Correlation of variable repeat number in the neck regions of DC-SIGN and DC-SIGNR with susceptibility to nasopharyngeal carcinoma in a Chinese population

Sisi Ning1,*, Mengwei Yao1,*, Yuan Wu1, Xunzhao Zhou1, Changtao Zhong1, Kui Yan1, Zhengbo Wei2, Ying Xie3,4

1Graduate School of Guangxi Medical University, Nanning, China; 2Department of Head and Neck Tumor Surgery, Affiliated Tumor Hospital of Guangxi Medical University, Nanning, China; 3Life Sciences Institute of Guangxi Medical University, Nanning, China; 4Key Laboratory for High-Incidence Tumor Prevention and Treatment, Guangxi Medical University, Ministry of Education, Nanning, China

*These authors contributed equally to this work

Objective: To evaluate the potential association of variations in the number of tandem repeats in the dendritic cell-specific intercellular adhesion molecule-3 grabbing nonintegrin (DC-SIGN) and dendritic cell-specific intercellular adhesion molecule-3 grabbing nonintegrin-related (DC-SIGNR) neck region with susceptibility to nasopharyngeal carcinoma (NPC).

Methods: Variations in the number of repeats in the genotypes and alleles in the neck region of DC-SIGN/DC-SIGNR were analyzed in 477 unrelated NPC patients and 561 cancer-free controls.

Results: Genotypes and alleles in the DC-SIGN neck region did not differ significantly between NPC patients and controls, but the 9-repeat genotype in the DC-SIGNR neck region was significantly more frequent among patients (OR 1.339, 95% CI 1.018–1.760, P=0.037). The association between this genotype and NPC remained significant after adjusting for sex, age, smoking history, and presence of immunoglobulin against Epstein–Barr virus viral capsid antigen (OR 1.625, 95% CI 1.134–2.329, P=0.0082).

Conclusion: These results suggest that genotypes/alleles in the DC-SIGN neck region are not associated with NPC susceptibility, whereas the 9-repeat variant in the neck region of DC-SIGNR may increase the risk of NPC.

Keywords: nasopharyngeal carcinoma, DC-SIGN, DC-SIGNR, genetic polymorphism, neck region repeat variation

Introduction

Nasopharyngeal carcinoma (NPC) is one of the most common malignant tumors in southern China, and its causes are poorly understood. Many studies indicate that genetic factors play an important role. Work from our group and others has suggested an association between NPC risk and single nucleotide polymorphism in the CD209 gene (Gene ID: 30835) encoding dendritic cell-specific intercellular adhesion molecule-3 grabbing nonintegrin (DC-SIGN). DC-SIGN is a C-type lectin expressed mainly on monocytes, immature dendritic cells (DCs), certain macrophages, and placental Hofbauer cells. The molecule, which contains intracellular, transmembrane, and extracellular domains, mediates host infection by various pathogens via a carbohydrate recognition domain (CRD) in the extracellular domain. This CRD recognizes high-mannose oligosaccharides on the surface of numerous pathogens, including HIV, Ebola virus, and cytomegalovirus. A so-called neck region in the extracellular domain consists of a variable number of 23-residue tandem repeats. The number of repeats in the
DC-SIGN neck region is usually 7 but can range from 4 to 8. It is possible that this polymorphism influences NPC risk.

Variations in the number of neck region repeats may affect the affinity of the CRD for pathogens because the neck region mediates the tetramerization of DC-SIGN on the host cell surface. Indeed, studies suggest that variations in the number of DC-SIGN neck region repeats may influence host susceptibility to many infectious diseases involving bacteria, virus, and parasites. For example, a heterozygous genotype in the DC-SIGN repeat region may reduce the risk of HIV-1 infection in a population from Seattle, USA. It is also possible that variation in the number of neck region repeats affects affinity for binding Epstein–Barr virus (EBV), which contains high-mannose glycoproteins like HIV and is strongly associated with NPC pathogenesis. Further work is needed to elucidate how polymorphism of DC-SIGN neck region repeats may contribute to NPC risk.

Such work should also examine whether polymorphism in the CLEC4M gene (Gene ID: 10332) influences NPC risk. This gene encodes dendritic cell-specific intercellular adhesion molecule-3 grabbing nonintegrin-related (DC-SIGNR) protein, which is another C-type lectin that shows 77% amino acid identity to DC-SIGN and also recognizes pathogens. DC-SIGNR is expressed mainly on lymphatic endothelial cells and lymph nodes, and it contains the same 3-domain structure as DC-SIGN, including a CRD and neck region. The neck region of DC-SIGNR is even more polymorphic than that of DC-SIGN, with the number of repeats varying from 3 to 9. Polymorphism of the neck region of DC-SIGNR appears to affect the ability of the CRD to bind virus, leading us to speculate that this polymorphism affects binding to EBV and therefore the risk of NPC.

Here, we examined whether polymorphism in the neck region repeats of DC-SIGN and DC-SIGNR is associated with the risk of NPC. We conducted a case–control study in the Chinese province of Guangxi, where NPC is endemic. The results may help clarify the genetic factors contributing to NPC, as well as suggest hypotheses into future studies of disease progression.

**Patients and methods**

**Study subjects**

The present study included 477 unrelated NPC patients and 561 cancer-free individuals. All patients were treated between July 2012 and July 2015 at the Affiliated Tumor Hospital of Guangxi Medical University, and their diagnosis of NPC was confirmed by pathology. Controls were recruited from individuals undergoing planned physical examinations at the Affiliated Tumor Hospital of Guangxi Medical University or the First Affiliated Hospital of Guangxi Medical University. None of the controls had a history of cancer. All patients and controls were residents of Guangxi province at the time of the study.

**Genomic DNA extraction**

Peripheral venous blood (3 mL) was collected from all subjects, and genomic DNA was extracted using the TIANamp Blood DNA Kit (Tiangen, Beijing, China). Extracted DNA was processed immediately for genotyping (see the "PCR-based genotyping of neck region repeat variations in DC-SIGN and DC-SIGNR" section) or stored at −20°C for future use.

**PCR-based genotyping of neck region repeat variations in DC-SIGN and DC-SIGNR**

The neck repeat region was amplified using PCR. For analysis of DC-SIGN polymorphism, primers 1F (5′-CCACTTTAGGGCAGGAC-3′) and 1R (5′-AGAAACACTCAACCACAA-3′) were used with the following thermocycling conditions: denaturation at 94°C for 3 minutes; 35 cycles of 94°C for 30 seconds, 60°C for 30 seconds, and 72°C for 1 minute; and 72°C for 5 minutes. For analysis of DC-SIGNR polymorphism, the primers 5′-TGTCACAGGTCCCTAGC-3′ (forward) and 5′-AGGACCCTCTTGATGTGCAGGAACTCACC-3′ (reverse) were used with the following thermocycling conditions: denaturation at 95°C for 5 minutes; 35 cycles of 94°C for 30 seconds, then 72°C for 55 seconds; and 72°C for 5 minutes.

Results of the PCR were analyzed on 2% agarose gels (Biowest, Madrid, Spain) stained with Gold View I. DNA bands were imaged using the GelDoc XR+ system (Bio-Rad Laboratories Inc., Hercules, CA, USA). Genotypes were determined from these gel images. Alleles were determined by direct sequencing of the amplified products (Thermo Fisher Scientific, Waltham, MA, USA). Alleles were expected to differ in the number of 69 bp tandem repeats. The expected possible amplicon lengths after amplification of the DC-SIGN neck region were 922 bp (8 repeats), 853 bp (7 repeats), and 646 bp (4 repeats). The expected possible amplicon lengths after amplification of the DC-SIGNR neck region were 734 bp (8 repeats), 665 bp (8 repeats), 596 bp (7 repeats), 527 bp (6 repeats), 458 bp (5 repeats), and 389 bp (4 repeats).

**Statistical analysis**

Data were analyzed using SPSS 17.0 (IBM, Chicago, IL, USA). Differences in frequencies of genotypes or alleles between NPC cases and controls were assessed for significance using the chi-squared test or Fisher’s exact test.
Multivariate logistic regression was used to estimate ORs with 95% CIs. When appropriate, ORs were adjusted by controlling for various combinations of age, sex, smoking history, and status of immunoglobulin A antibodies against EBV viral capsid antigen (EBV-VCA-IgA). Two-sided \( P<0.05 \) was defined as statistically significant.

**Ethics approval**
This study was approved by the Ethics Committee of Guangxi Medical University, and signed informed consent was obtained from all study subjects.

**Results**

**Demographics of patients and controls**
Patients and controls did not differ significantly in age \( (P=0.056) \), but they did differ significantly in sex distribution \( (P<0.001) \), smoking history \( (P<0.001) \), and presence of EBV-VCA-IgA \( (P<0.001; \text{Table 1}) \).

**Polymorphism in the DC-SIGN neck region**
Three DC-SIGN genotypes and 3 alleles in the neck region were identified among patients and cancer-free controls (Figure 1). All 3 alleles were confirmed by direct sequencing. The most frequent genotype in both groups was a homozygous “7/7” genotype (2 alleles encoding 7 tandem repeats). Frequencies of DC-SIGN genotypes and alleles did not differ significantly between patients and controls (Table 2).

**Association between DC-SIGNR genotypes and risk of NPC**
Analysis of genotype and allele frequencies in the DC-SIGNR neck region in patients and controls (Table 3) revealed high polymorphism: 10 genotypes and 4 alleles were identified (Figure 2). All 4 alleles were confirmed by direct sequencing. Six genotypes with frequencies <5% were...
merged into a single group to ensure adequate statistical power for intergroup comparisons. Comparison of patients and controls revealed no significant difference in genotype frequencies and only 1 significant difference in allele frequencies: the 9-repeat DC-SIGNR allele was significantly more frequent among patients than controls (OR 1.339, 95% CI 1.018–1.650, \( P=0.0082 \)).

**Association between 9-repeat DC-SIGNR and EBV-VCA-IgA status**

The frequency of genotypes containing the 9-repeat was marginally higher among controls positive for EBV-VCA-IgA than among controls negative for EBV-VCA-IgA (\( P=0.064 \); Table 5).

**Discussion**

DC-SIGN plays an important role in the occurrence of many diseases, including malignant tumors.\(^2\)\(^3\)\(^4\) Polymorphism in parts of DC-SIGN outside the neck region has been linked to NPC risk in a North African population and in Chinese populations from Guangdong and Guangxi.\(^1\)\(^2\)\(^3\)\(^4\) Whether the same is true of polymorphism in the neck region is unknown. Here, we provide the first evidence that it may not be associated with NPC risk, at least in a Chinese population. At the same time, we provide the first indications that the 9-repeat polymorphism in the neck region is associated with elevated disease risk. In fact, our results suggest that individuals carrying the 9-repeat are at 1.625-fold higher risk of NPC than those without the presence of EBV-VCA-IgA (OR 1.625, 95% CI 1.134–2.329, \( P=0.0082 \)).

**Table 3** Distribution of DC-SIGNR neck region genotypes and alleles in NPC patients and controls

<table>
<thead>
<tr>
<th>Genotype/allele</th>
<th>Controls (n)</th>
<th>Cases (n)</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5/5</td>
<td>28 (5.0)</td>
<td>15 (3.1)</td>
<td>0.927(^a)</td>
</tr>
<tr>
<td>6/5</td>
<td>13 (2.3)</td>
<td>9 (1.9)</td>
<td>–</td>
</tr>
<tr>
<td>6/6</td>
<td>4 (0.7)</td>
<td>1 (0.2)</td>
<td>–</td>
</tr>
<tr>
<td>9/5</td>
<td>10 (1.8)</td>
<td>13 (2.7)</td>
<td>–</td>
</tr>
<tr>
<td>9/6</td>
<td>6 (1.1)</td>
<td>9 (1.9)</td>
<td>–</td>
</tr>
<tr>
<td>9/9</td>
<td>13 (2.3)</td>
<td>15 (3.1)</td>
<td>–</td>
</tr>
<tr>
<td>7/5</td>
<td>110 (19.6)</td>
<td>83 (17.4)</td>
<td>0.362(^c)</td>
</tr>
<tr>
<td>7/6</td>
<td>62 (11.1)</td>
<td>41 (8.6)</td>
<td>0.187(^c)</td>
</tr>
<tr>
<td>7/7</td>
<td>206 (36.7)</td>
<td>183 (38.4)</td>
<td>0.585(^c)</td>
</tr>
<tr>
<td>9/7</td>
<td>109 (19.4)</td>
<td>108 (22.7)</td>
<td>0.205(^c)</td>
</tr>
<tr>
<td>Total</td>
<td>561</td>
<td>477</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Genotype</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>189 (16.8)</td>
<td>135 (14.1)</td>
</tr>
<tr>
<td>6</td>
<td>89 (7.9)</td>
<td>61 (6.4)</td>
</tr>
<tr>
<td>7</td>
<td>693 (61.8)</td>
<td>598 (62.7)</td>
</tr>
<tr>
<td>9</td>
<td>151 (13.5)</td>
<td>160 (16.8)</td>
</tr>
<tr>
<td>Total</td>
<td>1,122</td>
<td>954</td>
</tr>
</tbody>
</table>

**Notes:** Values are n (%). \(^a\)The number refers to the number of neck region repeats encoded by each genotype or allele. \(^c\)Chi-squared test for the merged group of 5/5, 6/5, 6/6, 9/5, 9/6, 9/9, each genotype frequency of which is <5%. \(^c\)Chi-squared test. \(^c\)Multiple logistic regression after controlling for age, sex, smoking history, and EBV-VCa-IgA status.

**Table 4** Distribution of DC-SIGNR genotypes with or without the 9-repeat in NPC patients and controls

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>OR (95% CI, ( P )-value)</th>
<th>Adjusted OR (95% CI, ( P )-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>With 9-repeat</td>
<td>145 (138) 1.339 (1.134–2.329, ( P=0.0082 ))</td>
<td>1.625 (1.018–2.650, ( P=0.0082 ))</td>
</tr>
<tr>
<td>Without 9-repeat</td>
<td>332 (423) (1.018–1.760, ( P=0.0082 ))</td>
<td>Total 477 (561) 0.037 (0.0082)</td>
</tr>
</tbody>
</table>

**Notes:** \(^a\)Chi-squared test. \(^b\)Chi-squared test. \(^c\)Multiple logistic regression after controlling for age, sex, smoking history, and EBV-VCa-IgA status.

**Table 5** Distribution of DC-SIGNR genotypes with or without the 9-repeat in controls positive or negative for EBV-VCa-IgA

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>EBV-VCa-IgA positive</th>
<th>EBV-VCa-IgA negative</th>
<th>( P )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>With 9-repeat</td>
<td>7 (46.67)</td>
<td>131 (23.99)</td>
<td>0.064(^d)</td>
</tr>
<tr>
<td>Without 9-repeat</td>
<td>8 (53.33)</td>
<td>415 (76.01)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>15</td>
<td>546</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** Values are n (%). \(^d\)Fisher’s exact test.
9-repeat. Our results provide what appears to be the first link between polymorphisms in the neck region of DC-SIGNR and malignant tumor risk.

The frequency of the 9-repeat is relatively high in a previously studied population from southern China (25%) and in the controls of the present population from the same region (24.6%, 138/561). However, it is rare in Caucasian (5%) and Indian (11%) populations. Similarly, NPC incidence is relatively high in southern China but low among Caucasians. These observations imply that the 9-repeat polymorphism may contribute to the risk of NPC and that the high incidence of this repeat in Guangxi may help explain why NPC is endemic in that region.

How the 9-repeat polymorphism in the DC-SIGNR neck region may affect NPC risk is unclear. Infection of B lymphocytes by EBV may play an important role in NPC onset. Given the high expression of DC-SIGNR on sinusoidal endothelial cells in the lymph node, where B-cell activation occurs, we propose that DC-SIGNR binds to glycans on the EBV surface and thereby traps EBV in the lymph nodes, where the virus interacts with cluster of differentiation 21 on the B-cell surface. The 9-repeat polymorphism may lead to tighter binding of EBV and therefore more efficient infection of B lymphocytes. The CRD binds multivalent ligands on pathogens with up to 7-fold greater affinity and selectivity when it is tetrameric than when it is monomeric. Shorter repeats appear to weaken the tetramer or prevent its formation altogether. Consistent with this hypothesis, the length of the neck region in DC-SIGNR appears to correlate with the protein’s ability to recognize and bind pathogens.

In addition to strengthening the CRD tetramer, the 9-repeat polymorphism may facilitate pathogen trapping in lymph nodes via a second mechanism. A longer neck region on DC-SIGNR has been proposed to distance the ligand-binding CRDs from potential glycans on the membrane of the host cell and bring the CRDs closer to glycans on the pathogen surface. Our results in the present study justify more detailed biochemical and structural analysis of how the 9-repeat version of DC-SIGNR may facilitate EBV infection of B lymphocytes, which may be a key driver of NPC onset.

Consistent with the ability of the 9-repeat polymorphism to facilitate EBV trapping, we found that the frequency of positive EBV-VCA-IgA was marginally higher among controls encoding the 9-repeat than among controls without the 9-repeat (P=0.064; Table 5). In addition, the 9-repeat is associated with increased HIV-RNA load among HIV-positive patients.

Limitations
The results of our study should be interpreted with caution in the light of its limitations. One is the relatively small sample from a single medical center, which may increase the risk of selection bias. Another limitation is the cross-sectional design of our study, which means that we cannot exclude the possibility that some of our controls will go on to develop NPC or other malignant cancers. Despite these limitations, our study provides the first data on the potential relationship between genetic variants in the DC-SIGNR neck region and NPC risk in a population from an endemic area.

Conclusion
Our case–control study suggests that polymorphism in the neck region of DC-SIGNR, but not DC-SIGN, may correlate with NPC onset. Further studies should verify and extend our findings by exploring the effects of the variable repeat regions on the structure and function of DC-SIGN and DC-SIGNR proteins.

Data availability
The raw data of our study can be obtained from the corresponding authors upon reasonable request.

Acknowledgments
This study was supported by the National Natural Science Foundation of China (81460415 and 81760361), the Guangxi Natural Science Foundation of China (2016GXNSFAA380096), and the Key Laboratory Foundation (GKE2018-07).

Author contributions
SN and MY performed most of the experiments, collected and analyzed all data, and contributed equally to this work. YW, XZ, CZ, and KY performed some of the experiments. ZW and YX designed the overall project. SN, ZW, and YX wrote the manuscript. All authors contributed toward data analysis, drafting and critically revising the paper, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

Disclosure
The authors report no conflicts of interest in this work.

References