Gene 678 (2018) 124–128

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Genetic variants within 17q12 are associated with the risk of cervical cancer in the Han Chinese population

Lingmin Hu[a,b,1], Meiqun Jia[b], Jing Zhou[b], Hongxia Ma[b], Guangfu Jin[b], Ni Li[c], Dong Hang[d,*], Zhibin Hu[b,e,*]

[a] The Affiliated Changzhou Maternity and Child Health Care Hospital of Nanjing Medical University, Department of Reproduction, Changzhou, Jiangsu, China
[b] Nanjing Medical University, School of Public Health, Department of Epidemiology, Nanjing, Jiangsu, China
[c] Chinese Academy of Medical Sciences, Peking Union Medical College, National Cancer Center/Cancer Hospital, Program Office for Cancer Screening in Urban China, Beijing, China
[d] Nanjing Medical University, Department of Epidemiology and Biostatistics, No.101 Longmian Ave, Jiangning District, Nanjing, China
[e] Nanjing Medical University, Prevention and Treatment, Collaborative Innovation Center For Cancer Personalized Medicine, Jiangsu Key Lab of Cancer Biomarkers, Nanjing, Jiangsu, China

ARTICLE INFO

Keywords:
Cervical cancer
Susceptibility
4q12
17q12
Variant

ABSTRACT

Chromosome 4q12 and 17q12 have been identified as two regions associated with susceptibility to cervical cancer in a genome-wide association study. To identify potential causal variants within these two regions, we conducted a case-control study including 1486 patients with cervical cancer and 1536 age-matched (± 5 years) healthy controls. Based on RegulomeDB database, 12 potentially functional variants were selected and then genotyped by using Sequenom MassARRAY. Univariate and multivariate logistic regression models were used to evaluate the associations. We found that the G allele of rs8076131 and the A allele of rs12150079 in 17q12 region were significantly associated with increased risk of cervical cancer (adjusted OR = 1.15, 95% CI = 1.02–1.30, P = 0.02 for rs8076131; adjusted OR = 1.19, 95% CI = 1.03–1.36, P = 0.02 for rs12150079). Individuals with 3–4 risk alleles of these two variants had 24% higher odds of cervical cancer than those without the risk alleles (OR = 1.24, 95% CI = 1.07–1.44, P < 0.01). The stratified analysis showed that the associations of rs8076131 and rs12150079 with cervical cancer risk were statistically significant in subgroups of older menarche age (> 16 years), more parities (≥2), nonsmokers, and having no family cancer history, but the test results for subgroup heterogeneity were not significant. The current study provides the evidence that rs8076131 and rs12150079 in 17q12 region may contribute to cervical cancer susceptibility in the Han Chinese population.

1. Introduction

Cervical cancer is a fatal disease threatening women's health worldwide. It represented the fourth most common female malignancy in 2012, with approximately 530,000 new cases annually, 85% of which occurred in developing countries (Ferlay et al., 2015). In China, the incidence and mortality of cervical cancer was 12.0 and 3.4 per 100,000 women, respectively (Chen et al., 2014). The absolute number of cervical cancer patients in China is enormous, and more importantly, an upward trend in incidence and mortality has been observed among younger women (Huang et al., 2016).

High-risk human papillomavirus (HR-HPV) infection has been recognized as the most important pathogenic factor of cervical cancer (de Sanjose et al., 2010). Although 80% of females may have a history of HPV infection before 50-year old, most cases are self-limited with clearance of virus and only a small fraction of women with HPV persistent infection develop cervical cancer (Schiffman et al., 2007), which suggests that individual genetic susceptibility play a key role in

Abbreviations: HR-HPV, high-risk human papillomavirus; GWAS, genome-wide association study; SNP, single nucleotide polymorphism; ENCODE, the Encyclopedia of DNA Elements; OR, odds ratio; MAF, minor allele frequency; LD, linkage disequilibrium; CI, confidence interval; ORMDL3, ORMDL sphingolipid biosynthesis regulator 3; NR1D1, nuclear receptor subfamily 1 group D member 1; ZPBP2, Zona pellucida binding protein 2; KRT222, keratin 222; CIN3, cervical intraepithelial neoplasia 3

* Corresponding authors at: Department of Epidemiology and Biostatistics, School of Public Health, Nanjing Medical University, No. 101 Longmian Ave, Jiangning District, Nanjing 211166, China.
E-mail addresses: hangdong@njmu.edu.cn (D. Hang), zhibin_hu@njmu.edu.cn (Z. Hu).
[1] Lingmin Hu and Meiqun Jia contributed equally to this work.

https://doi.org/10.1016/j.gene.2018.08.030
Received 26 February 2018; Received in revised form 28 June 2018; Accepted 6 August 2018
Available online 07 August 2018
0378-1119/

tumorigenesis of cervical cancer (de Freitas et al., 2012). Genetic variants within non-coding regulatory regions or protein-coding genes may change gene expression levels or protein functions, thereby affecting HPV-induced cervical pathogenesis (Das Ghosh et al., 2017).

A large number of genetic variants have been reported to be associated with cervical cancer in candidate gene studies, but many of the associations have failed to be independently replicated (Chen and Gyllensten, 2015). During the last decade, genome-wide association studies (GWAS), which scan the entire genome for common genetic variants, have been widely employed in case-control settings to identify genetic variants associated with cancer. Chromosome 4q12 and 17q12 have been identified as the two regions associated with cervical cancer risk (Shi et al., 2013). The results have been confirmed in Japanese and Polish populations (Miura et al., 2016; Lutkowska et al., 2017), suggesting that 4q12 and 17q12 may contain causal variants contributing to cervical cancer. However, the GWAS coverage is roughly one tag single nucleotide polymorphism (SNP) per 10,000 bp and may have insufficient ability to refine causal variants within these regions. Therefore, in the post-GWAS era, it is crucial to identify functional variants for understanding pathogenesis of cervical cancer and improving tailor prevention. To our knowledge, no fine-mapping study on 4q12 and 17q12 regions in Asians has been reported so far.

RegulomeDB represents a powerful bioinformatics tool that integrates high-throughput experimental data sets from the Encyclopedia of DNA Elements (ENCODE) project and other sources, as well as computational predictions to identify potentially functional variants in chromosome regions of interest (Boyle et al., 2012; ENCODE Project Consortium, 2012). It has a scoring system with categories ranging from 1 to 6 by multiple annotation information on methylation, chromatin structure, protein motifs, and transcription factor binding. The lower score indicates the stronger evidence that a functional variant could be located in a susceptible region.

Our hypothesis is that specific functional variants within 4q12 and 17q12 regions are associated with cervical cancer susceptibility. To confirm it, we screened potentially functional variants within 4q12 and 17q12 by using RegulomeDB, and genotyped them among a case-control study with 1486 cervical cancer cases and 1536 age-matched healthy controls.

2. Materials and methods

2.1. Study population

The criteria for participants' enrollment were previously described (Jiang et al., 2013). Briefly, a total of 1486 newly diagnosed and histologically confirmed cervical cancer patients were consecutively recruited from the First Affiliated Hospital of Nanjing Medical University and the Nantong Tumor Hospital in Jiangsu Province from March 2006 to December 2010. The 1536 controls were randomly chosen from a pool of > 30,000 individuals who participated in a community-based screening program for non-infectious diseases in Changzhou, Jiangsu Province during the same time. All the controls were frequency-matched to the cases by age (± 5 years). Both cases and controls had no self-reported cancer history and were genetically unrelated Han Chinese women. Each participant provided written informed consent and was face-to-face interviewed by trained interviewers at recruitment. A structured questionnaire was completed to obtain information on demographic characteristics, menstrual and reproductive history, and environmental exposures such as smoking history. Approximately 5 ml of venous blood sample was collected from each participant. This study was approved by the ethics committees of Cancer Institute and Hospital, Chinese Academy of Medical Sciences and Nanjing Medical University.

Based on the present sample size, assuming 80% statistical power, at an alpha of 0.05, we had the ability to detect an association with odds ratio (OR) of 1.23 if the minor allele frequency (MAF) was 0.50 and OR of 1.53 if the MAF was 0.05 in an additive model (NCSS-PASS 11 software).

Table 1
Demographic and clinical characteristics of 1356 cervical cancer patients and 1496 healthy controls.

Variable	Cases (%)	Controls (%)	*P*
Age, years			0.73
≤50	565 (41.7)	614 (41.0)	
> 50	791 (58.3)	882 (59.0)	
Menarche age, years[a]			< **0.01**
≤16	978 (73.6)	892 (59.6)	
> 16	351 (26.4)	604 (40.4)	
Parity[a]			
0–1	556 (41.6)	715 (48.5)	< **0.01**
≥2	781 (58.4)	759 (51.5)	
Menopausal status[a]			0.34
Premenopause	542 (40.8)	584 (39.0)	
Menopause	787 (59.2)	912 (61.0)	
Smoking history[a]			< **0.01**
Never	1279 (95.7)	1476 (98.7)	
Ever	58 (4.3)	20 (1.3)	
Family history of cancer[a]			0.16
No	1090 (81.6)	1189 (79.5)	
Yes	246 (18.4)	307 (20.5)	
Histology[a]			
Squamous cell carcinoma	1021 (91.8)		
Adenocarcinoma	68 (6.1)		
Adenosquamous carcinoma	23 (2.1)		
Stage[a]			
CIN3	8 (0.6)		
I	335 (26.7)		
II	703 (56.0)		
III	173 (13.8)		
IV	37 (2.9)		

Abbreviation: CIN3, cervical intraepithelial neoplasia 3. Bold font presents $P < 0.05$.

[a] The sum of numbers in strata was not equal to 1356 or 1496 because of missing data.

2.2. Variant selection

We only included variants with high-level functional evidence (i.e. RegulomeDB scores ranging from 1 to 2b). Based on the annotation data from UCSC database (GRCh37/hg19) and RegulomeDB database (http://www.regulomedb.org/), and the criteria of MAF > 0.05 and linkage disequilibrium (LD) < 0.8 in Han Chinese, we found 16 potentially functional variants within 4q12 and 17q12.

2.3. Genotyping and quality control

Genomic DNA was extracted from leukocyte pellets by standard phenol-chloroform extraction protocol. Genotyping was performed using Sequenom MassARRAY ® iPLEX assay according to the manufacturer's instructions without knowing the status of case and control. Four candidate variants were excluded because of the failure of primes design and the remaining 12 variants were genotyped. Each variant had a call rate > 90%. The samples with overall genotype completion rates < 90% were excluded, leaving 1356 cervical cancer cases and 1496 healthy controls in the final analysis.

Two blank controls (water) were used in each 384-well plate for quality control. Approximately 20% of samples were randomly selected for replication, showing a consistency of over 99%.

2.4. Statistical analysis

We used χ^2 test for categorical variables and Student's *t*-test for continuous variables to evaluate the difference of demographic characteristics between cases and controls. Deviation of genotype