Role of the Common Fragile Sites in Cancers with a Human Papillomavirus Etiology

Ge Gao David I. Smith

Division of Experimental Pathology, Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA

Keywords

Cervical cancer · Common fragile sites · Human papillomavirus · Oropharyngeal squamous cell carcinoma

Abstract

High-risk human papillomaviruses (HPVs) are known to be associated with different anogenital cancers including cervical, anal, penile, and vaginal cancers. They are also found to be responsible for the dramatic increases in oropharyngeal squamous cell carcinoma (OPSCC) observed in the United States and Europe. The model for how high-risk HPVs induce cancer formation comes from studies of cervical cancer which usually involves integration of the HPV into the human genome and subsequent changes due to induced chromosomal instability. Recent work, discussed here, however suggests that this model may not be completely correct. In addition, we summarize studies now done in OPSCC which demonstrate that the role of HPV in these cancers may be different from that in cervical cancer. Finally, we propose new models for how HPV may be involved in the formation of these 2 cancers.

© 2017 S. Karger AG, Basel

Human Papillomaviruses and Human Cancers

Human papillomavirus (HPV) is a double-stranded DNA virus from the papillomavirus family that can cause infections in the skin and mucous epithelial membranes. Over 170 types of HPV have been identified, and more than 40 of them are known to cause anogenital infections through sexual contact [de Villiers, 2013]. HPV genomes are approximately 8 kb in size and encode 6 early transcribed genes and 2 late transcribed genes. HPVs can be classified as low-risk (LR) viruses, which are only associated with the development of warts (papillomas), or high-risk (HR) viruses, as infection with them results in an increased risk for the development of cancer [Lacey et al., 2006; Parkin and Bray, 2006]. The difference between LR- and HR-HPVs is predominantly encoded within 2 genes that in HR-HPV subtypes are highly oncogenic. The 2 corresponding oncogenic proteins (HR-HPV E6 and E7) have been shown to promote the degradation of p53 and pRb by the ubiquitination-mediated pathway, thus leading to cell cycle checkpoint abrogation and genome instability [Dyson et al., 1989; Scheffner et al., 1990, 1993].

HR-HPVs are known to be associated with cancers of the cervix, vulva, vagina, penis, anus, and oropharynx [Anorlu, 2008; Chaturvedi et al., 2008]. Over 99% of cervical cancers have HPV infections, and the currently ac-

cepted model for the role of HPVs in the formation of cancers was developed from several decades of study of these cancers [Walboomers et al., 1999; zur Hausen, 2002]. Sexually active women are repeatedly exposed to HPV, and those women infected with HR-HPV strains whose immune system does not clear virus are at an increased risk of developing cervical cancer. HPV infects the basal epithelial layer in the cervix, and there the virus exists in an episomal form. The copy number of episomal HPV can actually be very high, and long-term infection with HPV can eventually lead to the virus integrating into the human genome. Viral integration appears to be associated with the development of invasive cervical cancer in most, but not all, cervical cancer cases. Studies in cervical cancer-derived cell lines and several primary cervical cancers found that after HPV integration there is a loss of the episomal copies of HPV. The site of integration frequently occurs within the HPV E2 gene which functions as a transcriptional repressor of the HPV oncogenes E6 and E7 [Badaracco et al., 2002].

Since not everyone infected with HR-HPV develops cancer, it suggests that additional genetic and genomic changes in the host cells are needed to facilitate cancer malignant progression. The accumulated cellular alterations could be due to the oncogenic roles of HR-HPV E6/E7, or could result from structural variations or gene expression alterations caused by the disruption of DNA sequences at or near the site(s) of HPV integration.

HR-HPV E6 and E7 Induced Genome Instability

Several lines of evidence have suggested that increased expression of the HPV E6 and E7 oncoproteins could lead to the disruption of genome integrity in host cells. First, HPV E7 can bind and degrade pRb and the related pRb family members p107 and p30 [Münger et al., 2001]. The pRb protein is a tumor suppressor that plays a critical role in negatively controlling the cell cycle at the G1 checkpoint, thus blocking S-phase entry and cell growth [Knudsen et al., 2000]. Loss of pRb leads to the unscheduled activation of cyclin/CDK2 complexes which could give rise to an increased risk for chromosome mis-segregation leading to aneuploidy which in turn could promote carcinogenic progression [Lundberg and Weinberg, 1998]. The E6 oncoprotein binds to and promotes the degradation of p53 which abrogates the G2/M checkpoint control thus leading to centrosome-related mitotic defects [Scheffner et al., 1990, 1993]. It has been shown that precise chromosome segregation in mitosis and coordinated cell

cycle progression play important roles in maintaining genomic stability [Shen, 2011]. Thus, the dysregulation of these 2 important cell cycle control proteins could further promote genome instability. In addition, expression of the HPV16 E6/E7 oncoproteins can result in the formation of multiple rearrangements, telomeric associations, and gross aneuploidy [White et al., 1994]. Furthermore, the HPV E6/E7 oncoprotein-expressing cells also show a decreased response to DNA damage repair [Kessiss et al., 1993; Song et al., 1998] which could be due to the loss of the function of both p53 and pRb. Increased DNA breakage or an impaired DNA repair pathway also raise the chance of integrating foreign DNA which may promote the viral DNA integrating into the host genome. Finally, microsatellite instability has also been observed in HPV-associated tumors [Rodriguez et al., 1998] suggesting that this could induce more subtle genome-wide alterations.

Persistent HPV infections can eventually result in an increased episomal copy number of HPV which could cause increased expression of the HPV E6 and E7 oncogenes. This would disrupt the host cell's genome integrity, which in concert with impaired DNA damage repair could contribute to cancer development. In cervical cancer, it was reported that over 80% of the HPV16-positive cancers and almost all HPV18-positive cancers have HPV integrated in the host genome [Cullen et al., 1991; Pirami et al., 1997]. As HPV integrations were found throughout the human genome based upon an analysis of a limited number of cell lines and primary tumors, the model suggested that the key event of integration was the increased expression of the 2 oncoproteins resulting from the disruption of the HPV E2 gene and that the sites of integration in the human genome were unimportant. However, our laboratory studied a larger number of HPV16 and HPV18 integrations, and we observed that over 50% of the HPV integrations actually occurred within one of the highly unstable common fragile site (CFS) regions. In addition, we also found that there was a hot spot for HPV18 integrations within one of the CFSs surrounding the *c-Myc* oncogene.

Common Fragile Sites

CFSs are regions of profound genomic instability that are found in all individuals [Glover et al., 1984]. When cells are cultured in the presence of the inhibitor of DNA polymerase alpha, aphidicolin, the CFSs are observed as discrete chromosomal regions that are either decondensed or broken in metaphase spreads of chromosomes

[Sutherland and Richards, 1995; Reshmi et al., 2007]. In order for a specific chromosomal region to be designated as a CFS, it must show this decondensation/breakage at a statistically significant frequency. There are over 90 of these chromosomal regions distributed throughout the human genome. CFSs are more fully characterized by cytogenetic techniques where large-insert clones (such as bacterial artificial chromosomes, BACs) are used as FISH probes on metaphases produced from cells grown in the presence of aphidicolin or other mitotic inhibitors. An individual BAC present within a CFS region can hybridize proximal, distal, or actually crossing the region of decondensation/breakage in different metaphases indicating that the CFS regions are considerably larger than the several hundred kilobase-long BAC clones. The full characterization of a CFS region involves the use of multiple BAC clones to define where it begins and ends. CFS regions that have been characterized have been found to range in size from less than 1 Mb to 10–15 Mb.

The biological functions of the CFSs are unknown but these chromosomal regions are found in all individuals, in contrast to the rare fragile sites. The 3 most frequently expressed CFSs are FRA3B (3p14.2), FRA16D (16q23.2), and FRA6E (6q26). Using the FISH-based assay to characterize these regions, we previously found that the full sizes of these regions were 4.0, 2.0, and 3.6 Mb, respectively. Our laboratory analyzed a number of the other CFS regions including FRA7G (300 kb), and FRA9E (9.0 Mb) [Huang et al., 1999; Callahan et al., 2003]. Rozier et al. [2004] characterized a CFS at human 4q22. This region of instability was found to span 15 Mb. Of the more than 90 CFS regions present in the human genome, however, only a small number has been fully characterized to determine precisely where they begin and end.

In spite of these chromosomal regions being so unstable, they are surprisingly conserved evolutionarily. The high conservation of the sequences within these regions and the fact that the homologous regions are also CFSs in other species suggest that there is an important biological function for these regions [Cha and Kleckner, 2002; Arlt et al., 2003; Lemoine et al., 2005].

Mechanism of CFS Instability

The reason why the CFS regions are so unstable remains elusive. Analysis of various CFS regions revealed that several features of these regions might contribute to their fragility. First, sequence analysis showed that the nucleotide sequences around the fragile sites have the po-

tential to form stable secondary DNA structures which might impede DNA replication elongation [Snow et al., 1994; Nancarrow et al., 1995]. Second, AT and CGG repeats which are often found to be expanded in rare fragile sites have enhanced torsional flexibility which could perturb replication progression as well as nucleosome assembly affecting the organization of the chromatin [Wang et al., 1996]. Several of the CFS regions also contain a number of AT or CGG repeats. However, 3 CFS regions, FRA3B, FRA7H, and FRA7G, did not show any expanded repeated sequence features that contribute to their fragility. Instead, it was revealed that they contain several regions having the potential to form unusual DNA structures with high flexibility, low stability, and non-B DNA-forming sequences which may affect their replication, condensation, and organization, thus leading to their fragility [Mishmar et al., 1999]. In addition, Minocherhomji et al. [2015] demonstrated that replication stress activates DNA repair synthesis in mitosis. The MUS81-EME1 structure-specific endonuclease promotes the appearance of chromosome gaps or breaks at CFSs at anaphase which leads to DNA damage in the subsequent G1 phase [Ying et al., 2013].

Large Genes in CFS Regions and Their Role in Cancer

The frequent deletions and other alterations of CFS regions observed in multiple cancers led to a search within these regions for important genes that might be the targets of these deletions. This then led to the discovery that there are a number of extremely large genes contained within these regions, and several of them have been demonstrated to function as important tumor suppressors involved in the development of multiple cancers.

FHIT

The first CFS region to be extensively characterized was the most active and most frequently observed CFS, FRA3B (3p14.2). This CFS is located in a chromosomal region that is frequently deleted in a number of different cancers and also contains a balanced reciprocal translocation t(3;8)(p14.2;q24.13) breakpoint found in a family which was predisposed to develop renal cell carcinoma. This translocation breakpoint was first identified in a 1.5-Mb yeast artificial chromosome that was derived from within the FRA3B region [Boldog et al., 1993].

Kay Huebner's laboratory identified the fragile histidine triad (*FHIT*) gene within FRA3B that spanned 1.5 Mb of genomic DNA, crossed the hRCC translocation

breakpoint, and had deletions in many different cancers [Ohta et al., 1996]. Surprisingly, the final processed transcript for this extremely large gene was only 1.1 kb in size. *FHIT* is not a typical tumor suppressor gene as many cancers had deletions in the very large introns of this gene that still left the coding regions intact. In spite of this, *FHIT* protein expression was decreased or absent in many different cancers. The mouse homolog, *Fhit*, is also a very large gene, and the chromosomal region surrounding this gene is a CFS in mice; hence, the large gene and the highly unstable region surrounding it were conserved [Matsuyama et al., 2003].

FHIT's tumor suppressor function was demonstrated in several cancer cell lines by both knock in and knock out experiments. Caspase dependent apoptosis, small apoptotic cell fractions, and the accumulation of the cells at S to G2 phase were observed within different cancer-derived cell lines when *FHIT* was introduced into them [Ishii et al., 2001]. *Fhit*^{+/-} and *Fhit*^{-/-} mice were also shown to be prone to develop tumors that were repressed with adenoviral transfected wild-type *FHIT* [Dumon et al., 2001]. *FHIT* was also reported to prevent the epithelial to mesenchymal transition. It is now thought that *FHIT* is a genome caretaker, and cells with loss of *FHIT* activity have spontaneous replication stress, increased dsDNA breaks, and chromosome instability, suggesting loss of *FHIT* could induce global genome instability [Saldivar et al., 2012]. It was also shown that loss of *FHIT* could increase the fragility of other CFS regions [Hosseini et al., 2013]. The deletion of the FRA3B region, aberrant transcripts, or the reduction or absence of *FHIT* protein expression occurs in almost 50% of all cancers. In lung squamous cell carcinoma, the loss of *FHIT* was observed in up to 90% of tumors [Sozzi et al., 1998] and in many premalignant lung lesions suggesting that carcinogens such as tobacco could cause genome instability and might actually be targeting the FRA3B region [Mao et al., 1997].

WWOX

The second most frequently observed CFS region, FRA16D, is within chromosomal band 16q23.2. Deletions within this band were observed in a number of different cancers, and about 25% of multiple myelomas had translocations within this same region [Krummel et al., 2000; Paige et al., 2000]. A second extremely large gene, called *WWOX* (1.1 Mb), was identified within the FRA16D CFS, and the protein encoded by *WWOX* is an oxidoreductase with 2 WW domains within it. This gene, like *FHIT*, has a very small final processed transcript of only 2.3 kb.

WWOX has a number of other similarities to *FHIT*. Deletions within the large region spanned by *WWOX* are seen in many different cancers, and the translocations observed in multiple myeloma frequently interrupt this gene. The mouse homolog of *WWOX*, *Wox1*, is highly conserved between humans and mice, and the chromosomal region surrounding *Wox1* was also found to be a highly unstable CFS [Krummel et al., 2002].

The loss of heterozygosity of the chromosomal region surrounding *WWOX*, homozygous deletions within this gene, aberrant *WWOX* transcripts, and promoter hypermethylation of *WWOX* have been reported in different cancers including esophageal squamous cell carcinomas, non-small cell lung cancer, prostate cancers, and pancreatic cancers [Kuroki et al., 2002, 2004; Yendamuri et al., 2003; Watson et al., 2004]. Both exogenous and endogenous *WWOX* restoration inhibited cell growth and reduced xenograft tumor growth in breast cancer cells [Iliopoulos et al., 2007]. In addition, *Wox1* knockout mice are more prone to form tumors [Aqeilan et al., 2007], demonstrating that *WWOX* is also a bona fide tumor suppressor like *FHIT*. It was also suggested that *WWOX* worked as an effector of hyaluronidase that could increase cancer cell sensitivity to tumor necrosis factor cytotoxicity [Chang et al., 2001].

PARK2

The third most frequently expressed CFS region is FRA6E (6q26). Contained within the middle of this region of instability is another very large gene, *PARK2* [Denison et al., 2003]. This 1.3-Mb gene was originally identified as being mutated in patients with autosomal recessive juvenile Parkinson's disease; hence, this was the first large CFS gene identified which is associated with neurological disease [Kitada et al., 1998]. Similar to *FHIT* and *WWOX*, *PARK2* also produces a small 3-kb final processed transcript. *PARK2* encodes an E3 ubiquitin-protein ligase, and somatic *PARK2* mutations could decrease its E3 ligase activity including its ability to ubiquitinate cyclin E which then results in mitotic instability [Imai et al., 2000]. Frequent deletions and other alterations of *PARK2* have been observed in multiple cancers including ovarian, breast, renal, and lung cancers.

CFSs and the Largest Human Genes

Since the 3 most frequently expressed CFS regions each span extremely large genes, and since they all function as tumor suppressors, we examined a list of the larg-

est known human genes to determine if any of the other CFS regions also comprised large genes that also could be playing important roles in cancer formation. There are actually 40 human genes that span at least 1 Mb of genomic DNA. Table 1 is a list of those 40 genes. Also contained in that list is the full size of the genomic region spanned by each of these very large genes, the total number of exons of each gene, and the size of the final processed transcript. There are a number of important observations that can be made from an examination of this list. First, there is no relationship between the full size of these genes and their final processed transcript. For example, the 1.5-Mb *FHIT* gene contains only 9 exons and actually has a very small final processed transcript of 1.1 kb. In contrast, the 2.1-Mb *CSMD1* gene contains 70 exons and has an 11.580-kb final processed transcript. Second, many of these large genes appear to play important roles in normal neurological development as it was observed that mutations in many of these genes are responsible for neurological defects. For example, deletion of the *Park2* gene and its neighbor, the *Park2* co-regulated gene, results in a mouse mutant with important neurological defects, Quaking, in addition to the role of *PARK2* in Parkinson's disease [Wilson et al., 2010]. Lurcher mice die shortly after birth because of a lack of mid- and hindbrain neurons, and these mice have deletions in the very large *Grid2* gene [Lalonde et al., 2003]. *Dab1* mutant mice have the neurological defect called Scrambler and have cerebellar hyperplasia and Purkinje cell ectopia [Rice et al., 1998]. A recent article showed that the very large CFS genes were preferential breakage sites in neural stem/progenitor cells suggesting a very interesting linkage between normal neural development and cancer [Wei et al., 2016]. However, the most important observation about these very large genes is that so many of them are derived from within the same chromosomal bands that are known to contain CFS regions. This does not definitively prove that these large genes are contained within those CFS regions, but based upon this colocalization, we began to examine a number of these very large genes in an attempt to determine which ones of them were actually derived from within the large CFSs that were localized in the same chromosomal bands.

In order to prove that a specific large gene was actually derived from within a CFS region, we identified large-insert BACs that span a portion of these genes. These BAC clones were fluorescently labeled and used as FISH probes on metaphase spreads from cells cultured in the presence of aphidicolin. We then examined a sufficient number of metaphases to determine where that specific

Table 1. The 40 large genes spanning more than 1 Mb and their chromosomal locations

Name	Chromosome	Closest CFS	Size, bp	Exons (FPT, bp)
1 <i>CNTNAP2</i>	7q35	FRA7I	2,304,258	25 (8,107)
2 <i>DMD</i>	Xp21.1	FRAXC	2,092,287	79 (13,957)
3 <i>CSMD1</i>	8p23.2	FRA8B	2,056,709	70 (11,580)
4 <i>LRP1B</i>	2q22.1	FRA2F	1,900,275	91 (16,556)
5 <i>CTNNA3</i>	10q21.3	FRA10D	1,775,996	18 (3,024)
6 <i>NRXN3</i>	14q24.3	FRA14C	1,691,449	21 (6,356)
7 <i>A2BP1/RBFOX1</i>	16p13.2		1,691,217	16 (2,279)
8 <i>DAB1</i>	1p32.3	FRA1B	1,548,827	21 (2,683)
9 <i>PDE4D</i>	5q11.2		1,513,407	17 (2,465)
10 <i>FHIT</i>	3p14.2	FRA3B	1,499,181	9 (1,095)
11 <i>FAM190A/CCSER1</i>	4q22.1		1,474,315	11 (5,833)
12 <i>GPC5</i>	13q31.3	FRA13D	1,468,199	8 (2,588)
13 <i>GRID2</i>	4q22.3	FRA4G	1,467,842	16 (3,024)
14 <i>DLG2</i>	11q14.1	FRA11F	1,463,760	23 (3,071)
15 <i>AIP1</i>	7q21.11	FRA7E	1,436,474	21 (6,795)
16 <i>DPP10</i>	2q14.1		1,402,038	26 (4,905)
17 <i>PARK2</i>	6q26	FRA6E	1,379,130	12 (2,960)
18 <i>ILIRAPL1</i>	Xp21.2	FRAXC	1,368,379	11 (2,722)
19 <i>PRKG1</i>	10q21.1	FRA10C	1,302,704	18 (2,213)
20 <i>EBI</i>	12q23.1	FRA12C	1,248,678	26 (3,750)
21 <i>CSMD3</i>	8q23.2	FRA8C	1,213,952	69 (12,486)
22 <i>ILIRAPL2</i>	Xq22.3		1,200,827	11 (2,985)
23 <i>AUTS2</i>	7q11.22	FRA7J	1,193,536	19 (5,972)
24 <i>DCC</i>	18q21.1	FRA18B	1,190,131	29 (4,608)
25 <i>GPC6</i>	13q31.3	FRA13D	1,176,822	9 (2,731)
26 <i>CDH13</i>	16q23.2	FRA16D	1,169,565	15 (3,926)
27 <i>ERBB4</i>	2q34	FRA2I	1,156,473	28 (5,484)
28 <i>ACCN1/ASIC2</i>	17q11.2		1,143,718	10 (2,748)
29 <i>CTNNA2</i>	2p12	FRA2E	1,135,782	18 (3,853)
30 <i>WDSUB1</i>	2q24		1,126,043	16 (2,132)
31 <i>DKFZp686H</i>	11q25	FRA11G	1,117,478	8 (6,830)
32 <i>PTPRT</i>	20q12		1,117,144	32 (12,680)
33 <i>WWOX</i>	16q23.2	FRA16D	1,113,013	9 (2,264)
34 <i>NRXN1</i>	2p16.3	FRA2D	1,109,951	21 (8,114)
35 <i>IGSF4D/CADM2</i>	3p12.1		1,109,105	13 (9,580)
36 <i>CDH12</i>	5p14.3	FRA5E	1,102,578	15 (4,167)
37 <i>PAR3L</i>	2q33.3	FRA2I	1,069,815	23 (4,176)
38 <i>PTPRN2</i>	7q36.3	FRA7I	1,048,712	22 (4,735)
39 <i>SOX5</i>	12p12.1		1,030,095	18 (4,492)
40 <i>TCBA1/NKAIN2</i>	6q22.31	FRA6F	1,021,499	8 (3,183)

CFS, common fragile site; FPT, final processed transcript.

BAC clone hybridized relative to the CFS region. If the BAC always hybridized proximal or distal to that breakage, then that BAC clone was proximal or distal to the CFS region. However, if the BAC clone actually crossed the region of decondensation/breakage in 1 or more metaphases, or if it sometimes hybridized proximal in one metaphase but then distal in another, we could conclude that this BAC, and the large gene, must have been derived from within that CFS region. Not all the very large genes were localized within a CFS. *A2BP1*, the ataxin-2 binding protein 1, spans 1.7 Mb within the short arm

of chromosome 16, but this chromosomal arm contains no CFSs. The known tumor suppressor *DCC* is a 1.2-Mb-long gene but it maps proximal to the CFS derived from the same chromosomal band.

However, many of the largest human genes were localized within CFS regions. This includes *DAB1*, *RORA*, *DMD*, and the very large adjacent gene *IL1RAPL1*. Several additional very large genes were localized within CFSs by other groups, including *GRID2* (in FRA4G), *ESRRG* (in FRA1H), and *NBEA* (in FRA13A). Thus, many of the largest human genes were localized within CFS regions and can thus be classified as large CFS genes.

Most of these large CFS genes also appear to play important roles in cancer development as deletions and other alterations within these genes are frequently observed in different cancers. However, to definitively prove that these genes function as tumor suppressors requires experimentation demonstrating that re-expression of these genes in vitro in cancer-derived cell lines lacking their expression suppresses tumor growth. More importantly, it also requires that mice with these genes knocked out are tumor prone and that replacing these genes in these mice suppresses those tumors. Hence, many of the large CFS genes are very attractive tumor suppressor candidates, but definitive proof that they actually function as tumor suppressors is currently lacking.

CFSs and Large CFS Genes in Cervical Cancer

Previously, we utilized the technique of restriction site oligonucleotide PCR to examine HPV integration events in a large number of cervical cancers. This technique is based upon the use of PCR primers derived from multiple positions within the HPV genome coupled with PCR primers specific for restriction endonuclease recognition sites. While we detected HPV16 and HPV18 integrations throughout the genome, they were not randomly distributed, as we found that 50% of the HPV16 integrations and 60% of the HPV18 integrations occurred within one of the CFS regions. We also showed that deletions and complex rearrangements frequently occurred in the cellular sequences targeted by the integrations, and the integrations clustered in FRA13C (13q22), FRA3B (3p14.2), and FRA17B (17q23) [Thorland et al., 2003]. Studies from other groups also suggested that chromosome bands that contain CFS regions are the preferred sites for HPV integrations in both cervix carcinoma-derived cell lines and clinical samples [Hidalgo et al., 2003; Dall et al., 2008; Matovina et al., 2009]. Furthermore, we also found a hot

spot of HPV18 integrations within one of the CFSs in chromosomal band 8q24 near the *c-Myc* proto-oncogene [Ferber et al., 2003].

A number of the integrations were also found to occur within very large genes, and we tested if they were localized within the closely mapped CFS regions. This enabled us to identify several additional very large CFS genes. One of the large CFS genes identified by this work was *LRP1B*. This is a very interesting gene which not only has deletions in a number of different cancers, but whose function demonstrates that it could play an important role in cancer development. Downregulation of *LRP1B* expression promotes cell migration via the RhoA/Cdc42 pathway and actin skeleton remodeling in renal cell cancer [Ni et al., 2013]. Furthermore, *LRP1B* deletion in high-grade serous ovarian cancers is associated with acquired chemotherapy resistance to liposomal doxorubicin [Cowin et al., 2012].

The HeLa cell line was derived from a highly aggressive HPV18-positive cancer cell, and whole-genome sequencing of the HeLa genome revealed that this cancer has HPV18 integrated within one of the CFS regions that surround the *c-Myc* oncogene. This integration event results in amplification of portions of the HPV genome between 8 and 32 times and also causes dramatic disruption of the human genome around that integration site [Adey et al., 2013]. This integration and amplification causes increased expression of *c-Myc*, demonstrating that the site of HPV integration within the human genome (at least in this cell line) does play an important role in the eventual cancer that develops. Thus, not only is HPV integration into the human genome potentially nonrandom, but the viral integration event can be quite disruptive and cause changes in the expression of important genes that could contribute to carcinogenesis.

Loss or aberrant expression of *FHIT*, *WWOX*, and *PARK2* were also observed in cervical cancer-derived cell lines and clinical samples. Aberrant *FHIT* transcripts and loss of *FHIT* protein expression are often found in primary cervical cancers and high-grade noninvasive lesions but not in normal cervical tissues [Wu et al., 2000], suggesting that *FHIT* inactivation can play an important role in cervical tumor progression. Reduced or absent *FHIT* protein is more frequently observed in invasive cancer than in high-grade squamous intraepithelial lesions suggesting it could play an important role in late stages of cancer progression [Connolly et al., 2000]. Moreover, abnormal *FHIT* expression was also reported as an independent poor prognostic factor for cervical cancer [Takizawa et al., 2003]. Protein expression of both *FHIT* and *WWOX*

were absent or reduced in several cell lines. Giarnieri et al. [2010] revealed that the expression of FHIT and WWOX protein were inversely correlated with precancerous versus invasive lesions, suggesting that altered expression of FHIT and WWOX could play a critical role in the progression of cervical neoplasia. WWOX was also found to induce apoptosis and inhibits proliferation in cervical cancer and cell lines [Qu et al., 2013]. Similarly, loss of heterozygosity of the region surrounding *PARK2* was reported in primary cervical tumor samples [Mehdi et al., 2011], and it was also revealed that *PARK2* could induce apoptotic cell death in TNF- α -treated cervical cancers [Lee et al., 2012]. HeLa cells also harbor a deleted *PARK2* gene and are resistant to TNF- α induced cell death [Franco et al., 2002].

HPV and Oropharyngeal Squamous Cell Carcinoma

HPV is also found to be responsible for the dramatically increased incidence of oropharyngeal squamous cell carcinomas (OPSCCs) which are a subset of head and neck cancers. These are cancers of the base of the tongue, tonsils, the soft palate, and the walls of the pharynx. The overall incidence of most types of head and neck cancer has been decreasing due to decreased tobacco consumption in the United States in the past several decades. However, the incidence of OPSCC cancers has actually been dramatically increasing, and currently around 80% of OPSCCs in the United States and 90% of those in Europe are HPV-positive [O'Regan et al., 2008; Boscolo-Rizzo et al., 2009; Marur et al., 2010]. This is most likely caused by changing sexual behaviors. In addition, HPV-positive OPSCCs differ from HPV-negative OPSCCs in tumor biology. Most importantly, HPV-positive OPSCCs have a much better clinical outcome than HPV-negative OPSCCs. Clinicians are currently attempting to determine if they could de-escalate therapies for patients with HPV-positive OPSCCs, hence most OPSCC patients are now tested for the presence of HPV sequences or the expression of HPV E6/E7 [Mirghani et al., 2015].

Based upon the currently accepted HPV-driven cervical cancer model, it was just assumed that HPV plays an identical role in the development of OPSCC as it does in cervical cancer. However, given the fact that most HPV-positive OPSCC patients also have a smoking and drinking history, HPV's role in the development of this type of cancer might be more complex. In addition, it has been shown that not all HPV-positive OPSCCs are the same as some OPSCCs have very low expression of HPV E6 and

E7, and these are thought to be due to latent HPV infections. Other OPSCCs have robust E6 and E7 expression, and these are thought to have active HPV infections. Thus, important questions are whether HPV plays different roles in different OPSCCs, what is the physical status of the virus in HPV-positive OPSCCs, and if and when HPV integrates into the human genome, where does the integration occur, both within the HPV genome and the human genome.

Our previous studies examining HPV integration in cervical cancer utilized restriction site PCR, but recent advances in DNA sequencing, specifically next-generation sequencing (NGS), is a much more powerful tool which enabled us to examine the physical status of HPV sequences in OPSCC. One specific application of NGS is mate-pair NGS (MP-Seq). This technology has the advantage to determine genomic structural variations at a very high resolution utilizing only 3–5 Gb of sequencing [Feldman et al., 2011; Murphy et al., 2012]. In addition, it could also detect HPV integration events which would be mate-pairs where one end of a mate-pair corresponds to human sequence and the other to an HPV-derived sequence [Gao et al., 2014].

We used MP-seq to study the HPV integration events in HPV-positive OPSCC and found that there were some distinct differences from what is usually observed in cervical cancer: (1) Only 30% of HPV-positive OPSCCs had viral integration. (2) The HPV integration event is not necessarily related to the level of HPV E6 and E7 onco-gene expression; thus, HPV can integrate into the human genome even in the presence of very low HPV E6 and E7 expression. (3) We did not observe any HPV integrations within the E2 gene. (4) Some OPSCCs had HPV integrated into the human genome but still had a high episomal copy number of HPV sequences. (5) The expression of HPV16 E6 and E7 was always concordant with E2 expression regardless of the HPV integration status in the OPSCC tumor samples [Gao et al., 2014].

Thus, our results demonstrate that the classically accepted cervical cancer model most likely does not apply to many OPSCCs. Seventy percent of OPSCCs develop in just the presence of episomal copies of HPV, which is true for only about 20% of cervical cancers. Although these cancers develop just in the presence of episomal copies of HPV, there are also differences. In cervical cancer, the episomal HPV copy number is always high in this small group of patients, while there are big variations of the HPV copy number found in OPSCC patients who only have episomal HPV present. In addition, the expression of HPV E6/E7 is independent of whether HPV is inte-

grated or not, and may not be dependent on whether the E2 gene is disrupted.

There are thus some profound differences between what we have observed in OPSCC and the classical accepted model for HPV-driven cervical cancer. However, there are also some similarities. When HPV integration occurred in OPSCC, we found that half of those integrations were within one of the CFS regions. We previously demonstrated that 60% of OPSCCs had decreased expression for 6 of the large CFS genes, and cancers with decreased expression of these 6 genes were more likely to recur. Hence, the CFS regions may have multiple interactions in HPV-driven cancers both in terms of being hot spots for integration (when it does occur) and also as regions that are particularly prone to being altered especially in genomically highly unstable HPV-positive cells.

A New Model for the Role of HPV in Cancer Development

Our work, and the work of others, has demonstrated that the sites of HPV integration in the human genome in different cancers, including cervical cancer, are nonrandom. The CFS regions, and the large genes contained within them, are important targets for HPV integrations as well as hot spots for genomic deletions. While the original model for HPV-driven cervical cancer involved frequent inactivation of the E2 gene retaining E6 and E7 intact, more recent studies utilizing NGS to analyze additional cell lines and both early and later stage primary cervical samples have revealed that integrations occur throughout the HPV genome and that the E2 region is not actually a hot spot for integrations. It now appears that the sites of integration in the HPV genome might be random, while the regions within the human genome that are disrupted by viral integration could be very important for the eventual cancers that develop.

Thus, we would like to suggest a new model for how HPV is involved in the development of cervical cancer. Sexually active women are repeatedly exposed to HPV, but those women who do not clear the virus are the ones at risk for developing cervical cancer. A subgroup of these women could develop cervical cancer just in the presence of a high episomal HPV copy number and the ensuing genomic instability resulting from the high expression of the 2 oncoproteins. However, long-term infection with HPV can also increase the probability of HPV integration into the human genome. These integration events can be quite disruptive to genes at or near the integration sites.

Once HPV has integrated, the increased genomic instability both within the region of integration and throughout the host genome could result in other alterations which could eventually lead to invasive disease. There might also be some clinical significance regarding where HPV has integrated and the structure of the actual integrations that we are beginning to explore.

The situation may be more complex with respect to the role of HPV in OPSCC. The majority of HPV-positive OPSCCs have just episomal copies of HPV present. Some of them have a low copy number of HPV, and how this causes cancer is presently unclear. It is possible that these cancers develop when there is a high copy number of HPV present but that after cancer formation some of them are cleared. However, another possibility is that in those patients who have a smoking history this, in addition to a low copy number of HPV sequences, could still promote additional alterations that could ultimately lead to cancer formation. Other OPSCCs have a high episomal copy number of HPV, and we propose that these cause cancer in a manner similar to that in cervical cancers that have no HPV integration. Then there are the 30% of OPSCCs that have HPV integrated into 1 or more positions within the human genome. We hypothesize that these may follow a model more similar to HPV's role in most cervical cancers and that integration could be targeting key genes, again such as the large CFS genes. Here too, we think that the sites of integration in the HPV genome are random, but that the sites in the human genome might not be.

HR-HPV is also found in a number of other cancers including 85% of anal cancers, and over half of vaginal, penile, and vulvar cancers. Currently little is known about the physical status of HPV in those cancers; hence it is too early to predict how HPV contributes to these cancers' development. More likely HPV may play different roles in the development of different cancers. These results suggest that much more work needs to be done, starting with using the powerful tools of NGS technologies to begin to answer these important questions.

Disclosure Statement

The authors have no conflicts of interest to declare.

References

- Adey A, Burton JN, Kitzman JO, Hiatt JB, Lewis AP, et al: The haplotype-resolved genome and epigenome of the aneuploid HeLa cancer cell line. *Nature* 500:207–211 (2013).
- Anorlu RI: What is the significance of the HPV epidemic? *Can J Urol* 15:3860–3865 (2008).
- Aqeilan RI, Trapasso F, Hussain S, Costinean S, Marshall D, et al: Targeted deletion of *Wwox* reveals a tumor suppressor function. *Proc Natl Acad Sci USA* 104:3949–3954 (2007).
- Arlt MF, Casper AM, Glover TW: Common fragile sites. *Cytogenet Genome Res* 100:92–100 (2003).
- Badaracco G, Venuti A, Sedati A, Marcante ML: HPV16 and HPV18 in genital tumors: significantly different levels of viral integration and correlation to tumor invasiveness. *J Med Virol* 67:574–582 (2002).
- Boldog FL, Gemmill RM, Wilke CM, Glover TW, Nilsson AS, et al: Positional cloning of the hereditary renal carcinoma 3;8 chromosome translocation breakpoint. *Proc Natl Acad Sci USA* 90:8509–8513 (1993).
- Boscolo-Rizzo P, Da Mosto MC, Fuson R, Frayle-Salamanca H, Trevisan R, Del Mistro A: HPV-16 E6 L83V variant in squamous cell carcinomas of the upper aerodigestive tract. *J Cancer Res Clin Oncol* 135:559–566 (2009).
- Callahan G, Denison SR, Phillips LA, Shridhar V, Smith DI: Characterization of the common fragile site FRA9E and its potential role in ovarian cancer. *Oncogene* 22:590–601 (2003).
- Cha RS, Kleckner N: ATR homolog Mec1 promotes fork progression, thus averting breaks in replication slow zones. *Science* 297:602–606 (2002).
- Chang NS, Pratt N, Heath J, Schultz L, Slevé D, et al: Hyaluronidase induction of a WW domain-containing oxidoreductase that enhances tumor necrosis factor cytotoxicity. *J Biol Chem* 276:3361–3370 (2001).
- Chaturvedi AK, Engels EA, Anderson WF, Gillison ML: Incidence trends for human papillomavirus-related and -unrelated oral squamous cell carcinomas in the United States. *J Clin Oncol* 26:612–619 (2008).
- Connolly DC, Greenspan DL, Wu R, Ren X, Dunn RL, et al: Loss of Fhit expression in invasive cervical carcinomas and intraepithelial lesions associated with invasive disease. *Clin Cancer Res* 6:3505–3510 (2000).
- Cowin PA, George J, Fereday S, Loehrer E, Van Loo P, et al: LRP1B deletion in high-grade serous ovarian cancers is associated with acquired chemotherapy resistance to liposomal doxorubicin. *Cancer Res* 72:4060–4073 (2012).
- Cullen AP, Reid R, Campion M, Lörincz AT: Analysis of the physical state of different human papillomavirus DNAs in intraepithelial and invasive cervical neoplasm. *J Virol* 65:606–612 (1991).
- Dall KL, Scarpini CG, Roberts I, Winder DM, Stanley MA, et al: Characterization of naturally occurring HPV16 integration sites isolated from cervical keratinocytes under non-competitive conditions. *Cancer Res* 68:8249–8259 (2008).
- Denison SR, Wang F, Becker NA, Schüle B, Kock N, et al: Alterations in the common fragile site gene *Parkin* in ovarian and other cancers. *Oncogene* 22:8370–8378 (2003).
- de Villiers EM: Cross-roads in the classification of papillomaviruses. *Virology* 445:2–10 (2013).
- Dumon KR, Ishii H, Fong LY, Zanesi N, Fidanza V, et al: *FHIT* gene therapy prevents tumor development in *Fhit*-deficient mice. *Proc Natl Acad Sci USA* 98:3346–3351 (2001).
- Dyson N, Howley PM, Münger K, Harlow E: The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. *Science* 243:934–937 (1989).
- Feldman AL, Dogan A, Smith DI, Law ME, Ansell SM, et al: Discovery of recurrent t(6;7)(p25.3;q32.3) translocations in ALK-negative anaplastic large cell lymphomas by massively parallel genomic sequencing. *Blood* 117:915–919 (2011).
- Ferber MJ, Thorland EC, Brink AA, Rapp AK, Phillips LA, et al: Preferential integration of human papillomavirus type 18 near the *c-myc* locus in cervical carcinoma. *Oncogene* 22:7233–7242 (2003).
- Franco DL, Nojek IM, Molinero L, Coso OA, Costas MA: Osmotic stress sensitizes naturally resistant cells to TNF-alpha-induced apoptosis. *Cell Death Differ* 9:1090–1098 (2002).
- Gao G, Johnson SH, Kasperbauer JL, Eckloff BW, Tombers NM, et al: Mate pair sequencing of oropharyngeal squamous cell carcinomas reveals that HPV integration occurs much less frequently than in cervical cancer. *J Clin Virol* 59:195–200 (2014).
- Giarnieri E, Zanesi N, Bottoni A, Alderisio M, Lukic A, et al: Oncosuppressor proteins of fragile sites are reduced in cervical cancer. *Cancer Lett* 289:40–45 (2010).
- Glover TW, Berger C, Coyle J, Echo B: DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. *Hum Genet* 67:136–142 (1984).
- Hidalgo A, Monroy A, Arana RM, Taja L, Vázquez G, Salcedo M: Chromosomal imbalances in four new uterine cervix carcinoma derived cell lines. *BMC Cancer* 3:8 (2003).
- Hosseini SA, Horton S, Saldivar JC, Miuma S, Stampfer MR, et al: Common chromosome fragile sites in human and murine epithelial cells and *FHIT*/FRA3B loss-induced global genome instability. *Genes Chromosomes Cancer* 52:1017–1029 (2013).
- Huang H, Reed CP, Mordí A, Lomberg G, Wang L, et al: Frequent deletions within FRA7G at 7q31.2 in invasive epithelial ovarian cancer. *Genes Chromosomes Cancer* 24:48–55 (1999).
- Iliopoulos D, Fabbri M, Druck T, Qin HR, Han SY, Huebner K: Inhibition of breast cancer cell growth in vitro and in vivo: effect of restoration of *Wwox* expression. *Clin Cancer Res* 13:268–274 (2007).
- Imai Y, Soda M, Takahashi R: Parkin suppresses unfolded protein stress-induced cell death through its E3 ubiquitin-protein ligase activity. *J Biol Chem* 275:35661–35664 (2000).
- Ishii H, Dumon KR, Vecchione A, Trapasso F, Mimori K, et al: Effect of adenoviral transduction of the fragile histidine triad gene into esophageal cancer cells. *Cancer Res* 61:1578–1584 (2001).
- Kessis TD, Slebos RJ, Nelson WG, Kastan MB, Plunkett BS, et al: Human papillomavirus 16 E6 expression disrupts the p53-mediated cellular response to DNA damage. *Proc Natl Acad Sci USA* 90:3988–3992 (1993).
- Kitada T, Asakawa S, Hattori N, Matsumine H, Yamamura Y, et al: Mutations in the *parkin* gene cause autosomal recessive juvenile parkinsonism. *Nature* 392:605–608 (1998).
- Knudsen KE, Booth D, Naderi S, Sever-Chroneos Z, Fribourg AF, et al: RB-dependent S-phase response to DNA damage. *Mol Cell Biol* 20:7751–7763 (2000).
- Krummel KA, Roberts LR, Kawakami M, Glover TW, Smith DI: The characterization of the common fragile site FRA16D and its involvement in multiple myeloma translocations. *Genomics* 69:37–46 (2000).
- Krummel KA, Denison SR, Calhoun E, Phillips LA, Smith DI: The common fragile site FRA16D and its associated gene *WWOX* are highly conserved in the mouse at Fra8E1. *Genes Chromosomes Cancer* 34:154–167 (2002).
- Kuroki T, Trapasso F, Shiraishi T, Alder H, Mimori K, et al: Genetic alterations of the tumor suppressor gene *WWOX* in esophageal squamous cell carcinoma. *Cancer Res* 62:2258–2260 (2002).
- Kuroki T, Yendamuri S, Trapasso F, Matsuyama A, Aqeilan RI, et al: The tumor suppressor gene *WWOX* at FRA16D is involved in pancreatic carcinogenesis. *Clin Cancer Res* 10:2459–2465 (2004).
- Lacey CJ, Lowndes CM, Shah KV: Chapter 4: Burden and management of non-cancerous HPV-related conditions: HPV-6/11 disease. *Vaccine* 24 Suppl 3:S3/S35–41 (2006).
- Lalonde R, Hayzoun K, Selimi F, Mariani J, Strazielle C: Motor coordination in mice with hot-foot, Lurcher, and double mutations of the *Grid2* gene encoding the delta-2 excitatory amino acid receptor. *Physiol Behav* 80:333–339 (2003).
- Lee K, Lee MH, Kang YW, Rhee KJ, Kim TU, Kim YS: Parkin induces apoptotic cell death in TNF- α -treated cervical cancer cells. *BMB Rep* 45:526–531 (2012).
- Lemoine FJ, Degtyareva NP, Lobachev K, Petes TD: Chromosomal translocations in yeast induced by low levels of DNA polymerase a model for chromosome fragile sites. *Cell* 120:587–598 (2005).

- Lundberg AS, Weinberg RA: Functional inactivation of the retinoblastoma protein requires sequential modification by at least two distinct cyclin-cdk complexes. *Mol Cell Biol* 18:753–761 (1998).
- Mao L, Lee JS, Kurie JM, Fan YH, Lippman SM, et al: Clonal genetic alterations in the lungs of current and former smokers. *J Natl Cancer Inst* 89:857–862 (1997).
- Marur S, D'Souza G, Westra WH, Forastiere AA: HPV-associated head and neck cancer: a virus-related cancer epidemic. *Lancet Oncol* 11:781–789 (2010).
- Matovina M, Sabol I, Grubisic G, Gasperov NM, Grce M: Identification of human papillomavirus type 16 integration sites in high-grade precancerous cervical lesions. *Gynecol Oncol* 113:120–127 (2009).
- Matsuyama A, Shiraishi T, Trapasso F, Kuroki T, Alder H, et al: Fragile site orthologs *FHIT/FRA3B* and *Fhit/Fra14A2*: evolutionarily conserved but highly recombinogenic. *Proc Natl Acad Sci USA* 100:14988–14993 (2003).
- Mehdi SJ, Alam MS, Batra S, Rizvi MM: Allelic loss of 6q25–27, the *PARKIN* tumor suppressor gene locus, in cervical carcinoma. *Med Oncol* 28:1520–1526 (2011).
- Minocherhomji S, Ying S, Bjerregaard VA, Bursumanno S, Aleliunaite A, et al: Replication stress activates DNA repair synthesis in mitosis. *Nature* 528:286–290 (2015).
- Mirghani H, Amen F, Blanchard P, Moreau F, Guigay J, et al: Treatment de-escalation in HPV-positive oropharyngeal carcinoma: ongoing trials, critical issues and perspectives. *Int J Cancer* 136:1494–1503 (2015).
- Mishmar D, Mandel-Gutfreund Y, Margalit H, Rahat A, Kerem B: Common fragile sites: G-band characteristics within an R-band. *Am J Hum Genet* 64:908–910 (1999).
- Münger K, Basile JR, Duensing S, Eichten A, Gonzalez SL, et al: Biological activities and molecular targets of the human papillomavirus E7 oncoprotein. *Oncogene* 20:7888–7898 (2001).
- Murphy SJ, Cheville JC, Zarei S, Johnson SH, Sikkink RA, et al: Mate pair sequencing of whole-genome-amplified DNA following laser capture microdissection of prostate cancer. *DNA Res* 19:395–406 (2012).
- Nancarrow JK, Holman K, Mangelsdorf M, Hori T, Denton M, et al: Molecular basis of p(CCG)n repeat instability at the FRA16A fragile site locus. *Hum Mol Genet* 4:367–372 (1995).
- Ni S, Hu J, Duan Y, Shi S, Li R, et al: Down expression of *LRP1B* promotes cell migration via RhoA/Cdc42 pathway and actin cytoskeleton remodeling in renal cell cancer. *Cancer Sci* 104:817–825 (2013).
- Ohta M, Inoue H, Cotticelli MG, Kastury K, Baffa R, 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).
- O'Regan EM, Toner ME, Finn SP, Fan CY, Ring M, et al: p16^{INK4A} genetic and epigenetic profiles differ in relation to age and site in head and neck squamous cell carcinomas. *Hum Pathol* 39:452–458 (2008).
- Paige AJ, Taylor KJ, Stewart A, Sgouros JG, Gabra H, et al: A 700-kb physical map of a region of 16q23.2 homozygously deleted in multiple cancers and spanning the common fragile site FRA16D. *Cancer Res* 60:1690–1697 (2000).
- Parkin DM, Bray F. Chapter 2: The burden of HPV-related cancers. *Vaccine* 24 Suppl 3:S11–25 (2006).
- Pirami L, Giachè V, Becciolini A: Analysis of HPV16, 18, 31, and 35 DNA in pre-invasive and invasive lesions of the uterine cervix. *J Clin Pathol* 50:600–604 (1997).
- Qu J, Lu W, Li B, Lu C, Wan X: WWOX induces apoptosis and inhibits proliferation in cervical cancer and cell lines. *Int J Mol Med* 31:1139–1147 (2013).
- Reshmi SC, Huang X, Schoppy DW, Black RC, Saunders WS, et al: Relationship between FRA11F and 11q13 gene amplification in oral cancer. *Genes Chromosomes Cancer* 46:143–154 (2007).
- Rice DS, Sheldon M, D'Arcangelo G, Nakajima K, Goldowitz D, Curran T: Disabled-1 acts downstream of Reelin in a signaling pathway that controls laminar organization in the mammalian brain. *Development* 125:3719–3729 (1998).
- Rodriguez JA, Barros F, Carracedo A, Mugicavan Herckenrode CM: Low incidence of microsatellite instability in patients with cervical carcinomas. *Diagn Mol Pathol* 7:276–282 (1998).
- Rozier L, El-Achkar E, Apiou F, Debatisse M: Characterization of a conserved aphidicolin-sensitive common fragile site at human 4q22 and mouse 6C1: possible association with an inherited disease and cancer. *Oncogene* 23:6872–6880 (2004).
- Saldivar JC, Miuma S, Bene J, Hosseini SA, Shibata H, et al: Initiation of genome instability and preneoplastic processes through loss of Fhit expression. *PLoS Genet* 8:e1003077 (2012).
- Scheffner M, Werness BA, Huibregtse JM, Levine AJ, Howley PM: The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell* 63:1129–1136 (1990).
- Scheffner M, Huibregtse JM, Vierstra RD, Howley PM: The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell* 75:495–505 (1993).
- Shen Z: Genomic instability and cancer: an introduction. *J Mol Cell Biol* 3:1–3 (2011).
- Snow K, Tester DJ, Kruckeberg KE, Schaid DJ, Thibodeau SN: Sequence analysis of the fragile X trinucleotide repeat: implications for the origin of the fragile X mutation. *Hum Mol Genet* 3:1543–1551 (1994).
- Song S, Gulliver GA, Lambert PF: Human papillomavirus type 16 E6 and E7 oncogenes abrogate radiation-induced DNA damage responses in vivo through p53-dependent and p53-independent pathways. *Proc Natl Acad Sci USA* 95:2290–2295 (1998).
- Sozzi G, Pastorino U, Moiraghi L, Tagliabue E, Pezzella F, et al: Loss of *FHIT* function in lung cancer and preinvasive bronchial lesions. *Cancer Res* 58:5032–5037 (1998).
- Sutherland GR, Richards RI: The molecular basis of fragile sites in human chromosomes. *Curr Opin Genet Dev* 5:323–327 (1995).
- Takizawa S, Nakagawa S, Nakagawa K, Yasugi T, Fujii T, et al: Abnormal Fhit expression is an independent poor prognostic factor for cervical cancer. *Br J Cancer* 88:1213–1216 (2003).
- Thorland EC, Myers SL, Gostout BS, Smith DI: Common fragile sites are preferential targets for HPV16 integrations in cervical tumors. *Oncogene* 22:1225–1237 (2003).
- Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, et al: Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 189:12–19 (1999).
- Wang YH, Gellibolian R, Shimizu M, Wells RD, Griffith J: Long CCG triplet repeat blocks exclude nucleosomes: a possible mechanism for the nature of fragile sites in chromosomes. *J Mol Biol* 263:511–516 (1996).
- Watson JE, Doggett NA, Albertson DG, Andaya A, Chinnaiyan A, et al: Integration of high-resolution array comparative genomic hybridization analysis of chromosome 16q with expression array data refines common regions of loss at 16q23-qter and identifies underlying candidate tumor suppressor genes in prostate cancer. *Oncogene* 23:3487–3494 (2004).
- Wei PC, Chang AN, Kao J, Du Z, Meyers RM, et al: Long neural genes harbor recurrent DNA break clusters in neural stem/progenitor cells. *Cell* 164:644–655 (2016).
- White AE, Livanos EM, Tlsty TD: Differential disruption of genomic integrity and cell cycle regulation in normal human fibroblasts by the HPV oncoproteins. *Genes Dev* 8:666–677 (1994).
- Wilson GR, Wang HX, Egan GF, Robinson PJ, Delatycki MB, et al: Deletion of the Parkin coregulated gene causes defects in ependymal ciliary motility and hydrocephalus in the *quaking^{viable}* mutant mouse. *Hum Mol Genet* 19:1593–1602 (2010).
- Wu R, Connolly DC, Dunn RL, Cho KR: Restored expression of fragile histidine triad protein and tumorigenicity of cervical carcinoma cells. *J Natl Cancer Inst* 92:338–344 (2000).
- Yendamuri S, Kuroki T, Trapasso F, Henry AC, Dumon KR, et al: WW domain containing oxidoreductase gene expression is altered in non-small cell lung cancer. *Cancer Res* 63:878–881 (2003).
- Ying S, Minocherhomji S, Chan KL, Palmi-Pallag T, Chu WK, et al: MUS81 promotes common fragile site expression. *Nat Cell Biol* 15:1001–1007 (2013).
- zur Hausen H: Papillomaviruses and cancer: from basic studies to clinical application. *Nat Rev Cancer* 2:342–350 (2002).