Published in final edited form as:

Neurogastroenterol Motil. 2019 April ; 31(4): e13550. doi:10.1111/nmo.13550.

Familial Chronic Megacolon Presenting in Childhood or Adulthood: Seeking the Presumed Gene Association

DR Michael Camilleri, MD,

Clinical Enteric Neuroscience Translational and Epidemiological Research (C.E.N.T.E.R.), Mayo Clinic, Rochester, Minnesota

Eric Wieben, PhD[^], Deborah Eckert, BSN,

Clinical Enteric Neuroscience Translational and Epidemiological Research (C.E.N.T.E.R.), Mayo Clinic, Rochester, Minnesota

Paula Carlson, BS,

Clinical Enteric Neuroscience Translational and Epidemiological Research (C.E.N.T.E.R.), Mayo Clinic, Rochester, Minnesota

Ralph Hurley O'Dwyer, MD,

Clinical Enteric Neuroscience Translational and Epidemiological Research (C.E.N.T.E.R.), Mayo Clinic, Rochester, Minnesota

Denys Gibbons, MD,

Clinical Enteric Neuroscience Translational and Epidemiological Research (C.E.N.T.E.R.), Mayo Clinic, Rochester, Minnesota

Andres Acosta, MD, PhD, and

Clinical Enteric Neuroscience Translational and Epidemiological Research (C.E.N.T.E.R.), Mayo Clinic, Rochester, Minnesota

Eric W. Klee, PhD[^]

[^]Center for Individualized Medicine, Mayo Clinic, Rochester, Minnesota

Abstract

Objective: We identified a pedigree over 5 generations with 49 members, some of whom had chronic megacolon presenting in adolescence or adulthood. We aimed to assess the genetic cause of chronic megacolon through clinical and DNA studies.

Address for correspondence: Michael Camilleri, M.D., Mayo Clinic, 200 First Street S.W., Charlton Bldg., Rm. 8-110, Rochester, Minnesota 55905, Phone: 507-266-2305, camilleri.michael@mayo.edu.

Authors' contributions:

Michael Camilleri: Clinical management of patients, review of clinical and radiological information and authorship

Andres Acosta: patient care and review of medical and radiological information with first author

Ralph Hurley O'Dwyer and Denys Gibbons: Collation, analysis and review of clinical information of affected members under supervision of senior author

Deborah Eckert: Consent of participants, questionnaire data collection

Paula Carlson: preparation of DNA samples, and DNA sequencing studies

Eric Wieben: supervision and analysis of DNA and Sanger sequencing studies

Eric Klee: supervision and analysis of exome DNA sequencing

Competing interests: None

Design: After ethical approval and informed consent, family members provided answers to standard bowel disease questionnaires, radiological or surgical records, and DNA (buccal mucosal scraping). Exome DNA sequencing of colon tissue or blood DNA from 7 family members with colon or duodenal dilatation, or no megacolon (n=1) was carried out. Sanger sequencing was performed in 22 additional family members to further evaluate candidate variants. The study focused on genes of potential relevance to enteric nerve (ENS) maturation and Hirschsprung's disease or megacolon, based on the literature [*GFRA1*, *NKX2-1*, *KIF26A*, *TPM3*, *ACTG2*, *SCN10A*, and *C17orf107 (CHRNE)*] and other genetic variants that co-segregated with megacolon in the 6 affected family members.

Results: Information was available in all except 5 members alive at time of study; among 30 members who provided DNA, 6 had definite megacolon, 1 megaduodenum, 7 significant constipation without bowel dilatation, and 16 normal bowel function by questionnaire. Among genes studied, *SEMA3F* (g.3:50225360A>G; c1873A>G) was found in 6/6 family members with megacolon. The *SEMA3F* gene variant was assessed as potentially pathogenic, based on M-CAP in silico prediction. *SEMA3F* function is associated with genes (*KIT* and *PDGFRB*) that impact intestinal pacemaker function.

Conclusion: Familial chronic megacolon appears to be associated with *SEMA3F*, which is associated with genes impacting enteric nerve or pacemaker function.

Abbreviated abstract:

We identified a pedigree over 5 generations with 49 members, some of whom had chronic megacolon presenting in adolescence or adulthood. We aimed to assess the genetic cause of chronic megacolon through clinical and DNA studies. Among genes studied, *SEMA3F* (g.3:50225360A>G; c1873A>G) was found in 6/6 family members with megacolon. *SEMA3F* function is associated with genes (*KIT* and *PDGFRB*) that impact intestinal pacemaker function. Familial chronic megacolon appears to be associated with *SEMA3F*, which is associated with genes impacting enteric nerve or pacemaker function.

Keywords

SEMA3F; congenital; pseudo-obstruction

INTRODUCTION

Chronic megacolon is characterized by increase in colon diameter or increase in colonic compliance,¹ and it presents most frequently in the neonatal period in association with well-defined mutations or syndromes. The development of the enteric nervous system (ENS) is relevant to congenital Hirschsprung's disease and neonatal intestinal pseudo-obstruction. Mutations in a number of genes including those encoding the RET (rearranged during transfection) receptor tyrosine kinase and endothelin receptor type B are involved in pathogenesis of Hirschsprung's disease.² RET deletion also leads to severe defects in ENS and kidney development; activating mutations are associated with multiple endocrine neoplasia (MEN2b), which is associated with megacolon and achalasia.³ However, the vast majority of patients observed in adult gastroenterology practice have sporadic megacolon unassociated with syndromes such as MEN2b.³ In children presenting with Hirschsprung's

disease, several chromosomal and related Mendelian syndromes may be associated including Down's syndrome, Waardenburg syndrome and other dominant sensorineural deafness, the Congenital Central Hypoventilation and Mowat-Wilson and other brain-related syndromes, and a number of other autosomal recessive syndromes including Shah-Waardenburg, Bardet-Biedl and Cartilage-hair hypoplasia, Goldberg-Shprintzen, as well as other syndromes related to cholesterol and fat metabolism.⁴ Such syndromes associated with megacolon in neonates or infants were not observed in a previous series of 24 patients with chronic megacolon reported from Mayo Clinic.¹

Molecules that act later in the development, maturation or terminal differentiation of the enteric neural crest cells and vagal axons that settle in the primordial gut may also lead to other disorders of gastrointestinal motility and sensation. Thus, disorders of motility may result from subtle defects in ENS development.⁵ In addition, the neuronal phenotypic expression is influenced by the enteric microenvironment,⁶ including extracellular signaling molecules such as glial cell-derived neurotrophic factor (GDNF) family of molecules (GDNF, neurturin, artemin and persephin). These molecules activate RET in a complex with members of the GDNF receptor- α (GFR α) protein family. Conditional ablation of GFR α 1, the high affinity receptor for GDNF, leads to colon aganglionosis due to caspase-independent neuronal death.⁷ Sequence variants in the genes for prokineticin receptors PROKR1 and PROKR2 may also result in aganglionosis.⁸

Since chronic megacolon is associated with increased colonic compliance, it is also relevant to assess genes associated with smooth muscle development which may result in hypoperistalsis. Mutations in the actin (*ACTG2*) gene are associated with pseudo-obstruction, microcolon, megacystis and may result from missense mutations at arginine residues,⁹ or mutations p.R178L and p.R178C,¹⁰ or p.(Gly269Glu) in *ACTG2*.¹¹

Table 1 provides a brief summary of the candidate genes associated in the literature with megacolon and congenital or developmental disorders of the enteric nervous system that do not result in the well-characterized phenotype of Hirschsprung's disease. Disorders of gastrointestinal motility may occur in families. For example, familial occurrence of Hirschsprung's disease is recorded in 18 twin pairs;¹² in 4331 index cases of Hirschsprung's disease, there was evidence of familial recurrence in 7.6%, although among cases with total colonic aganglionosis, 20% of cases were familial.¹³ Familial recurrence was observed usually in siblings (62%) and in multiple generations in 15% of families.¹³ Such patients almost invariably present in the neonatal period or infancy, rather than adolescence or adulthood.

We identified a pedigree over five generations with 49 members in whom there were several members with evidence of megacolon or significant constipation. The observations in this family are unique, given the observation of several affected family members. These patients presented usually after adolescence or well into adulthood. Our aim was to assess the genetic cause of these colonic motility disorders through study of the clinical manifestations and DNA from thirty consenting family members.

METHODS

Design

Two senior members of the family presented with chronic megacolon for clinical care and both volunteered their DNA to be submitted to the Mayo Clinic Bio-Bank and stored for future research at Mayo Clinic. A member of the family had volunteered to coordinate acquisition of postal addresses of family members willing to participate, as well as documentation of the entire family tree so that genetic results could be examined with the documented phenotype in each symptomatic and non-symptomatic member of the family.

After Mayo Clinic Institutional Review Board approval (IRB #15-004045), we obtained information on the phenotype through standard gastrointestinal symptom questionnaire¹⁴ and on DNA by buccal scrapings (Oragene-Dx, OGD-500/510, DNA Genotek Inc., Ottawa, ON, Canada K2K 1) from other consenting members of the family. Children under 18 years of age signed assent forms in addition to their parents' consent. In the two patients (III.SG, III.JR) for whom colonic tissue was available, no further DNA samples were collected. Patients were also invited to provide any medical records such as prior radiographs, barium enemas and abdominal CT scans.

Strategy for Identifying Genetic Variants

Variant calls were filtered to keep only calls with a quality score ≥ 20 , outside the 5% most exonically variable 100-base pair window regions in the 1000 genomes project, encoding for a nonsense, missense, frame-shifting, or splice-site alteration, and occurring at a minor allele frequency (MAF) of $\geq 0.5\%$ in 1000 genomes, NHLBI, and ExAC databases. A general, and not population specific, MAF threshold was used following standard rare disease testing dogma, as megacolon disease is not known to be enriched in any sub-population and, thus, it is expected to be rare in all populations. For example, conditions such as Hirschsprung's megacolon or Mowat-Wilson syndrome are identified (albeit in about 1 in 5000) in many ethnic and racial populations, based on the published literature.

Genetic Sequencing Strategy

The overall design of DNA sequence testing was as follows: Initially, we performed exome DNA sequencing on the two senior members of the family who presented clinically; exome sequencing identified candidate megacolon mutations in a fairly large number of genes in these first two individuals. From this, we narrowed down the candidate list based on expert knowledge (MC) focused on the association of candidate mechanisms or genes with migration, differentiation, maturation or survival of neuroblasts, as well as function or degeneration of the enteric nervous system and smooth muscle. This included *GDNF* (2), *GFRA1*, *NFATC4*, *NKX2-1* and *KIF26A* (2 variants), and 234 other genetic variants "one hop" from these genes based on bioinformatics analysis (Supplemental Table 1). Next, we used Sanger sequencing to screen for mutations identified in the exome data from the two family members in a large number of family members and for genes selected from the literature for their potential causative role in the observed familial megacolon [*GFRA1*, *NKX2-1*, *KIF26A* (2 variants), *TPM3*, *ACTG2*, *SCN10A*, and *C17orf107* (see Table 1)].

Lacking a definitive finding, based on the previously documented gene associations, we subsequently pursued exome DNA sequencing in all the affected members with definite megacolon, seeking to identify a new candidate gene/variant to explain the familial condition. These included the original two patients (SG and JR in the family pedigree, Figure 1) who presented for clinical care, 4 other members with megacolon (total n=6, that is, the sibling pair BC and EC as well as AR, DH in addition to JR and SG), and 1 senior (by age) member (EHB) with no symptoms and no bowel dilatation and, therefore, not likely to harbor the disease variant.

Variants in genes which completely co-segregated with the phenotype in the exome sequenced samples were selected for further analysis. Potential pathogenicity of each variant was predicted using CADD and M-CAP scores (detailed below). Each gene was also investigated in the literature to search for a known association with associated GI or neurological disease, and/or a relationship to enteric nervous system (ENS) development, differentiation or survival. Finally, for variants co-segregating with megacolon, as well as having documented evidence in the literature of effects on ENS development or function, and M-CAP or CADD scores suggestive of pathogenicity, we evaluated protein family members, as defined by UniProt, for known association with Hirschsprung's disease alone, megacolon, or ENS development. Resulting variant(s) from this search were Sanger sequenced in the additional family members.

Exome and In Depth DNA Sequencing Method and Bioinformatics Analysis

Patient DNA was captured by using Agilent SureSelectXT All Exome V4 and was sequenced by multiplexing four samples per lane on an Illumina HiSeq 2000. The resulting data were analyzed using an internally developed next-generation sequencing workflow. Briefly, FASTQ files were aligned to the hg19 reference genome by using Novoalign (VN:V2.07.13; www.novocraft.com) with the following options: -hdrhd off -v 120 -c4 -i PE 425,80 -x 5 -r Random. Realignment and recalibration were performed by use of Genome Analysis Toolkit [GATK] (VN:1.6 -7-g2be5704) Best Practices version 3 (<https://software.broadinstitute.org/gatk/>). Germline variations were called with GATK's Unified-Genotyper using default parameters. Variant quality score recalibration was also done with the following command line optimizations for SNVs, -an QD -an HaplotypeScore -an MQRankSum -an Read-PosRankSum -an FS -an MQ -an DP -nt 2 -an QD -an FS -an HaplotypeScore -an ReadPosRankSum-maxGaussians 4 -nt 2 -percent BadVariants 0.12 -std 10.0. VQSR "PASS" filter variants were selected for further analysis. VCF files were subsequently annotated by using Mayo Clinic's BioR annotation repository, which includes variant population frequencies, frequency of occurrence in 50 internal-to-Mayo control samples, function predictions of variant effects with SNP-Eff, an open source tool that annotates variants and predicts their effects on genes using an interval forest approach (http://snpeff.sourceforge.net/VCFannotationformat_v1.0.pdf), and SIFT (Sorting Intolerant from Tolerant, which predicts whether an amino acid substitution affects protein function) (<http://sift.jcvi.org/>).

Variant filtration and selection of candidates was carried out using Ingenuity Variant Analysis Software, as detailed in the Supplementary Material.

We considered that genes fulfilling the criteria for potential relevance may be interacting partners of genes that are associated with control in the development or differentiation of the enteric nervous system, based on the bioinformatics algorithm PCAN.¹⁵ Therefore, the protein-protein interaction characterization of SEMA3F was carried out using PCAN.¹⁵ The software was run with default parameters and refreshed data sources with the following download dates: Clinvar (01 July 2017), NCBI Gene Info (July 2017), HPO (July 2017), HPO phenotype annotation (April 2017), StringDB Protein Links (v10.5), and NCBI Gene2Ensembl (July 2017).

Combined Annotation Dependent Depletion (CADD)

We used CADD version 1.3 to assess the potential causative role of variants (as CADD scores) in the exomes within individuals or the family. CADD scores generally correlate with allelic diversity, pathogenicity of both coding and non-coding variants, experimentally measured regulatory effects, and highly rank causal variants within individual genome sequences.¹⁶

The threshold for pathogenicity was set at a score of >20. For all the variants, we also added the gnomAD minor allele frequency.

Mendelian Clinically Applicable Pathogenicity (M-CAP) Score

We also used M-CAP score to assess pathogenicity for rare missense mutations in the human genome.¹⁷ The threshold for pathogenicity was set at a score of >0.025.

RESULTS

Pedigree and Phenotype

We identified a pedigree over five generations with 49 members in whom there were several members with evidence of megacolon presenting usually after adolescence or well into adulthood (Figure 1). Among 30 members who provided DNA, 7 had definite megacolon (n=6) or megaduodenum (n=1); 7 others had significant gastrointestinal symptoms (predominantly constipation or alternating bowel function), 15 had no symptoms by questionnaire suggesting normal colonic function, and 1 had constipation in association with Down's syndrome. By family reports of deceased members or clinical or radiological evidence, there were 8 with megacolon (6 were males, and 3 were deceased). No information was available in five members of the family. None of the family members reported associated conditions that would suggest the presence of chromosomal and related Mendelian syndromes previously described in the literature in association with Hirschsprung's disease.⁴

Figure 2 shows abdominal imaging of six members of the family with evidence of intestinal or colonic dilatation.

Initial Exome and Sanger Sequencing Analysis

Initial molecular analysis by exome DNA sequencing of the two senior members of the family (III.SG, III.JR) with diagnosed megacolon identified 234 candidate variants which

co-segregated in these two individuals (Supplemental Table 1). A literature review of the genes in which these variants were identified resulted in the selection of 7 candidate genes encompassing 8 variants (Table 1) in which Sanger sequencing in all 30 family members was carried out (Table 2). These 8 variants included 4 non-synonymous coding changes (in *KIF26A*, *ACTG2*, *SCN10A*, and *CHRNE* [*C17orf107*]), 1 synonymous coding change in *KIF26A*, and 3 non-coding sequence changes [including a short tandem repeat (STR) in *NKX2-1*, and SNVs in the untranslated regions of *GFRA1* and *TPM3*]. Zero of the 8 candidate variants segregated uniquely with the megacolon phenotype in this family. Each variant was found, not only in a subset of affected family members, but also in multiple unaffected individuals (Table 2). This result suggested that it was unlikely the monogenic cause of this familial megacolon had been identified and, therefore, subsequent studies were pursued.

Expanded Cohort Exome Sequencing Analysis

Exome DNA sequencing was performed on an additional 5 family members, including 4 with megacolon from three generations (III.DH, IV.AR, V.BC, V.EC), and 1 unaffected family member (II.EHB). We were confident this individual was not likely to develop megacolon as she was asymptomatic at an advanced age. These additional exome sequencing data were combined with the initial exome sequencing (III.SG, III.JR) for subsequent candidate variant selection.

Following filtration for quality, impact, and co-segregation with phenotype, 12 candidate variants in 11 genes (*ATG16L1*, *ATN1*, *NISCH*, *NLRP9*, *PODN*, *PRSS3*, *MUC16*, *RFT1*, *SEMA3F*, *TPTE2* and *ZNF530*) were identified, as shown in Table 3.

A literature review of known function and disease association was carried out for these 11 genes (Table 4) to identify potential associations with megacolon or disorders of the enteric nervous system. Five genes (*MUC16*, *NLRP9*, *PODN*, *PRSS3*, *TPTE2*) encompassing 6 variants were found to have no known association with any associated GI or neurological disease. For each variant in these genes, CADD and M-CAP scores (Table 3) suggested the impact was likely benign, and these findings were not pursued further.

The 6 remaining genes (*ATG16L1*, *ATN1*, *NISCH*, *RFT1*, *SEMA3F*, *ZNF530*) have known associations to GI or neurological disease, but no direct association with megacolon or disorders of the enteric nervous system. Of these genes, *NISCH* and *RFT1*, had high M-CAP and CADD scores (>30), and *SEMA3F* had a high M-CAP score, suggesting these variants may have a pathogenic impact. These genes were further investigated for additional association with megacolon or ENS function. Of the 3, only *SEMA3F* belonged to a protein family where members *SEMA3A*, *SEMA3C* and *SEMA3D* have been implicated in Hirschsprung's disease. This association with disease affecting the enteric nervous system in the semaphorin protein family made this variant an interesting candidate to consider as causative for this familial chronic megacolon, and further Sanger sequencing was carried out.

Results of the Sanger sequencing for the entire family cohort for *SEMA3F* are shown in Supplemental Table 2. The *SEMA3F* variant results from exome sequencing were confirmed

by Sanger sequencing. The *SEMA3F* variant was also found in 2/7 family members with GI symptoms but no bowel dilation, and in 2 family members with no GI symptoms or bowel dilation (Supplemental Table 2). Gene variants were identified by Sanger sequencing in a small minority of currently unaffected members of the family. Since the phenotype often presents in adulthood in this family, it is conceivable that carriers of the gene variants may still develop megacolon in the future. Therefore, we considered that the gene variants may be present in family members who had not yet manifested the phenotype.

Genes Interacting with *SEMA3F* Gene

To better understand how *SEMA3F* may be involved in familial megacolon, we bioinformatically analyzed protein-protein interaction data to identify other genes encoding proteins associated with the functions of this gene. We conducted this assessment in order to identify potential co-factors that could be implicated in the derangement of the enteric nervous system. Figure 3 (upper panel) shows the results of this analysis for *SEMA3F*, and Table 5 (genes of relevance to ENS) and Supplemental Table 3 (genes that, at present, appear unrelated to ENS development, structure, function or disease) show a list of candidate genes functioning in association with *SEMA3F* gene. Based on the relevance of these interactions with likely proteins involved in enteric neural development or functions, we concluded the best candidate gene to be in association with this familial chronic megacolon is *SEMA3F*, based on this study.

DISCUSSION

This study documents the familial occurrence of the phenotype of chronically dilated segments of the gastrointestinal tract, most frequently megacolon, in at least six members and probably two deceased members of one family. Through the acquisition of phenotype information (symptoms and radiographs), as well as (mostly) buccal DNA samples from 30 affected or unaffected members, we embarked on an investigation of possible genetic variants predisposing to the disease. The occurrence of chronic megacolon in six members in one family is unusual in the absence of MEN2B syndrome which causes megacolon or achalasia.³ Similarly, the observation of so many affected members in one family contrasts with the scarcity of chronic megacolon presenting beyond infancy, which was diagnosed in 24 patients at Mayo Clinic over a 15-year period.¹ This is in contrast to Hirschsprung's disease which results in megacolon in the neonatal period or infancy and affects about 1 in 5000 live births.²

Our search for the causative gene has been inconclusive to date, despite several investigative strategies including targeted gene studies based on published variants in Hirschsprung's disease and congenital aganglionosis, Sanger sequencing and exome DNA sequencing. The most informative analysis from exome DNA sequencing revealed a number of potential candidates (shown in Tables 4 and 5) that co-segregated with disease phenotype. However, among these, the gene we considered most likely to be associated with ENS disorder is *SEMA3F*. *SEMA3F* has not been previously associated with definite megacolon, although SEMA signaling pathway has been associated with non-syndromic Hirschsprung's disease, specifically SEMA3A, SEMA3C and SEMA3D. There are several genes and pathways that

interact with *SEMA3F* and are of significant relevance to the colonization, migration, differentiation, and survival of neural crest cells that lead to development of the ENS. We used protein-protein interaction data to better characterize possibly relevant interactions; these are summarized in Table 5 and include the genes for receptor tyrosine kinase, platelet-derived growth factor receptor B (*PDGFRB*), nerve growth factor (*NGF*) and tyrosine hydroxylase (*TH*). In addition to development of ENS, these genes are critically involved in pacemaker function and adrenergic control of gut motility. At this point, these observations remain purely hypothesis-generating and require further investigation.

We observed variations in several members of this family as shown by diverse locations of the megacolon or dilated small intestine, the long history, and variable age at presentation. Thus, it appears that the site of poor ENS development is not stereotypical as usually observed in Hirschsprung's disease, in which the denervation affects the distal left colon. On the other hand, sporadic cases of chronic megacolon in adults may affect isolated segments of the colon, as has been described in hind gut dysgenesis^{18,19} or in proximal megacolon.²⁰ Thus, we envisage either failure of caudal migration of cranial neural crest cells, failure of rostral migration of sacral neural crest cells, or failure of the conditions locally at the site of involvement that normally facilitate colonization or maturation of the neural elements to form a normal ENS locally.

Our studies suggest the prime genetic candidate to be *SEMA3F* in this family with chronic megacolon. Two other genes which co-segregated with the megacolon phenotype in our exome data were *NISCH* and *RPT1*. Both variants in these genes had highly damaging in silico predictions, but no close association (themselves or through protein interactions explored with bioinformatics) with ENS development or maturation. Further investigation of *NISCH* and *RPT1* may be warranted in future studies; however, based on current knowledge, *SEMA3F* is the most relevant gene studied as it interacts with pathways involving genes critical to the development of intestinal pacemaker cells.

Although we sought variants in genes associated with genes that were previously associated with enteric neural dysfunction (*GFRA1*, *NKX2-1*, *KIF26A*, *TPM3*, *SCN10A*, *C17orf107*, and *ATG16L1*) or abnormal smooth muscle development or function, such as *ATCG2*, there was less convincing evidence of association of such genes with megacolon in this family. These observations lead us to propose hypothetical mechanisms that may contribute to the poor development of ENS in focal areas, resulting in dilation of the intestine or colon in several members of a single family. While several members of the family also had constipation or altered bowel function, they did not present dilated segments and only a minority (1 or 2 of 16) had genetic variants in *SEMA3F* on Sanger sequencing to suspect that their symptoms reflected incomplete penetrance of the gene.

Limitations

First, we performed exome DNA sequencing, which does not provide coverage of the entire coding region of the genome, leading to the possibility that the causative genetic variant was not assayed. Additionally, it is certainly possible that variations in regulatory regions exist and were undetected, which may be responsible for the phenotype. Indeed, it is also conceivable that such variations in regulatory regions in the genes previously reported in the

literature (and included in our study) were not evaluated for a potential role in this pedigree of familial megacolon. Finally, certain large insertions, duplication, and deletions, as well as, structural variations and other types of genetic defects are not adequately characterized by exome sequencing, leaving open the possibility that this condition was not caused by an SNV or small INDEL, but by another form of genetic defect.

Conclusion

While the gene segregating exclusively with this non-syndromic familial megacolon presenting beyond the neonatal period was not identified, we believe it is appropriate to follow up carriers of the *SEMA3F* variants in order to appraise its potential biological effects and institute steps to prevent development of symptoms (e.g., by aggressive treatment of constipation) or promptly diagnosing megacolon and offering curative colectomy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank the members of the family studied for providing medical information and biological samples. We are grateful to Drs. Sara Eleoff Van Durme and Jane Benson at Johns Hopkins Bayview Medical Center for sharing radiological studies and interpretation of the colon x-ray of one of the pediatric cases. We thank Mrs. Cindy Stanislav for excellent secretarial assistance.

Funding: Dr. Camilleri is supported by grants R01-DK115950 and R01-DK67071 from National Institutes of Health. This study was supported by a research award from the Center for Individualized Medicine at Mayo Clinic.

REFERENCES

1. O'Dwyer RH, Acosta A, Camilleri M, et al. Clinical features and colonic motor disturbances in chronic megacolon in adults. *Dig Dis Sci* 2015;60:2398–2407. [PubMed: 25868630]
2. Heuckeroth RO. Hirschsprung disease - integrating basic science and clinical medicine to improve outcomes. *Nat Rev Gastroenterol Hepatol* 2018;15:152–67. [PubMed: 29300049]
3. Gibbons D, Camilleri M, Nelson AD, et al. Characteristics of chronic megacolon among patients diagnosed with multiple endocrine neoplasia type 2B. *United Eur Gastroenterol J* 2016;4:449–54.
4. Moore SW. Chromosomal and related Mendelian syndromes associated with Hirschsprung's disease. *Pediatr Surg Int* 2012;28:1045–58. [PubMed: 23001136]
5. Gershon MD. Developmental determinants of the independence and complexity of the enteric nervous system. *Trends Neurosci* 2010;33:446–56. [PubMed: 20633936]
6. Gershon MD, Payette RF, Rothman TP. Development of the enteric nervous system. *Fed Proc* 1983;42:1620–5. [PubMed: 6131841]
7. Uesaka T, Jain S, Yonemura S, et al. Conditional ablation of GFR α 1 in postmigratory enteric neurons triggers unconventional neuronal death in the colon and causes a Hirschsprung's disease phenotype. *Development* 2007;134:2171–81. [PubMed: 17507417]
8. Ruiz-Ferrer M, Torroglosa A, Núñez-Torres R, et al. Expression of PROKR1 and PROKR2 in human enteric neural precursor cells and identification of sequence variants suggest a role in HSCR. *PLoS One* 2011;6:e23475. [PubMed: 21858136]
9. Wangler MF, Gonzaga-Jauregui C, Gambin T, et al. Heterozygous de novo and inherited mutations in the smooth muscle actin (*ACTG2*) gene underlie megacystis-microcolon-intestinal hypoperistalsis syndrome. *PLoS Genet* 2014;10:e1004258. [PubMed: 24676022]

10. Thorson W, Diaz-Horta O, Foster J, 2nd, et al. De novo ACTG2 mutations cause congenital distended bladder, microcolon, and intestinal hypoperistalsis. *Hum Genet* 2014;133:737–42. [PubMed: 24337657]
11. Klar J, Raykova D, Gustafson E, et al. Phenotypic expansion of visceral myopathy associated with ACTG2 tandem base substitution. *Eur J Hum Genet* 2015;23:1679–83. [PubMed: 25782675]
12. Henderson D, Zimmer J, Nakamura H, Puri P. Hirschsprung's disease in twins: a systematic review and meta-analysis. *Pediatr Surg Int* 2017;33:855–9. [PubMed: 28601901]
13. Mc Laughlin D, Puri P. Familial Hirschsprung's disease: a systematic review. *Pediatr Surg Int* 2015;31:695–700. [PubMed: 26179259]
14. Talley NJ, Phillips SF, Wiltgen CM, et al. Assessment of functional gastrointestinal disease: the bowel disease questionnaire. *Mayo Clin Proc* 1990;65:1456–79. [PubMed: 2232900]
15. Godard P, Page M. PCAN: phenotype consensus analysis to support disease-gene association. *BMC Bioinformatics* 2016;17:518. [PubMed: 27923364]
16. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 2014;46:310–5. [PubMed: 24487276]
17. Jagadeesh K, Wenger A, Berger M, Guturu H, Stenson P, Cooper D, Bernstein J, Bejerano G. M-CAP eliminates a majority of variants with uncertain significance in clinical exomes at high sensitivity. *Nat Genet* 2016;48:1581–6. [PubMed: 27776117]
18. Sanders KM, Kito Y, Hwang SJ, et al. Regulation of gastrointestinal smooth muscle function by interstitial cells. *Physiology* 2016;31:316–26. [PubMed: 27488743]
19. Sweetser S, Camilleri M. Clinical challenges and images in GI. Hindgut dysgenesis with megacolon. *Gastroenterology* 2008;134:1293 and 1635. [PubMed: 18471504]
20. Vijayvargiya P, Camilleri M. Proximal megacolon in an adult. *Clin Gastroenterol Hepatol* 2014;12:e83–4. [PubMed: 23811253]
21. Kapur RP, Gershon MD, Milla PJ, et al. The influence of Hox genes and three intercellular signalling pathways on enteric neuromuscular development. *Neurogastroenterol Motil* 2004;16, Suppl. 1:8–13. [PubMed: 15065997]
22. Que J, Okubo T, Goldenring JR, et al. Multiple dose-dependent roles for Sox2 in the patterning and differentiation of anterior foregut endoderm. *Development* 2007;134:2521–31. [PubMed: 17522155]
23. Zhou R, Niwa S, Homma N, et al. KIF26A is an unconventional kinesin and regulates GDNF-Ret signaling in enteric neuronal development. *Cell* 2009;139:802–13. [PubMed: 19914172]
24. Gattuso JM, Smith VV, Kamm MA. Altered contractile proteins and neural innervation in idiopathic megarectum and megacolon. *Histopathology* 1998;33:34–8. [PubMed: 9726046]
25. Hu S, Xu W, Miao X, et al. Sensitization of sodium channels by cystathionine β -synthetase activation in colon sensory neurons in adult rats with neonatal maternal deprivation. *Exp Neurol* 2013;248:275–85. [PubMed: 23834820]
26. Lorenzoni PJ, Scola RH, Kay CS, et al. Congenital myasthenic syndrome: a brief review. *Pediatr Neurol* 2012;46:141–8. [PubMed: 22353287]
27. Sanders KM, Ordög T, Ward SM. Physiology and pathophysiology of the interstitial cells of Cajal: from bench to bedside. IV. Genetic and animal models of GI motility disorders caused by loss of interstitial cells of Cajal. *Am J Physiol Gastrointest Liver Physiol* 2002;282:G747–56. [PubMed: 11960771]
28. Jabari S, de Oliveira EC, Brehmer A, et al. Chagasic megacolon: enteric neurons and related structures. *Histochem Cell Biol* 2014;142:235–44. [PubMed: 25059649]
29. Kazlauskas A PDGFs and their receptors. *Gene* 2017;614:1–7. [PubMed: 28267575]
30. Nair DG, Miller KG, Lourenssen SR, et al. Inflammatory cytokines promote growth of intestinal smooth muscle cells by induced expression of PDGF-R β . *J Cell Mol Med* 2014;18:444–54. [PubMed: 24417820]
31. Bautista-Cruz F, Nair DG, Lourenssen S, et al. Impaired platelet-derived growth factor receptor expression and function in cultured lower esophageal sphincter circular smooth muscle cells from W/W(v) mutant mice. *Can J Physiol Pharmacol* 2014;92:34–41. [PubMed: 24383871]

32. Tomuschat C, O'Donnell AM, Coyle D, et al. Reduction of hydrogen sulfide synthesis enzymes cystathionine- β -synthase and cystathionine- γ -lyase in the colon of patients with Hirschsprung's disease. *J Pediatr Surg* 2017 6 23 pii: S0022-3468(17)30370-6.
33. Kobayashi H, Yamataka A, Fujimoto T, et al. Mast cells and gut nerve development: implications for Hirschsprung's disease and intestinal neuronal dysplasia. *J Pediatr Surg* 1999;34:543-8. [PubMed: 10235318]
34. Kuroda T, Ueda M, Nakano M, et al. Altered production of nerve growth factor in aganglionic intestines. *J Pediatr Surg* 1994;29:288-92; discussion 292-3. [PubMed: 8176607]
35. Touloukian RJ, Morgenroth VH, 3rd, Roth RH. Sympathetic neurotransmitter metabolism in Hirschsprung's disease. *J Pediatr Surg* 1975;10:593-8. [PubMed: 241782]
36. Tomycz ND, Haynes RL, Schmidt EF, et al. Novel neuropathologic findings in the Haddad syndrome. *Acta Neuropathol* 2010;119:261-9. [PubMed: 19844731]
37. Stricker S, Rauschenberger V, Schambony A. ROR-family receptor tyrosine kinases. *Curr Top Dev Biol* 2017;123:105-42. [PubMed: 28236965]
38. Flanagan JG, Vanderhaeghen P. The ephrins and Eph receptors in neural development. *Annu Rev Neurosci* 1998;21:309-45. [PubMed: 9530499]
39. Memic F, Knoflach V, Morarach K, Sadler R, Laranjeira C, Hjerling-Leffler J, Sundström E, Pachnis V, Marklund U. Transcription and signaling regulators in developing neuronal subtypes of mouse and human enteric nervous system. *Gastroenterology* 2018;154:624-36. [PubMed: 29031500]

KEY POINTS

- Congenital megacolon is associated with defects in genes associated with enteric nerve or smooth muscle development (e.g., *GFRA1*, *NK82-1*, *KIF26A*, *TPM3*, *ACTG2*, *SCN10A*, and *C17orf107 (CHRNE)*). Some of these genes are also associated with Hirschsprung's disease, e.g. *GFRA1*. Familial or twin cases of Hirschsprung's disease are rare.
- Non-syndromic chronic megacolon after the neonatal period is a rare, sporadic disease.
- We identified a 5-generation family including 6 with chronic megacolon presenting beyond neonatal period. Clinical and DNA studies were conducted in 30 members to identify possible genetic variants.
- Most promising DNA variant is the G allele at rs50225360 in *SEMA3F*, which interacts with genes impacting enteric nervous system development or function (specifically, *KIT* and *PDGFRB*).
- This unique study of familial, chronic megacolon presenting post-infancy or -childhood provides insights on potential genetic mechanisms as well as an opportunity to identify susceptible family members, provide prompt diagnosis, and offer curative colectomy.

Figure 1.
DNA sequencing strategy

Figure 2.
Pedigree of the family showing 6 members in 5 generations with megacolon

Coronal images from Abdominal CT scans

Barium Enema

Plain Abdominal X-rays

Figure 3.

Examples of abdominal plain radiographs or CT images in coronal section to illustrate the large colonic diameters in 6 members of the family, including a 13 year-old with large diameter ascending colon

Figure 4.
Genes interacting with *SEMA3F*

Candidate genes identified by initial exome sequencing and associated in the literature with megacolon or variants that result in congenital or developmental disorders of the enteric nervous system, but excluding the well-characterized phenotype of Hirschsprung's disease

Table 1.

Gene Symbol	Name	Known effect	Ref. #
GFRA1	GDNF family receptor alpha 1	ENS development	21
NKX2-1	NK2 homeobox 1	Esophageal atresia	22
KIF26A	kinesin family member 26A	GDNF-Ret in ENS development	23
TPM3	tropomyosin 3	Megacolon	24
ACTG2	actin, gamma 2, smooth muscle, enteric	Visceral myopathy, megacolon;	9-11
SCN10A	sodium voltage-gated channel alpha subunit 10	Colon sensation	25
C17orf107; CHRNE	chromosome 17 open reading frame 107	Congenital myasthenia: cholinergic receptors*	26

Table 2: Sanger sequencing results of variants identified in the candidate genes selected by initial exome sequencing and literature review

Gene	GFRA1	NKX2-1	KIF26A	KIF26A	KIF26A	TPM3	ACTG2	SCN10A	C17orf107
Location	chr10: 118031734	chr14: 36986022	chr14: 104633244	chr14: 104646027	chr1: 154140007	chr2: 74143777	chr3: 38763774	chr17: 4805239	
Reference	G	C	G	G	A	G	A	G	
Predicted AA	UTR	UTR	P->P	C->Y	UTR	R->L	M->T	S->L	
EXAC Frequency	UTR	UTR	0.00144	0.00056	UTR	Novel	0.00022	0.00026	
Megacolon patients (2 senior members who presented for care of megacolon *)	G/C	15/15	G/A	G/A	A/A	G/G	A/A	G/G	
	G/G	15/15	G/A	G/A	A/A	G/T	A/G	G/G	
	G/G	14/15	G/G	G/G	A/A	G/T	A/A	G/G	
	G/G	14/15	G/G	G/G	A/A	G/T	A/A	G/G	
	G/C*	15/15	G/A	G/A	A/G	G/T	A/G	G/A	
	G/C*	15/15	G/A	G/A	A/G	G/T	A/G	G/A	
Megaduodenum	G/C	14/15	G/A	G/A	A/A	G/T	A/G	G/A	
Summary: Dilatation	4/7 G/C	4/7 15/15	5/7 G/A	4/7 G/A	5/7 A/A	6/7 G/T	4/7 A/G	4/7 G/G	
Family members with no GI symptoms	G/C	15/15	G/G	G/G	A/A	G/G	A/A	G/G	
	G/G	13/15	G/G	G/G	A/A	G/G	A/A	G/G	
	G/G	15/15	G/G	G/G	A/A	G/G	A/A	G/G	
	G/C	14/15	G/A	G/A	A/A	G/G	A/A	A/A	
	G/C	14/15	G/G	G/G	A/A	G/G	A/G	G/G	
	G/G	15/15	G/G	G/G	A/A	G/T	A/A	G/G	
	G/C	15/15	G/A	G/A	A/A	G/T	A/G	G/A	
	G/G	14/15	G/G	G/G	A/A	G/T	A/A	G/G	
	G/G	13/15	G/G	G/G	A/A	G/T	A/A	G/G	
	G/G	13/14	G/G	G/G	A/A	G/G	A/A	G/G	
	G/G	14/15	G/G	G/G	A/A	G/G	A/A	G/G	
	G/C	14/15	G/A	G/A	A/G	G/T	A/G	G/G	
	G/G	nd	G/G	G/G	A/A	G/T	A/A	G/G	
	G/C	14/15	G/G	G/G	A/A	G/G	A/G	G/G	
	G/G	14/14	G/A	G/A	A/A	G/G	A/G	G/G	

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Gene	<i>GFRAL</i>	<i>NKX2-1</i>	<i>KIF26A</i>	<i>KIF26A</i>	<i>KIF26A</i>	<i>TPM3</i>	<i>ACTG2</i>	<i>SCN10A</i>	<i>CL1orf107</i>
	G/G	13/15	G/G	G/G	G/G	A/A	G/G	A/A	G/G
Summary	6/16 G/C	6/16; 15/15	4/16 G/A	4/16 G/A	4/16 G/A	15/16 A/A	6/16 G/T	5/16 A/G	14/16 G/G
Family members with GI symptoms, without bowel dilatation	G/G	14/15	G/G	G/G	G/G	A/A	G/T	A/G	indel
	G/C	15/15	G/G	G/G	G/G	A/A	G/T	A/A	G/G
	G/G	15/15	G/G	G/G	G/G	A/A	G/T	A/A	G/G
	G/G	nd	G/G	G/G	G/G	A/A	G/T	A/A	G/G
	G/C	14/15	G/G	G/G	G/G	A/A	G/T	A/G	G/G
	G/G	nd	G/G	G/G	G/G	A/G	G/T	A/G	G/G
	G/G	13/15	G/G	G/G	G/G	A/A	G/T	A/A	G/G
Summary	2/7 G/C	2/7; 15/15	0/7 G/A	0/7 G/A	0/7 G/A	6/7 A/A	7/7 G/T	3/7 A/G	6/7 G/G

Table 3.

Variants identified by exome DNA sequencing co-segregating 100% with the phenotype megacolon (Ref. = reference; RS ID= Reference SNP cluster ID; Non-S=non-synonymous; MAF=minor allele frequency; gnomAD= genome Aggregation Database; NA= not available; pathogenicity threshold >0.025 for M-CAP and >20 for CADD)

Chr	Position	RS ID	Ref	Alt	Variation Type	Gene Symbol	Transcript ID	Transcript Variant	Protein Variant	gnomAD MAF	M-CAP score	CADD score
2	234173698	rs750820589	A	G	SNV	ATG16L1	NM_030803.6	c.550A>G	p.T68A	5.279e-5	0.012	14.28
12	7045904	rs781852614	NA	-	DEL	ATN1	NM_001940.3	c.1506_1508 delGCA	p.Q502del	0.001498	NA	11.57
3	52510604	rs142925971	G	A	SNV	NISCH	NM_007184.3	c.907G>A	p.D303N	0.0004149	0.063	31
19	8995955	NA	G	A	SNV	MUC16	NM_024690.2	c.41286C>T	p.S13762S	NA	NA	0.001
19	56249486	rs777225723	C	A	SNV	NLRP9	NM_176820.3	c.255G>T	p.W85C	0.001877	NA	13.7
1	53532545	rs149127843	T	C	SNV	PODN	NM_001199080.2	c.32+6T>C	---	0.003587	NA	4.628
3	53164370	rs148716754	G	C	SNV	RFT1	NM_052859.3	c.47C>G	p.S16C	0.0005529	0.079	32
3	50225360	rs774797321	A	G	SNV	SEMA3F	NM_004186.4	c.2170A>G	p.T724A	3.43e-5	0.033	12.45
9	33797969	rs149664918	T	A	SNV	PRSS3	NM_007343.3	c.514T>A	p.S172T	NA	0.014	5.774
9	33797978	rs145485932	G	A	SNV	PRSS3	NM_007343.3	c.523G>A	p.V175I	NA	0.023	1.079
13	20066994	rs78472618	T	C	SNV	TPTE2	NM_199254.2	c.115A>G	p.K39E	3.89e-5	0.005	0.025
19	58115769	rs76787024	C	T	SNV	ZNF530	NM_020880.4	c.155C>T	p.S52F	4.075e-6	0.003	16.14

Table 4.

Functional significance of gene variants identified on exome DNA sequencing that co-segregated with the familial phenotype of chronic megacolon or megaduodenum

Gene	Function	Associated GI or neurological disease	Relationship to ENS development, differentiation or survival
ATG16L1	Autophagy	Inflammatory bowel disease	None
ATN1	A triplet repeat disorder of atrophin 1	Rare neurodegenerative disorder with cerebellar ataxia, myoclonic epilepsy, choreoathetosis, and dementia	None
MUC16	Mucin 16, cell surface associated	None	none
NISCH	Nischarin: neuronal growth and differentiation in neural tissue	Regulatory protein involved in μ -opioid receptor trafficking	No relationship other than responsiveness to μ -opioid administration
NLRP9	NLR family nvrin domain containing 9 involved in immune system	None	None
PODN	Regulates cell proliferation and cell migration	None	None
PRSS3	Protease, Serine, 3	None	None
RFT1	Oligosaccharide transport	Congenital disorder of glycosylation	None
SEMA3F	Axon guidance in neural development; Negative regulation of nerve fibers; all 7 proteins SEMA (A to G) are expressed in embryonic mouse ENS	SEMA signaling pathway has been associated with non-syndromic Hirschsprung's disease, specifically SEMA3A, SEMA3C and SEMA3D (but not SEMA3B, SEMA3F, SEMA3G).	Expressed in embryonic mouse ENS; interaction with multiple genes involved in ENS pacemaker functions (specifically interstitial cells of Cajal [KIT positive] and electrical fibroblast-like cells [PDGFRB-positive]) and adrenergic neurons (tyrosine hydroxylase-positive)
TPTE2	A PTEN related gene	None	None
ZNF530	ZNFs Ubiquitously	None; however, variants in zinc finger family genes associated with Hirschsprung's disease and Mowat-Wilson syndrome (which includes Hirschsprung's phenotype)	Interaction with MEF2C genes which are expressed in ENS progenitor cells, and relevant to ENS development and differentiation

Table 5.

Genes functioning in association with *SEMA3F* and *ZNF530* genes and potentially affecting enteric nerves, based on literature

Gene and Full Name	Function	Disease Association/Phenotype
SEMA3F (semaphorin 3F) gene		
SEMA3F	axon guidance during neuronal development.	SEMA signaling pathway has been associated with non-syndromic Hirschsprung's, specifically SEMA3A, SEMA3C and SEMA3D (but not SEMA3B, SEMA3F, SEMA3G).
KIT (proto-oncogene receptor tyrosine kinase) and TEK receptor tyrosine kinase	Neural crest migration to primordial gut; precursors of gastrointestinal pacemakers, that regulate motility patterns (ref. 5,6): - interstitial cells: c-Kit(+) ICCs and - PDGFRα(+) cells	Reduced numbers of myenteric ICC or ICC networks; ²⁷ decreased densities of α-smooth muscle actin- and myenteric ICC in megacolon of Chagas disease; ²⁸ Genetic and animal models of GI motility disorders caused by loss of ICC (KIT) networks ²⁷
NOTCH1 neurogenic locus notch homolog	Notch signaling is a cell-fate determining factor in embryo development; dysregulation in many cancers e.g. hematological	Downregulation of Notch-1/Jagged-2 in human colon tissues from Hirschsprung's disease patients at aganglionosis segments
PDGFRβ platelet derived growth factor receptor beta	<i>PDGFRA</i> and <i>PDGFRB</i> genes are typical receptor tyrosine kinases; ²⁹ <i>PDGFRα</i> fibroblast-like cells are intestinal pacemakers; <i>PDGF-Rβ</i> ³⁰ induces growth of intestinal smooth muscle cells in vitro and in vivo; with inflammation, <i>PDGF-Rβ</i> may damage intrinsic inhibitory mechanisms and thus lead to hyperplasia.	Lack of <i>PDGFR</i> genes results in profound developmental defects; <i>PDGFRβ</i> expression and function impaired in LES circular smooth muscle from W/W(v) mice; ³¹ <i>PDGFRA</i> is associated with Hirschsprung's disease ³² and stromal tumors.
NGF nerve growth factor	Nerve growth stimulating activity and regulation of growth and the differentiation of sympathetic and certain sensory neurons	Hereditary sensory and autonomic neuropathy, type 5; Hirschsprung's disease; intestinal neuronal dysplasia; and aganglionosis ^{33,34}
TH tyrosine hydroxylase	Rate-limiting enzyme of catecholamine synthesis	Aganglionosis, Hirschsprung's, ENS development, Haddad (Hirschsprung's + Ondine) syndrome ^{35,36}