Identification of Variants in RET and IHH Pathway Members in a Large Family With History of Hirschsprung Disease

Sribudiani, Yunia; Chauhan, Rajendra K.; Alves, Maria M.; Petrova, Lucy; Brosens, Erwin; Harrison, Colin; Wabbersen, Tara; de Graaf, Bianca M.; Rugenbrink, Tim; Burzynski, Grzegorz

Published in:
Gastroenterology

DOI:
10.1053/j.gastro.2018.03.034

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2018

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 09-05-2020
Identification of Variants in RET and IHH Pathway Members in a Large Family With History of Hirschsprung Disease


BACKGROUND & AIMS: Hirschsprung disease (HSCR) is an inherited congenital disorder characterized by absence of enteric ganglia in the distal part of the gut. Variants in ret proto-oncogene (RET) have been associated with up to 50% of familial and 35% of sporadic cases. We searched for variants that affect disease risk in a large, multigenerational family with history of HSCR in a linkage region previously associated with the disease (4q31.3–q32.3) and exome wide. METHODS: We performed exome sequencing analyses of a family in the Netherlands with 5 members diagnosed with HSCR and 2 members diagnosed with functional constipation. We initially focused on variants in genes located in 4q31.3–q32.3; however, we also performed an exome-wide analysis in which known HSCR or HSCR-associated gene variants predicted to be deleterious were prioritized for further analysis. Candidate genes were expressed in HEK293, COS-7, and Neuro-2a cells and analyzed by luciferase and immunoblot assays. Morpholinos were designed to target exons of candidate genes and injected into 1-cell stage zebrafish embryos. Embryos were allowed to develop and stained for enteric neurons. RESULTS: Within the linkage region, we identified 1 putative splice variant in the lipopolysaccharide responsive beige-like anchor protein gene (LRBA). Functional assays could not confirm its predicted effect on messenger RNA splicing or on expression of the mab-21 like 2 gene (MAB21L2), which is embedded in LRBA. Zebrafish that developed following injection of the lrba morpholino had a shortened body axis and subtle gut morphological defects, but no significant reduction in number of enteric neurons compared with controls. Outside the linkage region, members of 1 branch of the family carried a previously unidentiﬁed RET variant or an in-frame deletion in the glial cell line derived neurotrophic factor gene (GDNF), which encodes a ligand of RET. This deletion was located 6 base pairs before the last codon. We also found variants in the Indian hedgehog gene (IHH) and its mediator, the transcription factor GLI family zinc ﬁnger 3 (GLI3). When expressed in cells, the RET-P399L variant disrupted protein glycosylation and had altered phosphorylation following activation by GDNF. The deletion in GDNF prevented secretion of its gene product, reducing RET activation, and the IHH-Q51K variant reduced expression of the transcription factor GLI1. Injection of morpholinos that target hh reduced the number of enteric neurons to 13% ± 1.4% of control zebrafish. CONCLUSIONS: In a study of a large family with history of HSCR, we identiﬁed variants in LRBA, RET, the gene encoding the RET ligand (GDNF), IHH, and a gene encoding a mediator of IHH signaling (GLI3). These variants altered functions of the gene products when expressed in cells and knockout of hh reduced the number of enteric neurons in the zebrafish gut.

Keywords: ENS; Neural Development; Genetic Causes of HSCR; Family Study.

Hirschsprung disease (HSCR) is a congenital disorder characterized by the absence of enteric ganglia in variable lengths of the distal gut. As a consequence, functional networks of neurons and glia, the intrinsic innervations of the gastrointestinal tract comprising the enteric nervous system (ENS), cannot be established, leading to intestinal obstruction by dysregulated smooth muscle contraction/relaxation.

HSCR is considered to be an inherited disease. This assumption is based on several lines of evidence, including familial occurrence (~5%), elevated risk of occurrence in

*Authors share co-ﬁrst authorship.

Abbreviations used in this paper: bp, base pair; ENS, enteric nervous system; GDNF, glial cell–derived neurotrophic factor; GFP, green ﬂuorescent protein; GLI1, GLI family zinc ﬁnger 1; GLI3, GLI family zinc ﬁnger 3; HEK, human embryonic kidney cells; Ihh, Hedgehog; hpf, hours post fertilization; HSCR, Hirschsprung disease; IHH, Indian Hedgehog; LRBA, lipopolysaccharide responsive beige-like anchor; MAB21L2, Mab-21-Like 2; mRNA, messenger RNA; Mut, Mutant; Neuro-2a, neuroblastoma; RET, Rearranged during Transfection; RT-PCR, reverse transcription polymerase chain reaction; WT, wild-type.

Most current article

© 2018 by the AGA Institute. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

0016-5085

https://doi.org/10.1053/j.gastro.2018.03.034
WHAT YOU NEED TO KNOW

BACKGROUND AND CONTEXT
Hirschsprung disease (HSCR) arises due to failure of the enteric neurons to colonize the gut. It is an inherited disorder, but the genetic cause is unknown in the majority of cases.

NEW FINDINGS
In a Dutch multigenerational family with history of HSCR, the authors identified variants in RET, GDNF, IHH and GLI3 that disrupt the function of their encoded proteins, contributing to disease development.

LIMITATIONS
The variants identified in this family are rare and unlikely to explain the missing heritability seen in the majority of HSCR cases.

IMPACT
This study confirms RET as the major HSCR gene and shows that a combination of rare variants in GDNF, IHH and GLI3, modulates clinical expression of the disease phenotype.

siblings (relative risk as high as 200), association with chromosomal abnormalities, and the existence of many naturally occurring animal models with colonic aganglionosis. The mode of inheritance can vary from dominant with reduced penetrance, mostly found in nonsyndromic familial HSCR cases, to recessive, in families with syndromic HSCR. Sporadic HSCR cases also have been reported and are believed to be multifactorial and polygenic in nature, suggesting the involvement of several genes in concert.

The search for genes involved in HSCR has been extensive and ranged from classical linkage to genome-wide association studies and candidate gene approaches. To date, mutations in approximately 20 genes have been identified. However, the REarranged during Transfection (RET) gene is still considered to be the major HSCR gene, as 50% of familial cases and 15% to 35% of sporadic cases carry a mutation in its coding or messenger RNA (mRNA) splicing regions. The RET locus (10q11) was the first one to be identified for HSCR, and, subsequently, coding variants in this gene were reported to give rise to a dominant form of the disease with incomplete penetrance (72% for male and 51% for female individuals). Moreover, all association studies conducted on sporadic HSCR cases showed the highest association with a low-penetrant variant present in intron 1 of RET (odds ratio = 2 when heterozygous and odds ratio = 20 when homozygous). Taken together, these genetic findings showed that RET variants can be high- or low-penetrant, but more importantly, they demonstrated that almost every patient with HSCR has a variant in RET. Thus, RET seems to be crucial in developing HSCR, but it is also likely that one or more modifier genes, as well as environmental factors, are involved in disease pathogenesis, even when a high penetrant RET mutation is found.

Linkage analysis conducted on 12 multiplex HSCR families corroborated this idea. Although 11 families showed linkage to the RET locus, only half of them (6) carried a severe RET coding mutation. Intriguingly, the remaining 5 families also showed linkage to 9q31, suggesting the involvement of a modifier gene at this locus. After these early findings, subsequent studies were conducted to search for modifier loci in HSCR. Sib-pair analysis resulted in the identification of 2 additional loci at 3p21 and 19q12, haplotype sharing in a large Mennonite kindred identified a new locus on 16q23, linkage analysis in a multigenerational Dutch family identified a locus at 4q31.3-q32.3, and genome-wide association studies also have identified loci at 7q21.11 and 8p12. However, combinations of distinct rare mutations resulting in HSCR are not frequently reported.

In this article, we focus on one family in which a 12.2-Mb interval suggestive for linkage was identified on 4q31.3-q32.3 (chr4: 142,197,646–158,353,484 [Hg19]), but no pathogenic variants in the known HSCR genes have been found. Based on the pedigree, incomplete penetrance of a disease-associated variant was expected, suggesting the involvement of several genes. In an attempt to identify the genetic cause of HSCR in this family, we have now used whole-exome sequencing to search for yet unidentified pathogenic rare variants or modifier genes. We determined segregation patterns for candidate variants, and performed in vitro and in vivo studies to test the involvement of the identified candidate genes in disease pathogenesis, revealing the complex genetic nature of HSCR.

Materials and Methods

Patient Information
A multigenerational Dutch family was included in this study. This family is composed of 5 individuals diagnosed with HSCR (IV-3, V-1, V-2, V-3, and V-4), and 2 diagnosed with functional constipation (III-1 and IV-2) (Figure 1). A detailed description of the phenotypes has been previously reported. Written informed consent was obtained from the parents for diagnostic analysis.

Exome Sequencing and Variant Prioritization
Two HSCR-affected individuals (V-1 and V-4) from different branches of the family were initially selected for exome sequencing. In a later stage of the study, IV-4 and IV-5 were also included (Figure 1). Three micrograms of DNA from each of the individuals was used. Details about execution and data analysis can be found in supplementary data.

Validation of Candidate Variants and Family Screening
Genomic DNA was isolated from peripheral blood leukocytes using a standard protocol previously described. Candidate variants were validated by Sanger sequencing as previously described. Segregation analysis was performed using family members for which DNA was available (II-2, III-1, III-2, IV-1, IV-2, IV-3, IV-4, IV-5, V-1, V-2, V-3, and V-4).
Vector Design and Site Direct Mutagenesis

Vectors used are described in detail in supplementary data.

Whole-Mount In Situ Hybridization for lrba, mab21l2, and ihh in Zebrafish

lrba (lipopolysaccharide responsive beige-like anchor), mab21l2 (Mab-21-Like 2), and ihh (Indian hedgehog) genes were amplified from total mRNA collected from zebrafish embryos at 48 hours post fertilization (hpf), by reverse transcription polymerase chain reaction (RT-PCR) using a One-Step RT-PCR Kit (Qiagen, Valencia, CA). Primers used are described in Supplementary Table 1. A detailed protocol can be found in supplementary data.

lrba, mab21l2, and ihh Morphant Analysis in Zebrafish

Two splice blocking morpholinos were designed to target exon 13 (AGTTGGTTTAGTCTCTTACCGAGAC) and exon 24 (ACTGCATACTAACCGAAGAAGAAGT) of lrba. The effectiveness of these morpholinos was confirmed by RT-PCR. A previously described translation blocking morpholino for mab21l2 (ACTGTAGACCGGAGTTTCGCAGTAC) was used (Gene Tools, Philomath, OR). A mab21l2 mutant line (au12 allele) was also analyzed. The ihh morpholino (GGAGACGCATTCCACCGCAAGCG) was designed to target the transcription start site of ihh, as previously described. Morphants were generated by injecting 100 μM of each morpholino into 1-cell-stage zebrafish embryos. Morphant/mutant and control embryos were allowed to develop until 120 hpf and were fixed and stained for ENS neurons using the HuC/D antibody (Invitrogen, Carlsbad, CA), as previously reported. A p53 control morpholino (Gene Tools) was coinjected in all morphant and control embryos, to suppress apoptotic effects induced as a secondary effect of the morpholinos, as described elsewhere. To determine the number of enteric neurons present, a 10-segment length of the gut to the vent was counted. The numbers in the text represent percent of control ± SEM for at least 5 separate embryos per morpholino/mutant genotype. Significance was determined by the Student t test with significance assessed when P < .0005.

Cell Culture and Transfections

Human embryonic kidney (HEK293) cells, COS-7 cells (CV-1 [simian] in origin, and carrying the SV40 genetic material), and control fibroblasts were cultured in Dulbecco’s modified Eagle’s minimal essential medium (GIBCO, Waltham, MA) containing 10% fetal bovine serum (GIBCO) and penicillin/streptomycin (GIBCO). The neuroblastoma cell line (Neuro-2a) (CCL-131; American Type Culture Collection, Manassas, VA) was cultured according to the protocol of the American Type Culture Collection. All cell lines were incubated at 37°C, and supplied with 5% CO₂. Transfection was performed as described before.

Exon Trap Assays

The exon trap assays were performed as described before. SD6 and SA2 primers are described in Supplementary Table 1.

Luciferase Assays

Neuro-2a cells were transfected with 1 μg of SV40-P or LRBA-wild-type (WT)/mutant (Mut) vectors and cotransfected with 10 ng of internal control, pRL-SV40-Renilla Luciferase (Promega, Madison, WI). Luciferase activity was measured and...
quantified as described before. SV40-E (without any promoter) was used as a negative control and RET-WT-enhancer was used as a positive control. Luciferase assays were performed in 3 independent, triplicate experiments (n = 9).

**Activation of the Indian Hedgehog Signaling**

HEK293 cells cultured in a 6-well plate were transiently transfected with pCMV-IHH-FLAG-WT/Mut. After 24 hours, the medium of transfected cells (conditioned medium) was collected and filtered using a 0.45-μm filter; 200,000 to 300,000 control human fibroblast cells were cultured in a 6-well plate for 24 hours. After this period, the medium was replaced by 1 mL fresh complete medium and 500 μL conditioned medium (containing secreted IHH-WT or IHH-Mut). Conditioned medium derived from nontransfected HEK293 cells was used as a negative control. Medium supplemented with 20 μM Purmaphamine (Calbiochem, San Diego, CA) was used as a positive control for activation of the Hedgehog (Hh) signaling: 500 μL conditioned medium was concentrated using an M-10 filter (Millipore, Bedford, MA) and used for Western blot to determine the levels of IHH-WT and IHH-Mut protein secreted into the medium.

**Glial Cell–Derived Neurotrophic Factor Stimulation and Western Blot**

HEK293 cells were transiently cotransfected with pCMV-RET-WT/Mut (P399L), pCMV-GFRα1, and pNE-green fluorescent protein (GFP). After 24 hours, cells were treated with 50 ng/mL glial cell–derived neurotrophic factor (GDNF) (PeproTech EC, London, UK) for 15 minutes. To test the effect of the GDNF deletion, conditioned medium collected from HEK293 cells transfected with a GDNF-WT or GDNF-Mut constructs was collected in a similar way as described for IHH, and used to stimulate HEK293 cells transfected with pCMV-RET-WT, pCMV-GFRα1, and pNE-GFP. An amount of 500 μL of conditioned medium was also concentrated as described for IHH, and used to determine the levels of GDNF-WT and GDNF-Mut protein secreted into the medium by Western blot. Cell lysis, protein quantification, and Western blot were performed as previously described. Primary and secondary antibodies used are described in Supplementary Table 2.

**RNA Isolation and qReal-time-PCR**

RNA isolation, complementary DNA preparation, and quantitative real-time (qRT) PCR are described in supplementary data.

**Statistical Analysis**

All results are expressed as the mean ± standard deviation or standard error of the mean. All data were analyzed using a 2-tailed Student t test or the χ² test. P < .05 was considered to be statistically significant.

**Results**

**A Putative Splice Variant in LRBA Was Found in the Linkage Interval**

Exome sequencing data collected from patients V-1 and V-4 were first analyzed to detect variants present in the linkage interval previously identified. Exons that were not totally covered within this region (7 exons), were Sanger sequenced. From the exome analysis, only 1 rare variant (Exome Aggregation Consortium: 0.002534, and Genome of the Netherlands database: 0.009), predicted to be deleterious was found: a putative splice variant affecting exon 20 of the LRBA gene (NM_001199282.2:c.2444A>G) (Table 1 and Supplementary Table 3). LRBA was also found to be expressed in mouse gut, leading us to consider it the best candidate gene for this family. Sanger sequencing confirmed the presence of the LRBA variant in all family members for which DNA was available (n = 11), and segregation patterns were determined (Figure 1; Supplementary Table 4).

**lrba Is Not Required for ENS Development in Zebrafish**

To investigate a possible role for LRBA in ENS development, we used the zebrafish as a model system. A single zebrafish ortholog for lrba was identified in an Ensemble gene search, which showed strong sequence similarity, as well as genome organization, to its human ortholog (82% homology). Whole-mount in situ hybridization revealed that lrba has a comparatively restricted expression pattern in zebrafish (Figure 2A). At 24 hpf, lrba expression was identified along the yolk sac boundary and weakly in the hindbrain. At 48 hpf, lrba was still weakly present in the hindbrain, and no apparent expression was detected elsewhere in the embryo (Figure 2A). A similar pattern of expression was detected at 72 and 96 hpf. However, at 72 hpf, lrba expression appeared in the intestinal bulb, and it was maintained at 96 hpf (Figure 2A). We also designed 2 different splice blocking morpholinos to suppress expression of this gene in zebrafish. Examination of lrba morphants at 120 hpf revealed a shortened body axis and subtle gut morphological defects. However, no significant reduction in the number of enteric neurons was detected when compared with controls, as the number of neurons in lrba morphants was 97.2% ± 4.8% of control (n = 17; Figure 2B).

**Lack of Splicing Effect and Enhancing Defects for the LRBA Variant**

The LRBA variant identified in this family (NM_001199282.2:c.2444A>G) is predicted to affect mRNA splicing of exon 20 by 1 of the 5 splice site prediction programs included in the Alamut splicing prediction module (http://www.interactive-biosoftware.com/alamut-visual/). To confirm pathogenicity of this variant, exon trap assays were performed, but no splice defect was detected. Similar-size bands of spliced product were observed in both the WT and Mut situations (Figure 3A).

Within intron 42 of LRBA, another gene called Mab-21-Like 2 (MAB21L2) is found (Figure 3B). A previous study has shown that expression of MAB21L2 can be controlled in a tissue-specific manner by several enhancer elements present within LRBA. This led us to hypothesize that exon 20 of LRBA might work as one of these enhancers, and that the variant identified in this gene might disturb expression of MAB21L2. Because the role of MAB21L2 in ENS...
development is unknown, we first investigated the expression pattern of this gene in zebrafish, by performing whole-mount in situ hybridization at different stages of embryonic development. *mab21l2* was already known to be strongly expressed in the hindbrain and cranial neural crest in zebrafish.\textsuperscript{25} Our results confirmed this expression pattern, and revealed strong expression of this gene in the pharyngeal arches, especially at 48 hpf (Supplementary Figure 1). We also observed a significant expression of *mab21l2* in the gut mesoderm from 48 hpf onward, which has not been reported before (Supplementary Figure 1). To investigate the importance of *MAB21L2* for ENS development, we initially used a morpholino-based approach to knock down this gene in zebrafish. Subsequently, we obtained a *mab21l2* mutant line (au12 allele).\textsuperscript{26} Morphant and mutant fish had identical phenotypes with defects in development of pharyngeal arches and intestinal smooth muscle, as previously reported.\textsuperscript{25} Critically, we detected a significant reduction in the number of enteric neurons in *mab21l2* homozygous mutants, which was 28.2% ± 3.8% of that seen in WT siblings (n = 12; Figure 3C). These results led us to conclude that *mab21l2* is required for ENS development in zebrafish, and that it might therefore also contribute to HSCR pathogenesis in humans. To further explore this possibility, we investigated whether the *LRBA* variant found in this family acts as a prospective enhancer element, thereby affecting *MAB21L2* expression. For this, a series of luciferase assays was performed using exon 20 of *LRBA* and its flanking regions, containing the WT or the Mut sequence (c.2444A>G). Our results showed that exon 20 of *LRBA* could indeed enhance the promoter activity of SV40 (Figure 3D). However, no difference was detected when the *LRBA* variant (c.2444A>G) was introduced (Figure 3D). Considering this result, we were unable to link *MAB21L2* to HSCR, as it is unlikely that the *LRBA* variant identified has an enhancing effect. However, to rule out the possibility that a mutation in *MAB21L2* was missed in the exome analysis, we Sanger sequenced all exons of *MAB21L2* and its regulatory regions (16 Kb upstream), in patients V-1 and V-4. No rare variants were identified that could be associated to the disease phenotype.

**Variants in RET, IHH, GLI3, and GDNF Were Detected Outside the Linkage Interval**

As we were unable to confirm pathogenicity of the variant found in *LRBA*, and could not find a link between *MAB21L2* and HSCR pathogenesis, we hypothesized that variants outside the linkage interval would be the ones determining disease development. Therefore, we focused on nonshared rare variants outside the linkage interval present in any of the 2 individuals sequenced (V-1 and V-4). A de novo analysis was also performed using the trio composed of IV-4, IV-5, and V-4. Initially, these 2 analyses aimed to find variants present in genes previously associated with HSCR.\textsuperscript{3,4} Moreover, variants were prioritized based on function and deleteriousness. With this approach, we found a previously unidentified rare variant in *RET* (NM_020975.4:c.1196C>T; p.L399Y) in patient V-1 (Table 1 and Supplementary Table 3). Segregation analysis showed

### Table 1. Rare variants identified in patients V-1 (a) and/or V-4 (b)

<table>
<thead>
<tr>
<th>Sample Gene</th>
<th>HGVS cDNA</th>
<th>Sample</th>
<th>Gene</th>
<th>HGVS protein</th>
<th>dbSNP</th>
<th>Location</th>
<th>Effect</th>
<th>Exon</th>
<th>HGS Variatin Society (ExAC) MAF</th>
<th>GoNL MAF</th>
<th>Allelic Fraction</th>
<th>Parent</th>
<th>ClinVar</th>
</tr>
</thead>
<tbody>
<tr>
<td>RET</td>
<td>NM_020975.4:c.1196C&gt;T</td>
<td>a, b</td>
<td>p.Pro399Leu</td>
<td>-</td>
<td>M</td>
<td>0</td>
<td>0.0015737</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>0</td>
<td>SCV000328919</td>
<td></td>
</tr>
<tr>
<td>NRP2</td>
<td>NM_201266.1:c.1000C&gt;A</td>
<td>a, b</td>
<td>p.Arg333Glu</td>
<td>-</td>
<td>F</td>
<td>0.0003475</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>0</td>
<td>SCV000328913</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LRBA</td>
<td>NM_001199282.2:c.2444A&gt;G</td>
<td>a, b</td>
<td>p.Leu815Ser</td>
<td>rs140666848</td>
<td>M; F</td>
<td>0.0025336</td>
<td>0.018</td>
<td>Yes</td>
<td>No</td>
<td>SCV000328914</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LRBA</td>
<td>NM_001190468.1:c.676_681delGGATGT</td>
<td>a, b</td>
<td>p.Gly226_227del</td>
<td>-</td>
<td>DN</td>
<td>0</td>
<td>0</td>
<td>Yes</td>
<td>No</td>
<td>SCV000328917</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OR1F1</td>
<td>NM_012360.1:c.47G&gt;A</td>
<td>a, b</td>
<td>p.Asp1183His</td>
<td>rs201361018</td>
<td>F</td>
<td>0.0001642</td>
<td>0 No</td>
<td>No</td>
<td>SCV000328921</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLUH</td>
<td>NM_015229.3:c.3547G&gt;A</td>
<td>a, b</td>
<td>p.Gly62Glu</td>
<td>rs41298555</td>
<td>F</td>
<td>0</td>
<td>No</td>
<td>Yes</td>
<td>SCV000328910</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLUH</td>
<td>NM_001199282.2:c.2444A&gt;G</td>
<td>a, b</td>
<td>p.Asn815Ser</td>
<td>rs140666848</td>
<td>M; F</td>
<td>0.0025336</td>
<td>0.018</td>
<td>Yes</td>
<td>No</td>
<td>SCV000328914</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PELP1</td>
<td>NM_014389.2:c.2696T&gt;</td>
<td>a, b</td>
<td>p.Met721Val</td>
<td>rs200062536</td>
<td>F</td>
<td>0.001568</td>
<td>0.006</td>
<td>No</td>
<td>No</td>
<td>SCV000328911</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PELP1</td>
<td>NM_014389.2:c.2161A&gt;G</td>
<td>a, b</td>
<td>p.Asn815Ser</td>
<td>rs140666848</td>
<td>M; F</td>
<td>0.0025336</td>
<td>0.018</td>
<td>Yes</td>
<td>No</td>
<td>SCV000328914</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAB21L2</td>
<td>NM_001199282.2:c.151C&gt;G</td>
<td>a, b</td>
<td>p.Pro707Ser</td>
<td>rs121917716</td>
<td>F</td>
<td>0.0001976</td>
<td>0.002</td>
<td>Yes</td>
<td>No</td>
<td>SCV000328910</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** Build hg19, # public databases are 1000 Genomes, ESP6500, and Genome of the Netherlands (GoNL). All variants are heterozygous. The variants in *RET* and *NRP2* are known deleterious variants. cDNA, complementary DNA; dbSNP, Single Nucleotide Polymorphism database; DN, de novo; E, exon; ExAC, Exome Aggregation Consortium; F, Father; HGVS, Human Genome Variation Society; IFD, in-frame deletion; M, Mother; MAF, minor allele frequency; MS, missense; PSE, putative splice effect.
that the 2 affected siblings of patient V-1 (V-2 and V-3), the unaffected mother (IV-2), and the affected maternal uncle (IV-3) also carry the same heterozygous \textit{RET} variant, whereas the grandmother (III-2) does not (Figure 1, Supplementary Table 4). Due to DNA unavailability, we were unable to confirm the presence of this variant in the grandfather (III-1). However, considering that both the mother (IV-2) and the grandfather (III-1) were reported to suffer from severe constipation in childhood, and the grandmother (III-2) had no intestinal complains, it is likely that this \textit{RET} variant was inherited from the grandfather (III-1).

For patient V-4, 2 rare variants were identified in 2 different genes: Indian hedgehog (\textit{IHH}) (NM_002181.3:c.151C>A; p.Q51K), and the GLI family zinc finger 3 (\textit{GLI3}) (NM_000168.5:c.2119C>T; p.P707S) (Table 1 and Supplementary Table 3). Segregation analysis showed that both variants were inherited from the father (Figure 1, Supplementary Table 4). The de novo analysis performed for patient V-4 also identified a heterozygous in-frame deletion in the Glial cell–derived neurotrophic factor gene (\textit{GDNF}) (NM_001190468.1:c.676_681delGGA TGT) (Figure 1, Table 1). No allelic frequencies of any of these variants were found in the available databases.

\textbf{RET-P399L Disturbs Protein Glycosylation and Affects Phosphorylation on GDNF Activation In Vitro}

To determine the effect of the \textit{RET} rare variant identified in the first branch of the family (c.1196C>T, p.P399L), we examined the glycosylation and phosphorylation status of the mutant protein and compared it with the WT. As expected, 2 bands were identified in the presence of the RET-WT–expressing vector (Figure 4A). The lower band (\sim 150 kDa) corresponds to the unglycosylated RET protein, whereas the upper one (\sim 170 kDa) is the glycosylated (mature) RET protein. In the presence of the RET-Mut (RET-P399L) expressing vector, only the lower band was detected, suggesting that this variant disturbs protein glycosylation (Figure 4A). RET phosphorylation was also investigated on GDNF stimulation, and in the presence of the Mut-expressing vector, RET phosphorylation was dramatically reduced (Figure 4A). These results confirm pathogenicity of the \textit{RET} variant identified.

\textbf{IHH-Q51K Disturbs Activation of Hedgehog Signaling In Vitro}

To study the effect of the \textit{IHH} variant identified in patient V-4 (c.151C>A, p.Q51K), we transiently transfected HEK293 cells with \textit{IHH-WT-FLAG} and \textit{IHH-Q51K-FLAG} vectors. Comparative expression levels of the precursor form of IHH-WT (\sim 46 kDa) and IHH-Q51K were found in the cell lysates and in the conditioned medium from transfected HEK293 cells (Figure 4B). However, a significant lower expression of the transcriptional target of Hh signaling, \textit{GLI1}, was identified by qreal time-PCR in fibroblasts cultured in the presence of conditioned medium containing the secreted form of mutant IHH (Figure 4C). This result confirms pathogenicity of the \textit{IHH} variant identified.

*\textit{ihh} Is Required for ENS Development in Zebrafish*

Transgenic zebrafish embryos \textit{Tg(-8.3phox2b:Kaede)} were injected with a morpholino designed to specifically target expression of \textit{ihh} to further study the involvement of this gene in ENS development. Morphant and uninjected control embryos were visualized at 120 hpf and several
differences were noticed. Morphant embryos showed a curved body, small eyes, and no swim bladder (Figure 4D). Moreover, a significant decrease in the number of enteric neurons was detected when compared with controls (Figure 4D). The number of enteric neurons in \textit{ihh} morphants was 13% ± 1.4% of that seen in controls (\( n = 23 \)), suggesting that \textit{ihh} is required for normal ENS development in zebrafish.

\textbf{De Novo Deletion in GDFN Leads to Reduced Levels of Secreted Protein and Results in Impaired RET Activation}

A heterozygous de novo in-frame deletion in \textit{GDFN} was identified in patient V-4 (NM_001190468.1:c.676–681delG-GATGT). Because this deletion affects 6 base pairs (bp) located just before the last codon of \textit{GDFN}, a change in RNA stability is expected based on the RNAfold online software (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi; Supplementary Figure 2 and Supplementary Table 5). To evaluate this effect, we performed qreal time-PCR on RNA isolated from HEK293 cells transfected with GDFN-WT-Myc-DDK and GDFN-Mut-Myc-DDK–expressing constructs. No significant effect on the mRNA levels was observed in the presence of the deletion (Figure 5A). To determine if the in-frame deletion identified impairs the function of GDFN, HEK293 cells transiently expressing RET and GFR-\( \alpha \)-1, were treated with conditioned medium containing GDFN-WT-Myc-DDK and GDFN-Mut-Myc-DDK. We observed that in the presence of the mutant protein, a decrease in RET expression and phosphorylation levels was observed when compared with the WT. This suggests that the deletion identified does affect the ability of GDFN to activate RET (Figure 5B). Moreover, we observed that the GDFN-Mut protein was
absent in the conditioned medium (~30 kDa), but it was still present inside the cells. The opposite situation was detected for GDNF-WT (Figure 5C). Based on our results, we concluded that the 6-bp deletion impairs secretion of GDNF, thus resulting in less RET activation.

Discussion

A complete understanding of the genetics of an inherited complex disease is a major challenge requiring substantial efforts. In this study, we used a combination of whole-exome sequencing and functional assays to find the underlying causes of HSCR in a multigenerational Dutch family.

Multiple Variants Contribute to HSCR

Finding multiple contributing variants seems logical for a disease with reduced penetrance, such as HSCR. Therefore, we were not surprised to find that 4 different genes appear to modulate disease expression in this family. In the first branch, a missense variant in RET was identified (Figure 1). This variant (c.1196C>T, p.P399L), was predicted to affect the extracellular domain of RET and result in RET dysfunction. Our in vitro studies confirmed this prediction, and showed that the variant identified was pathogenic, as it affected glycosylation and phosphorylation of RET (Figure 4A). A previous study of this family also reported that all 3 affected siblings (V-1, V-2, and V-3) inherited a common heterozygous RET risk haplotype from their father (IV-1).18 This haplotype is located in intron 1 of RET and has been shown to increase susceptibility for HSCR36 by affecting RET expression.33,37 Considering that the mother (IV-2) does not have HSCR despite carrying the pathogenic RET variant (c.1196C>T, p.P399L), it is logical to consider that the presence of the risk haplotype enhanced the penetrance of the RET variant, contributing to the development of the disease in patients V-1, V-2, and V-3.

In the second branch of this family, 2 missense variants located in IHH (NM_002181.3:c.151C>A) and GLI3 (NM_000168.5:c.2119C>T; p.P707S), and 1 de novo deletion in GDNF (c.676–681delGGATGT), have been found to underlie HSCR pathogenesis in patient V-4 (Figure 1). IHH and GLI3 encode members of the Hh pathway, whereas GDNF encodes a RET ligand. Hh signaling is known to be essential for the development of a variety of tissues and
organs and is required for normal ENS development in zebrafish.\textsuperscript{38} Despite a previous suspicion of the involvement of IHH in HSCR,\textsuperscript{39} the Hh signaling was only recently linked to this disease, when mutations in the GLI family of transcription factors, known as effectors of Hh signaling, were found in a series of patients with HSCR.\textsuperscript{40} Our functional studies support this involvement, as they confirmed the pathogenic nature of the IHH variant identified (Figure 4\textsuperscript{C}), and showed that the absence of ihh in zebrafish leads to an HSCR-like phenotype (Figure 4\textsuperscript{D}). The same effect has been previously observed in mice. However, only 50\% of Ihh knockout mice showed aganglionosis, suggesting that depletion of this gene is not fully penetrant, and disruption of additional genes is required for the intestinal phenotype observed.\textsuperscript{41} To date, it is still unclear how IHH affects ENS development, and further studies are required to determine if intestinal aganglionosis is due to a failure of migration of enteric neural crest cells from the vagal neural crest region into and along the gut tube, or whether IHH is required for proliferation of enteric neural crest cells once they enter the gastrointestinal tract. For GLI3, we found that the variant identified in patient V-4 and her father (c.2119C>T; p.P707S) has also been reported in patients with Greig cephalopolysyndactyly syndrome (MIM 175700), a rare disorder characterized by craniofacial abnormalities, polydactyly, and syndactyly of hands and feet.\textsuperscript{42} Previous studies have shown that this variant is pathogenic, as it leads to abnormal subcellular localization of GLI3 and reduced transcriptional activity.\textsuperscript{43} However, neither IV-4 nor V-4 have any of the features seen in patients with Greig cephalopolysyndactyly syndrome,\textsuperscript{18,44} leading us to conclude that this is a low-penetrance variant, likely requiring additional factors to modulate disease expression. Finally, a de novo variant in GDNF was also identified comprising an in-frame 6-bp deletion that led to the loss of 2 amino acids (c.676–681delGGATGT). Our results showed that this deletion has a pathogenic effect, as it impairs GDNF secretion and leads to reduced RET activation (Figure 5\textsuperscript{B} and 5\textsuperscript{C}). Mutations in GDNF have been previously reported in a few HSCR cases.\textsuperscript{20,21} However, it has been postulated that they are not sufficient to cause HSCR on their own, and require additional contributing factors.\textsuperscript{20,21,45} In this particular case, we hypothesize that the variants identified in IHH and GLI3 are these additional factors, especially because they are found in a heterozygote state in this family. Previously, we have proposed a model for disturbed ENS development, in which harmful and protective factors balance on a fulcrum representing a disease-specific genetic predisposition.\textsuperscript{14} In this model, mild variants that are harmless by themselves can lead to a disease phenotype if found together. For patient V-4, we believe that the deletion in GDNF is the one predisposing for HSCR, as it is the only variant present exclusively in patient V-4 and not in her healthy father (IV-4). However, it is the additive effect of the variants identified in IHH and GLI3 that triggers HSCR development in this patient.

Figure 5. De novo deletion in GDNF affects protein secretion and RET activation. (A) qreal time-PCR performed using RNA isolated from HEK293 cells transfected with GDNF-WT and GDNF-Mut constructs, shows that the 6-bp deletion identified has no effect on the levels of RNA present. (B) Western blot performed in HEK293 cells transiently expressing RET-WT, GFR-α1, and GFP, and grown in conditioned medium containing GDNF-WT and GDNF-Mut, showed that cells stimulated with the mutant GDNF have reduced levels of phosphorylated RET when compared with cells stimulated with the WT protein. Conditioned medium collected from untransfected HEK293 cells was used as negative control. Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was used as loading control and GFP as transfecting control. (B) Western blot analysis of HEK293 cells transiently expressing GDNF-WT-Myc and GDNF-Mut-Myc showed impaired secretion of the mutant protein (~30 kDa). –, absence of GDNF.
LRBA and MAB21L2

Based on the previously performed linkage analysis, we were expecting to find the causative gene for HSCR in this family on chromosome 4. Therefore, we initially focused our efforts on LRBA, as this was the only gene in the linkage region that showed expression in mouse gut. Our functional studies failed to confirm a direct involvement of LRBA in HSCR pathogenesis, and could not support a direct role for lrba in ENS development in zebrafish (Figures 2 and 3). Within LRBA, another gene can be found, MAB21L2, specifically located in intron 42 of LRBA. MAB21L2 is known to play a role in neural development, and here we show that this gene is required for ENS development in zebrafish (Figure 3C). Based on this evidence, MAB21L2 was considered to be a possible candidate gene for HSCR in this family, but because we could not identify any pathogenic variant in this gene in any of the affected members, and failed to show an effect of the LRBA variant identified on MAB21L2 expression, we were unable to link MAB21L2 to HSCR. Therefore, although we believe that MAB21L2 could play a role in HSCR pathogenesis based on its function, the risk allele on chromosome 4 for this family cannot be attributed to MAB21L2 or LRBA, and remains to be identified.

Consequences for Genetic Counseling

Complex inheritance in families with variable expression and incomplete penetrance is to be expected in HSCR. However, searching for multiple variants that in concert can explain disease variation and penetrance within such families is rare. Common practice in diagnostic laboratories is to search for mutations in the major known disease-associated gene. For HSCR, this means screening the RET gene. If a mutation is identified, the search for additional causing genes stops. However, in some families, this may not represent the full genetic etiology of the disease, leading to a miscalculation of the real genetic risk. Using the family described in this study as an example, the extensive genetic analysis was performed only because the RET variant (c.1196C>T, p.P399L) was missed in the initial screening. One could argue that for the branch in which this variant was found, the additional screen hardly

Figure 6. Schematic representation of the known and the newly identified HSCR genes.
adds any useful information, as RET probably determines most of the penetrance. However, in the RET negative branch, the additional genetic screening proved to be necessary. Finding a de novo GDNF deletion, in combination with 2 inherited variants in members of the Hh pathway (IHH and GLI3), changed genetic counseling, as we now predict a low recurrence risk for this branch of the family. Based on our findings, we believe that an extensive genetic screen can change genetic counseling of a complex genetic disease, especially if a de novo search is added. However, one should be cautious to counsel based only on the presence of a de novo variant, because it is difficult to assess the contribution of such variants to the overall disease risk.

Conclusions

HSCR is a complex disorder in which several genes are known to play a role (Figure 6). Although in 20% of the cases the genetic cause relies on the presence of a single deleterious mutation in a specific gene, for most patients it is likely that rare mutations affecting more than 1 gene are involved in disease pathogenesis. In this study, we report such a family, in which mutations in members of the major disease-associated pathway, RET and GDNF, in combination with mutations in GLI3 and in a previously unrelated HSCR gene, IHH, are likely to modulate the clinical expression of the disease phenotype (Figure 6). In addition, our results show that even familial cases can have a high genetic complexity, something that should be taken into account when counseling and performing genetic tests for disorders with a presumed multifactorial etiology.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of Gastroenterology at www.gastrojournal.org, and at https://doi.org/10.1053/j.gastro.2018.03.034.

References


Author names in bold designate shared co-first authorship.

Received October 12, 2017. Accepted March 19, 2018.

Reprint requests
Address requests for reprints to: Robert M. W. Hofstra, PhD, Department of Clinical Genetics, Erasmus University Medical Center, PO Box 2040, 3000CA Rotterdam, The Netherlands. e-mail: r.hofstra@erasusmc.nl.

Acknowledgments
The authors thank all members from the family described in this study. They also thank Dr Gang Ma from Shanghai Jiaotong University for kindly providing the pCMV-IHH-FLAG-WT vector, and Tom de Vries-Lentsch for preparing Figure 6.

The current affiliation of Lyudmila Petrova is RNA Biology Laboratory, National Cancer Institute, National Institutes of Health, Frederick, Maryland. The current affiliation of Grzegorz Burzynski is Department of Biotechnology and Molecular Biology, University of Opole, Opole, Poland.

Author contributions: Yunia Sribudiani, Rajendra K. Chauhan, Marij A. Alves, and Robert M.W. Hofstra designed and planned the experiments; Yunia Sribudiani, Rajendra K. Chauhan, Marij A. Alves, Lyudmila Petrova, Erwin Brosens, Colin Harrison, Tara Wabbersen, Bianca M. de Graaf, Tim Rügenbrink, Grzegorz Burzynski, and Jan Osinga prepared and executed the experiments; Erwin Brosens, Rutger W.W. Brouwer, and Wilfred F.J. van IJcken prepared and analyzed the sequencing data; Saskia M. Maas and Alice S. Brooks provided patient samples and important clinical information; Annelies de Klein, Bart J.L. Eggens, Alan J. Burns, and Robert M.W. Hofstra provided supervision and guidance; Yunia Sribudiani, Rajendra K. Chauhan, Marij A. Alves, Erwin Brosens, and Robert M.W. Hofstra interpreted the data and wrote the manuscript.

Conflicts of interest
The authors disclose no conflicts.

Funding
This study was supported by research grants from ZonMW (TOP-subsidie 40-00812-98-10042) and the Maag Lever Darm stichting to Robert M.W. Hofstra (WO09-82).
Supplementary Material and Methods

Exome Sequencing and Variant Prioritization

Three micrograms of DNA from each of the individuals was sheared using acoustic technology (Covaris, Inc, Woburn, MA). Target enrichment for V-1 and V-4 was performed with the SureSelect Human All Exon 50 Mb Targeted exome enrichment kit v4, and for the trio (IV-4, IV-5, and V-4) the Agilent Sureselect CRE capture kit (Agilent Technologies, Inc., Santa Clara, CA) was used. Captured fragments were sequenced (paired-end 101-bp read length) on the Illumina HiSeq2000 (Sureselect V4) and HiSeq2500 (CRE) sequencers (Illumina, San Diego, CA). De-multiplexing, alignment to the human genome build 19 (Hg19) reference genome, and curation of low-quality reads were done as described by our in-house developed NARWHAL pipeline.\(^1\) BAM-files were generated with SAMtools version 0.1.12a,\(^2\) and variant calling was performed with the Bayesian genotyper incorporated in the genome analysis toolkit version 1.2.9.\(^3\) Variant files generated of VCFv4 format were uploaded into Cartagenia Bench NGS version 5.0 (Cartagenia Inc, Boston, MA) for filtering with previously described settings.\(^4\)

Vector Design and Site Direct Mutagenesis

The genomic region of LRBA containing exon 20 and its flanking sequence (approximately 400 bp), was amplified from control and patient DNA to obtain WT and Mut (NM_001199282.2:c.2444A>G) alleles, respectively, using primers described in Supplementary Table 1 (LRBAF and LRBAR). PCR products obtained, LRBA-Enh-WT and LRBA-Enh-Mut, were inserted into the pCR 2.1-TOPO vector, subsequently digested with XhoI and KpnI restriction enzymes, and cloned into a pGL3-SV40 promoter (SV40-P) upstream of the luciferase gene (Promega, Madison, WI), to generate the pGL3-SV40p-Luc-LRBA-Enh-WT and pGL3-SV40p-Luc-LRBA-Enh-Mut vectors. The same LRBA PCR products, LRBA-Enh-WT and LRBA-Enh-Mut, were also directly cloned into the exon trapping vector pSPL3 (Invitrogen, Carlsbad, CA) to generate the pSPL3-LRBA-WT and the pSPL3-LRBA-Mut vectors. The same LRBA PCR products, LRBA-Enh-WT and LRBA-Enh-Mut, were also directly cloned into the exon trapping vector pSPL3 (Invitrogen, Carlsbad, CA) to generate the pSPL3-LRBA-WT and the pSPL3-LRBA-Mut vectors. The pRc/CMV-RET-WT vector\(^5\) and pCMV6-Entry-GDNF-Myc-DDK vector (OriGene, Rockville, MD) were used to create pCMV-IHH-FLAG-Mut (Q51K) and pCMV6-Entry-GDNF-Mut (Gly226-Cys227del)-Myc-DDK, respectively, by site-directed mutagenesis, according to the manufacturer’s instructions (Stratagene and New England Biolabs, Ipswich, MA). All inserts were Sanger-sequenced to confirm the presence of the WT and Mut variants, as well as the orientation of the inserted fragments. Primers used (RET-MutF; RET-MutR; IHH-MutF, IHH-MutR, GDNF-MutF, and GDNF-MutR) are described in Supplementary Table 1.

Whole-Mount In Situ Hybridization for lrba, mab21l2, and ihh in Zebrafish

lrba, mab21l2, and ihh genes were amplified from total mRNA collected from zebrafish embryos by RT-PCR. Amplified bands were gel-purified and sub-cloned into TOPO TA PCRII vector (Thermo Fisher, Waltham, MA). Digoxigenin-labeled antisense probes (Roche, Basel, Switzerland) were generated using SP6 polymerase (Roche) after linearizing the plasmid templates using NotI restriction enzyme (New England Biolabs). Embryos were collected and processed for whole-mount in situ hybridization as previously described.\(^7\) Digoxigenin-labeled probes were visualized with NBT/BCIP coloration reactions.

RNA Isolation and qReal time-PCR

RNA isolation was performed with the RNeasy mini kit (Qiagen, Valencia, CA) according to the manufacturer’s instructions. cDNA preparation was done with the iScript cDNA synthesis kit (Bio-Rad, Hercules, CA), using 1 μg RNA isolated from fibroblasts treated with Purmorphamine and conditioned medium containing IHH-WT and IHH-Mut. GLI1 expression levels were determined by quantitative real-time (qreal-time) Sybr Green PCR, using the 7300 Real-time PCR platform system (Applied Biosystems, Foster City, CA). The same procedure was used to determine levels of GDNF in HEK293 cells transfected with GDNF-WT and GDNF-Mut vectors. CLK2 was used as a housekeeping gene to normalize GLI1 expression levels, while GAPDH and ACTB were used for GDNF (primer details in Supplementary Table 1). qreal time-PCR data were analyzed using a method previously described,\(^9\) and presented as fold changes. These assays were performed in 3 independent triplicates (n = 9).
Supplementary Figure 1. *mab21l2* expression pattern in zebrafish. In situ hybridization showing that *mab21l2* has a strong expression in the hindbrain (black arrows) and pharyngeal arches (white arrows) through all time points (24–96 hpf). From 48 hpf onward, a strong expression is also detected in the gut mesoderm (*).

Supplementary Figure 2. A decrease in RNA stability is predicted in the presence of the de novo deletion in *GDNF*, by in silico analysis. Secondary structures of *GDNF* WT and Mut RNA determined using RNAfold software, showed that a change in both the minimum free energy (MFE) and the centroid secondary structures are predicted to occur in the presence of the deletion identified (arrowheads). Each color indicates the probability of individual nucleotides to participate in the structure, and range from the highest (red) to the lowest probability (blue-violet).
### Supplementary Table 1. List of Primers Used in This Study

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer (5′-3′)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRBAF</td>
<td>CCACATAACTTAAGGTTGATTC</td>
</tr>
<tr>
<td>LRBAR</td>
<td>GATATAAGGAGATGGGCTG</td>
</tr>
<tr>
<td>RETF</td>
<td>CTGGCCAGGCCATCTTGG</td>
</tr>
<tr>
<td>RETR</td>
<td>CGGAGTCACCATATGCAGATTACC</td>
</tr>
<tr>
<td>IHFF</td>
<td>ATCAGCCCAACGAGGACC</td>
</tr>
<tr>
<td>IHHR</td>
<td>CATCAGCCCAACGAGGACC</td>
</tr>
<tr>
<td>GLJ3F</td>
<td>AGTGCCAGCTCCTATACC</td>
</tr>
<tr>
<td>GLJ3R</td>
<td>GTTACAGCTCATTTTAGGACTTG</td>
</tr>
<tr>
<td>GDNFF</td>
<td>TTTCAACCTAATGCACTTTATTC</td>
</tr>
<tr>
<td>GDNFR</td>
<td>TGACCTGGAAGGCTAAGG</td>
</tr>
<tr>
<td>RET-MutF</td>
<td>CGTGTCGTCGCTGCTGCTGCAC</td>
</tr>
<tr>
<td>RET-MutR</td>
<td>GTGCAAGCTGACACGACGCCACG</td>
</tr>
<tr>
<td>IHH-MutF</td>
<td>CGTCGGCTACAAAGAAGTCAAGC</td>
</tr>
<tr>
<td>IHH-MutR</td>
<td>CTTGGGCGTGAACCTTGTAGCAGC</td>
</tr>
<tr>
<td>GDNF-MutF</td>
<td>ATCAGCGTGACGCCGCG</td>
</tr>
<tr>
<td>GDNF-MutR</td>
<td>ACACCTTTTAGCAGAAGTTCTTTAGAATATGG</td>
</tr>
<tr>
<td>SD6</td>
<td>TCTGAGTCACCTGGAACAACC</td>
</tr>
<tr>
<td>SA2</td>
<td>ATCTCAGTGGATTGTAAGG</td>
</tr>
<tr>
<td>lrbaF</td>
<td>CTTTTAGCAAGAAGGAGTGGGCTG</td>
</tr>
<tr>
<td>lrbaR</td>
<td>TCCAGCATGACTTCTGCTTCC</td>
</tr>
<tr>
<td>ihhF</td>
<td>GAATTTACAGGAACGAT</td>
</tr>
<tr>
<td>ihhR</td>
<td>GCTAATGCAGCAATCTCCA</td>
</tr>
<tr>
<td>mab21l2F</td>
<td>ATTCGTCCTCCGCTTCCAG</td>
</tr>
<tr>
<td>mab21l2R</td>
<td>TCCTCCAGTCGACCTCACC</td>
</tr>
<tr>
<td>GLI1qF</td>
<td>TCCCCATGACTCTGCCCG</td>
</tr>
<tr>
<td>GLI1qR</td>
<td>CCAGCATGTCAGTCAGA</td>
</tr>
<tr>
<td>GDNFqF</td>
<td>CGTACAGCTGACTCAAAT</td>
</tr>
<tr>
<td>GDNFqR</td>
<td>AGGAAGCAGTCCATTTGTT</td>
</tr>
<tr>
<td>CLK2qF</td>
<td>TCCTAGCAACCTTGGAGAGG</td>
</tr>
<tr>
<td>CLK2qR</td>
<td>TGATCTCAGGGCGACTCG</td>
</tr>
<tr>
<td>ACTBqF</td>
<td>AACCCGAGAAGATGGACC</td>
</tr>
<tr>
<td>ACTBqR</td>
<td>GCCAGAGGCGTACAGGATAG</td>
</tr>
<tr>
<td>GAPDHqF</td>
<td>CGACCTTCACCTCCCAT</td>
</tr>
<tr>
<td>GAPDHqR</td>
<td>TAAAGAGCAGCTGTGAGC</td>
</tr>
</tbody>
</table>

### Supplementary Table 2. List of Antibodies Used for Western Blot

<table>
<thead>
<tr>
<th>Antibodies</th>
<th>Host</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>RET</td>
<td>Rabbit</td>
<td>1:1000</td>
</tr>
<tr>
<td>p-RET</td>
<td>Rabbit</td>
<td>1:1000</td>
</tr>
<tr>
<td>Myc</td>
<td>Mouse</td>
<td>1:1000</td>
</tr>
<tr>
<td>β-Actin</td>
<td>Mouse</td>
<td>1:1000</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Mouse</td>
<td>1:10,000</td>
</tr>
<tr>
<td>GFP</td>
<td>Rabbit</td>
<td>1:2000</td>
</tr>
<tr>
<td>Flag</td>
<td>Mouse</td>
<td>1:1000</td>
</tr>
<tr>
<td>IRDye 800</td>
<td>Goat</td>
<td>1:10,000</td>
</tr>
<tr>
<td>IRDye 680</td>
<td>Goat</td>
<td>1:10,000</td>
</tr>
</tbody>
</table>
Supplementary Table 3. In Silico Prediction of the Pathogenic Nature of the Rare Variants Identified in Patients V-1 (a) and V-4 (b)

<table>
<thead>
<tr>
<th>Sample</th>
<th>Gene</th>
<th>Variant</th>
<th>PHAST score</th>
<th>GERP++ neutral rate</th>
<th>PhyloP score</th>
<th>SiPhy score</th>
<th>Mutation Taster</th>
<th>SIFT score</th>
<th>PolyPhen2</th>
<th>LRT prediction</th>
<th>Mutation Assessor</th>
<th>FATHMM score</th>
<th>BLOSUM62</th>
<th>Cadd Phred score</th>
</tr>
</thead>
<tbody>
<tr>
<td>a, b</td>
<td>RET</td>
<td>c.1196C&gt;T</td>
<td>-</td>
<td>5.13</td>
<td>1.151</td>
<td>10.524</td>
<td>1</td>
<td>0</td>
<td>0.856</td>
<td>Deleterious</td>
<td>1.955</td>
<td>-3.02</td>
<td>-3</td>
<td>27.5</td>
</tr>
<tr>
<td>a, b</td>
<td>NRP2</td>
<td>c.1000C&gt;T</td>
<td>0.9</td>
<td>5.91</td>
<td>1.505</td>
<td>15.056</td>
<td>1</td>
<td>0.02</td>
<td>0.929</td>
<td>Deleterious</td>
<td>2.555</td>
<td>-4.81</td>
<td>-3</td>
<td>35</td>
</tr>
<tr>
<td>a, b</td>
<td>PGRMC2</td>
<td>c.185G&gt;A</td>
<td>-</td>
<td>3.81</td>
<td>0.927</td>
<td>7.764</td>
<td>0.734</td>
<td>1</td>
<td>0.003</td>
<td>Neutral</td>
<td>-0.345</td>
<td>-1.13</td>
<td>-2</td>
<td>22.6</td>
</tr>
<tr>
<td>a, b</td>
<td>LRBA</td>
<td>c.2444A&gt;G</td>
<td>1</td>
<td>5.66</td>
<td>0.96</td>
<td>12.981</td>
<td>1</td>
<td>0.01</td>
<td>0.488</td>
<td>Deleterious</td>
<td>2.455</td>
<td>-0.12</td>
<td>1</td>
<td>25.2</td>
</tr>
<tr>
<td>a, b</td>
<td>OR1F1</td>
<td>c.477G&gt;A</td>
<td>1</td>
<td>4.97</td>
<td>2.456</td>
<td>16.064</td>
<td>0.94</td>
<td>0.01</td>
<td>0.997</td>
<td>Deleterious</td>
<td>3.54</td>
<td>5.95</td>
<td>-2</td>
<td>23.7</td>
</tr>
<tr>
<td>a, b</td>
<td>CLUH</td>
<td>c.3547G&gt;C</td>
<td>1</td>
<td>5.07</td>
<td>1.248</td>
<td>6.899</td>
<td>0.529</td>
<td>0.08</td>
<td>0.008</td>
<td>Neutral</td>
<td>0.625</td>
<td>-1.58</td>
<td>-1</td>
<td>21.7</td>
</tr>
<tr>
<td>a, b</td>
<td>PELP1</td>
<td>c.2696T&gt;C</td>
<td>1</td>
<td>4.42</td>
<td>0.564</td>
<td>0.625</td>
<td>1</td>
<td>0</td>
<td>1.5</td>
<td>Neutral</td>
<td>0.92</td>
<td>2.18</td>
<td>0</td>
<td>5.925</td>
</tr>
<tr>
<td>a, b</td>
<td>PELP1</td>
<td>c.2161A&gt;G</td>
<td>1</td>
<td>5.13</td>
<td>-0.013</td>
<td>3.9</td>
<td>1</td>
<td>21</td>
<td>0</td>
<td>Neutral</td>
<td>0.345</td>
<td>0.93</td>
<td>1</td>
<td>0.144</td>
</tr>
<tr>
<td>b</td>
<td>IHH</td>
<td>c.151C&gt;A</td>
<td>-</td>
<td>4.22</td>
<td>2.18</td>
<td>12.671</td>
<td>1</td>
<td>0</td>
<td>0.965</td>
<td>Deleterious</td>
<td>3.56</td>
<td>-6.03</td>
<td>1</td>
<td>25.2</td>
</tr>
<tr>
<td>b</td>
<td>GLI3</td>
<td>c.2119C&gt;T</td>
<td>1</td>
<td>5.82</td>
<td>1.468</td>
<td>14.65</td>
<td>1</td>
<td>0.01</td>
<td>0.925</td>
<td>Neutral</td>
<td>2.865</td>
<td>2.18</td>
<td>-1</td>
<td>28.8</td>
</tr>
<tr>
<td>b</td>
<td>GDNF</td>
<td>c.676_681delGGATGT</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>20.5</td>
</tr>
</tbody>
</table>

NOTE. The following thresholds were used to evaluate conservation: PhyloP ≥ 0.95; GERP++ ≥ 2; SiPhy ≥ 5; PHAST conservation score ≥ 0; Grantham distance ≥ 60. To predict deleteriousness the following thresholds were used: Mutation Taster ≥ 0.51; PolyPhen2 HumVar ≥ 0.909 (complex disease) or PolyPhen2 HDiv ≥ 0.956 (Mendelian disease); Mutation Assessor ≥ 1.91; FATHMM ≤ -1.56; SIFT ≤ 0.049; LRT: deleterious; BLOSUM62 < 0; CADD Phred ≥ 20 (mutations in the splice interval can have lower values). BLOSUM, Blocks Substitution Matrix; GERP, Genomic Evolutionary Rate Profiling; LRT, likelihood ratio test; PolyPhen, Polymorphism Phenotyping v2. -, unknown.
### Supplementary Table 4. Segregation Analysis of Candidate Variants Identified by Exome Sequencing in the Family Members

<table>
<thead>
<tr>
<th>Gene</th>
<th>II-2</th>
<th>III-2</th>
<th>IV-1</th>
<th>IV-2</th>
<th>IV-3</th>
<th>IV-4</th>
<th>IV-5</th>
<th>V-1</th>
<th>V-2</th>
<th>V-3</th>
<th>V-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRBA (c.2444A&gt;G)</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>RET (c.1196C&gt;T)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>IHH (c.151C&gt;A)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>GLI3 (c.2119C&gt;T)</td>
<td>NI</td>
<td>NI</td>
<td>–</td>
<td>–</td>
<td>NI</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>NI</td>
<td>NI</td>
<td>+</td>
</tr>
</tbody>
</table>

+, present; –, absent; NI, not investigated.

### Supplementary Table 5. Differences in Minimum Free Energies and Ensemble Diversity of Predicted Secondary Structures of GDNF WT and Mut RNA

<table>
<thead>
<tr>
<th></th>
<th>GDNF WT</th>
<th>GDNF Mut</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum free energy</td>
<td>–192.80 kcal/mol</td>
<td>–184.40 kcal/mol</td>
</tr>
<tr>
<td>Free energy of thermodynamic ensemble</td>
<td>–202.30 kcal/mol</td>
<td>–197.50 kcal/mol</td>
</tr>
<tr>
<td>Ensemble diversity</td>
<td>109.20</td>
<td>100.24</td>
</tr>
<tr>
<td>Minimum free energy (centroid secondary structure)</td>
<td>–177.30 kcal/mol</td>
<td>–171.10 kcal/mol</td>
</tr>
</tbody>
</table>
Supplementary References