ORIGINAL ARTICLE

A genome-wide association analysis of chromosomal aberrations and Hirschsprung disease

JOON SEOL BAE[1], INSONG KOH[1], HYUN SUB CHEONG, JEONG-MEEN SEO, DAE-YEON KIM, JUNG-TAK OH, HYUN-YOUNG KIM, KYUWHAN JUNG, JAE HOON SUL, WOONG-YANG PARK, JEONG-HYUN KIM, and HYOUNG DOO SHIN

SEOUL AND JEJU, REPUBLIC OF KOREA; CALIF, USA

Hirschsprung disease (HSCR) is a neurocristopathy characterized by the absence of intramural ganglion cells along variable lengths of the gastrointestinal tract. Although the *RET* proto-oncogene is considered to be the main risk factor for HSCR, only about 30% of the HSCR cases can be explained by variations in previously known genes including *RET*. Recently, copy number variation (CNV) and loss of heterozygosity (LOH) have emerged as new ways to understand human genomic variation. The goal of this present study is to identify new HSCR genetic factors related to CNV in Korean patients. In the genome-wide genotyping, using Illumina's HumanOmni1-Quad BeadChip (1,140,419 markers), of 123 HSCR patients and 432 unaffected subjects (total n = 555), a total of 8,188 CNVs (1 kb ~ 1 mb) were identified by CNVpartition. As a result, 16 CNV regions and 13 LOH regions were identified as associated with HSCR (minimum *P* = 0.0005). Two top CNV regions (deletions at chr6:32675155-32680480 and chr22:20733495-21607293) were successfully validated by additional real-time quantitative polymerase chain reaction analysis. In addition, 2 CNV regions (6p21.32 and 22q11.21) and 2 LOH regions (3p22.2 and 14q23.3) were discovered to be unique to the HSCR patients group. Regarding the large-scale chromosomal aberrations (>1 mb), 11 large aberrations in the HSCR patients group were identified, which suggests that they may be a risk factor for HSCR.

[1] These authors contributed equally to this work.

From the Laboratory of Translational Genomics, Samsung Genome Institute, Samsung Medical Center, Seoul, Republic of Korea; Department of Physiology, College of Medicine, Hanyang University, Seoul, Republic of Korea; Department of Genetic Epidemiology, SNP Genetics, Inc., Seoul, Republic of Korea; Division of Pediatric Surgery, Department of Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea; Department of Pediatric Surgery, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea; Department of Pediatric Surgery, Severance Children's Hospital, Yonsei University College of Medicine, Seoul, Republic of Korea; Department of Pediatric Surgery, Seoul National University Children's Hospital, Seoul, Republic of Korea; Department of Surgery, Jeju National University Hospital, Aran 13gil 15, Jeju, Republic of Korea; Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, Calif, USA; Semel Center for Informatics and Personalized Genomics, University of California, Los Angeles, Calif, USA; Research Institute for Basic Science, Sogang University, Seoul, Republic of Korea; Department of Life Science, Sogang University, Seoul, Republic of Korea.

Submitted for publication December 10, 2015; revision submitted May 12, 2016; accepted for publication June 6, 2016.

Reprint requests: Hyoung Doo Shin, Department of Life Science, Sogang University, Seoul 121-742, Republic of Korea; e-mail: hdshin@sogang.ac.kr or Jeong-Hyun Kim, Research Institute for Basic Science, Sogang University, Seoul 121-742, Republic of Korea; e-mail: soykm@sogang.ac.kr.

1931-5244/$ - see front matter

http://dx.doi.org/10.1016/j.trsl.2016.06.001

Although further replication in a larger cohort is needed, our findings may contribute to the understanding of the etiology of HSCR. (Translational Research 2016;■:1–10)

Abbreviations: HSCR = Hirschsprung disease; CNV = copy number variation; LOH = loss of heterozygosity; qPCR = quantitative polymerase chain reaction; S-HSCR = short-segment Hirschsprung disease; L-HSCR = long-segment Hirschsprung disease; TCA = total colonic aganglionosis; SNP = single nucleotide polymorphism; PCA = principal component analysis; GWAS = genome-wide association study; LRR = log R ratio; BAF = B allele frequency; FDR = False Discovery Rate; DGV = Database of Genomic Variants; CNVR = CNV region; CGH = comparative genomic hybridization

AT A GLANCE COMMENTARY

Bae JS, et al.

Background

Hirschsprung disease (HSCR, or aganglionic megacolon) is a congenital and heterogeneous disorder that leads to intestinal obstruction or chronic constipation. Although several confounding factors, including the *RET* proto-oncogene, for HSCR have been identified, only about 30% of the HSCR cases can be explained by variations in previously known genes. Recently, copy number variation (CNV) and loss of heterozygosity have emerged as new ways to understand human genomic variation and identify risk factors for human diseases.

Translational Significance

The present study identified 16 CNV regions and 13 loss of heterozygosity regions were identified as associated with HSCR. In additional real-time quantitative polymerase chain reaction analysis, 2 top CNV regions (deletions at chr6:32675155-32680480 and chr22:20733495-21607293) were successfully validated. Our results suggest that these new chromosomal aberrations associated with HSCR may be a risk factor for the disease.

INTRODUCTION

Hirschsprung disease (HSCR) is a rare congenital gastrointestinal malformation characterized by the absence of enteric ganglia in the development of the enteric nervous system, leading to an aganglionic mega colon. The incidence of HSCR is about 1 in 5,000 live births. Although 5 ~ 20% of cases are familial forms, HSCR can also occur sporadically. HSCR shows a sex bias with a male-female ratio of ~4:1 for short segment HSCR. HSCR patients are classified into 3 subtypes: short-segment (S-HSCR, ~80% of the cases), with aganglionosis as far as the rectum and a short portion /of the colon; long-segment (L-HSCR, ~15%), with the right transverse/ascending colon affected; and total colonic aganglionosis (TCA, ~5%), with the entire colon affected.[1,2]

HSCR is a heterogeneous disease, with variable phenotypes with respect to sex, the extent of a ganglionic segment, family history, and coexistence with other congenital abnormalities (such as Down's syndrome and congenital central hyperventilation). Although several confounding factors for HSCR have been identified (such as *RET* proto-oncogene, encoding a receptor tyrosine kinase), these genetic variations account for less than 30% of the patients.[2] Previous studies have identified several chromosomal aberrations (e.g., 2q22, 10q11.2/q21.2, distal 13q, 20p11.22-p11.23) associated with HSCR.[3-6] In addition, a recent study examined 67 HSCR-related markers, including 11 previously known neuro-developmental genes, and identified copy number variants (CNVs) at 3 loci of *MAPK10*, *ZFHX1B*, and *SOX2*.[7] Even more recently, a genome-wide copy number analysis in a Chinese population discovered neuregulin 3 (*NRG3*), located on 10q23.1, to be a contributing structural variant.[8]

Of late, the density of microarrays has dramatically increased. In addition, many reported or validated CNV regions have been included in the dense genome-wide microarray containing over 1-M single nucleotide polymorphisms (SNPs). The use of a high-density microarray in a CNV association study is very important in that it prevents the missing of disease-associated CNV regions because of a low-density microarray. Two previous HSCR genetic studies using a 500K SNP microarray and a 4 × 44 K CGH array discovered new CNV regions,[7,8] but the possibility that they missed additional causal CNV regions because of the low-density arrays cannot be ruled out.

Although alterations in loss of heterozygosity (LOH) are known to routinely occur in tumor cells, recent studies have shown that LOHs are more frequent in genomic regions harboring genes in nontumor individuals.[9] In the case of the CGH array, LOH cannot be detected because the platform generates only signal intensity values. Therefore, robust high-density SNP