Identification of susceptibility loci for cervical carcinoma by genome scan of affected sib-pairs

Malin T. Engelmark¹, Emma L. Ivansson¹, Jessica J. Magnusson¹, Inger M. Gustavsson¹, Anna H. Beskow¹, Patrik K.E. Magnusson¹,² and Ulf B. Gyllensten¹,*

¹Department of Genetics and Pathology, Section of Medical Genetics, Rudbeck Laboratory, University of Uppsala, S-751 85 Uppsala, Sweden and ²Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

Received June 20, 2006; Revised and Accepted October 6, 2006

Cervical cancer is caused by a combination of environmental and genetic risk factors. Infection by oncogenic types of human papillomavirus is recognized as the major environmental risk factor and epidemiological studies indicate that host genetic factors predispose to disease development. A number of genetic susceptibility factors have been proposed, but with exception of the human leukocyte antigen CHLA, class II, have not shown consistent results among studies. We have performed the first genomewide linkage scan using 278 affected sib-pairs to identify loci involved in susceptibility to cervical cancer. A two-step qualitative non-parametric linkage analysis using 387 microsatellites with an average spacing of 10.5 cM revealed excess allelic sharing at nine regions on eight chromosomes. These regions were further analysed with 125 markers to increase the map density to 1.28 cM. Nominal significant linkage was found for three of the nine loci [9q32 (maximum lod-score, MLS) =1.95, P < 0.002), 12q24 (MLS = 1.25, P < 0.015) and 16q24 (MLS = 1.35, P < 0.012)]. These three regions have previously been connected to human cancers that share characteristics with cervical carcinoma, such as esophageal cancer and Hodgkin’s lymphoma. A number of candidate genes involved in defence against viral infections, immune response and tumour suppression are found in these regions. One such gene is the thymic stromal co-transporter (TSCOT). Analyses of TSCOT single nucleotide polymorphisms further strengthen the linkage to this region (MLS = 2.40, P < 0.001). We propose that the 9q32 region contains susceptibility locus for cervical cancer and that TSCOT is a candidate gene potentially involved in the genetic predisposition to this disease.

INTRODUCTION

Cervical cancer is the second most common cancer worldwide among women with 493 000 cases reported in year 2002 (1). The development of cervical cancer is associated with both environmental and genetic risk factors. Infections by sexually transmitted forms of the human papillomavirus (HPV), such as the HPV16 and HPV18, are considered necessary for the development of cervical carcinoma (2,3). HPV encode two oncogenic proteins, denoted by E6 and E7, which bind to the tumour-suppressor protein p53 (p53) and retinoblastoma protein (pRb) and thereby interfere with the cell-cycle control (4,5).

On the basis of the recognition that HPV is the main environmental risk factor for cervical carcinoma, attempts have been made to develop both prophylactic and therapeutic HPV vaccines. The prophylactic HPV vaccines, based on non-infectious virus-like particles (VLPs), show promising results (6,7), although it has proved more difficult to develop therapeutic vaccines (8). A major problem in designing efficient vaccines is the incomplete understanding of the human immune defence against HPV. Cell-mediated immunity plays an important role, as is evident from the high prevalence of HPV infection and associated disease in immunosuppressed patients such as those infected with human immunodeficiency virus (9–11) and transplant recipients (12,13). In addition, activated cytotoxic T-lymphocytes (CTLs) (14) and HPV16 E6- and E7-specific CTLs (15) have been detected in squamous intraepithelial lesions and cervical cancer. HPV-specific CTLs has also been found in lymph nodes and tumours of cervical cancer (16).

*To whom correspondence should be addressed. Tel: +46 184714909; Fax: +46 184714931; Email: ulf.gyllensten@genpat.uu.se

© The Author 2006. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oxfordjournals.org
Most HPV infections are cleared spontaneously (2) but a small number of women develop persistent infections, indicating that host genetic risk factors affect the outcome of an HPV infection and thereby influence disease development. Persistent HPV infections are associated with high viral load (17,18) and a high risk of developing malignant cervical epithelial neoplasia (19,20). Biological first-degree relatives of cervical tumour patients have an almost 100% higher risk of developing the disease, in comparison to adoptive first-degree relatives of a case (21). The heritability of the disease is estimated to 27% and the non-shared environmental contribution to 73% (22). These results indicate that host genetic factors may be important in determining the outcome of an HPV infection. The identification of host genetic risk factors will contribute to an understanding of the aetiology of cervical cancer, such as the immune response to HPV infection, as well as provide general insight in host–pathogen interactions. Susceptibility factor candidates include genes encoding proteins involved in the presentation of foreign antigens, tumour suppression and metabolic pathway enzymes. Their importance has been examined in several case–control studies but the results have often been inconsistent between cohorts. A meta analysis of proposed cervical cancer susceptibility genes shows that only for p53 and the human leukocyte antigen (HLA) class II DQB1 and DRB1 genes associations were confirmed (23).

An alternative and unprejudiced strategy for identifying genetic risk factors is to scan the genome of familial materials, such as affected sib-pairs (ASPs) or multicase families, with densely spaced markers and search for regions with excess genetic sharing. This linkage-based analysis is a powerful method to identify genes involved in monogenic diseases (24) and has also been successful in mapping loci affecting complex diseases (25–27). However, linkage analysis has previously not been used to identify cervical cancer susceptibility loci, because of the lack of suitable family materials. We have identified a population-based cervical cancer material of ASPs, to our knowledge the only material of this kind available to genetic studies (28). Here, we report findings from the first linkage-based genomewide scan for cervical cancer susceptibility loci using ASPs.

**RESULTS**

**Genotyping of microsatellites**

In the first step of the genome scan, 363 of 387 microsatellites from the Cooperative Human Linkage Center (CHLC)/Weber Human Screening Set Version 6 (Research Genetics) were successfully genotyped. Of these 363 markers, 330 were genotyped in ≥ 95% of the 576 individuals, whereas for 33 markers, 59–94% of the genotypes were obtained. The two markers D17S939 and D17S802, previously linked to EV, a disease with HPV-infection aetiology (29), were genotyped in > 95% of all individuals. In the fine mapping, 116 of 125 microsatellites were successfully genotyped in ≥ 95% of the 576 individuals.

**Linkage analysis of microsatellites**

Nine marker loci located on a total of eight chromosomes (1p36.21, 3q13.12, 4q35.1, 5p15.33, 9q32, 12q24.32, 12q24.33, 16q24.1, 22q11.21) gave a maximum lod-score (MLS) > 1 in single-point and/or multipoint analysis of the ASPs in the first step of the genome scan (Table 1). In the single-point analysis, D3S3045 at 3q13.12 yielded a lod-score of MLS > 2 and seven markers (D1S1597, D4S408, D5S392, D9S930, D12S392, D16S539 and D22S446) gave single-point MLS > 1 (Table 1). Multipoint MLS calculations reached 1 for four markers (D3S3045, D4S408, D12S392 and GATA32F05) (Table 1). The multipoint MLSs were generally lower than single-point lod-scores, but otherwise the results from the two analyses were very similar (Fig. 1). No deviation from random allelic sharing was detected for the microsatellites D17S939 and D17S802.

The nine loci on chromosomes 1, 3, 4, 5, 9, 12, 16 and 22 were further studied with additional microsatellites positioned on the deCODE genetic map (30). All available markers within 10 cM from microsatellites yielding MLS > 1 in the initial genomewide scan were selected for fine mapping. This reduced the genetic map intervals to an average of 1.28 cM per marker. The increased marker information in the fine mapping strengthened the evidence of linkage for three regions: 9q31.1–9q33.1, 12q24.23–12q24.33 and 16q24.1–16q24.3. Nominal significant multipoint MLS were found at 9q32 (D9S1856 and D9S930, MLS = 1.95, P < 0.002), 12q24 (D12S1609 and D12S1679, MLS = 1.25, P < 0.015) and 16q24 (D16S539, D16S520 and D16S3048, MLS = 1.35, P < 0.012) (Fig. 2A–C). The MLS curve for the 9q32 region was relatively symmetrical and shows one major and one smaller peak. A single peak was found in the MLS curve for the 12q24 region. For the 16q24 locus, the MLS curve did not show a narrow peak, but a wider region of elevated MLS. The exact location of candidate region is therefore not possible to determine until additional markers spanning the 16q24 region are identified and genotyped in the ASPs. The evidence of linkage found in the first step of the genomewide scan was not strengthened for 1p36.21, 3q13.12 and 4q35.1 and was reduced for 5p15.33 and 22q11.21 when fine mapping increased marker density (data not shown).

**Bioinformatics of linked regions**

The three regions that gave positive linkage in the fine mapping were searched for genes that have previously been related to diseases with similar characteristics as cervical cancer. Ten human cancers are connected to the 9q31.1–33.1 region, either through linkage studies, loss of heterozygosity (LOH) analysis, or due to other somatic changes in tumours, namely multiple self-healing squamous epithelioma (31), nevoid basal cell carcinoma syndrome (32), T-cell leukaemia (33), colorectal cancers (34), Hodgkin’s lymphoma (35), esophageal cancer (36), bladder cancer (37), breast cancer (38), lung cancer (39) and gastrointestinal tumours (40). Two potential candidate genes are found at 9q31.3; ikappa-B kinase complex-associated protein (IKBAP), which is involved in inflammation, and tyrosine-protein phosphatase non-receptor type 3 (PTPN3), where somatic mutations have been found in colorectal cancers (34). At 9q32, five candidate genes for cervical cancer susceptibility with functions connected to T-lymphocyte immune defence
and inflammation are found; thymic stromal co-transporter (TSCOT), alpha-1 microglobulin precursor (AMBp), AT-hook transcription factor (AKNA) and tumour necrosis factor ligand superfamily members 8 (TNFSF8) and 15 (TNFSF15). A proposed candidate tumour-suppressor gene denoted deletions in esophageal cancer 1 (DECI) is also located at 9q32 (41). The physical locations of candidate genes at 9q31.3–32 are shown in Figure 3A.

The 12q24.31 region has shown inactivation in familiar hepatic adenomas (42) and recurrent breakpoints in high-grade B-cell lymphomas (43). The autoimmune disease systemic lupus erythematosus is located only 400 kb downstream of the TSCOT (44). Two potential candidate genes involved in T-lymphocyte activation, T-cell activation G protein-coupled receptor (TAGPcr) 81 (45) and interleukin 31 (IL-31) (46) are located at 12q24.31. The region also includes the BRI3BP, which encodes a cervical cancer protooncogene-binding protein (Uniprot/SWISSPROT Q8WY22). The physical locations of the candidate genes at 12q24.31–24.32 are shown in Figure 3B.

The 16q24 locus as a candidate region is questionable. However, 16q24 frequently shows LOH in breast (47), hepato-cellular (48), lung (49) and gastric cancer (50). Also, two genes in this region, cytochrome b (~245) alpha subunit (CYBa) and core-binding factor alpha subunit 2 translocated to 3 (CBFA2T3), are suggested tumour-suppressor genes (51). CBFA2T3 is likely to be involved in breast cancer (52).

**DISCUSSION**

Cervical carcinoma develops as a result of persistent infection with oncogenic types of HPV (2,3). Epidemiological studies suggest that genetic predisposition plays an important role in cervical cancer development (21,22). Attempts to identify host genetic factors predisposing to persistent infection and/or cancer development have, with some notable exceptions, proven difficult. The majority of studies have used case–control materials to examine single candidate genes, but have usually lacked sufficient power to detect but quite strong genetic effects (23). Our search for susceptibility loci, using a genome scan of ASP initially identified nine regions on eight chromosomes with excess allele sharing. Fine mapping with a denser set of microsatellites further strengthened the evidence of linkage for three chromosomal regions, 9q32, 12q24 and 16q24. The lod-scores in these regions reached nominal significance but not the formal threshold for genomewide significant MLS.

It is not unexpected that the MLS for these regions did not reach the level for genomewide significance. Almost all cancer
risk alleles identified so far have been highly penetrant and rare, but are unlikely to be responsible for a large proportion of cancer cases (56). The familial relative risk of cervical cancer in first-degree relatives is about 2 (21), similar to the relative risk of 1.5–2.5 for the most common cancers over all age groups (56). It is therefore unlikely that we would find common highly penetrant cancer susceptibility genes by linkage analysis in multifactorial diseases. Instead, a substantial fraction of the genetic susceptibility to cervical cancer may be due to the genes with weak-to-moderate effect, similar to the situation for other cancers and complex diseases.

A number of factors contribute to the difficulty of identifying susceptibility genes by linkage analysis. Many complex diseases, such as cancer, have a late onset, which makes it difficult to obtain samples from the parental generation. Lack of parental genotypes leads to decreased power for allele sharing methods, since identical by descent (IBD) status has to be inferred from identical by state status, leading to reduced precision in estimates of IBD. Loci with minor effects are also difficult to detect by genomewide scans, unless very large family materials are used. Simulations show that in order to have 90% power of detecting a gene

Figure 1. Distribution of non-parametric MLS allowing for dominance variance from the initial analysis of the whole genome of cervical cancer ASPs. (A) Single-point MLS. (B) Multipoint MLS. Genetic distance (in cM) is displayed along the bottom starting from pter on chromosome 1 and ending at qter on the X chromosome. The chromosome numbers are displayed at the top.
with a major effect ($\lambda_s > 3.5$) in a complex disease by linkage analysis, a minimum of 100 ASPs is required, assuming no recombination between the trait and marker locus ($\theta = 0$), completely informative markers, an additive model and a lod-score threshold of 3.0 (57). In reality, multiple loci usually affect the trait, interactions occur among loci (58,59), marker density is uneven and usually too low ($\theta > 0$) and some markers are not completely informative (57). Estimates indicate that at least 500 ASPs are needed to detect loci with a moderate effect ($1.23 \leq \lambda_s \leq 1.39$) and sample sizes of 1000–2000 pairs may be required to locate genes with small effects ($\lambda_s \leq 1.13$) (60).

Since several genes with low-to-moderate effects are likely to be involved in cervical cancer susceptibility, it will be important to study multilocus models and potential interactions between genes (61). Searching for interactions between multiple loci in genomewide association studies of complex diseases can be both computationally feasible and more powerful than single-gene analysis (62). Modelling of two-disease loci at different chromosomal locations, potentially interacting to produce an effect on disease susceptibility, has been performed in analyses of ASPs (58) and statistical methods to detect epistasis in humans have been developed (63). When software becomes available that can handle genetic interactions between multiple loci in different types of family materials, this should be applied to our ASP material. For example, it will be of interest to examine alleles in HLA class II loci, which are established risk factors for cervical cancer development (28,64), in relation to the loci identified in our genome scan.

Figure 2. Fine mapping analysis of loci giving significant non-parametric multipoint MLS when allowing for dominance variance (54) in ASPs with cervical cancer. (A) 9q32, (B) 12q24 and (C) 16q24. Genetic distance (in cM) is displayed along the bottom.
Our study has found linkage between 9q32, 12q24 and 16q24 and cervical cancer in ASPs. Several candidate genes involved in tumour suppression, inflammation and defence against viral infections are found in these regions. In addition, there are connections between these loci and other human cancers with similar characteristics as cervical carcinoma. The region surrounding 9q32 is especially interesting, since the highest multipoint lod-score (MLS = 1.95) in our genome-wide scan was found for the markers D9S1856 and D9S930 located in this region. Also, several potential genetic candidates, such as TSCOT and DEC1, are located in this region (Fig. 3A). TSCOT, which maps only 400 kb from D9S930, is suggested to be involved in antigen and HLA specificity during both thymal development and peripheral functions of T-cells (53,65,66). Since cellular immunity mediated by T-lymphocytes is critical for elimination of HPV infected cells (15,16) and the ability to respond to HPV antigens is related to the capacity of infected cells to present viral epitopes to T-lymphocytes, TSCOT is interesting in relation to disease development. Additional sharing analysis of TSCOT SNPs in the ASPs revealed linkage signals ranging up to MLS = 2.40, which further supports that this gene might be involved in cervical cancer susceptibility. Nevertheless, data from independent association studies as well as functional studies are needed to determine whether TSCOT is the actual susceptibility factor. It remains possible that our results reflect LD between the SNPs and the true susceptibility factor in 9q32. For instance, DEC1 is another interesting candidate located ~400 kb from D9S1855. DEC1 is often deleted in esophageal cell carcinomas (41), and in both cervical (2,3) and esophageal cancer (67), oncogenic HPV types are regarded as prominent risk factors.

At this time, prophylactic HPV vaccines have been successfully produced and are moving into phase III trials (6,7). However, there are concerns regarding these vaccines with regard to HPV type coverage, duration of protection, potential

![Figure 3. Physical locations of potential candidate genes for cervical cancer susceptibility. (A) 9q32 and (B) 12q24. Physical distance (in Mb) is displayed along the bottom.](image)

### Table 2. Description of identified TSCOT SNPs in the ASPs (n = 576)

<table>
<thead>
<tr>
<th>SNP</th>
<th>Reference id</th>
<th>Genomic positiona</th>
<th>TSCOT location</th>
<th>Allelesb</th>
<th>Major allele</th>
</tr>
</thead>
<tbody>
<tr>
<td>389</td>
<td>rs10739380</td>
<td>114 680 505</td>
<td>Promoter</td>
<td>A/T</td>
<td>A</td>
</tr>
<tr>
<td>565</td>
<td>—</td>
<td>114 680 681</td>
<td>Promoter</td>
<td>C/A</td>
<td>C</td>
</tr>
<tr>
<td>653</td>
<td>rs4978511</td>
<td>114 680 769</td>
<td>Promoter</td>
<td>A/G</td>
<td>G</td>
</tr>
<tr>
<td>723</td>
<td>rs4142853</td>
<td>114 680 839</td>
<td>Promoter</td>
<td>G/T</td>
<td>G</td>
</tr>
<tr>
<td>801</td>
<td>rs4142852</td>
<td>114 680 917</td>
<td>Promoter</td>
<td>A/G</td>
<td>A</td>
</tr>
<tr>
<td>2018</td>
<td>rs3802491</td>
<td>114 682 134</td>
<td>Exon 1</td>
<td>C/A</td>
<td>C</td>
</tr>
<tr>
<td>10115</td>
<td>—</td>
<td>114 690 231</td>
<td>Intron 3</td>
<td>C/G</td>
<td>C</td>
</tr>
<tr>
<td>10174</td>
<td>—</td>
<td>114 690 290</td>
<td>Intron 3</td>
<td>C/T</td>
<td>C</td>
</tr>
<tr>
<td>10564</td>
<td>rs4979178</td>
<td>114 690 680</td>
<td>Intron 3</td>
<td>C/G</td>
<td>G</td>
</tr>
<tr>
<td>10633</td>
<td>rs4979177</td>
<td>114 690 749</td>
<td>Intron 3</td>
<td>C/T</td>
<td>T</td>
</tr>
<tr>
<td>11104</td>
<td>rs4978510</td>
<td>114 691 220</td>
<td>Intron 3</td>
<td>C/G</td>
<td>G</td>
</tr>
<tr>
<td>11219</td>
<td>—</td>
<td>114 691 335</td>
<td>Intron 3</td>
<td>T/C</td>
<td>T</td>
</tr>
<tr>
<td>13136</td>
<td>rs2236600</td>
<td>114 693 252</td>
<td>Exon 4</td>
<td>C/G</td>
<td>C</td>
</tr>
<tr>
<td>13330</td>
<td>rs16917440</td>
<td>114 693 446</td>
<td>Exon 4</td>
<td>T/G</td>
<td>T</td>
</tr>
<tr>
<td>13920</td>
<td>—</td>
<td>114 694 036</td>
<td>3’-UTR</td>
<td>C/T</td>
<td>C</td>
</tr>
<tr>
<td>13923</td>
<td>rs10739379</td>
<td>114 694 039</td>
<td>3’-UTR</td>
<td>G/C</td>
<td>C</td>
</tr>
<tr>
<td>13926</td>
<td>—</td>
<td>114 694 042</td>
<td>3’-UTR</td>
<td>T/C</td>
<td>T</td>
</tr>
</tbody>
</table>

aAccording to draft genome sequence data (http://www.ensembl.org/index.html).
bFirst allele present in draft genome sequence data (http://www.ensembl.org/index.html).

### Table 3. Multipoint MLSs for TSCOT SNPs in the ASPs (n = 576)

<table>
<thead>
<tr>
<th>SNP</th>
<th>TSCOT location</th>
<th>MLSa</th>
<th>P-valueb</th>
</tr>
</thead>
<tbody>
<tr>
<td>723</td>
<td>Promoter</td>
<td>1.41</td>
<td>&lt;0.009</td>
</tr>
<tr>
<td>2018</td>
<td>Exon 1</td>
<td>2.12</td>
<td>&lt;0.002</td>
</tr>
<tr>
<td>10115</td>
<td>Intron 3</td>
<td>2.24</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>10174</td>
<td>Intron 3</td>
<td>2.40</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>11219</td>
<td>Intron 3</td>
<td>2.19</td>
<td>&lt;0.002</td>
</tr>
<tr>
<td>13136</td>
<td>Exon 4</td>
<td>2.03</td>
<td>&lt;0.002</td>
</tr>
<tr>
<td>13920</td>
<td>3’-UTR</td>
<td>2.36</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>13926</td>
<td>3’-UTR</td>
<td>2.36</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

aMultipoint analysis with D9S930 and D9S262, allowing for dominance variance (54).
bNominal P-values (54).
target age and sex groups and cost effectiveness (68). In therapeutic vaccine research, the focus has been directed towards mapping of HLA class I restricted epitopes of HPV16 and HPV18 E6 and E7 (69,70). Such vaccines have proven difficult to develop and peptide vaccination may be most efficient in individuals with pre-invasive disease, who are not immunocompromised and when the HLA genotype of the patient and the HPV genotype of the tumour is known (8). More knowledge is needed on the immune responses against HPV infections before a successful therapeutic vaccine is likely to become a reality. Further characterization of genes affecting the immune response against HPV infection and development of persistent infection will thus be important for the development of efficient HPV vaccines.

In summary, our genomewide scan of cervical cancer ASPs has revealed linkage to three chromosomal sites, 9q32, 12q24 and 16q24. Further characterization of the genetic susceptibility factors for cervical carcinoma, such as TSCOT and other loci located at 9q32, will provide important information for understanding the aetiology of tumour development.

MATERIALS AND METHODS

Family material

The Swedish Cancer Registry contains records of all cases of cancer since 1958. In order to identify families with several cases of cervical carcinoma, all cervical tumour cases, classified according to the criteria of the seventh revision of the International Classification of Diseases, born after 1940 and reported to the Swedish Cancer Registry before 1993, were localized. These individuals were also identified in the Swedish Multi-Generation Register (Statistics, Sweden), which resulted in the identification of families with different number of relatives diagnosed with cervical cancer. Among the diagnoses, dysplasia constituted 10%, in situ cervical carcinoma 85% and invasive cervical carcinoma 5% (22). In total, 1800 ASPs, 2855 affected mother–daughter pairs, 232 families with three cases and eight families with more than three cases were identified in the Swedish population using these registers and search criteria. The sib-pairs and mothers with at least one diagnosis of cervical carcinoma in situ were selected from the 1800 ASPs and these were invited to participate in the study. In total, 4145 women were invited and blood samples were obtained from 2135 participants (51%) belonging to 641 ASPs. A set of 278 ASPs, corresponding to 576 individuals, were chosen for the present study. The selection criteria and specific structure of the 278 ASPs has been described previously (28). The study was approved by the Research Ethics Committee of Uppsala University.

Genotyping of microsatellites

Extraction of DNA was performed from 4.5 ml whole blood by standard phenol chloroform procedures using GENE-PURE™ 341 Nucleic Acid Purification System (Applied Biosystems). The DNA concentration was determined by using spectrophotometries and the DNA samples were diluted to 5 ng/μl in 96-well plates. A two-stage genome scan was performed on the 576 DNA samples. In the first step, microsatellites from the CHLC/Weber Human Screening Set Version 6 (Research Genetics) were used. A total of 387 microsatellites were typed, spanning all 22 autosomes and the X chromosome. The markers were positioned according to the Marshfield genetic map with an average spacing of 10.5 cM across the genome (http://www2.marshfieldclinic.org/research/genetics). In addition to this Weber set, two extra microsatellites, D17S939 and D17S802, were typed due to their previous linkage to a disease with HPV infection aetiology (29). In the second step, regions on chromosomes 1, 3, 4, 5, 9, 12, 16 and 22 were further mapped with 125 additional markers. These microsatellites represent all markers available at the time, using the deCODE map, positioned within a 10 cM region from the marker showing the initial linkage (30). PCR amplification for microsatellite genotyping was performed using ABI PRISM™ 877 Integrated Thermal Cycler (Applied Biosystems) robots. Each PCR reaction of 5 μl contained 0.025 μl (5 U/μl) HotStarTaq™ DNA polymerase (QIAGEN), 0.5 μl 10× PCR buffer with MgCl2, 0.2 μl (20 mM) of dNTPs, 0.2 μl (25 mM) MgCl2, 0.01 μl (100 μM) forward primer, 0.01 μl (100 μM) reverse primer, 2.055 μl sterile H2O and 2.0 μl (5 ng/μl) DNA. All primers used in the genome scan were fluorescently labelled with HEX, 6-FAM and NED. The amplification was carried out with an initial step of 95°C for 5 min. This was followed by 35 cycles of 30 s at 95°C for denaturing, 45 s at 55°C for annealing and 90 s at 72°C for elongation. An elongation step at 72°C for 7 min was used as a final step. The PCR products were pooled with up to eight markers in the pools. From the pooled PCR product, 1.5 μl was mixed with 10 μl GeneScan™_500 ROX™ size standard (Applied Biosystems) diluted 1:56 with Hi-Di™ formamide (Applied Biosystems). This mixed pool was automatically loaded and fragments separated on an ABI PRISM™ 3700 DNA Analyser instrument (Applied Biosystems).

Linkage analysis of microsatellites

The allele binning and genotyping was performed using the GeneMapper™ software version 3.0 (Applied Biosystems). To minimize genotyping errors, two scientists performed the genotyping independently. The GAS package, version 2.0 (A. Young, Oxford University) was used to determine inconsistent inheritance.

Qualitative NPL analysis was performed using the MLS method (71). MLS values express deviation of IBD sharing among ASPs and are calculated using a likelihood ratio method with the null hypothesis of mendelian segregation expectations (z0 = 0.25, z1 = 0.5 and z2 = 0.25). Constraints on the observed IBD probabilities can also be used when calculating MLS values such that they reflect alternative genetic models. In contrast to the NPL statistics, the MLS is not affected by a bias towards the null hypothesis and generally gives more power to the ASP analysis when applying possible triangle constraints (72). A MLS of 0.74 corresponds to a nominal threshold of P < 0.05 and a MLS of 2.32 to a nominal P = 0.001, if calculated using the possible triangle constraint (allowing for dominance variance) (54). However, the appropriate criteria for genomewide significance is debatable. Calculations based on an infinitely dense map
with perfectly informative markers propose that suggestive and significant linkage is reached at MLS = 2.45 ($P = 7.4 \times 10^{-4}$) and MLS = 3.93 ($P = 2.2 \times 10^{-5}$), when allowing for dominance variance (73,74). However, these thresholds may be too conservative and an MLS = 3.2 is therefore more appropriate to use for genomewide significance when analyzing genetic sharing in ASP materials (75). In order to ensure that loci with a smaller effect are not missed in genome scans, it has been recommended to use even lower thresholds. For instance, MLS values greater than 1 are suggested to be indicative of a region of interest in genomewide scans (76). We applied this threshold in the initial scan when determining what regions to further investigate. The Genehunter 2.1 package (55) was used to determine maximum-likelihood proportions allowing for dominance variance and to calculate single-point and multipoint MLS values for the autosomes. Genehunter version 1.3 was used for the linkage analysis of the X chromosome. The allele frequencies for each marker loci were determined from the ASP material. Genetic distances from Marshfield genetic map (http://www2.marshfieldclinic.org/research/genetics) were used in the multipoint analysis of the initial scan. We applied the deCODE genetic map in the multipoint analysis of the fine mapping (30) because of improved precision of genetic distance estimates between markers.

Bioinformatics of linked regions

The regions showing increased MLS scores in the fine mapping analysis were searched for genes and disease associations with apparent connections to cervical cancer pathogenesis using the Ensembl Genome Browser (http://www.ensembl.org/index.html) and the Online Mendelian Inheritance in Man database (OMIM, http://www.ncbi.nlm.nih.gov/Omim/).

Identification of polymorphisms in TSCOT

The sequence analysis included the promoter region, exons, exon–intron borders, 3'UTR and most of the remaining intron sequences. Primer design was carried out using Primer-Express® 1.0 (Applied Biosystems) or Oligo® 6.0 (Molecular Biology Insights). PCR amplifications were performed in 20 μl reactions containing 1× PCR buffer (including 1.5 mM MgCl₂), 800 μM dNTP (200 μM each), 0.2 μM of each primer, 0.5 U HotStarTaq® Polymerase (Qiagen) and 10 ng of template. The assays were run on a PSQ™96MA instrument with enzymes and substrates from the Pyro Gold Reagent Kit (Biotage). Genotype calling for TSCOT SNPs was carried out using the SDS2.2 software (Applied Biosystems) when using TaqMan® SNP Genotyping and the PSQ™ 96MA 2.1 software (Biotage) when using the Pyrosequencing™ technology.

ACKNOWLEDGEMENTS

We are grateful to all women who generously participated in this study. Our work was supported by grants from the Swedish Cancer Society (grant number 3436-B9704XXA) and the Knut and Alice Wallenberg Foundation through the Wallenberg Consortium North. The Beijer Foundation is greatly acknowledged for sponsoring Patrik Magnusson’s contribution to this work.

Conflict of Interest statement. The authors declare no conflicts of interest.

REFERENCES


