Genetic polymorphisms of the human cytomegalovirus \textit{UL144} gene in colorectal cancer and its association with clinical outcome

Hsin-Pai Chen,1,2 Jeng-Kai Jiang,2,3 Chia-Hao Chan,4 Wan-Huai Teo,5 Chih-Yung Yang,6 Yen-Chung Chen,7 Teh-Ying Chou,8,9 Chi-Hung Lin10 and Yu-Jiun Chan1,4,11

Correspondence
Yu-Jiun Chan
yjchan@vghtpe.gov.tw

1Division of Infectious Diseases, Department of Medicine, Taipei Veterans General Hospital, Taipei, Taiwan
2School of Medicine, National Yang-Ming University, Taipei, Taiwan
3Division of Colon and Rectal Surgery, Department of Surgery, Taipei Veterans General Hospital, Taipei, Taiwan
4Institute of Public Health, School of Medicine, National Yang-Ming University, Taipei, Taiwan
5Department of Biotechnology and Laboratory Science in Medicine, School of Biomedical Science and Engineering, National Yang-Ming University, Taipei, Taiwan
6Department of Education and Research, Taipei City Hospital, Taipei, Taiwan
7Department of Pathology, National Yang-Ming University Hospital, Yilan, Taiwan
8Institute of Clinical Medicine, National Yang-Ming University, Taipei, Taiwan
9Division of Molecular Pathology, Department of Pathology and Laboratory Medicine, Taipei Veterans General Hospital, Taipei, Taiwan
10Institute of Microbiology and Immunology, School of Life Science, National Yang-Ming University, Taipei, Taiwan
11Division of Microbiology, Department of Pathology and Laboratory Medicine, Taipei Veterans General Hospital, Taipei, Taiwan

Human cytomegalovirus (HCMV) has been increasingly detected in colorectal cancer (CRC), and genetic polymorphisms in HCMV affect its pathogenesis. This study aimed to investigate HCMV genetic polymorphisms in CRC and its correlation with the clinical outcomes. We performed PCR and sequencing of a viral immunomodulatory gene, \textit{UL144}, in clinical isolates and CRC specimens. The nucleotide and amino acid sequences were aligned, and a phylogenetic tree was constructed. The clinical, pathological and survival data were compared among tumours with different \textit{UL144} genotypes. HCMV was detected in 49 (47.8 \%) of the tumour specimens. Genotype A predominated in 43 samples (22/43; 51.2 \%) with successful sequencing, followed by genotype B (13/43; 30.2 \%) and genotype C (8/43; 18.6 \%). The genotypic distribution was similar to that of the clinical isolates and those reported in other Asian populations. The amino acid sequence of genotype B was the most conserved. For stage II and III CRC patients with HCMV-positive tumours, disease-free survival (DFS) varied among the three major genotypes ($P=0.0046$). The presence of genotype B virus in the tumours was associated with a shorter DFS and independently predicted tumour recurrence in a multivariate Cox proportional hazards model (hazard ratio, 5.79; 95 \% confidence interval, 1.30–25.81; $P=0.021$). By reverse transcription PCR, tumour samples with genotype B viruses had the highest rate of \textit{UL144} expression. Our results suggest that genetic polymorphisms of HCMV \textit{UL144} are associated with clinical outcome in CRC and that HCMV may play an immunomodulatory role in the tumour microenvironment of CRC.
INTRODUCTION

Human cytomegalovirus (HCMV) is a herpesvirus that chronically infects its hosts. In the majority of immunocompetent hosts, the virus maintains a lifelong latency and is capable of modulating multiple immune and cellular functions, many of which are associated with oncogenesis (Castillo & Kowalik, 2002; Cinatl et al., 1996, 2004). HCMV has been detected in a number of human cancers, including breast cancer, cervical carcinoma, prostate adenocarcinoma, malignant glioma, Epstein–Barr virus-negative Hodgkin’s disease and Kaposi’s sarcoma (Cinatl et al., 2004). Recent evidence suggests that HCMV may play an oncomodulatory role in human cancers (Michaelis et al., 2011). In breast cancer, HCMV proteins have been found in the sentinel lymph nodes, raising the question of whether HCMV contributes to the metastasis (Taher et al., 2013). A very high prevalence of HCMV in both primary and metastatic tumours was found in breast and colon cancers with brain metastasis, while high-level expression of HCMV immediate-early proteins in tumours was associated with a poorer outcome (Taher et al., 2014). In malignant glioma and neuroblastoma, treatment with antiviral agents was associated with a favourable outcome (Söderberg-Naucler et al., 2013; Stragliotto et al., 2013; Wolmer-Solberg et al., 2013). All these data indicate that HCMV may play an important role in cancer behaviour and progression.

An increasing number of studies suggest that HCMV may be associated with colorectal cancer (CRC) (Chen et al., 2012; Harks et al., 2002). HCMV DNA has been detected in more than 40 % of CRC tumour specimens (Chen et al., 2012). The viral nucleic acids specifically localized to the cytoplasm of the neoplastic epithelium. In elderly CRC patients, we found that the tumoural presence of HCMV was associated with a decreased disease-free survival (DFS) and disparate inflammatory cytokine pathways (Chen et al., 2014), suggesting that HCMV may play an oncomodulatory role in the pathogenesis of CRC.

In clinical HCMV isolates, considerable genetic polymorphism exists and correlates with its clinical presentation (Pignatelli et al., 2010; Wu et al., 2011). Of the viral genes, UL144 is particularly notable because of its role in modulating the host immune response. UL144 is found exclusively in clinical HCMV strains and encodes a structural homologue of the herpesvirus entry mediator. It has been proposed that UL144 plays a role in virus-mediated immune evasion by transmitting inhibitory signals to downregulate T-cell responses (Benedict et al., 1999; Watts & Gommerman, 2005). This study aimed to investigate HCMV genetic polymorphisms in CRC tumour tissue and their correlation with clinical outcome.

RESULTS

Genotypic distribution of HCMV UL144 in tumour specimens of CRC

A total of 137 clinical isolates were collected. Among them, 131 were successfully PCR amplified and sequenced. Along with 34 reference sequences, three major genotypic clusters were identified in phylogenetic analysis.

The phylogenetic tree indicated that genotype B was a distinctive type, while genotypes A and C were more closely related (Fig. 1a). After excluding 15 duplicate isolates from the same patients, the frequency distribution of the remaining 116 isolates was 47.4 % (55/116) for genotype A, 29.3 % (34/116) for genotype B and 23.3 % (27/116) for genotype C (Table 1).

To analyse HCMV genotypes in CRC, 230 paired tumour and adjacent non-neoplastic specimens from 115 CRC patients were collected. Of these, 49 (42.6 %) tumour and 13 (11.3 %) adjacent non-neoplastic specimens were positive for HCMV by PCR. Sequencing was successful in 43 samples, but failed in six because of inadequate PCR products. Phylogenetic analysis also identified three major genotypic clusters (Fig. 1b). Genotype A was the most common in CRC (21/43; 48.8 %; Table 1), followed by genotype B (13/43; 30.2 %) and genotype C (8/43; 18.6 %). No recombinant types were identified in the CRC tumour samples. The genotypic distribution of UL144 in CRC was similar to that of the HCMV clinical strains in Taipei Veterans General Hospital (VGHTPE). There were seven pairs of specimens whose tumour and adjacent non-neoplastic specimens were both positive for HCMV; genotype B predominated in these samples (5/7; 71.4 %).

Sequence variation of amino acids in different types of tumour-associated HCMV

The predicted sequence of the UL144 protein contained 176 aa. There were four functional domains: the cysteine-rich domain (CRD) 1 from Cys25 to Cys56, CRD2 from Cys59 to Cys95, the transmembrane domain from Ser138 to Ala153, and the cytoplasmic tail between Ala156 and Leu176 (Lurain et al., 1999). In CRC samples, alignment of the amino acid sequences revealed that sequence variation in the UL144 gene resulted in substitution of 36.9 % (65/176) of the amino acids, with 92–95 % of substitutions occurring in the ectodomains CRD1 and CRD2 but particularly in CRD1 (Fig. S1, available in the online Supplementary Material). The extent of protein polymorphism varied significantly among the UL144 genotypes. The amino acid sequences of genotype B were the most conserved, with only 5.1 % variation in the amino acid sequences. Genotype C was the most polymorphic, with 27.8 % variation, while genotype A had 12.5 % variation.

Comparison of the CRC UL144 genotypes with clinical strains from different geographical regions

As the genotypic distribution of HCMV varies among different geographical regions (Pignatelli et al., 2003), we compared the CRC UL144 genotypes with those reported previously. In both the VGHTPE clinical isolates and CRC
Fig. 1. Neighbour-joining tree of human cytomegalovirus UL144 nucleotide sequences of (a) clinical isolates and (b) colorectal cancer specimens, containing 34 reference strains that are denoted by their GenBank accession numbers. *Bootstrap value >70%. Bars, nucleotide substitutions per site.
specimens, the genotypic distribution was significantly different from those of the American or European isolates (Table 1) (Arav-Boger et al., 2002, 2006; Bale et al., 2001; Heo et al., 2008; Lurain et al., 1999; Picone et al., 2005), where genotype B was predominant. In Asia, the genotypic distribution of the Japanese population was similar to that of the European population (Murayama et al., 2005; Tanaka et al., 2005; Yan et al., 2008), while the genotypic distribution of the Chinese population resembled the results of this study (Mao et al., 2007).

**Table 1.** Comparison of the genotypic distribution of HCMV UL144 from different sources and geographical regions

<table>
<thead>
<tr>
<th>Sample source</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>AB</th>
<th>AC</th>
<th>1b</th>
<th>Total</th>
<th>T</th>
<th>P (vs CRC)</th>
<th>P (vs clinical isolates)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical isolates (Taiwan)</td>
<td>55</td>
<td>34</td>
<td>27</td>
<td>116</td>
<td>0.387</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRC samples (Taiwan)</td>
<td>22</td>
<td>13</td>
<td>8</td>
<td>43</td>
<td>NA</td>
<td>0.387</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999 Lurain (USA)</td>
<td>10</td>
<td>23</td>
<td>4</td>
<td>14</td>
<td>4</td>
<td>45</td>
<td>0.008</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001 Bale (USA)</td>
<td>2</td>
<td>34</td>
<td>5</td>
<td>1</td>
<td>6</td>
<td>48</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002 Arav-Boger (USA)</td>
<td>2</td>
<td>51</td>
<td>14</td>
<td>8</td>
<td>1</td>
<td>76</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008 Heo (USA)</td>
<td>13</td>
<td>23</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>45</td>
<td>0.067</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005 Picone (France)</td>
<td>20</td>
<td>24</td>
<td>10</td>
<td>2</td>
<td>2</td>
<td>68</td>
<td>0.144</td>
<td>0.023</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006 Arav-Boger (Italy)</td>
<td>19</td>
<td>25</td>
<td>7</td>
<td>1</td>
<td>3</td>
<td>55</td>
<td>0.18</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005 Murayama (Japan)</td>
<td>15</td>
<td>21</td>
<td>5</td>
<td>1</td>
<td>42</td>
<td>0.234</td>
<td>0.024</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005 Tanaka (Japan)</td>
<td>21</td>
<td>24</td>
<td>7</td>
<td>3</td>
<td>52</td>
<td>0.319</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007 Mao (China)</td>
<td>33</td>
<td>15</td>
<td>19</td>
<td>3</td>
<td>70</td>
<td>0.274</td>
<td>0.099</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008 Yan (Japan)</td>
<td>24</td>
<td>25</td>
<td>10</td>
<td>59</td>
<td>0.428</td>
<td>0.211</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Presence of the HCMV UL144 genotype B was associated with a shorter DFS in stage II and III patients

Regardless of the tumoural presence of HCMV, patients with stage I disease had a favourable outcome, while stage IV patients had a poor outcome. For patients with stage II or III disease, a trend for unfavourable DFS was observed in those with HCMV-positive tumours, although this result was not statistically significant (P=0.178; Fig. 2a). Elderly patients aged ≥65 years, with stage II or III disease and HCMV-positive tumours, had a lower DFS rate (P=0.004; Fig. 2b). The DFS rates were not significantly different in patients aged <65 years (P=0.311; Fig. 2c).

The outcomes of stage II and III patients with HCMV-positive tumours were analysed, and the Kaplan–Meier curves revealed that DFS varied significantly according to the UL144 genotype. Genotype B virus in the tumours was associated with the lowest DFS rate, whereas those with genotype C viruses had a favourable outcome (P=0.0046; Fig. 3a). The DFS curves for each genotype were compared with patients whose tumours were negative for HCMV. This showed that patients with non-B genotypes were similar to those with HCMV-negative tumours, while the presence of genotype B virus in the tumours denoted a poorer outcome (Fig. 3b).

**Fig. 2.** Kaplan–Meier curves of DFS for all patients (n=71) (a), elderly patients (≥65 years; n=40) (b) and non-elderly patients (<65 years; n=31) (c) with stage II and III CRCs.
Fig. 3. Kaplan–Meier curves of DFS for stage II and III patients with HCMV-positive tumours (n=30). (a) Comparison of survival for the three genotypes. (b) Survival of each genotype compared with patients with HCMV-negative tumours.

Table 2. Demographic characteristics and underlying diseases of CRC patients with genotype B or non-genotype B HCMV in tumours

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Genotype B (n=12)</th>
<th>Genotype B (n=6)</th>
<th>P</th>
<th>Genotype B (n=29)</th>
<th>Genotype B (n=24)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male sex</td>
<td>2 (58.3)</td>
<td>4 (66.7)</td>
<td>0.47</td>
<td>21 (72.4)</td>
<td>17 (70.8)</td>
<td>1.00</td>
</tr>
<tr>
<td>Age (mean years ± SD)</td>
<td>68.6 ± 12.4</td>
<td>77.2 ± 6.7</td>
<td>0.76</td>
<td>69.9 ± 11.1</td>
<td>68.9 ± 11.8</td>
<td>0.04</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>4 (33.3)</td>
<td>3 (50.0)</td>
<td>0.05</td>
<td>2 (6.9)</td>
<td>2 (8.3)</td>
<td>0.04</td>
</tr>
<tr>
<td>Hypertension</td>
<td>3 (25.0)</td>
<td>3 (50.0)</td>
<td>1.00</td>
<td>8 (27.6)</td>
<td>6 (25.0)</td>
<td>0.33</td>
</tr>
<tr>
<td>Heart diseases</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>1.00</td>
<td>4 (13.8)</td>
<td>4 (16.7)</td>
<td>0.56</td>
</tr>
<tr>
<td>Lung diseases</td>
<td>3 (25.0)</td>
<td>0 (0.0)</td>
<td>0.45</td>
<td>3 (5.8)</td>
<td>1 (4.2)</td>
<td>1.00</td>
</tr>
<tr>
<td>Renal diseases</td>
<td>2 (4.7)</td>
<td>0 (0.0)</td>
<td>0.45</td>
<td>3 (5.8)</td>
<td>2 (8.3)</td>
<td>1.00</td>
</tr>
<tr>
<td>Liver diseases</td>
<td>1 (2.3)</td>
<td>1 (16.7)</td>
<td>0.37</td>
<td>0 (0.0)</td>
<td>1 (4.2)</td>
<td>0.17</td>
</tr>
<tr>
<td>Central nervous system diseases</td>
<td>5 (11.6)</td>
<td>5 (11.6)</td>
<td>1.00</td>
<td>3 (5.8)</td>
<td>2 (8.3)</td>
<td>0.041</td>
</tr>
<tr>
<td>History of polyp or CRC</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>1.00</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>1.00</td>
</tr>
<tr>
<td>History of other malignancy</td>
<td>1 (8.3)</td>
<td>1 (16.7)</td>
<td>0.16</td>
<td>4 (13.8)</td>
<td>1 (4.2)</td>
<td>0.17</td>
</tr>
<tr>
<td>HCMV in adjacent normal tissue</td>
<td>5 (41.7)</td>
<td>3 (50.0)</td>
<td>0.041</td>
<td>2 (6.9)</td>
<td>2 (8.3)</td>
<td>0.041</td>
</tr>
<tr>
<td>Ascending colon cancer</td>
<td>5 (41.7)</td>
<td>3 (50.0)</td>
<td>0.64</td>
<td>10 (34.5)</td>
<td>8 (33.3)</td>
<td>0.64</td>
</tr>
<tr>
<td>Rectal cancer</td>
<td>5 (41.7)</td>
<td>2 (33.3)</td>
<td>0.57</td>
<td>6 (20.7)</td>
<td>4 (16.7)</td>
<td>0.57</td>
</tr>
<tr>
<td>Poorly differentiated tumour</td>
<td>1 (3.4)</td>
<td>1 (16.7)</td>
<td>0.37</td>
<td>1 (8.3)</td>
<td>1 (4.2)</td>
<td>0.37</td>
</tr>
<tr>
<td>Significant inflammation in tumour</td>
<td>3 (25.0)</td>
<td>1 (16.7)</td>
<td>0.63</td>
<td>11 (37.9)</td>
<td>9 (37.5)</td>
<td>0.63</td>
</tr>
<tr>
<td>Pathological signs of early metastasis: VELIPI*</td>
<td>2 (6.9)</td>
<td>1 (16.7)</td>
<td>0.50</td>
<td>2 (16.7)</td>
<td>2 (8.3)</td>
<td>0.50</td>
</tr>
<tr>
<td>Advanced local invasion (T4)</td>
<td>1 (8.3)</td>
<td>1 (16.7)</td>
<td>0.37</td>
<td>1 (3.4)</td>
<td>1 (4.2)</td>
<td>0.37</td>
</tr>
<tr>
<td>Lymph node metastasis</td>
<td>4 (33.3)</td>
<td>3 (50.0)</td>
<td>1.00</td>
<td>12 (41.4)</td>
<td>10 (41.7)</td>
<td>1.00</td>
</tr>
<tr>
<td>Distant metastasis (stage IV)</td>
<td>2 (16.7)</td>
<td>2 (16.7)</td>
<td>NA</td>
<td>2 (6.9)</td>
<td>2 (8.3)</td>
<td>NA</td>
</tr>
</tbody>
</table>

*VELIPI denotes pathological features of early metastatic invasion, including vascular emboli, lymphatic invasion, and perineural invasion, alone or in combination.
UL144 genotype B independently predicts tumour recurrence in stage II and III patients

Patients were divided into genotype B and non-genotype B groups according to their tumour UL144 genotypes. We found that those with genotype B HCMV-positive tumours were older, more frequently diabetic and had a higher percentage of positive HCMV detection in both the tumour and adjacent non-neoplastic specimens (Table 2). There were no significant differences between the other clinical and pathological characteristics.

Univariate and multivariate Cox proportional hazards models were used to assess the risk of cancer recurrence in patients with stage II or III disease. Age was transformed into quartiles. Of the risk factors analysed, the presence of genotype B HCMV in the tumour independently predicted disease recurrence after surgery (Table 3), with a hazard ratio (HR) of 5.79 [95 % confidence interval (CI), 1.30–25.81; P=0.021]. Another independent predictor of disease recurrence was the presence of ‘VELIPI’, a pathological feature of early metastasis that includes vascular emboli, lymphatic invasion, and perineural invasion of the primary tumour (HR, 9.15; 95 % CI, 1.63–51.44; P=0.012).

UL144 is expressed most frequently in tumour samples with genotype B HCMV

We then investigated whether a type-specific gene expression of UL144 occurred in tumour tissues of CRC. RNAs were obtained from another 32 CRC specimens positive for HCMV. Among the three genotypes, CRC samples with genotype B virus had the highest rate of detectable UL144 RNA, although the expression levels of UL144 in CRC were similar among the three genotypes (Fig. 4).

A recent study has shown that the expression of UL144 in latently infected myeloid cells depends on the presence of GATA-2 transcription factor binding elements in the UL144 promoter region (Poole et al., 2013). In the 37 CRC samples analysed, 21 (56.8 %) lacked the predicted GATA-2 binding elements in the promoter region of UL144. A single GATA-2 binding site was identified in the remaining 16 (44.2 %) samples, including 10 (52.6 %) of the 19 genotype A, two (18.2 %) of the 11 genotype B, and four (57.1 %) of the seven genotype C samples. The presence or absence of the GATA-2 binding element was not correlated with the mRNA expression of UL144 (Table S1).

Table 3. Cox proportional hazards models for prediction of disease recurrence in stage II and III CRC patients

<table>
<thead>
<tr>
<th>Variable</th>
<th>Univariate Cox model</th>
<th>Multivariate Cox model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR</td>
<td>95 % CI</td>
</tr>
<tr>
<td>Genotype B in tumour (yes/no)</td>
<td>6.86</td>
<td>1.65–28.48</td>
</tr>
<tr>
<td>Age (quartile)</td>
<td>2.30</td>
<td>1.09–4.85</td>
</tr>
<tr>
<td>VELIPI (yes/no)*</td>
<td>11.13</td>
<td>2.18–56.97</td>
</tr>
</tbody>
</table>

*VELIPI denotes pathological features of early metastatic invasion, including vascular emboli, lymphatic invasion, and perineural invasion, alone or in combination.

Fig. 4. (a) Numbers of samples with detectable UL144 expression. (b) Relative expression levels of UL144 in three genotypes. The expression level of UL144 was normalized to that of the glyceraldehyde 3-phosphate dehydrogenase (GAPDH) gene in each sample.
DISCUSSION

The methods used to detect HCMV in cancer tissues have led to some controversial results. The specimen type, selection of the viral gene and primer design all critically influence the end result of PCR-based methods (Ranganathan et al., 2012). By combining PCR and gene sequencing, the results of the present study have provided unequivocal evidence of the existence of HCMV in CRC tumour tissues. This pairwise study of tumour and adjacent non-neoplastic specimens has confirmed reports that HCMV preferentially exists in the tumour tissue of CRC (Chen et al., 2012; Dimberg et al., 2013; Harkins et al., 2002).

The HCMV genetic polymorphism in CRC and its correlation with cancer outcome have never been investigated before. In the present study, for the first time to the best of our knowledge, we have identified the HCMV genotypic distribution in CRC tumour tissues. We found a similar genotypic distribution between the HCMV strains in the CRC and clinical isolates, suggesting that no specific selection occurred during viral entry into the tumour tissue. In immune-competent individuals, HCMV is carried latently in peripheral blood monocytes with limited viral gene expression (Sinclair & Sissons, 2006). The virus may circulate through blood to the CRC tumour tissue, where a suppressed immunity in the tumour microenvironment provides an opportunity for the virus to reactivate and subsist therein (Evans et al., 2006).

The genotypic distributions of HCMV in the CRC specimens and clinical isolates of our Taiwanese patients were different from those of American and European populations but were similar to the other Asian populations. The genetic polymorphisms of the clinical HCMV strains varied widely among the different geographical regions and may influence the results of studies on HCMV-related diseases (Chen et al., 2008; Pignatelli et al., 2003; Zipeto et al., 1998). The genotypic distribution of our Taiwanese patients was similar to that of the Chinese population, as they are ethnically and geographically closely related. It is possible that specific viral strains have circulated between these two populations as they are in close contact with each other; alternatively, host–pathogen co-evolution may have led to the selection of specific virus genotypes based on host susceptibility.

We were surprised to find that the survival rates varied considerably among the CRC UL144 genotypes. Of the three major genotypes, genotype B was phylogenetically distinctive, more frequently present in both the tumour and adjacent normal colon tissues, and had the most conserved amino acid sequences. UL144 has been found to bind the Ig superfamily member B- and T-lymphocyte attenuator (BTLA) on T-cells and transmits signals to inhibit T-cell proliferation (Cheung et al., 2005). In congenital HCMV infections, UL144 genotype B was associated with fewer pregnancy terminations, while genotype C was associated with loss of the fetus (Arav-Boger et al., 2006). These results seem to contradict the present study. Nevertheless, as UL144 plays a role in immune evasion, a genotype that is the most ‘immune tolerant’ may be beneficial in preserving the fetus but may be detrimental to the host considering the cancer pathophysiology. Expression of UL144 may assist the tumour to escape from anti-tumour immune surveillance of T-cells, which is now recognized as one of the major determinants of tumour recurrence and patient survival (Galon et al., 2006; Pages et al., 2005).

The next question is how different UL144 genotypes differentially influence the anti-tumour immunity and subsequently patients’ outcome. Comparison of the amino acid sequences revealed that CRD1 was highly variable among the three genotypes. CRD1 is the proposed interaction site between UL144 and BTLA (Watts & Gommerman, 2005). One may speculate that viruses of different UL144 genotypes encode proteins with differential binding affinities to BTLA on T-cells. However, the finding that UL144 from diverse clinical isolates maintains similar BTLA binding affinities despite the sequence variation in CRD1 argues against this speculation (Cheung et al., 2005). Another possibility is that the expression of UL144 varies among different genotypes. A recent study indicated that UL144 could be differentially expressed depending on the transcription factor binding sites in its promoter (Poole et al., 2013). It is noteworthy that the frequency of positive UL144 expression in CRC samples varied among genotypes, and UL144 was most frequently expressed in samples with genotype B viruses. We speculate that, in CRCs with genotype B viruses, UL144-mediated immune suppression occurs more frequently, contributes to immunosuppression of the tumour microenvironment, and subsequently leads to tumour recurrence.

The findings of the present study are novel and may inform future research in cancer pathophysiology. Nevertheless, there are some limitations to this study, one of which is the small cohort size, which led to wide CIs and imprecise estimates of HRs. Further large-scale studies are needed to confirm the findings of the present study. In addition, a causal inference could not be made on the association between genotypes and the outcome. There is the possibility that the tumour microenvironment was more favourable for specific genotypes of HCMV. Finally, we were unable to determine the genotype in six CRC samples because of inadequate PCR amplification products. It is not known whether these samples contained new genotypes or if they had any effect on the survival outcome.

In conclusion, the genotypic distribution of HCMV UL144 in CRC was similar to that of the clinical isolates. The tumoural presence of the genotype B virus was associated with a shorter DFS and independently predicted tumour recurrence in stage II and III patients. Further research is warranted to clarify whether UL144 plays an immunomodulatory role in the tumour microenvironment of CRC.

METHODS

Study population, specimens and data collection. This study was approved by the Institutional Review Board of VGHTPE. After
informed consent was obtained from all the patients. CRC specimens were collected and stored in the bank of residual surgical tissues at the VGH TPE Division of Colorectal Surgery, as described previously (Chen et al., 2012). Tumour and adjacent non-neoplastic specimens of CRC were randomly retrieved for PCR, and haematoxylin-and-eosin-stained specimen slides were re-examined by a pathologist to confirm the diagnosis.

Demographic, clinical and pathological data were extracted from the CRC database of the VGH TPE Division of Colorectal Surgery. These data had been recorded prospectively with regular updates, and included patient demographic characteristics, underlying diseases, surgical details, tumour and node metastasis stages, important pathological features, conditions associated with tumour recurrence and survival at the last follow-up. Patients were observed between the interval of cancer diagnosis and the last follow-up or death. Data were censored at the last follow-up for patients who had not shown relapse and for those who had died. For patients with tumour relapse, DFS was defined as the period from the date of surgery to the date of confirmed tumour relapse. For patients without detectable tumour relapse, DFS was defined as the period from the date of surgery to the date of the last follow-up.

**PCR of HCMV UL144.** Viral and host DNA were extracted using a commercial kit (QiAamp DNA Mini kit; Qiagen) according to the manufacturer’s instructions. For PCR, approximately 200 μl of the clinical isolate virus suspension or 20–25 mg frozen CRC tissue was used. Primers (sense, 5’-CTGATTACAAGCGGAGA-3′; antisense, 5’-CTGTGGTGCACCTGTAACG-3′) were designed to amplify the entire UL144 gene. The 20 μl PCR mixture contained 1 × reaction buffer, 500 μM dNTP mix, 1.5 mM Mg2+, 300 nM primer mix, 0.025 U Taq DNA polymerase (Invitrogen) and 2 μl DNA extract. The PCR was carried out with the following protocol: 95 °C for 1.5 min, 55 °C for 2 min and 72 °C for 1 min for 35 cycles, with an initial cycle of 95 °C for 5 min and a final cycle of 72 °C for 7 min.

**DNA sequencing, alignment and phylogenetic analysis.** The PCR products were sequenced by the National Yang-Ming University VYM Genome Research Centre. The nucleotide sequences were aligned and edited using BioEdit (version 7.0.7.1), and then translated into the amino acid sequences. Phylogenetic analysis was conducted with the PHYLIP (version 3.68) and MEGA4 programs. Neighbour-joining was used to reconstruct a tree with 1000 bootstraps. The neighbour-joining tree was built using the Kimura two-parameter model. A phylogenetic tree was reconstructed using UL144 reference sequences randomly selected from GenBank. Bootstrap values ≥70 % were considered definitive for significant clustering.

**Quantitative reverse transcription PCR (qRT-PCR) of UL144 and analysis of UL144 promoters.** Total RNA was extracted from CRC specimens with TRizol reagent (Invitrogen) according to the manufacturer’s protocol. Reverse transcription was carried out using 2 μl 10 × reaction buffer, 2 μl 10 × random hexamers, 0.8 μl 100 mM dNTP mix, 1 μl MultiScribe reverse transcriptase (Life Technology), 1 μl RNase Inhibitor (Applied Biosystems), 2 μg sample RNA and DEPC H2O to a total volume of 10 μl. Reverse transcription was carried out with the following protocol: 25 °C for 10 min, 37 °C for 120 min, heat inactivated at 85 °C for 5 min, and then chilled on ice. The cDNA was stored at −20 °C until use. The quantitative PCR mixture contained 12.5 μl 2 × SYBR Master Mix (Applied Biosystems), 2 μl cDNA, 0.25 μl primer mix (containing 0.1 μM each primer μl−1) and 23 μl H2O. Quantitative PCR was carried out with the following protocol: one cycle of 50 °C for 2 min, one cycle of 95 °C for 10 min, 40 cycles of 95 °C for 15 s, 60 °C for 30 s and 72 °C for 30 s, and a final extension at 72 °C for 10 min. The primers for the UL144 qRT-PCR were designed as follows: forward primer, 5′-AGGCC GTCCAACATCACAAGC-3′, and reverse primer, 5′-CGACCACCTTTCCTTGTTTG-3′. The primers for the human housekeeping gene glyceraldehyde 3-phosphate dehydrogenase (GAPDH) were as follows: forward primer, 5′-CTGGCCCCCTCTGTCTAGT-3′, and reverse primer, 5′-TCCAGATACCAAGTGTTCATG-3′. Expression levels of genes were quantified on the basis of intercalation of SYBR Green on an ABI 7000 real-time PCR system, using the SDS 7000 software (Life Technologies). Data normalization and analysis were accomplished using the comparative cycle threshold (ΔCt) method. Each replicate Ct was normalized to the Ct of GAPDH. The fold change in UL144 relative to the GAPDH endogenous control was calculated as 2−ΔΔCt.

The UL144 promoters of HCMV in CRC specimens were amplified by PCR with primers 5′-AAGCTTCTACGGAGGAA-3′ and 5′-TCT CGAGTATATGCCAT ACC-3′, as reported previously (Poole et al., 2013). PCR was carried out with the following protocol: 95 °C for 40 s, 55 °C for 40 s and 72 °C for 1 min for 35 cycles. The amplified products were sequenced by the National Yang-Ming University VYM Genome Research Centre.

**Statistical analysis.** A χ2 test with Yates’ correction or Fisher’s exact test was used to compare differences in categorical variables. Continuous variables were compared using a t-test for data that followed the normal distribution and a Mann–Whitney U test for data that did not follow the normal distribution. We used the Kaplan–Meier method to estimate patient survival. Significant differences in the survival curves were evaluated with a Mantel–Cox log-rank test. The Cox proportional hazards regression model with forward-stepwise selection procedures was used to identify risk factors for mortality after surgery. Variables with P < 0.1 in the univariate analysis were retained in the multivariate models. Patient age was transformed into quartiles to reduce right skewness in the multivariate analysis. P < 0.05 indicated a significant difference. All analyses were carried out by using SPSS 18.0 for Windows (IBM-SPSS).

**ACKNOWLEDGEMENTS**

This study was supported by Taipei Veterans General Hospital (grant nos V98C1-110 and V100C-006) and the National Science Council (grant nos NSC99-2314-B-075-015-MY2 and NSC99-3114-B-002-005). The computational facilities for this work were partly supported by the National Science Council, Taiwan (High-throughput Genome Analysis Core Facility) and the VYM Genome Research Centre, National Yang-Ming University. The authors declare no conflicts of interest.

**REFERENCES**


