Received: 29 January 2021 | Revised: 19 March 2021 | Accepted: 31 March 2021

DOI: 10.1002/ijc.33614

CANCER GENETICS AND EPIGENETICS

SHORT REPORT

Association of genomic variants at *PAX8* and *PBX2* with cervical cancer risk

Dhanya Ramachandran[1] | Yingying Wang[1] | Peter Schürmann[1] | Fabienne Hülse[1] | Qianqian Mao[1] | Matthias Jentschke[1] | Gerd Böhmer[2] | Hans-Georg Strauß[3] | Christine Hirchenhain[4] | Monika Schmidmayr[5] | Florian Müller[6] | Ingo Runnebaum[7] | Alexander Hein[8] | Martin Koch[8] | Matthias Ruebner[8] | Matthias W. Beckmann[8] | Peter A. Fasching[8] | Alexander Luyten[9,10] | Matthias Dürst[7] | Peter Hillemanns[1] | Thilo Dörk[1]

[1]Department of Gynaecology, Comprehensive Cancer Center, Hannover Medical School, Hannover, Germany

[2]IZD Hannover, Hannover, Germany

[3]Department of Gynaecology, University Clinics, Martin-Luther University, Halle-Wittenberg, Germany

[4]Department of Gynaecology, Clinics Carl Gustav Carus, University of Dresden, Dresden, Germany

[5]Department of Gynaecology, Technische Universität München, Munich, Germany

[6]Martin-Luther Hospital, Charite University, Berlin, Germany

[7]Department of Gynecology, Jena University Hospital, Friedrich -Schiller-University Jena, Jena, Germany

[8]Department of Gynecology and Obstetrics, Erlangen University Hospital, Comprehensive Cancer Center Erlangen-EMN, Friedrich Alexander University of Erlangen—Nuremberg (FAU), Erlangen, Germany

[9]Dysplasia Unit, Department of Gynecology and Obstetrics, Mare Klinikum, Kronshagen, Germany

[10]Department of Gynecology, Wolfsburg Hospital, Wolfsburg, Germany

Correspondence
Thilo Dörk, Hannover Medical School, Gynaecology Research Unit, Carl-Neuberg-St. 1, D-30625 Hannover, Germany.
Email: doerk.thilo@mh-hannover.de

Funding information
Bruno and Helene Jöster Foundation

Abstract

Cervical malignancy is triggered by human papillomavirus infection but the risk for cervical cancer has a hereditary component. From a recent Genome Wide Association Study meta-analysis, 2q14.1 (*PAX8*) and 6p21.32 (*PBX2*) have been proposed as novel cervical cancer susceptibility loci. We investigated the two main signals at these loci in an independent case-control series of 2578 cases with cervical dysplasia or carcinoma and 1483 healthy females. We find significant associations for both variants, rs10175462 at *PAX8* and rs2856437 at *PBX2*, with overall cervical disease (rs10175462: odds ratio [OR] 0.82, 95% confidence interval [CI] 0.74-0.91, $P = 2.4 \times 10^{-4}$; rs2856437: OR 1.52, 95% CI 1.14-2.02, $P = .004$). Both variants showed evidence of association with invasive

Abbreviations: cDNA, complementary DNA; CI, confidence interval; CIN, cervical intraepithelial neoplasia; CIS, carcinoma in situ; EDTA, ethylenediamine tetraacetic acid; eQTL, expression quantitative trait locus; GWAS, Genome Wide Association Study; HLA, human leukocyte antigen; HPV, human papillomavirus; HWE, Hardy Weinberg equilibrium; ICC, invasive cervical cancer; MAF, minor allele frequency; MHC, major histocompatibility complex; OR, odds ratio; RT-qPCR, reverse transcription followed by quantitative real time PCR; RFLP, restriction fragment length polymorphism; SNP, single nucleotide polymorphism.

Dhanya Ramachandran and Yingying Wang shared first authorship.

squamous cervical cancer (rs10175462: OR 0.80, 95% CI 0.68-0.94, $P = .006$; rs2856437: OR 1.56, 95% CI 1.03-2.36, $P = .036$) and with high-grade dysplasia (rs10175462: OR 0.79, 95%CI 0.70-0.90, $P = 1.9 \times 10^{-4}$; rs2856437: OR 1.58, 95% CI 1.15-2.17, $P = .005$). A combined analysis of high-grade dysplasia and invasive cervical cancer also showed significant associations for both variants (rs10175462: OR 0.81, 95% CI 0.73-0.91, $P = 2.4 \times 10^{-4}$; rs2856437: OR 1.57, 95% CI 1.18-2.10, $P = .002$). No association was detected for rs2856437 with low-grade dysplasia, while rs10175462 showed weak evidence of association ($P = .05$). RNA analyses in cervical samples revealed that *PAX8* transcripts were upregulated in HPV-positive lesions ($P = .008$) but this was not observed in the presence of the protective minor allele of rs10175462. The rs10175462 genotype also correlated with reduced levels of the lncRNA *PAX8-AS1* ($P < .001$). Taken together, our results extend the evidence for a link between genomic risk variants at the HLA region (*PBX2*) with cervical disease and support *PAX8* as the first consistent non-HLA cervical cancer susceptibility locus.

KEYWORDS
association study, cervical malignancy, eQTL, HPV infection, single nucleotide polymorphism

1 | INTRODUCTION

Cervical cancer, the fourth most prevalent form of cancer in females worldwide,[1] is induced by human papilloma-virus infection in almost all cases, with the high-risk HPV subtypes 16 and 18 being the major triggers of cervical cancer.[2] However, when HPV infection of the cervical epithelium results in cervical intraepithelial neoplasia (CIN), this can be transient or remain latent in many women, and only in a fraction of women with persistent infection of high-risk HPV, the neoplasia progresses to cervical cancer, indicating that additional factors are involved. There is an increased familial relative risk for cervical cancer,[3-5] and array heritability estimates are in the similar range as for colon or ovarian cancer.[6,7] Despite this evidence for hereditary genetic factors, only one genomic locus, HLA on chromosome 6, has been consistently associated with cervical cancer risk so far.[8-13] The HLA region modulates the host response to viral infection,[14,15] and it has been proposed that risk variants at this locus regulate genes that are important for cervical tumour immune evasion.

Recently, a large pan-cancer study has detected genomic variants and shared genetic basis across several cancers in the population-based UK Biobank and the Kaiser Permanente Genetic Epidemiology Research on Adult Health and Aging (GERA) cohorts.[7] In this meta-analysis, two variants have been reported to be specifically associated with cervical cancer at genome-wide significance: rs10175462 in *PAX8* on chromosome 2q14.1, and rs2856437 in *PBX2* on chromosome 6p21.32.[7] While rs2856437 appears to be an independent signal within the previously reported HLA region, the variant rs10175462, intronic to *PAX8*, represents a novel candidate risk region. *PAX8* encodes a member of the Pax family of transcription factors with oncogenic potential[16] that drives the development of epithelial cancers, including endometrial carcinoma[17] and ovarian carcinoma.[18]

What's new?

High-risk HPV is a major trigger of cervical cancer, but only some women are susceptible. Here, the authors investigated two genetic variants recently identified in genome-wide association studies as possibly related to cervical cancer risk. The variants are located in the genes *PAX8* and *PBX2*. This study tested the association in a German population containing 1122 cases of invasive cervical cancer, 1384 cases of cervical dysplasia, and 1483 controls. Both variants showed an association with cervical cancer and high-grade cervical dysplasia. RNA analysis also revealed that the *PAX8* variants influenced the amount of transcript found in cervical samples.

In the present work, we performed a genetic association study for these two variants to validate their role in a large case-control series for cervical cancer and dysplasia that had been established by the German Cervigen consortium, and we additionally tested the hypothesis that the signal represented by rs10175462 may regulate the RNA transcript levels of genes at the *PAX8* locus in cervical tissue.

2 | MATERIALS AND METHODS

2.1 | Patients

The German Cervigen Study has been described previously.[13] 4061 samples were used for the present case-control analysis, after

exclusion of patients with known non-European ancestry. We included 1122 cases with invasive cervical cancer and 1384 cases with cervical dysplasia (1037 CIN3, 220 CIN2, 127 CIN1) recruited from nine German hospitals in Hannover, Wolfsburg, Jena, Erlangen, Dresden, Halle, Munich, Berlin, and Bad Münder. Patients had been tested for their HPV status with either a diagnostic real-time PCR assay (Abbott) or, before 2015, with the Hybrid Capture Assay HC2 (Digene). For comparison, a total of 1483 healthy female controls were used that originated from the two centres in Hannover and Erlangen. These blood donors were unselected and cancer-free at the time of recruitment. Median age at diagnosis was 44 years (range 17-94 years) for patients with invasive cervical cancer and 31 years (range 16-79 years) for patients with cervical dysplasia, compared to a median age at recruitment of 32 years (range 18-86 years) for healthy female controls. The sample and histology distribution per centre have been described previously.[13] 5 mL peripheral venous EDTA blood was taken for genomic DNA extraction, and methanol-fixed cervical tissue smears were obtained from a smaller cohort of independent healthy participants without invasive cancer.

In the latter cohort of 317 cervical tissue samples, 91 were found to be HPV positive and 226 were HPV negative in a diagnostic real-time PCR assay that detects 14 carcinogenic HPV types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68) (Abbott). Further stratification by HPV sub-type shows that 38 samples were HPV16 positive, 11 samples were HPV18 positive, and 42 samples were infected by other strains of HPV. Of the 226 women with HPV-negative status, 205 had no sign for a cytologically or histologically detectable lesion ($r^2 = 0.39$). This "$HPV^{-}Lesion^{-}$" subgroup was analysed separately in comparison with the "$HPV^{+}Lesion^{+}$" subgroup that only contained HPV positive tissues with a documented lesion (CIN1-3 or PAPIII+, N = 64).

2.2 | SNP genotyping

Genomic DNA from peripheral white blood cells was extracted using the standard phenol-chloroform method. Fluidigm SNPtype assays with allele-specific probes labelled with FAM or HEX dyes were designed for genotyping of the two variants rs10175462 and rs2856437 (Fluidigm SNPtype assay IDs: GTA0267345 and GTA0267376). Both variants are purine transitions with MAFs of 0.38 and 0.04, respectively, in the GnomAD database. Two non-template controls were taken as negative controls. Example cluster plots are shown in Supplemental Figure 1. SNPtype genotyping results had already been obtained in this series for two previously tested signals at the HLA locus, rs9272117 (*HLA-DQA1*) and rs2844511 (*MICA/ HCP5*)[13] so that these genotypes were additionally used with rs2856437 in a multivariate analysis.

Genotyping for the *PAX8* variant rs10175462 was additionally validated using RFLP analysis in 1333 samples as an alternative probe-free method. For this purpose, we amplified a 154 bp fragment including rs10175462 with the primer pair 5′-GCCTAACCATGCCCTCTTAC-3′ and 5′-CAGTCAACATGGGCCTATGC-3′ and digested the PCR product with *Eco*NI (New England BioLabs) or with *Xag*I (Thermo Fisher Scientific). Both isoschizomers performed similarly well. Cleavage products at 115 bp and 39 bp were separated on 2% agarose gels and were visualised after GelRed staining on a UV transilluminator (Supplemental Figure 1). There was 99.1% concordance between SNPtype results and RFLP results in the case-control association study. For 11 discrepant samples with suboptimal clustering in the SNPtype assay, the RFLP results were taken for statistical analysis. Genotypes of the methanol-fixed cervical tissue samples that were used in the eQTL study showed a somewhat lower concordance of 92.1%. Also here, the RFLP results were taken in case of discrepant samples.

2.3 | Statistical analysis

Both the variants fulfilled Hardy-Weinberg equilibrium in Goodness-of-fit chi-square tests for both cases and controls. We then carried out logistic regression analyses to calculate odds ratios (ORs), *P*-values and 95% confidence intervals (CIs) under an additive model, with case-control status as the outcome and variant genotype as the predictor variables, using the STATA12 software package. ORs are given relative to the common homozygous (GG) genotypes for both variants (note that the hypothesis-generating study used the opposite definition[7]). We restricted all analyses to study participants with questionnaire-based European ancestry (defined as both parents being European). All analyses were age-adjusted. We further performed stratified analyses for the invasive, high-grade dysplasia or low-grade dysplasia case groups in comparison to all controls. For this purpose, we grouped CIN1 patients together with CIN2 cases at age < 30 years ($CIN2_{<30}$), and CIN2 cases at age ≥ 30 years ($CIN2_{\geq 30}$) together with CIN3 patients. A combined analysis was performed for CIN2 at age ≥ 30 together with CIN3 and with invasive cases. Two-sided *P*-values below 0.025 were considered significant in the main analyses and two-sided *P*-values below .005 were considered significant in the subgroup analyses. For rs2856437, we also performed a multivariate logistic regression analysis by incorporating two additional HLA variants into the model (rs2844511 and rs9272117) that we had previously genotyped with SNPtype assays on the same platform.[13]

2.4 | Transcript analysis

Total RNA was extracted from methanol-fixed cervical tissue samples of 317 healthy females who underwent routine HPV testing at Hannover Medical School using guanidinium-phenol-chloroform based extraction with Trizol reagent (peqGOLD TriFast).[13] At the same time, genomic DNA for SNP genotyping was extracted from these samples via the M24 SP robot (Abbott). 1ug RNA per sample was reverse transcribed into cDNA using the ProtoScript II First Strand cDNA Synthesis Kit (New England BioLabs). Fluidigm DeltaGene assays were designed for the genes *PAX8*, *PAX8-AS1*, and *PSD4* (Assay IDs GEP00114276, GEP00114265, and GEP00114264), and for housekeeping genes *B2M* and *RPL13A*, as reported previously.[13] The pre-amplified cDNA, after purification, was used for RT-qPCR analysis in

48x48 integrated fluidic circuit (IFC) plates on a BioMark HD real-time PCR instrument (Fluidigm). The stability of housekeeping genes was determined by qBASE+ (Biogazelle)[19] and GeNorm[20] and relative gene quantities were calculated with *B2M* and *RPL13A* as housekeeping controls. We obtained good-quality RT-qPCR results from 179 samples for *PAX8*, 184 samples for *PAX8-AS1* and 246 samples for *PSD4*.

Prior to statistical analysis, outliers were detected and excluded from each dataset via the ROUT method (1% false discovery rate) on Graphpad Prism v9.0.0. A *P*-value <.05 was considered significant in the Student's *t*-test (between two groups) and ANOVA (between 3 or more groups). Association of HPV status with gene expression was tested and for high-risk HPV subtypes 16 and 18, stratified analysis was conducted. Pearson correlation coefficients (R) were calculated for pairwise combinations of relative gene quantities using GraphPad Prism v9.0.0. In order to investigate the role of rs10175462 as an eQTL, relative gene quantities from cDNA samples were examined for association with their genotypes in the corresponding genomic DNA samples and examined overall as well as after HPV status based stratification.

3 | RESULTS

We obtained genotypes for rs10175462 at the *PAX8* locus from 3788 individuals (2524 cases and 1264 controls) and for rs2856437 at the *PBX2* locus from 3854 individuals (2445 cases and 1409 controls). Both variants showed significant associations with overall cervical disease status (rs10175462: OR 0.82, 95%CI 0.74-0.91, $P = 2.4 \times 10^{-4}$; rs2856437: OR 1.52, 95%CI 1.14-2.02, $P = .004$; Table 1). When cases were stratified by disease severity, there was evidence of association with invasive cervical cancer for both variants (rs10175462: OR 0.86, 95% CI 0.74-0.99, $P = .041$; rs2856437: OR 1.63, 95% CI 1.12-2.36, $P = .010$), and more specifically with squamous cervical cancer (rs10175462: OR 0.80, 95% CI 0.68-0.94, $P = .006$; rs2856437: OR 1.56, 95% CI 1.03-2.36, $P = .036$) (Table 1). There was also a significant association with high-grade dysplasia for both variants (rs10175462: OR 0.79, 95% CI 0.70-0.90, $P = 1.9 \times 10^{-4}$; rs2856437: OR 1.58, 95% CI 1.15-2.17, $P = .005$). Variant rs10175462, but not rs2856437, showed weak evidence of association with low-grade dysplasia (OR 0.814, 95% CI 0.66-1.00, $P = .054$). In a combined analysis of high-grade dysplasia and invasive cervical cancer, we found significant associations for both variants (rs10175462: OR 0.81, 95% CO 0.73-0.91, $P = 2.4 \times 10^{-4}$; rs2856437: OR 1.57, 95% CI 1.18-2.10, $P = .002$) (Table 1). Variant rs2856437 remained independently associated after adjustment for two previously tested risk factors at the HLA locus, rs2844511 and rs9272117,[13] in a multivariate regression analysis ($P = .024$).

Mining of the GTeX and SNiPA databases revealed that rs10175462 is a known expression quantitative trait locus (eQTL) for *PAX8*, *PAX8-AS1* and *PSD4* in several tissues.[21] To get further insight into a putative regulatory role of this novel locus, we tested the transcript levels of *PAX8* and two further candidate genes (*PAX8-AS1*, *PSD4*) after normalization to two housekeeping genes (*B2M* and *RPL13A*) by RT-qPCR from 317 cervical tissue smear samples. We found that *PAX8*, but not *PAX8-AS1* or *PSD4*, was upregulated in HPV-positive lesions compared to HPV-negative cervical tissue smears ($P = .008$, Figure 1A). Interestingly, the upregulation of *PAX8* was only observed for the main genotype but not for samples heterozygous or homozygous for the protective allele of rs10175462 (Figure 1B). When we stratified transcript levels by genotype, the rare allele of rs10175462 was also associated with markedly reduced steady-state levels of *PAX8-AS1* in a dose-dependent manner (ANOVA $P < .001$, independent of HPV status) while no such effect was detected for *PAX8* itself, and only a marginal decrease was observed for *PSD4* (Figure 2A-C).

TABLE 1 Association of rs10175462 and rs2856437 with case-control status

	rs10175462				rs2856437			
Stratum	N_{cases}	$N_{controls}$	**OR (95% CI)**	P_{reg}	N_{cases}	$N_{controls}$	**OR (95% CI)**	P_{reg}
CIN1 + $CIN2_{< 30}$	233	1264	0.81 (0.66–1.00)	.054	220	1409	1.07 (0.56-2.01)	.844
$CIN2_{\geq 30}$ + CIN3	1128	1264	0.79 (0.70-0.90)	1.9×10^{-4}	1078	1409	1.58 (1.15-2.17)	.005
Invasive	1099	1264	0.86 (0.74–0.99)	.041	1083	1409	1.63 (1.12–2.36)	.010
SC	692	1264	0.80 (0.68–0.94)	.006	685	1409	1.56 (1.03–2.36)	.036
AC	189	1264	0.95 (0.75-1.21)	.666	185	1409	1.53 (0.83-2.81)	.170
$CIN2_{\geq 30}$ + CIN3 + Invasive	2227	1264	0.81 (0.73-0.91)	2.4×10^{-4}	2161	1409	1.57 (1.18-2.10)	.002
Overall	2524	1264	0.82 (0.74-0.91)	2.4×10^{-4}	2445	1409	1.52 (1.14-2.02)	.004

Note: Stratified logistic regression analysis of rs10175462 (*PAX8*) and rs2856437 (*PBX2*) in the Cervigen case-control series. Cervical intraepithelial neoplasia was differentiated into low-risk dysplasia (CIN1 + $CIN2_{<30}$) and high-risk dysplasia ($CIN2_{\geq 30}$ + CIN3) groups. Invasive cervical cancer was stratified into squamous epithelial cell carcinoma (SC) or adenocarcinoma (AC). High-risk dysplasia ($CIN2_{\geq 30}$ + CIN3) and invasive cancer were further combined for joint analysis. Overall: Total number of cases, including further 64 histologically unclassified dysplasias.

Abbreviations: CI, confidence interval; N_{cases}, number of successfully genotyped cases per stratum; $N_{controls}$, number of successfully genotyped controls; OR, odds ratio for the minor allele, with healthy controls as the reference group; P_{reg}, *P* value from logistic regression analysis of cases within the stratum compared to all controls.

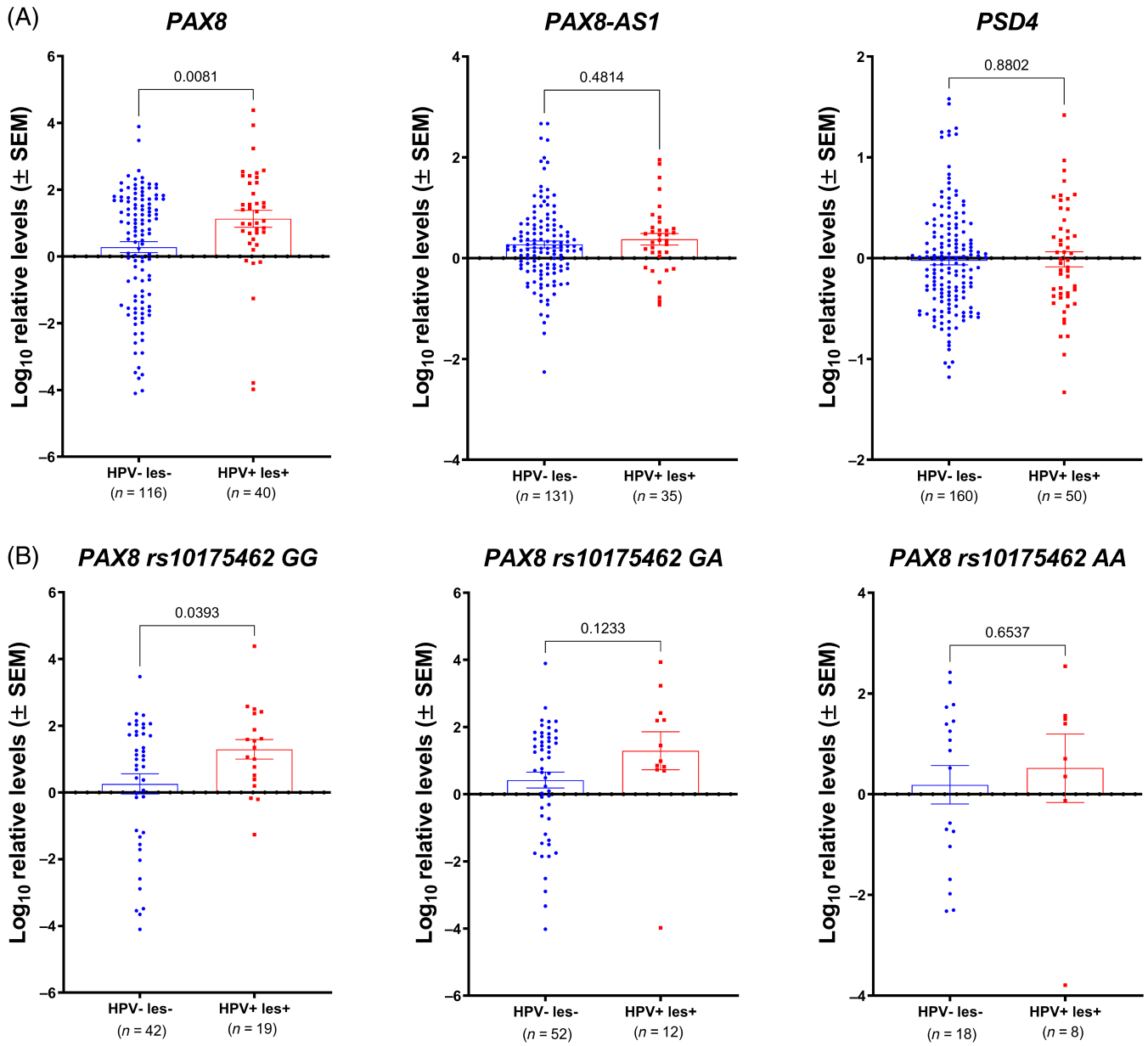

FIGURE 1 Association of *PAX8*, *PAX8-AS1* and *PSD4* RNA levels with HPV status. A, Transcript levels of *PAX8*, *PAX8-AS1* and *PSD4* were tested in comparing HPV negative cervical samples without lesion vs HPV-positive cervical samples with lesion. *P*-values were obtained from a *t* test between groups. B, Transcript levels of *PAX8* were tested after stratification into the three rs10175462 genotype groups. For each genotype, HPV negative cervical samples without lesion were compared to HPV-positive samples with lesion. *P*-values were obtained from a *t* test between groups

4 | DISCUSSION

Cervical cancer is known to have an increased familial relative risk[3-5] but only few hereditary risk factors have been uncovered so far. The HLA region of chromosome 6 has been replicated as a genetic risk factor in independent studies.[7-13] Of the two novel Genome Wide Association Study (GWAS) signals proposed by Rashkin et al,[7] the one represented by rs2856437 is located within the HLA gene cluster, in between two previously tested signals, rs9272117 (*HLA-DQA1*) and rs2844511 (*MICA/HCP5*).[13] However, it is not in linkage disequilibrium with either of them and remained significantly associated with cervical disease in a multivariate analysis. It is also embedded in an intron of a known oncogene, *PBX2*, and is located 5 kbp downstream of another proto-oncogene, *NOTCH4*. Variant rs2856437 and a closely correlated variant, rs2022059, are located within a promoter region (ENSR00001703115 in Ensembl) that is classified as a CTCF binding site. It is thus possible that they act on an insulator that regulates 3D chromatin structure. Our study did not clarify which gene(s) might be regulated by this locus but we clearly confirmed the significant association with cervical cancer, with the rare allele increasing the

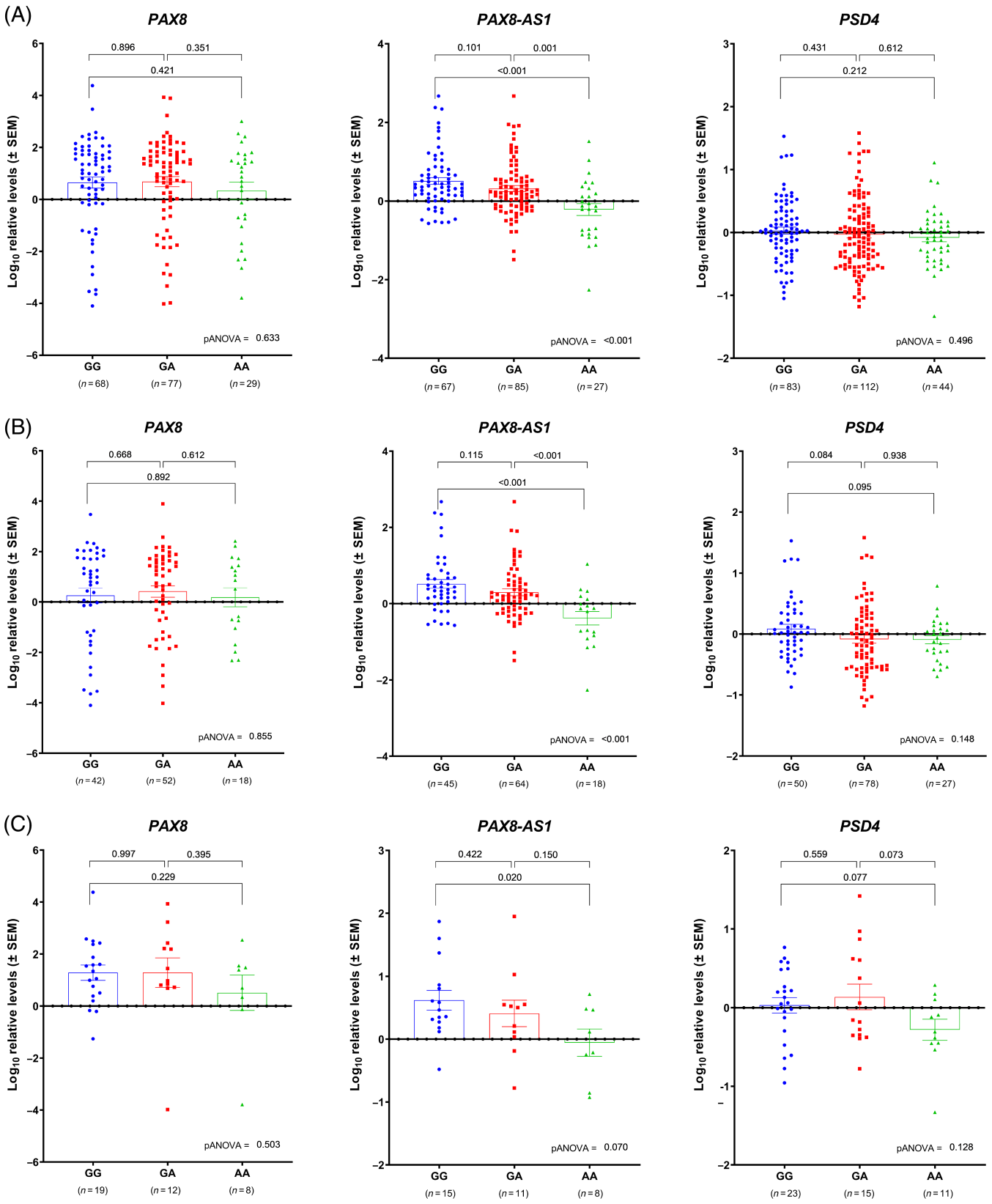

FIGURE 2 RNA levels of *PAX8*, *PAX8-AS1* and *PSD4* by rs10175462 genotype. A, Transcript levels of *PAX8*, *PAX8-AS1* and *PSD4* were associated with the genotypes of rs10175462 under an allelic dosage model, where each allele contributes to the transcript level in all samples. B, Transcript levels of *PAX8*, *PAX8-AS1* and *PSD4* were associated with the genotypes of rs10175462 under an allelic dosage model, like in A. But in this analysis the cases were restricted to HPV negative samples with no lesion. C, Transcript levels of *PAX8*, *PAX8-AS1* and *PSD4* were associated with the genotypes of rs10175462 under an allelic dosage model, like in A. But in this analysis the cases were restricted to HPV positive samples with documented lesions. *P*-values were obtained from a *t*-test between groups or from an ANOVA comparison across the three groups as indicated

risk by about two-thirds. This is the third independent variant in the risk region on chromosome 6p21.32 that has been replicated as a cervical cancer susceptibility variant in our German case-control series.

Importantly, the genome-wide meta-analyses of the population-based UK Biobank and the Kaiser Permanente GERA cohorts had also proposed a novel risk locus on chromosome 2q14.1 with the strongest risk variant rs10175462 being intronic to *PAX8*. Our study confirmed the association of rs10175462 with cervical malignancy, with the rare allele decreasing the risk by about 20%. This independent replication should be helpful towards establishing *PAX8* as a genetic susceptibility locus for cervical cancer. There is also evidence for rs10175462 being functionally relevant. In our transcript analyses in cervical tissue smears from women with different cytology and HPV status, the protective allele was associated with lower levels of lncRNA *PAX8-AS1*, suggesting a possible role of rs10175462 in regulating *PAX8-AS1* function throughout distinct stages of dysplasia. It is presently unknown which of the known *PAX8-AS1* isoforms are relevant for cancer and whether *PAX8-AS1* exerts an oncogenic role independent of *PAX8*. Interestingly, a haplotype in *PAX8-AS1* was reported which appeared associated with cervical cancer in a Han-Chinese population.[22] The two highly correlated markers used in that study, rs4848320 and rs1110839, are in modest linkage disequilibrium with the rs10175462 variant investigated here ($r^2 = 0.33$ and 0.34, respectively). It is likely that chromosome 2q14.1 variants constitute a risk locus for cervical cancer in non-European populations as well.

PAX8 itself is a tightly regulated transcription factor that shows increased immunostaining in several cancer tissues including cervical carcinomas.[23] It has been demonstrated that PAX8 transcriptionally regulates the E2F1 promoter directly, and E2F1 transcription is enhanced after RB1 depletion, thereby leading to persistent cell growth.[24] PAX8-positive cells appear to maintain the homeostasis of epithelium but, when acquiring *TP53* or *RB1* mutation, may generate precursors of epithelial lesions such as endometrial cancer.[17] This may similarly occur in HPV-infected cervical cells where we observed *PAX8* upregulation while TP53 and RB1 functions are blocked by viral proteins. It would seem in line with the protective effect of the minor allele of rs10175462 that its presence restrains HPV-associated *PAX8* upregulation. PAX8 also has been implicated as a master regulator of ovarian cancer progression[18] and may modulate the sensitivity of tumour cells against histone deacetylase antagonists.[25]

Both *PAX8* and *PAX8-AS1* are therefore excellent candidates for further studies of cervical cancer progression which may take the 2q14.1 genotype into account. However, rs10175462 is in strong linkage disequilibrium with 18 other highly correlated variants ($r^2 > 0.8$) within 15 kbp in this region. Thus, although the present study with is its relatively large sample size was able to confirm the association of this locus with cervical cancer, the disease-causing variant(s) and target gene(s) remain to be better defined. Additional work at the functional and epidemiological level (such as fine-mapping studies in further populations, luciferase assays, genome editing experiments, etc.) will be required to unambiguously identify the causal variant(s) in this set of candidates.

The successful replication of GWAS results confirms the value of a genome-wide approach in the study of virus-induced cancers and encourages additional GWA studies to determine the risk of progression in females with cervical dysplasia and/or HPV infection more precisely, and to eventually uncover potential targets. Our present data further corroborate chromosome 6p21.32-33 as consistent risk region and strongly support the chromosome 2q14.1 region (*PAX8*) as the first non-HLA genetic susceptibility locus for cervical cancer.

ACKNOWLEDGMENTS

We cordially thank all the patients, volunteers, and clinicians who participated in sampling and documentation. We thank the German Colposcopy Study Group under the lead of Karl-Ulrich Petry for their support in patient recruitment, Kristine Bousset and Eda Celik for support in database maintenance and helpful discussion, and Juliane Hoyer, Julia Enßen and Arif Ekici for support in DNA extraction. Our study was supported by the Bruno and Helene Jöster Foundation. Qianqian Mao was supported by a research stipend from the Shanghai New Medical Union Hospital, and Yingying Wang was supported by a research stipend from the Xuzhou Medical University.

CONFLICT OF INTEREST

Peter A. Fasching received grants from BioNTech and Cepheid, and personal fees from Novartis, Roche, Pfizer, Daiichi-Sankyo, Eisai, AstraZeneca, Merck Sharp & Dohme, Pierre Fabre, SeaGen, Hexal, Agenida and Lilly during the conduct of our study. This funding had no direct connection to the study presented here and did not influence our study design and results. The other authors have no disclosures and declare there are no conflicts of interest in the writing, preparing, or finalizing of our study.

ETHICS STATEMENT

Patient and control samples were obtained after informed consent and the study was approved by the Ethics committee of Hannover Medical School (Vote No. 441).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Peter A. Fasching https://orcid.org/0000-0003-4885-8471
Matthias Dürst https://orcid.org/0000-0003-1235-1150
Thilo Dörk https://orcid.org/0000-0002-9458-0282

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68:394-424.Updated online at: https://gco.iarc.fr/. Accessed June 2019.
2. Bruni L, Albero G, Serrano B, et al. Human papillomavirus and related diseases report—world. In: *ICO/IARC Inf Cent HPV Cancer (HPV Inf Centre) [Internet]* 2019;307. https://hpvcentre.net/statistics/reports/XWX.pdf.

3. Ahlbom A, Lichtenstein P, Malmström H, Feychting M, Hemminki K, Pedersen NL. Cancer in twins: genetic and nongenetic familial risk factors. *J Natl Cancer Inst*. 1997;89:287-293.
4. Hemminki K, Vaittinen P. Familial risks in in situ cancers from the family-cancer database. *Cancer Epidemiol Biomarkers Prev*. 1998;7:865-868.
5. Magnusson PKE, Lichtenstein P, Gyllensten UB. Heritability of cervical tumours. *Int J Cancer*. 2000;88:698-701.
6. Chen D, Cui T, Ek WE, Liu H, Wang H, Gyllensten U. Analysis of the genetic architecture of susceptibility to cervical cancer indicates that common SNPs explain a large proportion of the heritability. *Carcinogenesis*. 2015;36(9):992-998. https://doi.org/10.1093/carcin/bgv083.
7. Rashkin SR, Graff RE, Kachuri L, et al. Pan-cancer study detects genetic risk variants and shared genetic basis in two large cohorts. *Nat Commun*. 2020;11(1):4423. https://doi.org/10.1038/s41467-020-18246-6.
8. Leo PJ, Madeleine MM, Wang S, et al. Defining the genetic susceptibility to cervical neoplasia—A genome-wide association study. *PLoS Genet*. 2017;13:1-20.
9. Chen D, Gyllensten U. Lessons and implications from association studies and post-GWAS analyses of cervical cancer. *Trends Genet*. 2015;31:41-54.
10. Chen D, Enroth S, Liu H, et al. Pooled analysis of genome-wide association studies of cervical intraepithelial neoplasia 3 (CIN3) identifies a new susceptibility locus. *Oncotarget*. 2016;7:42216-42224.
11. Bao X, Hanson AL, Madeleine MM, et al. HLA and KIR associations of cervical neoplasia. *J Infect Dis*. 2018;218:2006-2015.
12. Safaeian M, Johnson LG, Yu K, et al. Human leukocyte antigen class I and II alleles and cervical adenocarcinoma. *Front Oncol*. 2014;4:1-8.
13. Ramachandran D, Schürmann P, Mao Q, et al. Association of genomic variants at the human leukocyte antigen locus with cervical cancer risk, HPV status and gene expression levels. *Int J Cancer*. 2020;147(9):2458-2468.
14. Tian C, Hromatka BS, Kiefer AK, et al. Genome-wide association and HLA region fine-mapping studies identify susceptibility loci for multiple common infections. *Nat Commun*. 2017;8(1):599. https://doi.org/10.1038/s41467-017-00257-5.
15. Kachuri L, Francis SS, Morrison M, et al. The landscape of host genetic factors involved in infection to common viruses and SARS-CoV-2. *Genome Med*. 2020;12(1):93. https://doi.org/10.1186/s13073-020-00790-x.
16. Maulbecker CC, Gruss P. The oncogenic potential of Pax genes. *EMBO J*. 1993;12(6):2361-2367.
17. Fu DJ, De Micheli AJ, Bidarimath M, et al. Cells expressing PAX8 are the main source of homeostatic regeneration of adult endometrial epithelium and give rise to serous endometrial carcinoma. *Dis Model Mech*. 2020;13(10):dmm.047035. https://doi.org/10.1242/dmm.047035.
18. Corona RI, Seo JH, Lin X, et al. Non-coding somatic mutations converge on the PAX8 pathway in ovarian cancer. *Nat Commun*. 2020;11(1):2020. https://doi.org/10.1038/s41467-020-15951-0.
19. Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol*. 2008;8:R19.
20. Vandesompele J, De Preter K, Pattyn F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*. 2002;3:research0034.1.
21. Arnold M, Raffler J, Pfeufer A, Suhre K, Kastenmüller G. SNiPA: an interactive, genetic variant-centered annotation browser. *Bioinformatics*. 2014;31(8):1334–1336. http://www.snipa.org. Accessed January 13, 2021.
22. Han J, Zhou W, Jia M, et al. Expression quantitative trait loci in long non-coding RNA PAX8-AS1 are associated with decreased risk of cervical cancer. *Mol Genet Genomics*. 2016;291:1743-1748.
23. Laury AR, Perets R, Piao H, et al. A comprehensive analysis of PAX8 expression in human epithelial tumors. *Am J Surg Pathol*. 2011;35(6):816-826. https://doi.org/10.1097/PAS.0b013e318216c112.
24. Li CG, Nyman JE, Braithwaite AW, Eccles MR. PAX8 promotes tumor cell growth by transcriptionally regulating E2F1 and stabilizing RB protein. *Oncogene*. 2011;30(48):4824-4834. https://doi.org/10.1038/onc.2011.190.
25. Shi K, Yin X, Cai MC, et al. PAX8 regulon in human ovarian cancer links lineage dependency with epigenetic vulnerability to HDAC inhibitors. *Elife*. 2019;8:e44306. https://doi.org/10.7554/eLife.44306.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Ramachandran D, Wang Y, Schürmann P, et al. Association of genomic variants at *PAX8* and *PBX2* with cervical cancer risk. *Int. J. Cancer*. 2021;149:893–900. https://doi.org/10.1002/ijc.33614