Gene 646 (2018) 227–233

Contents lists available at ScienceDirect

Gene

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

Research paper

Functional long non-coding RNAs associated with gastric cancer susceptibility and evaluation of the epidemiological efficacy in a central Chinese population

Fujiao Duan[a,b,d], Jicheng Jiang[a,b], Chunhua Song[a,b], Peng Wang[a,b], Hua Ye[a,b], Liping Dai[c], Jianying Zhang[a,b,c,*], Kaijuan Wang[a,b,*]

[a] College of Public Health, Zhengzhou University, Zhengzhou, Henan Province, China
[b] Key Laboratory of Tumor Epidemiology of Henan Province, Zhengzhou, Henan Province, China
[c] Institute of Medical and Pharmaceutical Sciences, Zhengzhou University, Henan Province, China
[d] Medical Research Office, Affiliated Cancer Hospital of Zhengzhou University, Zhengzhou, Henan, China

ARTICLE INFO

Keywords:
Gastric cancer
Long non-coding RNA
Single nucleotide polymorphism
Susceptibility
Epidemiological evaluation

ABSTRACT

Aim: To screen and validate the gastric cancer-associated long non-coding RNAs (lncRNAs) and their functional single nucleotide polymorphisms (SNPs).

Methods: Based on case-control design, we select the differentially expressed lncRNAs by bioinformation tools and validate SNPs in lncRNAs genes in population. Attributable risk percentage (*ARP*) and population attributable risk percentage (*PARP*) were used to assess the effect of epidemiology.

Results: Four lncRNAs with SNPs (*lnc-EVX1-3:3* (rs1859168), *lnc-MACC1-1:7* (rs3815254), *lnc-AMFR-1:1* (rs4784659) and *lnc-ZNF33B-2:1* (rs579501)) were selected to be validated in population. The unconditional multiple logistic regression based on the dominant (odds ratio, *OR* = 0.64, 95%confidence interval, 95%CI: 0.47–0.86) and recessive genetic model (*OR* = 1.77, 95%CI: 1.34–2.35) showed rs1859168 was significantly associated with lower risk of gastric cancer. The dominant (*OR* = 0.42, 95%CI:0.31–0.57) and additive (*OR* = 0.52, 95%CI:0.40–0.67) genetic model revealed that rs4784659 decreased the risk of gastric cancer. Similarly, the dominant (*OR* = 0.72, 95%CI: 0.52–0.98) and additive (*OR* = 0.73, 95%CI: 0.56–0.97) genetic model showed the individuals with rs579501 had reduced risk of gastric cancer. Whereas no statistical association between rs3815254 and gastric cancer. *ARP* and *PARP* for gastric cancer associated with rs1859168 in dominant model (56.25%, 33.05%), and in recessive model (43.50% and 29.37%). For rs4784659, *ARP* and *PARP* were 138.09% and 10.39% in dominant model, 92.31% and 8.76% in additive model. For rs579501, *ARP* and *PARP* were 38.89% and 4.03% in dominant model, 36.99% and 3.88% in additive model.

Conclusion: Our findings showed rs4784659, rs579501 and rs1859168 reduced the susceptibility of gastric cancer. From epidemiological perspective, the lncRNAs with SNPs attenuate the development of gastric cancer.

1. Introduction

Gastric cancer is a common and highly lethal malignancy, being the fourth most common cancer and the second leading cause of cancer death in the world (Torre et al., 2016). It is concerned worldwide with the highest estimated mortality rates in Eastern Asian (Ferlay et al., 2015). According to the updated cancer statistics, the incidence and mortality of gastric cancer have been both ranking second in the number of the cases and deaths of various cancers across China (Jing et al., 2012; Chen et al., 2016), with both the early diagnosis rate and 5-year survival rate in advanced gastric cancer < 20% (Zhu et al., 2013). Therefore, new biomarkers are needed to discriminate the high-risk patients with gastric cancer and consequently improve personalized cancer care.

Currently, substantial studies have demonstrated the roles of dysregulated functional long non-coding RNAs (lncRNAs) in human cancers (Mitra et al., 2012; Chi et al., 2017). LncRNAs are long RNAs (longer than 200 nucleotides) with limited protein coding potential

Abbreviations: lncRNA, long noncoding RNAs; SNP, single nucleotide polymorphisms; *ARP*, Attributable risk percentage; *PARP*, population attributable risk percentage; OR, odds ratio; 95%CI, 95%confidence interval; HWE, Hardy–Weinberg equilibrium; MDR, multifactor dimensionality reductions; PCR-RFLP, Polymerase chain reaction restriction fragment length polymorphism

* Corresponding authors at: College of Public Health, Zhengzhou University, Zhengzhou, Henan Province, China.
E-mail addresses: jianyingzhang@hotmail.com (J. Zhang), kjwang@163.com (K. Wang).

https://doi.org/10.1016/j.gene.2017.12.063
Received 16 November 2017; Received in revised form 22 December 2017; Accepted 29 December 2017
Available online 03 January 2018
0378-1119/

(Yoshimoto et al., 2016; S. Zhang et al., 2016) and play a vital role in the process of carcinogenesis that the expression and function of several key lncRNAs are affected by single-nucleotide polymorphisms (SNPs) (Tao et al., 2015; Bayram et al., 2015; Pan et al., 2016). Deregulation of lncRNAs contributes to the progression of cancer through modulating several cellular processes which are crucial to oncogenesis, including cell proliferation, migration and apoptosis (Li et al., 2013; Ricciuti et al., 2016). LncRNAs are expected to be the biomarkers for diagnosis and prognosis of gastric cancer (Fan et al., 2016). But the underlying mechanisms remain to be clarified.

The attributable risk (*AR*) is the difference between the risk encountered by individuals exposed to a particular factor and the risk encountered by individuals who are not exposed to it. This is the opposite to avoidable risk, and it measures the absolute effect of a cause. Attributable risk percentage (*ARP*) is a calculation that can be derived from attributable risk (Lacourt et al., 2014). It gives the portion of cases avoidable or attributable to this exposure in relation to all cases. *PAR* defined as the proportion of disease in the population that is attributable to a given risk factor (or set of risk factors), used to estimate the public health impact of that factor (Wu et al., 2015). The population attributable risk percentage (*PARP*) or population attributable fraction (*PAF*), which is defined as the reduction in incidence that would be achieved if the population had been entirely unexposed, compared with its current (actual) exposure pattern (Jiang et al., 2017).

The purpose of our study was to evaluate the relationship between the candidate lncRNA SNPs and gastric cancer susceptibility. *ARP* and *PARP* as epidemiological indexes were used to assess the effect of the gastric cancer–associated lncRNAs on the susceptibility of gastric cancer in central Chinese population to better understand the public health risk.

2. Materials and methods

2.1. Ethics statement

This study was received approval from the ethics committee of Zhengzhou University. Written informed consents were provided by all participants.

2.2. Study subjects

All subjects were genetically unrelated Chinese residents in the central region of China, consisting of 470 gastric cancer patients and 470 healthy controls. The patients were consecutively recruited from the First Affiliated Hospital of Zhengzhou University during December 2012 to December 2015. Eligible patients were histopathologically confirmed with primary gastric cancer before undergoing any treatment. Healthy controls were frequency-matched to cases on age (± 3 years) and randomly selected from unrelated community residents attending routine health checkups in the same hospital.

2.3. LncRNAs selection

Search in the GEO database and download the microarray data of gastric cancer (GSE50710, GSE53137, GSE58828) based on Chinese population. We managed to integrate several databases (Ensembl, refseq, lncRNAdb.v2.0, GENCODE.v24, NONCODE2016, LNCIpedia.v3.1, CCDS.v18), remove the redundancy and re-annotate the high-throughput microarrays by local BLASTN algorithm. Eventually, we selected the differentially expressed lncRNAs between gastric cancers and the normal by bioinformation methods and predicted the potential functions of the candidate lncRNAs, including *lnc-EVX1-3:3* (rs1859168), *lnc-MACC1-1:7* (rs3815254), *lnc-AMFR-1:1* (rs4784659) and *lnc-ZNF33B-2:1* (rs579501).

2.4. DNA extraction and genotyping

We genotyped four SNPs (*lnc-EVX1-3:3* rs1859168, *lnc-MACC1-1:7* rs3815254, *lnc-AMFR-1:1* rs4784659 and *lnc-ZNF33B-2:1* rs579501) from genomic DNA isolated from blood samples using the Qiagen DNA blood Mini Kit (Tiangen Biotech (Beijing) Co., Ltd.) according to the manufacture's instructions. Polymerase chain reaction restriction fragment length polymorphism (PCR-RFLP) was used to determine the genotypes of SNPs in lncRNAs genes. The primers used for PCR amplification were 5′-ACGACTGGGTCCCTCCTCA-3′ (forward) and 5′-TGGTGGCTACAACTCAATAC-3′ (reverse) for rs1859168, 5′-ATTGA TGAATACACGAGCAG-3′ (forward) and 5′-AAGGCAGAGTAGTAGAT GGT-3′ (reverse) for rs3815254, 5′- TGCTTGGTGCCAGGTCTA -3′ (forward) and 5′- GCCTTTCAGGGCTTGTCC-3′ (reverse) for rs4784659, 5′-CCTCAGTTTGGGAATTCAGTG -3′ (forward) and 5′- AAGCAAGCAG GCGAAGAA -3′ (reverse) for rs579501. PCR was performed in a final volume of 15 μl, which contained 7.5 μl 2 × Tap PCR MasterMix, 0.3 μl each primer, 1.0 μl DNA, 5.9 μl deionized water. The optimized reaction system was carried out for 8–12 h under the optimum temperature of the corresponding restriction enzyme. In addition, about 10% of samples were randomly selected from both cases and controls for duplicable genotyping and the results were 100% concordant. Direct sequencing (Beijing Genomics Institute (BGI) Sequencing, Beijing, China) was utilized for genotype confirmation, and the result was 100% concordant.

2.5. Epidemiological effect estimation

ARP is a calculation that can be derived from *AR*, it gives the portion of cases attributable (and avoidable) to this exposure in relation to all cases, and *PARP* measures estimate the impact of a causal factor in population that would be reduced if risk factor was removed.

The relative risk (*RR*) among the exposed portion of the population, divided by the *RR* in the unexposed portion of the population, gives a relative measure (*RR*) of the effect of a given exposure and approximates the *RR* if the occurrences are rare (*RR* ≈ *OR*) (Yao and Shi, 2003; Basu and Landis, 1995; Shi, 2001).

$$ARP = \mathrm{OR} - 1/\mathrm{OR} \times 100\%$$

$$PARP = P_e(\mathrm{OR}-1)/[P_e(\mathrm{OR}-1) + 1] \times 100\%$$

P_e was the mutation proportion of the control group.

2.6. Statistical analysis

The statistical analyses were conducted by SAS 9.1 (SAS Institute Inc., Cary, North Carolina, USA). The Hardy–Weinberg equilibrium (HWE) was evaluated using the goodness-of-fit Chi-square (χ^2) test. To evaluate the associations between the lncRNA SNPs and gastric cancers risk, multivariate logistic regression was used to compute odds ratios (ORs) and corresponding 95%confidence intervals (CIs), adjusted for age, gender, smoking, drinking and family history. Multifactor dimensionality reductions 2.0 (MDR, 2.0) (Moore et al., 2008) were applied to identify the possible of gene–gene and gene–environment interactions associated with gastric cancers risk, and determined the best model of interaction. The relationship between the number of SNP mutation and the risk of gastric cancer was analyzed by trend χ^2 test. *ARP* and *PARP* as epidemiological indexes were calculated according to the formula mentioned above to assess the effect of the gastric cancer–associated lncRNAs.

All *P* values were two-sided, and the data were considered to be statistically significant when $P < 0.05$.