Association between variant alleles of major histocompatibility complex class II regulatory genes and nasopharyngeal carcinoma susceptibility

Ping Zhou[a,b], Sha Liu[b], Nan-Nan Ji[c], Shuang Zhang[b], Peng Wang[b], Bing Lin[b], Ping Yang[b], Xian-Tao Lin[b], Yi-Zheng Cai[d], Zi-Ming Wang[e], Han Zhou[e], Shi-Yao Sun[e] and Xin-Bao Hao[a,f]

Major histocompatibility complex (*MHC*) class II regulatory genes play a paramount role in immune response that can exert a predominant influence on clinical outcome of Epstein–Barr virus infection consistently assumed as the main pathogenetic factor for nasopharyngeal carcinoma. To elucidate the relationship between allelic variants of *MHC* class II regulatory genes and susceptibility to nasopharyngeal carcinoma, a total of 28 polymorphic loci at *MHC* class II regulatory genes, involving *CIITA*, *CREB1*, *RFX* family genes (*RFX5*, *RFXAP,* and *RFXANK*), and *NFY* family genes *(NFYA*, *NFYB*, and *NFYC*), were genotyped by multiplex SNaPshot minisequencing in 137 patients with nasopharyngeal carcinoma and 107 healthy controls from the southern Chinese population. Allelic analysis disclosed that rs7404873, rs6498121, rs6498126, and rs56074043 shared correlations with nasopharyngeal carcinoma ($P_{trend} < 0.05$). Further, rs6498126 on *CIITA* was independently associated with the risk of developing nasopharyngeal carcinoma (CC vs. GG, odds ratio: 7.386, 95% confidence interval: 1.934–28.207, $P_{trend} < 0.01$). Conversely, rs7404873 on *CIITA* and rs56074043 on *NFYB* manifested epistatic interaction to decreased susceptibility of nasopharyngeal carcinoma (rs7404873, TT vs. GG, odds ratio: 0.256, 95% confidence interval: 0.088–0.740, $P_{trend} < 0.05$; rs56074043, AA vs. AG, odds ratio: 0.341, 95% confidence interval: 0.129–0.900, $P_{trend} < 0.05$). Additionally, bioinformatics analysis revealed that the three variants were transcriptional regulatory in function and might impact the expression of nearby genes. The findings suggested genetic variants on *MHC* class II regulatory genes contributed to nasopharyngeal carcinoma susceptibility and might provide new insights for screening high-risk population. *European Journal of Cancer Prevention* 29: 531–537 Copyright © 2020 The Author(s). Published by Wolters Kluwer Health, Inc.

European Journal of Cancer Prevention 2020, **29**:531–537

Keywords: alleles, disease susceptibility, major histocompatibility complex class II regulatory genes, nasopharyngeal carcinoma

[a]Nanjing Medical University, Nanjing, [b]Department of Radiotherapy, The First Affiliated Hospital of Hainan Medical University, [c]Department of Radiotherapy, The Second Affiliated Hospital of Hainan Medical University, [d]Department of Radiotherapy, Hainan Cancer Hospital, [e]Hainan Medical University and [f]Department of Hematology, Center for Translational Medicine, The First Affiliated Hospital, HMC Cancer Institute and Academician Research Workstation of Hainan, Hainan Medical University, Haikou, China

Correspondence to Xin-bao Hao, PhD, Nanjing Medical University, Nanjing 211166, China
Tel: + 86 0898 6653 8603; fax: +86 0898 6678 2831;
e-mail: haoxb@foxmail.com

Received 27 September 2019 **Accepted** 6 December 2019

Introduction

Nasopharyngeal carcinoma (NPC), an epithelial malignancy originating from nasopharynx, is a complex disease caused by an accumulation of Epstein–Barr virus (EBV) chronic infection, environmental risk factors, and host genetics in the multistep process of carcinogenesis (Chen *et al.*, 2019). Virological factor is one of the main tumorigenic causes of NPC; however, infection with the same EBV has been found to result in various clinical outcomes in different patient, suggesting that there is a strong immunogenic component to affect individual vulnerability to EBV infection (Hildesheim and Wang, 2012; Mahdavifar *et al.*, 2016). It has been generally established that cellular immunity against EBV infection primarily focus on major histocompatibility complex (MHC) class I-restricted CD8+ T cell after EBV penetration into epithelial cells or B lymphocytes (Antsiferova *et al.*, 2014). Recently, researchers have subscribed to the belief that class II MHC-restricted CD4+ T cell is also the most widely distributed contributory factor in effective anti-infection and antitumor immunity (Lam *et al.*, 2018). MHC class II molecule, an essential role for antigen presentation, has emerged as powerful immunoregulatory function in immune reaction and its deficiency may therefore engender autoimmune diseases, infectious diseases, and cancer (Marty Pyke *et al.*, 2018; Trier *et al.*, 2018; Gameiro *et al.*, 2019).

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (www.eurjcancerprev.com).

0959-8278 Copyright © 2020 The Author(s). Published by Wolters Kluwer Health, Inc. DOI: 10.1097/CEJ.0000000000000563

Expression of *MHC* class II gene is principally regulated at the transcriptional level by several transactivating genes, namely *MHC* class II regulatory genes, comprising *MHC* class II transactivator (*CIITA*), *CREB1*, *RFX* family genes (*RFX5*, *RFXAP*, and *RFXANK*), and *NFY* family genes (*NFYA*, *NFYB*, and *NFYC*) (Lochamy *et al.*, 2007; Sachini and Papamatheakis, 2017). During the transcriptional process of *MHC* class II gene, CIITA synergistic CREB1, RFX family, and NFY family assembles into a cell-specific multiprotein complex and coregulates *MHC* class II gene expression (Lochamy *et al.*, 2007). There is evidence to demonstrate that mutations on *MHC* class II regulatory genes are responsible for MHC class II molecule deficiency, subsequently bringing as a result to the bare lymphocyte (Masternak *et al.*, 1998). Similarly, it has been reported that variant alleles on *MHC* class II regulatory genes are correlated with susceptibility to immune diseases, infectious disease, and cancer (Zhang *et al.*, 2007; Martínez *et al.*, 2010; Steidl *et al.*, 2011).

On the basis of functional relevance and clinical manifestation of *MHC* class II regulatory genes in the pathogenesis of immune-related diseases, our hypothesis is that genetic variants on *CIITA*, *CREB1*, *RFX5*, *RFXAP*, *RFXANK*, *NFYA*, *NFYB*, and *NFYC* might impress effects on *MHC* class II gene expression, then modifying host immune reaction to EBV infection and contributing to NPC development. Thus, we conducted a pilot association analysis starting from NPC cases and healthy controls in the Hainan population in South China.

Participants and methods

Study participants

For the case-control study, 137 NPC cases and 107 healthy controls were recruited from the First Affiliated Hospital of Hainan Medical University, the Second Affiliated Hospital of Hainan Medical University, and Hainan Cancer Hospital between January 2018 and June 2019. All patients were pathologically confirmed. Healthy controls with no self-reported history of cancer were randomly selected from physical examination centers and were frequency-matched to the NPC patients by age (±5 years), gender, residential region, and ethnicity. Informed consent was obtained from all the subjects, and the study was performed with the approval of the ethical committee of Hainan Medical University (No. 2018-35).

At recruitment, personal information on demographic factors, medical history, tobacco, and family cancer history were collected. Individuals were considered smokers if they smoked up to 1 year (20 pack per year) before the date of cancer diagnosis or the date of interview for controls. Epstein–Barr virus capsid antigen (EB-VCA) IgA was positively defined as an EBV infection. Two or more patients with NPC in an immediate family were defined as family history.

Variants selection

In the first stage, genome variations on *CIITA*, *CREB1*, *RFX5*, *RFXAP*, *RFXANK*, *NFYA*, *NFYB*, and *NFYC* were chosen based on biological function score (score ≥ 5) on Regulome DB database (Boyle *et al.*, 2012). Then these functional variants were further selected according to genetic correlation with diseases in the previous genome-wide association studies or candidate genes studies. Afterwards, combined with Haploview 4.2 to construct linkage disequilibrium between candidate variants (parameter setting: Minimum MAF > 0.05, r2 < 0.8) on the basis of the population of Han Chinese in Beijing, China (CHB) in the International Genome Sample Resources 1000 genomes project thus excluded locus with strong evidence of linkage disequilibrium. Overall, 28 candidate variants on *MHC* class II regulatory genes were hypothesized to be associated with NPC and were selected for further genotyping (Table S1, Supplemental digital content 1, *http://links.lww.com/EJCP/A269*).

DNA preparation and PCR multiplex amplification

Genomic DNA was extracted from whole blood sample by commercial DNA extraction kits (Yaneng Bioscience Co., Ltd). Fifty-six primers were used for amplification of the mutation region (Table S2, Supplemental digital content 1, *http://links.lww.com/EJCP/A269*), and PCR amplification (Applied Biosystem, Thermo Fisher Scientific Company) was conducted in a reaction mixture that contained 1 μl DNA template, 3 μl deoxynucleotide, 5 μl 2× Taq PCR Master Mix, and 1 μl Primer Mix. The PCR conditions were as follows: initial denaturation at 95°C for 5 min, followed by denaturation at 94°C for 30 s, annealing at 56°C for 30 s, extension at 72°C for 30 s, and final extension at 72°C for 2 min. Five microliters of the PCR products were separated and visualized following electrophoresis on a 2% agarose gel (Figure S1, Supplemental digital content 1, *http://links.lww.com/EJCP/A269*). Shrimp alkaline phosphatase (SAP) and exonuclease I was used for post-PCR purification of the examined PCR products.

SNaPshot minisequencing reaction

The single base extension (SBE) reaction was performed in a reaction mixture with 3 μl purified multiplex PCR product, 1 μl pooled extension primers (Table S2, Supplemental digital content 1, *http://links.lww.com/EJCP/A269*), 0.5 μl SNaPshot Mix, and 0.5 μl dH_2O. The SNaPshot reaction conditions were as follows: 40 SBE cycles of initial denaturation at 95°C for 10 s, primer annealing at 52°C for 5 s, and primer extension at 60°C for 30 s. SAP was used for post-SBE purification of the SNaPshot reaction products. Two microliters of minisequencing reaction product was mixed with 8 μl of formamide containing 0.4% LIZ120 and was denatured at 95°C for 5 min. After cooling down to −20°C, the fluorescently labeled fragments were resolved by capillary electrophoresis on an ABI 3730XL Genetic Analyzer for

allele discrimination. The resulting data were analyzed with the GeneMapper v 4.0 software.

Statistical analysis

Allele and genotype frequencies for variants were determined by gene counting. The significance of deviations from Hardy–Weinberg equilibrium was tested by χ^2 test. *t*-test was used to assess age difference and χ^2 test was employed to evaluate the differences in gender, smoking, EBV infection, and family history between cases and controls. Correlations between variations and NPC risk were estimated by logistic regression analysis with adjustment for gender, age, smoking, EBV infection, and family history. Odds ratio (OR) and 95% confidence interval (CI) were used to measure the strength of association. All analysis was performed using SPSS software (version 24.0, SPSS Inc., Chicago, Illinois). A *P*-value of less than 0.05 was considered significant.

Bioinformatics analysis

Potential function of associated loci was analyzed by Regulome DB and HaploReg v2. Regulome DB is a database that annotates single-nucleotide polymorphism (SNP) with known and predicted regulatory elements (Boyle *et al.*, 2012). HaploReg v2 is a tool for exploring annotations of candidate regulatory SNPs at disease-associated loci and developing mechanistic hypotheses of the impact of noncoding variants on clinical phenotypes and normal variation (Ward and Kellis, 2012).

Results

Clinical and demographic characteristics

Table 1 shows that the median age was 56 ± 12 years in cases and 51 ± 9 year in controls, with no significant difference ($P_{trend} > 0.05$). No statistical difference was found in gender, smoking, and family history between the two groups ($P_{trend} > 0.05$), whereas more EBV infection were presented in NPC subjects than in control subjects (86.1% vs. 15.0%, $P_{trend} = 1.438 \times 10^{-28}$).

Table 1 Clinical and demographic characteristics of nasopharyngeal carcinoma patients and healthy controls

Variables	Cases (n = 137)	Controls (n = 107)	*P* value
Median age	56 ± 12	51 ± 9	0.126
Gender			0.080
Male	99 (72.3)	66 (61.7)	
Female	38 (27.7)	41 (38.3)	
Age (years)			0.136
≤45	33 (24.1)	35 (32.7)	
>45	104 (75.9)	72 (67.3)	
Smoking			0.604
No	75 (54.7)	55 (51.4)	
Yes	62 (45.3)	52 (48.6)	
EBV infection			1.438×10^{-28}
No	19 (13.9)	91 (85.0)	
Yes	118 (86.1)	16 (15.0)	
Family history			0.177
No	128 (93.4)	104 (97.2)	
Yes	9 (6.6)	3 (2.8)	

EBV, Epstein–Barr virus.

Association of allele frequency on *MHC* class II regulatory genes with nasopharyngeal carcinoma

The distribution of allele and genotype frequencies on *MHC* class II regulatory genes in each group is shown in Table 2. Variation at rs7197779, rs2551639, rs2709359, rs2709387, rs2233857, rs2233843, and rs62135502 has not yet been observed. Other genotype distributions were consistent with the existence of Hardy–Weinberg equilibrium ($P_{trend} > 0.05$). Allelic frequency of rs7404873 on *CIITA* for T and G were 53.3% and 46.7%, 62.6% and 37.4%, rs6498121 for T and G were 59.5% and 40.5%, 59.8% and 40.2%, and rs6498126 for C and G were 56.9% and 43.1%, 65.9% and 34.1% in the NPC populations and controls, respectively. In addition, allelic frequency of rs56074043 on *NFYB* for A and G were 90.4% and 9.6%, 85.0% and 15.0%. Based on logistic regression analysis with adjustment for gender, age, smoking, EBV infection, and family history, there was a strong linkage of allele G on rs6498126 with a high risk of NPC (OR: 1.855, 95% CI: 1.057–3.256, $P_{trend} < 0.05$), whereas allele G of rs7404873, rs6498121, and rs56074043 presented reduced susceptibility to NPC (rs7404873, OR: 0.472, 95% CI: 0.268–0.830, $P_{trend} < 0.001$; rs6498121, OR: 0.544, 95% CI: 0.611–0.962, $P_{trend} < 0.05$; rs56074043, OR: 0.349, 95% CI: 0.154–0.788, $P_{trend} < 0.05$). No evidence for genotype–phenotype correlation was observed between *CREB*, *RFX5*, *RFXANK*, *RFXAP*, *NFYA*, and *NFYC* and NPC risk ($P_{trend} > 0.05$).

Correlation of genotype frequency on *MHC* class II regulatory genes with nasopharyngeal carcinoma

In view of the logistic regression analysis with allele frequencies, we further analyzed the relevance between genotype frequencies of rs7404873, rs6498121, rs6498126, and rs56074043 and NPC occurrence. The genotype distribution of four loci is shown in Table 3. There was remarkable difference between NPC and control subjects in the distribution of CC/GG genotypes of rs6498126 (OR: 7.386, 95% CI: 1.934–28.207, $P_{trend} < 0.01$). Subjects bearing the GG homozygote on rs6498126 had an increased risk of NPC. Meanwhile, logistic regression analysis revealed a significant association of GG genotypes on rs7404873 and AG genotypes on rs56074043 with the low-risk of NPC (rs7404873, OR: 0.256, 95% CI: 0.088–0.740, $P_{trend} < 0.05$; rs56074043, OR: 0.341, 95% CI: 0.129–0.900, $P_{trend} < 0.05$). No relation between rs6498121 genotype and NPC risk was identified. Moreover, we explored the effects of demographic factors on genotype distribution of associated loci and the data manifested that there was no demographic factors linked closely to genotype distribution of rs7404873, rs6498126, and rs56074043 in NPC patients as shown in Table 4 ($P_{trend} > 0.05$).

Potential functional implication of associated loci

Two genetic loci, rs7404873 (Regulome DB score: 2b) and rs6498126 (Regulome DB score: 1f), are intronic region of

Table 2 Association between variants on major histocompatibility complex class II regulatory genes and risk of nasopharyngeal carcinoma in the Hainan population

							Genotype (137 cases and 107 controls)			Major allele frequency/minor allele frequency			
No.	CHR	rs ID	Gene	POS	Base change	HW	Common[a]	Heterozygous[a]	Rare[a]	Case	Control	OR (95% CI)[b]	*P* value[b]
1	16	rs4781011	*CIITA*	10881454	T>G	0.144	74.5/74.8	24.1/25.2	1.4/0	86.1/13.9	86.9/13.1	1.189 (0.544–2.602)	0.664
2	16	rs7404873	*CIITA*	10897592	G>T	0.055	32.8/43.0	40.9/38.3	26.3/18.7	53.3/46.7	62.6/37.4	0.472 (0.268–0.830)	0.009
3	16	rs6498121	*CIITA*	10898858	T>G	0.606	32.2/34.5	54.7/50.5	13.1/15.0	59.5/40.5	59.8/40.2	0.544 (0.611–0.962)	0.036
4	16	rs11647384	*CIITA*	10903432	A>G	0.721	29.9/31.8	46.7/47.7	23.4/20.6	53.3/46.7	55.6/44.4	1.046 (0.611–1.789)	0.871
5	16	rs4774	*CIITA*	10906991	G>C	0.735	76.6/72.0	21.9/25.2	1.5/2.8	89.2/12.8	84.6/15.4	0.822 (0.385–1.753)	0.611
6	16	rs7197779	*CIITA*	10909070	A>G	ND	100/100	ND	ND	ND	ND	ND	ND
7	16	rs6498126	*CIITA*	10910506	G>C	0.137	29.2/40.2	56.2/51.4	14.6/8.4	56.9/43.1	65.9/34.1	1.855 (1.057–3.256)	0.031
8	16	rs11074938	*CIITA*	10912686	A>G	0.628	32.8/24.3	42.3/52.3	24.9/23.4	54.0/46.0	50.5/49.5	1.696 (0.966–2.977)	0.066
9	16	rs8052440	*CIITA*	10912847	G>A	0.478	62.0/68.2	31.4/29.9	6.6/1.9	77.7/22.3	83.2/16.8	1.583 (0.788–3.180)	0.197
10	2	rs2551639	*CREB1*	207531463	A>G	ND	100/100	ND	ND	ND	ND	ND	ND
11	2	rs2709359	*CREB1*	207550413	A>G	ND	100/100	ND	ND	ND	ND	ND	ND
12	2	rs73056905	*CREB1*	207557025	G>A	0.536	49.6/52.3	38.7/38.4	11.7/9.3	69.0/31.0	71.5/28.5	0.979 (0.544–1.759)	0.943
13	2	rs10932201	*CREB1*	207561533	G>A	0.067	39.4/43.9	37.2/38.3	23.4/17.8	57.7/42.3	63.1/36.9	0.822 (0.473–1.428)	0.487
14	2	rs2709387	*CREB1*	207577371	G>A	ND	100/100	ND	ND	ND	ND	ND	ND
15	1	rs7552906	*RFX5*	151341298	A>G	0.715	46.7/43.0	40.2/43.9	13.1/13.1	66.8/33.2	65.0/35.0	0.844 (0.478–1.491)	0.560
16	1	rs1752387	*RFX5*	151341903	T>C	0.446	51.8/53.3	40.1/37.4	8.1/9.3	71.9/28.1	72.0/28.0	0.882 (0.483–1.609)	0.682
17	1	rs2233857	*RFX5*	151341942	T>A	ND	100/100	ND	ND	ND	ND	ND	ND
18	1	rs2233843	*RFX5*	151346257	C>T	ND	100/100	ND	ND	ND	ND	ND	ND
19	13	rs73531874	*RFXAP*	36820819	C>T	0.331	51.8/50.5	45.3/38.3	2.9/11.2	74.5/25.5	69.6/30.4	1.328 (0.712–2.474)	0.372
20	19	rs62135502	*RFXANK*	19192376	G>A	ND	100/100	ND	ND	ND	ND	ND	ND
21	19	rs34282046	*RFXANK*	19194090	G>C	0.883	97.8/97.22	2.2/2.8	0/0	98.9/1.1	98.6/1.4	0.999 (0.064–15.532)	1.000
22	19	rs2283626	*RFXANK*	19194156	A>G	0.099	49.6/45.8	46.7/48.6	3.7/5.6	73.0/27.0	70.1/29.9	0.862 (0.476–1.560)	0.623
23	19	rs1802498	*RFXANK*	19201687	C>G	0.321	70.8/76.6	24.8/20.6	4.4/2.8	83.2/16.8	86.9/13.1	1.640 (0.751–3.583)	0.215
24	6	rs4140578	*NFYA*	41074090	A>G	0.295	54.0/55.2	36.5/35.5	9.5/9.3	72.3/27.7	72.9/27.1	1.035 (0.567–1.890)	0.911
25	12	rs56074043	*NFYB*	104137402	A>G	0.289	83.8/71.0	14.0/28.1	2.2/0.9	90.4/9.6	85.0/15.0	0.349 (0.154–0.788)	0.011
26	1	rs657149	*NFYC*	40692486	A>C	0.713	30.1/32.7	55.9/50.5	14.0/16.8	57.4/42.6	57.9/42.1	1.065 (0.617–1.840)	0.820
27	1	rs12743517	*NFYC*	40759682	A>C	0.066	31.6/43.9	52.2/38.3	16.2/17.8	57.7/42.3	63.1/36.9	1.014 (0.587–1.751)	0.960
28	1	rs2361651	*NFYC*	40764542	G>A	0.439	30.2/31.8	56.6/52.3	13.2/15.9	58.5/41.5	57.5/42.5	1.006 (0.583–1.737)	0.983

P value of logistic regression analysis.
CHR, chromosome; CI, confident intervals; EBV, Epstein–Barr virus; HW, Hardy–Weinberg equilibrium; ND, no data; OR, odds ratios; POS, positon at bp.
[a]Frequency of the genotype in the patients/frequency of the genotype in the controls.
[b]The ORs and *P* values were adjusted for gender, age, smoking, EBV infection, and family history of nasopharyngeal carcinoma.

Table 3 The genotype frequency of four-associated variants in nasopharyngeal carcinoma patients and healthy controls

rs ID	Genotype	Cases (n = 137)	Controls (n = 107)	OR (95% CI)[a]	*P* value[a]
rs7404873	TT	45 (32.8)	46 (43.0)	1	
	TG	56 (40.9)	41 (38.3)	0.480 (0.169–1.360)	0.167
	GG	36 (26.3)	20 (18.7)	0.256 (0.088–0.740)	0.012
rs6498121	TT	44 (32.1)	37 (34.6)	1	
	TG	75 (54.8)	54 (50.4)	0.744 (0.233–2.380)	0.618
	GG	18 (13.1)	16 (15.0)	0.321 (0.092–1.123)	0.075
rs6498126	CC	40 (29.2)	43 (40.1)	1	
	CG	77 (56.2)	55 (51.5)	1.289 (0.556–2.984)	0.554
	GG	20 (14.6)	9 (8.4)	7.386 (1.934–28.207)	0.003
rs56074043	AA	114 (83.8)	76 (71.0)	1	
	AG	19 (14.0)	30 (28.1)	0.341 (0.129–0.900)	0.030
	GG	3 (2.2)	1 (0.9)	0.293 (0.027–3.209)	0.315

CI, confident intervals; EBV, Epstein–Barr virus; OR, odds ratios.
[a]The OR and *P* values were calculated by logistic regression analysis adjusted for gender, age, smoking, EBV infection, and family history of nasopharyngeal carcinoma.

CIITA, located within transcription factor binding sites, DNase peaks, and histone modifications as shown in the Regulome database. Rs7404873 was found in RUNX3 protein binding sites and Zfx binding motif when rs6498126 had eQTL effect on the expression of DEXI in monocytes and was speculated to bind Pax-8. Enquiry against HaploReg revealed that both of rs7404873 and rs6498126 were related to enhancer histone marks BLD. Meanwhile, rs7404873 altered the chromatin status for bindings of HNF4 and Zfx and rs6498126 changed the chromatin status for bindings of Pax-8, indicating that the two SNPs were highly accessible to regulatory factors (Table S3, Supplemental digital content 1, *http://links.lww.com/EJCP/A269*). The other SNP, rs56074043, is an intronic variant at *NFYB* and its Regulome DB score is 2c. Regulome DB survey disclosed that rs56074043 was located in a region with protein binding, Tcfe2a or RUNX2 motifs, histone modifications, and chromatin structures. Data mining at HaploReg also showed proteins bound in ChIP-Seq experiments for rs56074043 (Table S3, Supplemental digital content 1, *http://links.lww.com/EJCP/A269*).

Table 4 Association between demographic factors and genotype distribution in nasopharyngeal carcinoma patients

	rs7404873			rs6498126			rs56074043		
Factor/genotype	TT	TG + GG	*P* value	CC	CG + GG	*P* value	AA	AG + GG	*P* value
Gender			0.833			0.734			0.606
Female	13	25		10	27		32	5	
Male	32	67		30	70		82	17	
Age (years)			0.227			0.551			0.923
≤45	8	25		8	24		27	5	
>45	37	67		32	73		87	17	
Smoking			0.894			0.735			0.989
No	25	50		21	54		62	12	
Yes	20	42		19	43		52	10	
EBV infection			0.299			0.430			0.738
No	4	15		4	16		17	2	
Yes	41	77		36	81		97	20	
Family history			0.154			0.778			0.354
No	40	88		37	91		105	22	
Yes	5	4		3	6		9	0	

EBV, Epstein–Barr virus.

Discussion

EBV infection attached to be the most common causal agent of NPC is ubiquitous worldwide, affecting more than 95% of the worldwide population (Kanda *et al.*, 2019). However, only a very small percentage of infected population develops EBV-associated NPC and its incidence shows remarkably different geographic distributions, which is paradoxical in comparison to the widespread infection (Mahdavifar *et al.*, 2016). Besides, patients infected with the same EBV strain display various pathological processes and clinical outcomes (Tsang and Tsao, 2015). These results suggest that host immune reaction could play a pivotal role in the various outcomes of EBV infection and recent work has documented that immune-related genetics may be the contributory factor, intervening the individual susceptibility to EBV infection (Hildesheim and Wang, 2012; Chen *et al.*, 2019). In the genome-wide association studies, *HLA* genes residing at the *MHC* region have been widely recognized as genetic susceptible loci conferring NPC risk, which revealed that carriers of *HLA* risk-associated alleles may have differential viral antigen-presenting capacity and thus the ability to activate diverse antiviral immune responses, affecting the immune surveillance of viral infection and subsequently NPC susceptibility (Bei *et al.*, 2010; Chin *et al.*, 2015).

Recent studies have illuminated that class II MHC-restricted CD4+ T cell plays a key supporting role in anti-infective and antitumor immune responses by releasing various cytokines, thereafter amplifying the activity of CD8+ T cell and macrophages and inducing the activation and differentiation of B lymphocytes (Borst *et al.*, 2018). MHC class II molecule, the integral membrane glycoproteins, is constitutively expressed on the surface of professional antigen-presenting cells, including dendritic cells, B cells, macrophages, and thymic epithelial cells, where it captures antigenic peptides and presents them to T lymphocytes (Roche and Furuta, 2015). Consequently, MHC class II molecule is fundamental to immune reaction and its deficiency can conduce to autoimmune diseases, infectious diseases, and cancer. The principal genetic defect, however, is not attributed to *MHC* class II gene themselves (Masternak *et al.*, 1998). Instead, the affected genes involved *CIITA*, *CREB1*, *RFX* family genes (*RFX5*, *RFXAP*, and *RFXANK*), and *NFY* family genes (*NFYA*, *NFYB*, and *NFYC*) encoding transactivating factors that regulate expression of *MHC* class II gene and are responsible for the deficiency of *MHC* class II molecule (Masternak *et al.*, 1998). Reportedly, genetic mutation on *MHC* class II regulatory genes can impress effects on expression of *MHC* class II gene (Lochamy *et al.*, 2007). As for genotype–phenotype correlation, genetic variants on *CIITA*, *CREB1*, *NFYC*, *RFXANK*, and *RFXAP* have been related to squamous cell lung carcinoma, liver cancer, chronic hepatitis B, systemic lupus erythematosus, multiple sclerosis, lymphocyte syndrome, rheumatoid arthritis, and type 2 diabetes (Kishimoto *et al.*, 2005; Takenoyama *et al.*, 2006; Zhang *et al.*, 2007; Bronson *et al.*, 2011; Wang *et al.*, 2013; Xu *et al.*, 2018; Pereira *et al.*, 2019). Given its indispensable role in the expression of *MHC* class II gene, *MHC* class II regulatory genes may be the rudimentary candidates of susceptible genes for NPC. No data to date are available about the relationship between allele variants on *MHC* class II regulatory genes and susceptible traits to NPC.

This is the first study to address the role of allelic variants on *MHC* class II regulatory genes in nasopharyngeal tumorigenesis. We studied the sequence variability at *CIITA*, *CREB1*, *RFX5*, *RFXAP*, *RFXANK*, *NFYA*, *NFYB*, and *NFYC* in NPC patients and normal controls by SNaPshot analysis to evaluate whether there was association of genetic polymorphism with the etiology of NPC. *CIITA* genomic sequence variation represented by rs6498126 genotype GG was substantiated in relation to a high risk of NPC. On the contrary, genotype GG of rs7404873 on *CIITA* and genotype AG of rs56074043 on *NFYB* were

closely linked to a lower incidence of NPC. These findings lent support to the hypothesis that genetic variants on *MHC* class II regulatory genes had effects on NPC development. A similar scenario was the correlation of rs6498114 at *CIITA* with NPC risk in an extended genome-wide association study (Cui *et al*., 2016). The results denoted that *CIITA* and *NFYB* genes polymorphisms could impact NPC-associated *MHC* class II transcription, then interfere the presentation of EBV antigens to host immune cells and leading to NPC susceptibility.

The identification of variants (rs7404873, rs6498126, and rs56074043) for NPC susceptibility emphasizes that the contribution of *MHC* class II regulatory genes to NPC-related immune should not be ignored. Bioinformatics analysis indicated that rs7404873 and rs6498126 on *CIITA* were transcriptional regulatory and could bind protein, and thus might be the functional variation (Boyle *et al*., 2012; Ward and Kellis, 2012). Both rs7404873 and rs6498126 are located in *CIITA*, which implies that *CIITA* represents a key molecule in the regulation of *MHC* class II genetic expression and its mutation may greatly influence *MHC* class II gene expression. It is noteworthy that rs56074043 on *NFYB*, having a biologically active protein bound, might be transcriptional functional (Boyle *et al*., 2012; Ward and Kellis, 2012). Collectively, *MHC* class II regulatory gene variability may confer NPC susceptibility via regulating NPC-associated *MHC* class II gene expression. Nevertheless, further studies are awaited to clarify whether and how rs7404873 or other genetic variants disrupt the transcriptional regulation of *MHC* class II gene and then are correlated with NPC occurrence.

Conclusion

The current data conveyed a linkage between variants on *MHC* class II regulatory genes (*CIITA* and *NFYB*) and NPC susceptibility. Moreover, we noted that these genetic mutations might be functional regulators for *MHC* class II gene expression. Conclusively, these results suggest genetic variants on *CIITA* and *NFYB* contribution to NPC pathogenesis and provide new insights for screening populations at a high risk of NPC. It has to be acknowledged that the principal limitation of this analysis was that the sample size recruited is not enough. Further studies with larger sample sizes, more ethnic groups, and more geographic regions are needed to replicate the findings and rule out the confounding effects of population. Certainly, more efforts are required to investigate the functional mechanism of genetic variants.

Acknowledgements

This study was supported by Hainan Province Health and Family Planning Commission (no. 18A200027) and Hainan Science and Technology Association (no. QCXM201917).

We thank Physical Examination Centers of the First Affiliated Hospital of Hainan Medical University for providing personal information and blood sample of healthy controls, Clinical Laboratory of the First Affiliated Hospital of Hainan Medical University for providing detection of Epstein–Barr virus capsid antigen, and Department of Biochemistry and Molecular Biology of Hainan Medical University for providing biotechnology support. We appreciate all patients and their families for their participation in this study.

Conflicts of interest

There are no conflicts of interest.

References

Antsiferova O, Müller A, Rämer PC, Chijioke O, Chatterjee B, Raykova A, *et al*. (2014). Adoptive transfer of EBV specific CD8+ T cell clones can transiently control EBV infection in humanized mice. *Plos Pathog* **10**:e1004333.

Bei JX, Li Y, Jia WH, Feng BJ, Zhou G, Chen LZ, *et al*. (2010). A genome-wide association study of nasopharyngeal carcinoma identifies three new susceptibility loci. *Nat Genet* **42**:599–603.

Borst J, Ahrends T, Bąbała N, Melief CJM, Kastenmüller W (2018). CD4+ T cell help in cancer immunology and immunotherapy. *Nat Rev Immunol* **18**:635–647.

Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, *et al*. (2012). Annotation of functional variation in personal genomes using regulomedb. *Genome Res* **22**:1790–1797.

Bronson PG, Goldstein BA, Ramsay PP, Beckman KB, Noble JA, Lane JA., *et al*. (2011). The rs4774 CIITA missense variant is associated with risk of systemic lupus erythematosus. *Genes Immun* **12**:667–671.

Chen YP, Chan ATC, Le QT, Blanchard P, Sun Y, Ma J (2019). Nasopharyngeal carcinoma. *Lancet* **394**:64–80.

Chin YM, Mushiroda T, Takahashi A, Kubo M, Krishnan G, Yap LF, *et al*.; Malaysian NPC Study Group. (2015). HLA-A snps and amino acid variants are associated with nasopharyngeal carcinoma in Malaysian Chinese. *Int J Cancer* **136**:678–687.

Cui Q, Feng QS, Mo HY, Sun J, Xia YF, Zhang H, *et al*. (2016). An extended genome-wide association study identifies novel susceptibility loci for nasopharyngeal carcinoma. *Hum Mol Genet* **25**:3626–3634.

Gameiro SF, Ghasemi F, Barrett JW, Nichols AC, Mymryk JS (2019). High level expression of MHC-II in HPV+ head and neck cancers suggests that tumor epithelial cells serve an important role as accessory antigen presenting cells. *Cancers (Basel)* **11**:1129.

Hildesheim A, Wang CP (2012). Genetic predisposition factors and nasopharyngeal carcinoma risk: a review of epidemiological association studies, 2000-2011: rosetta stone for NPC: genetics, viral infection, and other environmental factors. *Semin Cancer Biol* **22**:107–116.

Kanda T, Yajima M, Ikuta K (2019). Epstein-Barr virus strain variation and cancer. *Cancer Sci* **110**:1132–1139.

Kishimoto M, Kohno T, Okudela K, Otsuka A, Sasaki H, Tanabe C, *et al*. (2005). Mutations and deletions of the CBP gene in human lung cancer. *Clin Cancer Res* **11**:512–519.

Lam JKP, Hui KF, Ning RJ, Xu XQ, Chan KH, Chiang AKS (2018). Emergence of CD4+ and CD8+ polyfunctional T cell responses against immunodominant lytic and latent EBV antigens in children with primary EBV infection. *Front Microbiol* **9**:416.

Lochamy J, Rogers EM, Boss JM (2007). CREB and phospho-CREB interact with RFX5 and CIITA to regulate MHC class II genes. *Mol Immunol* **44**: 837–847.

Mahdavifar N, Ghoncheh M, Mohammadian-Hafshejani A, Khosravi B, Salehiniya H (2016). Epidemiology and inequality in the incidence and mortality of nasopharynx cancer in Asia. *Osong Public Health Res Perspect* **7**:360–372.

Martínez A, Perdigones N, Cénit MC, Espino L, Varadé J, Lamas JR, *et al*. (2010). Chromosomal region 16p13: further evidence of increased predisposition to immune diseases. *Ann Rheum Dis* **69**:309–311.

Marty Pyke R, Thompson WK, Salem RM, Font-Burgada J, Zanetti M, Carter H (2018). Evolutionary pressure against MHC class II binding cancer mutations. *Cell* **175**:1991.

Masternak K, Barras E, Zufferey M, Conrad B, Corthals G, Aebersold R, *et al*. (1998). A gene encoding a novel RFX-associated transactivator is mutated in the majority of MHC class II deficiency patients. *Nat Genet* **20**:273–277.

Pereira VCSR, Fontes-Dantas FL, Paradela ER, Malfetano FR, Scherpenhuijzen SSB, Mansur LF, *et al*. (2019). Polymorphisms in the CIITA -168A/G (rs3087456) and CIITA +1614G/C (rs4774) may influence severity in multiple sclerosis patients. *Arq Neuropsiquiatr* **77**:166–173.

Roche PA, Furuta K (2015). The ins and outs of MHC class II-mediated antigen processing and presentation. *Nat Rev Immunol* **15**:203–216.

Sachini N, Papamatheakis J (2017). NF-Y and the immune response: dissecting the complex regulation of MHC genes. *Biochim Biophys Acta Gene Regul Mech* **1860**:537–542.

Steidl C, Shah SP, Woolcock BW, Rui L, Kawahara M, Farinha P, *et al*. (2011). MHC class II transactivator CIITA is a recurrent gene fusion partner in lymphoid cancers. *Nature* **471**:377–381.

Takenoyama M, Baurain JF, Yasuda M, So T, Sugaya M, Hanagiri T, *et al*. (2006). A point mutation in the NFYC gene generates an antigenic peptide recognized by autologous cytolytic T lymphocytes on a human squamous cell lung carcinoma. *Int J Cancer* **118**:1992–1997.

Trier N, Izarzugaza J, Chailyan A, Marcatili P, Houen G (2018). Human MHC-II with shared epitope motifs are optimal Epstein-Barr virus glycoprotein 42 ligands-relation to rheumatoid arthritis. *Int J Mol Sci* **19**:E317.

Tsang CM, Tsao SW (2015). The role of Epstein-Barr virus infection in the pathogenesis of nasopharyngeal carcinoma. *Virol Sin* **30**:107–121.

Wang J, Ma L, Weng W, Qiao Y, Zhang Y, He J, *et al*. (2013). Mutual interaction between YAP and CREB promotes tumorigenesis in liver cancer. *Hepatology* **58**:1011–1020.

Ward LD, Kellis M (2012). Haploreg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* **40**:D930–D934.

Xu Y, Song R, Long W, Guo H, Shi W, Yuan S, *et al*. (2018). CREB1 functional polymorphisms modulating promoter transcriptional activity are associated with type 2 diabetes mellitus risk in Chinese population. *Gene* **665**:133–140.

Zhang X, Hong X, Deng G, Bai X (2007). Single nucleotide polymorphisms and functional analysis of class II transactivator (CIITA) promoter IV in persistent HBV infection. *J Clin Virol* **40**:197–201.