressed in the neural crest cells (Wang et al., 2003; Ormestad et al., 2004). This fact can explain that Foxf1 expression was not detected in the gene expression profiling performed in this study. Further, expression of Foxf2 is regulated by endodermal hedgehog ligands - Ihh or Shh and Foxf2 increases Bmp4 production. Importantly, reduced dosage of Foxf2 can cause aganglionic megacolon and might be also involved in HSCR etiology in humans (Ormestad et al., 2005)

Of particular interest we found that several components of semaphorin signalling pathway, which guide migration and outgrowth of neurons were down-regulated in Gdnf treated NCSC's including semaphorin proteins sema3e, sema3d, their receptors and co-receptors, Nrp1, Plxnc1, Plxdc2 as well as intracellular mediator Crmp1 (Tab. 1). The Class III Sempahorin molecules are mostly known from their chemorepulsive properties (Kruger et al., 2005). Sema3e

and Sema3d were shown to inhibit the growth of retinal axons (Steinbach et al., 2002; Sakai and Halloran, 2006) Our data suggests Gdnf regulates negatively this pathway expression, which might indicate Gdnf-dependant neural fate differentiation of NCSC's.

Furthermore, we have found three genes – *Svep1*, *Klf4* and *Arhgef3*, which have human homologues that map to the previously described susceptibility loci on 3p21 and 9q31, which has been found in linkage studies performed on HSCR patients (Tab. 1) (Bolk et al., 2000; Gabriel et al., 2002). *Svep1* and *Klf4* are genes from the chromosome 9 region that were significantly down and up-regulated in cells treated with Gdnf, respectively. Svep1 is a newly described cell surface CAM protein, probably involved in initial cell adhesion process (Shur et al., 2006). Pre-blocking of Svep1 with the antibodies caused decreased attachment of the MBA-15 cells in culture (Shur et al., 2006). These findings might indicate pro-migratory effect of Gdnf stimulation. Klf4 (Kruppel like factor 4) is a transcription factor, highly expressed in the gut and to date it was mainly described as a potential tumor suppressor gene (Zhao et al., 2004). Nevertheless, its presence in the intestine might indicate an important role of Klf4 also in the ENS precursors. *Arhgef3*, which was down-regulated, encodes Rho guanine nucleotide exchange factor 3 and may form a complex with G proteins and stimulate Rho-dependent signalling resulting in migration, adhesion or axon growth (Bloch-Gallego et al., 2005), hence might very well be important for ENS development. These genes due to their localization and Ret related expression, might be considered as a good candidates genes in Hirschsprung disease.

None of the previously described genes involved in Hirschsprung disease (except one) were found to be significantly, differentially regulated in this data set. Surprisingly we did find that the Ret co-receptor - Gfrα1 had reduced expression in cells stimulated with Gdnf suggesting its expression is negatively regulated by the Ret signalling pathway. It can be hypothesized that

negative feedback loops exists in Ret signalling cascade and that Ret is able to actively down-regulate at least some of the components of the pathway.

In summary, there is an apparent, general trend in our data that genes that encode proteins that stimulate differentiation of the ENS progenitors and inhibit their migration and proliferation like RA signalling molecules, semaphorins or antagonists of Bmp proteins are generally down-regulated in NCSCs, which are stimulated by Gdnf. While genes that have the opposite effect like Fgf factors and receptors, Fox transcription factors, Hmga chromatin activators were up-regulated by Gdnf.

These findings resemble the responses of the ENS precursors in the early migratory stage, perhaps in the beginnings of gut colonization when Gdnf serves as a chemoattractant, stimulator of proliferation and migration.

**Table 1. List of genes differentially expressed in NCSCs upon Gdnf stimulation, belonging to signalling pathways important in ENS development**

| | Signalling pathway | Symbol | GeneBank accession | Description | Fold difference |
|---|---|---|---|---|---|
| 1 | Notch | Notch1 | Z11886 | Mmusculus notch-1 mRNA | ↓* 4,31 |
| 2 | | Hey1 | NM_010423 | hairy/enhancer-of-split related with YRPW motif 1 (Hey1), mRNA | ↓ 6,07 |
| 3 | | Hes1 | NM_008235 | hairy and enhancer of split 1 (Drosophila) (Hes1), mRNA | ↓ 3,25 |
| 4 | | Dlk1 | NM_010052 | delta-like 1 homolog (Drosophila) (Dlk1), mRNA | ↑* 18,83 |
| 5 | Wnt | Wnt2 | NM_023653 | wingless-related MMTV integration site 2 (Wnt2), mRNA | ↓ 4,00 |
| 6 | | Fzd1 | NM_021457 | frizzled homolog 1 (Drosophila) (Fzd1), mRNA | ↓ 4,19 |
| 7 | | Frzb | NM_011356 | frizzled-related protein (Frzb), mRNA | ↓ 3,20 |
| 8 | | Sfrp1 | NM_013834 | secreted frizzled-related sequence protein 1 (Sfrp1), mRNA | ↓ 4,33 |
| 9 | | Tle1 | AY155196 | truncated groucho protein GRG1-S (Tle1) mRNA, complete cds; alternatively spliced | ↑ 5,64 |
| 10 | Bmp | gremlin | NM_011825 | gremlin 2 homolog, cysteine knot superfamily (Xenopus laevis) (Grem2), mRNA | ↑ 3,69 |
| 11 | | Bmper | NM_028472 | BMP-binding endothelial regulator (Bmper), mRNA | ↑ 3,67 |
| 12 | Retinoic Acid | Rarβ | NM_011243 | retinoic acid receptor, beta (Rarb), mRNA | ↓ 3,27 |
| 13 | | Crabp1 | NM_013496 | cellular retinoic acid binding protein I (Crabp1), mRNA | ↓ 5,01 |
| 14 | | Hmga1 | NM_016660 | high mobility group AT-hook 1 (Hmga1), mRNA | ↑ 5,02 |
| 15 | | Hmga2 | NM_010441 | high mobility group AT-hook 2 (Hmga2), mRNA | ↑ 4,77 |
| 16 | Fgf | Fgf2 | NM_008006 | fibroblast growth factor 2 (Fgf2), mRNA | ↑ 31,06 |
| 17 | | Fgf7 | NM_008008 | fibroblast growth factor 7 (Fgf7), mRNA | ↑ 22,21 |
| 18 | | Fgfrl1 | NM_054071 | fibroblast growth factor receptor-like 1 (Fgfrl1), mRNA | ↑ 6,18 |
| 19 | Fox factors | Foxc1 | NM_008592 | forkhead box C1 (Foxc1), mRNA | ↑ 3,16 |
| 20 | | Foxc2 | NM_013519 | forkhead box C2 (Foxc2), mRNA | ↑ 6,30 |
| 21 | | Foxg1 | NM_008241 | forkhead box G1 (Foxg1), mRNA | ↑ 23,87 |
| 22 | | Foxf2 | NM_010225 | forkhead box F2 (Foxf2), mRNA | ↑ 10,05 |
| 23 | | Foxg1 | NM_008241 | forkhead box G1 (Foxg1), mRNA | ↑ 23,87 |
| 24 | semaphorins | Sema3e | NM_011348 | sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3E (Sema3e), mRNA | ↓ 6,93 |
| 25 | | Sema3d | NM_028882 | sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3D (Sema3d), mRNA | ↓ 4,63 |
| 26 | | Nrp1 | NM_008737 | neuropilin 1 (Nrp1), mRNA | ↓ 4,10 |
| 27 | | Plxnc1 | NM_018797 | plexin C1 (Plxnc1), mRNA | ↓ 4,25 |
| 28 | | Plxdc2 | NM_026162 | plexin domain containing 2 (Plxdc2), mRNA | ↓ 8,37 |
| 29 | | Crmp1 | NM_007765 | collapsin response mediator protein 1 (Crmp1), mRNA | ↓ 3,71 |
| 30 | HSCR susceptibility regions | Svep1 | NM_022814 | polydomain protein (Polydom), mRNA | ↓ 8,03 |
| 31 | | Klf4 | NM_010637 | Kruppel-like factor 4 (gut) (Klf4), mRNA | ↑ 5,47 |
| 32 | | Arhgef3 | NM_027871 | Rho guanine nucleotide exchange factor (GEF) 3 (Arhgef3), mRNA | ↓ 3,72 |

*↓ - indicates down-regulation upon Gdnf stimulation ↑ - indicates up-regulation upon Gdnf stimulation

## References

Artavanis-Tsakonas S, Rand MD, Lake RJ.
Notch signalling: cell fate control and signal integration in development.
Science. 1999 Apr 30;284(5415):770-6.

Barembaum M, Bronner-Fraser M.
Early steps in neural crest specification.
Semin Cell Dev Biol. 2005 Dec;16(6):642-6

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;97:268-73.

Chakravarti, A and Lyonnet S (2001). Hirschsprung's disease. In The Metabolic and Molecular Bases of Inherited Diseases (ed. C. R. Scriver, A. L. Beaudet, D. Valle, W. S. Sly, B. Childs, K. Kinzler and B. Vogelstein), pp. 6231-6255.
New York: McGraw-Hill.

Cornell RA, Eisen JS.
Notch in the pathway: the roles of Notch signalling in neural crest development.
Semin Cell Dev Biol. 2005 Dec;16(6):663-72

Diez del Corral R, Storey KG.
Opposing FGF and retinoid pathways: a signalling switch that controls differentiation and patterning onset in the extending vertebrate body axis.
Bioessays. 2004 Aug;26(8):857-69.

Durbec P, Marcos-Gutierrez CV, Kilkenny C, Grigoriou M, Wartiowaara K, Suvanto P, Smith D, Ponder B, Costantini F, Saarma M, et al.
GDNF signalling through the Ret receptor tyrosine kinase.
Nature. 1996 Jun 27;381(6585):789-93

Elworthy S, Pinto JP, Pettifer A, Cancela ML, Kelsh RN.
Phox2b function in the enteric nervous system is conserved in zebrafish and is sox10-dependent.
Mech Dev. 2005 May;122(5):659-69

Emison ES, McCallion AS, Kashuk CS, Bush RT, Grice E, Lin S, Portnoy ME, Cutler DJ, Green ED, Chakravarti A. A common sex-dependent mutation in a RET enhancer underlies Hirschsprung disease risk. Nature. 2005 Apr 14;434(7035):857-63.

Fu M, Lui VC, Sham MH, Pachnis V, Tam PK.
Sonic hedgehog regulates the proliferation, differentiation, and migration of enteric neural crest cells in gut.
J Cell Biol. 2004 Aug 30;166(5):673-84.

Gabriel SB, Salomon R, Pelet A, Angrist M, Amiel J, Fornage M, Attie-Bitach T, Olson JM, Hofstra R, Buys C, Steffann J, Munnich A, Lyonnet S, Chakravarti A
Segregation at three loci explains familial and population risk in Hirschsprung disease.
Nat Genet. 2002 31:89-93.

Garcia-Castro MI, Marcelle C, Bronner-Fraser M.
Ectodermal Wnt function as a neural crest inducer.
Science. 2002 Aug 2;297(5582):848-51

Gershon, M. D., Kirchgessner, A. L. and Wade, P. R. (1994). Functional
anatomy of the enteric nervous system. In Physiology of the Gastrointestinal
Tract, Vol. 1 (ed. L. R.Johnson), pp. 381-422. New York: Raven Press.

Giannini G, Di Marcotullio L, Ristori E, Zani M, Crescenzi M, Scarpa S, Piaggio G, Vacca A, Peverali FA, Diana F, Screpanti I, Frati L, Gulino A.
HMGI(Y) and HMGI-C genes are expressed in neuroblastoma cell lines and tumors and affect retinoic acid responsiveness.
Cancer Res. 1999 May 15;59(10):2484-92.

Sakai JA, Halloran MC.
Semaphorin 3d guides laterality of retinal ganglion cell projections in zebrafish. Development. 2006 Mar;133(6):1035-44.

Goldstein AM, Brewer KC, Doyle AM, Nagy N, Roberts DJ.
BMP signalling is necessary for neural crest cell migration and ganglion formation in the enteric nervous system.
Mech Dev. 2005 Jun;122(6):821-33.

Hansford LM, Marshall GM.
Glial cell line-derived neurotrophic factor (GDNF) family ligands reduce the sensitivity of neuroblastoma cells to pharmacologically induced cell death, growth arrest and differentiation.
Neurosci Lett. 2005 Dec 2;389(2):77-82.

't Hoen PA, de Kort F, van Ommen GJ, den Dunnen JT.
Fluorescent labelling of cRNA for microarray applications.
Nucleic Acids Res. 2003 Mar 1;31(5):e20

Honore SM, Aybar MJ, Mayor R.
Sox10 is required for the early development of the prospective neural crest in Xenopus embryos.
Dev Biol. 2003 Aug 1;260(1):79-96.

Kaestner KH, Knochel W, Martinez DE.
Unified nomenclature for the winged helix/forkhead transcription factors.
Genes Dev. 2000 Jan 15;14(2):142-6.

Kapur RP, Yost C, Palmiter RD.
A transgenic model for studying development of the enteric nervous system in normal and aganglionic mice.
Development. 1992 Sep;116(1):167-75

Kim J, Lo L, Dormand E, Anderson DJ.
SOX10 maintains multipotency and inhibits neuronal differentiation of neural crest stem cells. Neuron. 2003 Apr 10;38(1):17-31

Kleber M, Lee HY, Wurdak H, Buchstaller J, Riccomagno MM, Ittner LM, Suter U, Epstein DJ, Sommer L.
Neural crest stem cell maintenance by combinatorial Wnt and BMP signalling.
J Cell Biol. 2005 Apr 25;169(2):309-20.

Kruger GM, Mosher JT, Tsai YH, Yeager KJ, Iwashita T, Gariepy CE, Morrison SJ.
Temporally distinct requirements for endothelin receptor B in the generation and migration of gut neural crest stem cells.
Neuron. 2003 Dec 4;40(5):917-29.

Kruger RP, Aurandt J, Guan KL.
Semaphorins command cells to move.
Nat Rev Mol Cell Biol. 2005 Oct;6(10):789-800.

Kume T, Deng K, Hogan BL.
Murine forkhead/winged helix genes Foxc1 (Mf1) and Foxc2 (Mfh1) are required for the early organogenesis of the kidney and urinary tract.
Development. 2000 Apr;127(7):1387-95.

LaBonne C, Bronner-Fraser M.
Induction and patterning of the neural crest, a stem cell-like precursor population.
J Neurobiol. 1998 Aug;36(2):175-89

Lang D, Epstein JA.
Sox10 and Pax3 physically interact to mediate activation of a conserved c-RET enhancer. Hum Mol Genet. 2003 Apr 15;12(8):937-45

Mao Y, Lee AW.
A novel role for Gab2 in bFGF-mediated cell survival during retinoic acid-induced neuronal differentiation.
J Cell Biol. 2005 Jul 18;170(2):305-16.

Marcal N, Patel H, Dong Z, Belanger-Jasmin S, Hoffman B, Helgason CD, Dang J, Stifani S.
Antagonistic effects of Grg6 and Groucho/TLE on the transcription repression activity of brain factor 1/FoxG1 and cortical neuron differentiation.
Mol Cell Biol. 2005 Dec;25(24):10916-29.

Martynoga B, Morrison H, Price DJ, Mason JO.
Foxg1 is required for specification of ventral telencephalon and region-specific regulation of dorsal telencephalic precursor proliferation and apoptosis.
Dev Biol. 2005 Jul 1;283(1):113-27.

Mollaaghababa R, Pavan WJ.
The importance of having your SOX on: role of SOX10 in the development of neural crest-derived melanocytes and glia.
Oncogene. 2003 May 19;22(20):3024-34

Natarajan D, Marcos-Gutierrez C, Pachnis V, de Graaff E.
Requirement of signalling by receptor tyrosine kinase RET for the directed migration of enteric nervous system progenitor cells during mammalian embryogenesis.
Development. 2002 Nov;129(22):5151-60.

Ormestad M, Astorga J, Carlsson P.
Differences in the embryonic expression patterns of mouse Foxf1 and -2 match their distinct mutant phenotypes.
Dev Dyn. 2004 Feb;229(2):328-33.

Pattyn A, Morin X, Cremer H, Goridis C, Brunet JF.
The homeobox gene Phox2b is essential for the development of autonomic neural crest derivatives.
Nature. 1999 May 27;399(6734):366-70.

Schmidt C, Patel K.
Wnts and the neural crest.
Anat Embryol (Berl). 2005 Jun;209(5):349-55.

Steinbach K, Volkmer H, Schlosshauer B.
Semaphorin 3E/collapsin-5 inhibits growing retinal axons.
Exp Cell Res. 2002 Sep 10;279(1):52-61.

Taraviras S, Pachnis V.
Development of the mammalian enteric nervous system.
Curr Opin Genet Dev. 1999 Jun;9(3):321-7

Verheij J and Hofstra RM: Shah-Waardenburg' in "Molecular basis of inborn errors of development" (2004) edited by CJ Epstein, RP Erickson and A Wynshaw-Boris. Pp. 440-446

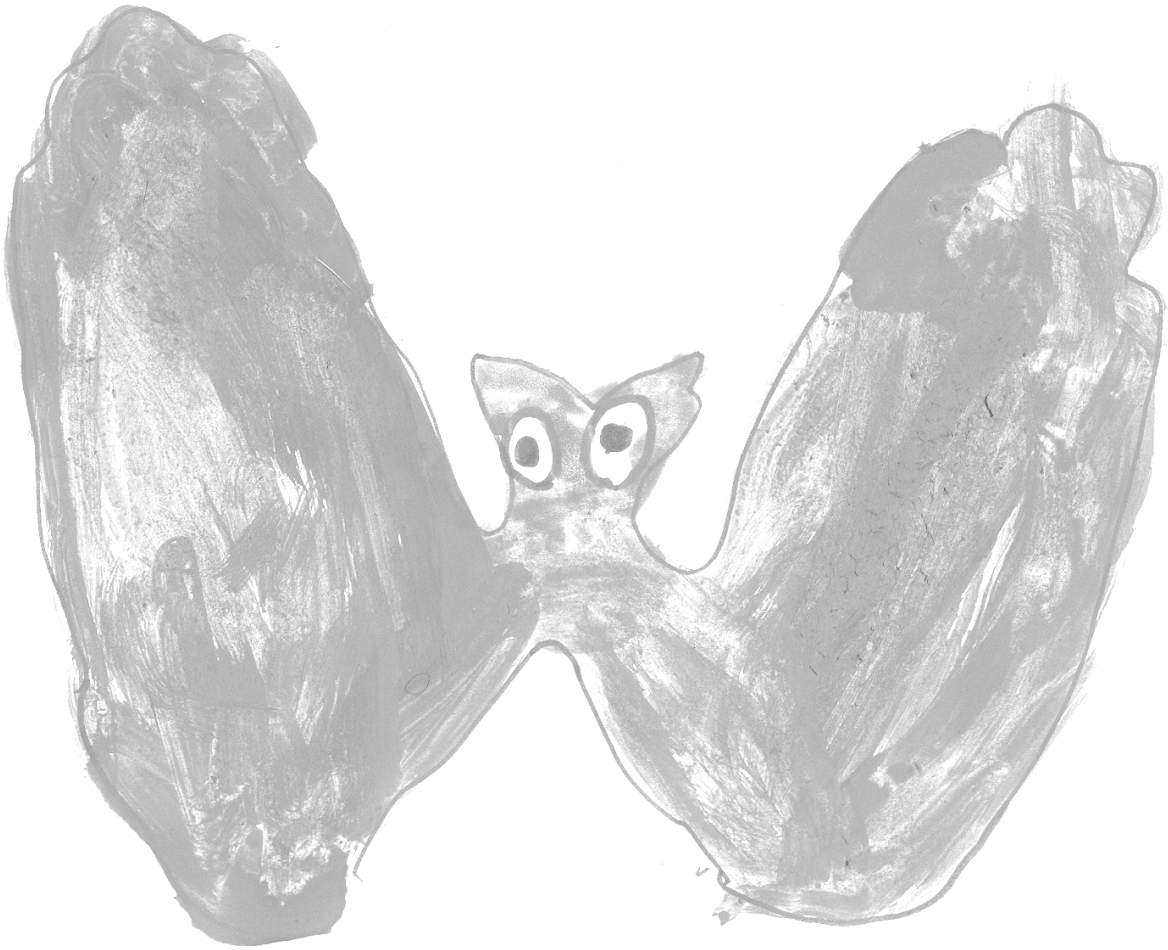
Villanueva S, Glavic A, Ruiz P, Mayor R.
Posteriorization by FGF, Wnt, and retinoic acid is required for neural crest induction.
Dev Biol. 2002 Jan 15;241(2):289-301

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 Mar;319(3):367-82

Young HM, Hearn CJ, Ciampoli D, Southwell BR, Brunet JF, Newgreen DF.
A single rostrocaudal colonization of the rodent intestine by enteric neuron precursors is revealed by the expression of Phox2b, Ret, and p75 and by explants grown under the kidney capsule or in organ culture.
Dev Biol. 1998 Oct 1;202(1):67-84.

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

Zhu L, Lee HO, Jordan CS, Cantrell VA, Southard-Smith EM, Shin MK.
Spatiotemporal regulation of endothelin receptor-B by SOX10 in neural crest-derived enteric neuron precursors.
Nat Genet. 2004 Jul;36(7):732-7



# Supplement

**Table 1 List of the genes significantly up-regulated after Gdnf treatment (expression fold difference ≥ 3)**

| | GeneBank accession | Description | Fold Diff |
|---|---|---|---|
| 1 | NM_008818 | placentae and embryos oncofetal gene (Pem), mRNA | 52,48 |
| 2 | NM_011940 | interferon activated gene 202B (Ifi202b), mRNA | 39,75 |
| 3 | NM_011414 | secretory leukocyte protease inhibitor (Slpi), mRNA | 38,57 |
| 4 | NM_008006 | fibroblast growth factor 2 (Fgf2), mRNA | 31,06 |
| 5 | NM_177380 | cytochrome P450, family 3, subfamily a, polypeptide 44 (Cyp3a44), mRNA | 25,68 |
| 6 | NM_008241 | forkhead box G1 (Foxg1), mRNA | 23,87 |
| 7 | AK008862 | adult male stomach cDNA, RIKEN full-length enriched library, clone:2210408O09 product:hypothetical Prokaryotic membrane lipoprotein lipid attachment site containing protein, full insert sequence | 23,51 |
| 8 | NM_008008 | fibroblast growth factor 7 (Fgf7), mRNA | 22,21 |
| 9 | NM_010743 | interleukin 1 receptor-like 1 (Il1rl1), mRNA | 21,16 |
| 10 | NM_027510 | RIKEN cDNA 3830403N18 gene (3830403N18Rik), mRNA | 20,77 |
| 11 | NM_007472 | aquaporin 1 (Aqp1), mRNA | 19,94 |
| 12 | NM_010052 | delta-like 1 homolog (Drosophila) (Dlk1), mRNA | 18,83 |
| 13 | NM_011725 | X-linked lymphocyte-regulated complex (Xlr), mRNA | 18,33 |
| 14 | NM_007820 | cytochrome P450, family 3, subfamily a, polypeptide 16 (Cyp3a16), mRNA | 17,90 |
| 15 | NAP113348-1 | Unknown | 16,43 |
| 16 | ENSMUST00000074104 | Unknown | 15,39 |
| 17 | XM_356393 | PREDICTED: similar to XMR (LOC382276), mRNA | 15,14 |
| 18 | NM_008216 | hyaluronan synthase 2 (Has2), mRNA | 15,03 |
| 19 | NM_146015 | epidermal growth factor-containing fibulin-like extracellular matrix protein 1 (Efemp1), mRNA | 14,81 |
| 20 | NM_011022 | ovary testis transcribed (Ott), mRNA | 14,74 |
| 21 | AK047983 | 16 days embryo head cDNA, RIKEN full-length enriched library, clone:C130026I21 product:hypothetical Sp100 domain containing protein, full insert sequence | 14,42 |
| 22 | NM_145467 | integrin, beta-like 1 (Itgbl1), mRNA | 14,27 |
| 23 | NM_007792 | cysteine and glycine-rich protein 2 (Csrp2), mRNA | 13,51 |
| 24 | NM_027519 | RIKEN cDNA 6330406I15 gene (6330406I15Rik), mRNA | 13,41 |
| 25 | D86232 | mRNA for Ly-6C variant, complete cds | 13,39 |
| 26 | BC053106 | actin-related protein 3-beta, mRNA (cDNA clone IMAGE:5685374), partial cds | 13,11 |
| 27 | AK043991 | 10 days neonate cortex cDNA, RIKEN full-length enriched library, clone:A830072J11 product:unknown EST, full insert sequence | 12,95 |
| 28 | NAP115231-1 | Unknown | 12,80 |
| 29 | NM_007598 | CAP, adenylate cyclase-associated protein 1 (yeast) (Cap1), mRNA | 12,77 |
| 30 | NM_010738 | lymphocyte antigen 6 complex, locus A (Ly6a), mRNA | 12,60 |
| 31 | NM_017396 | cytochrome P450, family 3, subfamily a, polypeptide 41 (Cyp3a41), mRNA | 12,49 |
| 32 | NM_017370 | haptoglobin (Hp), mRNA | 12,38 |
| 33 | AK019885 | 11 days pregnant adult female ovary and uterus cDNA, RIKEN full-length enriched library, clone:5031438N10 product:tumor necrosis factor receptor superfamily, member 9, full insert sequence | 11,58 |
| 34 | NM_007722 | chemokine orphan receptor 1 (Cmkor1), mRNA | 11,47 |

| | | | |
|---|---|---|---|
| 35 | NAP071160-1 | Unknown | 11,39 |
| 36 | NM_177346 | G protein-coupled receptor 149 (Gpr149), mRNA | 10,90 |
| 37 | NM_008926 | protein kinase, cGMP-dependent, type II (Prkg2), mRNA | 10,87 |
| 38 | NM_133775 | RIKEN cDNA 9230117N10 gene (9230117N10Rik), mRNA | 10,74 |
| 39 | NM_029614 | RIKEN cDNA 2310046G15 gene (2310046G15Rik), mRNA | 10,69 |
| 40 | AK031095 | 13 days embryo forelimb cDNA, RIKEN full-length enriched library, clone:5930404C23 product:weakly similar to pol polyprotein , full insert sequence | 10,23 |
| 41 | NM_008318 | integrin binding sialoprotein (Ibsp), mRNA | 10,08 |
| 42 | NM_010225 | forkhead box F2 (Foxf2), mRNA | 10,05 |
| 43 | NM_011311 | S100 calcium binding protein A4 (S100a4), mRNA | 10,01 |
| 44 | AB045716 | TISP22 mRNA, partial cds | 9,54 |
| 45 | NM_133654 | CD34 antigen (Cd34), mRNA | 9,12 |
| 46 | NM_010741 | lymphocyte antigen 6 complex, locus C (Ly6c), mRNA | 9,07 |
| 47 | NM_011620 | troponin T3, skeletal, fast (Tnnt3), mRNA | 8,87 |
| 48 | NM_177157 | GTP cyclohydrolase I feedback regulator (Gchfr), mRNA | 8,52 |
| 49 | NM_007945 | epidermal growth factor receptor pathway substrate 8 (Eps8), mRNA | 8,45 |
| 50 | BC023719 | , clone IMAGE:5347090, mRNA | 8,36 |
| 51 | NM_145078 | RIKEN cDNA 2610305D13 gene (2610305D13Rik), mRNA | 8,32 |
| 52 | NM_183124 | defensin beta 41 (Defb41), mRNA | 8,31 |
| 53 | AI747831 | ul02h07y1 Sugano mouse kidney mkia cDNA clone IMAGE:2064925 5' similar to gb:X04653 Mouse mRNA for Ly-6 alloantigen (MOUSE); | 8,27 |
| 54 | D13695 | mRNA for ST2L protein, complete cds | 8,12 |
| 55 | AK016231 | adult male testis cDNA, RIKEN full-length enriched library, clone:4930565N07 product:, clone MGC:8305 IMAGE:3593825, mRNA, complete cds, full insert sequence | 8,11 |
| 56 | NM_026594 | RIKEN cDNA 4930517K11 gene (4930517K11Rik), mRNA | 8,09 |
| 57 | NM_013473 | annexin A8 (Anxa8), mRNA | 8,07 |
| 58 | NM_008530 | lymphocyte antigen 6 complex, locus F (Ly6f), mRNA | 8,06 |
| 59 | NM_153526 | insulin induced gene 1 (Insig1), mRNA | 7,98 |
| 60 | NM_030736 | vomeronasal 1 receptor, D14 (V1rd14), mRNA | 7,78 |
| 61 | NM_022019 | dual specificity phosphatase 10 (Dusp10), mRNA | 7,70 |
| 62 | NM_172286 | RIKEN cDNA 6430548M08 gene (6430548M08Rik), mRNA | 7,63 |
| 63 | NAP061805-1 | Unknown | 7,60 |
| 64 | NM_025748 | deaminase domain containing 1 (Deadc1), mRNA | 7,56 |
| 65 | AK122503 | mRNA for mKIAA1400 protein | 7,12 |
| 66 | XM_484900 | PREDICTED: similar to hypothetical protein MGC37588 (LOC433337), mRNA | 7,04 |
| 67 | NM_028089 | cytochrome P450, family 2, subfamily c, polypeptide 55 (Cyp2c55), mRNA | 7,00 |
| 68 | BC004800 | procollagen, type IV, alpha 6, mRNA (cDNA clone IMAGE:3589442), partial cds | 6,86 |
| 69 | NM_009398 | tumor necrosis factor alpha induced protein 6 (Tnfaip6), mRNA | 6,73 |
| 70 | NM_173734 | RIKEN cDNA A930025J12 gene (A930025J12Rik), mRNA | 6,73 |
| 71 | NM_172463 | sushi, nidogen and EGF-like domains 1 (Sned1), mRNA | 6,70 |
| 72 | NM_145450 | cDNA sequence BC022687 (BC022687), mRNA | 6,62 |
| 73 | AK019568 | adult male testis cDNA, RIKEN full-length enriched library, clone:4930406H16 product:similar to RIBONUCLEASE/ANGIOGENIN INHIBITOR 2 , full insert sequence | 6,59 |
| 74 | NM_007818 | cytochrome P450, family 3, subfamily a, polypeptide 11 (Cyp3a11), mRNA | 6,55 |
| 75 | NM_022415 | prostaglandin E synthase (Ptges), mRNA | 6,53 |
| 76 | NM_138741 | serum deprivation response (Sdpr), mRNA | 6,52 |
| 77 | NM_178057 | high mobility group AT-hook 2 (Hmga2), mRNA | 6,43 |

| | | | |
|---|---|---|---|
| 78 | NM_016873 | WNT1 inducible signaling pathway protein 2 (Wisp2), mRNA | 6,40 |
| 79 | NM_009141 | chemokine (C-X-C motif) ligand 5 (Cxcl5), mRNA | 6,35 |
| 80 | NM_013519 | forkhead box C2 (Foxc2), mRNA | 6,30 |
| 81 | XM_142281 | PREDICTED: protein phosphatase with EF hand calcium-binding domain (Ppef), mRNA | 6,26 |
| 82 | NM_024433 | methylthioadenosine phosphorylase (Mtap), mRNA | 6,26 |
| 83 | NM_028561 | spermatogenesis associated glutamate (E)-rich protein 4b (Speer4b), mRNA | 6,25 |
| 84 | NM_054071 | fibroblast growth factor receptor-like 1 (Fgfrl1), mRNA | 6,18 |
| 85 | NM_028122 | solute carrier family 14 (urea transporter), member 1 (Slc14a1), mRNA | 6,14 |
| 86 | NM_021304 | abhydrolase domain containing 1 (Abhd1), mRNA | 6,14 |
| 87 | NM_009780 | complement component 4 (within H-2S) (C4), mRNA | 6,14 |
| 88 | NM_009338 | acetyl-Coenzyme A acetyltransferase 2 (Acat2), mRNA | 6,13 |
| 89 | XM_486437 | PREDICTED: similar to hypothetical protein MGC37588 (LOC434589), mRNA | 6,13 |
| 90 | NM_053152 | killer cell lectin-like receptor subfamily A, member 22 (Klra22), mRNA | 6,09 |
| 91 | NM_016778 | Bcl-2-related ovarian killer protein (Bok), mRNA | 5,93 |
| 92 | NM_177420 | phosphoserine aminotransferase 1 (Psat1), mRNA | 5,89 |
| 93 | NM_007616 | caveolin, caveolae protein 1 (Cav1), mRNA | 5,85 |
| 94 | NM_028903 | RIKEN cDNA 4933425F03 gene (4933425F03Rik), mRNA | 5,84 |
| 95 | NM_053150 | killer cell lectin-like receptor subfamily A, member 20 (Klra20), mRNA | 5,82 |
| 96 | NM_153151 | acetyl-Coenzyme A acetyltransferase 3 (Acat3), mRNA | 5,76 |
| 97 | NM_173011 | isocitrate dehydrogenase 2 (NADP+), mitochondrial (Idh2), mRNA | 5,74 |
| 98 | NM_020498 | lymphocyte antigen 6 complex, locus I (Ly6i), mRNA | 5,68 |
| 99 | NM_010235 | fos-like antigen 1 (Fosl1), mRNA | 5,67 |
| 100 | AY155196 | truncated groucho protein GRG1-S (Tle1) mRNA, complete cds; alternatively spliced | 5,64 |
| 101 | NM_009252 | serine (or cysteine) proteinase inhibitor, clade A, member 3N (Serpina3n), mRNA | 5,55 |
| 102 | NM_028535 | RIKEN cDNA 1700049E17 gene (1700049E17Rik), mRNA | 5,50 |
| 103 | NM_010637 | Kruppel-like factor 4 (gut) (Klf4), mRNA | 5,47 |
| 104 | AA726875 | vu36c10r1 Stratagene mouse Tcell 937311 cDNA clone IMAGE:1193490 5' | 5,45 |
| 105 | NM_008344 | insulin-like growth factor binding protein 6 (Igfbp6), mRNA | 5,44 |
| 106 | AK018881 | adult male testis cDNA, RIKEN full-length enriched library, clone:1700066C05 product:hypothetical protein, full insert sequence | 5,43 |
| 107 | NM_029931 | myeloid/lymphoid or mixed lineage-leukemia translocation to 3 homolog (Drosophila) (Mllt3), mRNA | 5,42 |
| 108 | AK079336 | 16 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:9630028E19 product:glycerol phosphate dehydrogenase 1, mitochondrial, full insert sequence | 5,39 |
| 109 | NAP095110-001 | Unknown | 5,36 |
| 110 | NM_007562 | basonuclin 1 (Bnc1), mRNA | 5,34 |
| 111 | AK032211 | adult male olfactory brain cDNA, RIKEN full-length enriched library, clone:6430501F08 product:E2A-PBX1-ASSOCIATED PROTEIN (FRAGMENT) homolog , full insert sequence | 5,32 |
| 112 | NM_011619 | troponin T2, cardiac (Tnnt2), mRNA | 5,29 |
| 113 | BC006733 | cDNA sequence BC037006, mRNA (cDNA clone IMAGE:3963643), partial cds | 5,28 |
| 114 | NM_011077 | phosphate regulating gene with homologies to endopeptidases on the X chromosome (hypophosphatemia, vitamin D resistant rickets) (Phex), mRNA | 5,28 |
| 115 | NM_007679 | CCAAT/enhancer binding protein (C/EBP), delta (Cebpd), mRNA | 5,27 |
| 116 | NM_026821 | DNA segment, Chr 4, Brigham & Women's Genetics 0951 expressed (D4Bwg0951e), mRNA | 5,25 |
| 117 | NM_008485 | laminin, gamma 2 (Lamc2), mRNA | 5,22 |
| 118 | AK005231 | adult male cerebellum cDNA, RIKEN full-length enriched library, clone:1500012F01 product:unknown EST, full insert sequence | 5,22 |

| | | | |
|---|---|---|---|
| 119 | NM_146237 | cDNA sequence BC024537 (BC024537), mRNA | 5,22 |
| 120 | AF334801 | OL-protocadherin isoform (Pcadh10) mRNA, complete cds; alternatively spliced | 5,20 |
| 121 | NM_013793 | killer cell lectin-like receptor, subfamily A, member 15 (Klra15), mRNA | 5,19 |
| 122 | NM_026062 | RIKEN cDNA 2900024C23 gene (2900024C23Rik), mRNA | 5,17 |
| 123 | NM_010107 | ephrin A1 (Efna1), mRNA | 5,16 |
| 124 | NM_028390 | anillin, actin binding protein (scraps homolog, Drosophila) (Anln), mRNA | 5,15 |
| 125 | AF084643 | histocompatibility antigen 60 mRNA, partial cds | 5,15 |
| 126 | NM_001001796 | paired related homeobox protein-like 1 (Prrxl1), mRNA | 5,15 |
| 127 | NM_175442 | RIKEN cDNA A630033H20 gene (A630033H20Rik), mRNA | 5,06 |
| 128 | NM_008987 | pentaxin related gene (Ptx3), mRNA | 5,05 |
| 129 | NM_007436 | aldehyde dehydrogenase family 3, subfamily A1 (Aldh3a1), mRNA | 5,05 |
| 130 | NM_008969 | prostaglandin-endoperoxide synthase 1 (Ptgs1), mRNA | 5,04 |
| 131 | AK078769 | 15 days embryo male testis cDNA, RIKEN full-length enriched library, clone:8030467N07 product:DUDULIN 2 homolog , full insert sequence | 5,03 |
| 132 | TC1262550 | Unknown | 5,03 |
| 133 | NM_016660 | high mobility group AT-hook 1 (Hmga1), mRNA | 5,02 |
| 134 | NM_007950 | epiregulin (Ereg), mRNA | 5,02 |
| 135 | NM_007695 | chitinase 3-like 1 (Chi3l1), mRNA | 4,99 |
| 136 | NM_014194 | killer cell lectin-like receptor, subfamily A, member 7 (Klra7), mRNA | 4,97 |
| 137 | BC024071 | serine (or cysteine) proteinase inhibitor, clade A, member 4, pseudogene 1, mRNA (cDNA clone IMAGE:5097263) | 4,95 |
| 138 | NM_010004 | cytochrome P450, family 2, subfamily c, polypeptide 40 (Cyp2c40), mRNA | 4,91 |
| 139 | NM_199148 | cDNA sequence BC051665 (BC051665), mRNA | 4,88 |
| 140 | AK015913 | adult male testis cDNA, RIKEN full-length enriched library, clone:4930527E24 product:weakly similar to XMR PROTEIN , full insert sequence | 4,84 |
| 141 | TC1244369 | IRX1_MOUSE (P81068) Iroquois-class homeodomain protein IRX-1 (Iroquois homeobox protein 1) (Homeodomain protein IRXA1), complete | 4,83 |
| 142 | NM_028176 | cytidine deaminase (Cda), mRNA | 4,83 |
| 143 | NM_010441 | high mobility group AT-hook 2 (Hmga2), mRNA | 4,77 |
| 144 | NM_011400 | solute carrier family 2 (facilitated glucose transporter), member 1 (Slc2a1), mRNA | 4,77 |
| 145 | NM_008536 | transmembrane 4 superfamily member 1 (Tm4sf1), mRNA | 4,75 |
| 146 | XM_134613 | PREDICTED: similar to hypothetical protein MGC37588 (LOC234907), mRNA | 4,72 |
| 147 | BC009664 | mRNA similar to myeloid cell nuclear differentiation antigen (cDNA clone MGC:7442 IMAGE:3489512), complete cds | 4,72 |
| 148 | XM_355224 | PREDICTED: hypothetical LOC381284 (LOC381284), mRNA | 4,71 |
| 149 | AB093239 | mRNA for mKIAA0417 protein | 4,70 |
| 150 | NM_027828 | RIKEN cDNA 9030611O19 gene (9030611O19Rik), mRNA | 4,67 |
| 151 | BC026762 | cDNA sequence BC026762, mRNA (cDNA clone IMAGE:4504054), containing frame-shift errors | 4,63 |
| 152 | NAP014889-001 | Unknown | 4,62 |
| 153 | XM_145511 | PREDICTED: heat shock protein, alpha-crystallin-related, B6 (Hspb6), mRNA | 4,59 |
| 154 | NM_144526 | RIKEN cDNA 6720460F02 gene (6720460F02Rik), mRNA | 4,55 |
| 155 | NM_012006 | cytosolic acyl-CoA thioesterase 1 (Cte1), mRNA | 4,55 |
| 156 | NM_020286 | pan hematopoietic expression (Phemx), mRNA | 4,53 |
| 157 | NM_026931 | RIKEN cDNA 1810011O10 gene (1810011O10Rik), mRNA | 4,52 |
| 158 | NM_027153 | pirin (Pir), mRNA | 4,49 |
| 159 | NM_175360 | oligonucleotide/oligosaccharide-binding fold containing 1 (Obfc1), mRNA | 4,46 |

| | | | |
|---|---|---|---|
| 160 | AK003102 | adult male heart cDNA, RIKEN full-length enriched library, clone:1010001B21 product:unclassifiable, full insert sequence | 4,46 |
| 161 | NM_018857 | mesothelin (Msln), mRNA | 4,45 |
| 162 | NM_033521 | lysosomal-associated protein transmembrane 4B (Laptm4b), mRNA | 4,45 |
| 163 | NM_028719 | copine IV (Cpne4), mRNA | 4,44 |
| 164 | XM_484575 | PREDICTED: similar to hypothetical protein MGC37588 (LOC433052), mRNA | 4,44 |
| 165 | NM_053178 | lipidosin (Lpd), mRNA | 4,43 |
| 166 | NM_016872 | vesicle-associated membrane protein 5 (Vamp5), mRNA | 4,36 |
| 167 | NM_144549 | tribbles homolog 1 (Drosophila) (Trib1), mRNA | 4,35 |
| 168 | NM_026377 | RIKEN cDNA 6330577E15 gene (6330577E15Rik), mRNA | 4,34 |
| 169 | NM_008778 | p21 (CDKN1A)-activated kinase 3 (Pak3), mRNA | 4,34 |
| 170 | NAP042901-1 | Unknown | 4,29 |
| 171 | TC1324768 | Q8N329 (Q8N329) MGC34132 protein, partial (10%) | 4,27 |
| 172 | NM_201360 | similar to Cytochrome P450 2D9 (CYPIID9) (P450-16-alpha) (CA) (Testosterone 16-alpha hydroxylase) (LOC380997), mRNA | 4,27 |
| 173 | AJ278734 | partial mRNA for hypothetical protein (ORF1), 759 BP | 4,25 |
| 174 | NM_145549 | cDNA sequence BC020077 (BC020077), mRNA | 4,19 |
| 175 | AK042477 | 3 days neonate thymus cDNA, RIKEN full-length enriched library, clone:A630095E13 product:unknown EST, full insert sequence | 4,18 |
| 176 | NM_011599 | transducin-like enhancer of split 1, homolog of Drosophila E(spl) (Tle1), mRNA | 4,17 |
| 177 | NM_008985 | protein tyrosine phosphatase, receptor type, N (Ptprn), mRNA | 4,16 |
| 178 | NM_010447 | heterogeneous nuclear ribonucleoprotein A1 (Hnrpa1), mRNA | 4,15 |
| 179 | NM_007736 | procollagen, type IV, alpha 5 (Col4a5), mRNA | 4,15 |
| 180 | AK019275 | 10, 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2810436B06 product:NG23 (G7D), full insert sequence | 4,13 |
| 181 | NM_011101 | protein kinase C, alpha (Prkca), mRNA | 4,12 |
| 182 | AK013903 | 12 days embryo head cDNA, RIKEN full-length enriched library, clone:3010033P07 product:SIMILAR TO RIBOSOMAL PROTEIN S9 (UNKNOWN) (PROTEIN FOR MGC:14341) (PROTEIN FOR MGC:2458) (PROTEIN FOR MGC:4138) homolog , full insert | 4,10 |
| 183 | NM_174995 | microsomal glutathione S-transferase 2 (Mgst2), mRNA | 4,10 |
| 184 | TC1364280 | ROA1_HUMAN (P09651) Heterogeneous nuclear ribonucleoprotein A1 (Helix-destabilizing protein) (Single-strand binding protein) (hnRNP core protein A1), partial (35%) | 4,10 |
| 185 | NM_138587 | DNA segment, Chr 6, Wayne State University 176, expressed (D6Wsu176e), mRNA | 4,06 |
| 186 | NM_010941 | NAD(P) dependent steroid dehydrogenase-like (Nsdhl), mRNA | 4,04 |
| 187 | NM_175327 | RIKEN cDNA B630019K06 gene (B630019K06Rik), mRNA | 4,04 |
| 188 | NM_021530 | solute carrier family 4 (anion exchanger), member 8 (Slc4a8), mRNA | 4,04 |
| 189 | NM_020583 | interferon-stimulated protein (Isg20), mRNA | 4,02 |
| 190 | NM_153800 | Rho GTPase activating protein 22 (Arhgap22), mRNA | 4,02 |
| 191 | XM_138240 | PREDICTED: similar to hypothetical protein A030003A19 (LOC238395), mRNA | 4,01 |
| 192 | NM_026632 | replication protein A3 (Rpa3), mRNA | 4,01 |
| 193 | AK090296 | 21 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:G630039L19 product:cytochrome P450, 2d22, full insert sequence | 3,99 |
| 194 | XM_486274 | PREDICTED: similar to hypothetical protein MGC37588 (LOC434448), mRNA | 3,99 |
| 195 | NM_011496 | aurora kinase B (Aurkb), mRNA | 3,98 |
| 196 | NM_013525 | growth arrest specific 5 (Gas5), mRNA | 3,97 |
| 197 | NM_007732 | procollagen, type XVII, alpha 1 (Col17a1), mRNA | 3,96 |
| 198 | AY358079 | unknown mRNA | 3,96 |
| 199 | NM_010849 | myelocytomatosis oncogene (Myc), mRNA | 3,95 |
| 200 | NM_023493 | camello-like 5 (Cml5), mRNA | 3,95 |

| | | | |
|---|---|---|---|
| 201 | NM_029376 | spermatogenesis associated glutamate (E)-rich protein 4a (Speer4a), mRNA | 3,92 |
| 202 | A_52_P100480 | Unknown | 3,91 |
| 203 | NM_144879 | RIKEN cDNA B130052G07 gene (B130052G07Rik), mRNA | 3,90 |
| 204 | NM_008815 | ets variant gene 4 (E1A enhancer binding protein, E1AF) (Etv4), mRNA | 3,90 |
| 205 | NM_026038 | RIKEN cDNA 2810055F11 gene (2810055F11Rik), mRNA | 3,89 |
| 206 | NM_028744 | phosphatidylinositol 4-kinase type 2 beta (Pi4k2b), mRNA | 3,89 |
| 207 | NM_026514 | CDC42 effector protein (Rho GTPase binding) 3 (Cdc42ep3), mRNA | 3,88 |
| 208 | NM_178595 | RIKEN cDNA 2210013M04 gene (2210013M04Rik), mRNA | 3,87 |
| 209 | NM_028817 | acyl-CoA synthetase long-chain family member 3 (Acsl3), mRNA | 3,86 |
| 210 | NM_023209 | PDZ binding kinase (Pbk), mRNA | 3,84 |
| 211 | NM_153529 | neuritin 1 (Nrn1), mRNA | 3,82 |
| 212 | NM_009096 | ribosomal protein S6 (Rps6), mRNA | 3,82 |
| 213 | NM_009505 | vascular endothelial growth factor A (Vegfa), mRNA | 3,81 |
| 214 | NM_025329 | RIKEN cDNA 0610012D17 gene (0610012D17Rik), mRNA | 3,81 |
| 215 | NM_009808 | caspase 12 (Casp12), mRNA | 3,80 |
| 216 | ENSMUST00000055154 | Unknown | 3,79 |
| 217 | NM_009533 | X-ray repair complementing defective repair in Chinese hamster cells 5 (Xrcc5), mRNA | 3,77 |
| 218 | ENSMUST00000074568 | Unknown | 3,76 |
| 219 | NM_018827 | cytokine receptor-like factor 1 (Crlf1), mRNA | 3,75 |
| 220 | NM_019823 | cytochrome P450, family 2, subfamily d, polypeptide 22 (Cyp2d22), mRNA | 3,74 |
| 221 | NM_178098 | RIKEN cDNA 4930486L24 gene (4930486L24Rik), mRNA | 3,74 |
| 222 | NM_026619 | glutathione S-transferase omega 2 (Gsto2), mRNA | 3,74 |
| 223 | AK045686 | adult male corpora quadrigemina cDNA, RIKEN full-length enriched library, clone:B230304B05 product:interleukin 1 receptor accessory protein, full insert sequence | 3,73 |
| 224 | NM_029865 | RIKEN cDNA 9430098E02 gene (9430098E02Rik), mRNA | 3,72 |
| 225 | NM_011565 | TEA domain family member 2 (Tead2), mRNA | 3,71 |
| 226 | XM_145565 | PREDICTED: hypothetical LOC233184 (LOC233184), mRNA | 3,70 |
| 227 | ENSMUST00000074320 | Unknown | 3,70 |
| 228 | ENSMUST00000040649 | Unknown | 3,69 |
| 229 | NM_175460 | nicotinamide nucleotide adenylyltransferase 2 (Nmnat2), mRNA | 3,68 |
| 230 | NM_025951 | phosphatidylinositol 4-kinase type 2 beta (Pi4k2b), mRNA | 3,67 |
| 231 | NM_175199 | heat shock 70kDa protein 12A (Hspa12a), mRNA | 3,66 |
| 232 | BC052177 | cDNA clone MGC:59472 IMAGE:6512126, complete cds | 3,66 |
| 233 | NM_009214 | spermine synthase (Sms), mRNA | 3,66 |
| 234 | NM_133719 | meteorin, glial cell differentiation regulator (Metrn), mRNA | 3,65 |
| 235 | NM_013794 | killer cell lectin-like receptor, subfamily A, member 16 (Klra16), mRNA | 3,65 |
| 236 | NM_010646 | killer cell lectin-like receptor subfamily A, member 12 (Klra12), mRNA | 3,65 |
| 237 | NM_022316 | SPARC related modular calcium binding 1 (Smoc1), mRNA | 3,64 |
| 238 | BM934075 | UI-M-CG0p-bmo-c-09-0-Ulr1 NIH_BMAP_Ret4_S2 cDNA clone UI-M-CG0p-bmo-c-09-0-Ul 5', mRNA sequence | 3,63 |
| 239 | ENSMUST00000048525 | Unknown | 3,63 |
| 240 | NM_011361 | serum/glucocorticoid regulated kinase (Sgk), mRNA | 3,62 |
| 241 | NM_009465 | AXL receptor tyrosine kinase (Axl), mRNA | 3,62 |

| | | | |
|---|---|---|---|
| 242 | AK081685 | 16 days embryo head cDNA, RIKEN full-length enriched library, clone:C130067A16 product:hypothetical protein, full insert sequence | 3,62 |
| 243 | NM_008828 | phosphoglycerate kinase 1 (Pgk1), mRNA | 3,61 |
| 244 | A_52_P932984 | Unknown | 3,60 |
| 245 | NM_010755 | v-maf musculoaponeurotic fibrosarcoma oncogene family, protein F (avian) (Maff), mRNA | 3,60 |
| 246 | AB023957 | EIG 180 mRNA for ethanol induced gene product, complete cds | 3,59 |
| 247 | AK012568 | 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2700085B09 product:high mobility group box 2, full insert sequence | 3,59 |
| 248 | NAP113430-1 | Unknown | 3,58 |
| 249 | NM_145978 | PDZ and LIM domain 2 (Pdlim2), mRNA | 3,58 |
| 250 | NM_008949 | proteasome (prosome, macropain) 26S subunit, ATPase 3, interacting protein (Psmc3ip), mRNA | 3,57 |
| 251 | BC032167 | phosphoglycerate kinase 1, mRNA (cDNA clone IMAGE:5354908) | 3,57 |
| 252 | NM_001005608 | integrin beta 4 (Itgb4), transcript variant 1, mRNA | 3,56 |
| 253 | NM_007707 | suppressor of cytokine signaling 3 (Socs3), mRNA | 3,55 |
| 254 | NM_207657 | RIKEN cDNA 5031410I06 gene (5031410I06Rik), mRNA | 3,55 |
| 255 | NM_177960 | isopentenyl-diphosphate delta isomerase (Idi1), mRNA | 3,54 |
| 256 | AI592979 | vu96g05x1 Stratagene mouse skin (#937313) cDNA clone IMAGE:1210040 3' | 3,54 |
| 257 | AK005772 | adult male testis cDNA, RIKEN full-length enriched library, clone:1700008I05 product:weakly similar to TCP11B PROTEIN , full insert sequence | 3,54 |
| 258 | AV264768 | AV264768 RIKEN full-length enriched, adult male testis (DH10B) cDNA clone 4930502O05 3' | 3,51 |
| 259 | NM_011058 | platelet derived growth factor receptor, alpha polypeptide (Pdgfra), mRNA | 3,51 |
| 260 | NM_019865 | ribosomal protein L36a (Rpl36a), mRNA | 3,51 |
| 261 | NM_011507 | succinate-Coenzyme A ligase, GDP-forming, beta subunit (Suclg2), mRNA | 3,50 |
| 262 | NM_026412 | DNA segment, Chr 2, ERATO Doi 750, expressed (D2Ertd750e), mRNA | 3,50 |
| 263 | NM_010455 | homeo box A7 (Hoxa7), mRNA | 3,50 |
| 264 | NM_011267 | regulator of G-protein signaling 16 (Rgs16), mRNA | 3,49 |
| 265 | NM_009760 | BCL2/adenovirus E1B 19kDa-interacting protein 1, NIP3 (Bnip3), mRNA | 3,49 |
| 266 | AK028491 | 0 day neonate skin cDNA, RIKEN full-length enriched library, clone:4631433N18 product:inositol 1,4,5-triphosphate receptor 3, full insert sequence | 3,49 |
| 267 | NM_013876 | ring finger protein 11 (Rnf11), mRNA | 3,48 |
| 268 | NM_022420 | G protein-coupled receptor, family C, group 5, member B (Gprc5b), mRNA | 3,47 |
| 269 | AK010192 | adult male tongue cDNA, RIKEN full-length enriched library, clone:2310076D10 product:hypothetical protein, full insert sequence | 3,47 |
| 270 | NM_172715 | hypothetical protein 4933408F15 (4933408F15), mRNA | 3,46 |
| 271 | NM_016748 | cytidine 5'-triphosphate synthase (Ctps), mRNA | 3,45 |
| 272 | NM_031384 | testis expressed gene 11 (Tex11), mRNA | 3,45 |
| 273 | A_52_P1013232 | Unknown | 3,44 |
| 274 | BC025226 | WD repeat and FYVE domain containing 1, mRNA (cDNA clone MGC:32262 IMAGE:5011404), complete cds | 3,43 |
| 275 | XM_488675 | PREDICTED: expressed sequence AI850995 (AI850995), mRNA | 3,43 |
| 276 | NM_016900 | caveolin 2 (Cav2), mRNA | 3,43 |
| 277 | NM_025800 | protein phosphatase 1, regulatory (inhibitor) subunit 2 (Ppp1r2), mRNA | 3,42 |
| 278 | NM_007971 | enhancer of zeste homolog 2 (Drosophila) (Ezh2), mRNA | 3,42 |
| 279 | NM_008176 | chemokine (C-X-C motif) ligand 1 (Cxcl1), mRNA | 3,40 |
| 280 | AK035814 | 16 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:9630007E23 product:unclassifiable, full insert sequence | 3,40 |
| 281 | NM_029508 | RIKEN cDNA 0610009F02 gene (0610009F02Rik), mRNA | 3,39 |

| | | | |
|---|---|---|---|
| 282 | U19596 | Cdk4 and Cdk6 inhibitor p18 protein mRNA, complete cds | 3,39 |
| 283 | NM_028444 | protein kinase C, delta binding protein (Prkcdbp), mRNA | 3,39 |
| 284 | NM_010884 | N-myc downstream regulated gene 1 (Ndrg1), mRNA | 3,39 |
| 285 | NM_080419 | immunoglobulin superfamily, member 8 (Igsf8), mRNA | 3,38 |
| 286 | NM_019654 | suppressor of cytokine signaling 5 (Socs5), mRNA | 3,37 |
| 287 | NM_027231 | RIKEN cDNA 1810060D16 gene (1810060D16Rik), mRNA | 3,36 |
| 288 | NM_023160 | camello-like 1 (Cml1), mRNA | 3,35 |
| 289 | NM_010274 | glycerol phosphate dehydrogenase 2, mitochondrial (Gpd2), mRNA | 3,35 |
| 290 | NM_019958 | regulator of G-protein signaling 17 (Rgs17), mRNA | 3,35 |
| 291 | AK013636 | adult male hippocampus cDNA, RIKEN full-length enriched library, clone:2900042B11 product:similar to HSPCO34 PROTEIN , full insert sequence | 3,35 |
| 292 | TC1364834 | SP10_HUMAN (P23497) Nuclear autoantigen Sp-100 (Speckled 100 kDa) (Nuclear dot-associated Sp100 protein) (Lysp100b), partial (3%) | 3,35 |
| 293 | NM_010517 | insulin-like growth factor binding protein 4 (Igfbp4), mRNA | 3,34 |
| 294 | NM_008288 | hydroxysteroid 11-beta dehydrogenase 1 (Hsd11b1), mRNA | 3,33 |
| 295 | XM_484297 | PREDICTED: similar to hypothetical protein MGC37588 (LOC432790), mRNA | 3,32 |
| 296 | NM_176835 | RIKEN cDNA 2810451A06 gene (2810451A06Rik), mRNA | 3,31 |
| 297 | NAP060490-1 | Unknown | 3,31 |
| 298 | BC013505 | abhydrolase domain containing 1, mRNA (cDNA clone MGC:19172 IMAGE:4224364), complete cds | 3,31 |
| 299 | NM_023223 | cell division cycle 20 homolog (S cerevisiae) (Cdc20), mRNA | 3,31 |
| 300 | A_52_P10216 40 | Unknown | 3,30 |
| 301 | NM_010807 | MARCKS-like protein (Mlp), mRNA | 3,29 |
| 302 | NM_025370 | RIKEN cDNA 1110018J18 gene (1110018J18Rik), mRNA | 3,28 |
| 303 | NM_178269 | RIKEN cDNA 2610019I03 gene (2610019I03Rik), transcript variant 2, mRNA | 3,28 |
| 304 | AK035884 | 16 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:9630013P17 product:unknown EST, full insert sequence | 3,27 |
| 305 | NM_213615 | RIKEN cDNA A530032D15Rik gene (A530032D15Rik), mRNA | 3,27 |
| 306 | NM_009098 | ribosomal protein S8 (Rps8), mRNA | 3,27 |
| 307 | NM_145150 | protein regulator of cytokinesis 1 (Prc1), mRNA | 3,27 |
| 308 | NM_178929 | Kazal-type serine protease inhibitor domain 1 (Kazald1), mRNA | 3,27 |
| 309 | NM_009372 | TG interacting factor (Tgif), mRNA | 3,27 |
| 310 | NM_007681 | centromere autoantigen A (Cenpa), mRNA | 3,25 |
| 311 | NM_023734 | protease inhibitor 16 (Pi16), mRNA | 3,24 |
| 312 | NM_008776 | platelet-activating factor acetylhydrolase, isoform 1b, alpha1 subunit (Pafah1b3), mRNA | 3,24 |
| 313 | BF149456 | uy65c04y1 McCarrey Eddy round spermatid cDNA clone IMAGE:3664422 5' | 3,23 |
| 314 | NM_016765 | dimethylarginine dimethylaminohydrolase 2 (Ddah2), mRNA | 3,23 |
| 315 | NAP026474-1 | Unknown | 3,23 |
| 316 | U61362 | groucho-related gene 1 protein (Grg1) mRNA, complete cds | 3,23 |
| 317 | NM_009378 | thrombomodulin (Thbd), mRNA | 3,22 |
| 318 | NM_134471 | kinesin family member 2C (Kif2c), mRNA | 3,22 |
| 319 | AK081810 | 16 days embryo head cDNA, RIKEN full-length enriched library, clone:C130078M18 product:fibroblast growth factor receptor 2, full insert sequence | 3,21 |
| 320 | NM_194347 | cDNA sequence AY358078 (AY358078), mRNA | 3,21 |
| 321 | AK017732 | 8 days embryo whole body cDNA, RIKEN full-length enriched library, clone:5730495F03 product:PROTEIN AD-016 (PROTEIN CGI-116) (X0009) homolog , full insert sequence | 3,20 |
| 322 | TC1277049 | BC065934 WD repeat and FYVE domain containing 1, isoform 1 {Homo sapiens;} , partial (68%) | 3,19 |

| | | | |
|---|---|---|---|
| 323 | NM_015774 | ERO1-like (S cerevisiae) (Ero1l), mRNA | 3,19 |
| 324 | NM_013765 | ribosomal protein S26 (Rps26), mRNA | 3,19 |
| 325 | NM_028760 | RIKEN cDNA 1200008O12 gene (1200008O12Rik), mRNA | 3,18 |
| 326 | NM_054048 | REST co-repressor 2 (Rcor2), mRNA | 3,18 |
| 327 | NM_010357 | glutathione S-transferase, alpha 4 (Gsta4), mRNA | 3,18 |
| 328 | NM_025446 | androgen-induced 1 (Aig1), mRNA | 3,18 |
| 329 | TC1308355 | Q7TP79 (Q7TP79) Aa2-245, partial (5%) | 3,16 |
| 330 | NM_008722 | nucleophosmin 1 (Npm1), mRNA | 3,16 |
| 331 | CB245554 | UI-M-FY0-cdt-i-24-0-UIr1 NIH_BMAP_FY0 cDNA clone IMAGE: 6834241 5', mRNA sequence | 3,16 |
| 332 | NM_008592 | forkhead box C1 (Foxc1), mRNA | 3,16 |
| 333 | NM_008966 | prostaglandin F receptor (Ptgfr), mRNA | 3,16 |
| 334 | NM_133641 | rhotekin (Rtkn), mRNA | 3,13 |
| 335 | AK020810 | adult retina cDNA, RIKEN full-length enriched library, clone:A930005F02 product:unknown EST, full insert sequence | 3,13 |
| 336 | AW913539 | uf54b07y1 Soares_mammary_gland_NMLMG cDNA clone IMAGE:1515157 5' | 3,13 |
| 337 | NM_008904 | peroxisome proliferative activated receptor, gamma, coactivator 1 alpha (Ppargc1a), mRNA | 3,13 |
| 338 | AK017223 | 10 days neonate intestine cDNA, RIKEN full-length enriched library, clone:5133401H06 product:unknown EST, full insert sequence | 3,13 |
| 339 | NM_198247 | SERTA domain containing 4 (Sertad4), mRNA | 3,12 |
| 340 | NM_009104 | ribonucleotide reductase M2 (Rrm2), mRNA | 3,12 |
| 341 | NM_009154 | sema domain, seven thrombospondin repeats (type 1 and type 1-like), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 5A (Sema5a), mRNA | 3,11 |
| 342 | TC1247017 | S85192 vascular endothelial growth factor {Homo sapiens;} , partial (45%) | 3,11 |
| 343 | NM_010942 | neuron specific gene family member 1 (Nsg1), mRNA | 3,11 |
| 344 | NM_009862 | cell division cycle 45 homolog (S cerevisiae)-like (Cdc45l), mRNA | 3,11 |
| 345 | NM_011595 | tissue inhibitor of metalloproteinase 3 (Timp3), mRNA | 3,10 |
| 346 | NM_009626 | alcohol dehydrogenase 7 (class IV), mu or sigma polypeptide (Adh7), mRNA | 3,10 |
| 347 | NM_008350 | interleukin 11 (Il11), mRNA | 3,10 |
| 348 | NM_008364 | interleukin 1 receptor accessory protein (Il1rap), transcript variant 1, mRNA | 3,10 |
| 349 | NM_018755 | plasma glutamate carboxypeptidase (Pgcp), mRNA | 3,10 |
| 350 | NM_019755 | proteolipid protein 2 (Plp2), mRNA | 3,09 |
| 351 | AK043877 | 10 days neonate cortex cDNA, RIKEN full-length enriched library, clone:A830044N22 product:ADRENERGIC RECEPTOR, ALPHA 1B, full insert sequence | 3,09 |
| 352 | NM_009344 | pleckstrin homology-like domain, family A, member 1 (Phlda1), mRNA | 3,09 |
| 353 | NM_009270 | squalene epoxidase (Sqle), mRNA | 3,08 |
| 354 | NM_009178 | ST3 beta-galactoside alpha-2,3-sialyltransferase 4 (St3gal4), mRNA | 3,08 |
| 355 | NM_023284 | cell division cycle associated 1 (Cdca1), mRNA | 3,08 |
| 356 | NM_011812 | fibulin 5 (Fbln5), mRNA | 3,07 |
| 357 | NM_009095 | ribosomal protein S5 (Rps5), mRNA | 3,07 |
| 358 | NM_027106 | arginine vasopressin-induced 1 (Avpi1), mRNA | 3,07 |
| 359 | AK004226 | 18-day embryo whole body cDNA, RIKEN full-length enriched library, clone:1110051A18 product:hypothetical protein, full insert sequence | 3,07 |
| 360 | NM_007855 | twist homolog 2 (Drosophila) (Twist2), mRNA | 3,07 |
| 361 | NM_021886 | centromere autoantigen H (Cenph), mRNA | 3,05 |
| 362 | NM_023476 | lipocalin 7 (Lcn7), mRNA | 3,05 |
| 363 | NM_007569 | B-cell translocation gene 1, anti-proliferative (Btg1), mRNA | 3,05 |
| 364 | NM_018796 | eukaryotic translation elongation factor 1 beta 2 (Eef1b2), mRNA | 3,05 |

| | | | |
|---|---|---|---|
| 365 | NM_022424 | fibronectin type III domain containing 4 (Fndc4), mRNA | 3,05 |
| 366 | NM_013509 | enolase 2, gamma neuronal (Eno2), mRNA | 3,04 |
| 367 | NM_009177 | ST3 beta-galactoside alpha-2,3-sialyltransferase 1 (St3gal1), mRNA | 3,04 |
| 368 | BC065150 | cDNA clone IMAGE:6856210, partial cds | 3,04 |
| 369 | NM_008234 | helicase, lymphoid specific (Hells), mRNA | 3,03 |
| 370 | NAP061485-1 | Unknown | 3,03 |
| 371 | ENSMUST00000071204 | Unknown | 3,03 |
| 372 | NM_178874 | transmembrane and coiled-coil domains 2 (Tmcc2), mRNA | 3,02 |
| 373 | AK078379 | 10 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:6530427L06 product:unknown EST, full insert sequence | 3,02 |
| 374 | NM_007815 | cytochrome P450, family 2, subfamily c, polypeptide 29 (Cyp2c29), mRNA | 3,02 |
| 375 | NM_029031 | carbohydrate kinase-like (Carkl), mRNA | 3,01 |
| 376 | BC011413 | mRNA similar to ribosomal protein S20 (cDNA clone MGC:6876 IMAGE:2651405), complete cds | 3,01 |
| 377 | BC027244 | SEH1-like (S cerevisiae, mRNA (cDNA clone MGC:27929 IMAGE:3584697), complete cds | 3,01 |
| 378 | NM_172086 | ribosomal protein L32 (Rpl32), mRNA | 3,00 |

**Table 2 List of the genes significantly down-regulated after Gdnf treatment (expressionfold difference ≥ 3)**

| | GeneBank accession | Description | Fold Diff |
|---|---|---|---|
| 1 | NM_008607 | matrix metalloproteinase 13 (Mmp13), mRNA | 21,30 |
| 2 | NM_008590 | mesoderm specific transcript (Mest), mRNA | 16,74 |
| 3 | NM_015784 | periostin, osteoblast specific factor (Postn), mRNA | 13,80 |
| 4 | NM_007802 | cathepsin K (Ctsk), mRNA | 11,64 |
| 5 | NM_009888 | complement component factor h (Cfh), mRNA | 11,54 |
| 6 | NM_010658 | v-maf musculoaponeurotic fibrosarcoma oncogene family, protein B (avian) (Mafb), mRNA | 10,99 |
| 7 | NM_011582 | thrombospondin 4 (Thbs4), mRNA | 9,92 |
| 8 | M29010 | Mouse complement factor H-related protein mRNA, complete cds, clone 3A4/5G4 | 9,88 |
| 9 | M29009 | Mouse complement factor H-related protein mRNA, complete cds, clone 9C4 | 9,38 |
| 10 | NM_010356 | glutathione S-transferase, alpha 3 (Gsta3), mRNA | 9,33 |
| 11 | NM_008438 | keratocan (Kera), mRNA | 9,09 |
| 12 | NM_023122 | glycoprotein m6b (Gpm6b), mRNA | 8,79 |
| 13 | AK045519 | adult male corpora quadrigemina cDNA, RIKEN full-length enriched library, clone:B230208H11 product:unknown EST, full insert sequence | 8,73 |
| 14 | NM_008524 | lumican (Lum), mRNA | 8,71 |
| 15 | NM_021281 | cathepsin S (Ctss), mRNA | 8,64 |
| 16 | NM_008597 | matrix gamma-carboxyglutamate (gla) protein (Mglap), mRNA | 8,63 |
| 17 | NM_138648 | oxidized low density lipoprotein (lectin-like) receptor 1 (Olr1), mRNA | 8,62 |
| 18 | BC041794 | fat tumor suppressor homolog (Drosophila), mRNA (cDNA clone IMAGE:4975442), partial cds | 8,57 |
| 19 | NM_009928 | procollagen, type XV (Col15a1), mRNA | 8,53 |
| 20 | NM_026162 | plexin domain containing 2 (Plxdc2), mRNA | 8,37 |
| 21 | NM_021451 | phorbol-12-myristate-13-acetate-induced protein 1 (Pmaip1), mRNA | 8,32 |
| 22 | M12660 | Mouse CFh locus, complement protein H gene, complete cds, clones MH(4,8) | 8,31 |
| 23 | NM_007670 | cyclin-dependent kinase inhibitor 2B (p15, inhibits CDK4) (Cdkn2b), mRNA | 8,29 |
| 24 | NM_009930 | procollagen, type III, alpha 1 (Col3a1), mRNA | 8,18 |
| 25 | NM_008605 | matrix metalloproteinase 12 (Mmp12), mRNA | 8,15 |
| 26 | NM_172781 | kelch-like 4 (Drosophila) (Klhl4), mRNA | 8,13 |
| 27 | NM_177715 | potassium channel tetramerisation domain containing 12 (Kctd12), mRNA | 8,12 |
| 28 | TC1323472 | Q86ZH4 (Q86ZH4) Probable CENTROMERE/MICROTUBULE BINDING PROTEIN CBF5, partial (5%) | 8,11 |
| 29 | NM_023422 | histone 1, H2bc (Hist1h2bc), mRNA | 8,09 |
| 30 | NM_022814 | polydomain protein (Polydom), mRNA | 8,03 |
| 31 | NM_025427 | RIKEN cDNA 1190002H23 gene (1190002H23Rik), mRNA | 8,03 |
| 32 | NM_178717 | relaxin 3 receptor 1 (Rln3r1), mRNA | 7,92 |
| 33 | NM_183136 | RIKEN cDNA C630041L24 gene (C630041L24Rik), mRNA | 7,89 |
| 34 | AK033431 | adult male colon cDNA, RIKEN full-length enriched library, clone:9030024J15 product:unknown EST, full insert sequence | 7,83 |
| 35 | NM_026835 | membrane-spanning 4-domains, subfamily A, member 6D (Ms4a6d), mRNA | 7,74 |
| 36 | NM_008083 | growth associated protein 43 (Gap43), mRNA | 7,70 |
| 37 | NM_026271 | RIKEN cDNA 1110018M03 gene (1110018M03Rik), mRNA | 7,64 |
| 38 | NM_022322 | tenomodulin (Tnmd), mRNA | 7,61 |
| 39 | BC062900 | sulfatase 2, mRNA (cDNA clone MGC:86096 IMAGE:6810085), complete cds | 7,57 |
| 40 | NM_007986 | fibroblast activation protein (Fap), mRNA | 7,47 |

| | | | |
|---|---|---|---|
| 41 | NM_008509 | lipoprotein lipase (Lpl), mRNA | 7,44 |
| 42 | NM_007643 | CD36 antigen (Cd36), mRNA | 7,38 |
| 43 | NM_009929 | procollagen, type XVIII, alpha 1 (Col18a1), mRNA | 7,38 |
| 44 | NM_019471 | matrix metalloproteinase 10 (Mmp10), mRNA | 7,38 |
| 45 | TC1324388 | RATALS delta-aminolevulinate synthase precursor {Rattus norvegicus;} , partial (5%) | 7,35 |
| 46 | BC003951 | insulin-like growth factor binding protein 5, mRNA (cDNA clone IMAGE:3487482), partial cds | 7,33 |
| 47 | BC087946 | cDNA clone MGC:107336 IMAGE:6591410, complete cds | 7,24 |
| 48 | BC049786 | aristaless 4, mRNA (cDNA clone MGC:59384 IMAGE:6506755), complete cds | 7,16 |
| 49 | NM_025711 | asporin (Aspn), mRNA | 7,04 |
| 50 | NM_139134 | chondrolectin (Chodl), mRNA | 6,95 |
| 51 | NM_011348 | sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3E (Sema3e), mRNA | 6,93 |
| 52 | NM_007649 | CD48 antigen (Cd48), mRNA | 6,89 |
| 53 | AF434663 | tumor suprressor in lung cancer 1 mRNA, complete cds | 6,89 |
| 54 | NM_010512 | insulin-like growth factor 1 (Igf1), mRNA | 6,89 |
| 55 | L23108 | CD36 antigen mRNA, complete cds | 6,86 |
| 56 | NM_009982 | cathepsin C (Ctsc), mRNA | 6,86 |
| 57 | NM_008584 | mesenchyme homeobox 2 (Meox2), mRNA | 6,84 |
| 58 | NM_010406 | hemolytic complement (Hc), mRNA | 6,83 |
| 59 | TC1262853 | Q8BWC9 (Q8BWC9) 16 days neonate heart cDNA, RIKEN full-length enriched library, clone:D830013H23 product:receptor (calcitonin) activity modifying protein 2, full insert sequence, partial (6%) | 6,82 |
| 60 | NM_019521 | growth arrest specific 6 (Gas6), mRNA | 6,81 |
| 61 | NM_145584 | spondin 1, (f-spondin) extracellular matrix protein (Spon1), mRNA | 6,79 |
| 62 | NM_009364 | tissue factor pathway inhibitor 2 (Tfpi2), mRNA | 6,75 |
| 63 | NM_008879 | lymphocyte cytosolic protein 1 (Lcp1), mRNA | 6,75 |
| 64 | NM_031195 | macrophage scavenger receptor 1 (Msr1), mRNA | 6,73 |
| 65 | NM_009593 | ATP-binding cassette, sub-family G (WHITE), member 1 (Abcg1), mRNA | 6,72 |
| 66 | NM_012008 | DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, Y-linked (Ddx3y), mRNA | 6,72 |
| 67 | NM_016968 | oligodendrocyte transcription factor 1 (Olig1), mRNA | 6,71 |
| 68 | AK018071 | 14 days embryo thymus cDNA, RIKEN full-length enriched library, clone:6130400C22 product:37 kDa leucine-rich repeat (LRR) protein, full insert sequence | 6,70 |
| 69 | NM_007993 | fibrillin 1 (Fbn1), mRNA | 6,67 |
| 70 | NM_028752 | RIKEN cDNA 0610039P13 gene (0610039P13Rik), mRNA | 6,64 |
| 71 | M64085 | Mouse spi2 proteinase inhibitor (spi2/eb1) mRNA, 3' end | 6,62 |
| 72 | NM_009777 | complement component 1, q subcomponent, beta polypeptide (C1qb), mRNA | 6,62 |
| 73 | NM_022431 | membrane-spanning 4-domains, subfamily A, member 11 (Ms4a11), mRNA | 6,59 |
| 74 | NM_019549 | pleckstrin (Plek), mRNA | 6,57 |
| 75 | NM_027209 | membrane-spanning 4-domains, subfamily A, member 6B (Ms4a6b), mRNA | 6,56 |
| 76 | NM_013703 | very low density lipoprotein receptor (Vldlr), mRNA | 6,51 |
| 77 | NM_028749 | N-acetylneuraminate pyruvate lyase (Npl), mRNA | 6,43 |
| 78 | NM_019419 | ADP-ribosylation factor-like 6 interacting protein 1 (Arl6ip1), mRNA | 6,43 |
| 79 | AK051421 | 12 days embryo spinal ganglion cDNA, RIKEN full-length enriched library, clone:D130047M14 product:hypothetical Fibronectin type III structure containing protein, full insert sequence | 6,43 |
| 80 | AK122570 | mRNA for mKIAA1916 protein | 6,42 |
| 81 | NM_013454 | ATP-binding cassette, sub-family A (ABC1), member 1 (Abca1), mRNA | 6,41 |
| 82 | NM_021792 | interferon inducible GTPase 1 (Iigp1), mRNA | 6,41 |

| | | | |
|---|---|---|---|
| 83 | NM_007807 | cytochrome b-245, beta polypeptide (Cybb), mRNA | 6,39 |
| 84 | AK028480 | 0 day neonate skin cDNA, RIKEN full-length enriched library, clone:4631424C05 product:macrophage scavenger receptor 1, full insert sequence | 6,39 |
| 85 | NM_009690 | CD5 antigen-like (Cd5l), mRNA | 6,36 |
| 86 | NM_175406 | ATPase, H+ transporting, V0 subunit D, isoform 2 (Atp6v0d2), mRNA | 6,36 |
| 87 | NM_011360 | sarcoglycan, epsilon (Sgce), mRNA | 6,20 |
| 88 | NM_008695 | nidogen 2 (Nid2), mRNA | 6,18 |
| 89 | NM_009890 | cholesterol 25-hydroxylase (Ch25h), mRNA | 6,17 |
| 90 | NM_023270 | ring finger protein 128 (Rnf128), mRNA | 6,15 |
| 91 | NM_008161 | glutathione peroxidase 3 (Gpx3), mRNA | 6,15 |
| 92 | BC013068 | proprotein convertase subtilisin/kexin type 5, mRNA (cDNA clone MGC:18501 IMAGE:4036159), complete cds | 6,13 |
| 93 | NM_016911 | sushi-repeat-containing protein (Srpx), mRNA | 6,11 |
| 94 | NM_011662 | TYRO protein tyrosine kinase binding protein (Tyrobp), mRNA | 6,09 |
| 95 | NM_007534 | B-cell leukemia/lymphoma 2 related protein A1b (Bcl2a1b), mRNA | 6,09 |
| 96 | NM_207676 | immunoglobulin superfamily, member 4A (Igsf4a), transcript variant 2, mRNA | 6,08 |
| 97 | NM_010423 | hairy/enhancer-of-split related with YRPW motif 1 (Hey1), mRNA | 6,07 |
| 98 | NM_009779 | complement component 3a receptor 1 (C3ar1), mRNA | 6,03 |
| 99 | NM_010724 | proteosome (prosome, macropain) subunit, beta type 8 (large multifunctional protease 7) (Psmb8), mRNA | 6,00 |
| 100 | NM_007572 | complement component 1, q subcomponent, alpha polypeptide (C1qa), mRNA | 5,98 |
| 101 | NM_010686 | lysosomal-associated protein transmembrane 5 (Laptm5), mRNA | 5,98 |
| 102 | AK031268 | 13 days embryo forelimb cDNA, RIKEN full-length enriched library, clone:5932402A14 product:odd Oz/ten-m homolog 3 (Drosophila), full insert sequence | 5,95 |
| 103 | BC020023 | glutaminyl-peptide cyclotransferase (glutaminyl cyclase), mRNA (cDNA clone MGC:27858 IMAGE:3491756), complete cds | 5,92 |
| 104 | NM_175663 | histone 1, H2ba (Hist1h2ba), mRNA | 5,90 |
| 105 | NM_008625 | mannose receptor, C type 1 (Mrc1), mRNA | 5,83 |
| 106 | NM_008311 | 5-hydroxytryptamine (serotonin) receptor 2B (Htr2b), mRNA | 5,83 |
| 107 | NM_178200 | histone 1, H2bm (Hist1h2bm), mRNA | 5,83 |
| 108 | NM_011443 | SRY-box containing gene 2 (Sox2), mRNA | 5,81 |
| 109 | AK019528 | 0 day neonate head cDNA, RIKEN full-length enriched library, clone:4833446K15 product:hypothetical protein, full insert sequence | 5,79 |
| 110 | AK053990 | 2 days pregnant adult female oviduct cDNA, RIKEN full-length enriched library, clone:E230011B21 product:unknown EST, full insert sequence | 5,79 |
| 111 | NM_175665 | histone 1, H2bk (Hist1h2bk), mRNA | 5,79 |
| 112 | NM_013532 | leukocyte immunoglobulin-like receptor, subfamily B, member 4 (Lilrb4), mRNA | 5,79 |
| 113 | NM_028595 | membrane-spanning 4-domains, subfamily A, member 6C (Ms4a6c), mRNA | 5,78 |
| 114 | NM_011095 | paired-Ig-like receptor B (Pirb), mRNA | 5,76 |
| 115 | NM_020008 | C-type (calcium dependent, carbohydrate recognition domain) lectin, superfamily member 12 (Clecsf12), mRNA | 5,75 |
| 116 | NM_009970 | colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage) (Csf2ra), mRNA | 5,75 |
| 117 | BB206048 | BB206048 RIKEN full-length enriched, 0 day neonate thymus cDNA clone A430071M17 3' | 5,73 |
| 118 | NM_009155 | selenoprotein P, plasma, 1 (Sepp1), mRNA | 5,72 |
| 119 | NM_181397 | RIKEN cDNA 2310015N21 gene (2310015N21Rik), mRNA | 5,69 |
| 120 | NM_007535 | B-cell leukemia/lymphoma 2 related protein A1c (Bcl2a1c), mRNA | 5,68 |
| 121 | NM_020561 | sphingomyelin phosphodiesterase, acid-like 3A (Smpdl3a), mRNA | 5,67 |
| 122 | AK014285 | 14, 17 days embryo head cDNA, RIKEN full-length enriched library, clone:3200001I04 product:hypothetical Cysteine-rich flanking region, C-terminal/Leucine-rich repeat/Leucine-rich repeat, typical subtype containing protein, full insert | 5,64 |

| | | | |
|---|---|---|---|
| 123 | NM_009849 | ectonucleoside triphosphate diphosphohydrolase 2 (Entpd2), mRNA | 5,64 |
| 124 | NM_207655 | epidermal growth factor receptor (Egfr), transcript variant 1, mRNA | 5,62 |
| 125 | NM_007403 | a disintegrin and metalloprotease domain 8 (Adam8), mRNA | 5,61 |
| 126 | NM_007574 | complement component 1, q subcomponent, gamma polypeptide (C1qg), mRNA | 5,60 |
| 127 | NM_206882 | histone 3, H2bb (Hist3h2bb), mRNA | 5,58 |
| 128 | NM_013489 | CD84 antigen (Cd84), mRNA | 5,57 |
| 129 | AK077477 | 8 days embryo whole body cDNA, RIKEN full-length enriched library, clone:5730419J04 product:insulin-like growth factor binding protein 3, full insert sequence | 5,55 |
| 130 | NM_007498 | activating transcription factor 3 (Atf3), mRNA | 5,52 |
| 131 | NM_026167 | kelch-like 13 (Drosophila) (Klhl13), mRNA | 5,51 |
| 132 | NM_007801 | cathepsin H (Ctsh), mRNA | 5,48 |
| 133 | NM_001002927 | preproenkephalin 1 (Penk1), mRNA | 5,47 |
| 134 | AK011113 | 13 days embryo liver cDNA, RIKEN full-length enriched library, clone:2510048K12 product:plexin C1, full insert sequence | 5,46 |
| 135 | NM_019955 | receptor-interacting serine-threonine kinase 3 (Ripk3), mRNA | 5,44 |
| 136 | NM_011670 | ubiquitin carboxy-terminal hydrolase L1 (Uchl1), mRNA | 5,41 |
| 137 | NM_027836 | membrane-spanning 4-domains, subfamily A, member 7 (Ms4a7), mRNA | 5,40 |
| 138 | NM_015811 | regulator of G-protein signaling 1 (Rgs1), mRNA | 5,40 |
| 139 | AK129139 | mRNA for mKIAA0429 protein | 5,39 |
| 140 | AK033780 | adult male epididymis cDNA, RIKEN full-length enriched library, clone:9230106K10 product:unknown EST, full insert sequence | 5,39 |
| 141 | NM_177337 | ADP-ribosylation factor-like 11 (Arl11), mRNA | 5,39 |
| 142 | NM_012011 | eukaryotic translation initiation factor 2, subunit 3, structural gene Y-linked (Eif2s3y), mRNA | 5,38 |
| 143 | NM_007651 | CD53 antigen (Cd53), mRNA | 5,38 |
| 144 | NM_010708 | lectin, galactose binding, soluble 9 (Lgals9), mRNA | 5,37 |
| 145 | NM_019413 | roundabout homolog 1 (Drosophila) (Robo1), mRNA | 5,36 |
| 146 | NM_134102 | phospholipase A1 member A (Pla1a), mRNA | 5,36 |
| 147 | NM_009911 | chemokine (C-X-C motif) receptor 4 (Cxcr4), mRNA | 5,34 |
| 148 | NM_010442 | heme oxygenase (decycling) 1 (Hmox1), mRNA | 5,33 |
| 149 | NM_008760 | osteoglycin (Ogn), mRNA | 5,31 |
| 150 | NM_010566 | inositol polyphosphate-5-phosphatase D (Inpp5d), mRNA | 5,30 |
| 151 | NM_009853 | CD68 antigen (Cd68), mRNA | 5,30 |
| 152 | NM_198885 | scleraxis (Scx), mRNA | 5,28 |
| 153 | NM_011854 | 2'-5' oligoadenylate synthetase-like 2 (Oasl2), mRNA | 5,28 |
| 154 | NM_010188 | Fc receptor, IgG, low affinity III (Fcgr3), mRNA | 5,27 |
| 155 | NM_183168 | pyrimidinergic receptor P2Y, G-protein coupled, 6 (P2ry6), mRNA | 5,27 |
| 156 | NM_021883 | tropomodulin 1 (Tmod1), mRNA | 5,26 |
| 157 | AI326608 | mm74d09y1 Stratagene mouse macrophage (#937306) cDNA clone IMAGE:534161 5' similar to gb:J05020 Mouse mast cell high affinity IgE receptor (MOUSE); | 5,26 |
| 158 | NM_027288 | mannosidase, beta A, lysosomal (Manba), mRNA | 5,26 |
| 159 | NM_030707 | macrophage scavenger receptor 2 (Msr2), mRNA | 5,25 |
| 160 | NM_019455 | prostaglandin D2 synthase 2, hematopoietic (Ptgds2), mRNA | 5,25 |
| 161 | NM_009549 | zinc finger protein 185 (Zfp185), mRNA | 5,24 |
| 162 | NM_175476 | Rho GTPase activating protein 25 (Arhgap25), mRNA | 5,23 |
| 163 | NM_184052 | insulin-like growth factor 1 (Igf1), mRNA | 5,23 |
| 164 | NM_023879 | retinitis pigmentosa GTPase regulator interacting protein 1 (Rpgrip1), mRNA | 5,21 |

| | | | |
|---|---|---|---|
| 165 | X15592 | Mouse ctla-2-beta mRNA, homolog to cysteine protease proregion | 5,21 |
| 166 | NM_007796 | cytotoxic T lymphocyte-associated protein 2 alpha (Ctla2a), mRNA | 5,21 |
| 167 | NM_023788 | melanoma antigen, family H, 1 (Mageh1), mRNA | 5,20 |
| 168 | AK083900 | 12 days embryo spinal ganglion cDNA, RIKEN full-length enriched library, clone:D130057B15 product:Friend leukemia integration 1, full insert sequence | 5,19 |
| 169 | TC1343649 | COAT_FMVD (P09519) Probable coat protein, partial (6%) | 5,19 |
| 170 | TC1344323 | CT2B_MOUSE (P12400) CTLA-2-beta protein precursor (Fragment), partial (98%) | 5,19 |
| 171 | BC016539 | cDNA clone IMAGE:3497440, with apparent retained intron | 5,18 |
| 172 | NM_020043 | neighbor of Punc E11 (Nope), mRNA | 5,18 |
| 173 | NM_023158 | chemokine (C-X-C motif) ligand 16 (Cxcl16), mRNA | 5,17 |
| 174 | NM_008481 | laminin, alpha 2 (Lama2), mRNA | 5,17 |
| 175 | NM_026656 | mucolipin 2 (Mcoln2), transcript variant 1, mRNA | 5,16 |
| 176 | NM_010819 | C-type (calcium dependent, carbohydrate recognition domain) lectin, superfamily member 8 (Clecsf8), mRNA | 5,15 |
| 177 | NM_009931 | procollagen, type IV, alpha 1 (Col4a1), mRNA | 5,15 |
| 178 | NM_134158 | immunoglobulin superfamily, member 7 (Igsf7), mRNA | 5,15 |
| 179 | NM_011539 | thromboxane A synthase 1, platelet (Tbxas1), mRNA | 5,13 |
| 180 | NM_178611 | leukocyte-associated Ig-like receptor 1 (Lair1), mRNA | 5,11 |
| 181 | NM_010747 | Yamaguchi sarcoma viral (v-yes-1) oncogene homolog (Lyn), mRNA | 5,10 |
| 182 | NM_010332 | endothelin receptor type A (Ednra), mRNA | 5,10 |
| 183 | AK033487 | adult male colon cDNA, RIKEN full-length enriched library, clone:9030224D03 product:hypothetical Immunoglobulin and major histocompatibility complex domain/Immunoglobulin C-2 type/Immunoglobulin-like/Immunoglobulin subtype containing | 5,09 |
| 184 | NM_016674 | claudin 1 (Cldn1), mRNA | 5,09 |
| 185 | NM_183204 | RIKEN cDNA C630023L15 gene (C630023L15Rik), mRNA | 5,07 |
| 186 | NM_198301 | cDNA sequence BC052328 (BC052328), mRNA | 5,07 |
| 187 | NM_173385 | cartilage intermediate layer protein, nucleotide pyrophosphohydrolase (Cilp), mRNA | 5,07 |
| 188 | NM_010726 | phytanoyl-CoA hydroxylase (Phyh), mRNA | 5,05 |
| 189 | NM_013590 | P lysozyme structural (Lzp-s), mRNA | 5,04 |
| 190 | BC038870 | , clone IMAGE:1348774, mRNA | 5,03 |
| 191 | BC013561 | cDNA clone IMAGE:3492058, partial cds | 5,02 |
| 192 | NM_013496 | cellular retinoic acid binding protein I (Crabp1), mRNA | 5,01 |
| 193 | NM_013470 | annexin A3 (Anxa3), mRNA | 4,98 |
| 194 | NM_008328 | interferon activated gene 203 (Ifi203), mRNA | 4,98 |
| 195 | AK020134 | 12 days embryo male wolffian duct includes surrounding region cDNA, RIKEN full-length enriched library, clone:6720458D04 product:receptor (calcitonin) activity modifying protein 2, full insert sequence | 4,97 |
| 196 | NM_172507 | SH3 domain binding glutamic acid-rich protein like 2 (Sh3bgrl2), mRNA | 4,97 |
| 197 | NM_008528 | B-cell linker (Blnk), mRNA | 4,97 |
| 198 | AK003401 | 18-day embryo whole body cDNA, RIKEN full-length enriched library, clone:1110004B06 product:unknown EST, full insert sequence | 4,95 |
| 199 | NM_007737 | procollagen, type V, alpha 2 (Col5a2), mRNA | 4,95 |
| 200 | NM_134066 | aldo-keto reductase family 1, member C18 (Akr1c18), mRNA | 4,93 |
| 201 | NM_134050 | RAB15, member RAS oncogene family (Rab15), mRNA | 4,91 |
| 202 | NM_016770 | folate hydrolase (Folh1), mRNA | 4,89 |
| 203 | NM_009255 | serine (or cysteine) proteinase inhibitor, clade E, member 2 (Serpine2), mRNA | 4,87 |
| 204 | NM_031185 | A kinase (PRKA) anchor protein (gravin) 12 (Akap12), mRNA | 4,86 |
| 205 | NM_021325 | CD200 receptor 1 (Cd200r1), mRNA | 4,86 |

| | | | |
|---|---|---|---|
| 206 | AK005381 | adult male cerebellum cDNA, RIKEN full-length enriched library, clone:1500041B16 product:hypothetical protein, full insert sequence | 4,86 |
| 207 | NM_011330 | small chemokine (C-C motif) ligand 11 (Ccl11), mRNA | 4,83 |
| 208 | NM_022029 | neurogranin (Nrgn), mRNA | 4,83 |
| 209 | NM_009696 | apolipoprotein E (Apoe), mRNA | 4,82 |
| 210 | NM_008131 | glutamate-ammonia ligase (glutamine synthase) (Glul), mRNA | 4,79 |
| 211 | NM_008981 | protein tyrosine phosphatase, receptor type, G (Ptprg), mRNA | 4,77 |
| 212 | NM_010258 | GATA binding protein 6 (Gata6), mRNA | 4,77 |
| 213 | NM_015786 | histone 1, H1c (Hist1h1c), mRNA | 4,73 |
| 214 | NM_008409 | integral membrane protein 2A (Itm2a), mRNA | 4,73 |
| 215 | NM_023061 | melanoma cell adhesion molecule (Mcam), mRNA | 4,72 |
| 216 | NAP108462-1 | Unknown | 4,71 |
| 217 | NM_011125 | phospholipid transfer protein (Pltp), mRNA | 4,70 |
| 218 | NM_011093 | paired-Ig-like receptor A6 (Pira6), transcript variant 2, mRNA | 4,70 |
| 219 | NM_031254 | triggering receptor expressed on myeloid cells 2 (Trem2), mRNA | 4,69 |
| 220 | NM_024223 | cysteine rich protein 2 (Crip2), mRNA | 4,69 |
| 221 | NM_020616 | predicted gene ICRFP703B1614Q56 (ICRFP703B1614Q56), mRNA | 4,69 |
| 222 | XM_135029 | PREDICTED: cDNA sequence BC023892 (BC023892), mRNA | 4,68 |
| 223 | NM_011638 | transferrin receptor (Tfrc), mRNA | 4,68 |
| 224 | NM_009876 | cyclin-dependent kinase inhibitor 1C (P57) (Cdkn1c), mRNA | 4,68 |
| 225 | AK129084 | mRNA for mKIAA0193 protein | 4,68 |
| 226 | NM 001005858 | similar to This ORF is capable of encoding 404aa which is homologous to two human interferon-inducible proteins, 54 kDa and 56 kDa proteins; ORF (LOC433243), mRNA | 4,66 |
| 227 | BC068138 | cDNA clone IMAGE:6838317, partial cds | 4,66 |
| 228 | BC057682 | cDNA clone IMAGE:3596174, partial cds | 4,64 |
| 229 | NM_028882 | sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3D (Sema3d), mRNA | 4,63 |
| 230 | NM_021452 | potassium large conductance calcium-activated channel, subfamily M, beta member 4 (Kcnmb4), mRNA | 4,63 |
| 231 | NM_010187 | Fc receptor, IgG, low affinity IIb (Fcgr2b), mRNA | 4,63 |
| 232 | NM_009848 | ectonucleoside triphosphate diphosphohydrolase 1 (Entpd1), mRNA | 4,61 |
| 233 | NM_011090 | paired-Ig-like receptor A3 (Pira3), mRNA | 4,61 |
| 234 | NM_198095 | bone marrow stromal cell antigen 2 (Bst2), mRNA | 4,60 |
| 235 | NM_008855 | protein kinase C, beta 1 (Prkcb1), mRNA | 4,59 |
| 236 | BC043308 | a disintegrin-like and metalloprotease (reprolysin type) with thrombospondin type 1 motif, 15, mRNA (cDNA clone IMAGE:5044493), partial cds | 4,57 |
| 237 | NM_010876 | neutrophil cytosolic factor 1 (Ncf1), mRNA | 4,56 |
| 238 | NM_021273 | creatine kinase, brain (Ckb), mRNA | 4,56 |
| 239 | NM_008546 | microfibrillar-associated protein 2 (Mfap2), mRNA | 4,55 |
| 240 | NM_011857 | odd Oz/ten-m homolog 3 (Drosophila) (Odz3), mRNA | 4,55 |
| 241 | NM_024474 | EMI domain containing 2 (Emid2), mRNA | 4,55 |
| 242 | NM_026838 | sushi-repeat-containing protein, X-linked 2 (Srpx2), mRNA | 4,55 |
| 243 | AK084717 | 13 days embryo heart cDNA, RIKEN full-length enriched library, clone:D330035F22 product:unknown EST, full insert sequence | 4,53 |
| 244 | NM_025311 | DNA segment, Chr 14, ERATO Doi 449, expressed (D14Ertd449e), mRNA | 4,50 |
| 245 | NM_027265 | RIKEN cDNA 2810004A10 gene (2810004A10Rik), mRNA | 4,50 |
| 246 | BC025535 | Fc receptor, IgG, high affinity I, mRNA (cDNA clone IMAGE:5249690), partial cds | 4,49 |
| 247 | NM_144538 | RAB3A interacting protein (rabin3)-like 1 (Rab3il1), mRNA | 4,48 |

| | | | |
|---|---|---|---|
| 248 | NM_053078 | DNA segment, human D4S114 (D0H4S114), mRNA | 4,48 |
| 249 | NM_007901 | endothelial differentiation sphingolipid G-protein-coupled receptor 1 (Edg1), mRNA | 4,48 |
| 250 | NM_023386 | RIKEN cDNA 5830458K16 gene (5830458K16Rik), mRNA | 4,47 |
| 251 | NM_007654 | CD72 antigen (Cd72), mRNA | 4,46 |
| 252 | AK129227 | mRNA for mKIAA0840 protein | 4,45 |
| 253 | NM_011087 | paired-Ig-like receptor A1 (Pira1), mRNA | 4,45 |
| 254 | NM_008331 | interferon-induced protein with tetratricopeptide repeats 1 (Ifit1), mRNA | 4,44 |
| 255 | NM_133211 | toll-like receptor 7 (Tlr7), mRNA | 4,43 |
| 256 | NM_010493 | intercellular adhesion molecule (Icam1), mRNA | 4,43 |
| 257 | NM_011150 | lectin, galactoside-binding, soluble, 3 binding protein (Lgals3bp), mRNA | 4,43 |
| 258 | NM_133969 | cytochrome P450, family 4, subfamily v, polypeptide 3 (Cyp4v3), mRNA | 4,43 |
| 259 | NM_146131 | pre-B-cell leukemia transcription factor interacting protein 1 (Pbxip1), mRNA | 4,42 |
| 260 | NM_010130 | EGF-like module containing, mucin-like, hormone receptor-like sequence 1 (Emr1), mRNA | 4,42 |
| 261 | NM_153505 | hematopoietic protein 1 (Hemp1), mRNA | 4,42 |
| 262 | NM_024406 | fatty acid binding protein 4, adipocyte (Fabp4), mRNA | 4,41 |
| 263 | NM_010186 | Fc receptor, IgG, high affinity I (Fcgr1), mRNA | 4,41 |
| 264 | NM_010422 | hexosaminidase B (Hexb), mRNA | 4,41 |
| 265 | NM_138683 | thrombospondin type 1 domain containing gene (Rspondin), mRNA | 4,40 |
| 266 | AK046600 | 4 days neonate male adipose cDNA, RIKEN full-length enriched library, clone:B430202I07 product:C-type (calcium dependent, carbohydrate-recognition domain) lectin, superfamily member 5, full insert sequence | 4,40 |
| 267 | BC052506 | cDNA clone MGC:62985 IMAGE:1245563, complete cds | 4,39 |
| 268 | NM_008152 | G-protein coupled receptor 65 (Gpr65), mRNA | 4,39 |
| 269 | NM_010809 | matrix metalloproteinase 3 (Mmp3), mRNA | 4,39 |
| 270 | NM_138672 | stabilin 1 (Stab1), mRNA | 4,39 |
| 271 | NM_178589 | tumor necrosis factor receptor superfamily, member 21 (Tnfrsf21), mRNA | 4,38 |
| 272 | NM_009197 | solute carrier family 16 (monocarboxylic acid transporters), member 2 (Slc16a2), mRNA | 4,38 |
| 273 | NM_033325 | lysyl oxidase-like 2 (Loxl2), mRNA | 4,38 |
| 274 | NM_011408 | schlafen 2 (Slfn2), mRNA | 4,35 |
| 275 | NM_138305 | adenylate cyclase 3 (Adcy3), mRNA | 4,35 |
| 276 | BC036146 | tissue factor pathway inhibitor, mRNA (cDNA clone MGC:37332 IMAGE:4975683), complete cds | 4,35 |
| 277 | AK054327 | 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched library, clone:E330015I04 product:PUTATIVE ANION TRANSPORTER homolog , full insert sequence | 4,35 |
| 278 | NM_010471 | hippocalcin (Hpca), mRNA | 4,34 |
| 279 | NM_008035 | folate receptor 2 (fetal) (Folr2), mRNA | 4,34 |
| 280 | NM_013834 | secreted frizzled-related sequence protein 1 (Sfrp1), mRNA | 4,33 |
| 281 | AK046802 | 10 days neonate medulla oblongata cDNA, RIKEN full-length enriched library, clone:B830011M21 product:ALPHA-2A ADRENERGIC RECEPTOR (ALPHA-2A ADRENOCEPTOR) (ALPHA-2AAR), full insert sequence | 4,33 |
| 282 | NM_009776 | serine (or cysteine) proteinase inhibitor, clade G, member 1 (Serping1), mRNA | 4,33 |
| 283 | NM_011337 | chemokine (C-C motif) ligand 3 (Ccl3), mRNA | 4,33 |
| 284 | NM_011909 | ubiquitin specific protease 18 (Usp18), mRNA | 4,32 |
| 285 | NM_023716 | RIKEN cDNA 2410129E14 gene (2410129E14Rik), mRNA | 4,32 |
| 286 | NM_022032 | PERP, TP53 apoptosis effector (Perp), mRNA | 4,32 |
| 287 | NM_011111 | serine (or cysteine) proteinase inhibitor, clade B, member 2 (Serpinb2), mRNA | 4,31 |
| 288 | NM_011576 | tissue factor pathway inhibitor (Tfpi), mRNA | 4,31 |
| 289 | NM_133212 | toll-like receptor 8 (Tlr8), mRNA | 4,31 |

| | | | |
|---|---|---|---|
| 290 | Z11886 | Mmusculus notch-1 mRNA | 4,31 |
| 291 | NM_153534 | adenylate cyclase 2 (Adcy2), mRNA | 4,29 |
| 292 | NM_053214 | myosin IF (Myo1f), mRNA | 4,28 |
| 293 | NM_013612 | solute carrier family 11 (proton-coupled divalent metal ion transporters), member 1 (Slc11a1), mRNA | 4,28 |
| 294 | NM_015783 | interferon, alpha-inducible protein (G1p2), mRNA | 4,28 |
| 295 | NM_008608 | matrix metalloproteinase 14 (membrane-inserted) (Mmp14), mRNA | 4,28 |
| 296 | NM_009199 | solute carrier family 1 (neuronal/epithelial high affinity glutamate transporter, system Xag), member 1 (Slc1a1), mRNA | 4,28 |
| 297 | NM_023184 | Kruppel-like factor 15 (Klf15), mRNA | 4,27 |
| 298 | NM_008144 | Bernardinelli-Seip congenital lipodystrophy 2 homolog (human) (Bscl2), mRNA | 4,27 |
| 299 | BC025841 | cDNA clone IMAGE:5149318, partial cds | 4,27 |
| 300 | NM_018797 | plexin C1 (Plxnc1), mRNA | 4,25 |
| 301 | NM_207675 | immunoglobulin superfamily, member 4A (Igsf4a), transcript variant 1, mRNA | 4,25 |
| 302 | NM_010501 | interferon-induced protein with tetratricopeptide repeats 3 (Ifit3), mRNA | 4,25 |
| 303 | NM_031257 | pleckstrin homology domain-containing, family A (phosphoinositide binding specific) member 2 (Plekha2), mRNA | 4,24 |
| 304 | NM_011355 | SFFV proviral integration 1 (Sfpi1), mRNA | 4,24 |
| 305 | NM_016865 | HIV-1 tat interactive protein 2, homolog (human) (Htatip2), mRNA | 4,24 |
| 306 | AK030569 | adult male pituitary gland cDNA, RIKEN full-length enriched library, clone:5330432M01 product:inferred: RIKEN cDNA 5930418K15 gene, full insert sequence | 4,23 |
| 307 | AK040921 | adult male aorta and vein cDNA, RIKEN full-length enriched library, clone:A530045K20 product:integrin alpha M, full insert sequence | 4,22 |
| 308 | AK129116 | mRNA for mKIAA0331 protein | 4,22 |
| 309 | NM_008013 | fibrinogen-like protein 2 (Fgl2), mRNA | 4,22 |
| 310 | AK042139 | 3 days neonate thymus cDNA, RIKEN full-length enriched library, clone:A630063B02 product:hypothetical SNF2 related domain , Helicase c-terminal domain, DEAD/DEAH box helicase containing protein, full insert sequence | 4,21 |
| 311 | BC038365 | nerve growth factor receptor (TNFR superfamily, member 16), mRNA (cDNA clone MGC:35588 IMAGE:5367638), complete cds | 4,21 |
| 312 | NM_011844 | monoglyceride lipase (Mgll), mRNA | 4,21 |
| 313 | NM_028071 | coactosin-like 1 (Dictyostelium) (Cotl1), mRNA | 4,21 |
| 314 | NM_020001 | C-type (calcium dependent, carbohydrate recognition domain) lectin, superfamily member 10 (Clecsf10), mRNA | 4,20 |
| 315 | NM_021457 | frizzled homolog 1 (Drosophila) (Fzd1), mRNA | 4,19 |
| 316 | NM_007400 | a disintegrin and metalloproteinase domain 12 (meltrin alpha) (Adam12), mRNA | 4,19 |
| 317 | NM_011260 | regenerating islet-derived 3 gamma (Reg3g), mRNA | 4,19 |
| 318 | BC021340 | poly (ADP-ribose) polymerase family, member 14, mRNA (cDNA clone MGC:29390 IMAGE:5065398), complete cds | 4,18 |
| 319 | NM_181318 | RasGEF domain family, member 1B (Rasgef1b), mRNA | 4,15 |
| 320 | NM_008973 | pleiotrophin (Ptn), mRNA | 4,14 |
| 321 | BC076584 | glutamate receptor, ionotropic, AMPA3 (alpha 3), mRNA (cDNA clone IMAGE:30620696), containing frame-shift errors | 4,14 |
| 322 | NM_029758 | DNA segment, Chr 12, ERATO Doi 553, expressed (D12Ertd553e), mRNA | 4,13 |
| 323 | NM_009994 | cytochrome P450, family 1, subfamily b, polypeptide 1 (Cyp1b1), mRNA | 4,13 |
| 324 | AJ006521 | mRNA for m1 muscarinic acetylcholine receptor protein, partial | 4,13 |
| 325 | NM_010185 | Fc receptor, IgE, high affinity I, gamma polypeptide (Fcer1g), mRNA | 4,12 |
| 326 | NM_177843 | hypothetical protein A830093M07 (A830093M07), mRNA | 4,12 |
| 327 | NM_173767 | RIKEN cDNA 3830422K02 gene (3830422K02Rik), mRNA | 4,11 |
| 328 | NM_010158 | KH domain containing, RNA binding, signal transduction associated 3 (Khdrbs3), mRNA | 4,10 |
| 329 | NM_025582 | RIKEN cDNA 2810405K02 gene (2810405K02Rik), mRNA | 4,10 |

| | | | |
|---|---|---|---|
| 330 | NM_133977 | transferrin (Trf), mRNA | 4,10 |
| 331 | NM_008737 | neuropilin 1 (Nrp1), mRNA | 4,10 |
| 332 | NM_008349 | interleukin 10 receptor, beta (Il10rb), mRNA | 4,09 |
| 333 | L20315 | MPS1 gene and mRNA, 3'end | 4,09 |
| 334 | NM_008245 | hematopoietically expressed homeobox (Hhex), mRNA | 4,08 |
| 335 | NM_009856 | CD83 antigen (Cd83), mRNA | 4,08 |
| 336 | NM_009303 | synaptogyrin 1 (Syngr1), transcript variant 1b, mRNA | 4,07 |
| 337 | NM_011925 | CD97 antigen (Cd97), mRNA | 4,07 |
| 338 | NM_030691 | immunoglobulin superfamily, member 6 (Igsf6), mRNA | 4,07 |
| 339 | NM_021342 | potassium voltage-gated channel, Isk-related subfamily, gene 4 (Kcne4), mRNA | 4,07 |
| 340 | NM_011691 | vav 1 oncogene (Vav1), mRNA | 4,06 |
| 341 | L02241 | Mouse protein kinase inhibitor (testicular isoform) mRNA, complete cds | 4,05 |
| 342 | NM_144882 | RIKEN cDNA 2810022L02 gene (2810022L02Rik), mRNA | 4,05 |
| 343 | NM_007409 | alcohol dehydrogenase 1 (class I) (Adh1), mRNA | 4,04 |
| 344 | NM_011204 | protein tyrosine phosphatase, non-receptor type 13 (Ptpn13), mRNA | 4,03 |
| 345 | NAP037326-1 | Unknown | 4,03 |
| 346 | NM_175499 | SLIT and NTRK-like family, member 6 (Slitrk6), mRNA | 4,02 |
| 347 | NM_019696 | carboxypeptidase X 1 (M14 family) (Cpxm1), mRNA | 4,01 |
| 348 | AK042798 | 7 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:A730025J02 product:unknown EST, full insert sequence | 4,00 |
| 349 | NM_023653 | wingless-related MMTV integration site 2 (Wnt2), mRNA | 4,00 |
| 350 | ENSMUST00000069384 | Unknown | 3,99 |
| 351 | NM_013689 | cytoplasmic tyrosine kinase, Dscr28C related (Drosophila) (Tec), mRNA | 3,99 |
| 352 | AK037205 | 6 days neonate skin cDNA, RIKEN full-length enriched library, clone:A030009O20 product:weakly similar to HYPOTHETICAL 484 KDA PROTEIN (FRAGMENT) , full insert sequence | 3,98 |
| 353 | NM_010877 | neutrophil cytosolic factor 2 (Ncf2), mRNA | 3,98 |
| 354 | NM_019417 | PDZ and LIM domain 4 (Pdlim4), mRNA | 3,98 |
| 355 | AK047180 | 10 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:B930032D14 product:hypothetical Zinc finger, C2H2 type containing protein, full insert sequence | 3,97 |
| 356 | AK052877 | 16 days neonate heart cDNA, RIKEN full-length enriched library, clone:D830017E01 product:protein kinase, cGMP-dependent, type I, full insert sequence | 3,97 |
| 357 | AK122468 | mRNA for mKIAA1209 protein | 3,97 |
| 358 | NM_031402 | LCCL domain containing cysteine-rich secretory protein 1 (Lcrisp1), mRNA | 3,96 |
| 359 | AK122347 | mRNA for mKIAA0694 protein | 3,96 |
| 360 | AF183946 | transmembrane glycoprotein cadherin-10 mRNA, complete cds | 3,96 |
| 361 | NM_027354 | RIKEN cDNA 2510040D07 gene (2510040D07Rik), mRNA | 3,95 |
| 362 | BC030896 | platelet-derived growth factor, D polypeptide, mRNA (cDNA clone MGC:31518 IMAGE:4489485), complete cds | 3,95 |
| 363 | AK084988 | 13 days embryo lung cDNA, RIKEN full-length enriched library, clone:D430021N24 product:hypothetical protein, full insert sequence | 3,95 |
| 364 | NM_010140 | Eph receptor A3 (Epha3), mRNA | 3,94 |
| 365 | NM_013646 | RAR-related orphan receptor alpha (Rora), mRNA | 3,93 |
| 366 | NM_007825 | cytochrome P450, family 7, subfamily b, polypeptide 1 (Cyp7b1), mRNA | 3,93 |
| 367 | NM_007912 | epidermal growth factor receptor (Egfr), transcript variant 2, mRNA | 3,93 |
| 368 | D87968 | Mouse mRNA for SHPS-1, complete cds | 3,92 |
| 369 | NM_146023 | ecotropic viral integration site 2b (Evi2b), mRNA | 3,92 |
| 370 | NM_027185 | differentially expressed in FDCP 6 (Def6), mRNA | 3,91 |

| | | | |
|---|---|---|---|
| 371 | AF543214 | strain C57BL/6 CD72 variant (Cd72) mRNA, partial cds | 3,91 |
| 372 | NM_010580 | integrin beta 5 (Itgb5), mRNA | 3,90 |
| 373 | A_52_P517668 | Unknown | 3,90 |
| 374 | A_51_P402908 | Unknown | 3,89 |
| 375 | NM_011607 | tenascin C (Tnc), mRNA | 3,88 |
| 376 | NM_011804 | cellular repressor of E1A-stimulated genes 1 (Creg1), mRNA | 3,87 |
| 377 | NM_011331 | chemokine (C-C motif) ligand 12 (Ccl12), mRNA | 3,87 |
| 378 | NM_026125 | RIKEN cDNA 1110035L05 gene (1110035L05Rik), mRNA | 3,87 |
| 379 | NM_172294 | sulfatase 1 (Sulf1), mRNA | 3,86 |
| 380 | NM_012043 | immunoglobulin superfamily containing leucine-rich repeat (Islr), mRNA | 3,86 |
| 381 | XM_358515 | PREDICTED: RIKEN cDNA 1110006O17 gene (1110006O17Rik), mRNA | 3,85 |
| 382 | BY592157 | BY592157 RIKEN full-length enriched, adult inner ear cDNA clone F930023E21 3', mRNA sequence | 3,85 |
| 383 | NM_023121 | guanine nucleotide binding protein (G protein), gamma transducing activity polypeptide 2 (Gngt2), mRNA | 3,85 |
| 384 | NM_019634 | transmembrane 4 superfamily member 2 (Tm4sf2), mRNA | 3,85 |
| 385 | NM_011157 | proteoglycan 1, secretory granule (Prg1), mRNA | 3,85 |
| 386 | NM_008587 | c-mer proto-oncogene tyrosine kinase (Mertk), mRNA | 3,85 |
| 387 | NM_007969 | extracellular proteinase inhibitor (Expi), mRNA | 3,85 |
| 388 | NM_011210 | protein tyrosine phosphatase, receptor type, C (Ptprc), mRNA | 3,85 |
| 389 | NM_021881 | quaking (Qk), mRNA | 3,85 |
| 390 | NM_146069 | RIKEN cDNA E430025L02 gene (E430025L02Rik), mRNA | 3,85 |
| 391 | NM_173442 | glucosaminyl (N-acetyl) transferase 1, core 2 (Gcnt1), mRNA | 3,84 |
| 392 | AK090301 | 10 days pregnant adult female ovary and uterus cDNA, RIKEN full-length enriched library, clone:G630042G04 product:hypothetical protein, full insert sequence | 3,84 |
| 393 | NM_028608 | GLI pathogenesis-related 1 (glioma) (Glipr1), mRNA | 3,84 |
| 394 | NM_134152 | leupaxin (Lpxn), mRNA | 3,84 |
| 395 | NM_176920 | RIKEN cDNA A930016D02 gene (A930016D02Rik), mRNA | 3,83 |
| 396 | NM_008572 | mast cell protease 8 (Mcpt8), mRNA | 3,83 |
| 397 | NM_133187 | RIKEN cDNA 1110032E23 gene (1110032E23Rik), mRNA | 3,82 |
| 398 | NM_009932 | procollagen, type IV, alpha 2 (Col4a2), mRNA | 3,81 |
| 399 | AK007352 | 10 day old male pancreas cDNA, RIKEN full-length enriched library, clone:1810006K23 product:hypothetical protein, full insert sequence | 3,81 |
| 400 | NM_178881 | expressed sequence C76746 (C76746), mRNA | 3,81 |
| 401 | NM_007577 | complement component 5, receptor 1 (C5r1), mRNA | 3,80 |
| 402 | AK085780 | 16 days neonate heart cDNA, RIKEN full-length enriched library, clone:D830007N24 product:IL-1RRP2 homolog , full insert sequence | 3,77 |
| 403 | NM_010764 | mannosidase 2, alpha B1 (Man2b1), mRNA | 3,77 |
| 404 | AF326559 | C57BL/6J protein tyrosine phosphatase, receptor type, delta A (Ptprd) mRNA, complete cds | 3,77 |
| 405 | NM_010634 | fatty acid binding protein 5, epidermal (Fabp5), mRNA | 3,77 |
| 406 | NAP120816-001 | Unknown | 3,77 |
| 407 | NM_019449 | unc-93 homolog B1 (C elegans) (Unc93b1), mRNA | 3,77 |
| 408 | NM_022018 | niban protein (Niban), mRNA | 3,76 |
| 409 | AK036882 | adult female vagina cDNA, RIKEN full-length enriched library, clone:9930021K24 product:2P DOMAIN K+ CHANNEL TWIK-2 homolog , full insert sequence | 3,74 |
| 410 | NM_007934 | glutamyl aminopeptidase (Enpep), mRNA | 3,73 |
| 411 | NM_009008 | RAS-related C3 botulinum substrate 2 (Rac2), mRNA | 3,72 |

| | | | |
|---|---|---|---|
| 412 | NM_027871 | Rho guanine nucleotide exchange factor (GEF) 3 (Arhgef3), mRNA | 3,72 |
| 413 | BC042790 | , Similar to IQ motif containing GTPase activating protein 2, clone IMAGE:3596508, mRNA, partial cds | 3,72 |
| 414 | NM_172893 | zinc finger CCCH type domain containing 1 (Zc3hdc1), mRNA | 3,71 |
| 415 | NM_011405 | solute carrier family 7 (cationic amino acid transporter, y+ system), member 7 (Slc7a7), mRNA | 3,71 |
| 416 | NM_007765 | collapsin response mediator protein 1 (Crmp1), mRNA | 3,71 |
| 417 | NM_008533 | lymphocyte antigen 78 (Ly78), mRNA | 3,71 |
| 418 | NM_026693 | gamma-aminobutyric acid (GABA-A) receptor-associated protein-like 2 (Gabarapl2), mRNA | 3,70 |
| 419 | NM_177204 | RIKEN cDNA D330017J20 gene (D330017J20Rik), mRNA | 3,70 |
| 420 | NM_011825 | gremlin 2 homolog, cysteine knot superfamily (Xenopus laevis) (Grem2), mRNA | 3,69 |
| 421 | NM_153131 | unc-5 homolog A (C elegans) (Unc5a), mRNA | 3,69 |
| 422 | AK122367 | mRNA for mKIAA0768 protein | 3,68 |
| 423 | TC1263746 | PRP3_HUMAN (O43395) U4/U6 small nuclear ribonucleoprotein Prp3 (Pre-mRNA splicing factor 3) (U4/U6 snRNP 90 kDa protein) (hPrp3), partial (3%) | 3,68 |
| 424 | NM_010696 | lymphocyte cytosolic protein 2 (Lcp2), mRNA | 3,68 |
| 425 | NM_007806 | cytochrome b-245, alpha polypeptide (Cyba), mRNA | 3,67 |
| 426 | NM_028472 | BMP-binding endothelial regulator (Bmper), mRNA | 3,67 |
| 427 | XM_136261 | PREDICTED: similar to hemicentin; fibulin 6 (LOC240793), mRNA | 3,67 |
| 428 | AK089569 | activated spleen cDNA, RIKEN full-length enriched library, clone:F830002F05 product:unclassifiable, full insert sequence | 3,67 |
| 429 | NM_144830 | cDNA sequence BC022145 (BC022145), mRNA | 3,66 |
| 430 | NM_019564 | protease, serine, 11 (Igf binding) (Prss11), mRNA | 3,66 |
| 431 | NM_007652 | CD59a antigen (Cd59a), mRNA | 3,66 |
| 432 | NM_019467 | allograft inflammatory factor 1 (Aif1), mRNA | 3,66 |
| 433 | NM_145463 | RIKEN cDNA 9430059P22 gene (9430059P22Rik), mRNA | 3,65 |
| 434 | Y13832 | mRNA for GT12 protein | 3,65 |
| 435 | NM_026627 | RIKEN cDNA 1700113O17 gene (1700113O17Rik), mRNA | 3,64 |
| 436 | AK053503 | 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone:E130103I17 product:similar to PTH-RESPONSIVE OSTEOSARCOMA B1 PROTEIN , , full insert sequence | 3,64 |
| 437 | AK047388 | 10 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:B930054M11 product:unknown EST, full insert sequence | 3,64 |
| 438 | NM_009804 | catalase (Cat), mRNA | 3,63 |
| 439 | NM_007742 | procollagen, type I, alpha 1 (Col1a1), mRNA | 3,62 |
| 440 | BC057027 | RIKEN cDNA 4933431N12 gene, mRNA (cDNA clone MGC:67187 IMAGE:6825869), complete cds | 3,62 |
| 441 | NM_133871 | interferon-induced protein 44 (Ifi44), mRNA | 3,61 |
| 442 | NM_008079 | galactosylceramidase (Galc), mRNA | 3,61 |
| 443 | X52264 | Mouse ICAM-1 mRNA for intercellular adhesion molecule-1 | 3,60 |
| 444 | NM_023044 | solute carrier family 15, member 3 (Slc15a3), mRNA | 3,60 |
| 445 | AK005731 | adult male testis cDNA, RIKEN full-length enriched library, clone:1700007K13 product:hypothetical protein, full insert sequence | 3,60 |
| 446 | NM_146102 | expressed sequence AU041783 (AU041783), mRNA | 3,60 |
| 447 | NM_010511 | interferon gamma receptor 1 (Ifngr1), mRNA | 3,59 |
| 448 | NM_008964 | prostaglandin E receptor 2 (subtype EP2) (Ptger2), mRNA | 3,59 |
| 449 | NM_027835 | interferon induced with helicase C domain 1 (Ifih1), mRNA | 3,59 |
| 450 | NAP042466-1 | Unknown | 3,59 |
| 451 | AK017277 | 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430405N12 product:unclassifiable, full insert sequence | 3,59 |
| 452 | NM_001001881 | RIKEN cDNA 2510009E07 gene (2510009E07Rik), mRNA | 3,59 |

| | | | |
|---|---|---|---|
| 453 | NM_181728 | ADP-ribosyltransferase 3 (Art3), mRNA | 3,58 |
| 454 | AK009769 | adult male tongue cDNA, RIKEN full-length enriched library, clone:2310043C01 product:tripartite motif protein 12, full insert sequence | 3,58 |
| 455 | NM_010424 | hemochromatosis (Hfe), mRNA | 3,57 |
| 456 | NM_007755 | cytoplasmic polyadenylation element binding protein 1 (Cpeb1), mRNA | 3,57 |
| 457 | AK020883 | adult retina cDNA, RIKEN full-length enriched library, clone:A930022H17 product:hypothetical PDZ domain (also known as DHR or GLGF) containing protein, full insert sequence | 3,57 |
| 458 | NAP007796-001 | Unknown | 3,57 |
| 459 | A_52_P2659 | Unknown | 3,57 |
| 460 | NM_019922 | cartilage associated protein (Crtap), mRNA | 3,56 |
| 461 | BC075677 | cDNA clone IMAGE:30619724, partial cds | 3,56 |
| 462 | NM_007706 | suppressor of cytokine signaling 2 (Socs2), mRNA | 3,56 |
| 463 | NM_011261 | reelin (Reln), mRNA | 3,55 |
| 464 | NM_009912 | chemokine (C-C motif) receptor 1 (Ccr1), mRNA | 3,55 |
| 465 | NM_031376 | phosphoinositide-3-kinase adaptor protein 1 (Pik3ap1), mRNA | 3,55 |
| 466 | NM_021704 | chemokine (C-X-C motif) ligand 12 (Cxcl12), transcript variant 1, mRNA | 3,55 |
| 467 | BC027288 | RIKEN cDNA 2300006M17 gene, mRNA (cDNA clone IMAGE:3589087), partial cds | 3,54 |
| 468 | NM_028732 | RIKEN cDNA 4632428N05 gene (4632428N05Rik), mRNA | 3,53 |
| 469 | NM_007646 | CD38 antigen (Cd38), mRNA | 3,53 |
| 470 | NM_008401 | integrin alpha M (Itgam), mRNA | 3,53 |
| 471 | NM_173749 | RIKEN cDNA E430002G05 gene (E430002G05Rik), mRNA | 3,53 |
| 472 | NM_178685 | protocadherin 20 (Pcdh20), mRNA | 3,51 |
| 473 | NM_011309 | S100 calcium binding protein A1 (S100a1), mRNA | 3,50 |
| 474 | NM_009099 | tripartite motif protein 30 (Trim30), mRNA | 3,50 |
| 475 | NM_194336 | macrophage activation 2 like (Mpa2l), mRNA | 3,49 |
| 476 | BC034073 | avian musculoaponeurotic fibrosarcoma (v-maf) AS42 oncogene homolog, mRNA (cDNA clone IMAGE:4221113), partial cds | 3,49 |
| 477 | AK045646 | adult male corpora quadrigemina cDNA, RIKEN full-length enriched library, clone:B230218B10 product:MESENCHYMAL STEM CELL PROTEIN DSC54 homolog , full insert sequence | 3,49 |
| 478 | NM_025405 | RIKEN cDNA 1110033J19 gene (1110033J19Rik), mRNA | 3,48 |
| 479 | NM_173451 | RIKEN cDNA 9330196J05 gene (9330196J05Rik), mRNA | 3,48 |
| 480 | NM_053246 | docking protein 4 (Dok4), mRNA | 3,48 |
| 481 | AK036010 | 16 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:9630026O22 product:SIMILAR TO VERY LOW DENSITY LIPOPROTEIN RECEPTOR homolog , full insert sequence | 3,48 |
| 482 | NM_023456 | neuropeptide Y (Npy), mRNA | 3,47 |
| 483 | XM_127381 | PREDICTED: hexokinase 3 (Hk3), mRNA | 3,47 |
| 484 | AK010908 | 13 days embryo liver cDNA, RIKEN full-length enriched library, clone:2510004M07 product:regional homology to PLECKSTRIN, SEC7 AND COILED/COIL DOMAINS 4 homolog , full insert sequence | 3,46 |
| 485 | X69620 | Mmusculus mRNA for inhibin beta-B subunit | 3,46 |
| 486 | NM_146139 | vav 3 oncogene (Vav3), mRNA | 3,46 |
| 487 | NM_007426 | angiopoietin 2 (Angpt2), mRNA | 3,46 |
| 488 | NM_013652 | chemokine (C-C motif) ligand 4 (Ccl4), mRNA | 3,46 |
| 489 | AK080287 | 3 days neonate thymus cDNA, RIKEN full-length enriched library, clone:A630031B20 product:weakly similar to HYPOTHETICAL 339 KDA PROTEIN , full insert sequence | 3,45 |
| 490 | TC1261738 | A8A1_MOUSE (P70704) Potential phospholipid-transporting ATPase IA (Chromaffin granule ATPase II) (ATPase class I type 8A member 1) , complete | 3,45 |
| 491 | AK077382 | 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430409C21 product:unknown EST, full insert sequence | 3,44 |

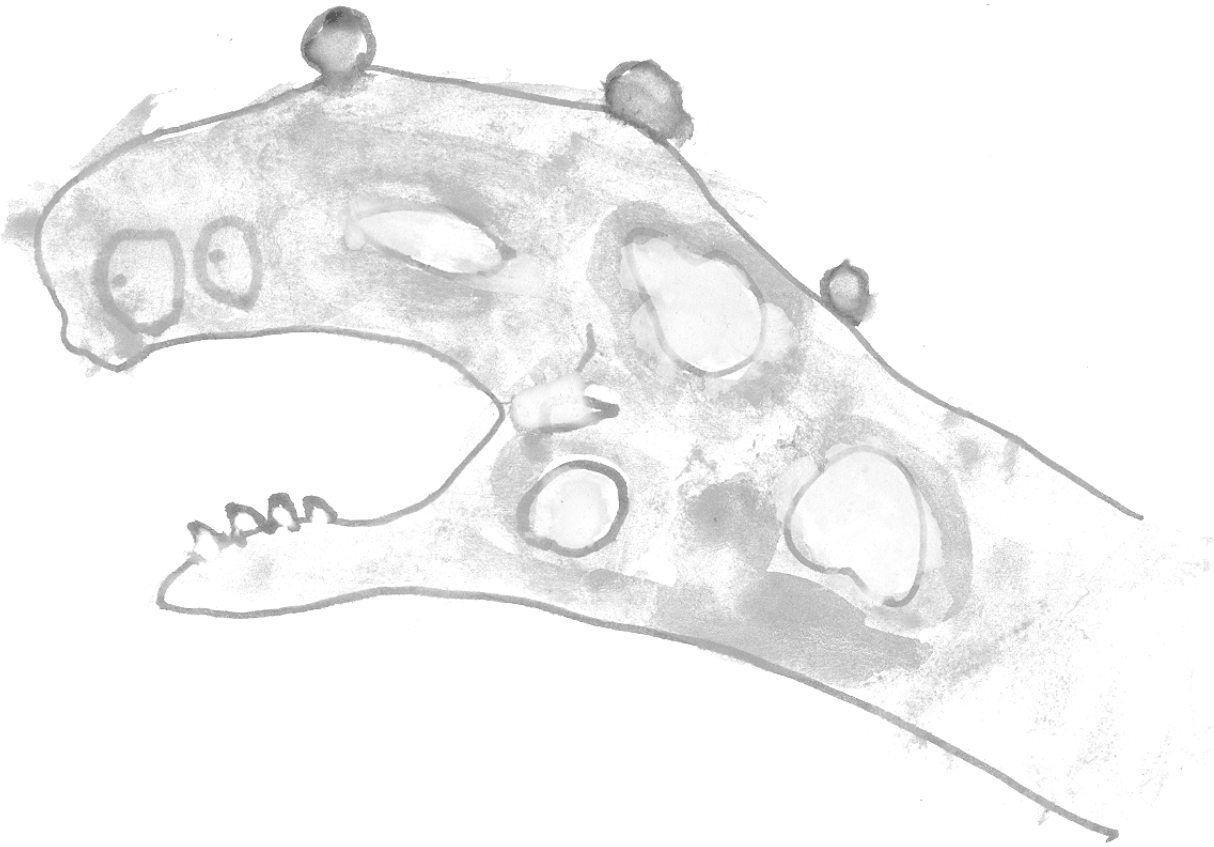
| | | | |
|---|---|---|---|
| 492 | NM_026465 | RIKEN cDNA 2010316F05 gene (2010316F05Rik), mRNA | 3,44 |
| 493 | NM_008048 | insulin-like growth factor binding protein 7 (Igfbp7), mRNA | 3,44 |
| 494 | NM_144915 | RIKEN cDNA E330036I19 gene (E330036I19Rik), mRNA | 3,44 |
| 495 | NM_176930 | RIKEN cDNA C130076O07 gene (C130076O07Rik), mRNA | 3,44 |
| 496 | NM_010496 | inhibitor of DNA binding 2 (Idb2), mRNA | 3,43 |
| 497 | NM_007406 | adenylate cyclase 7 (Adcy7), mRNA | 3,43 |
| 498 | AK017255 | 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430401H09 product:unclassifiable, full insert sequence | 3,43 |
| 499 | NM_009877 | cyclin-dependent kinase inhibitor 2A (Cdkn2a), mRNA | 3,43 |
| 500 | AK042112 | 3 days neonate thymus cDNA, RIKEN full-length enriched library, clone:A630059J03 product:inferred: Traf2 and NCK interacting kinase, splice variant 8 {Homo sapiens}, full insert sequence | 3,42 |
| 501 | BC013712 | cDNA sequence BC013712, mRNA (cDNA clone IMAGE:3155889), partial cds | 3,42 |
| 502 | CG869960 | XT0145 Sanger Institute Gene Trap Library pGT0lxf cDNA, mRNA sequence | 3,42 |
| 503 | NM_177909 | solute carrier family 9 (sodium/hydrogen exchanger), isoform 9 (Slc9a9), mRNA | 3,41 |
| 504 | NM_011206 | protein tyrosine phosphatase, non-receptor type 18 (Ptpn18), mRNA | 3,41 |
| 505 | AK032259 | adult male olfactory brain cDNA, RIKEN full-length enriched library, clone:6430512M05 product:hypothetical protein, full insert sequence | 3,41 |
| 506 | NM_153795 | cDNA sequence BC032204 (BC032204), mRNA | 3,41 |
| 507 | NM_178664 | RIKEN cDNA 6030413G23 gene (6030413G23Rik), mRNA | 3,40 |
| 508 | NM_010180 | fibulin 1 (Fbln1), mRNA | 3,39 |
| 509 | NM_023258 | PYD and CARD domain containing (Pycard), mRNA | 3,39 |
| 510 | AK052292 | 13 days embryo heart cDNA, RIKEN full-length enriched library, clone:D330022C23 product:unclassifiable, full insert sequence | 3,39 |
| 511 | NM_146937 | olfactory receptor 63 (Olfr63), mRNA | 3,39 |
| 512 | NM_013706 | CD52 antigen (Cd52), mRNA | 3,38 |
| 513 | NM_008505 | LIM domain only 2 (Lmo2), mRNA | 3,38 |
| 514 | NM_181277 | procollagen, type XIV, alpha 1 (Col14a1), mRNA | 3,38 |
| 515 | NM_144862 | LIM and senescent cell antigen like domains 2 (Lims2), mRNA | 3,38 |
| 516 | TC1264587 | Q9UA50 (Q9UA50) Engrailed-related 1 protein (Fragment), partial (13%) | 3,37 |
| 517 | AK045458 | adult male corpora quadrigemina cDNA, RIKEN full-length enriched library, clone:B230204J04 product:unknown EST, full insert sequence | 3,37 |
| 518 | NM_026439 | RIKEN cDNA 2610001E17 gene (2610001E17Rik), mRNA | 3,37 |
| 519 | NM_009983 | cathepsin D (Ctsd), mRNA | 3,37 |
| 520 | NM_019510 | transient receptor potential cation channel, subfamily C, member 3 (Trpc3), mRNA | 3,36 |
| 521 | BC049975 | cDNA sequence BC049975, mRNA (cDNA clone IMAGE:4921292), partial cds | 3,35 |
| 522 | NM_177583 | APH1B homolog (Celegans) (Aph1b), mRNA | 3,35 |
| 523 | BC040364 | dystrobrevin alpha, transcript variant 1, mRNA (cDNA clone MGC:25316 IMAGE:4505583), complete cds | 3,35 |
| 524 | TC1339880 | Unknown | 3,34 |
| 525 | NM_024465 | RIKEN cDNA 6330583M11 gene (6330583M11Rik), mRNA | 3,33 |
| 526 | NM_183390 | kelch-like 6 (Drosophila) (Klhl6), mRNA | 3,33 |
| 527 | NM_019929 | SMT3 suppressor of mif two 3 homolog 3 (yeast) (Sumo3), mRNA | 3,33 |
| 528 | NM_178748 | expressed sequence AU040377 (AU040377), mRNA | 3,32 |
| 529 | AK007777 | 10 day old male pancreas cDNA, RIKEN full-length enriched library, clone:1810044J04 product:B cell phosphoinositide 3-kinase adaptor, full insert sequence | 3,32 |
| 530 | AK032356 | adult male olfactory brain cDNA, RIKEN full-length enriched library, clone:6430526E05 product:CORTACTIN-BINDING PROTEIN 2 homolog , full insert sequence | 3,32 |
| 531 | TC1278013 | Q803L7 (Q803L7) Ras homolog gene family, member G, partial (5%) | 3,32 |

| | | | |
|---|---|---|---|
| 532 | NM_023635 | RAB27A, member RAS oncogene family (Rab27a), mRNA | 3,32 |
| 533 | NM_175316 | solute carrier organic anion transporter family, member 2b1 (Slco2b1), mRNA | 3,32 |
| 534 | NM_011303 | dehydrogenase/reductase (SDR family) member 3 (Dhrs3), mRNA | 3,30 |
| 535 | NM_010835 | homeo box, msh-like 1 (Msx1), mRNA | 3,30 |
| 536 | NM_015749 | transcobalamin 2 (Tcn2), mRNA | 3,29 |
| 537 | NM_007421 | adenylosuccinate synthetase like 1 (Adssl1), mRNA | 3,29 |
| 538 | AK031349 | 13 days embryo male testis cDNA, RIKEN full-length enriched library, clone:6030410K16 product:hypothetical Cysteine-rich region containing protein, full insert sequence | 3,28 |
| 539 | NM_134133 | RIKEN cDNA 2010002N04 gene (2010002N04Rik), mRNA | 3,28 |
| 540 | NM_008278 | hydroxyprostaglandin dehydrogenase 15 (NAD) (Hpgd), mRNA | 3,28 |
| 541 | NM_008908 | peptidylprolyl isomerase C (Ppic), mRNA | 3,28 |
| 542 | NM_020025 | UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase, polypeptide 2 (B3galt2), mRNA | 3,28 |
| 543 | NM_011243 | retinoic acid receptor, beta (Rarb), mRNA | 3,27 |
| 544 | NM_009242 | secreted acidic cysteine rich glycoprotein (Sparc), mRNA | 3,27 |
| 545 | XM_358544 | PREDICTED: RIKEN cDNA C630002N23 gene (C630002N23Rik), mRNA | 3,27 |
| 546 | NM_028773 | RIKEN cDNA 1200013B08 gene (1200013B08Rik), mRNA | 3,27 |
| 547 | BC023075 | RIKEN cDNA 6430411K14 gene, mRNA (cDNA clone IMAGE:5361628), containing frame-shift errors | 3,27 |
| 548 | NAP000805-003 | Unknown | 3,26 |
| 549 | NM_134038 | solute carrier family 16 (monocarboxylic acid transporters), member 6 (Slc16a6), mRNA | 3,26 |
| 550 | XM_484079 | PREDICTED: similar to keratin associated protein 9-1; keratin-associated protein 91; keratin-associated protein 9-1 (LOC432600), mRNA | 3,26 |
| 551 | NM_008741 | neuron specific gene family member 2 (Nsg2), mRNA | 3,26 |
| 552 | NM_008394 | interferon dependent positive acting transcription factor 3 gamma (Isgf3g), mRNA | 3,26 |
| 553 | NM_025829 | eukaryotic translation initiation factor 4E member 3 (Eif4e3), mRNA | 3,26 |
| 554 | NM_178440 | myosin IG (Myo1g), mRNA | 3,25 |
| 555 | NM_010169 | coagulation factor II (thrombin) receptor (F2r), mRNA | 3,25 |
| 556 | NM_133930 | cysteine-rich with EGF-like domains 1 (Creld1), mRNA | 3,25 |
| 557 | NM_008235 | hairy and enhancer of split 1 (Drosophila) (Hes1), mRNA | 3,25 |
| 558 | NM_010740 | complement component 1, q subcomponent, receptor 1 (C1qr1), mRNA | 3,25 |
| 559 | NM_009166 | sorbin and SH3 domain containing 1 (Sorbs1), mRNA | 3,25 |
| 560 | BC079850 | cDNA clone IMAGE:5706103 | 3,25 |
| 561 | BC058161 | CCAAT/enhancer binding protein (C/EBP), alpha, mRNA (cDNA clone MGC:65349 IMAGE:5053056), complete cds | 3,25 |
| 562 | NM_028186 | naked cuticle 2 homolog (Drosophila) (Nkd2), mRNA | 3,24 |
| 563 | NM_023056 | RIKEN cDNA 1810009M01 gene (1810009M01Rik), mRNA | 3,24 |
| 564 | NM_009606 | actin, alpha 1, skeletal muscle (Acta1), mRNA | 3,24 |
| 565 | NM_011518 | spleen tyrosine kinase (Syk), mRNA | 3,24 |
| 566 | NM_010181 | fibrillin 2 (Fbn2), mRNA | 3,24 |
| 567 | NM_011116 | phospholipase D3 (Pld3), mRNA | 3,24 |
| 568 | NM_029007 | expressed sequence AW125753 (AW125753), mRNA | 3,23 |
| 569 | NM_013880 | phospholipase C-like 2 (Plcl2), mRNA | 3,23 |
| 570 | NM_031159 | apolipoprotein B editing complex 1 (Apobec1), mRNA | 3,23 |
| 571 | BC057864 | RIKEN cDNA F730004D16 gene, mRNA (cDNA clone MGC:67746 IMAGE:5318856), complete cds | 3,21 |
| 572 | NM_019482 | pannexin 1 (Panx1), mRNA | 3,21 |
| 573 | NM_026416 | S100 calcium binding protein A16 (S100a16), mRNA | 3,21 |

| | | | |
|---|---|---|---|
| 574 | NM_133751 | RIKEN cDNA 1700060H10 gene (1700060H10Rik), mRNA | 3,20 |
| 575 | NM_133670 | sulfotransferase family 1A, phenol-preferring, member 1 (Sult1a1), mRNA | 3,20 |
| 576 | AK017266 | 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430404K19 product:hypothetical Fibrillar collagen C-terminal domain containing protein, full insert sequence | 3,20 |
| 577 | NM_011823 | G protein-coupled receptor 34 (Gpr34), mRNA | 3,20 |
| 578 | AF020313 | proline-rich protein 48 mRNA, partial cds | 3,20 |
| 579 | BC052446 | cDNA clone MGC:63400 IMAGE:6412797, complete cds | 3,20 |
| 580 | NM_011356 | frizzled-related protein (Frzb), mRNA | 3,20 |
| 581 | NM_025791 | RIKEN cDNA 0610006I08 gene (0610006I08Rik), mRNA | 3,19 |
| 582 | NM_011503 | syntaxin binding protein 2 (Stxbp2), mRNA | 3,19 |
| 583 | NM_016756 | cyclin-dependent kinase 2 (Cdk2), mRNA | 3,19 |
| 584 | NM_172648 | interferon activated gene 205 (Ifi205), mRNA | 3,19 |
| 585 | TC1247223 | P70232 (P70232) L1-like protein (Chl1-like protein), partial (53%) | 3,19 |
| 586 | AK005223 | adult male cerebellum cDNA, RIKEN full-length enriched library, clone:1500012A13 product:CLATHRIN COAT ASSEMBLY PROTEIN AP19 (CLATHRIN COAT ASSOCIATED PROTEIN AP19) (GOLGI ADAPTOR AP-1 19 KDA ADAPTIN) (HA1 19 KDA SUBUNIT) (CLATHRIN | 3,18 |
| 587 | NM_029612 | SLAM family member 9 (Slamf9), mRNA | 3,18 |
| 588 | NM_133193 | interleukin 1 receptor-like 2 (Il1rl2), mRNA | 3,18 |
| 589 | NM_199146 | tripartite motif protein 30-like (LOC209387), mRNA | 3,18 |
| 590 | NM_009899 | chloride channel calcium activated 1 (Clca1), mRNA | 3,18 |
| 591 | AK052809 | 10 days lactation, adult female mammary gland cDNA, RIKEN full-length enriched library, clone:D730020K15 product:HRIHFB2003 PROTEIN (FRAGMENT) homolog , full insert sequence | 3,18 |
| 592 | NM_013482 | Bruton agammaglobulinemia tyrosine kinase (Btk), mRNA | 3,18 |
| 593 | NM_010141 | Eph receptor A7 (Epha7), mRNA | 3,17 |
| 594 | NM_007779 | colony stimulating factor 1 receptor (Csf1r), mRNA | 3,17 |
| 595 | NM_178743 | solute carrier family 26, member 11 (Slc26a11), mRNA | 3,17 |
| 596 | NM_011196 | prostaglandin E receptor 3 (subtype EP3) (Ptger3), mRNA | 3,17 |
| 597 | BC079550 | cDNA clone MGC:90763 IMAGE:6852953, complete cds | 3,17 |
| 598 | NM_025359 | transmembrane 4 superfamily member 13 (Tm4sf13), mRNA | 3,17 |
| 599 | AK045478 | adult male corpora quadrigemina cDNA, RIKEN full-length enriched library, clone:B230206C19 product:unknown EST, full insert sequence | 3,16 |
| 600 | NM_172961 | 4-aminobutyrate aminotransferase (Abat), mRNA | 3,16 |
| 601 | NM_016919 | procollagen, type V, alpha 3 (Col5a3), mRNA | 3,16 |
| 602 | NM_010407 | hemopoietic cell kinase (Hck), mRNA | 3,16 |
| 603 | NM_008516 | leucine rich repeat protein 1, neuronal (Lrrn1), mRNA | 3,15 |
| 604 | NM_054098 | tumor necrosis factor, alpha-induced protein 9 (Tnfaip9), mRNA | 3,15 |
| 605 | NM_027840 | RIKEN cDNA 9130017C17 gene (9130017C17Rik), mRNA | 3,15 |
| 606 | NM_022325 | cathepsin Z (Ctsz), mRNA | 3,15 |
| 607 | NM_145515 | MAP/microtubule affinity-regulating kinase 1 (Mark1), mRNA | 3,15 |
| 608 | AK032603 | adult male olfactory brain cDNA, RIKEN full-length enriched library, clone:6430702L12 product:hypothetical protein, full insert sequence | 3,15 |
| 609 | NM_007781 | colony stimulating factor 2 receptor, beta 2, low-affinity (granulocyte-macrophage) (Csf2rb2), mRNA | 3,15 |
| 610 | NM_207246 | RAS, guanyl releasing protein 3 (Rasgrp3), mRNA | 3,15 |
| 611 | NM_177350 | collomin (Colm), mRNA | 3,14 |
| 612 | NM_010315 | guanine nucleotide binding protein (G protein), gamma 2 subunit (Gng2), mRNA | 3,14 |
| 613 | NM_010279 | glial cell line derived neurotrophic factor family receptor alpha 1 (Gfra1), mRNA | 3,13 |
| 614 | NM_009898 | coronin, actin binding protein 1A (Coro1a), mRNA | 3,13 |

| | | | |
|---|---|---|---|
| 615 | AK006128 | adult male testis cDNA, RIKEN full-length enriched library, clone:1700019L09 product:ATP-binding cassette, sub-family C (CFTR/MRP), member 3, full insert sequence | 3,12 |
| 616 | BC063327 | heparan sulfate 6-O-sulfotransferase 2, mRNA (cDNA clone MGC:74252 IMAGE:6531439), complete cds | 3,12 |
| 617 | AK020640 | adult male urinary bladder cDNA, RIKEN full-length enriched library, clone:9530073D23 product:unknown EST, full insert sequence | 3,12 |
| 618 | NM_175188 | membrane-associated ring finger (C3HC4) 1 (March1), mRNA | 3,12 |
| 619 | NM_001005508 | RIKEN cDNA 6030405P05 gene (6030405P05Rik), mRNA | 3,11 |
| 620 | BC058264 | serologically defined colon cancer antigen 33, mRNA (cDNA clone MGC:65389 IMAGE:6488400), complete cds | 3,11 |
| 621 | AK035376 | adult male urinary bladder cDNA, RIKEN full-length enriched library, clone:9530027C22 product:unclassifiable, full insert sequence | 3,11 |
| 622 | AK028875 | 10 days neonate skin cDNA, RIKEN full-length enriched library, clone:4732465K06 product:CXC chemokine ligand 16, full insert sequence | 3,11 |
| 623 | XM_357051 | PREDICTED: similar to sperm protein 3111; sperm protein SSP3111 (LOC383435), mRNA | 3,11 |
| 624 | AK020887 | adult retina cDNA, RIKEN full-length enriched library, clone:A930024K11 product:EUKARYOTIC TRANSLATION INITIATION FACTOR 2 ALPHA KINASE 1 (EC 271-) (HEME-REGULATED EUKARYOTIC INITIATION FACTOR EIF-2-ALPHA KINASE) (HEME-REGULATED | 3,11 |
| 625 | NM_011853 | 2'-5' oligoadenylate synthetase 1B (Oas1b), mRNA | 3,10 |
| 626 | AK051724 | 12 days embryo spinal ganglion cDNA, RIKEN full-length enriched library, clone:D130067L24 product:unknown EST, full insert sequence | 3,10 |
| 627 | NM_133838 | EH-domain containing 4 (Ehd4), mRNA | 3,09 |
| 628 | NM_033037 | cysteine dioxygenase 1, cytosolic (Cdo1), mRNA | 3,09 |
| 629 | NM_145133 | Traf2 binding protein (T2bp), mRNA | 3,08 |
| 630 | AK020213 | 15 days embryo male testis cDNA, RIKEN full-length enriched library, clone:8030469F12 product:unknown EST, full insert sequence | 3,07 |
| 631 | AK079320 | 16 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:9630021C19 product:unknown EST, full insert sequence | 3,07 |
| 632 | NM_199018 | START domain containing 8 (Stard8), mRNA | 3,07 |
| 633 | A_52_P335218 | Unknown | 3,07 |
| 634 | BY713104 | BY713104 RIKEN full-length enriched, 13 days embryo head cDNA clone 3110033I20 5' | 3,06 |
| 635 | NM_133198 | liver glycogen phosphorylase (Pygl), mRNA | 3,06 |
| 636 | NM_029000 | interferon inducible GTPase 1 (Iigs1), mRNA | 3,05 |
| 637 | TC1339296 | Q9ERK2 (Q9ERK2) Neprilysin-like peptidase gamma, partial (5%) | 3,05 |
| 638 | NM_197986 | RIKEN cDNA 1110007F12 gene (1110007F12Rik), mRNA | 3,05 |
| 639 | NM_007467 | amyloid beta (A4) precursor-like protein 1 (Aplp1), mRNA | 3,05 |
| 640 | NM_007470 | apolipoprotein D (Apod), mRNA | 3,03 |
| 641 | NM_026524 | Mid1 interacting protein 1 (gastrulation specific G12-like (zebrafish)) (Mid1ip1), mRNA | 3,03 |
| 642 | NM_145509 | RIKEN cDNA 5430435G22 gene (5430435G22Rik), mRNA | 3,03 |
| 643 | NM_198093 | engulfment and cell motility 1, ced-12 homolog (C elegans) (Elmo1), mRNA | 3,03 |
| 644 | NM_172633 | cerebellin 2 precursor protein (Cbln2), mRNA | 3,03 |
| 645 | NM_175254 | RIKEN cDNA 9330180L21 gene (9330180L21Rik), mRNA | 3,03 |
| 646 | NM_177836 | expressed sequence AW046396 (AW046396), mRNA | 3,03 |
| 647 | AK032736 | 12 days embryo male wolffian duct includes surrounding region cDNA, RIKEN full-length enriched library, clone:6720422L05 product:proprotein convertase subtilisin/kexin type 5, full insert sequence | 3,03 |
| 648 | NAP057003-1 | Unknown | 3,02 |
| 649 | NM_025422 | RIKEN cDNA 1110055L24 gene (1110055L24Rik), mRNA | 3,02 |
| 650 | NM_009308 | synaptotagmin 4 (Syt4), mRNA | 3,02 |
| 651 | NM_145134 | SPRY domain-containing SOCS box 4 (Ssb4), mRNA | 3,02 |
| 652 | NM_134032 | homeo box B2 (Hoxb2), mRNA | 3,02 |

| | | | |
|---|---|---|---|
| 653 | NM_018804 | synaptotagmin 11 (Syt11), mRNA | 3,02 |
| 654 | NM_009506 | vascular endothelial growth factor C (Vegfc), mRNA | 3,02 |
| 655 | AK003667 | 18-day embryo whole body cDNA, RIKEN full-length enriched library, clone:1110013L07 product:inferred: COTE1 PROTEIN {Homo sapiens}, full insert sequence | 3,01 |
| 656 | AB125594 | Skt-a mRNA for Sickle tail-a, complete cds | 3,01 |
| 657 | AK021081 | adult male corpus striatum cDNA, RIKEN full-length enriched library, clone:C030014I23 product:hypothetical protein, full insert sequence | 3,01 |
| 658 | BC060096 | RIKEN cDNA 3110041P15 gene, mRNA (cDNA clone IMAGE:6400746), partial cds | 3,01 |
| 659 | AK033138 | 15 days embryo male testis cDNA, RIKEN full-length enriched library, clone:8030445B08 product:WWP1 (FRAGMENT) homolog , full insert sequence | 3,01 |
| 660 | BC058110 | autism susceptibility candidate 2, mRNA (cDNA clone IMAGE:6835298), partial cds | 3,01 |
| 661 | NM_009881 | chromodomain protein, Y chromosome-like (Cdyl), mRNA | 3,01 |
| 662 | NM_146073 | zinc finger, DHHC domain containing 14 (Zdhhc14), mRNA | 3,01 |
| 663 | NM_011216 | protein tyrosine phosphatase, receptor type, O (Ptpro), mRNA | 3,00 |
| 664 | NM_026954 | tumor suppressor candidate 1 (Tusc1), mRNA | 3,00 |



## Summary

The aim of this thesis was to investigate and identify genetic fatcors involved in Hirschsprung disease (HSCR). HSCR is a congenital disorder characterized by intestinal obstruction due to an absence of enteric ganglia along variable lengths of the intestinal tract. Firstly, we determined the role of the major HSCR gene – *RET* plays in sporadic cases of the disease where no Coding Sequence (CDS) mutations were identified (Appendices 1 and 2). Secondly, we aimed at finding novel (*RET* modifying) loci that contribute to the HSCR phenotype (Appendices 3 and 4).

This thesis begins with the review of the current literature on HSCR (chapters 1-3). The first chapter of this overview deals with the genetic aspects of HSCR. Familiality, phenotypes (definition, classification) and epidemiology (incidence, sex ratio, segment length variations) as well as the heterogeneity and the model of gene segregation are discussed. Emphasis is laid on the lately discovered regulatory mutations in non-coding sequences at the *RET* locus. Their properties and contribution to HSCR are compared with CDS mutations among different risk categories present in HSCR. The second chapter of this review deals with developmental aspects of HSCR. A brief description is given of the individual stages of ENS development as well as the cells and tissues involved in this process. Subsequently, the most important signaling pathways are described, in particular the *RET* and *EDNRB* signalling routes and their interactions during ENS development. Furthermore, an attempt to combine the knowledge on molecular genetics, developmental aspects and cell biology.

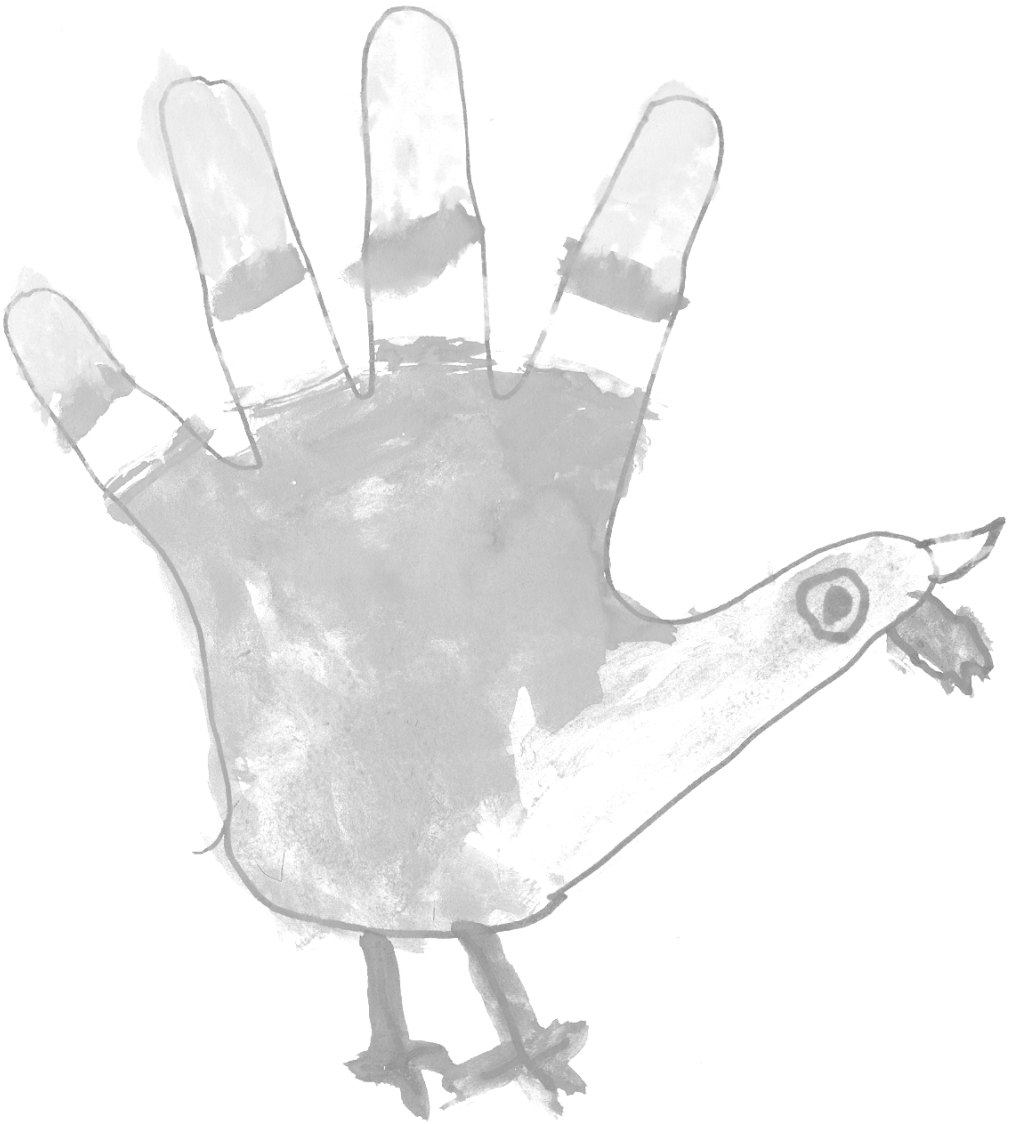
In the appendices our own experimental data (Appendices 1, 2 and 4), as well as a collaborative study (Appendix 3) are presented. Appendix 1 documents the finding, in the Dutch HSCR population, of a common disease-associated haplotype located in the 5'region of the *RET*

locus. An associated haplotype was identified, consisting of 6 markers and spanning 27 Kb (4 Kb 5' UTR and 23 Kb intron 1). The haplotype was present in 56% of our patients, but in only 16% of the controls. Furthermore, odds ratios for HSCR were increased in particular for patients homozygous for the risk haplotype (OR>20 for homozygotes, OR>2 for heterozygotes), suggesting a dose-dependent action of the mutation(s) present at this haplotype.

The work described in Appendix 2 is a continuation of the study described in Appendix 1. In order to more precisely define the HSCR-associated region and to identify a candidate disease-associated variant(s), we sequenced the shared common genomic region from 10 kb upstream the *RET* gene through intron 1 (in total 33 kb) in a patient homozygous for the common risk haplotype and in a control individual homozygous for the most common non-risk haplotype. A comparison of these sequences revealed eighty-four sequence differences. Eight of these variations proved to be in regions highly conserved between different vertebrates and within putative transcription factor binding sites. We, therefore, considered these as candidate disease-associated variants. Subsequent genotyping of these eight variants revealed a strong disease association for six of the eight markers. These six markers also showed the largest distortions in allele transmission. Interspecies comparison showed that only one of the six variations was located in a region also conserved in a non-mammalian species, making it the most likely candidate HSCR-associated variation.

In the collaborative study described in Appendix 3 we performed a genome-wide linkage study on a two-branched multigenerational Dutch HSCR family. We report the finding of suggestive linkage to the region 4q31-q32 (LOD – 2,7). The region spans 12.2 Mb and none of the genes previously implicated in HSCR map to this region. As the disease penetrance is low in the studied family, we hypothesized that this region might be necessary but not sufficient for presentation of the HSCR phenotype.

Appendix 4 describes a gene expression profiling study on cultured Neural Crest Stem Cells (NCSCs). NCSCs were isolated from E.11-E.14 mouse embryo intestines. Cells were cultured in selective media. After selection they were cultured under distinct conditions. Different culture conditions were applied, comparing a population of NCSCs stimulated with the RET ligand – GDNF, to cells in which *RET* was not activated. To identify genes differentially expressed upon RET stimulation, gene expression profiling was performed using Agilent Whole Mouse Genome Arrays. By this approach we are getting insight in the proteins that may be important in the HSCR etiology, assuming that these proteins are modifiers of the *RET*-dependent phenotype.



# Samenvatting

Dit proefschrift beschrijft de rol van genetische factoren in het ontstaan van de ziekte van Hirschsprung (ook wel afgekort als HSCR). HSCR is een aangeboren afwijking gekenmerkt door de afwezigheid van zenuwcellen in de darm (aganglionoses), met name in het laatste (distale) deel van de dikke darm. Door de afwezigheid van zenuwcellen ontbreekt de darmperistaltiek waardoor de darm vernauwt. Dat maakt dat voedsel niet kan passeren en zich ophoopt voor het aganglionaire gedeelte van de darm. Patiënten zijn dan ook gewoonlijk dusdanig ernstig geconstipeerd dat operatief moet worden ingegrepen.

Er zijn een aantal genen bekend die, als ze gemuteerd zijn bijdragen aan of de oorzaak zijn van het ontstaan van HSCR. Het belangrijkste gen is *RET*. Ik heb in dit promotie onderzoek gekeken naar de rol van *RET* in patiënten bij wie geen mutatie werd gevonden in de coderende sequentie van het gen. (Appendices 1 en 2). Verder is gezocht naar nieuwe (onbekende) genen die bij zouden kunnen dragen aan het ontstaan van HSCR (Appendices 3 en 4).

Dit proefschrift begint met een overzicht van de huidige literatuur over HSCR (Hoofdstukken 1-3). Het eerste hoofdstuk van dit overzicht bespreekt de genetische aspecten van de aandoening: het familiair voorkomen, de verschillende fenotypes (definities, classificatie), de epidemiologie (verspreiding, sekseverhouding, variaties in lengte van het aangedane darmsegment) genetische heterogeniteit, het feit dat mutaties in verschillende genen tot een zelfde fenotypisch beeld aanleiding kunnen geven, en het model van overerving. Nadruk wordt gelegd op de meest recente ontwikkelingen, namelijk de vaststelling van het voorkomen van zeer sterke associaties met niet coderende varianten op het *RET* locus. Varianten die bijdragen aan het fenotype in vrijwel alle patiënten. Hun eigenschappen en bijdrage aan HSCR worden vergeleken met coderende RET mutaties voor de verschillende risicocategorieën in de Nederlandse HSCR

populatie. Het tweede hoofdstuk van dit overzicht gaat over de ontwikkelingsaspecten van HSCR. Er wordt een korte beschrijving gegeven van de verschillende fasen van de ontwikkeling van het zenuwstelsel van de darm, het "enteric nervous system (ENS)" alsook van cellen en weefsels die betrokken zijn bij dit proces. Vervolgens worden de belangrijkste signaal transductie routes in het ontwikkelingsproces van het ENS beschreven, te weten de *RET* en de *EDNRB* signaal transductie routes, evenals hun wisselwerkingen gedurende de ENS ontwikkeling. Het derde hoofdstuk van het overzicht is een poging om onze kennis van de moleculaire genetica, van de ontwikkelingsaspecten en van de celbiologie met elkaar te verbinden.

Appendix 1 laat zien dat in de Nederlandse populatie van HSCR patiënten een gemeenschappelijk ziekte-gerelateerd haplotype voorkomt, wat betekent dat het in deze patiëntengroep om één (dan wel meerdere aan elkaar gekoppelde) veel voorkomende ziekte mutatie gaat. Het haplotype (of wel de mutatie) bleek aanwezig bij 56% van onze patiënten, terwijl dit bij de controlegroep slechts bij 16% bleek te zijn. Bovendien waren bijna alle patiënten homozygoot voor het haplotype (d.w.z. dat op beide chromosomen de mutatie voorkwam) en bleek dat de kans om de ziekte te ontwikkelen sterk afhing van het homozygoot zijn voor het haplotype (OR>20 voor homozygoten, OR>2 voor heterozygoten). Dit suggereert een dosis-afhankelijke werking van de mutatie(s) in dit hapoltype. Het gemeenschappelijk haplotype strekt zich uit over het 5' gebied van de *RET* locus en is 27 Kb groot (4 Kb 5' UTR en 23 Kb intron 1).

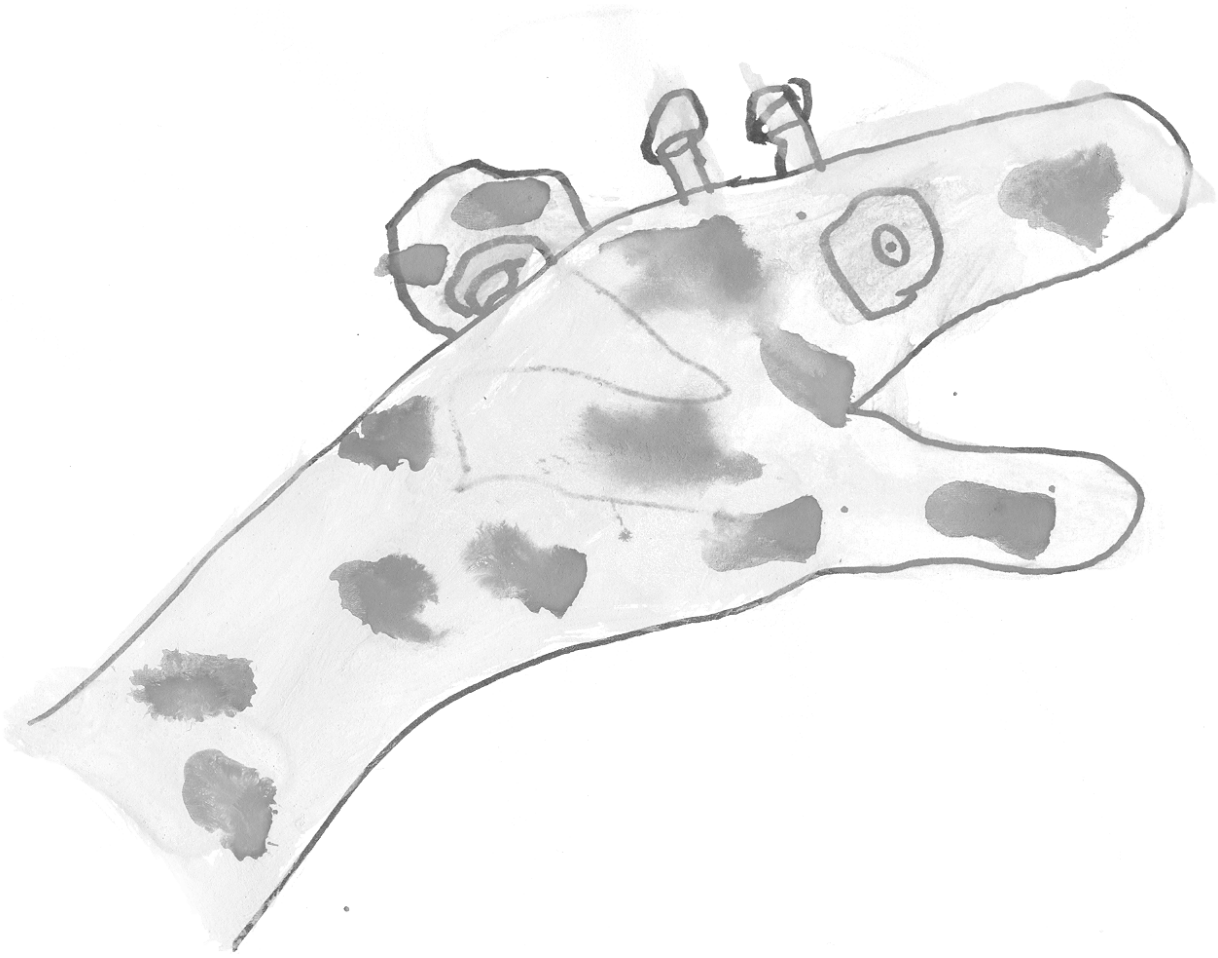
Het in Appendix 2 beschreven werk is een vervolg op het onderzoek dat in Appendix 1 is beschreven. Om het met HSCR geassocieerde gebied beter te definiëren en (een) kandidaat ziekte gerelateerde variant(en) te identificeren, is de basen volgorde bepaald van het gemeenschappelijk gedeelde gebied (vanaf 10kb 5'van *RET* tot en met exon 2, in totaal 33 kb).

Dit is gedaan in een patiënt homozygoot voor het risico haplotype en in een controle persoon homozygoot oor het niet-risico haplotype. Uit vergelijking van deze sequenties kwamen 84 verschillen. Acht van deze variaties bleken te liggen in gebieden die evolutionair geconserveerd zijn gebleven tussen verschillende gewervelde dieren (chimpansee, rat, muis en kip) en die bindingsplaatsen voor transcriptie factoren bevatten. Wij hebben daarom deze acht beschouwd als kandidaat ziekte-gerelateerde varianten. Haploytypering van deze acht varianten liet zien dat zes van de acht sterke associatie met de ziekte lieten zien. Slechts één van die zes variaties was gelegen in een gebied dat geconserveerd was tot in de kip. Dit maakte deze variant dan ook tot de meest aannemelijke kandidaat.

In het onderzoek dat is beschreven in Appendix 3, uitgevoerd in samenwerking met de groep van Prof. Dr. B. Oostra (Rotterdam) hebben we een genoomwijd koppelingsonderzoek uitgevoerd in een familie waarin HSCR zich in meerdere takken van de familie voordoet. We melden de vondst van een veelbelovende koppeling met het gebied 4q31-q32 (LOD – 2,7). Het gebied omvat 12.2 Mb en vele genen. Geen van de bekende HSCR genen is in dit gebied gelegen. Daar de ziekte-penetratie in de bestudeerde familie laag is veronderstelden we dat dit gebied nodig zou kunnen zijn, maar niet voldoende is voor het ontwikkelen van het HSCR fenotype.

Appendix 4 beschrijft het maken en analyseren van een gen-expressie profiel van gekweekte Neural Crest Stem Cells (NCSCs). Deze NCSCs zijn geïsoleerd uit darmen van muizenembryo's van 11 tot 14 dagen oud. Na kweek in een selectief medium, zijn ze onder verschillende condities verder gekweekt. Eén populatie van NCSCs werd gestimuleerd met de *RET* ligand *GDNF* en vervolgens vergeleken met cellen waarin *RET* niet was geactiveerd. Om genen te identificeren die verschillend tot expressie komen na RET stimulatie, zijn expressie-array analyses uitgevoerd met Agilent whole mouse genome arrays. De gevonden verschillen

worden beschreven en er worden verklaringen voor gezocht. Met deze benadering kan mogelijk inzicht worden verkregen in de eiwitten die belangrijk zijn voor de ontwikkeling van het ENS, alsook in de etiologie van HSCR



## Acknowledgments

At certain point of your Ph.D. studentship you have to realize that it is time to say ‘stop’, otherwise you’ll never finish, as there are always so many important things that still have to be done. Research is a continuous process without a true finish line but you have to draw one and start writing the things you have accomplished. This time has come for me as well and fortunately, as you can see, I’ve made it! Though, this book would not have been here without the help of many kind people, whom, hereby, I would like to thank warmly.

First of all, I would like to thank Robert, the supervisor of my Ph.D. studentship. I now have the opportunity to express the great respect I feel for your personally, your attitude to other people and the way you carry on the research. Thanks Robert for the great 5 years I had in Groningen at the Department of Medical Genetics. I am sure that the opportunities I was given will be of great help in my further scientific career. I thank also Charles Buys, my second promotor, for his help, in particular in writing the papers in beautifully, correct English but also for all the good word he always had for me. Next, I’d like to thank Jan for being such a great help, particularly in the beginning of my Ph.D. when I had to learn all the new methods and when I was finding myself in complete new environment. However, I’m also very much obliged to Jan for the entire thing he did for me throughout the years, for doing part of the experiments in my project and most of all for teaching my students. Thank you Jan!

My students are next “aan de buurt” – Krista and Agnes. Thank you both for being simply great students, you were of big help as can be seen in this thesis (special thanks to Agnes for the sequencing work!). It was always nice working with you both. I also thank all the other people from our Department as I was helped many times in different matters by many of you. Thanks to Ludolf for the sequencing lessons, sequencing itself and the many funny discussions we had. Thanks to Gerwin for his help with ABI’s. I always enjoyed talking to you especially on the

parties after a few… Thanks also to Pieter for carrying out some of the experiments. Then, many thanks for our secretaries for organizing all the formal things: Mentje, Ineke and Ria. Thanks to Edwin for taking care of some of my computer problems and many of you know I had a few… Damn windows '98 !$#@%, damn office '97 %@$#. At this point I wish you all a fast change to XP and the newest MS Office! Where Edwin is, there is also Arja and I wish to thank you for the many useful tips in finalizing my Ph.D. Similar thanks go to Marga. I was always scared watching Arja and Marga writing the thesis specially after working hours and on the weekends. Fortunately, I think I went through this period relatively easy and relaxed…Further, I'd like to thank Frans and Krista Kooi for their help in the array projects. Thanks also to Ivan for the collaboration on the "overlapping" projects and of course thanks for help in Western blotting, luc-assays, antibodies…

And last but certainly not least I want mention all the people with whom I had some more contact with than only at work. First, to our small Portuguese girls – Maria and Ana, it was great fun to meet, drink and talk to you. Big, big hug and kiss for being my paranimfen, thank you for arranging things for my defense!! I hope you both will like your 4 year stay here and finish your (excellent I'm sure) thesis on time. To Ana additionally, I wish a fast progress in learning Polish! And of course Reneé! I always liked you and I wish to thank you for all the help (like e.g. answering my stupid questions concerning Excel) and all the hilarious discussions we had. Then I'd like to thank Yunia, master student from Indonesia, Ramune, coming from Lithuania, a neighbor of Poland and Marina our great financial advisor from Russia! And I'm not forgetting of course my Chinese friend – Jianghua, who gave me so many great breaks at work, chatting about everything and nothing (also thanks for the lunches!). It was fun to know you all and do all the things we've done (mostly drinking but with a note of dancing and eating). I wish you all the best in your careers as well as in the private lifes.

At this point I want to acknowledge all the other members of The Department I was honoured to meet and whom I haven't listed above. It was really nice to spend these years with you all.

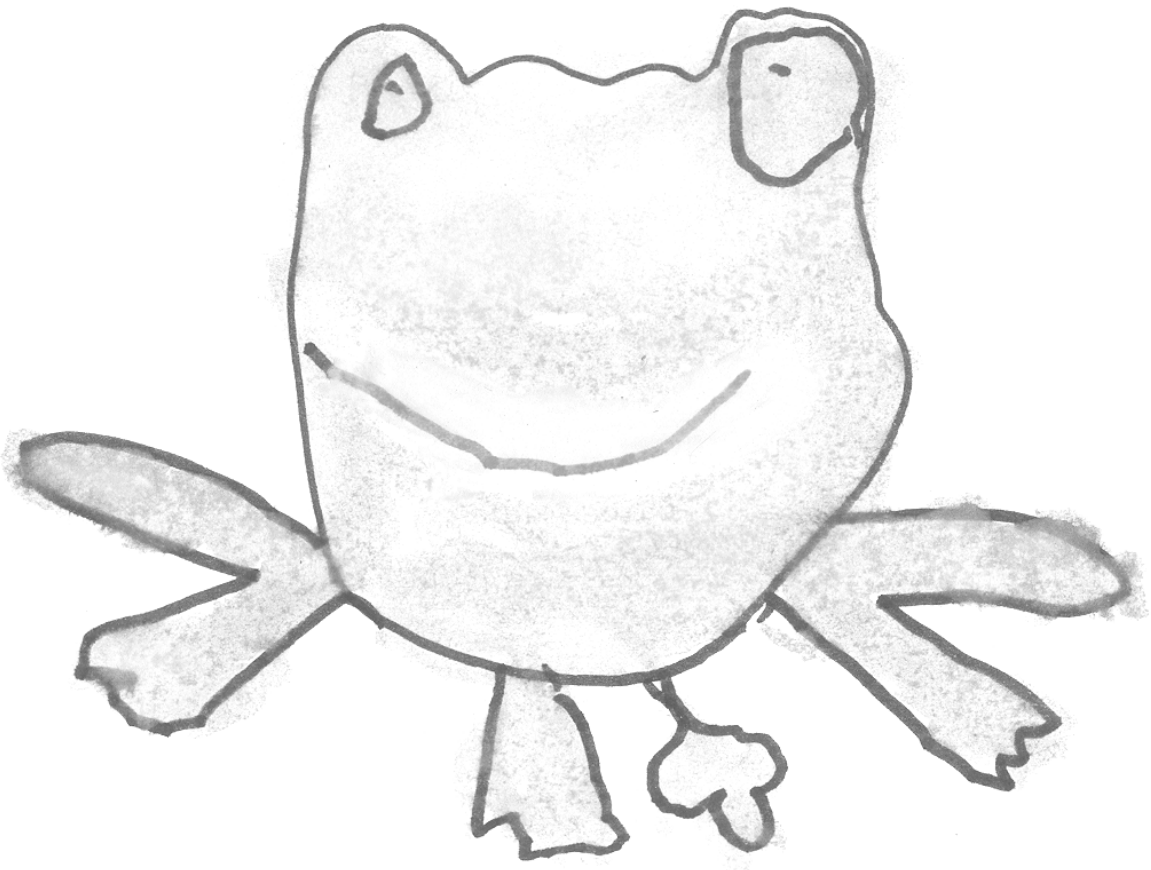
Furthermore, I thank all the people I was collaborating with during these years of my studentship. Thanks go to Anita Wiersema and Gerald de Haan from the Stem Cell Biology Department. It was a new but very nice experience for me to work with/on stem cells. Gerald, also thanks for being a member of my reading committee. Thanks to Alan Burns and Jean-Marie for being my hosts during these few weeks in London and supervising my work on chicken embryos(?!). Thanks to the Italian friends Francesca, Paola and Isa for sharing with us their scientific results, ideas and DNA's. Special thanks to Isa for agreeing and being (such a nice) member of the reading committee. I'd like to thank also our American collaborators – Aravinda, Eileen and Stacey for making ambitious, international projects possible and all the other members of the Hirschsprung consortium for giving me the opportunity of presenting the data at the ASHG in 2005.

Now, time has come to acknowledge and thank all my Polish friends from Groningen. Without them, this period of my Dutch life certainly would have been much "poorer" and completely different as most of my social life I spent with you. So, first I'd like to thank to Ozimki couple, Paulina and Lukasz. We've been together here almost from the beginning till the end. You already have a new, second, after-Groningen life in Poland, and I have to admit, you are doing great! Thank you both for all the parties, dinners, trips and more importantly, for all the kindness and the constant help. Next couple is Maciolki - Ada and Tomek plus their sons – Mateusz and Igor. Thank you for being always straightforward and honest. Thanks for the time we spent together as well as for the time our children did. Third couple is Mr. and Mrs. Lubelscy and their son, Pawel. I can just repeat my aforementioned words, it was always great going out with you,

spending holidays together, eating and talking altogether. Pawel has certainly become one of best friends of Natalia – Thank you so much! Subsequently, I thank Andrzej and Marta Lulko for interesting discussions, always having the freshest news on any subject and of course Marta's great kitchen! Not to forget our clones – Agnieszka and Grzesiek, most of the evenings spent with you were surprising in some way (enough to mention – playing puns), as was your whole life here. It was quite some experience to meet you. I hope you're doing well in the UK! At the end I want to acknowledge all the other Polish people I've met in Groningen and whom I'm not able to mention here. Big thanks, you've made my time better here.

The very end I have left for my family. First I'd like to thank my parents in law for the enormous support we've always had from your side! Many things would not have been possible without your help. And then, I want thank my father for bringing me up and educating me in many, many matters, especially when we lost our mom. I know I was often a horrible young man to cope with, but I think I'm on the right track now - thank you!

Very special thanks, hugs and kisses for my wife Agnieszka and my daughter Natalia. It is hard to imagine how life would be without you here! I'd like to thank Agnieszka for standing next to me throughout all these years here. I realize how difficult it sometimes is to live with a growler and mule like me (and then even in a foreign country!). Therefore, I thank you twice as much for supporting me always and most of all for the great job you do, bringing up our daughter. To Natalia I just want to say that I love you very much and I'm very happy to have you around.



# Publications

Plaza-Menacho I[*], **Burzynski GM[*]**, de Groot JW, Eggen BJ, Hofstra RM.
The RET receptor: genetics, signaling and therapeutics in human neural crest disorders.
Trends Genet (accepted), *equal contribution

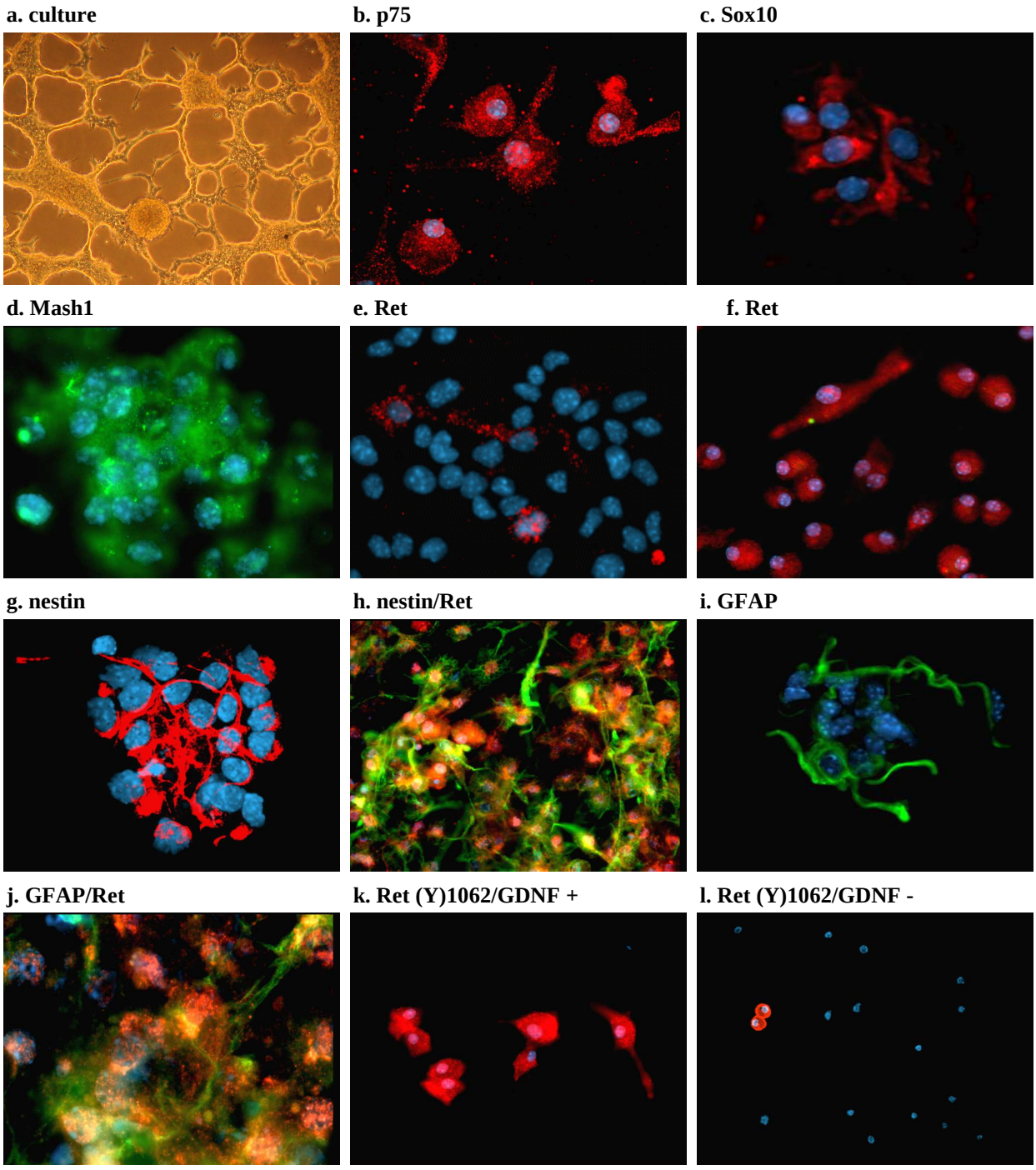
Brooks AS, Leegwater PA, **Burzynski GM**, Willems PJ, de Graaf B,van Langen I, Heutink P, Oostra BA, Hofstra RM, Bertoli-Avella AM. A novel susceptibility locus for Hirschsprung disease maps to 4q31.3-q32.3
J Med Genet 2006 Jul;43(7):e35

Brooks AS, Bertoli-Avella AM, **Burzynski GM**, Breedveld GJ, Osinga J, Boven LG, Hurst JA, Mancini GM, Lequin MH, de Coo RF, Matera I, de Graaff E, Meijers C, Willems PJ, Tibboel D, Oostra BA, Hofstra RM. Homozygous nonsense mutations in KIAA1279 are associated with malformations of both the central and the enteric nervous system.
Am J Hum Genet 2005 Jul;77(1):120-6

**Burzynski GM**, Nolte IM, Bronda A, Bos KK, Osinga J, Plaza Menacho I, Twigt B, Maas S, Brooks AS, Verheij JB, Buys CH, Hofstra RM. Identifying Candidate Hirschsprung Disease-Associated RET Variants.
Am J Hum Genet 2005 Mar 9;76(5): 850-858

Menacho IP, Koster R, van der Sloot AM, Quax WJ, Osinga J, van der Sluis T, Hollema H, **Burzynski GM**, Gimm O, Buys CH, Eggen BJ, Hofstra RM. 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 Mar 1;65(5):1729-37

**Burzynski GM**, Nolte IM, Osinga J, Ceccherini I, Twigt B, Maas S, Brooks A, Verheij J, Plaza Menacho I, Buys CH, Hofstra RM. 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 Aug;12(8):604-12




**Appendix 4, Figure 1. NCSC's growth in culture (a) and immunostainings with appropriate biomarkers. (b) Nerve Growth Factor (p75) Receptor – neural crest marker; (c) Sox10 transcription factor, crucial for neural crest induction, migration and differentiation; (d) Mash1, transcription factor that serves as an early neuronal marker; (e) staining with Ret H-300 antibody in 2nd day of culture; (f) staining with Ret H-300 antibody in 5th day of culture; (g) nestin, neuronal marker; (h) Ret/nestin double immunostaining; (i); Glial Fibrillary Acidic Protein (GFAP), glia cells marker; (j) Ret/GFAP double immunostaining; (k) staining with Ret (P)Y1062 antibody, cells stimulated with Gdnf; (l) staining with Ret (P)Y1062 antibody, cells unstimulated;**