rs12904 Polymorphism in the 3’UTR of EFNA1 is Associated with Colorectal Cancer Susceptibility in a Chinese Population

Ying-Ying Mao¹, Fang-Yuan Jing¹, Ming-Juan Jin¹, Ying-Jun Li¹, Ye Ding¹, Jing Guo¹, Fen-Juan Wang², Long-Fang Jiang², Kun Chen¹*

Abstract

Accumulated evidence has indicated that Ephrin A1 (EFNA1) is associated with angiogenesis and tumorigenesis in various types of malignancies, including colorectal cancer (CRC). In the current study, we performed an online search using the public microarray database to investigate whether EFNA1 expression might be altered in CRC tissues. We then conducted a case-control study including 306 subjects (102 cases and 204 well-matched controls) in Xiaoshan County to assess any association between genetic polymorphisms in EFNA1 and CRC susceptibility. Searches in the Oncomine expression profiling database revealed EFNA1 to be overexpressed in CRC tissue compared with adjacent normal tissue. The rs12904 G-A variant located in the 3’ untranscribed region (UTR) of EFNA1 was observed to be associated with CRC susceptibility. Compared with the AA homozygous genotype, those carrying GA genotype had a decreased risk of developing CRC (odds ratio (OR) =0.469, 95% confidence interval (CI): 0.225-0.977, and \( P =0.043 \)). The association was stronger among smokers and tea drinkers, however, no statistical evidence of interaction between rs12904 polymorphism and smoking or tea drinking on CRC risk was found. Our results suggest that EFNA1 is involved in colorectal tumorigenesis, and rs12904 A>G polymorphism in the 3’ UTR of EFNA1 is associated with CRC susceptibility. Larger studies and further mechanistic investigations are warranted to confirm our findings.

Keywords: EFNA1 - single-nucleotide polymorphism - colorectal cancer - susceptibility

Introduction

Colorectal cancer (CRC) is one of the most commonly diagnosed cancer types worldwide, with an estimation of more than 1.2 million new cases and 608,700 deaths occurred in 2008 (Jemal et al., 2011). The incidence rates of CRC are increasing rapidly in a number of countries historically at low risk, including Japan, Korea and China (Jemal et al., 2010). In Shanghai for instance, the incidence rates of CRC in both males and females increased by 50% from 1987 to 2002 (Center et al., 2009).

Environmental risk factors as well as acquired and constitutional genetic variations are considered as predominant contributors in the etiology of CRC. Modifiable risk factors include dietary factors and lifestyle-related factors such as smoking and moderate-to-heavy alcohol drinking (Huxley et al., 2009). In addition to these environmental factors, genetic variations of CRC-related genes contribute substantially to CRC risk as well. EFNA1 maps at 1q21-q22 on chromosome 1, and encodes for Ephrin A1 which is a glycosyl-phosphatidylinositol (GPI)-anchored ligand preferentially binding to Eph receptor tyrosine kinases. EFNA1 was originally isolated as a secreted protein in conditioned media from cultures of human umbilical vein endothelial cells treated with tumor necrosis factor-α (Holzman et al., 1990), and was found to play crucial roles in the migration and adhesion of cells during development (Poliakov et al., 2004). Recently, accumulated evidence has indicated that EFNA1 is associated with angiogenesis and tumorigenesis in various types of malignancies. Expression of EFNA1 was observed in various types of tumor cells and blocking EFNA1 with soluble EphA2-Fc decreased tumor-associated angiogenesis and consequently tumor progression (Brantley et al., 2002; Dobrzanski et al., 2004). In hepatocellular carcinoma (HCC), EFNA1 mRNA was overexpressed compared with adjacent nontumorous tissues, and the level of secreted EFNA1 in serum samples from HCC patients were significantly higher than those from healthy controls (Cui et al., 2010). Likewise, elevated expression of EFNA1 was observed in gastric cancer (Nakamura et al., 2005), bladder cancer (Abraham et al., 2006) and more recently, prostate cancer (Larkin et al., 2012).

Several studies investigating the role of EFNA1 in colorectal tumorigenesis yielded similar findings. Potla

¹Department of Epidemiology and Biostatistics, Zhejiang University School of Public Health, ²Center for Disease Control and Prevention of Xiaoshan, Hangzhou, China  *For correspondence: ck@zju.edu.cn
et al. (2002) found that decreased EFNA1 expression in HT29 colon carcinoma cells reduced the growth of the tumor cells in three-dimensional spheroid static cultures. Lips et al. (2008) and Shi et al. (2012) reported that EFNA1 mRNA expression increased from rectal adenoma to carcinoma. Furthermore, a recent study conducted by Yamamoto et al. suggested that EFNA1 was an independent prognostic factor for CRC, and its loss-of-function was associated with reduced proliferation and decreased invasion and migration of CRC cell lines (Yamamoto et al., 2013).

Given the apparent importance of EFNA1 in colorectal tumorigenesis, we hypothesized that genetic variations in EFNA1 gene may be associated with CRC susceptibility. In the current study, we first searched the public microarray database to investigate whether altered expression of EFNA1 can be observed in CRC tissues. We then conducted a case-control study to evaluate the association between genetic polymorphisms in EFNA1 and CRC susceptibility.

Materials and Methods

Expression analysis of EFNA1 in CRC tissues

We performed an online search using the Oncomine database (www.oncomine.org; last accessed on May 22, 2013) for expression array comparisons involving CRC tissues and normal controls. The key words used were: Gene “EFNA1”; Cancer type: “Colorectal cancer”; Analysis type: “Cancer vs. Normal Analysis”; and the results were filtered by $P \leq 0.005$ and fold change $\geq 2$. The search returned two arrays: the “Notterman Colon” which compares EFNA1 expression in colon adenocarcinoma and paired adjacent normal tissues, and the “Kaiser Colon” which compares EFNA1 expression in colorectal adenocarcinoma and normal colorectal tissues. Detailed information regarding the tissue sample collection and experimental protocols can be found in the Oncomine database or from the original publications (Notterman et al., 2001; Kaiser et al., 2007).

Study subjects

CRC patients were recruited from four local hospitals from May 2005 to October 2009. Eligible cases were incident and histologically confirmed CRC, living in Xiaoshan County for $\geq 20$ years, mentally competent to complete the interview and with no previous diagnosis of familial adenomatous polyposis, ulcerative colitis or Crohn’s disease. Healthy controls with no previous history of cancer were recruited in parallel and 1:2 matched to cases by age (± 5 years), gender, ethnicity, and residential area. Face-to-face interviews were conducted by trained interviewers, who administered a structured questionnaire relating to demographic characteristics (e.g., age, gender, body mass index (BMI), marital status, occupation and education level), family history of cancer, menopausal history, and lifestyle-related factors (e.g., diet, smoking, alcohol drinking and tea drinking). After interview, 2ml blood sample was collected into sodium citrate anticoagulant tubes and stored at -80°C for DNA isolation. A total of 102 cases and 204 well-matched controls with DNA samples available were included in the current study. All the study procedures were reviewed and approved by the local medical ethical committee and written informed consents were obtained from all study subjects before participation.

SNP selection and Genotyping

Genomic DNA was isolated from peripheral blood samples for each study subject using the modified salting-out procedure (Nasiri et al., 2005). Single nucleotide polymorphisms (SNPs) in the genomic region from 1,500bp upstream to 1,500bp downstream of EFNA1 with minor allele frequencies (MAF) of $> 0.1$ within the Han Chinese in Beijing (CHB) populations were identified from the dbSNP database (http://www.ncbi.nlm.nih.gov/projects/SNP/). Tag SNPs representing SNPs with pairwise correlation of $r^2 > 0.8$ were selected using the Tagger algorithm implemented in the haplview interface (Barrett, 2009). rs12904 which captures 100% alleles (Tag SNPs rs4971066, rs9297, rs4745 and rs12904) in EFNA1 with mean $r^2$ of 0.978 was chosen. Genotyping was performed using the Sequenom MassArray genotyping platform (Sequenom, Inc., San Diego, CA, USA). Blinded replicate samples (10% of the samples) and negative controls (one for each 96-well plate) were interspersed throughout the genotyping assays. The call rate for the SNP genotyped was $>99\%$, and the concordance rate for these quality control samples was 100%.

Statistical analyses

All statistical analyses were performed using the SAS statistical software, version 9.2 (SAS Institute, Cary, NC, USA), unless otherwise noted. Differences in the distribution of demographic and lifestyle characteristics between cases and controls were assessed using t-test for continuous variables and Pearson $\chi^2$-test for categorical variables. Departures from Hardy-Weinberg equilibrium (HWE) were tested among controls using goodness-of-fit $\chi^2$ analysis. Associations between rs12904 polymorphism and CRC risk were assessed by conditional multivariate logistic regression models including the following covariates: age, education level, occupation, cigarette smoking, alcohol drinking and tea drinking. Stratified analysis was also performed to explore potential gene-environment interactions and odds ratios (ORs) and 95% confidence intervals (CIs) were determined using unconditional logistic regression. Two-sided P-values of less than 0.05 were considered as statistically significant.

Results

We searched the Oncomine expression profiling database to determine whether EFNA1 expression was altered in patients with CRC relative to controls. We used the keywords “EFNA1” and “colorectal cancer” and conditional filters of $P \leq 0.005$ and fold change $\geq 2$. The search returned two arrays: the “Notterman Colon” and the “Kaiser Colon”. The study of “Notterman Colon” examined gene expression in colon adenocarcinoma and paired adjacent normal tissues. As shown in Figure 1A, EFNA1 expression was increased in 14 out of 18
Adenocarcinoma-adjacent normal tissue pairs \( (P < 0.005, \text{ fold change } = 2.588) \). Likewise, in the study of “Kaiser Colon” which examined gene expression in colon, cecum, retal-sigmoid and rectal adenocarcinoma and normal colon tissues, \textsc{efna1} expression was significantly lower in normal colorectal tissues than in adenocarcinoma tissues \( (P = 8.64 \times 10^{-5}, \text{ fold change } = 1.780-3.486) \). The median expression level of \textsc{efna1} and its interquartile range were shown for each subgroup

Table 1 summarizes the distributions of demographic characteristics and selected risk factors for the CRC case-

| Variables                | Cases (n = 102) | Controls (n = 204) | \( P \)-value | Adjusted OR (95% CI) | Adjusted OR (95% CI)* | Adjusted \( P \)-value*
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 65</td>
<td>50 49.0</td>
<td>120 58.8</td>
<td>0.104</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \geq 65 )</td>
<td>52 51.0</td>
<td>84 41.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>44 43.1</td>
<td>88 43.1</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females</td>
<td>58 56.9</td>
<td>116 56.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Educational level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illiterate</td>
<td>54 52.9</td>
<td>81 39.7</td>
<td>0.090</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary school</td>
<td>42 41.2</td>
<td>104 51.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle school and above</td>
<td>6 5.9</td>
<td>18 8.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>97 95.1</td>
<td>190 93.1</td>
<td>0.834</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>5 4.9</td>
<td>11 6.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Occupation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farmer</td>
<td>73 71.6</td>
<td>118 57.8</td>
<td>0.013</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>27 26.5</td>
<td>84 41.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cigarette smoking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>61 59.8</td>
<td>117 57.4</td>
<td>0.682</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>41 40.2</td>
<td>87 42.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol drinking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>57 55.9</td>
<td>119 58.3</td>
<td>0.683</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>45 44.1</td>
<td>83 41.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tea drinking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>53 52.0</td>
<td>118 57.8</td>
<td>0.306</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>49 48.0</td>
<td>85 41.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Subjects with missing information were not included in the control study. Because of 1:2 matching used in the study design, no significant differences in the distributions of age, gender, ethnicity and residential area between cases and controls were observed. CRC cases were more likely to be farmers compared with controls \( (P = 0.013) \), but no other demographic characteristics and lifestyle-related factors such as smoking, alcohol drinking and tea drinking status differed significantly between cases and controls. The tag SNP rs12904 located in the 3’ untranslated region (UTR) of \textsc{efna1} was chosen to represent all the variation in this gene region. No significant departure from Hardy-Weinberg equilibrium was observed among the controls \( (P = 0.677) \). The genotype distributions between cases and controls are listed in Table 2. In the unstratified analysis, the association between variant allele of rs12904 and CRC risk was found to be statistically significant. Compared with rs12904 AA homozygous genotype, those carrying GA genotype had a decreased risk of developing CRC \((OR=0.469, 95\%\ CI: 0.225-0.977)\). However, under the dominant model, rs12904 A>G was borderline associated
with a decreased CRC risk (OR=0.557, 95% CI: 0.290-1.070).

In the following stratified analyses by lifestyle-related factors such as smoking, alcohol drinking and tea drinking, only in the subgroups of smokers (GA vs. AA: OR=0.100, 95% CI: 0.014-1.977; GA+GG vs. AA: OR=0.098, 95% CI: 0.014-0.685) and tea drinkers (GA+GG vs. AA: OR=0.106, 95% CI: 0.012-0.975) did the decreased risk remain statistically significant (Table 3), though the subgroups had limited observations. No statistical evidence of interaction was found between rs12904 polymorphism and smoking or tea drinking on CRC risk.

Discussion

In the present study, we examined EFNA1 expression in CRC tissues using the public microarray database and assessed the genetic association between EFNA1 polymorphisms and CRC susceptibility. We found elevated EFNA1 expression in colorectal adenocarcinoma in comparison to normal controls, suggesting EFNA1’s involvement in colorectal tumorigenesis, and its potential role as a diagnostic biomarker in CRC characterization. In our case-control study, we found that rs12904 G/A variant was significantly associated with a decreased risk of developing CRC compared with AA genotype. In the following stratified analyses, such an effect was more evident in the subgroups of smokers and tea drinkers, which could be chance findings because of the limited observations in each subgroup. Interestingly, a recent case-control study conducted by Li et al. reported that EFNA1 rs12904 was associated with gastric cancer risk in a Chinese population (Li et al., 2012), which is in agreement with our finding that rs12904 G variant allele is protective against cancer susceptibility.

rs12904 resides in the 3'UTR of EFNA1 gene, which suggests that it may have the potential to interfere with EFNA1 mRNA stability and translation through altering miRNA:mRNA interactions. Although the role of miRNAs in CRC pathogenesis has been established by the identification of miRNA expression signatures that characterize normal and tumor phenotypes as well as numerous oncogenes and tumor suppressors as miRNA targets, few studies have investigated the role of SNPs in the miRNA binding sites in the etiology of CRC (Yang et al., 2009; Landi et al., 2012). The majority of miRNAs bind to target sequences located in the 3'UTR of mRNAs by base pairing, resulting in the cleavage of target mRNA or repression of their translation. When SNPs occur in the 3'UTR, they may modulate gene expression by altering miRNA target binding capacity, ultimately leading to
differences in the disease susceptibility (Nicoloso et al., 2010). As predicted by TargetScan (http://www.targetscan.org) and miRdSNP (http://mirdsnp.ccr.buffalo.edu), rs12904 A>G polymorphism may have the potential to affect the binding of three miR-200 family members miR-200c, miR-429 and miR-200b, which have been shown to play key roles in epithelial-mesenchymal transition (EMT) (Burk et al., 2008). miR-200c, miR-429 and miR-200b were highly expressed in colorectal cell lines, and altered expression of miR-200c and miR-429 were observed in CRC tissues, and moreover, high level of miR-200c was associated with poor prognosis in CRC (Cummins et al., 2006). However, only one study examined the function of rs12904 on miR-200c binding capacity in gastric cancer cell lines to date, which showed that rs12904 G>A change resulted in altered regulation of miR-200c on luciferase expression, and EFNA1 expression was significantly higher for rs12904 AA genotype than for AG or GG genotype (Li et al., 2012). Further studies are needed to elucidate the exact functional impact of EFNA1 rs12904 polymorphism and its interaction with miRNAs especially miR-200 family in colorectal tumorigenesis.

Although the present study, to our knowledge, is the first study on EFNA1 polymorphisms and CRC risk, our findings are best considered preliminary. Because of the limited sample size, the statistical power may not be adequate to detect weak gene-disease associations and gene-environment interactions. Moreover, the lack of supporting experimental information on the functional nature of rs12904 polymorphism precluded the possibility of corroborating the genetic association results with the findings from other analyses of our study. Therefore, larger studies are warranted to further assess the association between rs12904 and CRC risk and the exact function of the SNP in the etiology of CRC.

In summary, our results suggest that EFNA1 is involved in colorectal tumorigenesis, and rs12904 A>G polymorphism in the 3' UTR of EFNA1 is associated with decreased CRC susceptibility. Larger studies and further mechanistic investigations are warranted to confirm our findings.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Project No. 81072356). The author(s) declare that they have no competing interests.

References


