PAX8 is transcribed aberrantly in cervical tumors and derived cell lines due to complex gene rearrangements

EDUARDO LÓPEZ-URRUTIA[1,2], ABRAHAM PEDROZA-TORRES[2], JORGE FERNÁNDEZ-RETANA[2],
DAVID CANTU DE LEON[3], FERMÍN MORALES-GONZÁLEZ[4], NADIA JACOBO-HERRERA[5],
OSCAR PERALTA-ZARAGOZA[6], JORGE GARCÍA-MENDEZ[7], VERÓNICA GARCÍA-CASTILLO[1],
OSVALDO BAUTISTA-ISIDRO[1] and CARLOS PÉREZ-PLASENCIA[1,2]

[1]Laboratorio de Genómica Funcional, Unidad de Biomedicina, FES-Iztacala, UNAM, Tlalnepantla;
[2]Laboratorio de Genómica, Instituto Nacional de Cancerología; [3]Unidad de Investigaciones Biomédicas en Cáncer,
Instituto Nacional de Cancerología, Instituto de Investigaciones Biomédicas, UNAM, Tlalpan; [4]Subdirección Médica,
Instituto Jalisciense de Cancerología, Guadalajara, Jalisco; [5]Unidad de Bioquímica, Instituto de Ciencias Médicas y Nutrición,
Tlalpan; [6]Centro de Investigaciones en Enfermedades Infecciosas, Instituto Nacional de Salud Pública, Cuernavaca;
[7]Departamento de Posgrado y Educación Médica Continua, Instituto Nacional de Cancerología, Tlalpan, México

Abstract. The transcription factor PAX8, a member of the paired box-containing gene family with an important role in embryogenesis of the kidney, thyroid gland and nervous system, has been described as a biomarker in tumors of the thyroid, parathyroid, kidney and thymus. The PAX8 gene gives rise to four isoforms, through alternative mRNA splicing, but the splicing pattern in tumors is not yet established. Cervical cancer has a positive expression of PAX8; however, there is no available data determining which PAX8 isoform or isoforms are present in cervical cancer tissues as well as in cervical carcinoma-derived cell lines. Instead of a differential pattern of splicing isoforms, we found numerous previously unreported PAX8 aberrant transcripts ranging from 378 to 542 bases and present in both cervical carcinoma-derived cell lines and tumor samples. This is the first report of PAX8 aberrant transcript production in cervical cancer. Reported PAX8 isoforms possess differential transactivation properties; therefore, besides being a helpful marker for detection of cancer, PAX8 isoforms can plausibly exert differential regulation properties during carcinogenesis.

Introduction

The paired-box (PAX) family, initially described in *Drosophila* but highly conserved among vertebrates, encompasses nine

Correspondence to: Professor Carlos Perez-Plasencia, Cancer Genomics Laboratory, UB IMED, FES-Iztacala, UNAM, Los Reyes Iztacala, Av. San Fernando 2, Col. Sección XVI México, Tlalnepantla, DF 14080, México
E-mail: carlos.pplas@gmail.com

DNA-binding proteins that function as transcription factors. PAX proteins share a common structure: a DNA-binding paired domain (PAired boX, hence the name of the family), a PST homeodomain and an octapeptide; genes from this family differ in the length of the first two and the presence or absence of the last domain. During development, PAX genes are expressed in different regions of the forming embryo, thus they constitute important gene expression regulators (1,2). PAX8 has a fairly well understood role in development: it is associated to morphogenesis of the kidney, thyroid gland and nervous system (3,4), and its expression is regulated by alternative splicing (5).

The PAX8 gene, located at 2q13 and spanning 63 kb, is transcribed, processed and translated to give rise to six reported mRNAs: the originally described PAX8A, B, C, and D, which are in turn translated to the corresponding protein isoforms (6), and a more recently described transcript, PAX8F, that shares the PAX8A ORF and bears an extended 5'UTR (7). The existence of PAX8B has been questioned and its GenBank entry (NM_013951.3) is currently suppressed. All PAX8 isoforms bind DNA with similar affinity; however, PAX8A and PAX8B display higher transactivation potential (6).

Recent studies suggest a role for PAX8 in carcinogenesis that makes it a biomarker candidate; still, more studies are necessary to completely understand this role. So far, PAX8 has been associated to important cancer processes such as Retinoblastoma/E2F1 transcription (8,9) and telomerase activity (10). Another study recently showed that by inhibiting PAX8 expression with shRNAs in ovarian cancer cell lines, their viability diminished (11); this suggests a direct involvement of PAX8 in cell proliferation which, in turn, could represent a potential therapeutic target.

An important fact that links PAX8 to carcinogenesis is the PAX8/PPARγ genetic rearrangement found in 30-35% of follicular thyroid carcinoma and up to 10% of follicular thyroid adenoma (12-14). Hence, t(2;3)(q13;p25) fuses the promoter

and 5-coding portion of the transcription factor PAX8 that involves DNA binding domains to the full-length coding sequence of the nuclear receptor peroxisome proliferator-activated receptor-γ 1 gene (15).

As a biomarker, PAX8 is frequently detected in epithelial tumors of the thyroid, parathyroid, kidney, thymus and female genital tract (16,17). In thyroid tumors, PAX8 overexpression has been found in follicular and papillary adenomas and carcinomas (16,18,19). Wilms' tumors and nephrogenic adenomas, among other renal tumors, show PAX8 overexpression owing to its role during renal organogenesis (20-22). PAX8-expressing oviduct cells have been pointed out as the probable origin of ovarian and endometrial cancers, rendering PAX8 as an early gynecological cancer marker, associated with poor disease outcome (23,24). Only few groups have reported PAX8 overexpression in cervical cancer employing distinct methodologies (25-27).

A genome-wide search for PAX8 binding sites using ChIP-Seq technology was recently carried out on thyroid cells. The authors identified binding sites that show that PAX8 regulates genes involved in cell proliferation and differentiation, as well as binding sites located in non-promoter regions, including introns, that hint to a role as a distal transcriptional regulator (28). However, none of the aforementioned reports of PAX8 as a tumor marker or functional PAX8 studies specify whether its results concern any particular PAX8 isoform; therefore, we set out to investigate which isoform or isoforms were expressed in a set of locally advanced cervical tumor samples compared to normal cervical tissues. PAX8 isoforms could possibly activate different genes contributing to carcinogenesis, given their different *in vitro* transactivation properties (6).

We had hypothesized that, since splicing is highly altered in cancer (29), we would find a particular PAX8 splicing pattern in cervical carcinoma tumor samples and cervical carcinoma-derived cell lines, but RACE experiments revealed the expression of previously unreported aberrant transcripts instead. We also found evidence hinting of a complex gene rearrangement in the genomic DNA from the cell lines, providing a possible explanation for these transcripts. To our knowledge, this is the first study showing that PAX8 gene undergoes complex rearrangement processes that could contribute to establish the CC tumor phenotype.

Materials and methods

Expression data

Tissue samples. Patients were prospectively enrolled into the National Cancer Institute of Mexico (INCAN) tumor-banking protocol at the time of diagnosis. All patients included accepted and signed informed consent; institutional ethics and scientific board committees approved the protocol. Immediately after punch biopsy, tumor samples were split into three pieces, one for the pathologic confirmation of at least 80% of tumor cells, mandatory for this type of molecular profiles, and the remaining two for RNA and DNA isolation. RNA and DNA biopsies were frozen in liquid nitrogen until nucleic acid extraction. Eligibility criteria were i) patients with a confirmed pathologic diagnosis of CC staged IB2 up to IIIB (LACC); ii) biopsies with pathology report with >80% of tumors cells; hence, the genomic

analysis is mainly addressed for tumor cells; iii) age greater to 20 and less than 60 years; iv) high-quality DNA and RNA; v) no presence of comorbidities; vi) and without previous oncological treatment.

Healthy cervical tissues were obtained from patients who had undergone hysterectomy by uterine myomatosis. Inclusion criteria were: i) no previous cervical surgery (such as the loop electrosurgical excision procedure or cone biopsy), ii) no HPV infection, iii) no hormonal treatment, and iv) at last three previous negative Pap smears.

PAX8 mRNA expression. We designed two sets of primers, one for detection of PAX8 exons 3-4 and another for exons 10-11 in the mRNA (NM_003466.3; see Fig. 1). RNA was extracted from 18 tumoral and 20 healthy, HPV-free, cervical tissue samples using the TRIzol reagent (Life Technologies cat. # 15596-026), following the manufacturer's recommendations. Its integrity was verified through agarose gel electrophoresis and it was quantitated by spectrophotometry (OD260/280>1.9). For first strand synthesis, we mixed 500 ng RNA, 0.5 mM dNTPs, 10 mM DT, 1X First-strand buffer and 180 ng random hexamers, heated them to 42°C for 2 min; then we added 200 units of SuperScript II reverse transcriptase (Invitrogen cat. #18064-022) and incubated the reaction for 50 min at 42°C, followed by 15 min at 70°C. The resulting cDNAs was probed for DNA contamination by performing no-RT assays. Quantitative PCR (qPCR) reactions were performed in triplicate using LightCycler® 480 SYBR Green I Master Mix (Roche, cat. # 04 707 516 0081) in the LightCycler 480 instrument, following the manufacturer's recommendations (40 amplification cycles, Tm 58°C). End-point PCR was performed using the PCR Master mix 2X (Thermo Scientific, cat. #K0171) 0.5 μM de cada primer (35 cycles, Tm 56°C).

Western blotting. Total proteins were extracted from 80-85% confluent cell cultures. Culture media was removed and cells were rinsed twice with PBS, scraped off from culture dishes, and lysed using the RIPA lysis reagent (Santa Cruz Biotechnology) following the manufacturer's recommendations.

Total protein (50 μg) was mixed with Laemmli sample buffer, boiled, separated in 12% or 15% SDS-PAGE and transferred onto a Hybond-P PVDF membrane (Amersham-GE Healthcare). Membranes were probed overnight using a 1:500 (v/v) dilution of the appropriate antibody; for detection, 1:2500 (v/v) dilutions of HRP anti-rabbit or anti-mouse conjugate antibodies (Santa Cruz Biotechnology) were used. Finally, using the SuperSignal WestFemto chemiluminescent substrate (Thermo Scientific), the membranes were scanned in the C-Digit blot scanner (Li-Cor) and the images were analyzed in the associated ImageStudio software (LiCor). Membranes were stripped and re-probed for actin detection as a loading control. The commercial antibodies used were anti-PAX8 (Cell Signaling Technology #9857) and anti-β actin (SantaCruz Biotechnology sc-1616). A representative image from three independent experiments is shown.

RT-PCR, 5' RACE and Exon PCR. DNA and RNA were isolated from tumor samples or from the CaSki (ATCC, CRL1550) and SiHa (ATCC, HTB35) cells grown to approximately 80-85%

Table I. Primers used for 5' RACE and Exon PCR experiments.

Name	5'-3' Sequence	Nt position
Primers for 5' RACE		
1 5' RACE adapter	GCU GAU GGC GAU GAA UGA ACA CUG CGU UUG CUG GCU UUG AUG AAA	NA
2 RACE Outer primer	GCT GAT GGC GAT GAA TGA ACA CTG	NA
3 RACE Inner primer	CGC GGA TCC GAA CAC TGC GTT TGC TGG CTT TGA TG	NA
4 PAX8UTR3'RV	AGT CCT CCT GTT GCT CAG TCG CT	1561-1539[a]
5 PAX8 RV	CTA CAG ATG GTC AAA GGC CGT	1519-1499[a]
Primers for RT-PCR		
1 PAX8 FW	ATG CCT CAC AAC TCC ATC AGA	167-187[a]
2 PAX8 RV	CTA CAG ATG GTC AAA GGC CGT	1519-1499[a]
3 GAPDH FW	CCT CAA GAT CAT CAG CAA TGC CT	617-639[b]
4 GAPDH RV	TCA CGC CAC AGT TTC CCG GAG	781-761[b]
Primers for Exon PCR		
6 PAX8EX1Fw	GAT GCA GGC ATC GAA TCT C	399-417[c]
7 PAX8EX1Rv	ACG CTC TCG AGA TCC AAC C	639-621[c]
8 PAX8EX2Fw	ATC CCC ACC CAA ACT CCT AC	64185-64204[c]
9 PAX8EX2Rv	TCA GCT GGA GAA GTC AAG CC	64468-64449[c]
10 PAX8EX3Fw	TGT CTA AAG ACC CCA CCT GC	66487-66506[c]
11 PAX8EX3Rv	AGC CAG GCC TTT CTT GTC TC	66824-66805[c]
12 PAX8EX4Fw	GCC ATG AGT TCT CTT CCT CC	68332-68352[c]
13 PAX8EX4Rv	GGT ATG CTG AAG GGG AGG TG	68538-68519[c]
14 PAX8EX5Fw	ACT ACC CCA GAG TCA CCC AG	68961-68980[c]
15 PAX8EX5Rv	AAA GCC TCA GCA AAC TGC TC	69215-69196[c]
15 PAX8EX6Fw	TTT GGC CTA GAG CAT GAA TAG	69376-69396[c]
16 PAX8EX6Rv	GAG CAC AGG CTC ATT GGG AG	69692-69673[c]
17 PAX8EX7Fw	CCT AAG ACA CAG GCT CAG GG	74363-74382[c]
18 PAX8EX7Rv	AGC CAA GCT CTT CAG TCC C	74620-74602[c]
19 PAX8EX8Fw	CTT GTG CGT GTT CCC TCC	75738-77555[c]
20 PAX8EX8Rv	GTC TGC CCT GAG GAC CC	76060-76044[c]
21 PAX8EX9Fw	AGC TCA GGA GAG TGA GAT G	84030-84040[c]
22 PAX8EX9Rv	GTC CCA CCT TGC TCC AAT AC	84269-84250[c]
23 PAX8EX10Fw	CCT GCA TTG ATG CCC TTC	91115-91132[c]
24 PAX8EX10Rv	AGG TAA CCT TTG ACC CAC CC	91335-91316[c]

NA, not applicable. Nucleotide positions: relative to BC001060, [a]relative to NM_003466, [b]relative to NM_002046.5, [c]relative to NG_012384.

confluence, using the TRIzol reagent (Life technologies) following the manufacturer's recommendations.

First strand cDNA synthesis was carried out for 90 min at 42°C. Total RNA (5 µg) was mixed with 1 mM oligo dT, 0.25 mM of each dNTP, 1X First-strand buffer, 10 mM DTT and 200 units of M-MLV reverse transcriptase (Promega). This reaction (1 µl) was used as a template for 50 µl PCR reactions for amplification of PAX8 transcripts using the PAX8Fw and PAX8Rv primers (Table I) 0.2 µM each, 25 µM of each dNTP, 1X Herculase II buffer and 1 unit of Herculase II fusion DNA polymerase (Agilent).

Amplification conditions were 5 min pre-incubation at 95°C followed by 35 cycles of 20 sec at 95°C, 20 sec at 53°C and 70 sec at 68°C, followed by a final step of 4 min at 68°C. GAPDH transcripts were detected as a loading control, using the corresponding primers (Table I) 0.25 µM each, 0.2 mM of each dNTP, 2.5 mM MgCl2 and 0.75 units of GoTaq DNA Polymerase (Promega) in a 25 µl volume. The amplification conditions were: 5 min pre-incubation at 95°C, 35 cycles of 30 sec at 95°C, 20 sec at 57°C and 20 sec at 72°C, followed by a final elongation step of 7 min at 72°C.

Rapid amplification of cDNA ends (RACE) was performed using the First Choice RLM-RACE kit (Life Technologies-Ambion) according to the manufacturer's protocol. Briefly, 1 µg of RNA from HeLa cells, SiHa cells or tumor samples
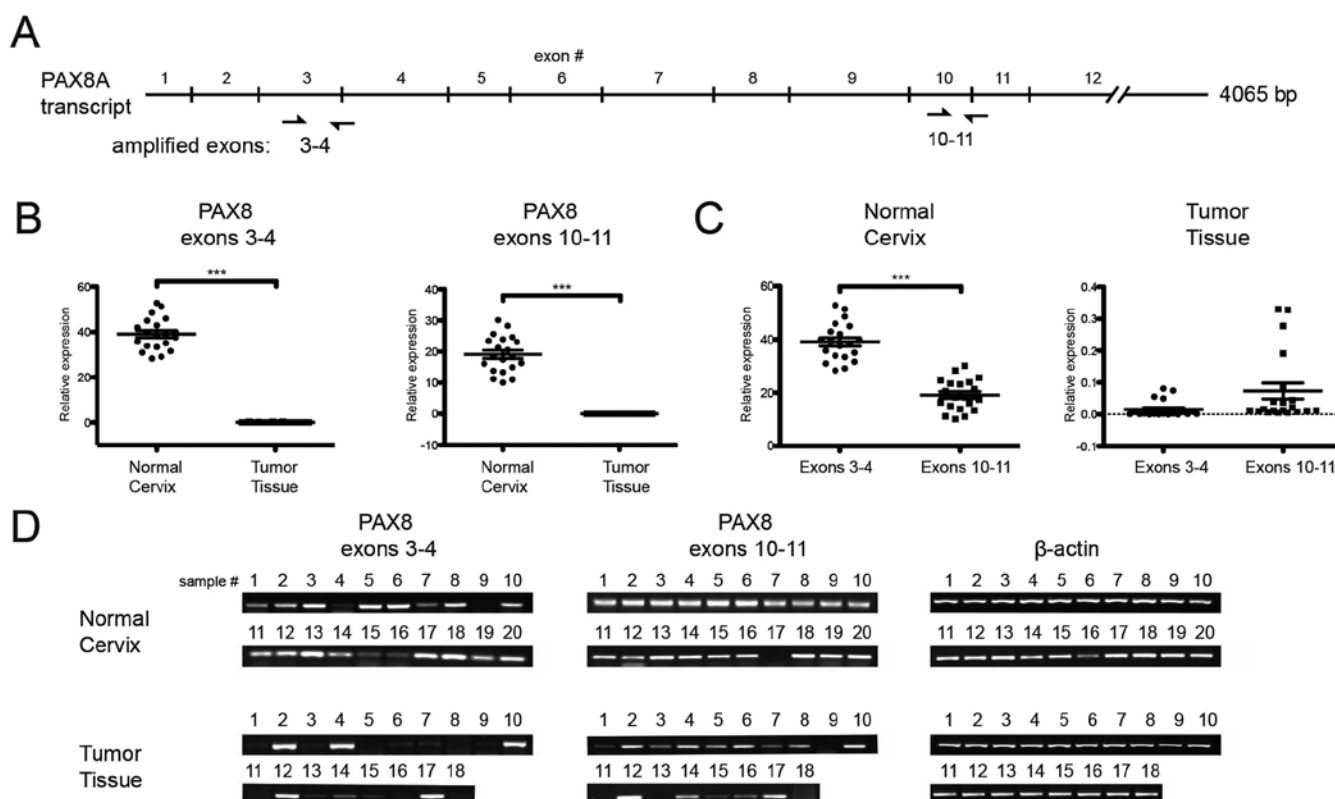
Figure 1. Detection of PAX8 in normal cervix and tissue samples. (A) Diagram showing the positions of the primer pairs used for detection of PAX8 exon amplification, along the full-length mRNA. Detailed nucleotide positions are found in Table I. (B) Relative expression of PAX8 exons 3-4 and 10-11 in normal cervix and tissue sample. Note the differential ratios in both regions. (C) Expression balance between PAX8 exons 3-4 and 10-11 in normal cervix and tissue sample. (D) PCR products showing higher detection of exons 10-11 in tumor samples, compared to normal cervix.

was dephosphorylated with Calf intestine Acid Phosphatase (CIP) to remove the available phosphates from rRNA, tRNA, uncapped and partial transcripts; the cap structure was subsequently removed from mature mRNAs with Tobacco Acid Phosphatase (TAP) so that only these transcripts could acquire the 5' RACE adapter (Table I) trough ligation with T4 RNA ligase. Next, a random-primed RT reaction was performed using M-MLV reverse transcriptase (Promega). The PAX8 transcripts were amplified from the cDNA pool through nested PCR, the first round was performed using the 5' RACE Outer Primer from the kit in combination with the PAX8UTR3'Rv primer, while the 5' RACE Inner Primer and the PAX8Rv primer were used for the second round (Table I). The resulting amplicons were resolved in ethidium bromide-stained 2% agarose gels.

PAX8 exons were amplified from genomic DNA with ad-hoc primers based on the PAX8 complete sequence (GenBank Gene ID: 7849; locus NG_012384). These primers were designed using the ExonPrimer server (ihg. helmholtz-muenchen.de/ihg/ExonPrimer.html) and are shown in Table I. PCR was performed using 25 ng gDNA template, 0.25 μM of each primer, 0.2 mM of each dNTP, 2.5 mM MgCl$_2$ and 0.75 units of GoTaq DNA Polymerase (Promega) in a 25 μl volume. The amplification conditions were as follows: 5 min pre-incubation at 92˚C followed by 35 cycles of 20 sec at 92˚C, 20 sec at 67˚C and 30 sec at 72˚C, followed by a final step of 7 min at 72˚C. All the resulting amplicons were resolved in ethidium bromide-stained 1.5% or 2% agarose gels.

Cloning and sequence analysis. The RACE product bands were excised from the agarose gels and purified using the QIAquick Spin kit (Qiagen) according to the manufacturer's protocol. After purification, the RACE products were cloned into the pGEM T-Easy plasmid (Promega). Plasmids from positive colonies were analyzed by restriction; each clone was sequenced in both chains using universal primers and the Big Dye Terminator Ready Reaction kit (Perkin-Elmer) and analyzed in the ABI PRISM 3130xl Genetic Analyzer System (Applied Biosystems). Sequence information was analyzed using the CLC Bio Main Workbench (CLC Bio; Qiagen) as well as the BLAT (genome.ucsc.edu) and BLAST (blast.ncbi. nlm.nih.gov) algorithms.

In order to search for sequence similarity among clones, we performed a CLUSTAL alignment and checked it manually. Based on this alignment, we searched for the most adequate model of evolution using the FindModel server (http://www. hiv.lanl.gov/content/sequence/findmodel/findmodel.html), and the GTR model was selected to perform a maximum likelihood phylogenetic analysis employing a 1000-replicate bootstrap analysis. Finally, a tree was constructed using the neighbor-joining method.

Results

PAX8 mRNA expression. Since splicing events remove exons 8-10 (6), we designed the two sets of primers depicted in Fig. 1A to assess the PAX8 mRNA expression and integrity. We amplified independently exons 3-4 and exons 10-11 from
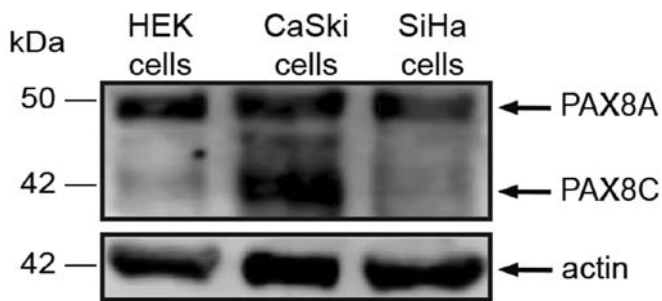
Figure 2. Immunodetection of PAX8 in samples from HEK (kidney tumor-derived), CaSki, and SiHa (cervical carcinoma-derived) cell lines. Actin was detected as a loading control.
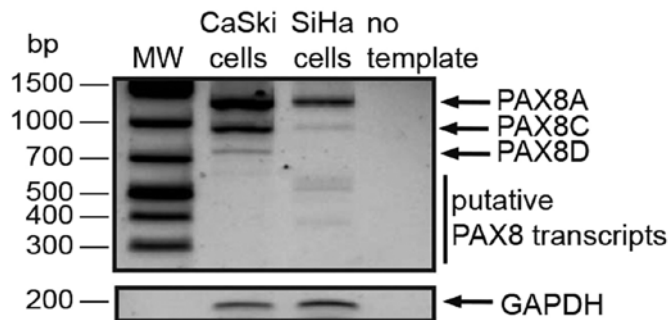

Figure 3. PAX8 ORF amplification. PAX8 alternative ORFs were amplified from CaSki or SiHa cells with similar results. The band intensities of the PAX8A/F ORFs remained constant among repetitions while the PAX8C and PAX8D band intensities varied slightly; a representative experiment is shown. GAPDH transcript was detected as a loading control.

Figure 4. PAX8 5' RACE results. Final nested PCR from 5' RACE experiments performed with RNA from (A) CaSki, (B) SiHa or (C) tumor samples.

20 normal cervix samples and 18 cervical tumor samples and found interesting results: Although overall messenger expression was lower in tumors (Fig. 1B), the exon 3-4:10-11 ratio (Fig. 1C) was different in tumors when compared to normal cervical samples; which was more evident upon agarose-gel separation of the amplicons (Fig. 1D). This observation suggested the presence of transcripts containing only the 3'-most end of the PAX8 mRNA that may or may not correspond to reported splicing products.

PAX8 protein expression. In order to detect the PAX8 isoforms possibly expressed in cervical carcinoma-derived cell lines, we selected an antibody raised against the carboxi-terminal region of PAX8, shared by its isoforms. Previously, PAX8 has been detected through immunocytochemistry but current cell imaging techniques are yet to discriminate between isoforms sharing structural motifs. Consistently with reported detection in HEK cells, we observed a 50 kDa band, apparently corresponding to PAX8A in CaSki and SiHa cells; a 43 kDa band was detected, corresponding to PAX8C (Fig. 2). The reported PAX8A and C transcripts bear exons 3-4 and 10-11, therefore, their presence was not likely to account for the imbalance between these mRNA regions in cervical tumors.

PAX8 aberrant transcripts. We next set out to identify the PAX8 transcripts possibly expressed in cervical carcinoma-derived cell lines at the mRNA level. The 1372-nt PAX8A/F
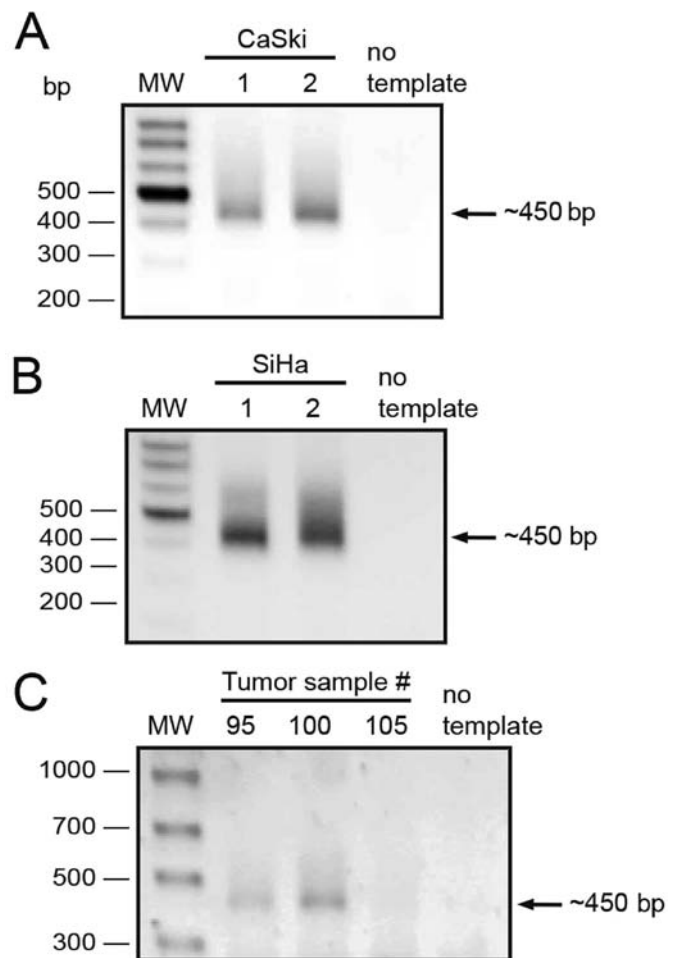
ORF was present in both CaSki and SiHa cells (Fig. 3), consistent with our protein detection results. The PAX8C and PAX8D transcripts were also detected in both cell lines. However, what really caught our interest were the 400-500 nt products that we found in both lines, more evident in the SiHa lane in Fig. 3.

Due to the fact that exons 10-11 were detected at higher levels in tumor samples, we reasoned there was a population of transcripts sharing the 3' end of the PAX8 ORF; so we set up a 5' RACE strategy using a reverse primer located 20 nt downstream from the end codon (PAX8UTR3'RV, Table I). We performed two independent 5' RACE experiments from CaSki and SiHa cell lines (Fig. 4A and B). A single band of about 450 nt was observed as product from all these experiments. Furthermore, to rule out the possibility that the detected transcripts were exclusive to cell lines, we performed 5' RACE experiments on three tumor samples and obtained similar ~450 bp products from two of them (Fig. 4C). All these 5' RACE products were excised from the gel and cloned into a T-protruding vector.

We expected to obtain a number of positive colonies carrying a 450 bp insert but instead, we surprisingly obtained colonies carrying differently sized inserts. Plasmids from these colonies were purified and sequenced. In total, we obtained 33 sequences that are shown in Fig. 5A, aligned
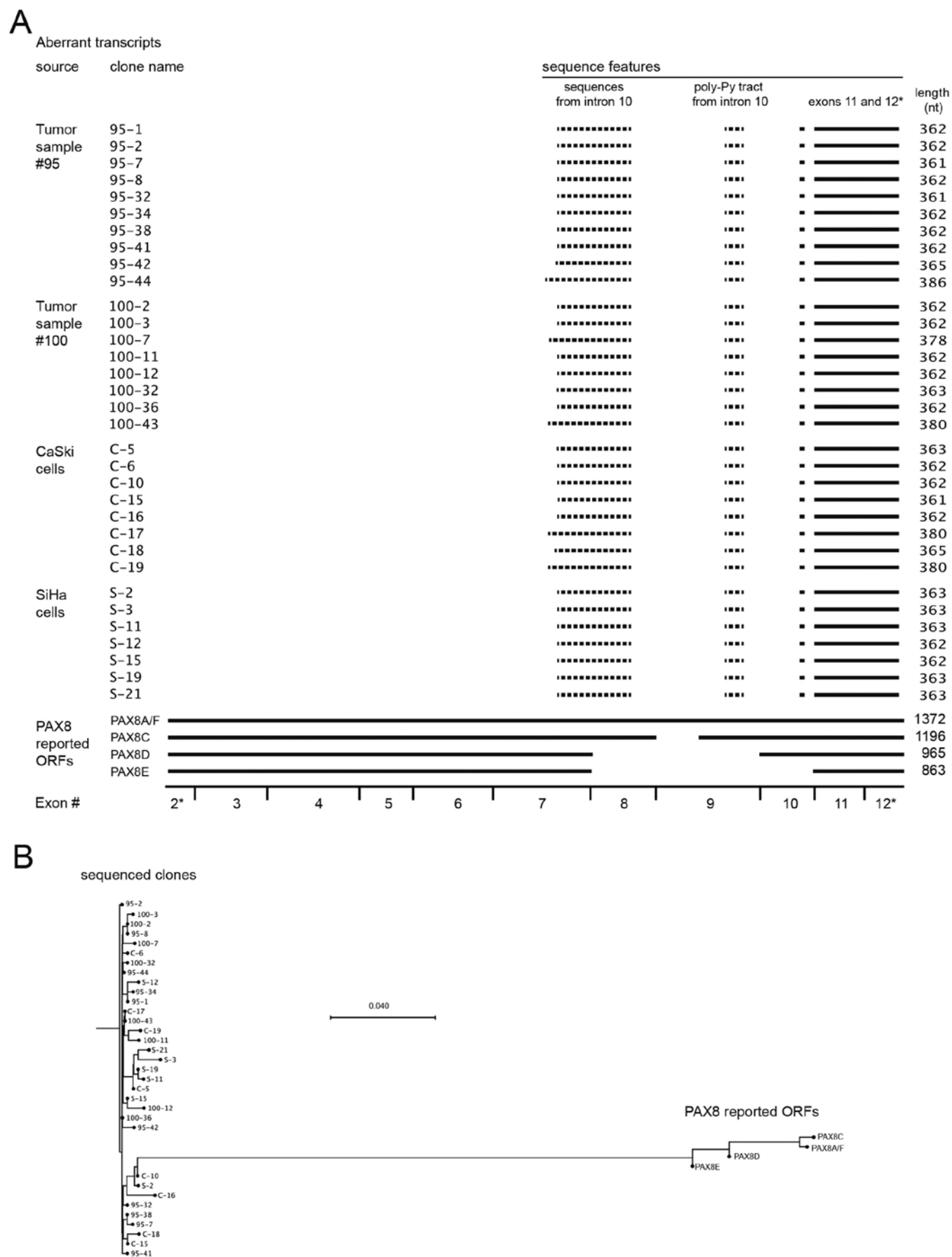
Figure 5. PAX8 aberrant transcripts. (A) Sequence alignment comparing 33 cloned PAX8 aberrant transcripts from CaSki (C-#) or SiHa (S-#) cells, or tumor samples (sample no.-#). The numbers correspond to the isolated clones; only those that resulted positive were analyzed. Solid lines represent identical sequences; dotted lines represent sequences that did not align to the PAX8 transcripts. The numbers below indicate the exon numbers; exons marked with * are partially translated. (B) Neighbor-joining tree based upon the sequence alignment. The scale represents the relative number of base changes.
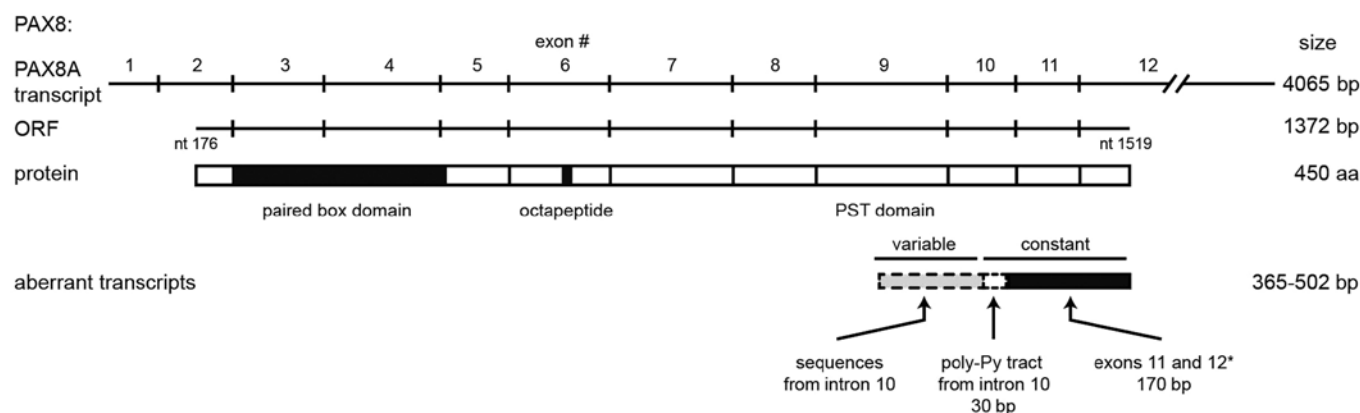
Figure 6. Comparison between PAX8 mRNA, protein and aberrant transcripts. Diagram showing the relative size and position of the PAX8 aberrant transcripts compared to the reported PAX8 mRNA, ORF and protein. Solid lines represent identical sequences; dotted lines represent sequences that did not align to the PAX8 transcripts.

Table II. Expected amplicon sizes from Exon PCR experiments.

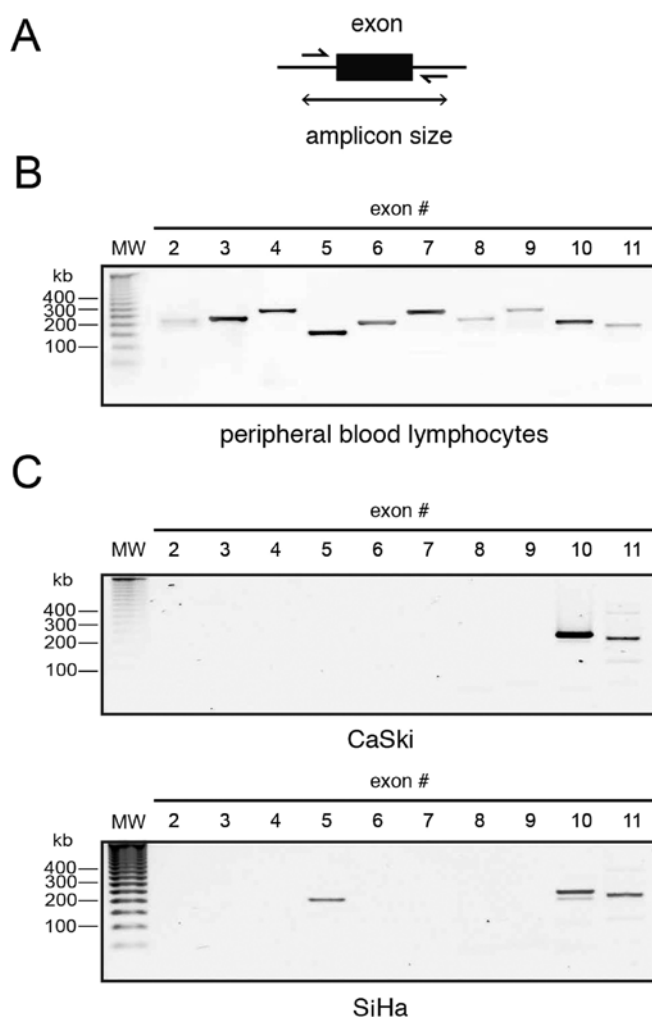| Exon no.	Exon PCR amplicon size (bp)
2	241
3	284
4	338
5	207
6	255
7	317
8	258
9	323
10	240
11	221

Figure 7. PAX8 exon amplification. (A) Representation of the rationale underlying the Exon PCR experiments. (B and C) Exon amplification results for peripheral blood lymphocytes and cervical carcinoma-derived cell lines, respectively.

with the four reported PAX8 transcripts. Solid lines represent identical sequences, whereas sequences represented by dotted lines did not align to the PAX8 transcripts. The sequences of all transcripts were submitted to GenBank under accessions KJ545852 to KJ545885. The clones obtained form tumor samples and cervical carcinoma-derived cell lines form a single pool of PAX8 aberrant transcripts. The reported transcripts and our sequenced clones clustered separately from each other, and no evident clustering was found within our clone group (Fig. 5B; Table II).

Upon analysis of the sequences from our collection of cloned transcripts we found that they only shared the 3'-most 170 nt with the PAX8 coding sequence. This region only represents two (11 and 12) of the ten exons that constitute the PAX8 ORF, encoded in exons 2-12 (see the representation in Fig. 6). Upstream from that region, the 33 clones share a 30-nt sequence from intron 10, which contains the polypyrimidine tract and is located 5 nt upstream of the py-AG splicing acceptor site, adjacent to exon 11. The remaining 5' region of the cloned transcripts was variable among them, comprising sequences that align to other regions of exon 10 with multiple single nucleotide mismatches. Performing a BLAST search with our clones produced variable results pointing at contigs

or chromosome assemblies (data not shown). Taken together, the results from both analyses suggested genomic instability in the PAX8 gene, because the aberrant transcripts that we were able to isolate contain sequences from different

non-contiguous genomic locations. Noteworthy, we did not find a definite sequence or pattern corresponding to either tumor samples or cell lines, reaffirming the idea that PAX8 produces similar aberrant transcripts in both cervical tumors and derived cell lines.

Some of the aberrant transcripts contain open reading frames, but none of them comprises its full length. If translated, these ORFs should lead to short (12-22 kDa) proteins sharing the carboxi-terminus and thus the PST transactivation domain from PAX8 (Fig. 6).

PAX8 exon amplification. Since all the detected transcripts conserved only the 3'-most introns, we wanted to know whether the missing introns were actually present in the genomic DNA of the cells. As an initial approach to assess the integrity of the PAX8 gene, we designed primers along the intronic regions flanking each exon (Fig. 7A) with the aid of the ExonPrimer software tool. We targeted only exons 2-11, since exon 1 is not translated in any reported PAX8 mRNA and exon 12, while protein coding, is more than 2.5 kb long. A validation experiment was performed using DNA isolated from peripheral blood lymphocytes from a healthy patient, it yielded the expected amplicon sizes, as seen in Fig. 7B.

The results using CaSki and SiHa cell DNA were very interesting: amplicons corresponding to exons 2-9 failed to amplify in both lines, save for exon 5 that was detected only in SiHa DNA (Fig. 7C). Notably, only exons 10 and 11, those identified in the aberrant transcripts, were detected through exon PCR. This suggests, together with the sequences present in the cloned transcripts, that the PAX8 gene may be altered in cervical carcinoma.

Discussion

In this study, we analyzed the expression levels of two regions of the PAX8 mRNA in a set of tumoral and healthy cervical tissue samples, and found an imbalance between their expression levels. This imbalance can be attributed to the presence of aberrant transcripts containing the 3'-most region of the mRNA. Previous work on female genital tract tumors has shown a consistent PAX8 expression in ovary tumors, such as non-mucinous carcinomas, including serous, endometroid, clear cell and transitional cell carcinomas (30,31). Only one investigation has reported PAX8 expression in carcinomas of uterine cervix, but the results are in the context of epithelial tumors (20).

We aimed to assess the possible differential splicing of the PAX8 transcripts in cervical carcinoma-derived cell lines and tumors. Upon performing these experiments, we found a completely different scenario, rather than the expected protein isoforms, we detected a single protein and several shorter putative isoforms. Consequently, we aimed to detect the corresponding shorter transcripts.

Notably, while we did detect a band of the expected molecular weight by western blotting, the full-length ORF was only amplified with a high sensitivity polymerase and was not detected through 5' RACE experiments (compare Figs. 3 and 4). This suggests a very low abundance of this transcripts and, thus, high translation efficiency of this mRNA. A high

translation efficiency has been already suggested for this transcript by Szczepanek-Parulska and co-workers (32), who amplified a 587/684 bp fragment to detect PAX8A/F and hypothesized that the extra 97 bases at the 5'UTR participate in translational regulation.

The 5' RACE experiments yielded a collection of transcripts ranging from 378 to 542 (Fig. 4), out of the roughly 1300 bases of the full-length PAX8 ORF. Shorter-than-expected transcripts from a 10-intron gene suggest alternative splicing, so it was puzzling when the sequencing results only revealed the two 3'-most coding exons from the ORF preceded by unknown and intronic sequences. The fact that similar transcripts were obtained from different independent sources strongly suggests that the aberrant transcripts that we found are product of a recurring process, present in both cervical carcinoma-derived cell lines and tumor samples.

The production of novel transcripts through alterations of alternative splicing is a known trait of cancer cells (29, reviewed in ref. 33) but only a few genes have been reported to produce aberrant transcripts containing sequences other than exons, such as FHIT and WWOX; this phenomenon has been attributed to their proximity to the chromosome fragile sites FRA3B and FRA16D, respectively (34,35). PAX8, in turn, is located at the 2q13 locus (36,37), close to the FRA2B fragile site (38,39), which prompted us to try and assess the integrity of the PAX8 gene.

Rather than trying to establish the genomic status, our exon amplification experiments show that it is a worthwhile venture. These results suggest that the genomic sequence of the PAX8 gene is at least partially disrupted, or that it harbors a considerable amount of mutations that impede amplification of exons 2 to 10. Noteworthy, these results seem to correspond to the sequences found in the aberrant transcripts, since only exons 10 and 11 were present in them, only these same exons were detectable through Exon PCR. The PAX8 gene has previously proven to be unstable in tumor cells, the (2;3)(q13;p25) translocation, detected in both follicular thyroid carcinomas and adenomas (40), leads to the formation of a chimeric PAX8-peroxisome proliferator-activated receptor (PPAR-γ) oncogene (15). Moreover, cervical carcinoma is prone to chromosome abnormalities as human papillomavirus has been demonstrated to produce specific chromosomal imbalances in transfected keratinocytes (41). In any case, we are conducting further research to establish the nature of the suggested genomic alterations in the PAX8 gene.

Taking the data together, we find it reasonable to speculate that the cDNA aberrations that we found are the result of genomic modifications characteristic of cervical carcinoma. Even so, it is notable that these probable genomic modifications are likely not present in every cell from our samples; otherwise we would not have been able to detect the full-length PAX8 in total protein extracts.

Here we report the presence of a number of aberrant transcripts produced by the PAX8 gene cervical cancer tissues. These transcripts are present in cervical carcinoma-derived cell lines and tumors and likely encode shorter isoforms. Further studies would shed light on the implications of these transcripts in carcinogenesis and their role as cause and/or consequence of other molecular phenomena.

Acknowledgements

References

1. Gruss P and Walther C: Pax in development. Cell 69: 719-722, 1992.
2. Goulding MD, Lumsden A and Gruss P: Signals from the notochord and floor plate regulate the region-specific expression of two Pax genes in the developing spinal cord. Development 117: 1001-1016, 1993.
3. Mansouri A, Hallonet M and Gruss P: Pax genes and their roles in cell differentiation and development. Curr Opin Cell Biol 8: 851-857, 1996.
4. Mansouri A, Goudreau G and Gruss P: Pax genes and their role in organogenesis. Cancer Res 59 (Suppl): S1707-S1710, 1999.
5. Poleev A, Fickenscher H, Mundlos S, Winterpacht A, Zabel B, Fidler A, Gruss P and Plachov D: PAX8, a human paired box gene: Isolation and expression in developing thyroid, kidney and Wilms' tumors. Development 116: 611-623, 1992.
6. Kozmik Z, Kurzbauer R, Dörfler P and Busslinger M: Alternative splicing of Pax-8 gene transcripts is developmentally regulated and generates isoforms with different transactivation properties. Mol Cell Biol 13: 6024-6035, 1993.
7. Szczepanek-Parulska E, Szaflarski W, Piątek K, Budny B, Jaszczyńska-Nowinka K, Biczysko M, Wierzbicki T, Skrobisz J, Zabel M and Ruchała M: Alternative 3'acceptor site in the exon 2 of human PAX8 gene resulting in the expression of unknown mRNA variant found in thyroid hemiagenesis and some types of cancers. Acta Biochim Pol 60: 573-578, 2013.
8. Miccadei S, Provenzano C, Mojzisek M, Natali PG and Civitareale D: Retinoblastoma protein acts as Pax 8 transcriptional coactivator. Oncogene 24: 6993-7001, 2005.
9. Li CG, Nyman JE, Braithwaite AW and Eccles MR: PAX8 promotes tumor cell growth by transcriptionally regulating E2F1 and stabilizing RB protein. Oncogene 30: 4824-4834, 2011.
10. Chen Y-J, Campbell HG, Wiles AK, Eccles MR, Reddel RR, Braithwaite AW and Royds JA: PAX8 regulates telomerase reverse transcriptase and telomerase RNA component in glioma. Cancer Res 68: 5724-5732, 2008.
11. Cheung HW, Cowley GS, Weir BA, Boehm JS, Rusin S, Scott JA, East A, Ali LD, Lizotte PH, Wong TC, et al: Systematic investigation of genetic vulnerabilities across cancer cell lines reveals lineage-specific dependencies in ovarian cancer. Proc Natl Acad Sci USA 108: 12372-12377, 2011.
12. Marques AR, Espadinha C, Catarino AL, Moniz S, Pereira T, Sobrinho LG and Leite V: Expression of PAX8-PPAR gamma 1 rearrangements in both follicular thyroid carcinomas and adenomas. J Clin Endocrinol Metab 87: 3947-3952, 2002.
13. French CA, Alexander EK, Cibas ES, Nose V, Laguette J, Faquin W, Garber J, Moore F Jr, Fletcher JA, Larsen PR, et al: Genetic and biological subgroups of low-stage follicular thyroid cancer. Am J Pathol 162: 1053-1060, 2003.
14. Dwight T, Thoppe SR, Foukakis T, Lui WO, Wallin G, Höög A, Frisk T, Larsson C and Zedenius J: Involvement of the PAX8/peroxisome proliferator-activated receptor gamma rearrangement in follicular thyroid tumors. J Clin Endocrinol Metab 88: 4440-4445, 2003.
15. Kroll TG, Sarraf P, Pecciarini L, Chen CJ, Mueller E, Spiegelman BM and Fletcher JA: PAX8-PPARgamma1 fusion oncogene in human thyroid carcinoma [corrected]. Science 289: 1357-1360, 2000.
16. Laury AR, Perets R, Piao H, Krane JF, Barletta JA, French C, Chirieac LR, Lis R, Loda M, Hornick JL, et al: A comprehensive analysis of PAX8 expression in human epithelial tumors. Am J Surg Pathol 35: 816-826, 2011.
17. Ordóñez NG: Value of PAX 8 immunostaining in tumor diagnosis: A review and update. Adv Anat Pathol 19: 140-151, 2012.
18. Nonaka D, Tang Y, Chiriboga L, Rivera M and Ghossein R: Diagnostic utility of thyroid transcription factors Pax8 and TTF-2 (FoxE1) in thyroid epithelial neoplasms. Mod Pathol 21: 192-200, 2008.
19. Scouten WT, Patel A, Terrell R, Burch HB, Bernet VJ, Tuttle RM and Francis GL: Cytoplasmic localization of the paired box gene, Pax-8, is found in pediatric thyroid cancer and may be associated with a greater risk of recurrence. Thyroid 14: 1037-1046, 2004.
20. Ozcan A, Shen SS, Hamilton C, Anjana K, Coffey D, Krishnan B and Truong LD: PAX 8 expression in non-neoplastic tissues, primary tumors, and metastatic tumors: A comprehensive immunohistochemical study. Mod Pathol 24: 751-764, 2011.
21. Tong G-X, Yu WM, Beaubier NT, Weeden EM, Hamele-Bena D, Mansukhani MM and O'Toole KM: Expression of PAX8 in normal and neoplastic renal tissues: An immunohistochemical study. Mod Pathol 22: 1218-1227, 2009.
22. Albadine R, Schultz L, Illei P, Ertoy D, Hicks J, Sharma R, Epstein JI and Netto GJ: PAX8 (+)/p63 (-) immunostaining pattern in renal collecting duct carcinoma (CDC): A useful immunoprofile in the differential diagnosis of CDC versus urothelial carcinoma of upper urinary tract. Am J Surg Pathol 34: 965-969, 2010.
23. Mhawech-Fauceglia P, Wang D, Samrao D, Godoy H, Pejovic T, Liu S and Lele S: Pair-Box (PAX8) protein-positive expression is associated with poor disease outcome in women with endometrial cancer. Br J Cancer 107: 370-374, 2012.
24. Wang Y, Wang Y, Li J, Yuan Z, Yuan B, Zhang T, Cragun JM, Kong B and Zheng W: PAX8: A sensitive and specific marker to identify cancer cells of ovarian origin for patients prior to neoadjuvant chemotherapy. J Hematol Oncol 6: 60, 2013.
25. Tacha D, Zhou D and Cheng L: Expression of PAX8 in normal and neoplastic tissues: A comprehensive immunohistochemical study. Appl Immunohistochem Mol Morphol 19: 293-299, 2011.
26. Kenny SL, McBride HA, Jamison J and McCluggage WG: Mesonephric adenocarcinomas of the uterine cervix and corpus: HPV-negative neoplasms that are commonly PAX8, CA125, and HMGA2 positive and that may be immunoreactive with TTF1 and hepatocyte nuclear factor 1-β. Am J Surg Pathol 36: 799-807, 2012.
27. Shukla A, Thomas D and Roh MH: PAX8 and PAX2 expression in endocervical adenocarcinoma in situ and high-grade squamous dysplasia. Int J Gynecol Pathol 32: 116-121, 2013.
28. Ruiz-Llorente S, Carrillo Santa de Pau E, Sastre-Perona A, Montero-Conde C, Gómez-López G, Fagin JA, Valencia A, Pisano DG and Santisteban P: Genome-wide analysis of Pax8 binding provides new insights into thyroid functions. BMC Genomics 13: 147, 2012.
29. David CJ and Manley JL: Alternative pre-mRNA splicing regulation in cancer: Pathways and programs unhinged. Genes Dev 24: 2343-2364, 2010.
30. Nonaka D, Chiriboga L and Soslow RA: Expression of pax8 as a useful marker in distinguishing ovarian carcinomas from mammary carcinomas. Am J Surg Pathol 32: 1566-1571, 2008.
31. Zhao L, Guo M, Sneige N and Gong Y: Value of PAX8 and WT1 immunostaining in confirming the ovarian origin of metastatic carcinoma in serous effusion specimens. Am J Clin Pathol 137: 304-309, 2012.
32. Szczepanek-Parulska E, Szaflarski W, Piątek K, Budny B, Jaszczyńska-Nowinka K, Biczysko M, Wierzbicki T, Skrobisz J, Zabel M and Ruchała M: Alternative 3' acceptor site in the exon 2 of human PAX8 gene resulting in the expression of unknown mRNA variant found in thyroid hemiagenesis and some types of cancers. Acta Biochim Pol 60: 573-578, 2013.
33. Kalnina Z, Zayakin P, Silina K and Linē A: Alterations of pre-mRNA splicing in cancer. Genes Chromosomes Cancer 42: 342-357, 2005.
34. Ohta M, Inoue H, Cotticelli MG, Kastury K, Baffa R, Palazzo J, Siprashvili Z, Mori M, McCue P, Druck T, et al: The FHIT gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers. Cell 84: 587-597, 1996.
35. Bednarek AK, Keck-Waggoner CL, Daniel RL, Laflin KJ, Bergsagel PL, Kiguchi K, Brenner AJ and Aldaz CM: WWOX, the FRA16D gene, behaves as a suppressor of tumor growth. Cancer Res 61: 8068-8073, 2001.
36. Stapleton P, Weith A, Urbánek P, Kozmik Z and Busslinger M: Chromosomal localization of seven PAX genes and cloning of a novel family member, PAX-9. Nat Genet 3: 292-298, 1993.

37. Karolchik D, Barber GP, Casper J, Clawson H, Cline MS, Diekhans M, Dreszer TR, Fujita PA, Guruvadoo L, Haeussler M, *et al*: The UCSC Genome Browser database: 2014 update. Nucleic Acids Res 42 (D1): D764-D770, 2014.
38. Hecht F, Ramesh KH and Lockwood DH: A guide to fragile sites on human chromosomes. Cancer Genet Cytogenet 44: 37-45, 1990.
39. IJdo JW, Baldini A, Wells RA, Ward DC and Reeders ST: FRA2B is distinct from inverted telomere repeat arrays at 2q13. Genomics 12: 833-835, 1992.
40. Cheung L, Messina M, Gill A, Clarkson A, Learoyd D, Delbridge L, Wentworth J, Philips J, Clifton-Bligh R and Robinson BG: Detection of the PAX8-PPARγ fusion oncogene in both follicular thyroid carcinomas and adenomas. J Clin Endocrinol Metab 88: 354-357, 2003.
41. Solinas-Toldo S, Dürst M and Lichter P: Specific chromosomal imbalances in human papillomavirus-transfected cells during progression toward immortality. Proc Natl Acad Sci USA 94: 3854-3859, 1997.