Posted on 26 Aug 2024 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.172467183.32321143/v1 — This is a preprint and has not been peer-reviewed. Data may be preliminary.

Human Papillomavirus and Merkel Cell Polyomavirus in Korean Patients with Non-Small Cell Lung Cancer: Evaluation and Genetic Variability of the Non-Coding Control Region

Hyoung-Tae Jin[1], Yong-Sun Kim[1], and Eun-Kyoung Choi[1]

[1]Hallym University College of Medicine

August 26, 2024

Abstract

Human papillomavirus (HPV) is an important causative factor of cervical cancer and is associated with non-small cell lung cancer (NSCLC). Merkel cell polyomavirus (MCPyV) is a rare and highly fatal cutaneous virus that can cause Merkel cell carcinoma (MCC). Although coinfection with oncogenic HPV and MCPyV may increase cancer risk, a definitive etiological link has not been established. Recently, genomic variation and genetic diversity in the MCPyV noncoding control region (NCCR) among ethnic groups has been reported. The current study aimed to provide accurate prevalence information on HPV and MCPyV infection/coinfection in NSCLC patients and to evaluate and confirm Korean MCPyV NCCR variant genotypes and sequences. DNA from 150 NSCLC tissues and 150 adjacent control tissues was assessed via polymerase chain reaction (PCR) targeting regions of the large T antigen (LT-ag), viral capsid protein 1 (VP1), and NCCR. MCPyV was detected in 22.7% (34 of 150) of NSCLC tissues and 8.0% (12 of 150) of adjacent tissues from Korean patients. The incidence rates of HPV with and without MCPyV were 26.5% (9 of 34) and 12.9% (15 of 116). The MCPyV NCCR genotype prevalence in Korean patients was 21.3% (32 of 150) for subtype I and 6% (9 of 150) for subtype IIc. Subtype I, a predominant East Asian strain containing 25 bp tandem repeats, was most common in the MCPyV NCCR dataset. Our results confirm that coinfection with other tumor-associated viruses is not associated with NSCLC. Although the role of NCCR rearrangements in MCPyV infection remains unknown, future studies are warranted to determine the associations of MCPyV NCCR sequence rearrangements with specific diseases.

Human Papillomavirus and Merkel Cell Polyomavirus in Korean Patients with Non-Small Cell Lung Cancer: Evaluation and Genetic Variability of the Non-Coding Control Region

Hyoung-Tae Jin[a], Yong-Sun Kim[a, b] and Eun-Kyoung Choi [a, c*]

[a]*Ilsong Institute of Life Science, Hallym University, Yeongdeungpo-gu, Seoul, 07247, Korea*

[b]*Department of Microbiology, College of Medicine, Hallym University, Chuncheon, Gangwon-do 24252, Korea*

[c]*Department of Biomedical Gerontology, Graduate School of Hallym University, Chuncheon, Gangwon-do 24252, Korea*

*Corresponding authors: ekchoi@hallym.ac.kr

Eun-Kyoung Choi, Ph.D.

Ilsong Institute of Life Science, Hallym University, 55 Beodeunaru-ro, Yeongdeungpo-gu, Seoul 07247, Republic of Korea.

Tel.: +82-2-6923-8255, Fax: +82-2-6923-8203

Posted on 26 Aug 2024 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.172467183.32321143/v1 — This is a preprint and has not been peer-reviewed. Data may be preliminary.

E-mail: ekchoi@hallym.ac.kr

Full-length article word count – 3787

Abstract word count – 265

ABSTRACT

Human papillomavirus (HPV) is an important causative factor of cervical cancer and is associated with non-small cell lung cancer (NSCLC). Merkel cell polyomavirus (MCPyV) is a rare and highly fatal cutaneous virus that can cause Merkel cell carcinoma (MCC). Although coinfection with oncogenic HPV and MCPyV may increase cancer risk, a definitive etiological link has not been established. Recently, genomic variation and genetic diversity in the MCPyV noncoding control region (NCCR) among ethnic groups has been reported. The current study aimed to provide accurate prevalence information on HPV and MCPyV infection/coinfection in NSCLC patients and to evaluate and confirm Korean MCPyV NCCR variant genotypes and sequences. DNA from 150 NSCLC tissues and 150 adjacent control tissues was assessed via polymerase chain reaction (PCR) targeting regions of the large T antigen (LT-ag), viral capsid protein 1 (VP1), and NCCR. MCPyV was detected in 22.7% (34 of 150) of NSCLC tissues and 8.0% (12 of 150) of adjacent tissues from Korean patients. The incidence rates of HPV with and without MCPyV were 26.5% (9 of 34) and 12.9% (15 of 116). The MCPyV NCCR genotype prevalence in Korean patients was 21.3% (32 of 150) for subtype I and 6% (9 of 150) for subtype IIc. Subtype I, a predominant East Asian strain containing 25 bp tandem repeats, was most common in the MCPyV NCCR dataset. Our results confirm that coinfection with other tumor-associated viruses is not associated with NSCLC. Although the role of NCCR rearrangements in MCPyV infection remains unknown, future studies are warranted to determine the associations of MCPyV NCCR sequence rearrangements with specific diseases.

Keywords: Merkel cell polyomavirus, **Human papillomavirus** , Non-small cell lung cancer, Noncoding control region, Genetic variability, Asian/Koran genotype

1. INTRODUCTION

Smoking is a well-established risk factor for lung cancer, which is the leading cause of cancer-related deaths worldwide, including in Korea.[1,2] Notably, the incidence of lung cancer in women who have never smoked is on the rise globally, and there is an increasing prevalence of adenocarcinoma (AD) compared to squamous cell carcinoma (SCC).[3-5] Several studies have proposed that other etiological factors in addition to smoking contribute to the pathogenesis of lung cancer. These include infection with an oncogenic type of human papillomavirus (HPV), which is recognized as the primary cause of cervical cancer.[6-8] Rezazadeh et al.[9] reported an association between HPV infection and non-small cell lung cancer (NSCLC) (i.e., 20-25% of NSCLC patients).

Polyomaviruses, which belong to the family *Polyomaviridae* , are small nonenveloped DNA viruses that are widespread in nature and are suspected to be causative factors in various human tumors.[10] Moreover, infection by these viruses may lead to a variety of disorders when they persist for a lifetime and undergo reactivation in immunocompromised hosts.[11]According to Feng et al. (2008), Merkel cell polyomavirus (MCPyV), which is an oncogenic etiology of Merkel cell carcinoma (MCC), is rare, aggressive and directly linked to human skin cancer.[12]Because MCC and small cell lung cancer (SCLC) share similar histological characteristics, specifically the presence of small cells similar to those found in neuroendocrine carcinoma, studies have explored the incidence of MCPyV infection in both SCLC and NSCLC.[13-15] The MCPyV infection rate in NSCLC patients ranges from 4.7% to 38.5%.[15,16-18] However, the prevalence of MCPyV in control lung tissues and adjacent normal lung tissues is either lower than that found in PCR-positive lung tissues[18], or MCPyV is not detectable.[13,19] Although Joe et al.[20] investigated the correlation between HPV and MCPyV in NSCLC, a direct association between NSCLC and oncogenic HPV and MCPyV has not yet been clearly established.

The MCPyV genome is divided into three functional domains: an early gene region, including the large T antigen (LT), small T antigen (sT), 57kT antigen (57kT), alternative T antigen open reading frame (ALTO), and microRNA (miRNA); a noncoding control region (NCCR) containing the viral origin of replication (ORI); and late gene regions, including the capsid protein (VP1-3).[21] MCPyV genotypes associated with the early region, LT and sT regions and the late region encoding viral protein 1 (VP1) play important roles in viral genome replication, transcription and tumorigenesis,[22] and phylogenetic analyses based on these genes have enabled geographically related MCPyV genotype classification.[23] In general, the LT and VP1 regions have been well studied, whereas knowledge of the NCCR region is relatively poor. In the JC polyomavirus (JCV) and BK polyomavirus (BKV), rearranged NCCR variants are known to be associated with human disease,[24,25] but information on the MCPyV NCCR is limited.

Hashida et al.[26] classified five genotypes from people with different racial backgrounds and geographic locations depending on the occurrence of NCCR rearrangement. Additionally, NCCR rearrangement analysis of the NCCR strain type has been used to assess whether NCCR genetic variability in this virus has an impact on its pathogenicity.[27,28] Although it is possible to speculate on the correlation between genetic variability in NCCRs and disease, the pathogenesis and mechanisms underlying these mutations remain unknown.

Thus, the present study was designed to obtain additional data for use in resolving the inconsistencies and ambiguities discussed above and to explore aspects (i.e., analysis of the MCPyV NCCR) that have yet to be studied in Korea. The specific objectives of this study were as follows: to compare of the prevalence and frequency of MCPyV between NSCLC patients and adjacent nonlung cancer patients; to investigate the correlation between coinfection/infection with MCPyV and HPV in patients with NSCLC; and to analyze and compare specific genotypes in the Korean population through MCPyV DNA sequence analysis, particularly the classification of MCPyV NCCR genotypes and evaluation of MCPyV NCCR genetic variability, owing to the lack of MCPyV NCCR information.

2. MATERIALS AND METHODS

not-yet-known not-yet-known

not-yet-known

unknown

2.1 Study Design and Sample Preparation

NSCLC specimens (n=150) and corresponding adjacent nonlung cancer tissues (n=150 were collected from 3 medical centers: Korea University Guro Hospital (40 samples from 20 patients), Pusan University Hospital (140 samples from 70 patients), and Jeonbuk National University Hospital (120 samples from 60 patients). We also collected basic clinical data for the NSCLC patients, including age, sex, histological type, cancer stage and smoking status (Table 1). All frozen specimens were collected after informed consent was obtained, and this study was approved by the Ethics Committee of Hallym University of Korea (HIRB-2016-018-6-RMMMC-C-C).

not-yet-known not-yet-known

not-yet-known

unknown

2.3 Full-Length Sequences of MCPyV Full-length amplified products were sequenced using 13 overlapping primer sets and constructed based on the MCC350 sequence.12 PCR of the full-length sequence was accomplished with DiaStarTM Taq polymerase

2.4 DNA Sequence Analyses

Posted on 26 Aug 2024 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.172467183.32321143/v1 — This is a preprint and has not been peer-reviewed. Data may be preliminary.

PCR products were prepared for sequencing using a QIAquick Gel Extraction Kit (Qiagen). DNA sequencing of the purified PCR products was performed using a model ABI3730XL sequencer with a BigDye® Terminator v3.1 Cycle Sequencing Kit (ABI, Foster City, CA, USA). Viral MCPyV genotypes were identified by direct sequence analysis. The obtained DNA sequence data were BLAST searched at the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi). All MCPyV sequence alignments were performed using BioEdit version 7.1.9 (www.mbio.ncsu.edu/bioedit/bioedit.html) and compared with the six reference sequences of North American (MCC350, EU375803.1, MCC339, and EU375804.1), Japanese (TKS and FJ 464337), Swedish, French, Italian (MKL-1 and FJ173815), Chinese (HB039C and KC571692.1) and Korean (KIB and MK561422.1) MCPyV isolates.[12,31-34]

2.5 Statistical analysis

Statistical analyses of the data were performed using R software version 2.11.1. The formula for calculating the 95% confidence intervals of the proportions of molecular diagnoses was determined based on a binomial distribution. In all tests, P¡0.05 was considered to indicate statistical significance.

3. RESULTS

3.1 Prevalence of MCPyV in NSCLC and Adjacent Non-NSCLC Tissues

In the present study, MCPyV DNA was detected in 22.7% (34 of 150) of NSCLC tissues and 8.0% (12 of 150) of adjacent non-NSCLC tissues from Korean patients (Table 1 and Figure 1). Among the 150 patients enrolled, 80 had AD, 63 had SCC, and 7 had other NSCLCs (including pleomorphic carcinomas). Among the 80 ADs, 22 (27.5%) tested positive for MCPyV sequences; among the 63 SCCs, 10 (15.9%) tested positive. Notably, among the 34 positive-NSCLC samples, MCPyV DNA was detected in 64.7% (22 of 34) of the ADs and 29.4% (10 of 34) of the SCCs. Among the 12 positive-adjacent nonlung cancer samples, MCPyV DNA was detected in 91.7% (11 of 12) of the AD samples and 8.3% (1 of 12) of the SCC samples.

However, no significant differences in clinical outcomes related to viral detection, including age, sex, histology, stage, or smoking status, were detected among the clinicopathological parameters investigated (Table 1).

3.2 HPV detection in non-small cell lung cancer with MCPyV and non-MCPyV

We next investigated whether MCPyV-infection in NSCLC is associated with clinical outcomes in HPV-infected lung cancer patients. The prevalence of HPV-infection with MCPyV and non-MCPyV was 26.5% (9 of 34) and 12.9% (15 of 116), respectively (Table 2 and Supplementary Table S2). Among the 34 MCPyV-positive samples, HPV DNA sequences were detected in 6 (27.3%) of 22 ADs, in 2 (20%) of 10 SCCs and in 1 (50%) of 2 other NSCLC subtypes (Table 3). Among the 9 HPV-positive (with MCPyV) samples, eight were the high-risk type, namely, HPV-16; another high-risk type detected was HPV-33 (Table 2). Similarly, of 15 HPV-positive (with non-MCPyV) samples, HPV-16 (n=12) was the main type found, with the proportion of HPV-18 (n=3) being relatively small (Table 3 and Supplementary Table S2). There was no statistically significant difference in HPV DNA positivity between the MCPyV and non-MCPyV groups (P=0.06).

3.3 MCPyV DNA Sequence Analysis in Non-Small Cell Lung Cancer

Compared to the six representative reference strains, the Korean strains carry four specific mutations in LT-ag: a T> C substitution at position 1,501 in JL93; a T> C substitution at position 1,466; a G> A substitution at position 1,489; and a T> C substitution at position 1,572 in JL95 (Figure 2 and Supplementary Figure S1). In addition, the Korean strains showed two specific mutations in the VP1 gene: a T > G substitution at position 4,152 and a G > A substitution at position 4,153 in three samples (BL1, BL4, and BL28). At positions

Posted on 26 Aug 2024 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.172467183.32321143/v1 — This is a preprint and has not been peer-reviewed. Data may be preliminary.

4,362 and 4,368, two point mutations were found to be present in a mixed form: G > A (North American/European/TKS strains) and A > G (China/Korean strains) (Figure 3 and Supplementary Figure S2).

3.4 MCPyV NCCR Analyses of Non-Small Cell Lung Cancer from Korean Patients

To date, five genotypes have been reported. In this study, we examined tumor samples obtained from Korean lung cancer patients to determine the MCyV NCCR. The prevalence rates for MCPyV NCCR genotypes in Korean patients with NSCLC were 27.3% (41 of 150) for all identified genotypes, 21.3% (32 of 150) for subtype I, and 6% (9 of 150) for subtype IIc (Table 4 and Figure 4), with no other genotypes detected.

In this study, we found a few point mutations in the MCPyV NCCR: C > T at position 5,153 and G > A at position 5,168 in BL13; C > T at position 5,190, G > A at position 5,203 and C > T at position 5,238 in BL28; and A > G at position 5,239 in JL150 (Figure 5).

not-yet-known not-yet-known

not-yet-known

unknown

3.5 Sequencing of the Full-length MCPyV Genome from Non-Small Cell Lung Cancer

The complete MCPyV genome from MCPyV-positive NSCLCs was sequenced as described in a previous study (31). However, PCR amplification of full-length MCPyV genomes was successful for only one of the 34 MCPyV-positive samples, and the abovementioned Korean-specific mutations have not been reported in other strains. Analysis of the MCPyV sequence revealed it to be 98% to 99% identical to that of the MCV strains 350 and KIB. In addition, the full-length MCPyV genome is 5,417 bp long, the same length as that of KIB. The MCPyV genomes reported in this study have been deposited under GenBank accession number OQ616757.

CONFLICTS OF INTEREST STATEMENT

The authors have no conflicts of interest to declare.

ACKNOWLEDGMENTS

The biospecimens and data used in this study were obtained from the Biobanks of *Korea University Guro Hospital*, *Pusan National* University Hospital, *Jeonbuk National* University Hospital, and Chungbuk National University Hospital, which are members of the Korea Biobank Network.

Hosted file

`FIGURE 1.pptx` available at https://authorea.com/users/819917/articles/1218708-human-papillomavirus-and-merkel-cell-polyomavirus-in-korean-patients-with-non-small-cell-lung-cancer-evaluation-and-genetic-variability-of-the-non-coding-control-region

Hosted file

`FIGURE 2.pptx` available at https://authorea.com/users/819917/articles/1218708-human-papillomavirus-and-merkel-cell-polyomavirus-in-korean-patients-with-non-small-cell-lung-cancer-evaluation-and-genetic-variability-of-the-non-coding-control-region

Hosted file

`FIGURE 3.pptx` available at https://authorea.com/users/819917/articles/1218708-human-papillomavirus-and-merkel-cell-polyomavirus-in-korean-patients-with-non-small-cell-

Posted on 26 Aug 2024 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.172467183.32321143/v1 — This is a preprint and has not been peer-reviewed. Data may be preliminary.

Posted on 26 Aug 2024 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.172467183.32321143/v1 — This is a preprint and has not been peer-reviewed. Data may be preliminary.

lung-cancer-evaluation-and-genetic-variability-of-the-non-coding-control-region

Hosted file

FIGURE 4.pptx available at https://authorea.com/users/819917/articles/1218708-human-papillomavirus-and-merkel-cell-polyomavirus-in-korean-patients-with-non-small-cell-lung-cancer-evaluation-and-genetic-variability-of-the-non-coding-control-region

Hosted file

FIGURE 5.ppt available at https://authorea.com/users/819917/articles/1218708-human-papillomavirus-and-merkel-cell-polyomavirus-in-korean-patients-with-non-small-cell-lung-cancer-evaluation-and-genetic-variability-of-the-non-coding-control-region

Hosted file

FIGURE LEGENDS.docx available at https://authorea.com/users/819917/articles/1218708-human-papillomavirus-and-merkel-cell-polyomavirus-in-korean-patients-with-non-small-cell-lung-cancer-evaluation-and-genetic-variability-of-the-non-coding-control-region

Hosted file

TABLE 1.doc available at https://authorea.com/users/819917/articles/1218708-human-papillomavirus-and-merkel-cell-polyomavirus-in-korean-patients-with-non-small-cell-lung-cancer-evaluation-and-genetic-variability-of-the-non-coding-control-region

Hosted file

TABLE 2.doc available at https://authorea.com/users/819917/articles/1218708-human-papillomavirus-and-merkel-cell-polyomavirus-in-korean-patients-with-non-small-cell-lung-cancer-evaluation-and-genetic-variability-of-the-non-coding-control-region

Hosted file

TABLE 3.doc available at https://authorea.com/users/819917/articles/1218708-human-papillomavirus-and-merkel-cell-polyomavirus-in-korean-patients-with-non-small-cell-lung-cancer-evaluation-and-genetic-variability-of-the-non-coding-control-region

Hosted file

TABLE 4.doc available at https://authorea.com/users/819917/articles/1218708-human-papillomavirus-and-merkel-cell-polyomavirus-in-korean-patients-with-non-small-cell-lung-cancer-evaluation-and-genetic-variability-of-the-non-coding-control-region