FULL PAPER

BJC

British Journal of Cancer (2013) 108, 629–637 | doi: 10.1038/bjc.2012.567

Keywords: Merkel cell polyomavirus; lung cancer; integration; mutation; Japanese survey

Detection of Merkel cell polyomavirus with a tumour-specific signature in non-small cell lung cancer

Y Hashida[1], M Imajoh[1], Y Nemoto[1], M Kamioka[2], A Taniguchi[3], T Taguchi[4], M Kume[5], K Orihashi[5] and M Daibata*[,1]

[1]*Department of Microbiology and Infection, Kochi Medical School, Kochi University, Nankoku, Kochi 783-8505, Japan;* [2]*Department of Laboratory Medicine, Kochi Medical School, Kochi University, Nankoku, Kochi 783-8505, Japan;* [3]*Department of Hematology and Respiratory Medicine, Kochi Medical School, Kochi University, Nankoku, Kochi 783-8505, Japan;* [4]*Division of Human Health and Medical Science, Graduate School of Kuroshio Science, Kochi University, Nankoku, Kochi 783-8505, Japan and* [5]*Department of Surgery II, Kochi Medical School, Kochi University, Nankoku, Kochi 783-8505, Japan*

Background: We searched for a viral aetiology for non-small cell lung cancer (NSCLC), focusing on Merkel cell polyomavirus (MCPyV).

Methods: We analysed 112 Japanese cases of NSCLC for the presence of the MCPyV genome and the expressions of RNA transcripts and MCPyV-encoded antigen. We also conducted the first analysis of the molecular features of MCPyV in lung cancers.

Results: PCR revealed that 9 out of 32 squamous cell carcinomas (SCCs), 9 out of 45 adenocarcinomas (ACs), 1 out of 32 large-cell carcinomas, and 1 out of 3 pleomorphic carcinomas were positive for MCPyV DNA. Some MCPyV DNA-positive cancers expressed large T antigen (LT) RNA transcripts. Immunohistochemistry showed that MCPyV LT antigen was expressed in the tumour cells. The viral integration sites were identified in one SCC and one AC. One had both episomal and integrated/truncated forms. The other carried an integrated MCPyV genome with frameshift mutations in the *LT* gene.

Conclusion: We have demonstrated the expression of a viral oncoprotein, the presence of integrated MCPyV, and a truncated *LT* gene with a preserved retinoblastoma tumour-suppressor protein-binding domain in NSCLCs. Although the viral prevalence was low, the tumour-specific molecular signatures support the possibility that MCPyV is partly associated with the pathogenesis of NSCLC in a subset of patients.

Lung cancer remains the leading cause of cancer-related mortality worldwide. Although cigarette smoking is the predominant risk factor for the development of lung cancer, several studies have shown that the proportion of lung cancers diagnosed in never smokers has increased in recent times (Thun *et al*, 2006; Wakelee *et al*, 2007; Yano *et al*, 2008; Dela Cruz *et al*, 2011). Recent studies have also indicated that the incidence of lung cancer in never smokers seems to vary geographically, with greater prevalence in certain parts of the world, such as Asia (Thun *et al*, 2008; Yano *et al*, 2008; Dela Cruz *et al*, 2011). These findings suggest that factors other than smoking may be implicated in development of lung cancer.

Infectious agents have been proposed to contribute to the pathogenesis of lung cancer (Engels, 2008). Accumulating epidemiological evidence suggests that oncogenic viruses have important roles in the development of lung cancer. For instance, human papillomavirus (HPV) infection has been observed in association with cases of lung cancer in many studies worldwide, suggesting a role for HPV as an aetiological agent of lung cancer (Aguayo *et al*, 2007; Klein *et al*, 2009; Rezazadeh *et al*, 2009;

*Correspondence: Dr M Daibata; E-mail: daibatam@kochi-u.ac.jp

Received 29 August 2012; revised 29 October 2012; accepted 21 November 2012; published online 15 January 2013

© 2013 Cancer Research UK. All rights reserved 0007 – 0920/13

Prabhu *et al*, 2012). However, the results are discrepant among research groups (Brouchet *et al*, 2005) and underline some geographical variations (Miyagi *et al*, 2000; Aguayo *et al*, 2010; Goto *et al*, 2011).

In 2008, a new human tumour virus, designated Merkel cell polyomavirus (MCPyV), was reported in Merkel cell carcinomas (MCCs), a neuroendocrine carcinoma (NEC) of the skin (Feng *et al*, 2008). Recent studies have demonstrated the presence of MCPyV DNA in other types of tumours and in a subset of haematological malignancies, including chronic lymphocytic leukaemia (Dworkin *et al*, 2009; Kassem *et al*, 2009; Katano *et al*, 2009; Pantulu *et al*, 2010; Murakami *et al*, 2011; Imajoh *et al*, 2012a, b; Rollison *et al*, 2012). Although the route of MCPyV transmission has not been established, MCPyV DNA fragments have been detected in lower respiratory tract secretions (Babakir-Mina *et al*, 2010). If the respiratory tract is exposed continuously to this oncogenic virus, it is conceivable that cancer may develop in this region.

Based on the histological similarities between MCC and pulmonary NEC, including SCLC and large-cell NEC (LCNEC), the possible association between MCPyV and pulmonary NEC has been studied. Two German groups showed the presence of MCPyV DNA in a subset of SCLCs, with detection rates of 6.7% (2 out of 30) and 38.9% (7 out of 18; Andres *et al*, 2009; Helmbold *et al*, 2009). In contrast, no evidence of a causal relationship between MCPyV and pulmonary LCNEC has been reported (Busam *et al*, 2009; Duncavage *et al*, 2009; Schmitt *et al*, 2011). The prevalence of MCPyV in non-small cell lung cancers (NSCLCs) has not been investigated well. Recently, North and South American groups detected MCPyV DNA in 16.7% (5 out of 30) and 4.7% (4 out of 86) of NSCLCs, respectively, using PCR (Joh *et al*, 2010; Gheit *et al*, 2012). However, these epidemiological studies were all undertaken in Americans and Europeans, and the prevalence of MCPyV in lung cancers in an Asian cohort has not yet been reported. In this study, we investigated the possible viral aetiology of NSCLCs by examining a series of 112 Japanese patients for both the presence of specific MCPyV DNA and the expressions of viral RNA transcripts and virally encoded protein. We also looked for integrated/mutated forms of these MCPyV-positive NSCLCs. As far as we know, this is the first study to show the expression of MCPyV antigen and its integrated/mutated forms, which are recognised as tumour-specific signatures of MCPyV, in a specific cancer other than MCC.

MATERIALS AND METHODS

Patients and samples. This study included 112 Japanese patients with primary resectable NSCLC, including 32 squamous cell carcinomas (SCCs; denoted SCC1–32), 45 adenocarcinomas (AC1–45), and 32 large-cell carcinomas, including 19 LCNECs (LCC1–13 and LCNEC1–19) and 3 pleomorphic carcinomas (PL1–3; Table 1). The median ages of the patients were 71 years (range, 48–89) for SCC, 68 years (range, 36–86) for AC, 72 years (range, 60–83) for large-cell carcinoma, and 69 years (range, 66–73) for PL. There were 25 male and 7 female SCC patients, 24 male and 21 female AC patients, 28 male and 4 female large-cell carcinoma patients, and 3 male PL patients. Of the 112 lung cancer patients, 85 (75.9%) were smokers, including current and former smokers, and 27 (24.1%) were never smokers. All patients with NSCLC did not have apparent history of exposure to known or suspected lung carcinogens including asbestos, radon, and indoor air pollutants (Sun *et al*, 2007; Samet *et al*, 2009; Dela Cruz *et al*, 2011; Couraud, *et al*, 2012). The clinical stages of the patients are shown in Table 1. Surgically resected frozen tumour samples were collected from the archives of Kochi University Hospital. Frozen tumour specimens from eight cases of malignant mesothelioma and five cases of thymoma were also included in the study. This study was approved by the Ethics Committee of Kochi Medical School, Kochi University.

Table 1. Statistical associations between the presence of MCPyV DNA and the clinicopathological parameters

		MCPyV DNA (%)		
	All	Positive	Negative	*P*-value
	112			
Age				0.115
<65	21	1 (4.8)	20 (95.2)	
⩾65	91	19 (20.9)	72 (79.1)	
Sex				0.586
Male	80	13 (16.3)	67 (83.7)	
Female	32	7 (21.9)	25 (78.1)	
Histology				0.022[a]
SCC	32	9 (28.1)	23 (71.9)	
AC	45	9 (20.0)	36 (80.0)	
Large-cell carcinoma	32	1 (3.1)	31 (96.9)	
PL	3	1 (33.3)	2 (66.7)	
Staging				0.103
I–II	54	14 (25.9)	40 (74.1)	
III–IV	41	5 (12.2)	36 (87.8)	
Unknown	17	1 (5.9)	16 (94.1)	
Smoking status				0.085
Smokers	85	12 (14.1)	73 (85.9)	
Never smokers	27	8 (29.6)	19 (70.4)	

Abbreviations: AC = adenocarcinoma; MCPyV = Merkel cell polyomavirus; PL = pleomorphic carcinoma; SCC = squamous cell carcinoma.
[a] *P*-value indicates statistical significant correlation ($P<0.05$).

Detection of MCPyV with standard PCR. DNA was isolated with the DNeasy Blood and Tissue Kit (Qiagen, Tokyo, Japan). PCR was conducted with 200 ng of DNA using the AmpliTaq Gold 360 Master Mix (Life Technologies, Tokyo, Japan) and 0.4 μM each primer in a total volume of 50 μl. To detect the MCPyV large T antigen (*LT*) and viral protein 1 (*VP1*) genes, three primer sets, LT1, LT3, and VP1, were used as described previously (Feng *et al*, 2008). The β-globin gene (*HBB*) was amplified to confirm the presence of PCR-amplifiable DNA. The primer sequences are listed in Supplementary Table 1. Water containing all the PCR components except template DNA was used as the PCR-negative control. The reaction conditions included denaturation at 95°C for 10 min, followed by amplification with 40 cycles of 95°C for 30 s, 59°C for 30 s, and 72°C for 30 s, with a final extension for 7 min at 72°C. The PCR products were then separated electrophoretically on 2.0% agarose gels, visualised with ethidium bromide staining, and photographed. The PCR products were purified with the High Pure PCR Product Purification Kit (Roche Diagnostics, Tokyo, Japan) and then sequenced directly with the ABI Prism BigDye Terminator v1.1 Cycle Sequencing Kit (Life Technologies). The sequenced products were analysed using a model 3130 Genetic Analyser (Applied Biosystems, Tokyo, Japan). The nucleotide (nt) sequences obtained were aligned and edited with the BioEdit software (Ibis Biosciences, Carlsbad, CA, USA).

Quantitative real-time PCR. The standard PCR-positive samples were subjected to a further quantitative real-time PCR analysis. The reaction was performed in triplicate on a StepOnePlus thermocycler (Life Technologies). Primers and probes were prepared to amplify the MCPyV small T antigen gene (*ST*) and

the human RNase P gene (*RPP30*; Imajoh *et al*, 2012a). The reaction mixture was prepared in a total volume of 20 μl containing 200 ng of DNA, the TaqMan gene expression master mix (Life Technologies), 900 nM each primer (Supplementary Table 1), and 250 nM dual-labelled probe. The reaction conditions were 95°C for 10 min, followed by 40 cycles of 95°C for 15 s and 60°C for 1 min. A standard PCR was performed using the same primers and the PCR product was cloned into the pMD20-T vector (TaKaRa Bio, Shiga, Japan). We prepared 6-fold serial dilutions using 10 ng of the cloned plasmid DNA to generate a standard curve and from this, we calculated the copy number in each sample. The viral DNA load was defined as the viral DNA copies per RNase P gene copy, which represented the viral copy number per cell.

Reverse transcription–PCR. Total RNA was extracted from the frozen biopsy specimens with the High Pure RNA Tissue Kit (Roche Diagnostics). The total RNA was treated with DNase to avoid the amplification of viral DNA. The RNA (1 μg) was reverse transcribed using the SuperScript III First-Strand Synthesis System (Life Technologies). An aliquot (1 μl) of the reverse transcription (RT) reaction mixture was used for the subsequent PCR amplification. The primer sequences used to determine the *LT* and *VP1* gene expression are listed in Supplementary Table 1. The β-globin gene was amplified to confirm the presence of PCR-amplifiable cDNA.

Immunohistochemistry. To detecte MCPyV antigen, immunohistochemistry was performed on formalin-fixed paraffin-embedded tissue sections using a mouse monoclonal antibody, CM2B4 (IgG2b isotype), raised against exon 2 of the MCPyV LT antigen (Shuda *et al*, 2009). Samples sectioned at 4 μm were deparaffinised and rehydrated. Heat-induced epitope retrieval was performed using EDTA antigen retrieval buffer (pH 9.0). Endogenous peroxidase activity was blocked with 3% hydrogen peroxide. Following their incubation in blocking solution, the slides were incubated with the CM2B4 antibody diluted 1 : 100. After sufficient washes, horseradish peroxidase-conjugated goat anti-mouse immunoglobulin antibody was applied as the secondary antibody. After further washes, the binding of the primary antibody was detected with a biotin-free tyramide signal amplification system (Dako, Tokyo, Japan). Finally, the sections were counterstained with haematoxylin. The specificity of staining with CM2B4 was controlled by testing an isotype-matched control mouse IgG2b in parallel. The results of the immunohistochemical analysis of the tumour and adjacent non-neoplastic cells were scored as – (negative), 1+ (<10% immunoreactive cells), 2+ (10–50% immunoreactive cells), or 3+ (>50% immunoreactive cells).

Analysis of the MCPyV integration sites. The integration sites of MCPyV were investigated using the detection of integrated papilloma sequences (DIPS)–PCR technique, as described previously (Sastre-Garau *et al*, 2009). This method allows the amplification of the junctions between viral and cellular genomes (Luft *et al*, 2001). Briefly, after the DNA was digested with the *Taq* I restriction enzyme, the DNA fragments obtained were ligated to enzyme-specific adaptors. The ligated fragments were subjected to PCR amplification using viral- and adaptor-specific primers (Supplementary Table 1). The PCR products were purified and sequenced directly, as described above. The integration sites were determined by submitting the sequences to the databases of the National Center for Biotechnology Information and analysing them with the Basic Local Alignment Search Tool for genomic localisation.

Sequencing analysis of the MCPyV *LT* gene. The DNA sequences of the viral genomes from nt positions 151–3102 (based on Genbank sequence MCC350; EU375803), which include the entire *LT* gene, were determined by PCR using different combinations of six primer sets. This was followed by a direct sequence analysis of the amplified products. The primer sequences are listed in Supplementary Table 1.

Statistical analysis. The statistical correlations between the variables were analysed with Fisher's exact test. A P-value of <0.05 was considered significant.

RESULTS

Detection of the MCPyV genome by standard PCR and quantification of viral loads by real-time PCR. Specimens from 112 NSCLCs were screened for the presence of MCPyV DNA by standard PCR with three primer sets (LT1, LT3, and VP1) commonly used to detect the MCPyV *LT* and *VP1* genes (Feng *et al*, 2008). Table 2 summarises the clinical parameters for each PCR-positive patient, including the histological type of lung cancer, age, sex, clinical stage, and smoking status, and representative results are shown in Figure 1. Of the 32 SCCs, MCPyV DNA was detected in 3 tumours with the LT1 primers, 7 tumours with the LT3 primers, and 4 tumours with the VP1 primers. Of the 45 ACs, MCPyV was found in 7 tumours with the LT1 primers, 8 tumours with the LT3 primers, and 5 tumours with the VP1 primers. Of the 32 large-cell carcinomas, MCPyV DNA was detected in only one LCNEC with the LT3 primers. The LT3 and VP1 primer sets each detected the viral genome in one of three patients with PL. All the PCR-positive products were purified and sequenced, and we confirmed that they were MCPyV-specific DNA. The β-globin gene was amplified consistently in all the samples. Overall, nine SCCs (28.1%), nine ACs (20.0%), one LCNEC (3.1%), and one PL were positive for MCPyV when either the LT1, LT3, or VP1 primers were used. In total, 20 of 112 lung cancers (17.9%) harboured MCPyV DNA. Of these, six tumours (SCC15, SCC27, AC16, AC35, AC37, and AC39) contained MCPyV DNA sequences amplifiable by all three primer sets. None of the DNAs from malignant mesotheliomas and thymomas yielded MCPyV amplicons when these primer sets were used.

It was possible to obtain tumour tissues and the matching non-neoplastic tissues from our archive for three cases (SCC15, AC35, and AC39) and the DNAs were amplified in parallel, with the cycles that amplified the β-globin gene to plateau levels (Figure 1). Merkel cell polyomavirus genomes were only found in the tumour tissues of cases SCC15 and AC35, whereas the LT1 primers detected the MCPyV sequence in the non-neoplastic tissue of case AC39, although with only a faint band compared with the intensive bands found in the tumour tissues.

The MCPyV-DNA-positive samples were subjected to quantitative real-time PCR. The viral DNA loads ranged from 0.0001 to 0.026 copies per cell (Table 2).

The MCPyV DNA detection rates in SCCs and ACs were significantly higher than that in the large-cell carcinomas (Table 1). No significant association was found between the presence of MCPyV and the other clinicopathological features analysed. However, MCPyV was more frequently detected in never smokers than in smokers, although this difference was not statistically significant.

Expression of transcripts from the MCPyV *LT* gene. Specimens from 10 of the 20 MCPyV-DNA-positive tumours were suitable for RNA extraction. The expression of the MCPyV *LT* (nt positions 910–1152, corresponding to exon 2) and *VP1* (nt positions 3786–4137) transcripts was examined at the RNA level by RT–PCR (Figure 2). The cDNAs were amplified with the cycles that amplified the β-globin cDNA to the plateau level. Four of ten samples expressed the *LT* gene transcript, whereas no *VP1* gene transcript was found in any samples. The specific amplification of the *LT* gene transcripts was confirmed by direct sequencing.

Table 2. Summary of clinicopathology data and results of PCRs and immunohistochemistry in MCPyV-DNA-positive tumours

					PCR			Real-time PCR	RT—PCR		Immunohistochemistry	
Case	Age	Sex	Clinical stage	Smoking status	LT1	LT3	VP1	MCPyV DNA load (copies per cell: $\times 10^{-3}$)	LT-RNA	VP1-RNA	Tumour	Adjacent tissue
SCC3	73	F	IA	N	+	–	–	0.6	NT		2+	–
SCC15[a]	78	M	IB	F	+	+	+	26.0	+	–	3+	–
SCC17	74	M	IB	F	–	–	+	0.4	–	–	–	–
SCC21	71	M	IIA	C	–	+	–	NT	NT		2+	1+
SCC22	73	M	IB	N	–	+	–	0.1	–	–	–	–
SCC23	71	M	U	N	–	+	+	0.5	NT		NT	
SCC24	69	M	IV	F	–	+	–	NT	NT		2+	–
SCC27	75	F	IA	N	+	+	+	1.5	–	–	1+	1+
SCC32	83	M	IA	N	–	+	–	0.3	NT		–	–
AC7	78	M	IB	F	+	+	–	1.1	NT		2+	–
AC8	59	M	IIIA	F	+	–	+	0.6	NT		1+	1+
AC15	70	F	IIIB	N	–	+	–	0.2	+	–	1+	1+
AC16	74	F	IB	N	+	+	+	0.7	–	–	2+	1+
AC34	73	M	IIIA	F	–	+	–	0.8	NT		1+	1+
AC35	71	F	IIB	F	+	+	+	4.6	NT		3+	–
AC37	67	M	IIIA	F	+	+	+	0.7	–	–	1+	–
AC39	66	F	IIB	F	+	+	+	2.9	+	–	3+	1+
AC43[b]	81	F	IIB	N	+	+	–	0.5	+	–	3+	1+
LCNEC27	79	M	IIA	F	–	+	–	NT	NT		–	–
PL2	68	M	IIB	F	–	+	+	0.8	–	–	–	–

Abbreviations: AC = adenocarcinoma; C = current smokers; F = female; F = former smokers; LCNEC = large-cell neuroendocrine carcinoma; LT = large T; M = male; MCPyV = Merkel cell polyomavirus; N = never smokers; NT = not tested; PL = pleomorphic carcinoma; RT = reverse transcription; SCC = squamous cell carcinoma; U = unknown; VP = viral protein.
[a]MCPyV integration was found at chromosome 5q23.
[b]MCPyV integration was found at chromosome 11q25.

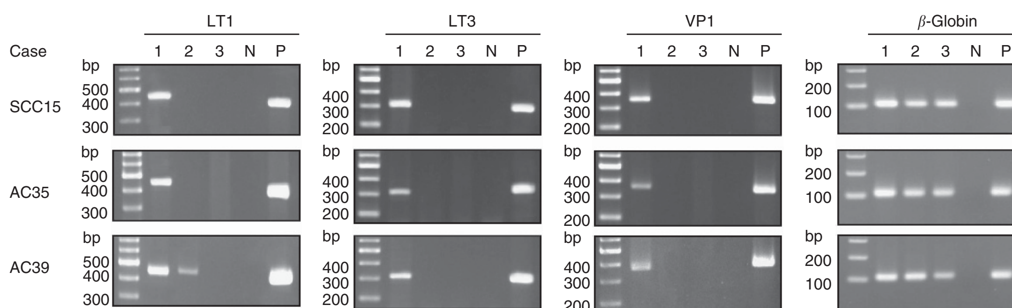

Figure 1. Representative results for the detection of MCPyV genomes by PCR. The LT1, LT3, and VP1 primers produced amplicons with sizes of 440, 309, and 352 bp, respectively. The DNA samples were subjected to amplification in parallel with the housekeeping gene β-globin (110 bp), which was detected in all samples. Lane 1: DNA extracted from tumours; lanes 2 and 3: DNAs extracted from two different sites in non-neoplastic tissues; lane N: water as the negative control; and lane P: positive control DNA from MCPyV-positive MCC tumour. Case numbers and molecular weight markers are shown on the left.

Expression of the MCPyV LT antigen. To examine the expression of the MCPyV LT antigen and to evaluate its localisation, immunohistochemistry was performed using the monoclonal antibody CM2B4 (Shuda *et al*, 2009). Strong diffuse or speckled strong nuclear signals were observed in the tumour cells, indicating that these cancer cells expressed the LT antigen. Although in some cases, weak staining was also observed in a small fraction of the adjacent non-neoplastic components, the tumour cells displayed stronger nuclear immunoreactivity. The semi-quantitative immunohistochemical findings are summarised in Table 2, and representative results are shown in Figure 3. The isotype-matched negative control antibody for CM2B4 showed no immunoreactivity in either the

tumour cells or the surrounding normal components, suggesting that nonspecific staining of the tissues is unlikely.

Integration of the MCPyV genome. The quantity and quality of the DNA are critical for the DIPS–PCR analysis, and the DNAs from 11 MCPyV-DNA-positive samples were suitable for this analysis. The integration site was identified in two cases (SCC15 and AC43) with primers designed to identify either the 3′ or 5′ virus–host junction. In case SCC15, the 5′ virus–host junction was located in the MCPyV *LT* gene at nt position 2738 and the viral DNA sequence was inserted into the long arm of chromosome 5 (5q23.1). In case AC43, the 3′ virus–host junction was located in the MCPyV *LT* gene at nt position 1801 and the viral DNA sequence was inserted into the long arm of chromosome 11 (11q25).

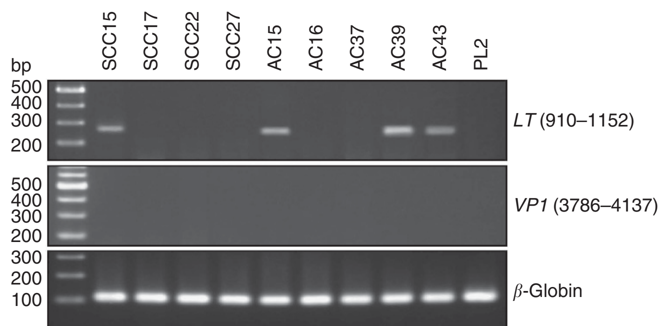

Figure 2. Expression of the *LT* and *VP1* gene transcripts. DNase-treated RNAs were reverse transcribed and the cDNAs were PCR amplified with primers complementary to the regions corresponding to the *LT* or *VP1* genes. All cDNAs were also subjected to amplification in parallel with the housekeeping gene β-globin, which was expressed at similar levels in all samples. Case numbers are indicated on the top. Molecular weight markers are shown on the left.

DNA sequencing analysis of the MCPyV *LT* gene. DNAs from four tumours (SCC15, AC35, AC39, and AC43) were subjected to a sequence analysis of the full-length *LT* gene at nt positions 151–3102. According to GeneBank data, the wild-type non-tumour-derived MCPyV strain Appendix206 (JN038578) has a lysine at amino-acid position 216 within a penta-amino-acid retinoblastoma tumour-suppressor protein (Rb)-binding motif (LFCDK) encoded by exon 2 of the *LT* gene, whereas its substitution with a glutamate (LFCDE) is found in the MCC-tumour-derived MCC350 strain (EU375803). The MCPyV strains found in our four patients were consistent with the MCC350 strain and contained the LxCxE motif (Figure 4). The psycho motif is also known to modulate Rb activities (White *et al*, 2009). Merkel cell polyomavirus has a unique psycho motif that is separated by a spacer region (amino-acid positions 83–102 and 210–218; Johnson, 2010). Several amino-acid differences between Appendix206 and our MCPyV strains were found in the spacer region. Of these, the MCPyV strains in our samples shared the same amino-acid substitution at position 112 (H→Y), and substitutions at positions 121 (S→P) and 156 (P→S) with the MCC350 strain.

The full-length *LT* gene sequence could be amplified for cases SCC15, AC35, and AC39. Although several non-synonymous mutations, resulting in amino-acid substitutions, were present at the C terminus of LT in cases SCC15 and AC35, no mutations causing stop codons were observed (Figure 4). In contrast, the *LT* gene in case AC43 had a frameshift mutation arising from a 46-base pair (bp) deletion at nt positions 1611–1656, which generated several stop codons. These mutations occurred downstream from the Rb-binding domain and caused a truncated exon 2, which encodes the LT helicase.

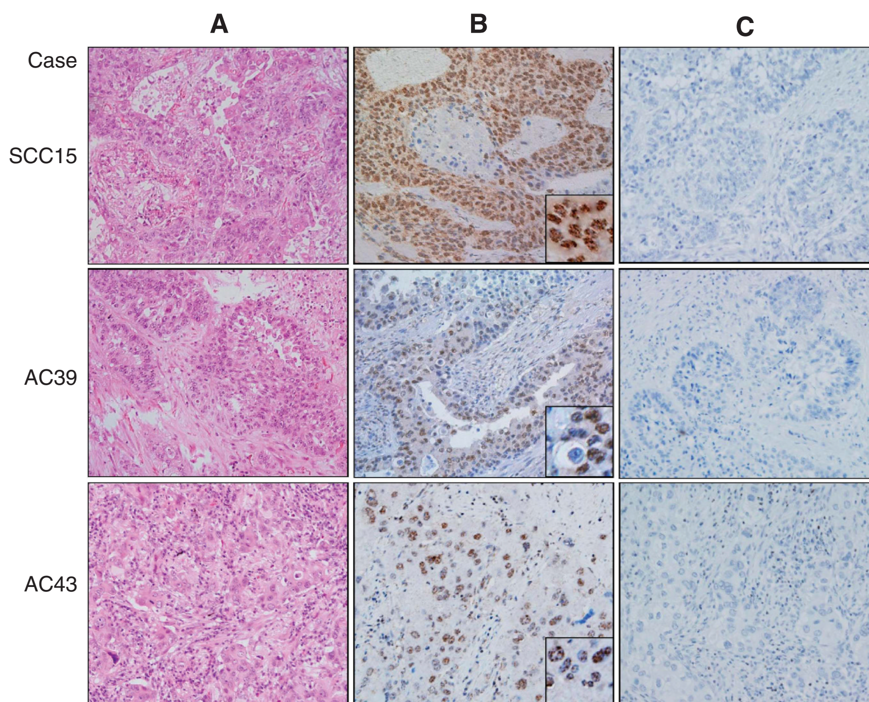

Figure 3. Immunohistochemical detection of the MCPyV LT antigen. (**A**) Haematoxylin and eosin staining of histological specimens containing tumour cells (original magnification ×200). (**B**) Immunohistochemical analysis with the CM2B4 monoclonal antibody, showing immunoreactivity in the tumour cells. Insets show higher magnification views of the tumour cells. (**C**) Immunohistochemical analysis with an isotype-matched negative control antibody (mouse IgG2b) for CM2B4, showing no immunoreactivity.

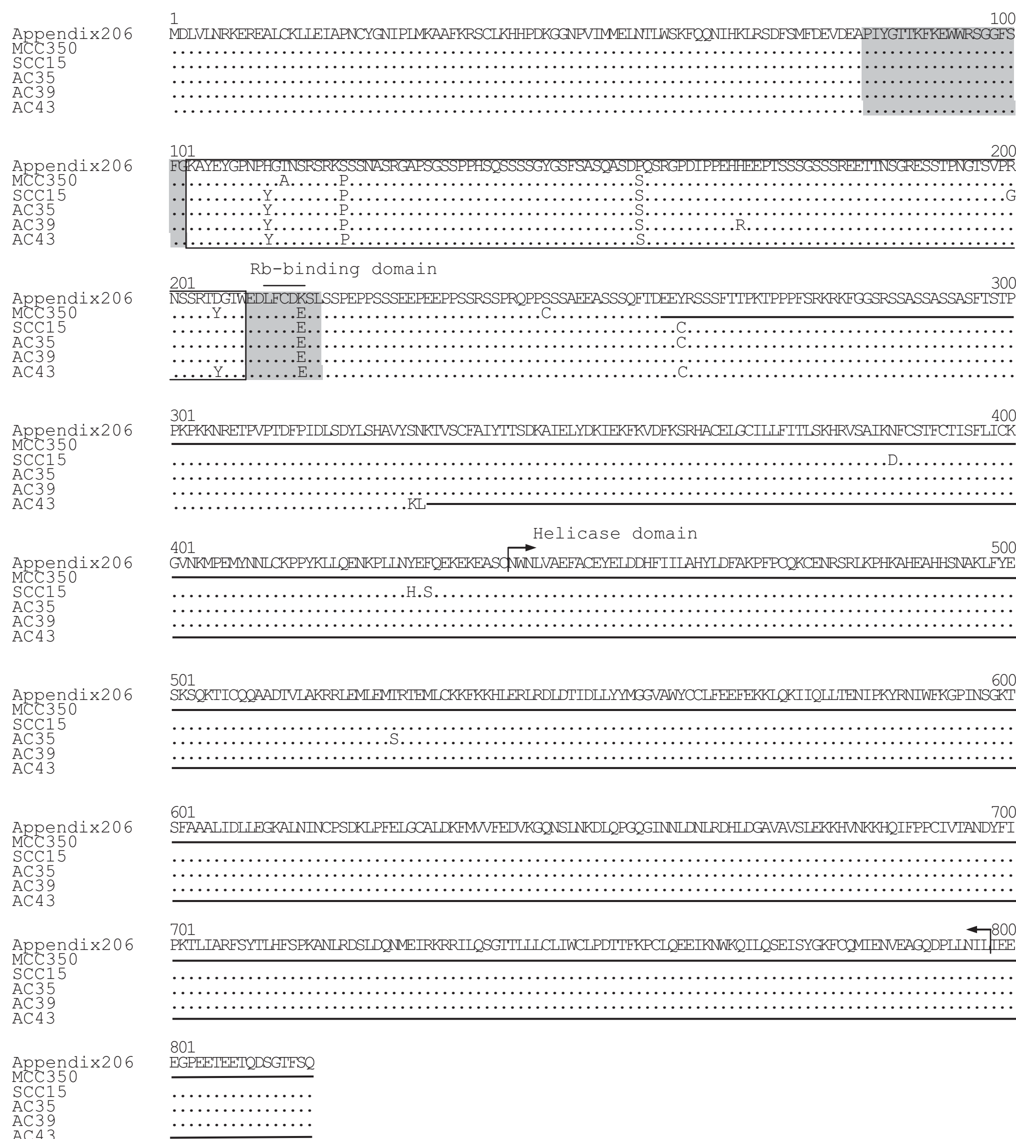

Figure 4. Amino-acid sequence alignment of the MCPyV LT antigen. The *LT* genes in the MCPyV-positive tumours were predicted to encode 817 amino-acid proteins. The amino-acid sequences were compared with reference sequences from the non-tumour-derived MCPyV isolate Appendix206 (GenBank accession number JN038578) and the MCC-tumour-derived isolate MCC350 (GenBank accession number EU375803). The position of the LxCxE motif, which is essential for Rb binding, is shown by the upper line. The psycho motif, which influences Rb activities, is represented with grey boxes. This motif, interrupted by amino-acids at positions 103–209, is shown in a white box with black lines. The arrows indicate the position of the helicase domain. The bold lines indicate the truncated regions of the LT antigen found in MCC350 and the MCPyV strain in tumour AC43. The numbers indicate the amino-acid positions.

DISCUSSION

This study provides the first evidence of the prevalence of MCPyV in NSCLCs in an Asian population. We used three primer sets (LT1, LT3, and VP1) that are commonly used to detect MCPyV by PCR. In good concordance with previous reports (Kassem *et al*, 2008), the LT3 primers showed the highest sensitivity. In total, 20 of 112 lung cancers (17.9%) harboured MCPyV DNA. Recently, a group detected MCPyV DNA in 5 of 30 (16.7%) NSCLCs from North America by PCR using the LT3 primers (Joh *et al*, 2010), which is a prevalence rate close to our result. Another report identified MCPyV DNA in 4 of 86 (4.7%) NSCLCs from Chilean subjects (Gheit *et al*, 2012). The differences in the detection rates may be explained in part by geographic epidemiological variations in patients with NSCLCs or merely by the technical approaches used. The absence of MCPyV we found in LCNECs and mesotheliomas is consistent with previous reports (Busam *et al*, 2009; Duncavage *et al*, 2009; Bhatia *et al*, 2010; Schmitt *et al*, 2011). Moreover, our results do not support a direct association between MCPyV and thymoma.

Smoking is the major risk factor for lung cancer. As expected, most subjects in our study group were smokers (75.9%), but one fourth of the patients were never smokers. In the United States and Europe, studies have estimated that $\sim$15–20% of lung cancers in women and 2–9% of lung cancers in men occur in never smokers (Boffetta *et al*, 2001; Wakelee *et al*, 2007), whereas in countries in East and South Asia, $\sim$60–80% of women and 10–15% of men with lung cancer are never smokers (Wakelee *et al*, 2007; Couraud *et al*, 2012). In this study, 68.8% (22 out of 32) of women and 6.3% (5 out of 80) of men with lung cancer were never smokers. Thus, smoking behaviour cannot fully explain the epidemiological

characteristics of lung cancer. Our findings show a trend toward a higher detection rate of MCPyV in never smokers than in smokers, although further studies are required to confirm this result. Similarly, other studies have reported higher incidences of HPV DNA in lung cancers from Asian cohorts, especially among never smokers (Cheng *et al*, 2001; Dela Cruz *et al*, 2011).

To define MCPyV as another infectious agent associated with lung cancer, MCPyV DNA positivity alone is not sufficiently specific to establish its aetiological role because MCPyV DNA fragments were detected in the lower respiratory tract when a high-sensitive PCR assay was used (Babakir-Mina *et al*, 2010). Several studies have demonstrated that the expression of the MCPyV LT antigen is essential for the oncogenesis of MCPyV-positive MCC (Shuda *et al*, 2008; Houben *et al*, 2010, 2012a; Sihto *et al*, 2011). In this context, we next investigated the expression of the *LT* gene at the RNA and protein levels. We found that 4 of 10 MCPyV-DNA-positive tumours had detectable levels of *LT* gene transcripts, whereas *VP1* gene transcripts were undetectable in all the samples analysed. During viral replication, MCPyV, like other polyomaviruses, displays an orderly gene expression cascade in which the *LT* gene transcript is expressed first (early gene transcription), followed by the expression of the *VP1* gene (late gene transcription; Pastrana *et al*, 2009; Feng *et al*, 2011). However, loss of the viral replication capacity is a common feature of virus-associated tumours (zur Hausen, 2008). In most MCPyV-positive MCCs, viral replication is hampered, and the *LT* gene but not the *VP1* gene is constitutively expressed (Feng *et al*, 2008; Shuda *et al*, 2008; Sastre-Garau *et al*, 2009). Based on this knowledge, it is plausible that our MCPyV strains that expressed the *LT* gene transcripts but not *VP1* gene transcripts had no viral replication activity. We used immunohistochemistry to study the expression and localisation of the MCPyV LT antigen. The localisation of strong immunoreactivity in nuclei of the tumour cells indicates the expression of the MCPyV LT antigen in the lung cancer cells.

Another oncogenic mechanism of MCPyV involves the binding capacity of the LT antigen to the tumour-suppressor protein Rb through the conserved LxCxE motif (Shuda *et al*, 2008). The LxCxE motif was preserved in the MCPyV strains we analysed, as has been observed in the MCC-tumour-derived isolate MCC350. Thus, the viruses, which had a conserved Rb-binding domain in the *LT* gene, may have a role in the tumorigenesis process. Our MCPyV strains also contained a conserved psycho motif, which modulates Rb activities (White *et al*, 2009). The psycho motif of MCPyV shows limited homology to those of other polyomavirus LT antigens and is uniquely interrupted by 107 amino acids (Johnson, 2010). Compared with the wild-type strain Appendix206, the spacer has some amino-acid substitutions, and two of these are shared by MCC350 and our MCPyV strains. The potential roles of these substitutions remain to be clarified.

The integration of MCPyV into the host cellular genome is also considered as a key element in oncogenesis. We next investigated the physical status of MCPyV using the DIPS–PCR technique. Merkel cell polyomavirus integration was demonstrated in two tumours (SCC15 and AC43), in which the viral genomes were inserted into chromosome 5 and 11, respectively. Although the chromosomal integration sites identified so far in MCPyV-positive MCCs differ from one another, one study has reported that chromosome 5 is a preferred integration site (Martel-Jantin *et al*, 2012). Importantly, the integration sites found in our NSCLC samples were localised within exon 2 of the *LT* gene, downstream from the Rb-binding domain. This confirms the results of previous studies in which most MCPyV-positive MCCs harboured the integrated viral genome at the 3′ end of the *LT* gene (Sastre-Garau *et al*, 2009; Laude *et al*, 2010). In tumour SCC15, the virus–host junction was located in the *LT* gene region (nt 2738), interrupting the helicase domain (nt positions 1947–3017), generating an integrated/truncated LT. Interestingly, the full-length *LT* gene sequence was amplified by direct PCR sequencing in this case and found to encode a non-truncated protein. This finding suggests the coexistence of a truncated/integrated form and an episomal form of MCPyV. This phenomenon has also been observed in some MCPyV-positive MCCs (Laude *et al*, 2010; Martel-Jantin *et al*, 2012). Tumour AC43 carried both an integrated MCPyV and frameshift mutations that preserved the Rb-binding domain but truncated the oncoprotein before the helicase domain, as in the tumour-specific molecular signatures observed in most MCPyV-positive MCCs (Shuda *et al*, 2008). Such truncations preserve the transformation capacity of LT through Rb sequestration, but eliminate viral DNA replication. In fact, our RT–PCR assay showed the expression of the *LT* gene but the absence of *VP1* gene expression, as discussed above.

Taken together, this study provides the first demonstration of not only the detection of MCPyV DNA but also the expressions of both *LT* RNA transcripts and LT antigen in NSCLCs. Furthermore, this study demonstrates the integrated/mutated forms of MCPyV for the first time in a specific cancer other than MCC. Viral integration and the acquisition of mutations within LT are common features of MCPyV-positive MCCs and are thought to be a prerequisite for tumorigenesis. Although the prevalence of MCPyV and the viral loads in our NSCLCs were low compared with those in MCPyV-positive MCCs, our findings suggest a potential role of MCPyV in the pathogenesis of NSCLC in a subset of patients. It can be hypothesised that the continuous exposure of the tissues of the lower respiratory tract to MCPyV may accidentally cause viral integration, together with the expression of the LT antigen, thereby leading to the oncogenic transformation of the infected host cells via MCPyV-associated tumorigenic pathways (Nikitin and Luftig, 2012). Alternatively, viral integration itself might indirectly promote additional cellular changes during the multistep carcinomagenesis process, which allow the outgrowth of the tumour. Recently, new evidence has been presented that only MCPyV is required for tumour initiation, but that additional cellular gene mutations during tumour progression render LT expression dispensable for MCC carcinogenesis (Houben *et al*, 2012b). At present, it is unknown whether the MCPyV DNA is lost in the majority of cells after the establishment of the tumour, which would explain the small number of MCPyV-positive NSCLCs. Another important question to be addressed is whether the prevalence and integration frequency of MCPyV in NSCLCs are influenced by ethnic factors and/or by exposure to risk factors such as smoking. Therefore, the factors that contribute to the infection and persistence of MCPyV in the lower airways must be defined. Our novel findings should stimulate further worldwide epidemiological and virological studies to determine the pathogenetic relevance of MCPyV in NSCLCs.

ACKNOWLEDGEMENTS

This work was supported by a Grant-in-Aid for Scientific Research from the Japanese Ministry of Education, Culture, Science, and Technology of Japan to MD.

REFERENCES

Aguayo F, Anwar M, Koriyama C, Castillo A, Sun Q, Morewaya J, Eizuru Y, Akiba S (2010) Human papillomavirus-16 presence and physical status in lung carcinomas from Asia. *Infect Agent Cancer* **5**: 20.

Aguayo F, Castillo A, Koriyama C, Higashi M, Itoh T, Capetillo M, Shuyama K, Corvalan A, Eizuru Y, Akiba S (2007) Human

papillomavirus-16 is integrated in lung carcinomas: a study in Chile. *Br J Cancer* **97**: 85–91.

Andres C, Ihrler S, Puchta U, Flaig MJ (2009) Merkel cell polyomavirus is prevalent in a subset of small cell lung cancer: a study of 31 patients. *Thorax* **64**: 1007–1008.

Babakir-Mina M, Ciccozzi M, Lo Presti A, Greco F, Perno CF, Ciotti M (2010) Identification of Merkel cell polyomavirus in the lower respiratory tract of Italian patients. *J Med Virol* **82**: 505–509.

Bhatia K, Modali R, Goedert JJ (2010) Merkel cell polyomavirus is not detected in mesotheliomas. *J Clin Virol* **47**: 196–198.

Boffetta P, Järvholm B, Brennan P, Nyrën O (2001) Incidence of lung cancer in a large cohort of non-smoking men from Sweden. *Int J Cancer* **94**: 591–593.

Brouchet L, Valmary S, Dahan M, Didier A, Galateau-Salle F, Brousset P, Degano B (2005) Detection of oncogenic virus genomes and gene products in lung carcinoma. *Br J Cancer* **92**: 743–746.

Busam KJ, Jungbluth AA, Rekthman N, Coit D, Pulitzer M, Bini J, Arora R, Hanson NC, Tassello JA, Frosina D, Moore P, Chang Y (2009) Merkel cell polyomavirus expression in merkel cell carcinomas and its absence in combined tumors and pulmonary neuroendocrine carcinomas. *Am J Surg Pathol* **33**: 1378–1385.

Cheng YW, Chiou HL, Sheu GT, Hsieh LL, Chen JT, Chen CY, Su JM, Lee H (2001) The association of human papillomavirus 16/18 infection with lung cancer among nonsmoking Taiwanese women. *Cancer Res* **61**: 2799–2803.

Couraud S, Zalcman G, Milleron B, Morin F, Souquet PJ (2012) Lung cancer in never smokers–a review. *Eur J Cancer* **48**: 1299–1311.

Dela Cruz CS, Tanoue LT, Matthay RA (2011) Lung cancer: epidemiology, etiology, and prevention. *Clin Chest Med* **32**: 605–644.

Duncavage EJ, Le BM, Wang D, Pfeifer JD (2009) Merkel cell polyomavirus: a specific marker for Merkel cell carcinoma in histologically similar tumors. *Am J Surg Pathol* **33**: 1771–1777.

Dworkin AM, Tseng SY, Allain DC, Iwenofu OH, Peters SB, Toland AE (2009) Merkel cell polyomavirus in cutaneous squamous cell carcinoma of immunocompetent individuals. *J Invest Dermatol* **129**: 2868–2874.

Engels EA (2008) Inflammation in the development of lung cancer: epidemiological evidence. *Expert Rev Anticancer Ther* **8**: 605–615.

Feng H, Kwun HJ, Liu X, Gjoerup O, Stolz DB, Chang Y, Moore PS (2011) Cellular and viral factors regulating Merkel cell polyomavirus replication. *PLoS One* **6**: e22468.

Feng H, Shuda M, Chang Y, Moore PS (2008) Clonal integration of a polyomavirus in human Merkel cell carcinoma. *Science* **319**: 1096–1100.

Gheit T, Muñoz JP, Levican J, González C, Ampuero S, Parra B, Gaggero A, Corvalán AH, Meneses M, Tommasino M, Aguayo F (2012) Merkel cell polyomavirus in non-small cell lung carcinomas from Chile. *Exp Mol Pathol* **93**: 162–166.

Goto A, Li CP, Ota S, Niki T, Ohtsuki Y, Kitajima S, Yonezawa S, Koriyama C, Akiba S, Uchima H, Lin YM, Yeh KT, Koh JS, Kim CW, Kwon KY, Nga ME, Fukayama M (2011) Human papillomavirus infection in lung and esophageal cancers: analysis of 485 Asian cases. *J Med Virol* **83**: 1383–1390.

Helmbold P, Lahtz C, Herpel E, Schnabel PA, Dammann RH (2009) Frequent hypermethylation of RASSF1A tumour suppressor gene promoter and presence of Merkel cell polyomavirus in small cell lung cancer. *Eur J Cancer* **45**: 2207–2211.

Houben R, Adam C, Baeurle A, Hesbacher S, Grimm J, Angermeyer S, Henzel K, Hauser S, Elling R, Bröcker EB, Gaubatz S, Becker JC, Schrama D (2012a) An intact retinoblastoma protein-binding site in Merkel cell polyomavirus large T antigen is required for promoting growth of Merkel cell carcinoma cells. *Int J Cancer* **130**: 847–856.

Houben R, Grimm J, Willmes C, Weinkam R, Becker JC, Schrama D (2012b) Merkel cell carcinoma and Merkel cell polyomavirus: evidence for hit-and-run oncogenesis. *J Invest Dermatol* **132**: 254–256.

Houben R, Shuda M, Weinkam R, Schrama D, Feng H, Chang Y, Moore PS, Becker JC (2010) Merkel cell polyomavirus-infected Merkel cell carcinoma cells require expression of viral T antigens. *J Virol* **84**: 7064–7072.

Imajoh M, Hashida Y, Nemoto Y, Oguri H, Maeda N, Furihata M, Fukaya T, Daibata M (2012a) Detection of Merkel cell polyomavirus in cervical squamous cell carcinomas and adenocarcinomas from Japanese patients. *Virol J* **9**: 154.

Imajoh M, Hashida Y, Taniguchi A, Kamioka M, Daibata M (2012b) Novel human polyomaviruses, Merkel cell polyomavirus and human polyomavirus 9, in Japanese chronic lymphocytic leukemia cases. *J Hematol Oncol* **5**: 25.

Joh J, Jenson AB, Moore GD, Rezazedeh A, Slone SP, Ghim S, Kloecker GH (2010) Human papillomavirus (HPV) and Merkel cell polyomavirus (MCPyV) in non small cell lung cancer. *Exp Mol Pathol* **89**: 222–226.

Johnson EM (2010) Structural evaluation of new human polyomaviruses provides clues to pathobiology. *Trends Microbiol* **18**: 215–223.

Kassem A, Schöpflin A, Diaz C, Weyers W, Stickeler E, Werner M, Zur Hausen A (2008) Frequent detection of Merkel cell polyomavirus in human Merkel cell carcinomas and identification of a unique deletion in the VP1 gene. *Cancer Res* **68**: 5009–5013.

Kassem A, Technau K, Kurz AK, Pantulu D, Löning M, Kayser G, Stickeler E, Weyers W, Diaz C, Werner M, Nashan D, zur Hausen A (2009) Merkel cell polyomavirus sequences are frequently detected in nonmelanoma skin cancer of immunosuppressed patients. *Int J Cancer* **125**: 356–361.

Katano H, Ito H, Suzuki Y, Nakamura T, Sato Y, Tsuji T, Matsuo K, Nakagawa H, Sata T (2009) Detection of Merkel cell polyomavirus in Merkel cell carcinoma and Kaposi's sarcoma. *J Med Virol* **81**: 1951–1958.

Klein F, Amin Kotb WF, Petersen I (2009) Incidence of human papilloma virus in lung cancer. *Lung Cancer* **65**: 13–18.

Laude HC, Jonchère B, Maubec E, Carlotti A, Marinho E, Couturaud B, Peter M, Sastre-Garau X, Avril MF, Dupin N, Rozenberg F (2010) Distinct merkel cell polyomavirus molecular features in tumour and non tumour specimens from patients with Merkel cell carcinoma. *PLoS Pathog* **6**: e1001076.

Luft F, Klaes R, Nees M, Dürst M, Heilmann V, Melsheimer P, von Knebel Doeberitz M (2001) Detection of integrated papillomavirus sequences by ligation-mediated PCR (DIPS-PCR) and molecular characterization in cervical cancer cells. *Int J Cancer* **92**: 9–17.

Martel-Jantin C, Filippone C, Cassar O, Peter M, Tomasic G, Vielh P, Brière J, Petrella T, Aubriot-Lorton MH, Mortier L, Jouvion G, Sastre-Garau X, Robert C, Gessain A (2012) Genetic variability and integration of Merkel cell polyomavirus in Merkel cell carcinoma. *Virology* **426**: 134–142.

Miyagi J, Tsuhako K, Kinjo T, Iwamasa T, Hirayasu T (2000) Recent striking changes in histological differentiation and rate of human papillomavirus infection in squamous cell carcinoma of the lung in Okinawa, a subtropical island in southern Japan. *J Clin Pathol* **53**: 676–684.

Murakami M, Imajoh M, Ikawa T, Nakajima H, Kamioka M, Nemoto Y, Ujihara T, Uchiyama J, Matsuzaki S, Sano S, Daibata M (2011) Presence of Merkel cell polyomavirus in Japanese cutaneous squamous cell carcinoma. *J Clin Virol* **50**: 37–41.

Nikitin PA, Luftig MA (2012) The DNA damage response in viral-induced cellular transformation. *Br J Cancer* **106**: 429–435.

Pantulu ND, Pallasch CP, Kurz AK, Kassem A, Frenzel L, Sodenkamp S, Kvasnicka HM, Wendtner CM, zur Hausen A (2010) Detection of a novel truncating Merkel cell polyomavirus large T antigen deletion in chronic lymphocytic leukemia cells. *Blood* **116**: 5280–5284.

Pastrana DV, Tolstov YL, Becker JC, Moore PS, Chang Y, Buck CB (2009) Quantitation of human seroresponsiveness to Merkel cell polyomavirus. *PLoS Pathog* **5**: e1000578.

Prabhu PR, Jayalekshmi D, Pillai MR (2012) Lung cancer and human papilloma viruses (HPVs): examining the molecular evidence. *J Oncol* **2012**: 750270.

Rezazadeh A, Laber DA, Ghim SJ, Jenson AB, Kloecker G (2009) The role of human papilloma virus in lung cancer: a review of the evidence. *Am J Med Sci* **338**: 64–67.

Rollison DE, Giuliano AR, Messina JL, Fenske NA, Cherpelis BS, Sondak VK, Roetzheim RG, Iannacone MR, Michael KM, Gheit T, Waterboer T, Tommasino M, Pawlita M (2012) Case-control study of Merkel cell polyomavirus infection and cutaneous squamous cell carcinoma. *Cancer Epidemiol Biomarkers Prev* **21**: 74–81.

Samet JM, Avila-Tang E, Boffetta P, Hannan LM, Olivo-Marston S, Thun MJ, Rudin CM (2009) Lung cancer in never smokers: clinical epidemiology and environmental risk factors. *Clin Cancer Res* **15**: 5626–5645.

Sastre-Garau X, Peter M, Avril MF, Laude H, Couturier J, Rozenberg F, Almeida A, Boitier F, Carlotti A, Couturaud B, Dupin N (2009) Merkel cell carcinoma of the skin: pathological and molecular evidence for a causative role of MCV in oncogenesis. *J Pathol* **218**: 48–56.

Schmitt M, Höfler D, Koleganova N, Pawlita M (2011) Human polyomaviruses and other human viruses in neuroendocrine tumors. *Cancer Epidemiol Biomarkers Prev* **20**: 1558–1561.

Shuda M, Arora R, Kwun HJ, Feng H, Sarid R, Fernández-Figueras MT, Tolstov Y, Gjoerup O, Mansukhani MM, Swerdlow SH, Chaudhary PM, Kirkwood JM, Nalesnik MA, Kant JA, Weiss LM, Moore PS, Chang Y

(2009) Human Merkel cell polyomavirus infection I. MCV T antigen expression in Merkel cell carcinoma, lymphoid tissues and lymphoid tumors. *Int J Cancer* **125**: 1243–1249.

Shuda M, Feng H, Kwun HJ, Rosen ST, Gjoerup O, Moore PS, Chang Y (2008) T antigen mutations are a human tumor-specific signature for Merkel cell polyomavirus. *Proc Natl Acad Sci USA* **105**: 16272–16277.

Sihto H, Kukko H, Koljonen V, Sankila R, Böhling T, Joensuu H (2011) Merkel cell polyomavirus infection, large T antigen, retinoblastoma protein and outcome in Merkel cell carcinoma. *Clin Cancer Res* **17**: 4806–4813.

Sun S, Schiller JH, Gazdar AF (2007) Lung cancer in never smokers–a different disease. *Nat Rev Cancer* **7**: 778–790.

Thun MJ, Hannan LM, Adams-Campbell LL, Boffetta P, Buring JE, Feskanich D, Flanders WD, Jee SH, Katanoda K, Kolonel LN, Lee IM, Marugame T, Palmer JR, Riboli E, Sobue T, Avila-Tang E, Wilkens LR, Samet JM (2008) Lung cancer occurrence in never-smokers: an analysis of 13 cohorts and 22 cancer registry studies. *PLoS Med* **5**: e185.

Thun MJ, Henley SJ, Burns D, Jemal A, Shanks TG, Calle EE (2006) Lung cancer death rates in lifelong nonsmokers. *J Natl Cancer Inst* **98**: 691–699.

Wakelee HA, Chang ET, Gomez SL, Keegan TH, Feskanich D, Clarke CA, Holmberg L, Yong LC, Kolonel LN, Gould MK, West DW (2007) Lung cancer incidence in never smokers. *J Clin Oncol* **25**: 472–478.

White MK, Johnson EM, Khalili K (2009) Multiple roles for Puralpha in cellular and viral regulation. *Cell Cycle* **8**: 1–7.

Yano T, Miura N, Takenaka T, Haro A, Okazaki H, Ohba T, Kouso H, Kometani T, Shoji F, Maehara Y (2008) Never-somoking nonsmall cell lung cancer as a separate entity: clinicopathologic features and survival. *Cancer* **113**: 1012–1018.

zur Hausen H (2008) A specific signature of Merkel cell polyomavirus persistence in human cancer cells. *Proc Natl Acad Sci USA* **105**: 16063–16064.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported License.

Supplementary Information accompanies this paper on British Journal of Cancer website (http://www.nature.com/bjc)