Five Gastric Cancer-Susceptibility Loci Identified by Genome-Wide Association Studies

Hiroe Ono¹ and Norihisa Saeki²*

¹Pharmacovigilance Department, Data Science Center, Clinical Information Division, EPS Corporation, Koraku 2-3-19, Bunkyo-ku, Tokyo 112-0004, Japan
²Division of Genetics, National Cancer Center Research Institute, Tsukiji 5-1-1, Chuo-ku, Tokyo 104-0045, Japan

Abstract

Gastric cancer is one of the most common cancers worldwide, especially in Asia, and ranks fourth in cancer death worldwide. Recently, in order to identify genes related to gastric cancer susceptibility, three genome-wide association studies were conducted using Japanese or Chinese populations, which revealed 5 gastric cancer-susceptibility loci: chromosome 1q22, 3q13.31, 5p13.1, 8q24.3 and 10q23. Further statistical and biological studies unveiled the causal candidate genes for the association in each locus: Mucin 1 (MUC1), encoding a membrane protein, in 1q22; zinc finger and BTB domain containing 20 (ZBTB20), a transcription factor, in 3q13.31; protein kinase, AMP-activated, alpha 1 catalytic subunit (PRKAA1), a ser/thr protein kinase, and/or prostaglandin E receptor 4 (PTGER4), a prostaglandin E2 receptor, and other 3 genes, in 5p13; Prostate stem cell antigen (PSCA), a membrane protein, in 8q24; phospholipase C, epsilon 1 (PLCE1), a phospholipase, in 10q23. Among these genes, MUC1 and PSCA were biologically investigated to obtain some rational for their relation to gastric carcinogenesis; however, the function of the other genes in cancer development is yet to be uncovered. As the genes identified through the genome-wide association studies have a strong potential to open a new door in cancer research, further studies are anticipated to be performed to elucidate their function and molecular pathways for application of the products to the therapy and prevention of gastric cancer.

Keywords: Helicobacter pylori; Adenocarcinoma; Single nucleotide polymorphisms

Introduction

Gastric cancer (GC) is a common cancer worldwide and the fourth neoplasm in cancer death worldwide [1]. In Asia, its incidence is quite high; for example, the age-standardized incidence rate for Japanese males is 62.1/100000 [2]. Adenocarcinoma, a major type of GC, is classified into intestinal-type GC (IGC) and diffuse-type GC (DGC) [3]. Helicobacter pylori (HP) infection is a definite carcinogen for GC carcinogenesis, especially for IGC. Consequently, it is anticipated that, by eradication of HP infection, the incidence of IGC can be reduced. On the contrary, DGC development is less influenced than IGC by HP infection, and its patients are relatively young and exhibit no male predominance, suggesting genetic factors are more influential than by environmental factors.

Single nucleotide polymorphisms (SNPs) are genetic variations present in the human genome: every 1000 nucleotides on average. Recently, genetic factors of many types of diseases were detected by conducting case-control association studies using SNPs as a genetic marker. At present, a number of genes were reported as GC genetic factors, some of which were identified by a genome-wide scanning, i.e., a genome-wide association study (GWAS), and others by a candidate-gene approach. Genes identified through GWAS have an especially strong potential to open a new door in research, because GWAS conducted in the manner of genetic statistics with SNP, a highly reliable marker. At present, a number of genes were reported as GC genetic susceptibility gene. PSCA is a glycosylphosphatidylinositol (GPI)-anchored membrane protein with an unknown biological function [7]. It was originally reported as the gene up-regulated in prostate cancer [8]. It is also up-regulated in many other types of cancers including urinary bladder cancer, renal cell carcinoma, hydatidiform mole, ovarian mucinous tumor, pancreatic cancer, non-small cell lung cancer and glioma, in which PSCA seems to promote tumor progression [9]. On the other hand, down-regulation of the gene was reported only in esophageal, gastric and gallbladder cancer (GBC) [9]. In those cancers, PSCA seems to act in suppression of tumor progression, and its growth inhibition activity on GC and GBC cells has actually been demonstrated [4,10].

In the gastric epithelium, PSCA is expressed in the isthmus/neck region in which rapidly amplifying pre-pit cells are present to support the rapid turn-over of mucus-secreting pit cells. This expression pattern is quite interesting because it is thought that the initial lesion of DGC arises in the isthmus/neck region. It is speculated that PSCA has a role in regulating the cell growth of the pre-pit cells, and that reduction of its function results in abnormal cell division and carcinogenesis. A SNP, rs2294008, in the gene was revealed to be a functional SNP affecting the transcriptional activity of the PSCA promoter, and may be a causal of the relationship between the gene and GC. It was also shown that the rs2294008 determines the position of the translation initiation codon; the T allele makes itself a part of the codon encoding the first methionine (ATG). On the contrary, the C allele replaces the encoded amino acid from methionine to threonine (ACG), resulting in...
a change of the first methionine position. The T allele associated with the risk for GC has a negative effect on the promoter activity in gastric, urinary bladder and gallbladder cancer cell lines [4,10,11]. Therefore, it is likely that people possessing the T allele have a lower amount of PSCA protein in the organs than those possessing the C allele. However, recent study showed that normal and malignant urinary bladder tissues from people with the T allele contained more PSCA transcripts than those from people with the C allele [12]. The discrepancy between the in vitro reporter assay and in vivo expression data further suggests a complexity of the PSCA regulation, which may be influenced by tissue-specific transcriptional factors and DNA methylation, as shown in gastric and gallbladder cancer cell lines [10]. As PSCA seems to have contextual functions dependent on tissue types and pathological states, the functional effect of rs2294008 may also be contextual.

The relation between PSCA and GC is demonstrated in ethnic populations other than Japanese, including Korean, Chinese, Tibetan and Caucasian [9,13]. The gene was also indentified as a susceptibility gene of urinary bladder cancer [11].

### 1q22

The GWAS conducted within the Japanese population also identified 1q22 as a GC-related locus [3]. This locus harbors 13 SNPs in strong linkage disequilibrium (LD) spanning 5 genes, and the Mucin 1 (MUC1) gene is most likely a causal for the association between this loci and GC [14]. MUC1 is a membrane-bound protein [15]. After being translated, a single MUC1 peptide was cleaved to N-terminal and C-terminal subunits, designated as MUC1-N and MUC1-C, respectively, but both subunits are localized to the cell membrane in the apical side of the epithelial cells. MUC1-C has a transmembrane domain and a cytoplasmic tail (CT), in which the latter is involved in subcellular signal transduction. On the other hand, MUC1-N present on the cell surface has multiple glycosylation sites and is thought to act as protection against many types of insults [16]. This protective function may be what contributes to the suppression of GC development.

In addition to the GWAS conducted in Japan [4], the two GWAS on the Chinese population and imputation analyses on Chinese case-control samples also listed 1q22 as a candidate for GC-related locus (rs4072037; OR = 0.75, P = 1.2 × 10^{-4}; Table 1). This locus was identified by the GWAS on the ethnic Chinese population (2240 GC cases and 2115 controls) and simultaneously identified as an esophageal cancer-related locus [6]. In GC, the association was stronger in cardia GC (rs2274223: p = 4.19 × 10^{-4}; OR = 1.57, 95% CI 1.40-1.76) and almost non-existent in non-cardia GC. Although rs2274223 locates in the nucleolar complex associated 3 homolog (NOCL3) gene, LD analyses and recombination hot spot analyses indicate that phospholipase C, epsilon 1 (PLCE1), which resides next to NOCL3 in the locus, is a causal for the association.

PLCE1 is a member of the phospholipase family that catalyzes the hydrolysis of polyphosphoinositides to generate the second messengers, which are involved in a cascade of intracellular responses that result in cell growth and differentiation and gene expression. In addition to its phospholipase C catalytic activity, this enzyme has an N-terminal domain with guanine nucleotide exchange (GEE) activity and is regulated by small monomeric GTPases of the Ras and Rho families and by heterotrimeric G proteins [23]. PLCE1 is a novel R-Ras effector mediating the function of R-Ras on the actin cytoskeleton and membrane protrusion [24].

A study of PLCE1-knockout mice demonstrated a crucial role of PLCE1 in Ras oncogene-induced de novo carcinogenesis: the mice showed a delayed onset and markedly reduced incidence of skin squamous tumors induced by initiation with 7,12-dimethylbenz(a)anthracene followed by promotion with 12-O-tetradecanoylphorbol-13-acetate, and the papillomas formed in the mice did not undergo malignant progression into carcinomas [25]. Apc(Min/+)- mice, carrying an inactivated allele of the adenomatous polyposis coli gene, exhibited higher resistance to spontaneous intestinal tumorigenesis in the PLCE1(-/-) genetic background compared with those with intact PLCE1 [26]. Low-grade adenomas of the PLCE1(-/-)/Apc(Min/+)- mice exhibited accelerated apoptosis, reduced cellular proliferation, marked attenuation of tumor angiogenesis and reduction in expression of vascular endothelial growth factor. In contrast, high-grade adenomas of the mice exhibited marked attenuation of tumor-associated inflammation without significant differences in apoptosis and proliferation. Therefore, PLCE1 seems to plays crucial roles in intestinal tumorigenesis through two distinct mechanisms, augmentation of angiogenesis and inflammation, depending on the tumor stage. It was also reported that silencing of PLCE1 decreased the invasion ability of bladder cancer-derived T24 cells, which might be through down regulation of MMP and BCL2 gene expression [27]. On the other hand, the gene was reported to have a tumor suppressive activity in colorectal cancer: it was down-regulated in 42% (21/50) of colorectal cancer.

---

**Table 1: Five gastric-cancer susceptibility loci identified by GWAS.**

<table>
<thead>
<tr>
<th>Locus</th>
<th>Representative SNP (major/minor allele)</th>
<th>Odds ratio (95% CI)</th>
<th>P-value</th>
<th>Ethnic</th>
<th>Cancer type</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1q22</td>
<td>rs2070803 (G/A)</td>
<td>1.63 (1.33-1.98)†</td>
<td>1.2 × 10^{-4}</td>
<td>Japanese</td>
<td>Diffuse</td>
<td>[4]</td>
</tr>
<tr>
<td>3q13.31</td>
<td>rs9841504 (C/G)</td>
<td>0.76 (0.69-0.83)†</td>
<td>1.7 × 10^{-4}</td>
<td>Chinese</td>
<td>Non-cardia</td>
<td>[5]</td>
</tr>
<tr>
<td>5p13.1</td>
<td>rs13361707 (T/C)</td>
<td>1.41 (1.32-1.49)†</td>
<td>7.6 × 10^{-4}</td>
<td>Chinese</td>
<td>Non-cardia</td>
<td>[5]</td>
</tr>
<tr>
<td>8q24.3</td>
<td>rs2976392 (A/G)</td>
<td>1.62 (1.38-1.89)†</td>
<td>1.1 × 10^{-4}</td>
<td>Japanese</td>
<td>Diffuse</td>
<td>[4]</td>
</tr>
<tr>
<td>10q23</td>
<td>rs2274223 (A/G)</td>
<td>1.31 (1.19-1.43)†</td>
<td>8.40 × 10^{-9}</td>
<td>Chinese</td>
<td>Cardia, non-cardia</td>
<td>[6]</td>
</tr>
</tbody>
</table>

CI: confidence interval, †additive model, P- additive model
tissues and its overexpression inhibited the proliferation of colon cancer cells in vivo as well as in vitro [28]. The gene might have a contextual function in tumorigenesis, depending on cancer types.

The association between PLCE1 and GC was also demonstrated in other studies. An association study on an eastern Chinese population (1059 GC cases and 1240 controls) revealed that a significant higher risk for gastric adenocarcinoma was associated with rs2274223 variant G allele (adjusted OR = 1.35, 95% CI = 1.14-1.60 for AG+GG vs. AA) and rs1187870 variant C allele (adjusted OR = 1.26, 95% CI = 1.05-1.50 for CG+CC vs. GG) [29]. In another study on the Chinese population, genotyping of SNP rs2274223 of 940 gastric cancer patients revealed that the AA genotype survived for a significantly shorter time than those carrying the AG and GG genotypes (log-rank P = 0.046) and that this significance was enhanced in the dominant model (AA vs. AG/GG, log-rank P = 0.014) [30].

3q13.31

A GWAS on non-cardia GC using a Chinese descendant population (discovery phase: 1006 cases and 2273 controls; replication phase: 3288 cases and 3609 controls) identified rs13361707 as a GC susceptibility locus [5]. A SNP rs9841504, showing correlation to GC with statistical significance in the GWAS, locates between two recombinant hot spots within the zinc finger and BTB domain containing 20 (ZBTB20) gene, and all the SNPs in strong LD with rs9841504 are confined within the gene, suggesting the gene is a causal of the relation.

ZBTB20 encodes a transcription factor whose function in the central nervous system was recently well-documented [31-33]. In the field of cancer research, it was reported that ZBTB20 expression level was significantly elevated in hepatocellular carcinoma (HCC) tissues and over expressed ZBTB20 protein in HCC was significantly associated with vein invasion and the recurrence or metastasis rates, which were markedly greater in HCC cases with higher ZBTB20 expression than in cases with lower ZBTB20 expression [34]. Intriguingly, in spite of these findings, ZBTB20 was reported as a negative regulator of α-fetoprotein (AFP), a tumor marker for HCC [35]; however, regulation of AFP gene expression is a complex process mediated by a number of transcriptional activators and repressors that bind the AFP gene [36].

5p13.1

This locus was identified by the GWAS on the Chinese population, simultaneously with 3q13.31 [5]. The recombination rate analyses and LD analyses on this locus exhibited a critical region for the association, which harbors 5 genes, protein kinase, AMP-activated, alpha 1 catalytic subunit (PRKAA1), prostaglandin E receptor 4 (PTGER4), Ribosomal Protein L37 (RPL37) small nucleolar RNA, C/D box 72 (SNORD72) and tetratricopeptide repeat domain 33 (TTC33). Three SNPs with the lowest P-value, less than 1×10^-10, for the association locate adjacent to PRKAA1, TTC33 and PTGER4 in strong LD, more than 0.8 of r², with the index SNP rs13361707; however, there is also a SNP rs10065570 (OR = 0.76, P = 9.80×10^-7) in moderate LD (0.6 < r² < 0.8) with the index SNP at the 5′ side of RPL37 [5]. Therefore, the candidates for the gene causal for the relation might expand to RPL37 and SNORD72, in addition to PRKAA1, TTC33 and PTGER4; meaning further studies are required for its identification.

RPL37 encodes a member of L37E ribosomal protein (RP) that is a component of the 60S subunit. Recently, it has been suggested that regulation of protein synthesis through the RP-p53-Mdm2 pathway may make a significant contribution to protecting cells against DNA damage-initiated oncogenesis [37-39]. The tumor suppressor p53 is critical in inducing cell cycle arrest in response to DNA damage, and p53 also plays a role in linking cell division with cell growth by sensing nuclear stress: when there is stress to ribosome biogenesis, and thus protein synthesis, several ribosomal proteins (RPs) such as RPL11, RPL3 and RPL5 have been shown to interact with Mdm2 and inhibit its E3 ubiquitin ligase activity towards p53. It was demonstrated that depletion of endogenous L37 led to an increase in p53 protein levels and its downstream targets p21 and Mdm2, and that silencing of L37 induced a decrease in S-phase cells, suggesting activation of p53-mediated cell cycle arrest [39].

SNORD72 is a gene for small nucleolar RNA (snoRNA) which belongs to small non-coding RNAs. Recently, significant involvement of snoRNAs in carcinogenesis has been suggested. Some snoRNAs exhibit differential expression patterns in a variety of human cancers and demonstrate their capability to affect cell transformation, tumorigenesis, and metastasis [40].

PRKAA1 protein is a catalytic subunit of AMP-activated protein kinase (AMPK), which plays a key role in regulating cellular energy metabolism. In response to reduction of intracellular ATP levels, AMPK activates energy-producing pathways and inhibits energy-consuming processes, which result in the inhibition of protein, carbohydrate and lipid biosynthesis, and the inhibition of cell growth and proliferation. The association between PRKAA1 and GC was also demonstrated in the Korean population, where it was more strongly associated with intestinal type (OR = 1.39, 95% CI = 1.22-1.58, P = 3.77×10^-7) rather than diffuse type [41]. In addition, it was also reported to be related to other cancers: breast cancer [42], colorectal cancer [43] and cervical cancer [44], which may be through its influence on the metabolism of fatty acids, carbohydrates, hormone and so on.

PTGER4 encodes a member of the G-protein coupled receptor family, which is a receptor for prostaglandin E2 (PGE2). It was reported that Cyclooxygenase-2 (COX-2) plays an important role in gastric tumorigenesis through PGE2 signaling which transactivates an epidermal growth factor receptor (EGFR) [45]. However, it might have a protective function for gastric mucosa. In mice, it was suggested that a PTGER4-selective agonist may prevent indomethacin-induced gastric lesions and promote healing of existing and indomethacin-aggravated gastric ulcers, via promoting proliferation and survival of mucous epithelial cells [46]. In addition, it was reported that Helicobacter pylori (H. pylori) induces cyclooxygenase-2 (COX-2) expression in the stomach, which may play an important role in resistance to HCl/ EtOH damage in H. pylori-infected mice by activating PGE2 receptors [47]. The expression sites of PGE2 receptors in the gastrointestinal tract were well elucidated in rats [48]. In the stomach, PTGER1 mRNA was detected in gastric muscle layers, whereas PTGER3 and PTGER4 expression was mainly present in the gastric mucosal layer containing epithelial cells. In gastric epithelium, parietal cells were found to have both PTGER3 and PTGER4 expression. At lower concentrations, PGE2 inhibited gastric acid secretion by parietal cells, probably through PTGER4. At higher concentrations, however, it stimulated it. On the other hand, mucous secreting cells, i.e., pit cells, possessed only PTGER4 mRNA. Moreover, the SNPs tagging the PTGER4 were reported to be associated with ulcerative colitis in an Italian cohort [49], and also with Crohn disease [50,51], suggesting it has a role in inflammation in the gastrointestinal tract.

TTC33 encodes an osmosis-responsive factor but its function is yet to be documented.
Future Perspective

To date, three GWAS have been conducted and have identified 5 loci related to GC susceptibility. The studies performed with statistical methodology revealed several genes which are novel in the field of GC research. These genes have a strong potential to open a new door in cancer research. At present, we know only the 5 loci, but in the future, GWAS may detect other GC-related loci with less association, by reinforcing detection power with larger case-control sets. Further studies are anticipated to elucidate the function of the GC susceptibility genes and their molecular pathways, and also the interaction between the genes and environmental factors, in order to apply the harvest in research to the therapy and prevention of gastric cancer.

Acknowledgement

This work was supported by a Grant-in-Aid for Scientific Research (KAKENHI) by the Japan Society for the Promotion of Science (No. 23501327).

References


This article was originally published in a special issue, Gastrointestinal Cancer handled by Editor(s). Dr.Aliasger Amin, James Cook University Hospital Middlesbrough, United Kingdom