IJC
International Journal of Cancer

Genomic aberrations in cervical adenocarcinomas in Hong Kong Chinese women

Tony K.H. Chung[1], Paul Van Hummelen[2], Paul K.S. Chan[3], Tak Hong Cheung[1], So Fan Yim[1], Mei Y. Yu[4], Matthew D. Ducar[2], Aaron R. Thorner[2], Laura E. MacConaill[2], Graeme Doran[5], Chandra Sekhar Pedamallu[6,7], Akinyemi I. Ojesina[6,7], Raymond R.Y. Wong[8], Vivian W. Wang[9], Samuel S. Freeman[6], Tat San Lau[1], Joseph Kwong[1], Loucia K.Y. Chan[1], Menachem Fromer[10], Taymaa May[11], Michael J. Worley Jr.[12], Katharine M. Esselen[12], Kevin M. Elias[12], Michael Lawrence[13], Gad Getz[13], David I. Smith[9], Christopher P. Crum[14], Matthew Meyerson[2,6,7], Ross S. Berkowitz[12] and Yick Fu Wong[1,12]

[1] Department of Obstetrics & Gynaecology, The Chinese University of Hong Kong, Hong Kong, People's Republic of China
[2] Center for Cancer Genome Discovery, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA
[3] Department of Microbiology, The Chinese University of Hong Kong, Hong Kong, People's Republic of China
[4] Department of Anatomical & Cellular Pathology, The Chinese University of Hong Kong, Hong Kong, People's Republic of China
[5] Department of Pathology, Harvard Medical School, Boston, MA
[6] Cancer Program, The Broad Institute of MIT and Harvard University, Cambridge, MA
[7] Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA
[8] Harvard Medical School, Pediatric Surgical Laboratories, Massachusetts General Hospital, Boston, MA
[9] Department of Pathology and Laboratory Medicine, Mayo Clinic, Rochester, MN
[10] Division of Psychiatric Genomics, Mount Sinai School of Medicine, New York, NY
[11] Division of Gynecologic Oncology, Princess Margaret Cancer Center, Toronto, ON, Canada
[12] Department of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA
[13] Cancer Genomics Informatics and Computational Biology, The Broad Institute of Harvard and MIT, Cambridge, MA
[14] Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA

Cancer Cell Biology

Although the rates of cervical squamous cell carcinoma have been declining, the rates of cervical adenocarcinoma are increasing in some countries. Outcomes for advanced cervical adenocarcinoma remain poor. Precision mapping of genetic alterations in cervical adenocarcinoma may enable better selection of therapies and deliver improved outcomes when combined with new sequencing diagnostics. We present whole-exome sequencing results from 15 cervical adenocarcinomas and paired normal samples from Hong Kong Chinese women. These data revealed a heterogeneous mutation spectrum and identified several frequently altered genes including *FAT1*, *ARID1A*, *ERBB2* and *PIK3CA*. Exome sequencing identified human papillomavirus (HPV) sequences in 13 tumors in which the HPV genome might have integrated into and hence disrupted the functions of certain exons, raising the possibility that HPV integration can alter pathways other than p53 and pRb. Together, these provisionary data suggest the potential for individualized therapies for cervical adenocarcinoma based on genomic information.

Key words: cervical adenocarcinoma, genomic alternations, HPV

Additional Supporting Information may be found in the online version of this article.

Author Contributions: Y.F.W., T.K.H.C., R.S.B., P.V.H., C.P.C., L.E.M. and M.M. conceived the study. Y.F.W., T.K.H.C., P.V.H. and R.S.B. directed the study. Y.F.W., T.K.H.C., P.V.H., P.K.S.C., R.S.B., M.M., K.M.El., D.I.S., R.R.Y.W., G.D., V.W.W., M.J.W., K.M.Es., T.M., A.I.O and C.S.P. wrote the manuscript with the assistance and final approval of all authors. A.R.T. performed experiments. M.D.D., C.S.P., A.I.O., S.S.F. and M.F. performed computational analysis. G.G. and M.L. guided statistical analysis. T.H.C., S.F.Y., T.S.L., J.K. and L.K.Y.C. contributed collection of samples. M.Y.Y. was involved in histological review of samples.

Conflict of interest: M.M. is a shareholder in and consultant for Foundation Medicine, Inc.

T.K.H.C., P.V.H. and P.K.S.C contributed equally to this work

Grant sponsor: The Department of Obstetrics & Gynaecology, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong
DOI: 10.1002/ijc.29456
History: Received 12 Aug 2014; Accepted 12 Jan 2015; Online 27 Jan 2015

Correspondence to: Yick Fu Wong or Ross S. Berkowitz, Department of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital, Harvard Medical School, 75 Francis St, Boston, MA 02115, USA. Tel.: +1-617-732-8840, Fax: +1-617-738-5124, E-mail: yfwong@rics.bwh.harvard.edu (or) Ross_berkowitz@dfci.harvard.edu

What's new?

Cervical cancer is almost always associated with infections with human papilloma virus (HPV) but additional genetic mutations are required for carcinogenesis. Here the authors performed whole-exome sequencing in Hong Kong women and identified several novel recurrently mutated genes in cervical adenocarcinomas that were not present in normal adjacent tissue. Notably, HPV sequences were detected in several tumors pointing to chromosomal integration of the virus. These studies may support the development of targeted molecular therapies of cervical cancer patients in East Asia and other parts of the world.

Cervical cancer is the fourth most common cancer in women worldwide with an estimated 528,000 new cases and 266,000 deaths in 2012.[1] In China alone, there were an estimated 62,000 new cases and 30,000 deaths in the same year.[1] Cervical cancer presents as one of the two morphologically distinct histologic entities, cervical squamous cell carcinoma or cervical adenocarcinoma. Although the relative rate of cervical squamous cell carcinoma has been steadily declining for several decades, the incidence of cervical adenocarcinoma appears to be increasing.[2] It remains controversial as to whether adenocarcinoma of the cervix is associated with a significantly worse prognosis than squamous cell carcinoma of the cervix. However, both histologic subtypes are associated with poor long-term survival in the setting of advanced stage disease.[3] Cervical cancer is almost always associated with infection from oncogenic types of human papillomavirus (HPV). HPV18 is more preferentially found in adenocarcinoma, whereas HPV16 is associated with both squamous cell carcinoma and adenocarcinoma.[4] However, HPV infection alone is insufficient for malignant transformation; other genetic events independent or in conjunction with HPV infection are required.[5] We set on to explore the genetic alterations underlying cervical adenocarcinoma by performing a whole-exome sequencing study on tumor/normal pairs from Hong Kong Chinese women.

Material and Methods

Clinical material

Snap-frozen primary tumor tissues and matched blood samples from 15 Hong Kong Chinese patients with cervical adenocarcinoma prior to any treatment were obtained from the Biosample Repository at the Department of Obstetrics & Gynaecology, The Chinese University of Hong Kong, Hong Kong. All patients provided informed consent. Cryomolds prepared from tumor specimens were laser-capture microdissected to reach tumor cell purity to 90% or more. A pathologist reviewed hematoxylin–eosin-stained sections of banked tumor tissues to verify that they were representative of the original histological classification and to locate tumor cells for microdissection. Genomic DNA isolated from peripheral blood was collected from each patient as a corresponding control. This was a retrospective study approved by the The Joint Chinese University of Hong Kong—New Territories East Cluster Clinical Research Ethics Committee.

DNA extraction and identity testing

Genomic DNA was isolated from microdissected tumor cells and matched blood cells using the Qiagen DNeasy Tissue and Blood Kit (Valencia, CA) according to the manufacturer's instructions. High-molecular-weight DNAs were eluted in TE buffer. To ensure that the tumor specimens and the corresponding blood samples were collected from the same individual, a finger printing analysis was performed to confirm the sample identity. For this analysis, the sequence of 48 SNPs from each sample was compared to each other sample using the R-package SPIA.[6] Finger printings from all 15 tumor/normal pairs were matched.

Whole-exome sequencing

Prior to library preparation, DNA was fragmented by sonication (Covaris, Woburn, MA) to 150 bp and further purified using Agencourt AMPure XP beads. In brief, 50 ng of size-selected DNA was then ligated to specific adaptors during library preparation (Illumina TruSeq, Illumina, San Diego, CA). Each library was made with sample-specific barcodes, quantified by qPCR (Kapa Biosystems, Woburn, MA) and two libraries were pooled to a total of 500 ng for exome enrichment using the Agilent SureSelect hybrid capture kit (Whole Exome_v2, 44 Mb; Agilent Technologies, Santa Clara, CA). Several captures were pooled further and sequenced in one or more lanes to a final equivalent of two exomes per lane on a Hiseq 2000 (Illumina, San Diego, CA). Sequence alignment, demultiplexing and variant calling, including non-synonymous somatic single nucleotide variants (SSNVs) and insertions/deletions (InDels), were performed using PICARD, GATK tools, MuTect and IndeLocator as described previously.[7] Significantly mutated genes were identified using MutSigCV and genes with false discovery rates (FDRs, q-values) of <0.1 were reported as significantly mutated.[8]

Validation of candidate mutations by Sequenom genotyping

We validated a subset of mutations by mass spectrometric genotyping based on the Sequenom MassARRAY® technology (Sequenom, San Diego, CA) using a multibase homogeneous Mass-Extend (hME) as described previously.[9]

Copy-number analysis

Copy-number analysis was performed using Nexus7.0 (BioDiscovery, Hawthorne, CA) after calculating read counts using the ngCGH tool. Somatic copy-number alteration (SCNA) was called with the following NGS settings: significance threshold = 1E − 9; max contiguous probe spacing or 1,000 Kbp; minimum number of probes per segment = 3.

SCNA gains had a log ration of >0.4 and were called high gain if >1.1. Single copy loss threshold was −0.6 and big loss was −1. GISTIC analysis was used to determine statically significant recurrent SCNA using the following parameters: lesion amplitude threshold =4 and *q*-value FDR was set at 0.1. Only SCNA regions that were larger than 10% of the chromosome and <99% overlapping with common SCNA variants based on DGV database (http://dgv.tcag.ca/dgv/app/home) were retained.[10]

HPV infection status

The Linear Array HPV Genotyping Test (Roche Molecular Systems, Pleasanton, CA) was used to detect and type HPV present in tissue samples. The exome sequencing data were also analyzed for the presence of HPV sequences integrated into exons (Table 1). Putative HPV integration sites were identified using the PathSeq algorithm, which performs computational subtraction of human reads, followed by alignment of residual reads to a combined database of the human and HPV reference genomes (including >150 HPV types) obtained from the NCBI sequence database.[11] This phenomenon results in the identification of reads mapping with high confidence to HPV genomes. Chimeric human and HPV read pairs were identified by extracting the pair mates of HPV reads and aligning the paired end reads to the combined human and HPV reference genome database, using BWA (http://biowulf.nih.gov/apps/bwa.html). The chimeric read pairs, in which one read maps to the human genome and the mate maps to the HPV genome, represent integration sites.

Results

Tumor/normal pairs from 15 Hong Kong Chinese women between 34 and 76 years of age at the time of diagnosis with early stage cervical adenocarcinomas (seven Stage I and eight Stage II according to the International Federation of Gynecology and Obstetrics [FIGO] staging system) were included in our study. All tumor samples were positive for one or more HPV types by the Linear Array HPV Genotyping Test (Roche Molecular Systems, Pleasanton, CA) (Table 1).

Whole-exome sequencing was performed to an average target coverage of 182× and 88% of the targets had a minimum coverage of 30×. Using each case's corresponding peripheral blood control sample to filter out germline events and sequencing artefacts using MuTect and Indelocator,[12,13] a total of 2,900 nonsynonymous somatic mutations (median of 70 mutations per sample; range, 37–1,260) were identified across the 15 tumors (Supporting Information Fig. S1). There were 2,558 nonsynonymous SSNVs including 2,346 missense mutations, 172 nonsense mutations and 40 splice mutations in 2,182 genes as well as 342 InDels in 333 genes in this set of tumors. A summary of somatic alterations is shown in Figure 1 (and compared to cervical cancer data on the cBio portal[14] shown in Supporting Information Fig. S2). As a validation, we analyzed a subset of 18 mutations by mass spectrometric genotyping in all patients' tumor and blood samples. All variants were confirmed except for *PIK3CA* p.E545K that was detected at 6% allele fraction by sequencing (Table 2).

Table 1. Clinical stage and HPV infection status of 15 cervical adenocarcinomas

Tumor sample code	Age (years)	Clinical stage[1]	HPV genotyping[2]	Integrated HPVs at exons[3]
S645	76	IIa	HPV18	HPV 18 positive
S652	54	Ib	HPV16,18	HPV negative
S658	34	Ib	HPV16	HPV 16 positive
S671	52	Ib	HPV16	HPV negative
S755	45	Ib	HPV16,18	HPV 18 positive
S770	55	IIa	HPV18	HPV 18 positive
S823	73	IIa	HPV16	HPV 16 positive
S835	50	IIb	HPV18	HPV 18 positive
S848	54	IIb	HPV16	HPV 16 positive
S867	71	IIb	HPV16	HPV 16 positive
S883	66	IIb	HPV58	HPV 58 positive
S928	56	IIb	HPV45	HPV 45 positive
S961	51	Ib	HPV16,58	HPV 58 positive
S1004	55	Ib	HPV16	HPV 16 positive
S1009	44	Ib	HPV16	HPV 16 positive

[1]According to the International Federation of Gynecology and Obstetrics (FIGO) staging system.
[2]HPV type determined by Linear Array HPV Genotyping Test (Roche Molecular Systems, Pleasanton, CA).
[3]HPV positive: HPV sequence present in exome sequence data; HPV negative: no HPV sequence identified from exome sequence data.

Because of the limited numbers of samples available for our study, we took two analytic approaches to analyze the data for significantly mutated genes: a simple enumeration as well as a statistical approach, MutSigCV. There were 23 nonsynonymous mutations where each was observed in at least 3 out of 15 cases (Supporting Information Table S1). Three tumors contained frame-shift deletions in *ARID1A* (AT-rich interactive domain 1A). Mutations in *PIK3CA* were also detected in three tumors at commonly observed nucleotide positions, E453K, E545K and H1047R, respectively.[15] Except for the E545K and H1047R substitutions in *PIK3CA*, the nonsynonymous mutations recurrently mutated in the other 22 genes have not yet been reported in cervical adenocarcinoma in the COSMIC database_v71 (http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/). However, most of these mutations have been previously reported for other tumor types in COSMIC and are predicted by PolyPhen2[16] analyses to have functional significance, either "probably damaging" or "possibly damaging," based on false-positive cut-off rates of 5 and 10%, respectively (Table 2). In addition, two tumors harbored activating mutations in S310F and S310Y in the extracellular domain of ERBB2 that may predict sensitivity to anti-Her2-targeted therapy.[17]

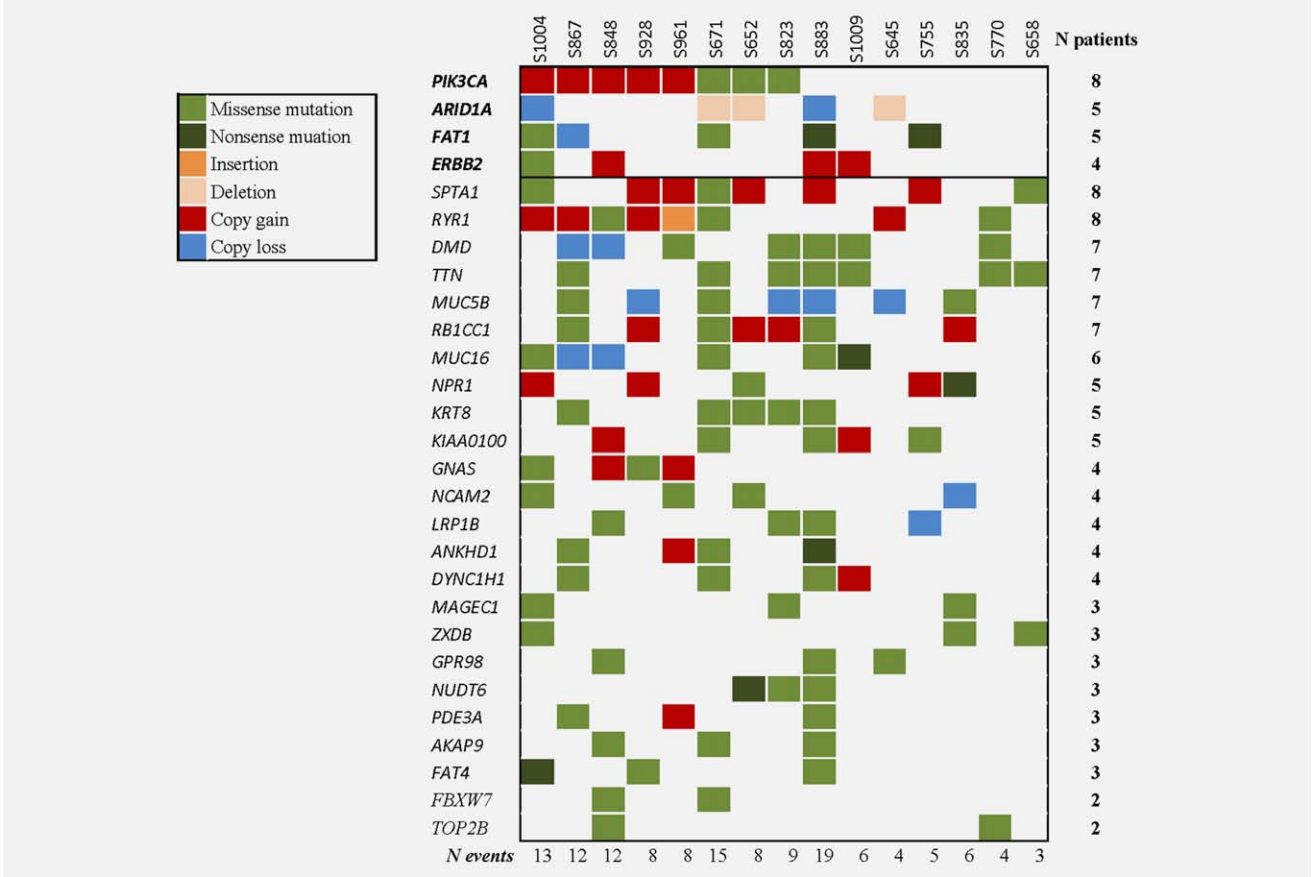

Figure 1. Summary of somatic alterations in genes that were recurrently mutated and that had recurrent copy-number gains or losses. Genomic alteration subtypes are denoted by color. If multiple mutations were found in a gene in a single sample, only one is shown in the order as listed in the color legend from top to bottom. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Statistical analysis of mutation significance using MutSigCV showed that the most significantly mutated gene in the 15 whole-exome sequenced samples is *FAT1*, with significant recurrence at a FDR of <0.1 (Table 3). Of the four tumors that harbored mutations of *FAT1*, two harbored missense and two nonsense mutations (Table 3). Other genes with mutation significance of $q < 0.25$ included the guanylate cyclase gene *NPR1*, the phosphodiesterase gene *PDE3A*, the well-known tumor suppressor gene *FBXW7*, and the topoisomerase gene *TOP2B* (Table 3).

The exome sequencing data were also analyzed for copy-number changes. This revealed a median of 72 SCNA regions per sample (range, 2–556) encompassing a median of 3,151 genes per tumor with 12 statistically significant recurrent SCNA regions (data not shown). Significant recurrent gains were identified in chromosomes 1q, 3q, 8q, 11p, 17q, 19q and 20q, whereas recurrent deletions were found in chromosomes 11 and 16 (Table 4). Similar to the previous findings by us in a cohort of predominantly squamous cervical cancer samples,[18] three tumors showed an amplification of *ERBB2* on chromosome 17q, which encodes the ErbB2 receptor tyrosine kinase protein, with amplification to >16 copies in sample S848 (Fig. 2), lower level amplifications in sample S883 that contained the S310F mutation and in sample S1009. In addition, six tumors had copy gains of *PIK3CA*, whereas two tumors had a copy loss of *ARID1A*, with one of these tumors harboring both a frame-shift deletion and a copy loss. Furthermore, owing to paucity of copy-number variations (CNVs) in the *ARID1A* and *FAT1* genes in the cervical cancer data currently available in cBio (Supporting Information Fig. S2), we therefore interrogated the data set produced by Ojesina *et al.*[18] for relationships between copy-number changes and gene expression in the highlighted genes. There was no clear correlation between copy number and gene expression for these genes, in part because they are so highly expressed in most of the tumors (Supporting Information Fig. S3).

Interestingly, the exome sequencing identified HPV sequences from 13 out of the 15 adenocarcinomas (Table 1) in which the HPV genome might have integrated into and disrupted the continuity, and hence the functions, of certain exons. Nine samples identified as positive for HPV DNA were interrogated by whole-exome sequencing and typing results were matched with those from the LINEAR ARRAY® HPV Genotyping Test. Furthermore, we found that two

Table 2. Mutations and their allele fraction detected by whole-exome sequencing and validated by hME[1]

Subject ID	Gene	Protein change	Genome position (hg19)	Base change	Tumor fraction	Sequenom genotyping	Predicted functional significance by PolyPhen2.0	Present in COSMIC
S867	*ALK*	p.D466N	chr2:29551234	C>T	0.22	Validated	Probably Damaging	No
S671	*CLTC*	p.V804A	chr17:57751126	T>C	0.27	Validated	Possibly damaging	No
S671	*CTNNB1*	p.S33A	chr3:41266100	T>G	0.28	Validated	Probably damaging	Yes
S652	*CTNNB1*	p.S37F	chr3:41266113	C>T	0.45	Validated	Probably damaging	Yes
S883	*ERBB2*	p.S310F	chr17:37868208	C>T	0.56	Validated	Probably damaging	Yes
S1004	*ERBB2*	p.S310Y	chr17:37868208	C>A	0.42	Validated	Probably damaging	Yes
S848	*FBXW7*	p.R479P	chr4:153247366	C>G	0.36	Validated	Probably damaging	Yes
S671	*FBXW7*	p.R505C	chr4:153247289	G>A	0.23	Validated	Probably damaging	Yes
S671	*PIK3CA*	p.E453K	chr3:178928079	G>A	0.21	Validated	Possibly damaging	Yes
S652	*PIK3CA*	p.H1047R	chr3:178952085	A>G	0.37	Validated	Possibly damaging	Yes
S823	*PIK3CA*	p.E545K	chr3:178936091	G>A	0.06	Not validated	Probably damaging	Yes
S652	*PTEN*	p.R130G	chr10:89692904	C>G	0.77	Validated	Probably damaging	Yes
S671	*PTEN*	p.R233*	chr10:89717672	C>T	0.37	Validated	–	Yes
S867	*TP53*	p.Q167*	chr17:7578431	G>A	0.60	Validated	–	Yes
S835	*TP53*	p.R248Q	chr17:7577538	C>T	0.46	Validated	Probably damaging	Yes
S645	*ARID1A*	p.T1360fs	chr1:27100366	AC>A	0.94	Validated	–	No
S652	*ARID1A*	p.G1515fs	chr1:27101262	GCTCTGCCCCC>G	0.60	Validated	–	No
S671	*ARID1A*	p.Y1324fs	chr1:27100175	AC>A	0.35	Validated	–	Yes

[1]All variants were analyzed in all 30 samples and only those observed by sequencing were validated except for *PIK3CA* p.E545K. The functional significance of nonsense and frameshift mutations was not determined. However, these mutations will result in a change in protein structure and are traditionally believed to be loss-of-function in nature.

tumors had chimeric read pairs in which one read mapped to the HPV genome, whereas its pairmate mapped to the human genome (highly suggestive of an HPV integration event). The chimeric read pairs involved HPV16 and the *STARD3* gene in tumor S848 (Fig. 2), and HPV18 and *SERPINB1* in tumor S835. Sample S848 also had the highest *ERBB2* amplification.

Discussion

At present, data on the spectrum of genomic alterations in cervical adenocarcinoma are limited. In our study, we observed that although there were many heterogeneous mutations detected in the data set of 15 cervical adenocarcinomas, very few of these mutations were observed in multiple tumors. *ARID1A* (also known as *BAF250*) encodes a key member of the SWI/SNF chromatin-modeling complex and it has been reported to be frequently mutated in a wide variety of cancer types including ovarian, bladder, gastric, breast and endometrial.[19] *ARID1A* has been reported to regulate the PI3K pathway.[20] *PIK3CA*, a member of the PI3K/AKT/mTOR pathway, is the most frequently mutated oncogene in human cancers.[20,21] It also appeared as a frequently altered gene in this set of tumors with SSNVs detected in three tumors and copy gains in six tumors. These findings are in keeping with recent reports of *PIK3CA* point mutations in 10 out of 40 (25%) cases and 2 out of 24 (8%) cases of cervical adenocarcinoma in women from the USA, and in Norwegian and Mexican women, respectively.[18,22] In contrast, the recently reported mutations in *ELF3* and *CBFB*[18] were not identified in our study, possibly because of the low sample size. *PIK3CA* mutations can cause neoplastic transformation and promote cancer progression; they have been found to predict response to PI3K/AKT/mTOR axis inhibitors in patients with advanced cancers. The response rate is significantly higher for patients with *PIK3CA* mutations treated with PI3K/AKT/mTOR pathway inhibitors than for those without documented mutations.[23] It was shown in an *in vitro* study that the activated PI3K/Akt/COX-2 pathway might induce resistance to radiation in human cervical adenocarcinoma HeLa cells.[24] Understanding the function of *PIK3CA*, confirming whether it is an effective target of molecular targeted therapy, and exploring the mutation of *PIK3CA* as a predictive biomarker in both gene therapy and radiation therapy in cervical adenocarcinoma will be important areas for future work.[25]

Four tumors harbored missense or nonsense mutations and one additional tumor had a partial deletion in *FAT1*, which is located at 4q35 and encodes a cadherin-like protein that is able to suppress cancer cell growth *in vitro* and *in vivo* by binding β-catenin and antagonizing its nuclear localization. Inactivation of *FAT1 via* mutation has been demonstrated to promote Wnt signaling and tumorigenesis.[26] Recurrent somatic mutations of *FAT1* have been identified in

Table 3. Significantly mutated genes in 15 cervical adenocarcinomas using MutSig CV analysis

Gene	codelen	nncd	nsyn	nmis	nnon	nspl	*p*–value	*q*-value
FAT1	13,871	0	0	2	2	0	3.52e − 06	0.06
NPR1	3,270	0	0	1	1	0	1.91e − 05	0.17
PDE3A	3,488	0	0	2	0	0	5.52e − 05	0.24
FBXW7	2,580	0	0	2	0	0	5.93e − 05	0.24
TOP2B	5,006	0	0	2	0	0	6.65e − 05	0.24
NRAP	5,359	0	0	2	0	0	0.00015	0.42
STXBP5	3,564	0	0	1	1	0	0.00018	0.42
CGNL1	3,981	0	0	1	1	0	0.00023	0.42
NBR1	2,981	0	0	2	0	0	0.00025	0.42
BRCA2	10,361	0	1	1	0	1	0.00029	0.42
STAT5A	2,457	0	0	2	0	0	0.00029	0.42
DAAM1	3,337	0	0	2	0	0	0.00029	0.42
FOXK2	2,015	0	0	2	0	0	0.00033	0.42
MUC5B	17,492	0	0	3	0	1	0.00036	0.42
TRPA1	3,464	0	0	1	1	0	0.00036	0.42
ASH1L	9,003	0	0	3	0	0	0.00037	0.42
GPR64	3,161	0	0	1	1	0	0.00041	0.44
METTL3	1,787	0	0	2	0	0	0.00046	0.44
NIPBL	8,642	0	0	1	1	0	0.00048	0.44
ACACA	8,227	0	0	1	1	0	0.00049	0.44
ALG13	3,662	0	0	0	1	1	0.00050	0.44
AP1B1	2,938	0	0	1	1	0	0.00069	0.57
LRP1	13,987	1	0	1	0	1	0.00076	0.57
CADPS	4,251	0	0	1	1	0	0.00077	0.57
PIK3CA	3,287	0	0	3	0	0	0.00079	0.57
MICAL1	3,300	1	1	2	0	0	0.00081	0.57
YEATS2	4,389	0	0	1	0	1	0.00084	0.57
CCBL1	1,317	0	0	2	0	0	0.00092	0.60
RB1CC1	4,873	0	1	3	0	0	0.0010	0.65
TICAM1	2,139	0	0	2	0	0	0.0011	0.67

Abbreviations: codelen: number of coding nucleotides covered by sequencing; nncd: number of noncoding mutations; nsyn: number of synonymous mutations; nmis: number of missense mutations; nnon: number of nonsense mutations; nspl: number of splice site mutations; *q*-value: FDR.

a fraction of glioblastomas, colorectal cancers and head and neck cancers, but not in cervical adenocarcinoma to date.[26]

We interrogated the cBio Portal to determine the frequencies of all the genes shown in Figure 1 with respect to the provisional data set produced by the TCGA project on cervical cancers. These data (Supporting Information Fig. S2) show that genomic alterations in *PIK3CA*, *ERBB2*, *FAT1* and *ARID1A* are present in 20/36 (56%) of cervical tumors in the cBio portal. In addition, the cBio analysis revealed that 31/36 (86%) tumors in the provisional TCGA cervical set had alterations in the total set of genes highlighted in our manuscript (Supporting Information Fig. S2). Furthermore, the relative frequencies of alterations by gene are similar across both data sets.

The lack of correlation between copy number and gene expression for these genes does not preclude functional significance for these alterations because (as it is also shown in Supporting Information Fig. S3) the well-known tumor suppressor gene *TP53* exhibits a similar lack of correlation between copy number and gene expression data.

One well-documented oncogenic pathway mediated by integration of HPV into human genome is the disruption of viral gene E2, which leads to a loss of control over the expression of the viral oncogenes E6 and E7, resulting in interruption of the normal tumor suppressing role of p53 and pRb.[27] In our study, we detected HPV sequences in the exome data set in 13 out of 15 tumors analyzed, including two tumors with chimeric HPV–human read pairs. This

Table 4. Recurrent SCNAs as detected by GISTIC across 15 cervical adenocarcinomas[1]

Region	Extended region	Type	Q-bound	G-Score	% CNV overlap	Chromosome fraction of extended region	Samples with multiple copy variants
chr1:152,127,775-152,128,753	chr1:152,127,126-249,250,621	Gain	0.07	5.27	0	0.39	652,755,835,883, 961,928,1004
chr3:168,834,334-169,485,851	chr3:105,238,768-198,022,430	Gain	0.07	5.45	0	0.47	823,848,928,961,1004
chr8:130,742,503-131,249,098	chr8:47,012,755-146,364,022	Gain	0.07	4.24	0	0.68	652,835,928,961
chr11:34,250,664-35,198,065	chr11:0-42,703,258	Gain	0.07	4.09	0	0.32	823,848
chr17:37,815,234-38,024,545	chr17:25,376,007-81,195,210	Gain	0.06	6.33	8	0.69	848,883,1009
chr19:44,535,917-44,661,061	chr19:27,888,389-59,128,983	Gain	0.07	4.93	0	0.53	823,848,867,928,1004
chr20:30,686,771-31,138,993	chr20:29,847,524-63,025,520	Gain	0.07	4.74	0	0.56	823,848,883,961,1004
chr11:47,744,758-47,840,863	chr11:35,198,065-50,379,652	Loss	0.05	4.32	0	0.11	645,848
chr16:81,145,770-81,272,651	chr16:46,505,069-90,354,753	Loss	0.05	4.50	2	0.49	645,658,835,1004

[1]Only statistically significant SCNA regions that were >10% of the chromosome and <99% overlapping with common SCNA variants based on DGV database are shown.

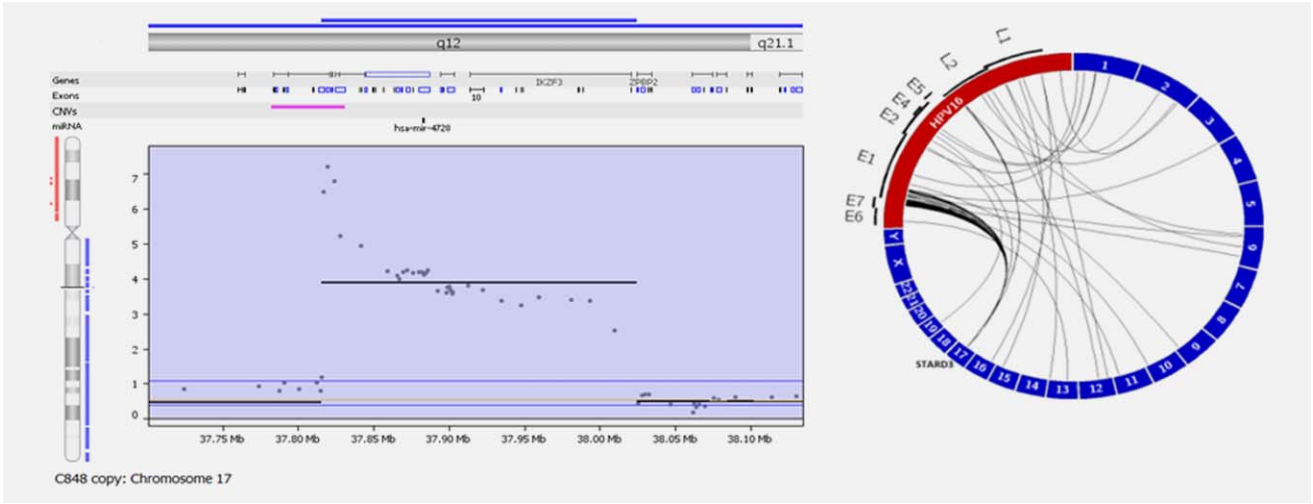

Figure 2. Amplification of region containing *ERBB2* and *STARD3*. Sample S848 had a large amplification of >16 copies and an integration site of HPV in *STARD3* (circus plot on right). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

suggests, in addition to the well-known E6-p53 and E7-pRb pathways, HPV may mediate oncogenic progression by direct impairment of host gene expression as a result of host gene disruption by viral integration. This hypothesis is in line with recent findings that HPV integration is associated with high expression of target genes.[18] In our study, we found two gene transcripts (*SERPINB1* and *STARD3*) were involved as chimeric pairs with HPV and thus their encoding functions may be impaired. *SERPINB1* is a member of the serpin family of protease inhibitors which maintain homeostasis by neutralizing overexpressed protease activity.[28] *SERPINB1* plays a key role in protecting tissues at inflammatory sites, whereas its role in carcinogenesis remains to be established. *STARD3* is a membrane protein involved in cholesterol transport.[29] Interestingly, the highest level of *ERBB2* was detected in the sample with chimeric *STARD3*–HPV reads. Although we could not verify the link between HPV integration and *ERBB2* amplification in this case, it is worth noting that *STARD3* is colocated with *ERBB2* on the genome and is often coamplified together in breast cancer.[30] *ERBB2* amplification can result in overexpression of its encoded protein ErbB2. This protein has an intracellular tyrosine kinase domain that regulates several pathways associated with cell survival and proliferation. It has been shown that HPV16 E6 can regulate ErbB2.[31] There have been several studies linking the amplification of the *ERBB2* with chemoresistance and poor overall survival. Tyrosine kinase inhibitors and immunotherapy with monoclonal antibodies targeting ErbB2 hold promise for patients harboring these aggressive neoplasms.[32] Furthermore, amplification of *ERBB2* was frequently detected in cervical

adenocarcinoma in Hong Kong Chinese women in our study. Similar amplification was also found in 2 out of 24 cervical adenocarcinomas from Norwegian and Mexican women.[18] Overall, we demonstrated the potential of using exome sequencing to reveal alternative oncogenic pathways mediated by viral integration.

Conclusions

In conclusion, we have identified a set of candidate genes that were recurrently altered by mutation and/or copy-number changes in cervical adenocarcinoma, most notably *FAT1*, *ARID1A*, *PIK3CA* and *ERBB2*. The complexity and heterogeneity of the protein-altering somatic mutational spectrum across the 15 tumors was impressive. As an increasing number of studies point to the importance of tumor context and the unique mutation profile of a tumor to predict specific drug response,[33,34] further characterization of the genomes of cervical adenocarcinomas is warranted and will enable more robust and biologically meaningful classification of these tumors.

References

1. GLOBOCAN 2012. Cervical Cancer Incidence, Mortality and Prevalence Worldwide in 2012. International Agency for Research on Cancer, WHO. Available at: http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx. Accessed May 22, 2014.
2. Adegoke O, Kulasingam S, Virnig B. Cervical cancer trends in the united states: a 35-year population-based analysis. *J Womens Health* 2012;21:1031–7.
3. Steren A, Nguyen HN, Averette HE, et al. Radical hysterectomy for stage IB adenocarcinoma of the cervix: The University of Miami experience. *Gynecol Oncol* 1993;48:355–9.
4. Bulk S, Berkhof J, Bulkmans NW, et al. Preferential risk of HPV16 for squamous cell carcinoma and of HPV18 for adenocarcinoma of the cervix compared to women with normal cytology in the Netherlands. *Br J Cancer* 2006;94:171–5.
5. Wang SS, Hildesheim A. Chapter 5: viral and host factors in human papillomavirus persistence and progression. *J Natl Cancer Inst Monogr* 2003;31:35–40.
6. Demichelis F, Greulich H, Macoska JA, et al. SNP panel identification assay (SPIA): a genetic-based assay for the identification of cell lines. *Nucleic Acids Res* 2008;36:2446–56.
7. Hettmer S, Teot LA, van Hummelen P, et al. Mutations in hedgehog pathway genes in fetal rhabdomyomas. *J Pathol* 2013;231:44–52.
8. Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 2013;499:214–8.
9. Lee J, van Hummelen P, Go C, et al. High-throughput mutation profiling identifies frequent somatic mutations in advanced gastric adenocarcinoma. *PLoS One* 2012;7:e38892
10. Mermel CH, Schumacher SE, Hill B, et al. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol* 2011;12:R41
11. Kostic AD, Ojesina AI, Pedamallu CS, et al. PathSeq: software to identify or discover microbes by deep sequencing of human tissue. *Nat Biotechnol* 2011;29:393–6.
12. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* 2013;31:213–9.
13. Banerji S, Cibulskis K, Rangel-Escareno C, et al. Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature* 2012;486:405–9.
14. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 2013;6:pl1.
15. Burke JE, Perisic O, Masson GR, et al. Oncogenic mutations mimic and enhance dynamic events in the natural activation of phosphoinositide 3-kinase p110α (PIK3CA). *Proc Natl Acad Sci USA* 2012; 109:15259–64.
16. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248–9.
17. Herter-Sprie GS, Greulich H, Wong KK. Activating mutations in ERBB2 and their impact on diagnostics and treatment. Front Oncol 2013;3:86.
18. Ojesina AI, Lichtenstein L, Freeman SS, et al. Landscape of genomic alterations in cervical carcinomas. *Nature* 2014;506:371–5.
19. Liang H, Cheung LW, Li J, et al. Whole-exome sequencing combined with functional genomics reveals novel candidate driver cancer genes in endometrial cancer. *Genome Res* 2012;22:2120–9.
20. Rudd ML, Price JC, Fogoros S, et al. A unique spectrum of somatic PIK3CA (p110alpha) mutations within primary endometrial carcinomas. *Clin Cancer Res* 2011;17:1331–40.
21. German S, Aslam HM, Saleem S, et al. Carcinogenesis of PIK3CA. *Hered Cancer Clin Pract* 2013;11:5
22. Wright AA, Howitt BE, Myers AP, et al. Oncogenic mutations in cervical cancer: genomic differences between adenocarcinomas and squamous cell carcinomas of the cervix. *Cancer* 2013;119:3776–83.
23. Janku F, Tsimberidou AM, Garrido-Laguna I, et al. PIK3CA mutations in patients with advanced cancers treated with PI3K/AKT/mTOR axis inhibitors. *Mol Cancer Ther* 2011;10:558–65.
24. Xia S, Zhao Y, Yu S, et al. Activated PI3K/akt/COX-2 pathway induces resistance to radiation in human cervical cancer HeLa cells. *Cancer Biother Radiopharm* 2010;25:317–23.
25. Mohseni M, Park BH. PIK3CA and KRAS mutations predict for response to everolimus therapy: now that's RAD001. *J Clin Invest* 2010;120:2655–8.
26. Morris LGT, Kaufman AM, Gong Y, et al. Recurrent somatic mutation of FAT1 in multiple human cancers leads to aberrant Wnt activation. *Nat Genet* 2013;45:253–61.
27. Doorbar J. Molecular biology of human papillomavirus infection and cervical cancer. *Clin Sci* 2006;110:525–41.
28. Zhao P, Hou L, Farley K, et al. SerpinB1 regulates homeostatic expansion of IL-17+ γδ and CD4+ Th17 cells. J Leukoc Biol 2014;95:521–30.
29. Reitz J, Gehrig-Burger K, Strauss JF 3rd, et al. Cholesterol interaction with the related steroidogenic acute regulatory lipid-transfer (START) domains of StAR (STARD1) and MLN64 (STARD3). FEBS J 2008;275:1790–802.
30. Jacot W, Fiche M, Zaman K, et al. The HER2 amplicon in breast cancer: topoisomerase IIA and beyond. *Biochim Biophys Acta* 2013;1836:146–57.
31. Narisawa-Saito M, Handa K, Yugawa T, et al. HPV16 E6-mediated stabilization of ErbB2 in neoplastic transformation of human cervical keratinocytes. Oncogene 2007;26:2988–96.
32. English DP, Roque DM, Santin AD. HER2 expression beyond breast cancer: therapeutic implications for gynecologic malignancies. *Mol Diagn Ther* 2013;17:85–99.
33. Ross JS, Ali SM, Wang K, et al. Comprehensive genomic profiling of epithelial ovarian cancer by next generation sequencing-based diagnostic assay reveals new routes to targeted therapies. *Gynecol Oncol* 2013;130:554–9.
34. Wheler J, Hong D, Swisher SG, et al. Thymoma patients treated in a phase I clinic at MD Anderson Cancer Center: responses to mTOR inhibitors and molecular analyses. *Oncotarget* 2013;4:890–8.