

ORIGINAL ARTICLE

# The Amerindian mtDNA haplogroup B2 enhances the risk of HPV for cervical cancer: de-regulation of mitochondrial genes may be involved

Mariano Guardado-Estrada<sup>1</sup>, Ingrid Medina-Martínez<sup>1</sup>, Eligia Juárez-Torres<sup>1</sup>, Edgar Roman-Bassauré<sup>2</sup>, Luis Macías<sup>1</sup>, Ana Alfaro<sup>1</sup>, Avissai Alcántara-Vázquez<sup>3</sup>, Patricia Alonso<sup>3</sup>, Guillermo Gomez<sup>2</sup>, Fernando Cruz-Talonia<sup>2</sup>, Luis Serna<sup>2</sup>, Sergio Muñoz-Cortez<sup>4</sup>, Manuel Borges-Ibañez<sup>4</sup>, Ana Espinosa<sup>1</sup>, Susana Kofman<sup>5</sup> and Jaime Berumen<sup>1,6</sup>

Although human papillomavirus (HPV) infection is the main causal factor for cervical cancer (CC), there are data suggesting that genetic factors could modulate the risk for CC. Sibling studies suggest that maternally inherited factors could be involved in CC. To assess whether mitochondrial DNA (mtDNA) polymorphisms are associated to CC, HPV infection and HPV types, a case-control study was performed in the Mexican population. Polymorphism of mtDNA D-loop was investigated in 187 CC patients and 270 healthy controls. HPV was detected and typed in cervical scrapes. The expression of 29 mitochondrial genes was analyzed in a subset of 45 tumor biopsies using the expression microarray ST1.0. The Amerindian haplogroup B2 increased the risk for CC (odds ratio (OR)=1.6; 95% confidence interval (CI): 1.05–2.58) and enhanced 36% (OR=208; 95% CI: 25.2–1735.5) the risk conferred by the HPV alone (OR=152.9; 95% CI: 65.4–357.5). In cases, the distribution of HPV types was similar in all haplogroups but one (D1), in which is remarkable the absence of HPV18, a very low frequency of HPV16 and high frequencies of HPV45, HPV31 and other HPV types. Two mtDNA genes (mitochondrial aspartic acid tRNA (*MT-TD*), mitochondrial lysine tRNA (*MT-TK*)) could be involved in the increased risk conferred by the haplogroup B2, as they were upregulated exclusively in B2 tumors ( $P < 0.01$ , *t*-test). Although the association of mtDNA with CC and HPV infection is clear, other studies with higher sample size will be needed to elucidate the role of mtDNA in cervical carcinogenesis.

*Journal of Human Genetics* (2012) 57, 269–276; doi:10.1038/jhg.2012.17; published online 23 February 2012

**Keywords:** Amerindian haplogroups; cervical cancer; HPV; mtDNA; MT-tRNA; polymorphism

## INTRODUCTION

Human papillomavirus (HPV) is the main causal factor for cervical cancer (CC) and it is found in almost 100% of the carcinomas.<sup>1</sup> Despite the high prevalence of HPV infection in the general population, only few women infected will develop cancer.<sup>2</sup> Host factors seem to be critical for the resolution or evolution of HPV infection to a neoplastic lesion and then to an invasive cancer. Recently, it has been suggested that genetic factors could modulate the risk for CC caused by HPV infection.<sup>3</sup> In family studies, the hereditary component has been found relatively high, from 27 to 64%,<sup>4,5</sup> despite the disease has a strong environmental etiological factor. Moreover, it has been reported that first-degree relatives of women with CC have up to threefold greater risk for CC.<sup>6,7</sup> The major histocompatibility complex was the first and the most studied loci associated to CC.<sup>3,8,9</sup> However, there are polymorphisms in these loci associated to the disease in some

populations but not in others.<sup>9</sup> Several polymorphisms in non-HLA immune response-related molecules, xenobiotic metabolizing cell cycle and DNA repair genes have also been studied in CC.<sup>10–12</sup> Most of the odd ratios (ORs) resulted in those studies ranged from low to moderate (1.1–5), numbers that are common for low-penetrance alleles in multifactorial diseases.<sup>13</sup> These data suggest that CC is a complex disease and that some genes could participate at different stages of its development. Interestingly, most genes associated to other cancers are tumor type-specific and different alleles are associated to distinct populations.<sup>13</sup> In the case of CC, specific alleles have been associated to particular populations, but not replicated in others.<sup>14</sup>

On the other hand, in a family study, the risk for CC was higher for maternal than for paternal half-siblings.<sup>5,7</sup> These data could suggest that in addition to nuclear genes, specific maternal genetic factors may

<sup>1</sup>Unidad de Medicina Genómica, Hospital General de México, Facultad de Medicina, Universidad Nacional Autónoma de México, Mexico City, Mexico; <sup>2</sup>Servicio de Oncología, Hospital General de México, Mexico City, Mexico; <sup>3</sup>Servicio de Patología, Hospital General de México, Mexico City, Mexico; <sup>4</sup>Servicio de Ginecología, Hospital General de México, Mexico City, Mexico; <sup>5</sup>Departamento de Genética, Hospital General de México, Facultad de Medicina, Universidad Nacional Autónoma de México, Mexico City, Mexico and <sup>6</sup>Departamento de Medicina Experimental, Facultad de Medicina, Universidad Nacional Autónoma de México, Mexico City, Mexico  
Correspondence: Dr J Berumen, Unidad de Medicina Genómica, Hospital General de México, Dr Balmis 148, Colonia Doctores, Delegación Cuauhtémoc, Mexico City DF 06726, México. E-mail: jaimeberumen@hotmail.com

Received 2 September 2011; revised 25 January 2012; accepted 27 January 2012; published online 23 February 2012

be involved. In contrast to nuclear genome, the human mitochondrial DNA (mtDNA) is exclusively maternally inherited and presents a higher mutation rate. The mitochondria have main roles in energy production in the form of ATP synthesis, apoptosis and regulation of reactive oxygen species production.<sup>15</sup> Several studies have found a direct relation between mitochondrial metabolism and cancer.<sup>16</sup> Somatic mutations in the mitochondrial genome have been reported in various types of malignancies as breast cancer, thyroid, prostate, ovarian carcinomas, among others.<sup>17</sup> Interestingly, some of these mutations also have been found as polymorphisms in populations worldwide.<sup>18</sup> For instance, Asian and West Eurasian mitochondrial haplogroups have been associated with prostate, thyroid, breast and renal carcinomas.<sup>16,19,20</sup> The mitochondrial haplogroups present in different populations have characteristic polymorphisms, although there are few data of functional differences among them.<sup>21,22</sup> In the Mexican mestizo population predominates the Amerindian mtDNA haplogroups A2, B1, C1 and D1, which came from Asian lineages.<sup>23</sup> Currently, there are no data concerning the association of the Amerindian mitochondrial haplogroups with CC and their functional differences in the cervical tumors.

In this report, a case-control study was performed to assess whether mtDNA polymorphisms are associated to CC and if the mitochondrial genetic background modifies the risk conferred by the HPV. In addition, the level of expression of 29 mtDNA genes was measured to analyze whether the different risks conferred by the mitochondrial haplogroups are associated to the activation or inactivation of particular genes.

## MATERIALS AND METHODS

### Subjects and samples

A total of 187 patients with invasive CC diagnosed at the Department of Oncology of the Hospital General de México in Mexico City were recruited. Patients were obtained sequentially during the period from November 2003 through April 2005 and represent about 80% of the patients with newly diagnosed CC in this period because of the restrictive inclusion criteria of our study (no previous treatment, to be incident cases, born in Mexico with Mexican ancestry of two generations). Among the 187 CCs, 86.1% were squamous cell carcinomas, 12.3% were adenocarcinomas, 0.5% were adenosquamous carcinomas and 1.1% were undifferentiated carcinomas. A total of 270 age-matched control subjects (within two years of the patient's age; range 24–78 years) were selected from women attending the CC Screening Program of the Hospital during the same period as the case patients. When a case was identified and confirmed, one or two age-matched control subjects with a normal Pap smear was chosen from the screening program. The hypervariable region 1 and 2 (HVR1 and HVR2) mtDNA polymorphism from the control group was published previously.<sup>23</sup> The participation rate of case and control subjects was close to 95%, which is consistent with previous studies in Mexico.<sup>24</sup> To avoid population stratification between cases and controls, both groups were recruited at the same Hospital, which attends people without social security. The patient subjects received a complete clinical evaluation. Tumor staging was carried out according to the last international revised protocol for gynecological cancer.<sup>25</sup> Directed-biopsy specimens were obtained by colposcopy from the patients and divided into two equal parts. The two fragments were snap-frozen in dry ice and stored at  $-80^{\circ}\text{C}$  until analysis. For HPV detection and typing, a scrape from the endocervix and ectocervix was collected in patients and controls with a cytobrush, and the cells were suspended in a vial with extraction buffer and stored at  $-20^{\circ}\text{C}$  until analysis. From most patients, both cervical scrapes and tumor biopsy were explored. For HPV16 detection, agreement was 98% between these types of specimens. To analyze the mtDNA polymorphism, blood samples were drawn from cases and controls. Analysis of mitochondrial gene expression was performed in RNA extracted from fresh tumor biopsies of 45 cases. A total of 17 normal epithelial samples were used as control for the analysis of gene expression. These samples were obtained from patients undergoing hysterectomy, due to myomatosis, at

the Gynecology Service in the Hospital General de México. These patients had a normal cervix diagnosed by colposcopy and cytology. The epithelium was dissected with a scalpel immediately after receiving a cervix fragment from the operating room. Subsequently, it was snap-frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until being used. An informed written consent was obtained from all participants and the study protocol was approved by the local ethic and scientific committees.

### DNA and RNA isolation

DNA was purified from cervical scrapes, blood samples and biopsy specimens with PureLink genomic DNA kit (Invitrogen, Carlsbad, CA, USA) and maintained at  $-20^{\circ}\text{C}$  until analysis. The total RNA was isolated using TRIzol reagent (Invitrogen), according to the manufacturer's protocol. The quality of RNA was confirmed by agarose gel electrophoresis, as demonstrated by the presence of intact ribosomal RNA (28s and 18s bands).

### mtDNA control region sequencing

The full control region (1261 bp) was amplified by polymerase chain reaction (PCR) from 500 ng of the whole genomic DNA obtained from the blood samples. The mix was carried out in a 25  $\mu\text{l}$  reaction containing 2 mM  $\text{MgCl}_2$ , 1  $\times$  PCR buffer (50 mM KCl, 10 mM Tris-HCl and pH 8.3), 200  $\mu\text{M}$  of each dATP, dCTP, dGTP and dTTP, 1  $\mu\text{M}$  of each primer and 1 U of *Taq* Platinum DNA polymerase (Invitrogen). Reactions were carried out in a GeneAmp PCR System 2720 (Applied Biosystems, Carlsbad, CA, USA). After denaturing for 2 min at  $94^{\circ}\text{C}$ , 32 cycles were performed by setting station 1 at  $94^{\circ}\text{C}$  for 45 s, station 2 at  $60^{\circ}\text{C}$  for 1 min and station 3 at  $72^{\circ}\text{C}$  for 2 min. Finally, the PCR reaction was kept at  $72^{\circ}\text{C}$  for another 7 min. Both DNA strands of HVR1 (positions 16 024–16 365) and HVR2 (positions 73–340) were sequenced using the fluorescent cycle-sequencing method (BigDye Terminator Ready Reaction Kit; Applied Biosystems). Sequence analysis was performed using an ABI PRISM 3130xl genetic analyzer system (Applied Biosystems). Data were analyzed with SeqScape software. The sequence and positions of the forward (F) and reverse (R) primers for PCR and sequencing were as described previously.<sup>23</sup>

### Detection and HPV typing

The HPV detection was performed by PCR using universal primers located in HPV L1 genes *MY09/MY11*, *GP5+/6+* and *L1C1* as described previously.<sup>26–28</sup> The *HBB* gene was used as an internal control to assess the quality of DNA. The HPV types were identified by sequencing the amplified bands in positive samples using the fluorescent cycle-sequencing method (BigDye Terminator Ready Reaction Kit; Applied Biosystems). Sequence analysis was performed using an ABI PRISM 3130xl genetic analyzer system (Applied Biosystems). Each band sequenced from the HPV-positive samples was analyzed through the FASTA sequence similarity tool.<sup>29</sup> The average percentage of identity of HPV types detected was 98.7% (range=91–100%) compared with the reference sequences.

### mtDNA gene expression analysis

The gene expression profile was explored in 45 cases and 17 healthy cervical epithelium controls, using Human Gene 1.0 ST oligonucleotide microarray (Affymetrix, Santa Clara, CA, USA). This array contains 33 297 probe sets that correspond to approximately 20 741 genes of the Gene Reference database, including 29 mtDNA genes. A total of 300 ng of RNA of each sample was used for the synthesis of cDNA. It was carried out with SuperScript II reverse transcriptase and oligo(dT) primer, containing a T7 RNA polymerase promoter, using the GeneChip WT cDNA synthesis kit (Affymetrix). Then, an *in vitro* transcription amplification was carried out overnight using the GeneChip Amplification Kit (Affymetrix). The cRNA was random-primed, including dUTP, and single-stranded DNA was fragmented with the uracil DNA glycosylase, followed by exonuclease 1. Fragmented DNA was then labeled using terminal deoxy-nucleotidyl transferase and biotinylated nucleotides (GeneChip Terminal Labeling Kit; Affymetrix). A hybridization cocktail was prepared, including the labeled target DNA and control probes for hybridization. The microarrays were hybridized during 16 h at  $45^{\circ}\text{C}$  and 60 r.p.m., and then washed and stained with streptavidin phycoerythrin conjugate in a Gene

Chip fluidics station 450. Finally, the chips were scanned using a Gene Chip Scanner 3000. Array hybridization, scanning and image analysis were carried out according to the manufacturer's protocols (Affymetrix GeneChip Expression Assay manual). To assess the quality of the experiments, the following parameters were used: the expression of the exogenous polyA controls, the presence of the oligo B2 used to make grid alignments and the values of the area under the curve above 0.8. Only those microarrays with optimal quality controls were then analyzed. Microarrays were normalized using the Robust Multichip Average algorithm in the Affymetrix expression console. The values of the normalized intensity were referred as units of intensity. Only the normalized intensities from 29 mtDNA probes were used for the analysis. The mean of normalized signals were compared between the whole cases and controls and among the cases grouped by Amerindian haplogroups against the whole controls. Genes were selected using cutoff values of  $P < 0.05$  with a  $t$ -test and fold change (cancer/control)  $\geq 1.5$  for upregulated genes or  $\leq 0.67$  for downregulated genes. The analysis was performed with the FlexArray software (<http://genomequebec.mcgill.ca/FlexArray>). The raw microarray data have been deposited in NCBI's Gene Expression Omnibus repository (<http://www.ncbi.nlm.nih.gov/geo/>) website and are available under the accession number GSE29570 through the following link: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=nnsxvougwkasixw&acc=GSE29570>.

### Data analysis

The mtDNA HVRI and HVR2 sequences of each individual were analyzed and haplogroups were characterized with the haplogrep tool (<http://haplogrep.uibk.ac.at/>) to assure that the haplogroups were correctly assigned (Supplementary Table 1). The haplogrep tool identifies the mtDNA haplogroup sublineages that belonged to the roots from haplogroups A2, B4, B2, C1, D1 and D4 (see Supplementary Table 1). However, for comparisons throughout the paper, the samples were grouped by the mtDNA lineages A2, B2, C1 and D1 (fourth column of Supplementary Table 1), which are commonly reported in populations from Latin America.<sup>23</sup> The risk was estimated as OR, either for HPV infection or mtDNA haplogroups, using unconditional logistic regression adjusted by age when HPV infection was considered in the model. The OR was also calculated for HPV infection in the samples stratified by the mtDNA Amerindian haplogroups. In the case of haplogroup D1, where all cases were HPV positive, an approximate OR was calculated considering one case as negative (see Table 3). The HPV frequencies in the mtDNA haplogroups were compared with the  $\chi^2$  test, Yate's adjusted or Fisher's exact test. Differences were considered significant with a  $P$ -value  $< 0.05$ . All statistical tests were carried out using the software SPSS ver. 17. The GenBank accession numbers of the HVRI and HV2 sequences are in Supplementary Table 1.

## RESULTS

### HPV and CC

The HPV DNA was found in 180 of 187 (96.3%) CC samples and in 42 of 270 (15.6%) cervical epithelial samples of healthy women. In cancer samples, HPV16 (46.5%) had the highest frequency, followed by HPV18 (13.4%), HPV45 (8.6%), HPV31 (4.8%), HPV58 (3.7%), HPV59 (3.2%), HPV33 (2.7%), HPV56 (2.7%), HPV52 (2.1%), HPV51 (2.1%), HPV35 (1.1%) and HPV66 (1.1%; see Table 1). The cumulative frequency of these 12 common viral types in women with CC was 92%. Other HPVs, with a frequency  $< 1\%$ , were found in the tumor sample, including HPV types 73, 53, 39, 6, 61, 68, 11 and 69, which account for the remaining 4.3% of HPV-positive samples (Supplementary Table 1). In the control women, HPV16 was also the most common viral type (4.8%), followed by HPV types 31, 58, 53, 18, 56, 66, 6, 61, 71, 45, 68, 29, 32, 70, 90 and 97, all of them with a frequency  $< 1.5\%$ .

The risk for CC associated to HPV infection was calculated by unconditional logistic regression and is shown in Table 1. The OR of all HPVs together was 152.9 (95% confidence interval (CI): 65.4–357.5). For individual HPV types, the highest OR was conferred by HPV45 (23.3; 95% CI: 3.0–178.0), followed by HPV18 (20.0; 95%

**Table 1** Frequency of 12 HPV types in cervical cancer and normal controls and its associated risk for cervical cancer

HPV type	Controls % (n), N=270	Cases % (n), N=187	OR (CI 95%) <sup>a</sup>	P-value <sup>b</sup>
HPV16	4.8 (13)	46.5 (87)	16.8 (8.9–31.5)	<0.01
HPV18	0.7 (2)	13.4 (25)	20.0 (4.6–86.0)	<0.01
HPV45	0.4 (1)	8.6 (16)	23.3 (3.0–178.0)	<0.05
HPV31	1.1 (3)	4.8 (9)	4.8 (1.2–18.2)	<0.05
HPV58	1.1 (3)	3.7 (7)	3.8 (0.96–15.0)	<0.05
HPV59	0.0 (0)	3.2 (6)	—	
HPV33	0.0 (0)	2.7 (5)	—	
HPV56	0.7 (2)	2.7 (5)	3.8 (0.73–20.3)	0.05
HPV52	0.0 (0)	2.1 (4)	—	
HPV51	0.4 (1)	2.1 (4)	6.4 (0.71–59.1)	0.1
HPV35	0.0 (0)	1.1 (2)	—	
HPV66	0.7 (2)	1.1 (2)	1.7 (0.24–12.7)	0.5
Other HPVs <sup>c</sup>	5.6 (15)	4.3 (8)	0.55 (0.37–0.82)	0.6
HPV positive	15.6 (42)	96.3 (180)	152.9 (65.4–357.5)	<0.05
HPV negative	84.4 (228)	3.7 (7)	1	

Abbreviations: CI, confidence interval; HPV, human papillomavirus; OR, odds ratio.  
<sup>a</sup>OR adjusted for age.

<sup>b</sup>Unconditional regression logistic  $P$ -value.

<sup>c</sup>This group included HPVs 6, 11, 39, 53, 61, 68, 69 and 73 in cases and HPVs 6, 29, 32, 53, 61, 68, 70, 71, 90 and 97 in controls.

CI: 4.6–86.0) and HPV16 (16.8; 95% CI: 8.9–31.5). However, the calculated risk of HPV16 and 18 had a  $P$ -value much lower than and confidence intervals narrower than HPV45 (see Table 1). It is worth noting that the OR of these three viruses is much higher than the OR of the other HPVs followed in order of frequency.

### mtDNA polymorphism and CC

A total of 192 polymorphic sites was identified in the HVRI and HVR2, including cases and controls. From these polymorphisms, 98 were found in both groups, 39 in the cancer group and 55 in healthy women. These polymorphisms belonged to 351 different haplotypes, 133 of which were found exclusively in cases and 188 only in controls (Supplementary Table 1). The mtDNA haplogroups were identified using the haplogrep tool as described in Materials and methods. The Amerindian mitochondrial ancestry predominated in the total sample with a proportion of 93.8%, followed by European (4.2%) and African (2.0%) ancestry. The frequencies of the Amerindian haplogroups A2, B2, C1 and D1 in the global sample were 48.4, 21.2, 18.1 and 6.1%, respectively. The most common haplotype had a frequency of 3.2% and belonged to the haplogroup A2 (Supplementary Table 1). The European mtDNA haplogroups identified were the haplogroups H, J, K, R, T and U, whereas the African included the haplogroups L2 and L3. The frequency of mtDNA haplogroups in healthy women and with CC is shown in Table 2. The frequency of haplogroup A2 was higher in control than in cancer patients (51.1% versus 44.4%), whereas the opposite was found for the haplogroup B2 (17.8% in controls vs 26.2% in cases). However, only the difference observed for haplogroup B2 was statistically significant ( $P=0.04$ , Pearson's  $\chi^2$ ) and the OR was 1.64 (95% CI: 1.05–2.58), suggesting that this haplogroup increased the risk for CC. It has been previously reported that the frequency of Amerindian haplogroups could vary across geographical region according to admixture history.<sup>23</sup> Nevertheless, the frequency of haplogroup B2 differed similarly between cases and controls across geographical regions (Supplementary Table 2).

The alleles of 10 polymorphic positions in HVRI are linked with particular haplogroups (Supplementary Table 3). Thus, the allelic

**Table 2** Frequency of mtDNA haplogroups in cervical cancer cases and normal controls and the risk for cervical cancer

mtDNA	Controls % (n), N=270	Cases % (n), N=187	OR 95% CI <sup>a</sup>	P-value <sup>b</sup>
<i>Haplogroup<sup>c</sup></i>				
A2	51.1 (138)	44.4 (83)	0.73 (0.52–1.11)	0.18
B2	17.8 (48)	26.2 (49)	1.64 (1.05–2.58)	0.04
C1	18.5 (50)	17.6 (33)	0.94 (0.58–1.53)	0.92
D1	5.5 (15)	7.0 (13)	1.27 (0.59–2.73)	0.8
European	5.2 (14)	2.7 (5)	0.50 (0.17–1.41)	0.27
African	1.8 (5)	2.1 (4)	1.15 (0.31–4.37)	0.72

Abbreviations: CI, confidence interval; mtDNA, mitochondrial DNA; OR, odds ratio.

<sup>a</sup>OR calculated by univariate unconditional logistic regression.

<sup>b</sup>Pearson  $\chi^2$  test P-value corrected for continuity.

<sup>c</sup>The Amerindian mtDNA haplogroups were grouped according to column 4 in Supplementary Table 1.

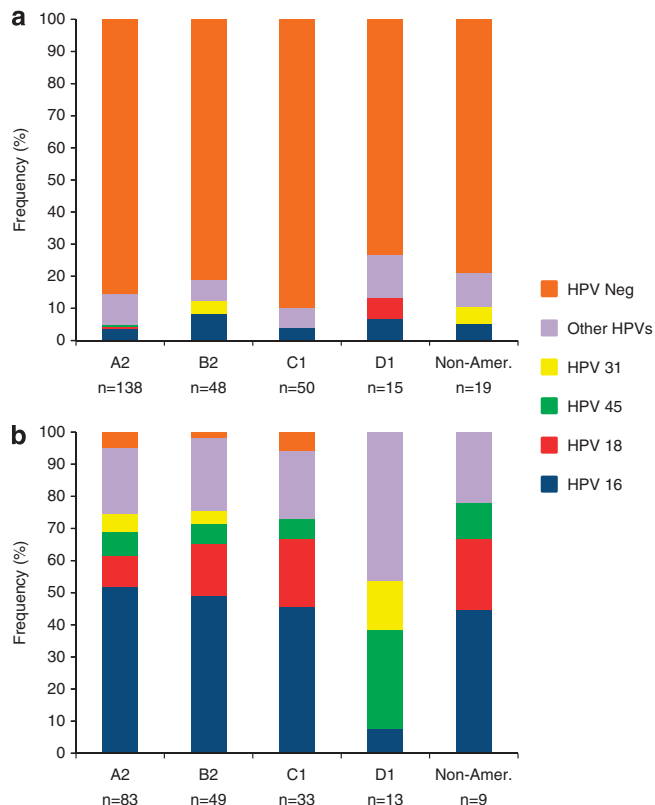
distribution of those individual positions was compared between cases and controls (Supplementary Table 4). In three positions (16362, 16183 and 16189), the allele distribution was different and statistically significant between the groups. The polymorphism 16362C, which is present mainly in haplogroup A2, was associated with controls, whereas the alleles 16183C and 16189C, present mainly in haplogroup B2, were associated with cancer patients (Supplementary Table 4).

#### Interaction of HPV infection with mtDNA polymorphism in CC risk

To identify whether the distribution of HPV viral types are related to human mtDNA ancestry, the frequencies of HPVs 16, 18, 45, 31 and other HPV types were compared among the four Amerindian mtDNA haplogroups in cases and controls. Interestingly, in cancer patients the distribution of HPV types was quite similar in all haplogroups but one (D1), in which is remarkable the absence of HPV18, a very low frequency of HPV16 and a high frequency of HPV45, HPV31 and other non-common HPV types ( $P < 0.01$ , Fisher's exact test; Figure 1b). In control patients, the haplogroup D1 showed the highest positivity for the HPVs, and it was 1.7-fold times higher than the positivity in the whole control sample (26.6% versus 15.6%;  $P > 0.05$ , Fisher's exact test). This increase was accounted mainly by the increase of HPV18 and other HPV types (Figure 1a). To evaluate whether mtDNA ancestry potentiate or attenuate the risk for CC of HPV infection, a stratified analysis by Amerindian mtDNA haplogroups was performed. The risk for CC conferred by the HPV infection alone (OR=152.9; 95% CI: 65.4–357.5; Table 1) increased in women with haplogroup B2 (OR=208; 95% CI: 25.2–1735.5;  $P < 0.0001$ , unconditional regression bivariate). This increase was even higher, nearly 50%, with the polymorphisms linked to B2 (like 16183C and 16189C; Table 3). In contrast, an estimation of the risk for CC conferred by the HPV infection would decrease at least sixfold in women with haplogroup D1, mainly by the increased frequency of HPV infection in the control group up to 26.6% (Figure 1a). The risk for CC conferred by the HPV infection was not influenced by haplogroups A2 and C1 (Table 3). However, these results should be taken cautiously owing to the small sample size of this study. To better confirm this results, a bigger sample size could be helpful to better define the interaction between HPV and mtDNA polymorphisms.

#### mtDNA gene expression analysis

If a polymorphism confers a higher risk for the disease, a change of gene expression could be expected in genes directly linked with it.



**Figure 1** Distribution of HPV types 16, 18, 45 and 31 and other HPV types among mtDNA Amerindian haplogroups in healthy controls (a) and CC patients (b). The Amerindian mtDNA haplogroups were grouped according to column 4 in Supplementary Table 1. The number of individuals in each haplogroup is beneath the frequency bar. Non-Amerindian includes European and African haplogroups. Other HPVs include viral types 6, 11, 33, 35, 39, 51, 52, 53, 56, 58, 59, 61, 66, 69, 68 and 73 in cases and HPVs 6, 29, 32, 51, 53, 56, 61, 66, 68, 70, 71, 90 and 97 in controls. Color codes are the same for cases and controls.

**Table 3** Cervical cancer risk associated to HPV infection among women with Amerindian mtDNA haplogroups and the polymorphisms 16362C, 16183C and 16189C

mtDNA	Frequency of HPV % (positive/explored)		OR (CI 95%) <sup>§</sup>
	Controls	Cases	
<i>Haplogroup<sup>b</sup></i>			
A2	14.5 (20/138)	95.1 (79/81)	160.6 (46.2–558.3)*
B2	18.8 (9/48)	98.0 (48/49)	208 (25.1–1735.5)**
C1	10.0 (5/50)	94.1 (31/33)	148.7 (25.4–868.4)*
D1	26.6 (4/15)	100.0 (13/13)	> 26.2 <sup>c</sup>
<i>Alleles</i>			
16362C	14.8 (24/162)	96.6 (90/94)	149.6 (47.8–468.2)*
16183C	18.0 (9/50)	98.0 (49/50)	222.9 (26.9–1843.2)**
16189C	16.1 (9/56)	98.3 (57/58)	295.6 (36–2425)**

Abbreviations: CI, confidence interval; HPV, human papillomavirus; mtDNA, mitochondrial DNA; OR, odds ratio.

<sup>a</sup>ORs were adjusted by age of women.

<sup>b</sup>The Amerindian mtDNA haplogroups were grouped according to column 4 in Supplementary Table 1.

<sup>c</sup>In the case of haplogroup D1, where all cases were HPV positive, an approximate OR was calculated considering one case as negative.

\* $P < 0.05$ .

\*\* $P < 0.01$ .

Therefore, it was investigated whether the expression level of some mitochondrial genes changed in tumor samples of the haplogroup B2 compared with control epitheliums and tumors with other haplogroups. A total of 29 mtDNA genes were analyzed for changes in the gene expression, comparing a subset of 45 HPV16-positive carcinomas with 17 healthy cervical epitheliums. Seven genes (*MT-TG*, *MT-TH*, *MT-TL2*, *MT-TC*, *MT-TY*, *MT-TN* and *MT-TA*) were found overexpressed in the whole tumor sample as compared with the control sample ( $P < 0.05$ ; Student's *t*-test; Table 4). When the analysis was performed by individual haplogroups, the difference was statistically significant only for the genes *MT-TH*, *MT-TC* and *MT-TY* in A2 tumors. However, no differences were found among the haplogroups (see Supplementary Figure S1A-C). On the other hand,

two different genes were found upregulated only in tumors from haplogroup B2 (mitochondrial aspartic acid tRNA (*MT-TD*), mitochondrial lysine tRNA (*MT-TK*); Table 4 and Figure 2a and b;  $P < 0.01$ ; *t*-test). Comparing among the tumors, the mean genetic expression of *MT-TD* and *MT-TK* genes was also significantly higher in haplogroup B2 compared with haplogroup A2 ( $P < 0.05$ , Student's *t*-test; Figures 2a and b). It is worth mentioning that these two genes belong to the same mtRNA transcript (transcript 16 in H strand; Figure 2d). On the other hand, although in this transcript, the cytochrome oxidase sub-unit II (*MT-CO2*) gene is flanked by these two tRNA genes, the genetic expression did not change compared with controls or among the cancer groups.

**Table 4** Fold change of mtDNA gene expression in the whole cases or grouped by mtDNA haplogroups

mtDNA gene	Fold change <sup>a</sup>				
	mtDNA haplogroups <sup>b</sup>				All tumors, N=45
	A2, N=23	B2, N=9	C1, N=9	D1, N=4	
<i>H-strand</i>					
<i>MT-RNR 1</i>	0.9	1.0	0.9	0.8	0.9
<i>MT-TV</i>	1.1	1.3	1.1	1.5	1.2
<i>MT-RNR 2</i>	1.0	1.1	1.1	1.0	1.0
<i>MT-ND 1</i>	1.1	1.2	1.1	1.1	1.1
<i>MT-TM</i>	1.1	1.4	1.1	1.4	1.2
<i>MT-ND2</i>	1.0	1.3	1.0	0.8	1.0
<i>MT-CO1</i>	1.0	1.1	1.0	1.0	1.0
<i>MT-TD</i>	1.0	2.1*	1.3	1.1	1.3
<i>MT-CO2</i>	1.0	1.1	1.1	1.1	1.1
<i>MT-TK</i>	1.2	1.8**	1.2	1.3	1.3
<i>MT-CO3</i>	1.0	1.1	1.2	1.1	1.1
<i>MT-ATP8</i>	1.0	1.1	1.2	1.1	1.1
<i>MT-TG</i>	1.6	2.4	2.1	1.0	1.8*
<i>MT-ND3</i>	1.0	1.1	1.0	1.0	1.0
<i>MT-ND4</i>	0.9	1.1	1.1	1.1	1.0
<i>MT-TH</i>	1.6*	1.5	1.5	1.3	1.5**
<i>MT-TS2</i>	1.2	1.1	1.2	1.2	1.2
<i>MT-TL2</i>	1.6	2.0	1.7	0.9	1.6*
<i>MT-ND5</i>	1.0	1.1	1.0	1.0	1.0
<i>MT-CYB</i>	1.0	1.1	1.1	1.0	1.0
<i>MT-TT</i>	1.3	1.7	0.9	0.9	1.3
<i>L-strand</i>					
<i>MT-TF</i>	0.9	1.1	0.9	1.1	1.0
<i>MT-TP</i>	1.3	1.4	1.2	1.2	1.3
<i>MT-TQ</i>	1.4	1.5	1.4	1.1	1.4
<i>MT-ND6</i>	1.2	1.3	1.4	1.5	1.3
<i>MT-TS1</i>	1.4	1.3	1.5	1.2	1.4
<i>MT-TC</i>	1.7*	1.6	1.6	1.2	1.6**
<i>MT-TY</i>	1.7*	1.6	1.6	1.2	1.6**
<i>MT-TN</i>	1.5	1.6	1.7	1.3	1.6**
<i>MT-TA</i>	1.5	1.4	1.6	1.3	1.5**
<i>MT-TE</i>	1.3	1.4	1.2	1.3	1.3

<sup>a</sup>Fold change: Mean probe intensity from cases divided by mean probe intensity from controls ( $n=17$ ).

<sup>b</sup>The Amerindian mtDNA haplogroups were grouped according to column 4 in Supplementary Table 1.

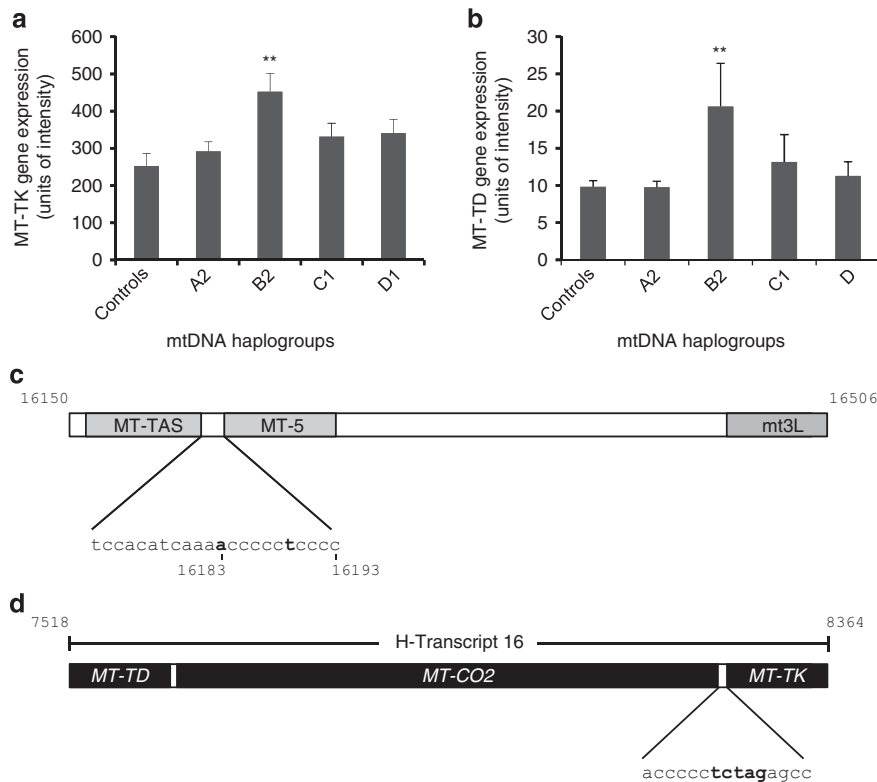
\*  $P$ -value  $< 0.05$ .

\*\* $P$ -value  $< 0.01$ , both calculated with the *t*-test.

## DISCUSSION

This is the first study in which an association between Amerindian mtDNA polymorphisms with CC and HPV viral types has been demonstrated. The Amerindian haplogroup B2 increased the risk for CC (OR=1.6) and showed an additive effect over the risk conferred by the HPV. Although the increment of the OR was modest, the upregulation of two mtDNA genes (*MT-TD*, *MT-TK*), exclusively in B2 tumors, strengthen its association with CC. The Amerindian mtDNA ancestry predominated in both healthy and cancer women, which is expected owing to the admixture history in Mexican mestizo population.<sup>23</sup> The mtDNA haplogroups diverged into three major branches: the paragroup L, which comprises the sub-Saharan African populations and the macrohaplogroups M and N, which represents the Eurasian populations.<sup>21</sup> Under this mtDNA classification, the Amerindian haplogroups B2 and A2 belong to the macrohaplogroup N, in addition to the West Eurasian haplogroups H, J, K, T, W, I, V and U. In contrast, the macrohaplogroup M contains the Amerindian haplogroups C1 and D1 with haplogroups M and G.<sup>21</sup> Each mtDNA haplogroup is distributed in defined groups of population, and the polymorphism within a population is given in the context of specific haplogroups.

Although mtDNA has been associated with different type of cancers, the association with CC is scarcely known. The few studies reported focused on specific positions instead of the haplogroups or haplotypes. In this study, the association of haplogroup B2 with CC, together with the association of different mtDNA polymorphism with other malignancies in diverse populations, suggests that associated mtDNA polymorphism is population and cancer specific.<sup>16</sup> For instance, in a Chinese population, the haplogroup D4a increased the risk of thyroid cancer (OR=3.0) and haplogroup D5 increased the risk of breast cancer (OR=3.1).<sup>20</sup> The haplogroup U raise the risk for prostate (OR=1.95) and renal cancer (OR=2.52) in white North Americans,<sup>19</sup> but not in European populations for prostate cancer.<sup>30</sup> Conversely, in these populations, the haplogroups I, H and the super-haplogroup UK were associated to an increased risk for breast, colon and vulvar cancers, respectively.<sup>16,31</sup> Whereas most of mtDNA polymorphism increased the cancer risk, others declined the risk in certain types of cancer. In West Eurasian populations, the risk for endometrial cancer is reduced with haplogroup H, which has the highest frequency in these populations.<sup>32</sup> In this study, the specific alleles associated with CC were those related to haplogroups B2 and A2. Alleles 16183C and 16189C were associated to an increased risk for CC. The allele 16189C is almost representative of haplogroup B2, although it can also be found in other haplogroups owing to homoplasic mutations.<sup>21</sup> For instance, in a previous report in a Chinese population, the allele 16189C, linked to haplogroups other than B2, was found to increase the risk for endometrial cancer, but not for CC.<sup>33</sup> This polymorphism is located into a microsatellite (16 184–16 193), and when the transi-



**Figure 2** Comparison of the mean expression intensity of genes upregulated and the topographic relation among the relevant B2 polymorphisms and some mtDNA sequence motifs. *MT-TD* (a) and *MT-TK* (b). Vertical bars show the mean expression intensity  $\pm$  s.e. The number of samples explored for each group is the same as in Table 4. The differences between B2 tumors and controls or A2 tumors were significant,  $P < 0.01$ ,  $t$ -test (\*\*). (c) Schematic representation of the *cis*-elements flanking the microsatellite 16184–16193. A partial D-loop region is depicted, spanning from positions 16150–16506. *Cis* elements (*MT-TAS*, *MT-5*, *mt3L*) are represented as gray boxes. Sequence shows the microsatellite 16183–16193, which include the polymorphic positions 16183 (C>A) and 16189 (C>T), shown in bold. (d) Genes surrounding the 9 bp intergenic deletion (*MT-CO2* and *MT-TK*) in haplogroup B2. The transcript 16 coding region with the *MT-TD*, *MT-CO2* and *MT-TK* genes are depicted. The intergenic sequence between *MT-CO2* and *MT-TK* is shown and the nine nucleotides that are deleted in haplogroup B2 are underlined. The mTFA consensus sequence is shown in bold.

tion C>T occurs, a long stretch of unstable cytosines is created.<sup>33,34</sup> This microsatellite is flanked by two regulatory sequences, a TAS (termination-associated sequence) element implicated in mtDNA replication and an MT5 element involved in mtDNA transcription (Figure 2c).<sup>15</sup> The allele 16183C has not been previously associated with any malignancy, but the effect could be similar to 16189C allele, as this position also belongs to the haplogroup B2 and is located into the same microsatellite. In contrast to the haplogroup A2, the allele 16362C, mainly exclusive of this haplogroup, was found to be protective for CC (OR=0.63;  $P$ -value  $< 0.02$ ). Haplogroup A2 is the most common haplogroup in Mexican as well as in other Latin American mestizo populations.<sup>35,36</sup>

Our results, in conjunction with other studies, shed light on the importance of the genetic background in the development of CC and HPV infection. Some nuclear loci have also been associated to variations in HPV positivity and distribution of viral types in different populations. For instance, in African populations the genotype –652 6N del/del from the *CASP8* gene was associated with high-risk HPV infection only in black women.<sup>37</sup> In a study performed in Latin America, where the functional polymorphisms of *MBL2* gene were analyzed, the deficient-producer genotypes were more frequent in women infected with high-risk viral types.<sup>38</sup> In Japanese women with cervical neoplasias, the *GSTT1*-null genotype had a greater frequency in those infected with HPV16/18 than with other HPVs.<sup>39</sup> In addition to those studies, in this work, the distribution of HPV types was

different for the haplogroup D1 in CC compared with other haplogroups. HPV45 and HPV31 predominate in D1, whereas HPV16 and HPV18 predominated in haplogroups A2, B2 and C1. As well, in healthy women, the D1 subset showed the highest HPV positivity and haplogroups B2 and C1 did not show infections with HPV18. Although the sample size is rather low for some haplogroups, results suggest the susceptibility to HPV infection, and viral types could be different among Amerindian mitochondrial haplogroups. If these results are confirmed in a greater sample, women with haplogroup D1 would not be protected by the available preventive vaccines against HPV16 and HPV18.

These different associations with cancer and viral infection suggest that individual haplogroups have distinct responses to environmental exposures. The phenotypic effects of some polymorphisms in mtDNA haplogroups have been related to mitochondria dynamics, oxygen consumption, production of reactive oxygen species and regulation of mitochondrial matrix pH.<sup>22,40</sup> It is not known whether mtDNA polymorphisms have been originated as a mechanism for adaptation of human beings to circumvent the environment across human history, as the divergence and spread out of modern humans from Africa. Although the HPV is the main factor for CC development, these results suggest that mtDNA genetic variants, like haplogroup B2, can modulate the risk for the disease and viral infection. An interaction between mitochondria and HPV has been demonstrated. It has been reported that the HPV protein E1<sup>A</sup>E4 binds to the mitochondria

external membrane to induce apoptosis and complete the viral cycle.<sup>41</sup> In addition, there is evidence that the HPV oncoprotein E7 can protect HaCat keratinocytes from reactive oxygen species damage.<sup>42</sup> Although the knowledge of mitochondria and its contribution to cancer is growing, further studies will clarify the role of mitochondria in HPV infection and CC.

The polymorphisms associated to cancer behave as recessive mutation with the subsequent loss of function, and they have been identified mainly in tumor suppressor genes or genes involved in DNA repair. Under this model, a second hit is necessary to inactivate the functional allele at the same locus.<sup>13</sup> In the case of mtDNA polymorphisms associated to cancer, they might behave always as dominant, but with either loss or gain of function.<sup>16,43</sup> However, when they have been tested in cell lines or tumoral models, all mtDNA mutations identified in cancer samples behaved dominant and with a gain of function.<sup>44,45</sup> Over 72% of the mitochondrial mutations reported in cancer have also been found as genetic variants in human populations.<sup>17</sup> Therefore, it is proposed that mitochondrial polymorphisms could mimic that the somatic mtDNA mutations originated in the tumor.<sup>18</sup>

The fact that over 20% of explored mtDNA genes (7 out of 29) were de-regulated in tumor samples, compared with the healthy normal epithelium, suggests that the mitochondria participate actively in the tumor process. It is worth mentioning that all de-regulated genes were upregulated and codify for tRNAs. However, none of them seems to be deregulated in any particular haplogroup. For instance, tRNA genes *MT-TH*, *MT-TC* and *MT-TY* were found associated with A2 tumors, when they were compared with the controls, but not when they were compared with the other tumor haplogroups.

The two upregulated genes in B2 tumors (*MT-TD*, *MT-TK*) belong to the same transcript (transcript 16),<sup>15</sup> and flank the gene of the *MT-CO2*. Interestingly, just after the end of this gene and before the *MT-TK*, a deletion of 9bp is found in a non-coding region (NC7; see Figure 2d).<sup>15</sup> Although this deletion was first used to identify the former haplogroup B, it has also been reported associated to some A2 haplogroups.<sup>21,46</sup> On the other hand, the unchanged expression of *MT-CO2* gene could be explained by differences in the post-transcriptional regulation. In contrast with the *MT-TK* and *MT-TD*, the *MT-CO2* transcript is a messenger RNA and its lifespan could be shorter than for the tRNAs.<sup>47</sup> Also, it is not known whether this deletion could have an advantage or be involved in RNA processing.<sup>15</sup> However, the fact that overexpressed genes (*MT-TD*, *MT-TK*) are physically linked with this deletion suggests some functional relation between them. For instance, the deletion includes a consensus motif (TCTAG) for mTFA close to the *MT-TK* gene. mTFA is a transcription factor that is involved in packing and maintenance of mtDNA.<sup>48</sup> The tRNAs are very important for the synthesis of proteins participating in the oxidative phosphorylation and deleterious mutations in these genes have serious consequences in the respiratory chain.<sup>49</sup> Higher expression levels of both nuclear and mitochondrial tRNAs, in breast cancer and cell lines derived from this tumor, have been reported.<sup>50</sup> The greater rate of tRNA synthesis could be a regulatory process to satisfy the higher translational demands in growing tumors. Indeed, one of the upregulated genes found in these tumor samples, the *MT-TL2*, has been reported as one of the most used for protein synthesis in the mitochondria.<sup>51</sup> In contrast, upregulation of MT-tRNAs is proposed as a mechanism to prevent apoptosis by the tumor cell.<sup>52</sup>

Although the results presented in this work showed an association of the mtDNA with CC, these results should be confirmed in a case-control study with a higher sample size. Besides, the mtDNA would be only one of the multiple loci that could be implicated in the risk for

CC. Further studies are needed to elucidate the mechanism related to haplogroup B2 and cancer predisposition.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

This work was supported by the National Council of Science and Technology (CONACYT), grant numbers 8135/A1, 24341 (to JB) and 80680 (to SK), and the National University of Mexico (UNAM), grant number SDI.PTID.05.2 (to JB). MGE and IMM were recipients of a scholarship from CONACYT.

**Author contributions:** All authors read and approved the final manuscript. GEM sequenced the HVRI and HVR2 control region in cases and controls, analyzed the sequence data, performed the statistics and drafts the manuscript. MMI evaluated global gene expression with HG 1.0 ST expression microarrays and analyzed the gene expression data. JTE and MMI carried out the detection and HPV typing. MA purified DNA. EA collected the control epithelium samples and purified RNA. RBE, MCS, and BIM evaluated clinically the cancer women and provided the biopsies, cervical scrapes and blood samples from the patients. GG, CTF and SL clinically evaluated the healthy women and provided the blood and cervical samples from control women. AA, AAV and AP confirmed the histopathological diagnosis in the biopsies from women with cancer. KS helped with data analysis, discussion and the manuscript. BJ design and coordinated the project, integrated and perform part of the data analysis, participate in the discussion and rewrote the manuscript.

- 1 Munoz, N., Bosch, F. X., de Sanjose, S., Herrero, R., Castellsague, X., Shah, K. V. *et al.* Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N. Engl. J. Med.* **348**, 518–527 (2003).
- 2 Schlecht, N. F., Platt, R. W., Duarte-Franco, E., Costa, M. C., Sobrinho, J. P., Prado, J. C. M. *et al.* Human papillomavirus infection and time to progression and regression of cervical intraepithelial neoplasia. *J. Natl. Cancer Inst.* **95**, 1336–1343 (2003).
- 3 Wang, S. S. & Hildesheim, A. Chapter 5: viral and host factors in human papillomavirus persistence and progression. *JNCI Monogr.* **2003**, 35–40 (2003).
- 4 Magnusson, P. K. E., Lichtenstein, P. & Gyllenstein, U. B. Heritability of cervical tumours. *Int. J. Cancer* **88**, 698–701 (2000).
- 5 Hemminki, K. & Chen, B. Familial risks for cervical tumors in full and half siblings: etiologic apportioning. *Cancer Epidemiol. Biomarkers Prev.* **15**, 1413–1414 (2006).
- 6 Zelmanowicz Ade, M., Schiffman, M., Herrero, R., Goldstein, A. M., Sherman, M. E., Burk, R. D. *et al.* Family history as a co-factor for adenocarcinoma and squamous cell carcinoma of the uterine cervix: results from two studies conducted in Costa Rica and the United States. *Int. J. Cancer* **116**, 599–605 (2005).
- 7 Couto, E. & Hemminki, K. Heritable and environmental components in cervical tumors. *Int. J. Cancer* **119**, 2699–2701 (2006).
- 8 Engelmark, M., Beskow, A., Magnusson, J., Erlich, H. & Gyllenstein, U. Affected sib-pair analysis of the contribution of HLA class I and class II loci to development of cervical cancer. *Hum. Mol. Genet.* **13**, 1951–1958 (2004).
- 9 de Araujo Souza, P. S., Sichero, L. & Maciag, P. C. HPV variants and HLA polymorphisms: the role of variability on the risk of cervical cancer. *Fut. Oncol.* **5**, 359–370 (2009).
- 10 Wang, S. S., Gonzalez, P., Yu, K., Porras, C., Li, Q., Safaeian, M. *et al.* Common genetic variants and risk for HPV persistence and progression to cervical cancer. *PLoS One* **5**, e8667 (2010).
- 11 Thakur, N., Hussain, S., Kahaar, I., Tabassum, R., Nasare, V., Tiwari, P. *et al.* Genetic variant of CCND1: association with HPV-mediated cervical cancer in Indian population. *Biomarkers* **14**, 219–225 (2009).
- 12 Ivansson, E. L., Gustavsson, I. M., Magnusson, J. J., Steiner, L. L., Magnusson, P. K. E., Erlich, H. A. *et al.* Variants of chemokine receptor 2 and interleukin 4 receptor, but not interleukin 10 or Fas ligand, increase risk of cervical cancer. *Int. J. Cancer* **121**, 2451–2457 (2007).
- 13 Fletcher, O. & Houlston, R. S. Architecture of inherited susceptibility to common cancer. *Nat. Rev. Cancer* **10**, 353–361 (2010).
- 14 Yang, Y.-C., Chang, T.-Y., Lee, Y.-J., Su, T.-H., Dang, C.-W., Wu, C.-C. *et al.* HLA-DRB1 alleles and cervical squamous cell carcinoma: experimental study and meta-analysis. *Hum. Immunol.* **67**, 331–340 (2006).
- 15 Fernandez-Silva, P., Enriquez, J. A. & Montoya, J. Replication and transcription of mammalian mitochondrial DNA. *Exp. Physiol.* **88**, 41–56 (2003).

- 16 Bayona-Bafaluy, M. P., López-Gallardo, E., Montoya, J. & Ruiz-Pesini, E. Maternally inherited susceptibility to cancer. *Biochim. Biophys. Acta* **1807**, 643–649 (2011).
- 17 Brandon, M., Baldi, P. & Wallace, D. C. Mitochondrial mutations in cancer. *Oncogene* **25**, 4647–4662 (2006).
- 18 Singh, K. K. & Kulawiec, M. Mitochondrial DNA polymorphism and risk of cancer. *Methods Mol. Biol.* **471**, 291–303 (2009).
- 19 Booker, L. M., Habermacher, G. M., Jessie, B. C., Sun, Q. C., Baumann, A. K., Amin, M. *et al.* North American white mitochondrial haplogroups in prostate and renal cancer. *J. Urol.* **175**, 468–473 (2006).
- 20 Fang, H., Shen, L., Chen, T., He, J., Ding, Z., Wei, J. *et al.* Cancer type-specific modulation of mitochondrial haplogroups in breast, colorectal and thyroid cancer. *BMC Cancer* **10**, 421 (2010).
- 21 Maca-Meyer, N., Gonzalez, A., Larruga, J., Flores, C. & Cabrera, V. Major genomic mitochondrial lineages delineate early human expansions. *BMC Genet.* **2**, 13 (2001).
- 22 Pello, R., Martín, M. A., Carelli, V., Nijtmans, L. G., Achilli, A., Pala, M. *et al.* Mitochondrial DNA background modulates the assembly kinetics of OXPHOS complexes in a cellular model of mitochondrial disease. *Hum. Mol. Genet.* **17**, 4001–4011 (2008).
- 23 Guardado-Estrada, M., Juárez-Torres, E., Medina-Martínez, I., Wegier, A., Macias, A., Gomez, G. *et al.* A great diversity of Amerindian mitochondrial DNA ancestry is present in the Mexican mestizo population. *J. Hum. Genet.* **54**, 695–705 (2009).
- 24 Lazcano-Ponce, E., Herrero, R., Muñoz, N., Cruz, A., Shah, K. V., Alonso, P. *et al.* Epidemiology of HPV infection among Mexican women with normal cervical cytology. *Int. J. Cancer* **91**, 412–420 (2001).
- 25 Shepherd, J. H. Revised FIGO staging for gynaecological cancer. *BJOG* **96**, 889–892 (1989).
- 26 Manos, M. M., Ting, Y., Wright, D. K., Lewis, A. J., Broker, T. R. & Wolinski, S. M. in *The Use of Polymerase Chain Reaction Amplification for the Detection of Genital Human Papillomaviruses* (eds Furth, M. & Greaves, M. F.) 209–214 (Cold Spring Harbor, New York, 1989).
- 27 van den Brule, A. J., Meijer, C. J., Bakels, V., Kenemans, P. & Walboomers, J. M. Rapid detection of human papillomavirus in cervical scrapes by combined general primer-mediated and type-specific polymerase chain reaction. *J. Clin. Microbiol.* **28**, 2739–2743 (1990).
- 28 Yoshikawa, H., Kawana, T., Kitagawa, K., Mizuno, M., Yoshikura, H. & Iwamoto, A. Detection and typing of multiple genital human papillomaviruses by DNA amplification with consensus primers. *Cancer Sci.* **82**, 524–531 (1991).
- 29 Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J. *et al.* A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res.* **38**, W695–W699 (2010).
- 30 Mueller, E. E., Eder, W., Mayr, J. A., Paulweber, B., Sperl, W., Horninger, W. *et al.* Mitochondrial haplogroups and control region polymorphisms are not associated with prostate cancer in Middle European Caucasians. *PLoS One* **4**, e6370 (2009).
- 31 Klemba, A., Kowalewska, M., Kukwa, W., Tonska, K., Szybinska, A., Mossakowska, M. *et al.* Mitochondrial genotype in vulvar carcinoma—cuckoo in the nest. *J. Biomed. Sci.* **17**, 73 (2010).
- 32 Czarnecka, A., Klemba, A., Semczuk, A., Plak, K., Marzec, B., Krawczyk, T. *et al.* Common mitochondrial polymorphisms as risk factor for endometrial cancer. *Int. Arch. Med.* **2**, 33 (2009).
- 33 Liu, V. W. S., Wang, Y., Yang, H.-J., Tsang, P. C. K., Ng, T.-Y., Wong, L.-C. *et al.* Mitochondrial DNA variant 16189T>C is associated with susceptibility to endometrial cancer. *Hum. Mutat.* **22**, 173–174 (2003).
- 34 Wang, Y. U. E., Liu, V. W. S., Ngan, H. Y. S. & Nagley, P. Frequent occurrence of mitochondrial microsatellite instability in the D-loop region of human cancers. *Ann. N Y Acad. Sci.* **1042**, 123–129 (2005).
- 35 Barbosa, A., Antonio, L., Dalmo, d.S., Valdir, A. & Luiz, B. Mitochondrial DNA control region polymorphism in the population of Alagoas State, North-Eastern Brazil. *J. Forensic Sci.* **53**, 142–146 (2008).
- 36 Salas, A., Acosta, V., Ivarez-Iglesias, M., Cerezo, C., Phillips, M. V. & Carracedo, L. The mtDNA ancestry of admixed Colombian populations. *Am. J. Hum. Biol.* **20**, 584–591 (2008).
- 37 Chatterjee, K., Williamson, A.-L., Hoffman, M. & Dandara, C. CASP8 promoter polymorphism is associated with high-risk HPV types and abnormal cytology but not with cervical cancer. *J. Med. Virol.* **83**, 630–636 (2011).
- 38 Guimaraes, V., Guimaraes, R., Brandao, L., Baldez da Silva, M.F.P.T., Milanese, M., Segat, L. *et al.* Association between MBL2 gene functional polymorphisms and high-risk human papillomavirus infection in Brazilian women. *Hum. Immunol.* **69**, 273–278 (2008).
- 39 Ueda, M., Toji, E., Nunobiki, O., Sato, N., Izuma, S., Torii, K. *et al.* Germline polymorphisms of glutathione-S-transferase GSTM1, GSTT1 and p53 codon 72 in cervical carcinogenesis. *Hum. Cell* **23**, 119–125 (2010).
- 40 Martínez-Redondo, D., Marcuello, A., Casajús, J. A., Ara, I., Dahmani, Y., Montoya, J. *et al.* Human mitochondrial haplogroup H: the highest VO2max consumer—Is it a paradox? *Mitochondrion* **10**, 102–107 (2010).
- 41 Raj, K., Berguerand, S., Southern, S., Doorbar, J. & Beard, P. E1<sup>E4</sup> protein of human papillomavirus type 16 associates with mitochondria. *J. Virol.* **78**, 7199–7207 (2004).
- 42 Shim, J. H., Kim, K. H., Cho, Y. S., Choi, H. S., Song, E. Y., Myung, P. K. *et al.* Protective effect of oxidative stress in HaCAT keratinocytes expressing E7 oncogene. *Amino Acids* **34**, 135–141 (2008).
- 43 Solaini, G., Sgarbi, G. & Baracca, A. Oxidative phosphorylation in cancer cells. *Biochim. Biophys. Acta* **1807**, 534–542 (2011).
- 44 Koshikawa, N., Hayashi, J.-I., Nakagawara, A. & Takenaga, K. Reactive oxygen species-generating mitochondrial DNA mutation up-regulates hypoxia-inducible factor-1 alpha gene transcription via phosphatidylinositol 3-kinase-Akt/protein kinase C/histone deacetylase pathway. *J. Biol. Chem.* **284**, 33185–33194 (2009).
- 45 Ishikawa, K. & Hayashi, J.-I. A novel function of mtDNA: its involvement in metastasis. *Ann. N Y Acad. Sci.* **1201**, 40–43 (2010).
- 46 Achilli, A., Perego, U. A., Bravi, C. M., Coble, M. D., Kong, Q.-P., Woodward, S. R. *et al.* The phylogeny of the four Pan-American mtDNA haplogroups: implications for evolutionary and disease studies. *PLoS One* **3**, e1764 (2008).
- 47 Shutt, T. E. & Shadel, G. S. A compendium of human mitochondrial gene expression machinery with links to disease. *Environ. Mol. Mutagen.* **51**, 360–379 (2010).
- 48 Chen, X. J. & Butow, R. A. The organization and inheritance of the mitochondrial genome. *Nat. Rev. Genet.* **6**, 815–825 (2005).
- 49 Boulet, L., Karpati, G. & Shoubridge, E. A. Distribution and threshold expression of the tRNA(Lys) mutation in skeletal muscle of patients with myoclonic epilepsy and ragged fibers (MERRF). *Am. J. Hum. Genet.* **51**, 1187–1200 (1992).
- 50 Pavon-Eterod, M., Gomes, S., Geslain, R., Dai, Q., Rosner, M. R. & Pan, T. tRNA over-expression in breast cancer and functional consequences. *Nucleic Acids Res.* **37**, 7268–7280 (2009).
- 51 Swire, J., Judson, O. P. & Burt, A. Mitochondrial genetic codes evolve to match amino acid requirements of proteins. *J. Mol. Evol.* **60**, 128–139 (2005).
- 52 Mei, Y., Yong, J., Liu, H., Shi, Y., Meinkoth, J., Dreyfuss, G. *et al.* tRNA binds to cytochrome c and inhibits caspase activation. *Mol. Cell* **37**, 668–678 (2010).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)