Gene 579 (2016) 1–7

Contents lists available at ScienceDirect

Gene

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

Research paper

Screening of intracerebral hemorrhage associated allele combinations at different loci using a novel association analysis

Liping Gai, Cui Sun, Weijian Yu, Hui Liu *

College of Medical Laboratory, Dalian Medical University, China

ARTICLE INFO

Article history:
Received 7 May 2015
Received in revised form 28 October 2015
Accepted 15 December 2015
Available online 24 December 2015

Keywords:
Heredity
Intracerebral haemorrhage (ICH)
Genetic analysis

ABSTRACT

Background: Genetic research has progressed along with scientific and technological developments. However, it is difficult to identify frequency differences in the allele combination at cross-loci.
Objective: The purpose of this study was to examine the relationship between the presence of specific allele combinations of short tandem repeat (STR) loci and the onset of intracerebral hemorrhage (ICH) using a novel methodology.
Methods: DNA samples were collected from patients with ICH, who were adult population. There were a total of 51 Chinese patients (102 chromosomes), comprising 30 males and 21 females. Alleles from short tandem repeat (STR) loci were determined using the STR Profiler Plus PCR amplification kit (15 STR loci). Statistically significant differences between observed and expected frequencies of allele combinations were identified. To further determine allele combinations related to the disease, analyses of patient age at disease onset for those carrying a specific allele combination were conducted. Finally, cross-validation of the two sets of analytical results was carried out.
Results: A total of 1550 pairwise combinations were obtained by computer counting, of which eight pairs of alleles showed significant differences between the observed and expected frequencies ($p < 0.05$, from 0.006 to 0.042). The p value for the cross-validation analysis was less than 0.05 for two pairs of alleles (D13S317-11 and vWA-17, $p = 0.021$; D7S820-13 and D2S1338-18, $p = 0.023$).
Conclusions: The study identified each population had a unique gene distribution and that distribution followed certain rules. ICH onset may be associated with this allele combinations (D13S317-11 and vWA-17; D7S820-13 and D2S1338-18). The new methodology used in this study could enable additional discoveries pertaining to the relationship between specific allele combinations at different loci and the onset of complex diseases.

1. Introduction

Intra cerebral hemorrhage (ICH) is a complex disease. It refers to the non-traumatic brain parenchyma hemorrhage (Wang et al., 2012). Once it happened, this will cause the patients' brain tissue injured. As a result, the patients may have action obstacle, language disorder, disorders of consciousness, memory disorders, and sphincter disturbances, etc. The morbidity and mortality rate of ICH is very high. It is seriously affected to survival and quality of life of the quinquagenarian.

ICH is mostly caused by cerebrovascular diseases and hypertension. Both cerebrovascular diseases and hypertension have important genetic basis. Usually, they are disease of multifactorial inheritance (Rosell and Vilalta, 2011). The genetic basis of ICH is made up of numerous micro effect genes. So, it is almost impossible to explain the genetic mechanism of ICH through a few strong effect genes. We believe that the micro effect genes related ICH widely exists in the genome. According to this assumption, we randomly selected widely distributed nucleotides sequences in the genome as a point, to explore the micro effect genes related ICH in genome, and to amplify the effects of micro effects gene using the method of allele combinations. So that we can evaluate the genetic predisposition of an individual intracerebral hemorrhage or further explore related genes to expound on genetic mechanism of ICH.

Typically, studies of allele combinations begin by calculating the Hardy–Weinberg equilibrium for a gene locus, and then look for a correlation between specific alleles 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 at one specific locus. Future studies will undoubtedly discover additional relationships between specific allele combinations at different loci and disease onset (Bugiani et al., 2010; Jun and Yong-Xi, 2015; Chen et al., 2010; Wu et al., 2009).

Abbreviation list: CT, computed tomography; CODIS, combined DNA index system; ICH, intracerebral hemorrhage; MRI, magnetic resonance imaging; PCR, polymerase chain reaction; STR, short tandem repeat; T1WI, T1 weighted imaging in magnetic resonance imaging; T2WI, T2 weighted imaging in magnetic resonance imaging.

* Corresponding author at: College of Medical Laboratory, Dalian Medical University, Dalian 116044, PR China.
E-mail address: liuhui60@sina.com (H. Liu).

http://dx.doi.org/10.1016/j.gene.2015.12.031
0378-1119/

2. Materials and methods

2.1. Subjects

This research is to determine whether the genes associated with disease by observing a certain gene combination. Known STRs are randomly distributed in the genome. If the combination of alleles on different STR loci is not a random distribution, then it is concluded that the no random distribution of STR combination is caused due to illness. The STR adopted in this study is often used in forensic STR, it is a random distribution known in the crowd. So it is reasonable to omit the specimen collection and detection to the control population.

This study was an observational, retrospective case-control study. These are continuous collection of cases. 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.

Chinese patients were recruited from the Dalian Third Hospital and the Affiliated Hospital of Dalian University in China between May 2008 and May 2010. The patients with ICH were recruited from the Dalian Third Hospital and the Affiliated Hospital of Dalian University in China between May 2008 and May 2010. ICH was diagnosed according to World Health Organization (WHO) criteria (1989) Recommendations on stroke prevention, diagnosis, and therapy. Combined with brain imaging (computed tomography and/or magnetic resonance imaging), ICH was defined as a sudden, focal, neurological deficit with ICH seen on brain imaging. The diagnosis is made jointly by neurologists and imaging experts. Exclusion criteria were as follow: (1) disturbances of blood coagulation, e.g., thrombocytopenia, hepatitis; (2) traumatic intracranial hemorrhage; (3) intracranial or general infection; (4) complicated with serious heart, liver, renal, or lung disease or functional failure; (5) intracranial aneurysm or arteriovenous malformation complicated with hemorrhage; (6) the patients whose medical records did not clearly record the age at first onset of ICH.

A total of 51 Chinese patients (102 chromosomes) with ICH, comprising 30 males and 21 females, were recruited in this study. The mean age of patients at first onset of ICH, as indicated by the medical records, was 61.9 ± 15.8 years.

The short tandem repeat (STR) adopted in this study is often used in forensic STR, it is a random distribution known in the crowd. So it is reasonable to use expected frequency of the combination from two alleles as the control and omit the specimen collection and detection to the control population.

2.2. Sample preparation and genotyping

Microsatellite or short tandem repeats (STRs) consist of tandem repeated DNA units ranging from 2 to 6 nucleotides (Liping et al., 2015; Edwards et al., 1991; Edwards et al., 1992). Because of the variable number of highly polymorphic tandem repeats in humans, examination of STRs is a powerful tool in human genetics. Forensic laboratories currently use the following STR loci listed in the US national Combined DNA Indexing System (CODIS) database.

DNA was extracted from 3 ml of peripheral blood using the Chelex 100 procedure). 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. They are simplified as D3, VWA, FGA, D8, D21, D18, D5, D13, D16, D6, D12, CS, D7, D2, and D19. 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 Bio systems). The Genescan Analysis 2.1 software (Applied Bio systems) 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 Genotype Software package (Perkin Elmer).

We selected these loci for this study because reliable genotyping assays for these STR loci are commercially available. Genotyping was carried out at a half-atomised condition in accordance with international standards and the results were assessed using standard software; therefore, error due to artificial performance and judgment was greatly reduced (Buse et al., 2003; Moretti et al., 2001; Budowle and Sprecher, 2001). The current study aimed to examine the relationship between the presence of specific CODIS–STR allele combinations at different loci and the onset of ICH, using a novel methodology.

2.3. Allele frequencies at a locus

In heterozygote, two STR alleles (a, b) are present at a single STR locus, whereas one STR allele (a, a) at a single locus is identical as homozygote. In the sample library we can see the heterozygote and homozygote samples. See Fig.1.

Allele frequency should be calculated as follows:

Allele frequency at a locus = [number of alleles / (number of subjects × 2)] × 100%, where number of subjects × 2 is equal to the number of chromosomes.

Allele frequencies were compared with a control population that consisted of individuals from the same geographic region as the patients in this study (n = 231) (Huang et al., 2010). The constituent ratio of allele frequencies in the ICH group was compared with that in the control population using a chi-square test. If a locus was determined to be significant, then the significant differences in allele frequencies between the ICH group and the control group were examined using a chi-square test.

2.4. Count and frequencies of allele 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. In the current study we used computer software to test for the random distribution of pairwise specific allele combinations at different loci. See Fig.2.

The combination rules of loci are:

$$C(n, m) = \frac{n!}{m!(n-m)!}$$

This means to select m loci from the total loci numbers n (n = 15, m = 2), e.g. combination loci (D8S1179, D21S11).

To any combination loci, the combination rules of alleles are:

$$P(k, r) = \frac{k!}{(k-r)!}$$

This means to select r alleles from the total alleles k (k is decided by the source samples, r = 2) e.g. the combination alleles (8, 10) at the combination loci (D8S1179, D21S11). Because each locus has two different chromosomes, one each from the father and mother, so the rules will be:

$$P(k, r)P(2, 1)P(2, 1) = 4\frac{k!}{(k-r)!}$$

To all the loci combinations and allele combinations, the system used the mathematical methods for statistical analyses to assess the