Common variant in 6q26-q27 is associated with distal colon cancer in an Asian population

R Cui, Y Okada, S G Jang, J L Ku, J G Park, Y Kamatani, N Hosono, T Tsunoda, V Kumar, C Tanikawa, N Kamatani, R Yamada, M Kubo, Y Nakamura, K Matsuda

ABSTRACT

Background and aim Colorectal cancer (CRC) is a multifactorial disease with both environmental and genetic factors contributing to its development. The incidence of CRC is increasing year by year in Japan. Patients with CRC in advanced stages have a poor prognosis, but detection of CRC at earlier stages can improve clinical outcome. Therefore, identification of epidemiological factors that influence development of CRC would facilitate the prevention or early detection of disease.

Methods To identify loci associated with CRC risk, we performed a genome-wide association study (GWAS) for CRC and sub-analyses by tumour location using 1583 Japanese CRC cases and 1898 controls. Subsequently, we conducted replication analyses using a total of 4809 CRC cases and 2973 controls including 225 Korean subjects with distal colon cancer and 377 controls.

Results We identified a novel locus on 6q26-q27 region (rs7758229 in SLC22A3, p=7.92×10^{-9}, OR of 1.28) that was significantly associated with distal colon cancer. We also replicated the association between CRC and SNPs on 8q24 (rs6983267 and rs7837328, p=1.51×10^{-8} and 7.44×10^{-8}, ORs of 1.18 and 1.17, respectively). Moreover, we found cumulative effects of three genetic factors (rs7758229, rs6983267, and rs4939827 in SMAD7) and one environmental factor (alcohol drinking) which appear to increase CRC risk approximately twofold.

Conclusions We found a novel susceptible locus in SLC22A3 that contributes to the risk of distal colon cancer in an Asian population. These findings would further extend our understanding of the role of common genetic variants in the aetiology of CRC.

INTRODUCTION

Colorectal cancer (CRC) is the third most common cancer and the fourth leading cause of cancer death worldwide. The recent development of novel drugs and therapeutics has remarkably improved overall survival, and the 5-year survival rate is around 90% for patients who are diagnosed at stage I. However, the prognosis of patients with CRC in an advanced stage is still disappointing. Hence, identification of epidemiological factors that influence development of the disease would facilitate its prevention or early detection and subsequently provide better prognosis.

Family history is acknowledged to be one of the strong risk factors, and an approximately twofold increased risk for CRC was observed among patients who have a first-degree relative with CRC. Thus, nearly 15% of patients with CRC have a positive family history of disease. Although inherited susceptibility was thought to account for ~35% of all CRC, high-risk germline mutations in APC, DNA mismatch repair genes, MUTYH, SMAD4, BMPR1A, and LKB1 account for <6% of all cases. Therefore, the remaining heritable CRC risk (approximately 30%) would be caused by the combination of common variants with modest effects.

Recent genome-wide association studies (GWASs) among European populations have identified a subset of CRC susceptibility loci in 8q24, 8q23.3 (EIF3H), 10p14, 11q23, 15q13, 13q21 (SMAD7), 14q22.2 (BMP4), 16q22.1 (CDH1), 19q13.1 (RHPN2), and 20p12.3. However, no GWAS for CRC in an Asian population.
population has been performed. To explore the variants that predispose to the disease in an Asian population, we conducted a GWAS for CRC and sub-analyses by tumour location.

METHODS

Study population
In this study, we used a total of 6167 CRC cases and 4494 control subjects from a Japanese population. CRC patients were categorised into two main groups (colon and rectal cancers) according to the tumour location. For the colon cancer, we further categorised into proximal colon (caecum, ascending colon, and transverse colon) and distal colon (descending colon, and sigmoid colon) cancers. Characteristics of each cohort are shown in Table 1. Case samples of the Japanese population were obtained from BioBank Japan (http://biobankjp.org/). Control DNA samples in the screening stage were obtained from healthy volunteers (n=904, 74.4% males, mean age at diagnosis=52.5 years, SD±14.3) as well as from BioBank Japan (n=994, kelo-dosis, chronic hepatitis B, pulmonary tuberculosis, and drug rash). To increase the power to detect genetic factors related with CRC, we used cases with higher hereditary predisposition in our screening stage. CRC cases who developed the disease at a younger age (<60 years old) or had at least one first-degree relative with a history of CRC were enriched in our screening stage. The genome-wide significance threshold is 2.54×10⁻⁸ (0.05/391749×5). The statistical powers to detect a variant with OR of 1.5 and MAF of 0.2 were estimated to be 62.6%, 99.9% and 94.2% (screening stage, first and second replication) for CRC, and 19.1%, 95.8%, and 77.2% for distal colon cancer, respectively. ORs and CIs were calculated using the major allele as a reference. Since alcohol intake of more than two standard drinks (28 g of pure alcohol per day) was shown to increase the risk of CRC, we classified the subjects into three categories: non-drinkers (0–1 g/day of alcohol), light drinkers (1–28 g/day of alcohol), or heavy drinkers (>28 g/day of alcohol). Multiple logistic regression analysis was used to assess the contributions of the confounding factors with the R package (version 2.8.1). Age and gender were designated as regulatory factors, and the following explanatory variables were included in the analysis: alcohol consumption status (0=non-drinker, 1=light drinker, 2=heavy drinker), tobacco smoking (0=never smoker, 1=smoker) and rs7758229 genotype (0=GG, 1=TT).

SNP genotyping and quality control
Platforms used in each stage are shown in Table 1. Analyses using Illumina Beadchip (Illumina, San Diego, California, USA) were conducted for the subjects with cancers or diabetes from control groups. These samples were genotyped 1583 Japanese individuals with CRC and 1898 control subjects. These samples were genotyped using the Illumina Human610-Quad BeadChip. These findings indicated that different genotyping methods are not likely to cause the inflation of association in our study.

In the screening stage, we genotyped 1595 CRC cases and 1905 control subjects. These samples were genotyped using the Illumina Human610-Quad BeadChip in cases and the Illumina HumanHap550v3 BeadChip in controls. In fact, genotype concordance between these two BeadChips was 99.99% among 182 duplicate samples, indicating a low possibility of genotype error. The samples with a call rate of < 0.98 were excluded from our analysis (12 cases and five controls). Then we applied SNP quality control as follows: call rate ≥0.99 in cases and controls, Hardy–Weinberg p=1×10⁻⁷ in controls. Finally, 496531 common SNPs between Human610-Quad Beadchip and HumanHap550v3 Genotyping BeadChip on autosomal chromosomes passed the quality control criteria. We selected 391749 SNPs with minor allele frequency (MAF) of ≥0.05 in either case or control samples for further analyses, considering statistical power in the replication analyses.

Statistical analysis
In the screening stage, the associations between each SNP and CRC were assessed using the Cochran–Armitage trend test. Thirty-six SNPs that exhibited false discovery rate Q value ≤0.2 (p≤2.5×10⁻⁵) were further analysed using an independent cohort consisting of 3099 CRC cases and 1777 controls. Sub-analyses by tumour location were also performed applying the same criteria (p≤2.5×10⁻⁵). The significance thresholds were set to be 0.05 in the first, second and third replication study, and 1.27×10⁻⁵ (0.05/391749×5) in the meta-analysis. When we take account of the sub-group analyses for multiple testing correction, the genome-wide significance threshold is 2.54×10⁻⁸ (0.05/391749×5) in cases and controls, indicating a low possibility of genotype error. The samples with a call rate of < 0.98 were excluded from our analysis (12 cases and five controls). Then we applied SNP quality control as follows: call rate ≥0.99 in cases and controls, Hardy–Weinberg p=1×10⁻⁷ in controls. Finally, 496531 common SNPs between Human610-Quad Beadchip and HumanHap550v3 Genotyping BeadChip on autosomal chromosomes passed the quality control criteria. We selected 391749 SNPs with minor allele frequency (MAF) of ≥0.05 in either case or control samples for further analyses, considering statistical power in the replication analyses.

RESULTS

To identify common variants that influence CRC risk, we genotyped 1583 Japanese individuals with CRC and 1898...
control individuals (supplementary figure 1A). To overcome the relatively low statistical power in the screening stage, patients under 60 years old who developed cancer or who had a positive family history were preferentially enrolled in the screening stage, as mentioned in the methods section. We also performed sub-analysis by tumour location (colon, proximal colon, distal colon and rectal cancer) to explore common variants that predispose to some subsets of CRC (supplementary figure 1B–E). Application of the Cochran–Armitage trend test to all the tested SNPs indicated that the genomic inflation factor λ was 1.05, 1.04, 1.02, 1.03 and 1.02 for the colorectal, proximal colon, distal colon and rectal cancers, respectively (supplementary figure 2A–E). For CRC, the inflation factor λ was 1.05 in our study. Since the inflation factor is highly dependent on sample size,28 19 we calculated it by using 1000 cases and 1000 controls. As a result, the inflation factor becomes as small as 1.03, implying a low possibility of false positive associations due to population stratification.

SNPs with a false discovery rate Q value ≤0.2 (p≤2.5×10−5) in the screening stage were considered as candidates. Thus, 36 SNPs for CRC, 27 SNPs for colon cancer, 20 SNPs for proximal colon cancer, 18 SNPs for distal colon cancer, and nine SNPs for rectal cancer were selected for further analyses (supplementary tables 2—6). In the first replication study, we genotyped these candidate SNPs using independent cohorts consisting of up to 3099 cases and 1777 controls (table 1). Two SNPs for CRC (table 2) and one SNP for distal colon cancer (table 3) exhibited a p value lower than 0.05. Then we analysed these SNPs in 1485 CRC cases, 489 distal colon cancer cases, and 819 controls, respectively. As a result, all three SNPs showed a p value lower than 0.05 in the second replication study. Consequently, we identified a significant association between CRC and SNPs on 8q24 (rs6983267 and rs7837328, p≤1.51×10−8 and 7.44×10−3, respectively; table 2), which have been reported to be associated with CRC in studies of Caucasian subjects.8 In addition, we identified a novel locus (rs7758229 on 6q26-q27) that was significantly associated with distal colon cancer. Meta-analysis of all stages showed a Mantel–Haenszel p value of 1.07/84 (OR of 1.28, P=0.286) and of 1.29 (table 3). Since rs7758229 did not associate with any disease that was used in the control group in our study (supplementary figure 4), the case-mix cohort is not likely to affect the association between rs7758229 and distal colon cancer.

For this novel locus, we conducted a replication study using 225 Korean subjects with distal colon cancer and 377 controls. Although the association was not significant, we observed a similar trend in the samples from the Korean subjects (p=0.286 with OR of 1.16) and Mantel–Haenszel p value for independence had improved from 1.07/84 to 7.92/10−9 (OR of 1.28, P=0.20; table 3) when we conducted a meta-analysis of the Japanese and Korean study with a fixed-effects model.
Interestingly, rs7758229 exhibited a much stronger effect on the risk of distal colon cancer among younger populations and the patients with a family history of the disease (supplementary figure 5). Taken together, this association appears to be true.

SNP rs7758229 is located within an intron 5 of SLC22A3 (the solute carrier family 22, member 3), one of the organic cation transporter genes. Organic cation transporters are critical for the elimination of some drugs and environmental toxins. This SNP is located within a recombination hot spot between two linkage disequilibrium blocks spanning an approximately 350-kb region on chromosome 6q26-q27 (figure 1). To further investigate this candidate region, we conducted imputation analysis using a screening stage genome-wide dataset. As a result, many SNPs within the SLC22A3 gene locus indicated strong association (supplementary figure 5), suggesting a possible role of SLC22A3 in the pathogenesis of distal colon cancer. To identify the causative variant(s) that might alter the function or expression level in the pathogenesis of distal colon cancer. To identify the causative variant(s) that might alter the function or expression level in the pathogenesis of distal colon cancer, we re-sequenced using genomic DNA from 48 cases and 48 healthy controls. As a result, we identified three novel SNPs consisting of one non-synonymous SNP (novel v1, Serine106-Glycine) in exon 1 and two synonymous SNPs (novel v2, novel v3) in exon 3 and 8 (supplementary table 7). In addition to these three novel SNPs, 19 tag SNPs were genotyped using 1916 distal colon cancer cases (screening stage, the first and the second replication studies) and 1818 controls (screening stage). Although three SNPs in intron 1 (rs847442) and intron 5 (rs3125636 and rs3106164) exhibited suggestive associations with distal colon cancer; no variants indicated stronger association (p = 3.51 × 10⁻⁶, OR of 1.21, respectively) after adjustment of age and sex.

Table 2: The combined results of screening stage, first replication and second replication analyses for colorectal cancer

<table>
<thead>
<tr>
<th>SNP</th>
<th>Allele (1/2)</th>
<th>Chr/gene</th>
<th>Stage</th>
<th>Case</th>
<th>Control</th>
<th>OR (95% CI)</th>
<th>p Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs61833267</td>
<td>(T/G)</td>
<td>8q24/Not in gene</td>
<td>Screening stage</td>
<td>593</td>
<td>757</td>
<td>233</td>
<td>38.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>First replication</td>
<td>1214</td>
<td>1476</td>
<td>408</td>
<td>36.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Second replication</td>
<td>540</td>
<td>701</td>
<td>239</td>
<td>39.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Combination</td>
<td>2347</td>
<td>2934</td>
<td>880</td>
<td>38.09</td>
</tr>
<tr>
<td>rs7837328</td>
<td>(A/G)</td>
<td>8q24/Not in gene</td>
<td>Screening stage</td>
<td>209</td>
<td>750</td>
<td>624</td>
<td>36.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>First replication</td>
<td>366</td>
<td>1448</td>
<td>1283</td>
<td>35.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Second replication</td>
<td>215</td>
<td>688</td>
<td>580</td>
<td>37.69</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Combination</td>
<td>790</td>
<td>2886</td>
<td>2487</td>
<td>36.23</td>
</tr>
</tbody>
</table>

*p Values derived from Cochran–Armitage trend tests.
† OR of minor allele from 2×2 allele frequency table.
MAF, minor allele frequency

Table 3: Results of replication studies and meta-analyses for distal colon cancer

<table>
<thead>
<tr>
<th>SNP</th>
<th>Allele (1/2)</th>
<th>Gene</th>
<th>Stage</th>
<th>Case</th>
<th>Control</th>
<th>OR (95% CI) *</th>
<th>p Value†</th>
<th>p Heterogeneity‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs7758229</td>
<td>(T/G)</td>
<td>SLC22A3</td>
<td>Screening stage</td>
<td>53</td>
<td>232</td>
<td>292</td>
<td>97</td>
<td>641</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>First replication</td>
<td>62</td>
<td>313</td>
<td>472</td>
<td>96</td>
<td>603</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Second replication</td>
<td>30</td>
<td>191</td>
<td>268</td>
<td>33</td>
<td>295</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Meta-analysis (Japanese)§</td>
<td>29</td>
<td>191</td>
<td>268</td>
<td>33</td>
<td>295</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Third replication</td>
<td>9</td>
<td>85</td>
<td>130</td>
<td>15</td>
<td>124</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Meta-analysis (Asian)§</td>
<td>29</td>
<td>191</td>
<td>268</td>
<td>33</td>
<td>295</td>
</tr>
</tbody>
</table>

*OR of minor allele from 2×2 allele frequency table.
† p Values derived from Cochran–Armitage trend tests.
‡ Results of Breslow–Day tests.
§Meta-analyses of Japanese and Asian populations, respectively. OR and p value for independence test were calculated by the Mantel–Haenszel method.
approximately twofold higher risk of developing CRC compared with individuals with a score of 2 (figure 2B). These results indicate that genetic and environmental factors play a crucial role in the development of CRC.

DISCUSSION

CRC that arises proximal or distal to the splenic flexure exhibits differences in incidence according to age, gender and ethnicity. For example, distal colon cancers predominantly occur in white males, while proximal colon cancers are frequent among older African–American females.21 22 The hereditary familial forms of CRC, familial adenomatous polyposis (FAP) and hereditary non-polyposis CRC (HNPCC) also exhibit markedly different clinical features.23 Nearly 100% of individuals with FAP will develop CRC in the distal colon.24 In contrast, approximately 70% of large bowel tumours in individuals with HNPCC arise in the proximal colon.25 In addition, mutations in TP53 are approximately 1.5- to 3-fold more frequent in distal colon cancer compared with proximal colon cancer.26 Recently, the incidence of colon cancer in Japan has been increasing.27 In addition, CRC occurs more frequently in the proximal colon and less frequently in the rectum among the Japanese–American population in Hawaii compared to native Japanese in Japan.28 These facts suggest that different genetic and environmental factors contribute to the pathogenesis of rectal, distal and proximal colon cancer, respectively.

To date, many studies have shown the associations between various polymorphisms and the risk of CRC, but few have analysed cancer risk by tumour location. In this study, we have identified a novel locus on 6q26-q27, tagged by rs7758229 in SLC22A3 which was significantly associated with distal colon cancer in Asians. The 6q26-q27 region contains four genes, including SLC22A2, SLC22A3, LPAL2 and LPA. Imputation analysis of 6q26-q27 region indicated that SNPs around SLC22A3 revealed strong associations. Interestingly, SNP rs9364554 in intron 5 of SLC22A3 was shown to associate with prostate cancer in Caucasian populations.29 Thus SLC22A3 is likely to be associated with multiple cancers.

SLC22A3 is a member of the organic cation transporter family that is highly expressed in liver, kidney, intestine and brain.30

Table 4 Logistic regression analysis

<table>
<thead>
<tr>
<th></th>
<th>OR (95% CI)</th>
<th>p Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs7758229</td>
<td>1.31 (1.20 to 1.43)</td>
<td>5.61E-09†</td>
</tr>
<tr>
<td>Alcohol consumption</td>
<td>1.21 (1.12 to 1.31)</td>
<td>1.53E-06†</td>
</tr>
<tr>
<td>Smoking</td>
<td>0.92 (0.80 to 1.05)</td>
<td>2.10E-01</td>
</tr>
</tbody>
</table>

ORs were adjusted for age and sex.
*p Values derived from the Wald test.
†The p value of less than 0.05 was considered statistically significant.
This family members play a critical role for the transport of cationic drugs, toxins, and endogenous metabolites. SLC22A3 is expressed in many cancer cell lines, such as colorectal and kidney cancers, and its expression level is correlated with the sensitivity to chemotherapeutic agents. Since several toxins or endogenous metabolites such as lipopolysaccharide and linoleic acid metabolite were shown to induce tumour formation, SLC22A3 might be involved in colorectal tumorigenesis through the clearance of some carcinogen.

Numerous studies have indicated that alcohol drinking could be positively associated with CRC risk. In addition, alcohol intake increases the risk of distal colon cancer and/or rectal cancer than proximal colon cancer. Furthermore, the effect of alcohol drinking is stronger among Asian populations because of their relatively high prevalence of a slow-metabolising aldehyde dehydrogenase variant. Similarly, our study validated the hypothesis that alcohol consumption was strongly associated with distal colon cancer in the Japanese population.

In summary, we have identified a novel susceptible locus in SLC22A3 that contributes to a risk of distal colon cancer. The incidence of CRC was increasing with the Westernisation of lifestyle and dietary habit in Japan. However, only three of nine CRC susceptibility loci from studies of Caucasians exhibited a p value of less than 0.05 in our study. These results suggest an interesting racial diversity between Asians and Caucasians in the colorectal pathogenesis. Although further functional studies are essential, our findings extend the understanding of the cumulative role of genetic and environmental factors in the colorectal carcinogenesis in the Asian population.

**REFERENCES**

Common variant in 6q26-q27 is associated with distal colon cancer in an Asian population

R Cui, Y Okada, S G Jang, J L Ku, J G Park, Y Kamatani, N Hosono, T Tsunoda, V Kumar, C Tanikawa, N Kamatani, R Yamada, M Kubo, Y Nakamura and K Matsuda

Gut published online January 17, 2011

Updated information and services can be found at:
http://gut.bmj.com/content/early/2010/12/31/gut.2010.215947

These include:

Supplementary Material
Supplementary material can be found at:
http://gut.bmj.com/content/suppl/2011/01/24/gut.2010.215947.DC1

References
This article cites 39 articles, 5 of which you can access for free at:
http://gut.bmj.com/content/early/2010/12/31/gut.2010.215947#BIBL

Open Access
This is an open-access article distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license. See: http://creativecommons.org/licenses/by-nc/2.0/ and http://creativecommons.org/licenses/by-nc/2.0/legalcode.

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections

Open access (404)
Colon cancer (1547)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/