Genome Analysis of Latin American Cervical Cancer: Frequent Activation of the PIK3CA Pathway

Hong Lou1, Guillermo Villagran2, Joseph F. Boland3, Kate M. Im4, Sarita Polo2, Weiyn Zhou3, Ushie Odey5, Elijia Juárez-Torres6, Ingrid Medina-Martínez6, Edgar Roman-Basaure7, Jason Mitchell3, David Roberson3, Julie Sawitzke1, Lisa Garland1, Maria Rodríguez-Herrera6, David Wells1, Jennifer Troyer1, Francisco Castillo Pinto2, Sara Bass3, Xijun Zhang3, Miriam Castillo2, Bert Gold4, Hesler Morales2, Meredith Yeager3, Jaime Berumen6, Enrique Álvarez5, Eduardo Gharzouzi2, and Michael Dean4

Abstract

Purpose: Cervical cancer is one of the most common causes of cancer mortality for women living in poverty, causing more than 28,000 deaths annually in Latin America and 266,000 worldwide. To better understand the molecular basis of the disease, we ascertained blood and tumor samples from Guatemala and Venezuela and performed genomic characterization.

Experimental Design: We performed human papillomavirus (HPV) typing and identified somatically mutated genes using exome and ultra-deep targeted sequencing with confirmation in samples from Mexico. Copy number changes were also assessed in the exome sequence.

Results: Cervical cancer cases in Guatemala and Venezuela have an average age of diagnosis of 50 years and 5.6 children. Analysis of 675 tumors revealed activation of PIK3CA and other PI3K/AKT pathway genes in 31% of squamous carcinomas and 24% of adenocarcinoma and adenosquamous tumors, predominantly at two sites (E542K, E545K) in the helical domain of the PIK3CA gene. This distribution of PIK3CA mutations is distinct from most other cancer types and does not result in the in vitro phosphorylation of AKT. Somatic mutations were more frequent in squamous carcinomas diagnosed after the age of 50 years. Frequent gain of chromosome 3q was found, and low PIK3CA mutation fractions in many tumors suggest that PI3K mutation can be a late event in tumor progression.

Conclusions: PI3K pathway mutation is important to cervical carcinogenesis in Latin America. Therapeutic agents that directly target PI3K could play a role in the therapy of this common malignancy. Clin Cancer Res; 21(23); 5360–70. ©2015 AACR.

Introduction

Human papillomavirus (HPV) causes more than 90% of cervical cancer, one of the most common malignancies in women worldwide (1–3). While 70% to 90% of infections are cleared by the immune system, persistent HPV infections can lead to high-grade cervical intraepithelial neoplasia (CIN) and cervical cancer (4). HPV replicates episomally but can integrate; integration is more frequent in late-stage lesions and cancer and is associated with genome alteration and instability (5, 6). The viral E6 and E7 oncoproteins inhibit both TP53 and RB1 proteins altering cell-cycle control, apoptosis, and DNA repair.

The HPV16 and HPV18 types are the most oncogenic and account for 60% to 70% of cervical cancers; however, at least 12 other high-risk types can also be found in cervical tumors (7–9). HPV is both highly prevalent and highly infectious, being transmitted through multiple forms of sexual contact, and most males and females acquire one or more infections in their lifetime. Because of the long latency of development of cervical cancer, typically 10 to 15 years, there is opportunity to identify precancers and eliminate them before the appearance of invasive cancer (4). The use of the Papanicolaou (Pap) test has reduced cervical cancer mortality by up to 80% in countries that have used screening (3). However, women living in poverty, with inadequate health care, receive either no screening or poorly controlled screening, and nearly 90% of cervical cancer mortality occurs in low- and middle-income countries (LMIC), and within minority populations in higher income countries (3, 10).
Translational Relevance

Cervical cancer is one of the most common cancers in women worldwide with more than 80% of deaths occurring in women living in poverty. Although precancerous lesions and local malignancy is curable, invasive and/or metastatic tumors have poor survival. In invasive tumors from 3 Latin American countries, we identify common activation of the PI3K pathway. Up to 33% of tumors have mutations in the PIK3CA gene, predominantly at 2 specific helical domain sites, E542K and E545K, and rarely in the kinase domain. The PI3K pathway can signal through the AKT1 and mTOR serine-threonine kinases; however, PIK3CA helical mutations may activate alternative pathway(s). Our data are relevant to the development and application of therapeutic strategies for invasive cervical tumors.

Many Latin American countries have high incidence and mortality from cervical cancer [8, 414 deaths annually in Brazil; 4,769 in Mexico, and 27,000 in the region overall (http://globocan.iarc.fr)]. Several Latin American countries have large minority and/or indigenous populations that through poverty, discrimination, rural isolation, and/or language barriers suffer from health disparities. Indigenous populations in Guatemala comprise 40% of the inhabitants and speak more than 2 languages, Mexico has 75 indigenous groups and Venezuela has 40 recognized indigenous peoples.

Until recently, molecular genetic studies of cervical tumors have been limited to candidate gene tests and studies of cell lines such as HeLa (11). Because the E6 protein of high-risk types causes the degradation of TP53, cervical tumors are one of the few cancers with a low level of mutation in the TP53 gene (12). The highly pathogenic HPV16 and HPV18 types can immortalize many cell types, but additional lesions are required for transformation (13). The PIK3CA gene pathway has also been shown to be frequently mutated (14) with a 31% mutation frequency in U.S. cervical tumors. PIK3CA encodes an enzyme converting PIP3 back to PIP2 and is a tumor suppressor gene. PIK3CA is one of the most frequently mutated genes in cancer, with most mutations occurring in the kinase domain and resulting in constitutive activation. The helical domain of PIK3CA is also frequently mutated at residues E542 and E545 and this domain is thought to mediate interactions with an inhibitory subunit and may signal through other kinases such as SGK3 (15, 16).

Genome sequencing of cervical tumors with viral integration revealed widespread genome rearrangement and specific sites of integration (17). A comprehensive study of 100 Norwegian and 15 Mexican tumors with genome, exome, and RNA-seq analyses identified frequent driver genes, common chromosomal alterations, and integration sites (18). A genome sequencing study of HPV integration in Chinese cervical tumors revealed diverse integration sites, and a study of 15 cervical adenocarcinomas from Hong Kong identified frequently mutated genes (19, 20).

To further understand the molecular basis of cervical cancer in high-incidence countries, we undertook an analysis of cervical cancers in Guatemala and Venezuela with validation from a cohort from Mexico to identify frequently mutated genes and correlate mutations with histologic type, HPV type, and age of cervical cancer onset.

Materials and Methods

Sample collection

Samples were obtained under Institutional Review Board (IRB) approval with informed consent, a standardized questionnaire on sociodemographic characteristics, reproductive and contraceptive history, smoking, and Pap smear history. Protocols were approved by the Ethical Review Committees in Guatemala, Venezuela, and Mexico and the Office of Human Research Subjects, NIH. Subjects from Guatemala (296 FIGO stage I–IV cancers) and Venezuela (39 with FIGO stage I–IV, and 38 CIN 1–3) were collected from 2011 to 2013 and from Mexico from 2003 to 2007. All patients were referred for suspected invasive cervical cancer, all consenting patients were included, and only pregnant subjects and cancer-free women were excluded. Subjects from Venezuela included subjects with a positive Pap smear and biopsy (CIN 1–CIN 3). Surgically resected cervical tumors tissues were stored in RNA later (QIAGEN) at −20°C until extraction (Supplementary Fig. S1).

DNA and RNA extraction and HPV genotyping

DNA and RNA were extracted from the cervical cancer tissues (5–10 mg) using the AllPrep DNA/RNA Micro Kit (QIAGEN) as described by the manufacturer. For HPV type determination, DNA samples were amplified by PCR using Broad-Spectrum GP5+/BSGP5+ forward primers with inosine at selected sites and GP6+/reverse primers along with the β-globin MS3/MS10 primers (226 bp) as a control for DNA quality (21). Four hundred nanomolars of each forward primer were used with ZymoEaq PreMix (Zymo Research). A 10-minute denaturation step (95°C) was followed by 40 cycles of amplification (Perkin-Elmer thermocycler). Each cycle was 94°C for 20 seconds, 38°C for 30 seconds, and 71°C for 80 seconds and final elongation for 5 minutes. Ramping rates for the Mastercycler were: 1.8°C/s from 94°C to 38°C, 2.0°C/s from 38°C to 71°C, and 2.8°C/s from 71°C to 94°C. Each PCR experiment included HeLa DNA as positive and HEK293 and/or C33A as negative controls and sample lacking template DNA. Positive samples by gel electrophoresis were sequenced on an ABI 3730XL and analyzed by assembly and trimming in SeqMan (DNASTAR) followed by BLAST search (NCBI).

NextGen sequencing of HPV

To resolve multiple types on the Ion Torrent PGM, the BSGP6+ primers were tailed with Ion Express barcodes and Ion Torrent A adapter and BSGP5+ primers tailed with the P1 adapter (Supplementary Table S1). β-globin MS3/MS10 primers were included in the reaction to control for amplification, without sequencing adapters or barcodes. When samples were PCR-negative for HPV but amplified for β-globin (by gel electrophoresis), a second PCR reaction was performed using 2× DNA input material. Positive PCR products were quantitated on the Caliper GX, normalized and pooled for sequencing on the Ion Torrent PGM as per manufacturer’s instructions. Briefly, normalized, pooled libraries...
were amplified via emulsion PCR using the One Touch v2, enriched on the ES2, and sequenced for 520 cycles on the PGM. An average of 6,000 reads was obtained for each sample. There was an 83% concordance between the PCR and next-generation sequencing methods (including both HPV+ and HPV− samples), and another 6% of samples had mixed HPV infection by next generation with Sanger sequencing detecting one of the types. A total of 8% of samples had abundant reads of more than one HPV type by next-generation sequencing.

Exome sequencing
DNA preparation. A 1-μg aliquot of high-molecular-weight DNA (determined by Picogreen, Invitrogen) was used in TargetSeq exome v2 capture process with enzymatic shearing (Ion Shear Plus Reagents Kit, Life Technologies) to a target size range of 135 to 165 bp.

Library preparation for TargetSeq exome capture. Sheared gDNA followed the TargetSeq protocol for ligation, nick repair, purification, size selection, and final amplification. For the ligation and nick repair, a master mix consisting of 10 μL 10× Ligase Buffer, 10 μL L A and P1 adapters, 2 μL dNTP mix, 41 μL nuclease free water, 4 μL DNA ligase, and 8 μL nick repair polymerase. PCR conditions were as follows: 25°C for 15 minutes and 98°C for 20 minutes, and a hold at 4°C. The amplified material was cleaned with Ampure XP reagent (Agencourt) according to the TargetSeq v2 protocol, with DNA elution in 20 μL of Low TE. Amplified sample libraries were size selected (Pippen Prep instrument, Sage Science) and processed according to the TargetSeq protocol. The Pippen Prep was set to elute “Tight” at 220 bp. The eluted size selected samples were cleaned with Ampure XP reagent according to the TargetSeq v2 protocol and DNA eluted in 30 μL of Low TE. The final amplification of the size-selected fragment libraries was performed using a master mix consisting of 200 μL Platinum PCR Supermix High Fidelity and 20 μL Library Amplification Primer Mix. PCR cycling conditions were: 95°C for 5 minutes, and 8 cycles of 95°C for 15 seconds, 58°C for 15 minutes, and 72°C for 1 minute. The amplification of the PCR and next-generation sequencing methods (including both HPV+ and HPV− samples), and another 6% of samples had mixed HPV infection by next generation with Sanger sequencing detecting one of the types. A total of 8% of samples had abundant reads of more than one HPV type by next-generation sequencing.

Exome sequencing
DNA preparation. A 1-μg aliquot of high-molecular-weight DNA (determined by Picogreen, Invitrogen) was used in TargetSeq exome v2 capture process with enzymatic shearing (Ion Shear Plus Reagents Kit, Life Technologies) to a target size range of 135 to 165 bp.

Library preparation for TargetSeq exome capture. Sheared gDNA followed the TargetSeq protocol for ligation, nick repair, purification, size selection, and final amplification. For the ligation and nick repair, a master mix consisting of 10 μL 10× Ligase Buffer, 10 μL L A and P1 adapters, 2 μL dNTP mix, 41 μL nuclease free water, 4 μL DNA ligase, and 8 μL nick repair polymerase. PCR conditions were as follows: 25°C for 15 minutes and 98°C for 20 minutes, and a hold at 4°C. The amplified material was cleaned with Ampure XP reagent (Agencourt) according to the TargetSeq v2 protocol, with DNA elution in 20 μL of Low TE. Amplified sample libraries were size selected (Pippen Prep instrument, Sage Science) and processed according to the TargetSeq protocol. The Pippen Prep was set to elute “Tight” at 220 bp. The eluted size selected samples were cleaned with Ampure XP reagent according to the TargetSeq v2 protocol and DNA eluted in 30 μL of Low TE. The final amplification of the size-selected fragment libraries was performed using a master mix consisting of 200 μL Platinum PCR Supermix High Fidelity and 20 μL Library Amplification Primer Mix. PCR cycling conditions were: 95°C for 5 minutes, and 8 cycles of 95°C for 15 seconds, 58°C for 15 seconds, 70°C for 1 minute, purified with Ampure XP reagent according to the TargetSeq v2 protocol, eluted in 25 μL of Low TE, quantitated and qualitatively assessed on the Agilent Bioanalyzer.

Ampliseq exome sequencing. Aliquots of 100 ng of gDNA from cervical tumor and normal samples were processed according to the standard protocol for Ampliseq target amplification and library preparation. Each tumor and normal library were pooled for the template emulsion prep and sequenced as a pair using the Proton P1 chip and Ion Torrent Proton Sequencer (Thermo Fisher Scientific). Each run produced over 10 Gb of sequence data and had an average depth of coverage surpassing 100×. The data were aligned using TMAP (https://github.com/iontorrent/TS/tree/master/Analysis/TMAP), and variants were called using Torrent Suite Variant Caller (TSVC) onboard the Proton Sequencer (Thermo Fisher Scientific). The data were processed offline through a custom analysis workflow using the aligned reads and a dual variant calling process, TSVC, and a modified GATK variant caller optimized for Proton data (22).

Copy number analysis
The copy number variation (CNV) analysis was performed using tumor and matched normal sequence. A log2 ratio between tumor and normal was calculated on the basis of the ngCGH algorithm (https://github.com/seandavi/ngCGH), using the tumor and normal BAM files. Genomic windows are defined from blocks of 1000 reads in the normal sample and then the number of reads in the tumor is quantified. A ratio is made between the number of reads in the tumor and the number of reads in the normal. Finally, a log2 transformation is applied to each ratio, and the entire vector of the results is centered by subtracting the median to make the median of the log2 ratios zero.

The log2 ratios were imported to Nexus Copy Number Discovery Edition Version 7.5 (BioDiscovery, Inc.; http://www.biodiscovery.com). The Fast Adaptive States Segmentation Technique (FASST2) segmentation method (BioDiscovery, Inc.) was used to make CNV calls. A significance threshold of 1.0E−5 was used to adjust the sensitivity of the FASST2 segmentation algorithm. A minimum number of 20 amplicons per segment were used to eliminate small CNVs. Cutoffs of 0.2/−0.2 were used for gain/loss and for high gain/high loss were set to 0.6/−1.0, and the results tabulated (Supplementary Fig. S4).

Targeted gene sequencing
A targeted, multiplex PCR primer panel was designed using the custom Ion Ampliseq Designer v1.2 (Life Technologies). The primer panel covered 12 kb of sequence including the coding region of 8 genes—HRAS, CTNNB1, KRAS, STK11, CDKN2A, PIK3CA, PTEN, and TP53 (average coverage, 99%; average amplicon size, 225 bp). Sample DNA (tumor or tumor/ normal pairs) was amplified and libraries prepared following the Ion Ampliseq Library Preparation protocol (Life Technologies). Individual samples were barcoded, pooled, applied to chips, and sequenced on the Ion Torrent PGM Sequencer using the Ion PGM Template OT2 200 and Ion PGM Sequencing 200v2 kits. Mean read length after sequencing was 116 bp, and 94% of amplicons gave an average coverage of greater than 50 reads per sample.
Sequence alignment and mutation prediction

Resulting sequence reads were aligned to the human reference genome version hg19 using the TMAP aligner (Life Technologies) and single-nucleotide variants (SNV) were called using the Genome Analysis Tool Kit (GATK; refs. 22, 23) and the Torrent Variant Caller (TSVC, Life Technologies), and small insertions and deletions were called using the TSVC. The NIH Biowulf Cluster was used for additional variant annotation. All mutations in Supplementary Tables S1 and S2 were manually examined in IGV (24) to confirm an adequate number of mutant reads in both directions and to eliminate false-positives. Selected sites were manually examined to identify potential false-negative predictions. For the E542K and E545K sites, a minimum of 100 reads and 3% mutant reads were necessary to call the sample mutation positive. For a subset of samples, targeted sequencing was performed on both tumor and normal DNA to confirm that reported mutations were somatic.

PIK3CA mutation verification

Primers for the PIK3CA gene were described (25). The amplification (WGQIACEN) required 20 ng cervical cancer tumor or normal blood DNA and standard PCR conditions at an annealing temperature of 63°C for 10 cycles and 58°C for 30 cycles, respectively. Big Dye v.3.1 chemistry (ABI) sequencing reactions were performed on a 3730 Genetic Analyzer (ABI) and chromatograms were analyzed using Sequencher, v.4.8 (GeneCodes) and Mutation Surveyor (Softgenetics). Exons 9 and 20 were sequenced in the majority of the Venezuela and Guatemala tumor samples and selected mutations in other gene regions were also validated.

Quantitative RT real-time PCR

One microgram of total RNA was reverse-transcribed into cDNA using the Transcriptor First Strand cDNA Synthesis Kit (Roche) with oligonucleotide (dT)18 primer according to the manufacturer’s instructions. Real-time PCR was performed for E6 and E7 HPV transcripts, using gene- and type-specific primers, in the presence of SYBR green, in HPV16- and 18-positive tumors. Relative expression levels of E6 and E7 were multiplied by 1,000. The E6 primers detect the full-length E6/E7 transcript producing principally E6 protein and the E7 primers detect transcripts expressing E6, E6’1, E6’2, and E7 proteins (Supplementary Fig. S10; ref. 26). For normalization, the expression of a 144-bp β-actin fragment was used. Each experiment included HeLa cDNA and Ca-Ski cDNA as positive and C-33A as negative controls. The relative mRNA expression level of PIK3CA was calibrated with HUC (Human Universal Control from Clontech). Primers and probes for the PIK3CA gene (Hs00907966_m1) and β-actin gene (Hs99999903_m1) were from Applied Biosystems.

Site-directed mutagenesis and phosphorylation determination

Site-directed mutations were introduced into a full length PIK3CA expression vector (Origene) and confirmed by sequencing using the QuickChange II XL Kit (Agilent). The helical domain mutations E542K, E545K, E542K/E545K, E542Q/E545K, Q546R, D549H of p110α were compared with the kinase domain mutation H11047R. The empty vector, wild-type PIK3CA, and mutant PIK3CA expression constructs were transiently transfected into U2OS cells using HylitMax (DojinDo). The U2OS cells were serum-starved in DMEM containing 0.5% FBS overnight and treated with 0.1 μmol/L Calcuclin A (Cell Signaling Technology) in 0% FBS for 30 minutes before cell lysis. Whole-cell protein (25 μg) was separated in a 4% to 12% NuPAGE Bis–Tris gel, transferred to polyvinylidene difluoride (PVDF) membrane (Invitrogen). The primary antibodies were monoclonal rabbit- and phospho-AKTSer473, phospho-AKTThr308, total AKT (Cell Signaling Technology) at a dilution of 1:8,000 to 1:10,000, rabbit anti-p110α antibody, and β-actin was used as a control. HRP-conjugated anti-rabbit IgG was used as secondary antibody (Cell Signaling Technology).

Statistical analyses

Mann–Whitney U, 2-tailed t test, 1-way ANOVA, Kruskal–Wallis, Pearson χ2 test, and Fisher exact test statistical analyses were performed using GraphPad Prism version 5 for Windows. P < 0.05 was regarded to be statistically significant.

Results

Exome and targeted gene sequencing

To determine the HPV types and molecular characteristics of Latin American women with cervical tumors, subjects were prospectively enrolled and tumor tissue and blood collected at the Instituto de Cancerología (Guatemala) and the Hospital Central Universitario (Venezuela). A comparison sample set of 330 subjects from Mexico was also included (Supplementary Fig. S1). There was no significant difference between squamous carcinoma and adeno- and adenosquamous tumors in age at collection, reproductive factors, HPV type, or smoke exposure; however, adeno- and adenosquamous tumors are more often diagnosed at stage I (P = 0.061, Table 1). To identify potential cancer genes, 23 Guatemalan cervical cancer and corresponding normal blood DNAs were subjected to exome sequencing.

Table 1. Patient demographics and disease characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total (n = 511)</th>
<th>Squamous (n = 426)</th>
<th>Adenocarcinoma/</th>
<th>Adenosquamous (n = 65)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age at collection</td>
<td>52 ± 13.2</td>
<td>52 ± 13.2</td>
<td>49.3 ± 12.4</td>
<td>0.055*</td>
<td></td>
</tr>
<tr>
<td>Mean age at menarche</td>
<td>13.3 ± 1.5</td>
<td>13.3 ± 1.5</td>
<td>15.4 ± 1.6</td>
<td>0.210</td>
<td></td>
</tr>
<tr>
<td>Mean pregnancies</td>
<td>5.9 ± 3.3</td>
<td>6.0 ± 3.3</td>
<td>5.3 ± 3.0</td>
<td>0.0061b</td>
<td></td>
</tr>
<tr>
<td>Mean age at first birth</td>
<td>18.5 ± 2.7</td>
<td>18.6 ± 2.6</td>
<td>17.7 ± 2.9</td>
<td>0.010c</td>
<td></td>
</tr>
<tr>
<td>HPV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPV16</td>
<td>252 (51.5)</td>
<td>221 (51.5)</td>
<td>31 (51.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPV18</td>
<td>48 (9.8)</td>
<td>38 (8.9)</td>
<td>10 (16.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPV35</td>
<td>38 (7.8)</td>
<td>33 (7.7)</td>
<td>5 (8.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All others</td>
<td>151 (30.8)</td>
<td>137 (31.9)</td>
<td>14 (23.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>122 (23.3)</td>
<td>96 (20.8)</td>
<td>26 (40.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>215 (40.7)</td>
<td>193 (41.9)</td>
<td>20 (31.2)c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>173 (33.0)</td>
<td>157 (34.1)</td>
<td>16 (25.0)d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>16 (3.0)</td>
<td>14 (3.0)</td>
<td>2 (3.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guatemala</td>
<td>208 (39.2)</td>
<td>192 (41.2)</td>
<td>16 (24.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mexico</td>
<td>323 (60.8)</td>
<td>274 (58.8)</td>
<td>49 (75.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cooking method</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gas</td>
<td>57 (22.3)</td>
<td>25 (22.1)</td>
<td>3 (27.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wood</td>
<td>155 (60.6)</td>
<td>74 (65.5)</td>
<td>6 (54.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gas and wood</td>
<td>44 (17.2)</td>
<td>44 (17.4)</td>
<td>5 (41.1)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTE: The table includes all samples with known pathology that are squamous carcinoma, adenosquamous, or adenosquamous carcinoma. Means are followed by SD and numbers by percentages.

*aPearson χ2 test.

*bFisher exact test.

*cP < 0.001.
Variants in genes predicted to be mutated in 2 or more tumors and in known somatic cancer genes (18, 27, 28) are shown (Fig. 1). The tumors contained a predominance of C-T and T-C mutations (Supplementary Fig. S2). Predicted somatic mutations were found in known cancer genes including PIK3CA, RB1, TP53, MAPK1, HRAS, KRAS, TSC1, BRCA1, BRCA2, BAP1, and ATM. The tumor and normal exome sequence was also used for copy number determination and as seen previously (18, 29), nearly all Guatemalan tumors tested have 3 to 5 copies of chromosome 3q, 43% have gain of 5q, and 14 of 23 (61%) have chromosome loss at 17p containing the TP53 loci (Figs. 1 and 2). In total, 8 of 23 cervical tumors show extensive chromosomal rearrangement (>100 chromosome breaks; CHR re), gain of chromosome 3q, and loss of 17p are shown.

Figure 1. Driver gene mutations in 23 Guatemalan cervical cancers. Genes frequently mutated or amplified in other cancers are indicated, with their gene names at the left and the mutation percentages (%) on the right. The predominant HPV type and pathology is indicated at the top (unlabeled tumors are squamous cell carcinomas). Below the main section, the presence of chromosome rearrangement (>100 chromosome breaks, CHR re), gain of chromosome 3q, and loss of 17p are shown.

Distribution of mutations in the PIK3CA gene

Somatic PIK3CA mutations were detected in 33% (91 of 280) of carcinomas from Guatemala, 28% (11 of 40) from Venezuela, and 28% (91 of 325) from Mexico. Only 2 of 27 CIN grade 1, 2, or 3 lesions from Venezuela (one CIN2 and one CIN3) had PIK3CA mutations (8%), indicating that these mutations occur predominantly in malignant tumors. When tumors were divided by histological type, 155 of 499 squamous cell carcinomas (31%), 16 of 67 adenocarcinomas (24%), and 5 of 21 adenosquamous carcinomas (24%) have a PIK3CA mutation. A total of 95% of all PIK3CA mutations were located in the ABD, ABD–RBD linker, C2, and helical domains of PIK3CA, whereas mutations in the kinase domain were rare (Fig. 3A and B). In fact, 2 specific mutations (E542K, E545K) account for 81% of the PIK3CA mutations in Mexico and 76% in Guatemala (Fig. 3), and helical domain mutations are significantly more common in squamous cell versus adenocarcinomas (P = 0.017, Supplementary Fig. S5). Helical PIK3CA nucleotides (500× average, Supplementary Figs. S1 and S4). Predicted variants are shown (Supplementary Tables S2 and S3). Mutations in HRAS, KRAS, CDRN2A, and CTNNB1 were infrequent; however, genes in the PI3K pathway were frequently mutated. Specifically, activating mutations were common in PIK3CA, and inactivating mutations in PTEN and STK11 were also found (Table 2 and Supplementary Tables S2 and S3).
mutations are more frequent in cervical and bladder cancer as compared with breast, endometrial, or intestinal tumors ($P < 0.0001$; Fig. 3C). Several tumors had more than one PIK3CA or PTEN mutation. For adjacent mutations, their presence on either the same DNA strand (cis) or different strands (trans) could be determined, and both cis and trans examples were found (Supplementary Fig. S6).

Interestingly, despite tumors having 70% or greater tumor cells by pathologic examination, many tumors with PIK3CA mutations have a low percentage of mutant reads. A total of 79% of tumors with a PIK3CA mutation have between 4 and 33% mutant reads, and this is constant across grades I–IV (78%–83%), suggesting that there is considerable tumor heterogeneity. Targeted sequence on tumor and normal DNA and Sanger sequencing was used to confirm selected PIK3CA mutations were somatic variants and the approximate mutated allele fraction was comparable to that obtained from next-generation sequencing (Supplementary Fig. S7). The high frequency of E542K and E545K mutations allowed the determination of mutation fraction for each individual mutation and for double mutants (Supplementary Fig. S8A and S8B). E542K is found 31% of the time at a mutation fraction greater than 20%, whereas 51% of E545K-containing tumors have more than 20% mutant reads ($X^2 = 5.5; P = 0.019$). However, for the 6 E542K/E545K double mutants, the 2 mutations are always on different haplotypes and E542K is usually more prevalent (Supplementary Fig. S8B).

Several PIK3CA mutations described here have not previously been reported (delN107, E1034Q in Vtacxuela; T229I, Q861R, K942M, and V952G in Guatemala; and V146L, M299V, delH419_C420, and G914R in Mexican tumors), whereas the R38H, R88Q, K111E, K111N, E453K, E542K, E545K, and H1047R variants are documented somatic gain-of-function alleles (refs. 25, 30–32; Fig. 3A). Both combined PIK3CA mutations and overall mutations are statistically significantly increased in patients diagnosed at a later age and are less common in HPV18- and HPV45-positive tumors ($P = 0.0016$; Supplementary Fig. S9 and Supplementary Table S4). While adenocarcinomas occur with a younger average age and a lower tumor stage and may confound these relationships, squamous cell subjects diagnosed before the age of 50 years have significantly fewer PIK3CA mutations ($P = 0.0001$). Comparison of our Latin American data with data from other countries (14, 18, 33; COSMIC database) demonstrates differences in frequency of PIK3CA mutations in individual domains, as well as the E542K, E545K, and H1047R sites (Fig. 3C and Supplementary Table S5).

### Table 2. Summary of mutations in targeted gene panel

<table>
<thead>
<tr>
<th>Gene</th>
<th>Guatemala</th>
<th>%</th>
<th>Venezuela CIN</th>
<th>%</th>
<th>Venezuela cervical cancer</th>
<th>%</th>
<th>Mexico</th>
<th>%</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIK3CA</td>
<td>91/280</td>
<td>33%</td>
<td>2/30</td>
<td>7%</td>
<td>11/40</td>
<td>28%</td>
<td>91/325</td>
<td>28%</td>
<td>30%</td>
</tr>
<tr>
<td>TP53</td>
<td>15/280</td>
<td>5.4%</td>
<td>0/30</td>
<td>0%</td>
<td>2/40</td>
<td>5%</td>
<td>15/325</td>
<td>5%</td>
<td>5.0%</td>
</tr>
<tr>
<td>STK11</td>
<td>11/280</td>
<td>3.9%</td>
<td>0/30</td>
<td>0%</td>
<td>2/40</td>
<td>5%</td>
<td>6/325</td>
<td>2%</td>
<td>2.9%</td>
</tr>
<tr>
<td>PTEN</td>
<td>14/280</td>
<td>5.0%</td>
<td>0/24</td>
<td>0%</td>
<td>2/40</td>
<td>5%</td>
<td>22/325</td>
<td>7%</td>
<td>5.9%</td>
</tr>
<tr>
<td>KRAS</td>
<td>2/280</td>
<td>0.7%</td>
<td>0/24</td>
<td>0%</td>
<td>1/40</td>
<td>3%</td>
<td>12/325</td>
<td>4%</td>
<td>2.3%</td>
</tr>
<tr>
<td>HRAS</td>
<td>2/280</td>
<td>0%</td>
<td>0/24</td>
<td>0%</td>
<td>0/40</td>
<td>0%</td>
<td>5/325</td>
<td>2%</td>
<td>1.1%</td>
</tr>
<tr>
<td>CDKN2A</td>
<td>0/280</td>
<td>0%</td>
<td>1/24</td>
<td>3%</td>
<td>1/40</td>
<td>32%</td>
<td>0/325</td>
<td>0%</td>
<td>0.2%</td>
</tr>
<tr>
<td>CTNNB</td>
<td>0/280</td>
<td>0%</td>
<td>1/24</td>
<td>3%</td>
<td>1/40</td>
<td>3%</td>
<td>0/325</td>
<td>0%</td>
<td>0.2%</td>
</tr>
<tr>
<td>PIK3CA + PTEN</td>
<td>105/280</td>
<td>38%</td>
<td></td>
<td></td>
<td>13/40</td>
<td>33%</td>
<td>108/325</td>
<td>33%</td>
<td>33%</td>
</tr>
<tr>
<td>All PI3K</td>
<td>111/280</td>
<td>40%</td>
<td></td>
<td></td>
<td>15/40</td>
<td>38%</td>
<td>104/325</td>
<td>32%</td>
<td>36%</td>
</tr>
</tbody>
</table>

NOTE: Mutations in targeted genes are shown for each country along with the percentage and sum of PIK3CA and PTEN (PIK3CA + PTEN) and PIK3CA + PTEN and STK11 (all PI3K). Total excludes Venezuela CIN.
HPV type and PIK3CA gene expression

It has been shown that the HPV E6 and E7 proteins can activate the PI3K/AKT pathway (34). We determined the mRNA expression levels of the PIK3CA and HPV E6 and E7 transcripts in 65 HPV16+ tumors. The mRNA levels for PIK3CA were higher in PIK3CA mutation–positive tumors than in PIK3CA WT tumors (P = 0.029), and HPV E6 and E7 expression was also elevated in PIK3CA-mutant tumors (P = 0.040, Supplementary Fig. S10A and S10B).

Phosphorylation of AKT by specific PIK3CA mutations

To determine whether specific PIK3CA mutations lead to increased AKT phosphorylation in vitro, U2OS cells were transfected with mutant constructs and used for western blot analysis. Total PIK3CA (p110α), AKT, β-actin, and p-AKT (Ser473) were constant. However, the kinase domain mutation H1047R led to increased p-AKT at Thr308; however, neither E542K, E545K nor the double mutant significantly increased p-AKT at Thr308 levels (Fig. 4).

Discussion

Our study combines exome and targeted sequencing to examine the relationship between genetic mutations and cervical cancer in three countries with high incidence and mortality. We identified frequent PIK3CA gene mutations in Latin American cervical tumors that had a distinct distribution from those found in most other cancers (25, 30). Latin American tumors have a similar PIK3CA mutation frequency as U.S. tumors (33% vs. 31%) and display a somewhat higher prevalence in squamous carcinomas (Supplementary Table S5). However, the PIK3CA mutation frequency is significantly lower in Swedish tumors (8.2%; $\chi^2 = 35, 2.89 \times 10^{-5}$; refs. 14, 18). Most PIK3CA mutations are located in the ABD and ABD–RBD linker and helical domain, especially in the Arg532 and Arg535 residues in the helical domain, particularly in squamous carcinomas. Recent structural studies of the PIK3CA protein (p110α) indicate that the helical domain of PIK3CA acts as a scaffold for the assembly of the kinase domain. In contrast, we found only 6 mutations within the kinase domain in cervical cancer. The most common PIK3CA mutation in all tumors, H1047R, which has been associated with an increased response rate to PI3K/AKT/mTOR inhibitors (39), is
very rare in our study, consistent with U.S. tumors (14). The distribution of mutations in PIK3CA/p110α suggests that there is a selective advantage to disrupting the ABF and ABF-RBD linker interactions in cervical cancers. Interestingly, a similar pattern was recently identified in bladder and HPV+ oral cancer with more helical domain mutations than kinase domain mutations, and a role for APOBEC has been proposed (38, 40–42).

Our study has several limitations: (i) the small size of the mutation discovery cohort (24 tumors of different histology and HPV type) would not allow the discovery of all frequently mutated genes or copy number changes, and without paired-end whole-genome sequence, we cannot fully evaluate the high rearrangement tumors for chromothripsis or chromoplexy (43); (ii) we lack a cancer-free cohort to explore risk factors for invasive cancer (apart from HPV infection); (iii) our data on tumor mutation heterogeneity lack a direct assessment of tumor purity and would require a detailed microdissection, immunohistochemistry, and/or single-cell sequencing component to fully explore; and (iv) to date, we have limited treatment outcome data to understand how molecular events may predict survival. However, we feel that data on tumors from high prevalence countries are important in understanding cervical cancer worldwide.

PI3K signaling has been shown to be important in HPV transformation (44). In this study, we show that PIK3CA mRNA expression is (41) elevated in tumors with PIK3CA mutations and associated with upregulated E7 mRNA expression at least in HPV16-positive tumors. The PIK3CA gene is on chromosome 3q in a region known to be duplicated in most invasive cervical cancers and in 19 of 22 of the tumors we report (29). Our data indicate that in approximately 20% of tumors with PIK3CA mutation (5% of all tumors), mutation is an early event and the 3q duplication involves the mutant PIK3CA allele, whereas in the remaining tumors, the mutation occurs on the nonduplicated allele or after duplication (Supplementary Fig. S11).

The presumed late occurring PIK3CA mutations could represent tumor evolution or represent the accumulation of somatic passenger mutations driven perhaps by HPV induction of APOBEC enzymes (42). Given the roles of PI3K in cellular proliferation, metastasis, cell cycle, and cell survival, PIK3CA mutations may impart a more aggressive and treatment-resistant phenotype (44, 45). However, it should be pointed out that our study is limited in having few early-stage lesions, no normal cervical tissue control, and exome and targeted sequencing can miss mutations, especially insertions/deletions (46). Sampling of multiple areas of the same tumor as well as primary and metastatic lesions from the same patients could address this issue.

While kinase domain mutations in PIK3CA have been shown to result in elevated AKT phosphorylation and mTOR activation, consistent with other reports, helical domain mutations do not (47). Several groups have documented that the AKT-related kinase SGK3 is activated by PIK3CA helical domain mutations (15, 16). Interestingly, in a clinical trial of AKT inhibitors, cervical and other cancers patients with the E542K mutation had a shorter median progression-free survival than those with the H1047R mutation (14, 39), suggesting that this subtype may be more refractory to therapy. In addition, a recent trial showed PIK3CA to be a marker of poor response (48). The data from our study and others would suggest that agents directly targeting PIK3C would have the best chance of benefit.

Our data are consistent with data (18) that cervical tumors have a relatively low level of driver gene mutations and group with glioblastoma, ovarian, breast, kidney, and acute myelogenous leukemia as tumors with an average of 2 to 4 driver gene mutations (28). Given that HPV expresses the 2 viral oncoproteins, E6 and E7, this is expected. However, RB1 and TP53 are mutated in a subset of tumors suggesting that further inactivation of these genes is sometimes required. In squamous tumors, these additional drivers are more frequently found in tumors in older women.

While HPV infection is the dominant risk factor in cervical cancer, tobacco use is a major co-factor in developed countries (49) and is variable within the subjects in this study. Interestingly, tobacco use is very low in Guatemala, but exposure to wood smoke is high (Table 1). Whether cooking with wood in the home contributes to cervical cancer similar to tobacco will require further study. We find that the distribution of HPV types is similar in the 3 populations studied, but that the PIK3CA and overall mutation load is higher in patients diagnosed after the age of 50 years and in HPV16+ patients (Supplementary Fig. S9, data not shown). HPV18 and HPV45 negative tumors are also more common in older patients (Supplementary Fig. S9B).

In summary, Latin American cervical tumors have a high frequency of mutations in the PIK3CA gene, especially at the E542 and E545 residues in the helical domain. PIK3CA kinase domain mutations more extensively phosphorylate AKT and are independent of RAS activity (47, 50). Furthermore, clinical trials of PI3K/AKT/mTOR inhibitors have shown that patients with the H1047R mutation in the PIK3CA kinase domain had better response (39). Therefore, these 2 classes of PIK3CA mutations, while both displaying activating/gain-of-function properties, denote functionally distinct classes of cancers. Our data add to the literature published to date demonstrating that although PIK3CA mutations are common in cervical cancer, they are not predicted to respond well to AKT/mTOR targeted agents. However, comprehensive surveys of other high prevalence countries are needed to fully understand the worldwide heterogeneity of this disease.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors’ Contributions
Conception and design: S. Polo, M. Dean
Writing, review, and/or revision of the manuscript: H. Lou, G. Villagran, S. Polo, W. Zhou, E. Roman-Basaure, J. Mitchell, J. Troyer, F.C. Pinto, S. Bass, X. Zhang, M. Yeager, J. Berumen, E. Álvarez, M. Dean
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): S. Polo, E. Juárez-Torres, I. Medina-Martínez, E. Roman-Basaure, D. Roberson, J. Sawitzke, L. Garland, J. Troyer, S. Bass, X. Zhang, H. Morales, J. Berumen, E. Álvarez, M. Dean
Study supervision: C. Villagran, S. Polo, F.C. Pinto, E. Gharzouzi, M. Dean
Other (figure construction for the mutation spectrum figure and wrote code to produce this visual): B. Gold
Activation of PIK3CA in Latin American Cervical Cancer

Grant Support

The work was supported in part by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research, and by Leidos Biomedical Research, Inc., under contract # HHSN261200800001E and the National University of Mexico [www.unam.mx], grant number SDI PTED 05.2 [to J. Benumem].

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received July 23, 2014; revised May 11, 2015; accepted May 14, 2015; published OnlineFirst June 16, 2015.

References


