Analysis of the entire HLA region in susceptibility for cervical cancer: a comprehensive study

M Zoodsma, I M Nolte, M Schipper, E Oosterom, G van der Steege, E G de Vries, G J te Meerman, A G J van der Zee

**Background**: Infection with human papillomavirus (HPV) is the main cause of cervical cancer and its precursor lesion, cervical intraepithelial neoplasia (CIN). Variability in host immunogenetic background is important in determining the overall cellular immune response to HPV infections.

**Objective**: To determine whether the HLA-DQ or HLA-DR genes, or others in their vicinity, are associated with cervical cancer.

**Methods**: Markers covering the entire HLA region were genotyped in a large sample of CIN and cervical cancer patients and in controls (311 CIN, 695 cervical cancer, 115 family controls, and 586 unrelated controls).

**Results**: Two markers were associated with susceptibility to cervical neoplasia, G511525 and MICA. G511525, close to the region containing the HLA-DQ and HLA-DR genes, was most strongly associated, showing a decrease in frequency of allele 221 from 6.7% to 3.3% in patients with squamous cell cancer (SCC). An association was found for MICA (allele 184) with SCC (odds ratio (OR) = 1.31 (95% confidence interval, 1.13 to 1.53)); homozygotes, OR = 1.48 (1.06 to 2.06)). No associations were observed with adenocarcinoma or CIN.

**Conclusions**: There is an association of the region containing the HLA-DQ and HLA-DR genes with the risk of developing squamous cell carcinoma. An increased risk was observed for carriers of allele 184 of the MICA locus, in particular for homozygotes, suggesting a recessive effect.

Cervical cancer is the second most common cancer in women worldwide. Infection with oncogenic types of human papillomavirus (HPV) is the main causal factor of cervical cancer and its precursor lesion, cervical intraepithelial neoplasia (CIN). Although many women are infected with HPV, only a minority will develop CIN or cervical cancer. Cellular immunity may be critical in the elimination of HPV-harbouring cells. The ability to respond to HPV antigens appears to be related to the capacity of infected cells to present viral epitopes (E6 and E7 oncoproteins) effectively to T cells. The variability in host immunogenetic background, such as the human leucocyte antigen (HLA) class I or II type antigens, is an important variable in determining the overall cellular immune response to HPV infections. The MHC region encodes for the HLA molecules, encompasses 3.6 Mb, and is divided into three regions (centromeric to telomeric): class II (1.1 Mb), class III (0.7 Mb), and class I (1.8 Mb). The principal function of the highly polymorphic HLA molecules is to bind "non-self" peptide fragments, so that they can be optimally presented to cytotoxic T lymphocytes and natural killer cells.

Many studies have analysed different HLA II type antigens in CIN and cervical cancer susceptibility. These were recently reviewed and analysed in a pooled analysis which showed an association with various DQA1-, DQB1, and DRB1 alleles. A few studies on HLA I type antigens also revealed associations with HLA-B alleles, but other studies did not find any associations with HLA-A, HLA-B, or HLA-C. Recently, Engelmark et al. analysed 576 individuals in a sibling pair study and found both linkage and association for the DPB1, DQB1, and DRB1 loci. As there is a strong linkage disequilibrium (LD) in the HLA region, it may well be that other genes in the HLA region are responsible for the association rather than the ones mentioned above. Thus an investigation of the whole HLA region by an extended set of markers should give an answer to this question. In our research centre, comparable studies have been carried out in testicular, breast, and colorectal cancer. The study on breast cancer susceptibility revealed an association with the HLA class III subregion. No associations were found between the HLA region and testicular or colorectal cancer.

In this study, we analysed the involvement of the HLA region in CIN and cervical cancer susceptibility. Markers covering the entire HLA region were genotyped in a large sample of CIN and cervical cancer patients and in controls. As well as allele, genotype, and haplotype association analyses, we also used the haplotype sharing statistic (HSS), which analyses differences between haplotype sharing in this region among patients and controls. HSS extracts extra information from phase as compared with association analysis.

**METHODS**

**Patients and controls**

All patients and controls were part of the white founder population of the northern Netherlands. Patients participated in a population-based study that aimed to detect CIN and cervical cancer susceptibility genes. Between November 1999 and August 2002 all CIN and cervical cancer patients who visited our outpatient clinic were asked by their physician to participate. Through the patients, family members (preferably parents or child and spouse) were also asked to participate. To obtain more power for the association analyses, patients from our outpatient clinic who were included in earlier studies and agreed to the use of their serum in follow-up studies were also included in the present study. In addition, 563 spouses of cases of two other comparable studies in different malignancies were included to complement the family-based controls. All DNA samples and data in this study were handled anonymously, and individuals were aware that.

**Abbreviations**: AC, adenocarcinoma; CIN, cervical intraepithelial neoplasia; HPV, human papillomavirus; HSS, haplotype sharing statistic; LD, linkage disequilibrium; SCC, squamous cell carcinoma; SNP, single nucleotide polymorphism.
they would not be informed about individual test results. The study was approved by the medical ethics committee of the hospital. All included subjects gave written informed consent.

**Genotyping**

DNA was either extracted from 20 ml EDTA blood following standard procedures or from serum by using the QiAamp DNA Midi Kit (Qiagen, Valencia, California, USA) and was stored at −80°C. Subsequently, 22 polymorphic microsatellite markers and two single nucleotide polymorphisms (SNPs) in the HLA region on 6p21 were genotyped in all patients and controls. For detailed marker, primers, and probe information see table 1. All markers and SNPs used are located in the 3.6 Megabase (Mb) HLA region, in particular in the HLA class II region. The marker order was determined by using sequence data, as published by the MHC sequencing consortium.2 In addition, for the genes located in this region and the positions relative to the markers, both the NCBI map (http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/map-search) and the Celera maps (Celera Genomics, Rockville, Massachusetts, USA) were used.

Microsatellite markers were amplified by polymerase chain reaction (PCR), in final reaction volumes of 10 μl, each containing ~25 ng DNA, 0.5 units Taq DNA polymerase (Amersham Pharmacia Biotech, Uppsala, Sweden), 0.2 mM dNTP (Roche Diagnostics, Mannheim, Germany), 2.5 mM MgCl2, 10 mM Tris-HCl, 50 mM KCl (Amersham Pharmacia Biotech, Uppsala, Sweden), 0.2 mM dNTP (Roche Diagnostics, Mannheim, Germany), 2.5 mM MgCl2, 10 mM Tris-HCl, 50 mM KCl (Amersham Pharmacia Biotech, Uppsala, Sweden), 0.2 mM containing 2.5 μl Milliq and 0.2 μl ET-400R size standard (Amersham Pharmacia Biotech), and separated on a MegaBACE 1000 capillary sequencer (Amersham Pharmacia Biotech) according to the manufacturer’s protocol. Results were analysed using Genetic Profiler version 2.0 (Amersham Pharmacia Biotech).

For the SNPs, reactions were carried out in 5 μl volumes and contained 25 ng DNA, 1× TaqMan Universal PCR Master Mix (Applied Biosystems), 100 nM of each primer, and 900 nM of each probe. Cycling conditions on the ABI prism 7900 HT (Applied Biosystems) were two minutes at 50°C, 10 minutes at 95°C, followed by 40 cycles of 15 seconds at 92°C and one minute at 60°C. End point fluorescence was measured immediately after cycling. Alleles were assigned using SDS 2.0 software (Applied Biosystems).

Scoring of the alleles was blinded for affection status (thus patient or control) and family structure.

**Statistical methods**

After genotyping the 22 markers and two SNPs in all participants, they were first tested for quality by the Hardy–Weinberg equilibrium test. Markers that failed this test were not included in the analyses.

Single locus allele and genotype association analyses were carried out. For these analyses, the allele and genotype counts, respectively, were compared between patients and controls using a χ² test if the expected counts were more than five, and a Fisher’s exact test if otherwise. LD between the microsatellite markers was determined and haplotype frequencies were estimated by the expectation-maximisation (EM) algorithm using the software package “Arlequin”.17 For these analyses, only individuals

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Marker data and primer sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Locus</strong></td>
<td><strong>Marker name</strong></td>
</tr>
<tr>
<td>1</td>
<td>D6S001</td>
</tr>
<tr>
<td>2</td>
<td>D6S002</td>
</tr>
<tr>
<td>3</td>
<td>DNRNGCA</td>
</tr>
<tr>
<td>4</td>
<td>RN56C0A</td>
</tr>
<tr>
<td>5</td>
<td>D6S2565</td>
</tr>
<tr>
<td>6</td>
<td>D6S2658</td>
</tr>
<tr>
<td>7</td>
<td>TAP1</td>
</tr>
<tr>
<td>8</td>
<td>D6S2444</td>
</tr>
<tr>
<td>9</td>
<td>D6S2662</td>
</tr>
<tr>
<td>10</td>
<td>WJ13</td>
</tr>
<tr>
<td>11</td>
<td>G51185</td>
</tr>
<tr>
<td>12</td>
<td>D6S2447</td>
</tr>
<tr>
<td>13</td>
<td>WJ11</td>
</tr>
<tr>
<td>14</td>
<td>D6S2665</td>
</tr>
<tr>
<td>15</td>
<td>D6S2670</td>
</tr>
<tr>
<td>16</td>
<td>D6S273</td>
</tr>
<tr>
<td>17</td>
<td>D6S2671</td>
</tr>
<tr>
<td>18</td>
<td>TNFα</td>
</tr>
<tr>
<td>19</td>
<td>D6S2672</td>
</tr>
<tr>
<td>20</td>
<td>MICA</td>
</tr>
<tr>
<td>21</td>
<td>D6S2694</td>
</tr>
<tr>
<td>22</td>
<td>D6S2696</td>
</tr>
<tr>
<td>23</td>
<td>D6S2710</td>
</tr>
<tr>
<td>24</td>
<td>D6S265</td>
</tr>
<tr>
<td>25</td>
<td>D6S2707</td>
</tr>
<tr>
<td>26</td>
<td>D6S1683</td>
</tr>
</tbody>
</table>
In the single locus allele and genotype association analyses, all patients were included, but for HSS—which requires phased haplotypes—only those with at least one family member were used.

Significant associations are defined by a probability (p) value below 0.05/24 = 0.0021—that is, at a 5% level corrected for multiple tests at 24 markers. The corrected 95% CI corresponds to an uncorrected 99.79% CI.

**RESULTS**

Hardy–Weinberg equilibrium testing showed that the observed genotype frequencies of two markers (D6S2662 and D6S2447) differed highly significantly from the expected ones. They were therefore excluded from the analyses.

In all 311 CIN patients, 695 cervical cancer patients, 115 family based controls, and 586 unrelated controls were included. Figure 1 shows the single locus allele association analysis of the microsatellite markers and the SNPs. Two markers had a significant association with susceptibility to cervical neoplasia, namely G511525 (p = 0.0019) and MICA (p = 0.00098). The allele and genotype frequencies for G511525 (allele 221) and the MICA (allele 184) markers, the ORs and 95% CIs among cases and controls, and the different histological subgroups are shown in table 2. The genetic effect was strongest at G511525, showing a decrease in allele/caller frequency of allele 221 from 6.7%/11.7% among controls to 3.9%/7.1% among all cases (OR for the 221 allele = 0.56 (95% CI, 0.36 to 0.89); OR for 221 carriers = 0.58 (0.34 to 0.99)). Subgroup analysis showed that this difference was most pronounced among patients with SCC (3.3%/5.6%: OR for the 221 allele = 0.47; OR for 221 carriers = 0.44). The corresponding frequencies among either CIN or AC patients did not differ significantly from controls.

Allele 184 at MICA, and in particular homozygotes, gave a higher risk of cervical carcinoma (OR for the 184 allele = 1.25 (95% CI, 1.00 to 1.56); OR for 184 homozygotes = 1.46 (1.01 to 2.12)). Again, this effect was strongest among SCC cases (OR for the 184 allele = 1.31 (1.03 to 1.67); OR for 184 homozygotes = 1.51 (1.02 to 2.24)).

HSS on a subset of the analysed markers (see Methods) did not show a difference in mean haplotype sharing between patients and controls (fig 2). Of note, because this method requires phase haplotypes, the set of patient haplotypes in this analysis was small (n = 190). Two five marker haplotype association analyses were carried out—one for TAP1-D6S2444-G511525-D6S2665-D6S2670 and one for TNFα-D6S2672-MICA-D6S2673-D6S2694. With this haplotype method, no specific haplotypes were found associated with cervical cancer risk; in particular, there were none with the associated alleles 221 at G511525 and 184 at MICA (data not shown).

**DISCUSSION**

To our knowledge, this is the first study to investigate the entire HLA region in a large sample of patients with cervical neoplasia. The aim of our study was to find out whether the HLA-DQ or HLA-DR genes or, because of strong LD in the HLA region, other genes in the vicinity were responsible for the association with the disease. Our study shows that a decreased risk of developing SCC is associated with marker G511525. G511525 is located in the region containing the HLA-DQ and HLA-DR genes, which have previously been implicated in susceptibility to CIN and cervical cancer. As G511525 was the most strongly associated with SCC, our study establishes the role of the HLA-DQ and HLA-DR genes in SCC development. Our results are also consistent with the finding of Engelmark et al and indicate that associations can be found with non-HLA markers that are as strong as those found with HLA alleles.

![Association analysis](image-url)
We did not find an association between HLA and CIN or adenocarcinoma. Lack of association with the latter might reflect the relative small sample size of this subgroup. However, results from a recent study also suggested that genetic susceptibility appears to play a role particularly in invasive squamous cell carcinoma of the cervix, while such an association is not as clearly evident for adenocarcinoma. This supports a previous assertion that these two histological subtypes of cervical cancer—while sharing HPV infection as a common aetiological agent—have different aetiological cofactors. The association pointed towards the relative small sample size of this subgroup. Only large case–control studies with large numbers of samples and a small sample size in this subgroup, or of the reasons why the negative result for adenocarcinoma might be a result of the small sample size in this subgroup, or of the reasons mentioned above. Recently, a study into the role of the MICA gene in the development of squamous cell carcinoma is likely. Our negative result for adenocarcinoma might be a result of the small sample size in this subgroup, or of the reasons mentioned above. Recently, a study into the role of the MICA gene in the development of squamous cell carcinoma is likely. The negative result for adenocarcinoma might be a result of the small sample size in this subgroup, or of the reasons mentioned above. Recently, a study into the role of the MICA gene in the development of squamous cell carcinoma is likely. Our negative result for adenocarcinoma might be a result of the small sample size in this subgroup, or of the reasons mentioned above.

We also found an association for the MICA marker (allele 184) with squamous cell carcinoma. The association pointed in particular to patients who are homozygote for the 184 allele, which suggests a recessive effect. Although HSS did not reveal a significant association, the highest p values obtained by this method were also observed around this marker. The absence of any significance probably reflects lack of power, as HSS requires phased haplotypes and our dataset contained only 118 patients with at least one relative who could be used for this analysis. Ghaderi et al. studied this locus previously and found no association with either CIN or cervical cancer, but only 58 cervical cancer patients and 78 CIN patients were included in those studies. We attribute their negative result to lack of power. As our case–control sample size was much larger, a role of the MICA gene in the development of squamous cell carcinoma is likely. Our negative result for adenocarcinoma might be a result of the small sample size in this subgroup, or of the reasons mentioned above. Recently, a study into the role of the MICA gene in the development of squamous cell carcinoma is likely. The negative result for adenocarcinoma might be a result of the small sample size in this subgroup, or of the reasons mentioned above.

Conclusions
Our study further establishes an association of the region containing the HLA-DQ and HLA-DR genes in the development of cervical cancer, a denser set of SNPs, ideally all SNPs located in the region, should be analysed, because there is still no statistical evidence that the associations are really caused by HLA variations. Only large scale association studies with large numbers of samples and a high density of markers can pinpoint the most commonly associated genetic variation. Functional arguments for the causal role of an SNP need to be complemented with statistical evidence that the SNPs indicated have the strongest association with cervical cancer.

We also found an association for the MICA marker (allele 184) with squamous cell carcinoma. The association pointed in particular to patients who are homozygote for the 184 allele, which suggests a recessive effect. Although HSS did not reveal a significant association, the highest p values obtained by this method were also observed around this marker. The absence of any significance probably reflects lack of power, as HSS requires phased haplotypes and our dataset contained only 118 patients with at least one relative who could be used for this analysis. Ghaderi et al. studied this locus previously and found no association with either CIN or cervical cancer, but only 58 cervical cancer patients and 78 CIN patients were included in those studies. We attribute their negative result to lack of power. As our case–control sample size was much larger, a role of the MICA gene in the development of squamous cell carcinoma is likely. Our negative result for adenocarcinoma might be a result of the small sample size in this subgroup, or of the reasons mentioned above. Recently, a study into the role of the HLA region in breast cancer development showed an association of the HLA class III subregion with increased breast cancer risk. This suggests that there might be a role of the MICA gene in the development of squamous cell carcinoma is likely. Our negative result for adenocarcinoma might be a result of the small sample size in this subgroup, or of the reasons mentioned above. Recently, a study into the role of the HLA region in breast cancer development showed an association of the HLA class III subregion with increased breast cancer risk. This suggests that there might be a role of the MICA gene in the development of squamous cell carcinoma is likely. Our negative result for adenocar

**Table 2** Odds ratios and (corrected) 95% confidence intervals for the association of the allele/genotype of the markers with cervical neoplasia risk

<table>
<thead>
<tr>
<th>Locus</th>
<th>Allele</th>
<th>Controls OR (95% CI)</th>
<th>All patients OR (95% CI)</th>
<th>CIN OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS11525</td>
<td>221</td>
<td>127/1887 (6.7%) 1</td>
<td>76/1945 (3.9%) 0.56 (0.36 to 0.89)</td>
<td>23/522 (4.4%) 0.64 (0.31 to 1.30)</td>
</tr>
<tr>
<td>MICA</td>
<td>184</td>
<td>996/1945 (50.9%) 1</td>
<td>1073/1925 (55.7) 1.20 (0.98 to 1.46)</td>
<td>281/518 (54.2%) 1.13 (0.83 to 1.53)</td>
</tr>
</tbody>
</table>

**Genotype**

<table>
<thead>
<tr>
<th>Locus</th>
<th>Allele</th>
<th>All carcinomas OR (95% CI)</th>
<th>SCC OR (95% CI)</th>
<th>AC OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS11525</td>
<td>221</td>
<td>48/1322 (3.6%) 0.52 (0.31 to 0.89)</td>
<td>33/1012 (3.3%) 0.47 (0.25 to 0.86)</td>
<td>15/310 (4.8%) 0.70 (0.30 to 1.67)</td>
</tr>
<tr>
<td>MICA</td>
<td>184</td>
<td>738/1390 (54.6%) 1.25 (1.00 to 1.56)</td>
<td>576/994 (57.9) 1.31 (1.03 to 1.67)</td>
<td>162/306 (52.8%) 1.07 (0.73 to 1.56)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Locus</th>
<th>Allele</th>
<th>All carcinomas OR (95% CI)</th>
<th>SCC OR (95% CI)</th>
<th>AC OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS11525</td>
<td>221</td>
<td>42/659 (6.4%) 0.51 (0.28 to 0.95)</td>
<td>28/504 (5.6%) 0.44 (0.22 to 0.90)</td>
<td>14/155 (9.0%) 0.75 (0.29 to 1.91)</td>
</tr>
<tr>
<td>MICA</td>
<td>184</td>
<td>222/649 (34.2) 1.46 (1.01 to 2.12)</td>
<td>173/496 (34.9%) 1.51 (1.02 to 2.24)</td>
<td>49/153 (32.0%) 1.33 (0.73 to 2.41)</td>
</tr>
</tbody>
</table>

**Significant numbers in bold.**

**AC**, adenocarcinoma; **CI**, confidence interval; **CIN**, cervical intraepithelial neoplasia; **HZ**, homozygous; **OR**, odds ratio; **SCC**, squamous cell carcinoma.

Figure 2 Haplotype sharing statistics.

---

We did not find an association between HLA and CIN or adenocarcinoma. Lack of association with the latter might reflect the relative small sample size of this subgroup. However, results from a recent study also suggested that genetic susceptibility appears to play a role particularly in invasive squamous cell carcinoma of the cervix, while such an association is not as clearly evident for adenocarcinoma. This supports a previous assertion that these two histological subtypes of cervical cancer—while sharing HPV infection as a common aetiological agent—have different aetiological cofactors, such as different prevalent subtypes of HPV.

To further investigate the role of the HLA-DQ and HLA-DR genes in the development of cervical cancer, a denser set of SNPs, ideally all SNPs located in the region, should be analysed, because there is still no statistical evidence that the associations are really caused by HLA variations. Only large scale association studies with large numbers of samples and a high density of markers can pinpoint the most commonly associated genetic variation. Functional arguments for the causal role of an SNP need to be complemented with statistical evidence that the SNPs indicated have the strongest association with cervical cancer.

We also found an association for the MICA marker (allele 184) with squamous cell carcinoma. The association pointed in particular to patients who are homozygote for the 184 allele, which suggests a recessive effect. Although HSS did not reveal a significant association, the highest p values obtained by this method were also observed around this marker. The absence of any significance probably reflects lack of power, as HSS requires phased haplotypes and our dataset contained only 118 patients with at least one relative who could be used for this analysis. Ghaderi et al. studied this locus previously and found no association with either CIN or cervical cancer, but only 58 cervical cancer patients and 78 CIN patients were included in those studies. We attribute their negative result to lack of power. As our case–control sample size was much larger, a role of the MICA gene in the development of squamous cell carcinoma is likely. Our negative result for adenocarcinoma might be a result of the small sample size in this subgroup, or of the reasons mentioned above.

Recently, a study into the role of the HLA region in breast cancer development showed an association of the HLA class III subregion with increased breast cancer risk. This suggests that there might be a role of the MICA gene in the development of squamous cell carcinoma is likely. Our negative result for adenocarcinoma might be a result of the small sample size in this subgroup, or of the reasons mentioned above. Recently, a study into the role of the HLA region in breast cancer development showed an association of the HLA class III subregion with increased breast cancer risk. This suggests that there might be a role of the MICA gene in the development of squamous cell carcinoma is likely. Our negative result for adenocarcinoma might be a result of the small sample size in this subgroup, or of the reasons mentioned above. Recently, a study into the role of the HLA region in breast cancer development showed an association of the HLA class III subregion with increased breast cancer risk. This suggests that there might be a role of the MICA gene in the development of squamous cell carcinoma is likely. Our negative result for adenocarcinoma might be a result of the small sample size in this subgroup, or of the reasons mentioned above.
ACKNOWLEDGEMENTS

We thank all patients and their family members who donated their DNA to this project. This work was supported by grant RUG-99-1878 of the Dutch Cancer Society.

Authors’ affiliations

M Zoodsma, A G van der Zee, Department of Gynaecology, University Medical Centre Groningen, Groningen, Netherlands
I M Nolte, M Schipper, E Oosterom, G van der Steege, Department of Medical Biology, University Medical Centre Groningen
E E de Vries, Department of Medical Oncology, University Medical Centre Groningen
Gerard J te Meerman, Department of Medical Genetics, University Medical Centre Groningen

Competing interests: none declared

Correspondence to: Professor A G van der Zee, Department of Gynaecology, University Medical Centre, Hanzeplein 1, PO Box 30.001, 9700 RB Groningen, Netherlands; A.G.van.der.zeee@umcg.nl

Received 7 March 2005
Revised 12 April 2005
Accepted 14 April 2005

REFERENCES


Vandenveltde C, De Door M, Van Beers D. HLA-DQBl*03 and cervical intraepithelial neoplasia grades I–III. Lancet 1993;341:442.


