Cancer Epidemiology 50 (2017) 39–45

Contents lists available at ScienceDirect

Cancer Epidemiology

The International Journal of Cancer Epidemiology, Detection, and Prevention

journal homepage: www.canceepidemiology.net

Genetic association of telomere length with hepatocellular carcinoma risk: A Mendelian randomization analysis

Yue Cheng[a,1], Chengxiao Yu[a,1], Mingtao Huang[a], Fangzhi Du[a], Ci Song[a], Zijian Ma[a], Xiangjun Zhai[b], Yuan Yang[c], Jibin Liu[d], Jin-Xin Bei[e], Weihua Jia[e], Guangfu Jin[a], Shengping Li[e], Weiping Zhou[c], Jianjun Liu[f], Juncheng Dai[a,*], Zhibin Hu[a,*]

[a] Department of Epidemiology and Biostatistics, Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center For Cancer Personalized Medicine, School of Public Health, Nanjing Medical University, Nanjing 211166, China
[b] Department of Infectious Diseases, Jiangsu Province Center for Disease Prevention and Control, Nanjing, China
[c] Eastern Hepatobiliary Surgery Hospital, Second Military Medical University, Shanghai, China
[d] Tumor Biobank, Nantong Tumor Hospital, Nantong, China
[e] Department of Hepatobiliary Oncology, State Key Laboratory of Oncology in South China, Sun Yat-Sen University Cancer Center, Guangzhou, China
[f] Department of Human Genetics, Genome Institute of Singapore, A*STAR, Singapore, Singapore

ARTICLE INFO

Article history:
Received 3 June 2017
Received in revised form 21 July 2017
Accepted 24 July 2017
Available online 7 August 2017

Keywords:
Telomere length
HBV-related hepatocellular carcinoma
Mendelian randomization analysis

ABSTRACT

Background: Observational studies show an association between telomere length and Hepatocellular carcinoma (HCC) risk, but the relationship is controversial. Particularly, it remains unclear whether the association is due to confounding or biases inherent in conventional epidemiological studies. Here, we applied Mendelian randomization approach to evaluate whether telomere length is causally associated with HCC risk.

Methods: Individual-level data were from HBV-related HCC Genome-wide association studies (1,538 HBV positive HCC patients and 1,465 HBV positive controls). Genetic risk score, as proxy for actual measured telomere length, derived from nine telomere length-associated genetic variants was used to evaluate the effect of telomere length on HCC risk.

Results: We observed a significant risk signal between genetically increased telomere length and HBV-related HCC risk (OR = 2.09, 95% CI 1.32–3.31, $P = 0.002$). Furthermore, a U-shaped curve was fitted by the restricted cubic spline curve, which indicated that either short or long telomere length would increase HCC risk ($P = 0.0022$ for non-linearity test). Subgroup analysis did not reveal significant heterogeneity between different age, gender, smoking status and drinking status groups.

Conclusions: Our results indicated that a genetic background that favors longer or shorter telomere length may increase HBV-related HCC risk—a U-shaped association.

1. Introduction

Primary liver cancer is a major health problem around the world, which is the second leading cause of cancer death in men worldwide [1]. According to the latest statistics of China, it was estimated that 466,100 new liver cancer cases and 422,100 liver cancer related deaths had occurred in 2015 [2]. Hepatocellular carcinoma (HCC) accounts for 85%–90% of the primary liver cancer. Studies proved that the etiology of HCC is complicated, HBV, HCV, environmental and genetics risk factors together involved in the process [1]. Chronic hepatitis B virus (HBV) infection is a major reason of HCC, however, the contributing factors of development from chronic hepatitis B to HCC are not fully understood.

Telomeres are short repetitive AGGGTT sequences that protect the ends of chromosomes from degradation and shorten during each round of cell division [3,4]. Because of the "end replication problem", telomeres would shorten after each DNA replication and each cell division in normal cells. Excessive telomere shortening may lead to cellular senescence, genetic instability and apoptosis

Abbreviations: HCC, hepatocellular carcinoma; HBV, hepatitis B virus; HCV, hepatitis C virus; TL, telomere length; MR, Mendelian randomization; GRS, genetic risk score; GWAS, Genome-wide association studies; MAF, minor allele frequency; HWE, Hardy-Weinberg equilibrium; OR, odds ratios; IVW, inverse-variance weighting; CI, confidence interval; LD, linkage disequilibrium.

* Corresponding authors.
E-mail addresses: djc@njmu.edu.cn (J. Dai), zhibin_hu@njmu.edu.cn (Z. Hu).

[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.canep.2017.07.011
1877-7821/

[5]. Excessively long telomere length (TL) and upregulated telomerase activity may result in immortalized cells with unlimited potential for growth and proliferation [6,7].

Recent studies found longer TL increased the risk of hepatitis B virus-related HCC [8,9]. On the other hand, Plentz's data suggested that hepatocyte telomere shortening could increase the risk of HCC development [10]. Accumulating evidence revealed the association between TL and HCC risk, however, whether the observed association reflected a causal effect was still confused. For example, the relative TL was measured after the HCC development, which may result in "reverse causation bias", because TL measured after HCC development may have been altered by HCC and thus cannot accurately inform the association between TL and subsequent HCC risk. Moreover, the association may be confounded by insufficient adjustment for exposures that affect both HCC risk and telomere length, e.g. smoking [11–13].

Mendelian randomization (MR) is an alternative way to estimate the causal effect, which is a form of instrumental variable analysis whereby selected genetic variants related to a specific exposure of interest are utilized to statistically evaluate a causal hypothesis between the exposure and an outcome [14,15]. As genotype is presumed to be randomly allocated during the process of meiosis and genotypes precede phenotype, MR addresses the issue of reverse causality [16]. Confounding bias can be avoided by selecting genetic factors that are exclusively associated with telomere length in MR.

To clarify the causal effects for HCC risk, we used leukocyte telomere length-associated genetic variants as proxy for actual measured TL to investigate the association between telomere length and HBV-related HCC risk using a case-control study (1,538 HBV-positive HCC patients and 1,465 HBV-positive controls). Specifically, nine telomere length-related variants were considered and genetic risk score (GRS) was used to estimate the causal effect for HBV-related HCC risk.

2. Materials and methods

2.1. Study subjects

Participants and data were derived from our previous HBV-related HCC Genome-wide association studies (GWAS) (1,538 HBV positive HCC patients and 1,465 HBV positive controls). Briefly, all participants were Han Chinese from East and South China. HCC diagnosis was verified by a pathological examination and/or α-fetoprotein elevation (>400 ng/ml) combined with imaging examination, while HBV-positive controls were screened for being HBV persistent carriers [17]. Demographic characteristics of the participants were shown in Supplementary Table 1. All the subjects were provided informed consent at recruitment and the study was approved by the institutional review board of Nanjing Medical University.

2.2. Selection of TL-associated genetic variants

TL-related variants were systematically selected according to the following criteria: (i) the reported association reaching genome-wide association significance level (i.e. $P \leq 5 \times 10^{-8}$); (ii) the variants having a minor allele frequency (MAF) of at least 1% in Chinese population. Information of MAF was obtained from the 1000 Genomes Project CHB subjects (Phase 3). Imputed SNPs had IMPUTE2 INFO scores >0.8, which indicated that the imputation had a high degree of accuracy. (iii) *P*-value from an exact test of Hardy-Weinberg equilibrium (HWE) in controls $> 1 \times 10^{-5}$. Finally, nine SNPs associated with leukocyte telomere length (rs10936599, rs2736100, rs7675998, rs4387287, rs8105767, rs755017, rs11125529, rs3027234, rs412658) were included for further evaluation [18–20].

2.3. Instrumental variable for telomere length

To predict leukocyte telomere length with these 9 telomere length associated SNPs, GRSs were calculated using the following formula:

$$GRS_i = \sum_{j=1}^{9} w_j x_{ij}$$

where x_{ij} is number of long telomere length related alleles for the *j*th SNP of the *i*th subject (=0, 1 or 2) and w_j is the weight or coefficient for the *j*th SNP, originating from the previously published telomere length association articles [18–20]. The β-estimates were all scaled to kb of telomere length per long allele (Supplementary Table 2). The GRS could be used as an instrumental variable to represent the weighed number of telomere length increasing alleles, and then predicted the leukocyte telomere length of each individual in the Mendelian randomization analysis. Several conditions are necessary for these Mendelian randomization effect estimates to have a causal interpretation: (1) the TL-associated variants need to be associated with TL in leukocytes, (2) the TL-associated variants are not associated with other factors (confounders) that influence both TL and HCC risk and (3) the TL-associated SNPs only affect HCC risk through their effects on TL, i.e. there are no alternative causal pathways by which the SNPs influence HCC risk [21].

2.4. Statistical analysis

Linkage disequilibrium analysis was conducted among the 1000 Genomes Project CHB subjects (n = 103) by plink1.9.

Aggregate test was applied to compare a null model having only gender, age, smoking status and drinking status with an expanded model that included all nine telomere length-associated variants. Chi-square test was used to assess statistical significance of aggregations.

The causal effects were estimated with our SNP collections using GRS method, that is, individual-level data were aggregated into a univariate score (GRS) to calculate odds ratios (OR) and 95% confidence intervals (95%CI) with logistic regression models adjusted with age, gender, smoking status and alcohol-drinking status, as well as the first principal component [22].

To determine if the analysis was robust to the choice of weights used in the genetic risk score analysis, we evaluated the casual association of TL and HCC risk using an unweighted GRS. Besides the GRS approach, we also used the inverse-variance weighting method (IVW) to estimate the effect of telomere length on risk of HBV-related HCC. The inverse-variance weighting method is a summary statistic based method, and causal estimation was calculated from each instrumental variant using summarized data [22]. The method is described in greater details by Burgess et al. [23]. As to the LD, the weaker SNP was removed from GRS and summarized data for IVW to assess the effect again.

To assess whether the MR assumptions are valid, especially the second and third assumptions, we used the "gtx" package (v0.0.8) in R software to conduct tests of heterogeneity.

In addition, to further examine the relationship between TL and HBV-related HCC, we categorized TL-associated GRS into five groups based on its quintile distribution among all participants. Unconditional logistic regression was used to estimate the odds ratio (OR) and 95% confidence interval (CI) for HBV-related HCC risk in each GRS quintile, with an adjustment for age, gender, smoking and drinking status. At the same time, we performed