Gene 540 (2014) 32–36

Contents lists available at ScienceDirect

Gene

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

A new design without control population for identification of gastric cancer-related allele combinations based on interaction of genes

Liu Hui [a,*], Gai Liping [b], Yu Jian [a], Yu Laisui [a]

[a] College of Medical Laboratory, Dalian Medical University, China
[b] Dept. of Physics, Dalian Medical University, China

ARTICLE INFO

Article history:
Accepted 17 February 2014
Available online 26 February 2014

Keywords:
Heredity
Gastric cancer
Genetic analysis

ABSTRACT

Objective: The purpose of this study was to develop a novel approach without control population to examine the relationship between the presence of specific allele combinations at different loci with the onset of gastric cancer.
Methods: DNA samples were collected from patients with gastric cancer. Alleles from short tandem repeat loci were determined using the STR Profiler Plus PCR amplification kit (15 STR loci). The observed and expected frequencies of specific allele combinations were calculated; statistically significant allele combinations were identified by comparing the observed frequency with the expected frequency. The age at disease onset of patients carrying a specific allele combination was further analysed; allele combinations related to the gastric cancer were effectively identified from the large number of possible allele combinations by cross-validation of the 2 sets of analytical results.
Results: A total of 2209 pairwise combinations were obtained by computer counting, of which 11 pairs of genes showed significant differences between the observed and expected frequencies ($p < 0.05$). The p value for the cross-validation was also less than 0.05 for 2 pair of alleles (D8S1179-16 and D5S818-13; D2S1338-23 and D6S1043-11).
Conclusion: Gastric cancer onset may be associated with these allele combinations. The new methodology without control group will enable additional discoveries pertaining to the relationship between specific allele combinations at different loci and the onset of complex diseases.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Gastric cancer is a typical complex disorder, to which a hitherto unknown number of genes contribute by interacting with each other and the environment (Nobili et al., 2011; Resende et al., 2011; Yamashita et al., 2011). The identity of the genes involved in gastric cancer has remained elusive in spite of the rapid pace of developments in molecular technology and the increase in genome sequence information (Ha et al., 2011; Shi et al., 2011; Yoshida et al., 2010). Recent research has largely excluded the possibility that genes with major causative effects exist, however, genetic liability almost certainly results from the combined effects of multiple susceptibility and resistance loci; most studies to date have been under-equipped to detect such effects (Gong et al., 2013; Li et al., 2013; Torres-Jasso et al., 2013).

The differences between the frequency distributions of the gene in the cancer and control groups are usually observed in studies of cancer-related allele. However, it is difficult to obtain the healthy individuals without any gastric cancer-related gene as a control group. Lack of high quality control group may be partly attributed to that inconsistent findings and weak signals of associated-cancer gene studies. Therefore, we developed a new study design without control population for the identification of gastric cancer-related genes.

Alleles on the same or different loci will be randomly distributed. If an observed combination of alleles is not randomly distributed within a cancer group, that combination can be determined as the interaction of alleles and to be associated with the cancer. Thus, specific allele combinations could be identified without the control population.

Typically, studies of genetic combined effects proceed by first calculating the Hardy–Weinberg equilibrium (HWE) for a gene locus and then looking for a correlation between specific allele combination and the occurrence of the disease of interest. However, the number of gene combinations that are correlated with disease onset is not necessarily limited to the allele combinations of one specific locus. Unfortunately, there is very little information for discovering additional relationships between specific allele combinations at different loci and cancer onset.

Microsatellite or short tandem repeats (STR) consist of tandem repeats of DNA units ranging from two to six nucleotides (Edwards et al., 1991, 1992). Because of the variable number of highly polymorphic tandem repeats in humans, the investigation of STR is a powerful

Abbreviations: CODIS, the US National Combined DNA Indexing System; HWE, Hardy–Weinberg equilibrium; STR, short tandem repeats.

* Corresponding author at: College of Medical Laboratory, Dalian Medical University, Dalian 116044, China.
E-mail addresses: liuhui60@sina.com, immunology@dlmedu.edu.cn (L. Hui).

http://dx.doi.org/10.1016/j.gene.2014.02.033
0378-1119/© 2014 Elsevier B.V. All rights reserved.

tool in human genetics (Moretti et al., 2001). We selected STR loci listed in the US National Combined DNA Indexing System (CODIS) database, which forensic laboratories currently use, for analysis as reliable genotyping assays are commercially available. We hypothesized that there is an extensive range of gastric cancer-associated genes in the genome; therefore, we randomly selected 15 CODIS STR loci to screen for gastric cancer-related allele combinations.

2. Materials and methods

2.1. Subjects

A total of 75 Chinese patients with unselected gastric cancer (as evidenced by surgical intervention and pathologic findings), including 47 males and 28 females, were recruited from the Dalian Municipal Friendship Hospital, Dalian Third Hospital and Dalian Fifth Hospital in China. All the subjects were unrelated ethnic Han Chinese individuals. None of the patients had received preoperative radiation therapy or chemotherapy.

The age of onset should be recorded. The mean age at first onset of gastric cancer based on the medical records was 58.8 ± 11.7 years. The patients, whose medical records did not clearly record the age at first onset of gastric cancer, were excluded from the study.

Fifty unselected liver cancer patients (33 males and 17 females) who underwent hepatectomy at our university hospital were also enrolled for comparative observation. The liver cancer patients were from the same geographic region as gastric cancer in this study. The mean age at the first onset of liver cancer was 59.3 ± 11.7 years based on the medical records.

The experiments were conducted in accordance with the Declaration of Helsinki. The blood samples taken were part of the usual care of the subjects, rather than for research purposes alone. The Institutional Ethics Committee of Dalian Medical University approved the study and waived the need for written informed consent from the participants due to the observational nature of the study.

2.2. Sample preparation and genotyping

DNA was extracted from 3 ml of peripheral blood using the Chelex 100 procedure (Moretti et al., 2001). The 15 STR loci of D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317, D16S539, D6S1043, D12S391, CSF1PO, D7S820, D2S1338, and D19S433 were co-amplified using AmpF/STR Sinofiler™ PCR Amplification Kit (Applied Biosystems, USA) according to the manufacturer's recommendations. A 9600 Perkin Elmer thermal cycler was used for amplification. Amplification products were heat denatured at 95 °C for 5 min and chilled for 5 min in an ice-water bath prior to capillary electrophoresis on an ABI 310 automated sequencer (Applied Biosystems). The Genescan Analysis 2.1 software (Applied Biosystems) was used to determine fragment sizes. Alleles were identified by comparing the amplified fragments with the allelic ladders included in the reagent set. The alleles were labelled according to international nomenclature using the Genotyper Software package (Perkin Elmer).

2.3. Allele frequencies at a locus

In heterozygotes, two STR alleles are present at a single STR locus, whereas both STR alleles in homozygotes are identical at a single locus.

Allele frequency should be calculated as follows:

$$\text{Allele frequency at a locus} = [\text{number of alleles}/(\text{number of subjects} * 2)] * 100\%,$$

where number of subjects ∗ 2 is equal to the number of chromosomes.

A Chi-square test was used to determine whether the distribution of STR genotypes followed the Hardy–Weinberg equilibrium.

2.4. Count of gene combinations at different loci

There are many genes on different chromosomes, with many possible combinations of those genes, so combinations that are correlated with disease presence are unlikely to be observed manually. Therefore, we used computer software to test for the random distribution of pairwise combinations of specific alleles at different loci.

Alleles at 15 STR loci were recorded in Excel to establish a database. Using Microsoft macro-programming techniques, statistical sorting was conducted on genes distributed in the database. Possible pairwise locus combinations were identified using a multi-layer recycling method. Progressive scans were conducted on all samples to quantify all pairwise allele combinations. Information, including the combination results and the frequency of allele combinations at different loci, was sequentially stored in a newly generated pairwise locus gene combination database.

2.5. Calculation of the observed and expected gene combination frequencies of different loci

Theoretically, alleles on different chromosomes are randomly distributed. If an observed combination of alleles is not randomly distributed within a disease group, that combination can be determined to be associated with the disease.

Alleles at loci L1 and L2 were set as a, b and c, d, respectively. If the observed gene combination was a and c (ac), then there would be four gene combination models, as shown in Table 1.

For any combination model, the frequency of the ac combination should be the number of ac combinations divided by the number of people observed with that combination. Because there are four ac combination models the overall frequency of ac combinations should be the mean of the frequencies of the four models. Therefore, the observed frequency (FO) of a specific gene combination should be:

$$\mathrm{FO} = [(\mathrm{NO}/\mathrm{n})/4] * 100\%,$$

where NO is the total number of combinations of a specific gene observed and n is the number of people observed.

The expected frequency (FE) of the ac combination should be:

$$\mathrm{FE} = \mathrm{fa} * \mathrm{fc},$$

where fa and fc are the gene frequencies of a and c, respectively.

2.6. Comparison of the observed and expected frequencies of gene combinations

Statistical analysis was not conducted for samples in which the number of gene combinations was less than 5. Statistical analyses and other comparison were conducted only on effective combinations (i.e., where the number of gene combinations is ≥5). In the disease group, the observed and expected frequencies of gene combinations were compared using the binomial test.

Table 1
Gene combination models for alleles a and c at different loci.

Loci 1	Loci 2	Model name
a (father), b (mother)	c (father), d (mother)	
a, *	c, *	Model 1
a, *	*, c	Model 2
*, a	c, *	Model 3
*, a	*, c	Model 4

Download English Version:

https://daneshyari.com/en/article/2816535

Download Persian Version:

https://daneshyari.com/article/2816535

Daneshyari.com