BIOMEDICAL AND ENVIRONMENTAL SCIENCES **19**, 392-398 (2006)

Molecular Cloning and Preliminary Analysis of a Fragile Site Associated Gene

YI-WEN CAO[†,*], CHUAN-LU JIANG[‡,*], AND TAO JIANG[#,1]

[†]Department of Echography, People's Armed Police General Hospital of Heilongjiang Province, Harbin 150076, Heilongjiang, China; [‡]Department of Neurosurgery, the Second Affiliated Hospital, Harbin Medical University, Harbin 150001, Heilongjiang, China; [#]Department of Neurosurgery, Neurosurgery Institute and Tiantan Hospital, PUMC and CAMS, Beijing 100050, China

Objective To analyze the molecular colning of a fragile site-associated gene. **Methods** Genomic Chinese hamster ovary (CHO) DNA library was constructed using high molecular weight CHO DNA partially digested with *Mbo*I restriction enzyme from cultured CHO cells. Screening of genomic DNA library followed the established procedures. Genomic CHO in the positive clones was sequenced. Appropriate primers were designed for the reverse transcriptase-polymerase chain reactions (RT-PCR). The RT-PCR products were cloned into a pCRII TOPO vector and confirmed by DNA sequencing. Antibodies were prepared using synthetic peptides as antigens by immunizing the rabbits. Immunohistochemical analyses were performed to evaluate the expression of the novel gene in different tissues. **Results** To investigate the molecular mechanism underlying the initial events of *mdr1a* amplification, we cloned 1q31 fragile site DNA. Strikingly, we found that this fragile site contained a novel gene which was designated as a fragile site-associated (FSA) gene. FSA encoded an unusually large mRNA of ~16 kb. Full-length human FSA cDNA was cloned. FSA mRNA was expressed in many cultured cells and tissue types. Immunohistochemical analyses also revealed an expression pattern of the encoded proteins in postmitotic, well-differentiated epithelial compartments of many organs, including colon, mammary glands, ovary, prostate, and bladder. **Conclusion** FSA plays an important role in regulating mammalian epithelial cell growth and differentiation.

Key words: Fragile site; DNA amplification; Epithelial differentiation

INTRODUCTION

DNA amplification is frequently observed in drug-resistant and cancer cells. Recent studies have demonstrated that DNA amplification involves the McClintock's breakage-fusion-bridge (BFB) mechanism[1]. According to this model, an initial break of a chromatid bearing the selected gene leads to isochromatid fusion of the uncapped sister chromatids after replication. The resultant dicentric chromosome forms an anaphase bridge between the centromeres. As the two centromeres move toward opposite poles during telophase of the cell cycle progression, an additional break between centromere and the selected gene of the dicentric chromosome occurs. The resultant chromosome contains a duplication chromosomal segment harboring extra-copy of the target gene, thus completing the initial round of DNA amplification.

The BFB mechanism underscores the importance of the initial chromosomal break as a trigger for the DNA amplification process. However, the molecular mechanisms underlying the initial break remain largely unknown. It was reported that breakage of the common fragile site 1q31 was involved in the BFB-mediated amplification of the multidrug resistance *mdr1a* gene in many independently established multidrug resistant CHO cells[2]. The involvement of fragile sites in the amplification of other drug resistance genes in drug-resistant variants has also been observed[3]. Moreover, Hellman *et al.*[4] showed that amplification of the Met oncogene in human glioma cells also involved the BFB mechanism, with the initial break at the common fragile site FRA7G, suggesting that common fragile sites play a role in DNA amplification.

To elucidate the molecular basis of the role of chromosomal fragile sites and DNA amplification, we cloned the CHO chromosomal 1q31 fragile site DNA and identified a novel gene which was

[1]Correspondence should be addressed to Tao JIANG, Department of Neurosurgery, Institute of Neurosurgery and Tiantan Hospital, Capital University of Medical Sciences, 6 Tiantan Xili, Beijing 100050, China. Tel: 86-10-13801126903. E-mail: jiangtao369@sohu.com

Biographical note of the first authors: Yi-Wen CAO, majoring in physical diagnosis. Tel: 86-451-82387879. E-mail: nanni_2000@163.com. Chuan-Lu JIANG, M. D., Ph. D., professor. Tel: 86-451-86605088. E-mail: Chuanlu1965@yahoo.com.cn

[*] Equal contributors to this article.

0895-3988/2006
CN 11-2816/Q

designated as a fragile site-associated (FSA) gene that was located in this fragile site. FSA encoded an mRNA spanning ~16 kb, probably by one of the largest mammalian mRNAs. Proteins encoded by FSA were expressed in many postmitotic, well-differentiated epithelial compartments. These results suggest that FSA plays an important role in regulating cell growth and differentiation in mammalian epithelial compartments.

METHODS

Cell Cultures

CHO TI6e cells[5] and the multidrug-resistant MDR (CHO) cell lines VCR0.2, VBR0.5, and ADR0.2[2] have been described previously. Human colorectal cancer cell lines HCT-116, HT-29, HT-15, BC, and cisplatin-resistant variant, brain tumor cell lines, glioma cell line, U373, A172, and U118, small cell lung cancer cell lines SR3A, SR3A-14, have been previously described[6]. Cells were maintained in Dulbecco's modified Eagle's medium (Sigma) supplemented with 10% fetal calf serum (Life Technologies, Inc.) in a humidified incubator containing 5% CO_2.

Construction of Genomic CHO DNA Library

High molecular weight CHO DNA was prepared from cultured cells, partially digested with *Mbo*I restriction enzyme, and size-fractionated by sucrose gradient centrifugation. Fragment sizes with 10-15 kb were collected and ligated to λDASHII vector arms and packaged (Sratagene, La Jolla, CA.). The packaged library was plated on the Xl-1 blue *Escherichia coli* strain. Screening of genomic DNA library followed the established procedures. Genomic CHO in the positive clones was sequenced.

Molecular Cloning of FSA cDNA

The CHO sequences from the phage clones were used to search the database (http://www.ncbi.nlm.nih.gov/BLAST/) and a human cDNA sequence (KIAA1109 or AB029032) sharing three exonic sequences was obtained. The sequence in KIAA1109 was used to search the overlapping, adjacent, and neighboring sequences in an additional database: (http://www.celeradiscoverysystem.com). Appropriate primers were designed for the reverse transcriptase-polymerase chain reaction (RT-PCR) using poly A^+ RNA prepared from normal human heart RNA as template and reagents from Promega (Madison, WI). The RT-PCR products were cloned into a pCRII TOPO vector (Strategene) and confirmed by DNA sequencing.

Northern Blot Hybridization

Total RNA was extracted from cultured cells and tissues using RNA STATE-60 (Friendswood, TX). Poly A^+ RNA was prepared using a mRNA preparation kit (Qiagen, Valencia, CA). Northern blot hybridization was carried out under the conditions as described previously[7].

Preparation of Antibodies

Two anti-FSA polyclonal antibodies were prepared using synthetic peptides as antigens (CHDDNSSDKDREDSITY and CALQDEKEKKGK DKEEH). Each peptide was conjugated with KLH using the maleimide chemistry. Rabbits were immunized and hyperimmune serum was processed over the respective immunosorbents to capture antibodies specific for each peptide using the procedure previously described (Bethyl Laboratory, Montgomery, TX).

Immunohistochemical Analyses

Immunohistochemistry was performed on paraffin-embedded tissue sections using the avidin-biotin-peroxidase complex method with the prepared primary antibodies. In brief, deproteinized sections were heated in a microwave oven for 12 min to retrieve the antigens. Endogenous peroxidase was blocked using 0.6% hydrogen peroxide in methanol for 20 min. Tissue sections were treated with 10% fetal bovine serum for 10 min. Primary antibody was added to slides and incubated at room temperature for 10 min. After washed in PBS, the sections were incubated with biotinylated anti-rabbit immunoglobulins (DAKO, Carpinteria, CA) for 10 min, followed by incubation with streptavidin peroxidase conjugate (DAKO). The color reaction was developed using liquid DAB substrate pack (BioGenex, San Ramon, CA). Sections were counterstained with hematoxylin.

RESULTS

Molecular Cloning of CHO Chromosomal 1q31 Fragile Site DNA

A CHO-human somatic hybrid line containing a segment of human chromosome 3 and a full complement of CHO chromosomes were treated with a fragile site inducer, aphidicolin. The treated cells were transfected with *pSV$_2$neo* plasmid DNA followed by G418 selection. The *pSV$_2$neo* integration sites in the positive transfectants were determined by fluorescence *in situ* hybridization (FISH). It was found that many of the integration sites were known

fragile sites. One of the clones, designated as TI6e, contained only one integration at the CHO chromosomal 1q31 fragile site involved in amplification of *mdr1a* in MDR CHO cells.

To clone 1q31 fragile site DNA, we prepared a TI6e genomic DNA library in phage λDASH II vector. The library was screened using *neo* sequence as a probe. Two phage clones with overlapping CHO DNA sequences were obtained. We sequenced the entire CHO DNA linked to the integrated *neo* sequence. A 4937-bp CHO DNA sequence was determined. The overall sequence was 64% A^+T. At least 10 highly A/T-rich clusters and homopurine/homopurimidine segments of ~30 bp long were present throughout the sequence. These sequences were potential hotspots for the formation of non-B-DNA structure or triple-helix structures[8-10]. The presence of these unusual DNA structures was characteristic of the fragile DNA found in many human common fragile sites[11], *i.e.*, FRA3B[12-14], FRA7G[10] and FRA7H[15].

BLAST searches revealed that the sequenced 1q31 fragile site DNA contains three exonic sequences sharing >90% similarity with those in human KIAA1109 cDNA (GeneBank No. AB029032). The 6377-bp KIAA cDNA sequence in the database contains an open reading frame (orf) with the putative TAA stop codon at nucleotide 5874. Using the human KIAA1109 sequence as a probe in fluorescent *in situ* hybridization (FISH) analysis, we mapped the hybridization signal to the Chinese hamster 1q31 fragile site (Fig. 1), and found that the sequences obtained from the phage library were located in the identified fragile site.

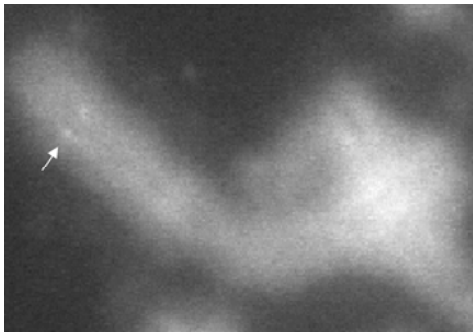

FIG. 1. (A), FISH localization of *FSA* sequence on CHO chromosomes. Arrow points to fragile site 1q31.

Molecular Cloning of FSA cDNA

We performed Northern blot hybridization using KIAA1109 cDNA as a probe to the poly A^+ RNA prepared from CHO and TI6e cells. A hybridization signal corresponding to ~16 kb mRNA was observed in both cell lines (Fig. 2A). These results demonstrated that an expressed gene encoding a mature mRNA of ~16 kb was located on the CHO chromosomal 1q31 fragile site. Similar hybridization signals were observed in many human cultured cell lines, including small cell lung cancer (SA3A and SR3A-14), colorectal cancer cell lines (BC, CP2.0) (Fig. 2A). We designated the encoded gene as a fragile site-associated (FSA) gene.

To investigate the tissue-specific expression of FSA, we carried out Northern blot hybridizations to RNA from various human tissues. We used commercial nitrocellular membranes containing 2 μg each of human poly A^+ RNA from various organs, including heart, brain, placenta, testis, prostate, lung, liver, pancreas, small and large intestines (Clontech, Palo Alto, CA, USA). A hybridization signal corresponding to ~16 kb mRNA in smear background was visible only on some tissues, *i.e.*, heart, brain, placenta, testis, and prostate, indicating that FSA mRNA in these samples were substantially degraded, although β-actin mRNA remained relatively intact (Fig. 2B). These results suggested that FSA mRNA was relatively unstable compared with β-actin mRNA, even when the differences in size between the two mRNAs were taken into account.

Since it was difficult to maintain the good quality for human tissue biopsy samples from operation room to laboratory, we prepared poly A^+ RNA from various mouse tissues. Northern blot hybridization showed a ~16 kb predominant band in all the tissues investigated, with elevated levels of expression in the testis and ovary (Fig. 2C).

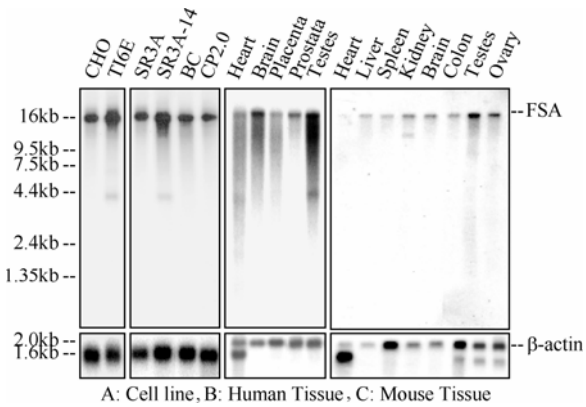

FIG. 2. Determination of FSA mRNA levels: Northern blots showing the hybridization of ~16 kb mRNA signals in cultured cell lines (A), human tissues (B), and mouse tissues (C). and mouse (E) tissues.

After completion of the human genome sequence, we cloned human FSA cDNA. Our strategy for cloning the human FSA cDNA is schematically

presented in Fig. 3. We searched the human genomic databases from NCBI and Celera. Several sequences were found to be overlapping or adjacent to the KIAA1109 (Fig. 3). Using this sequence information, we designed primers in RT-PCR for cloning the full-length FSA cDNA. However, we found that some of these sequences were alternative splicing variants and contained poly A^+ sequences, including AL137384, AK025057, AL137532, and BC018095, while others contained cloning artifacts with sequences that were not co-linear with the sequence in the FSA mRNA, after the appropriate primers for RT-PCR reactions were finally sorted out (Fig. 3). These allowed us to clone sequences in many gaps that were not in the databases (Fig. 3, dashed lines). A human FSA cDNA sequence of 15 595 bp was finally constructed. The entire mRNA sequence was confirmed by DNA sequencing from many overlapping RT-PCR clones using the indicated primers (Fig. 3, lower panel). Moreover, these RT-PCR products all hybridized to the ~16 kb RNA signal in Northern blotting analyses (data not shown). The nucleotide sequence and deduced amino acid sequences were not shown in this paper (data are available on request).

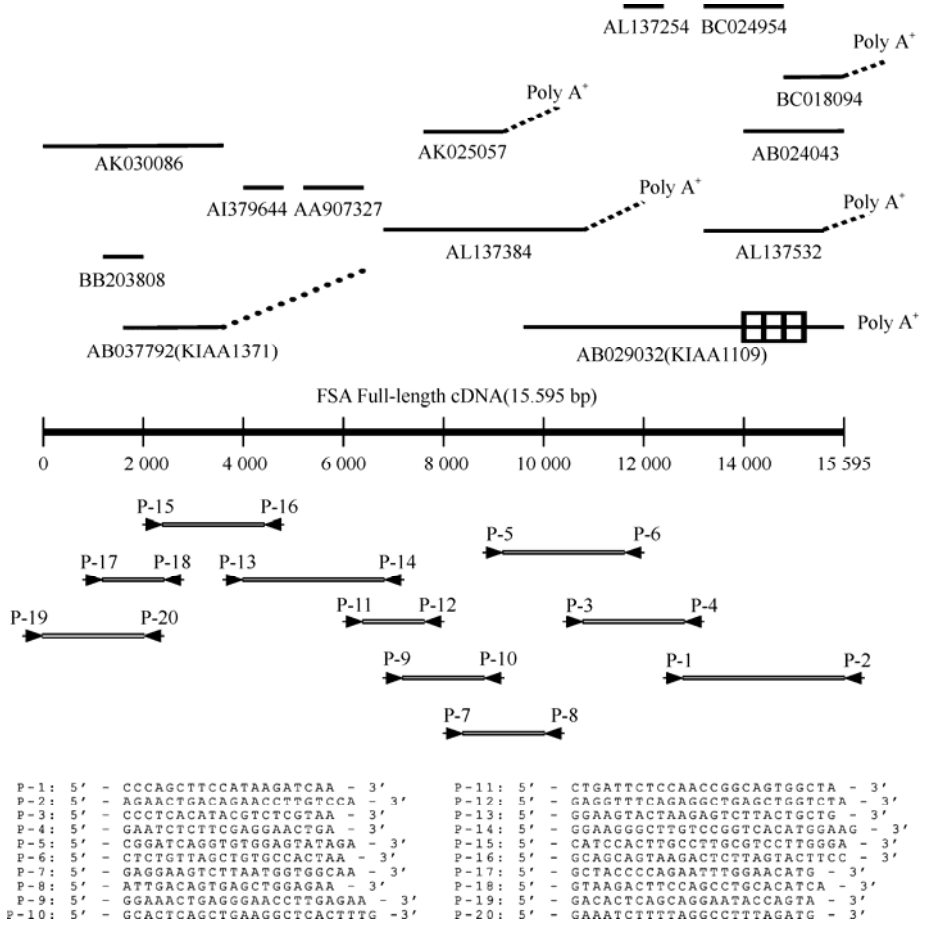

FIG. 3. Cloning strategy for the 16 kb-human FSA cDNA. Available DNA sequences in the database with the access numbers indicated are shown above the cDNA line. Broken lines indicated that the sequences are not present in the FSA cDNA, perhaps due to cloning artifacts or alternative splicing. KIAA1109 was the sequence detected by the hamster 1q31 fragile DNA, with the three exons indicated in boxes. Below the cDNA line is the RT-PCR cloning strategy using the primers as indicated. The PCR products were cloned into plasmid vector pCRII-TOPO and sequenced.

Immunohistochemical Analyses of FSA Expression in Tissues

To study the expression of FSA in the postmitotic compartments, antibodies against FSA were used to stain a series of paraffin-embedded tissue sections containing histologically normal tissues adjacent to tumor cells from various human biopsies. We first stained human testis sections. The expression of FSA protein was found at all stages except for the spermatogonia located at the periphery associated with the testis sack (Fig. 4A). We focused specifically

on epithelial tissues because they often undergo cell renewal and contain both proliferative and postmitotic, well-differentiated compartments. We focused on intestinal and epidermal epithelia which undergo continuous and rapid renewal. Epithelial cell renewal in the adult colon is initiated by the multipotent stem cells located at the base of mucosal invaginations known as crypts of Lieberkuhn[16]. In colonic crypt, the proliferating compartment is present near the crypt base. The transit amplifying cells occupy roughly two thirds of the height of the crypt, and the postmitotic differentiated cells line the upper part of the crypts[17]. Immunohistochemical studies showed that the postmitotic differentiated cells, or the upper part of the crypt was stained strongly, whereas the crypt located at the base of mucosal were relatively weak (Fig. 4B). These results showed that FSA was preferentially expressed in the postmitotic, well-differentiated colonic epithelial cells.

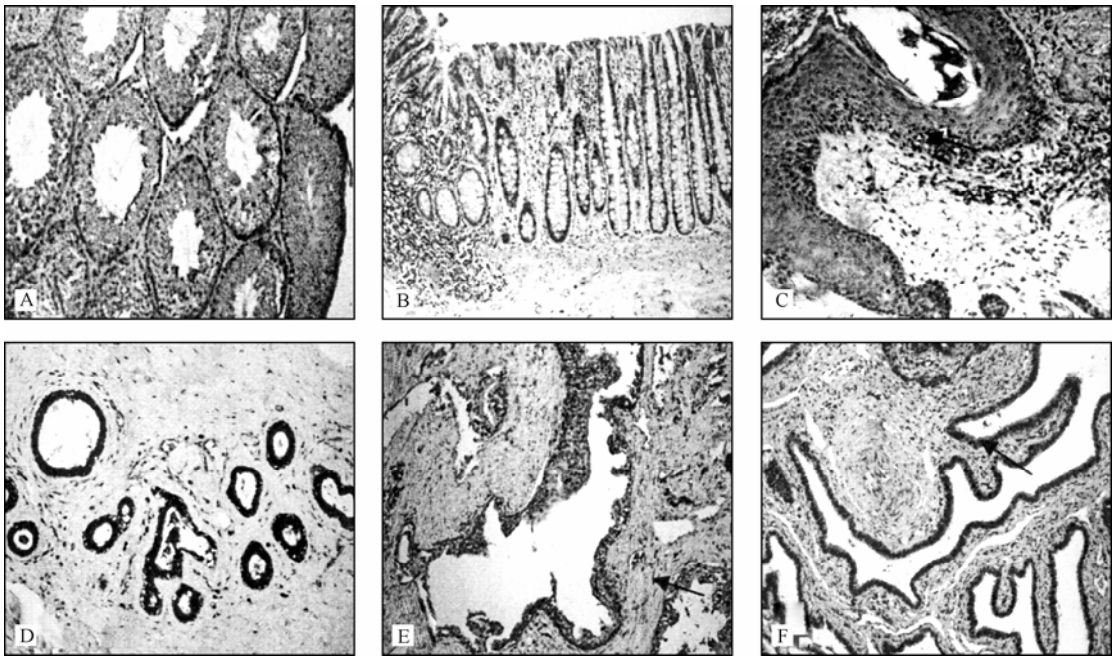

FIG. 4. Immunohistochemical staining using anti-FSA antibodies. (A), testis; (B), colon; (C), skin; (D), breast; (E), prostate; and (F), ovary. All are normal tissues adjacent to their respective tumor lesions.

Immunohistochemical staining was also carried out in human skin sections. In the epidermis, dividing keratinocytes are normally restricted to the basal layer. As cells withdraw from the cell cycle, they move outward and differentiate, producing postmitotic spinous layers, granular layer and dead stratum corneum cells[18]. Our immunohistochemical staining of human epidermis detected weak staining on the basal layer. Gradually increased expression of FSA outward in accordance with the differentiation program was seen (Fig. 4C). Immunohistochemical staining was also applied to other human tissues, i.e., breast, prostate, and ovary. As shown in Figs. 4D-F, respectively, the antibodies preferentially stained the epithelial compartments exposing to outer surface of the lining. It is noteworthy that in most cases, the basal layers of the epithelial linings representing proliferating population also showed reduced intensitivies of staining. These immunohistochemical results suggested that FSA expression was associated with a mature, well-differentiated epithelial population and that expression of the bicistronic FSA was associated with the mature, well-differentiated, epithelial compartments of many tissues, and that FSA played an important role in epithelial differentiation.

DISCUSSION

Molecular cloning of the CHO chromosomal 1q31 fragile site identified a transcription unit which was termed as FSA. The FSA gene encoding an mRNA of ~16 kb, probably by the largest mammalian mRNA was also identified. The polycistronic mRNA is abundant in prokaryotic and viral genomes, but rare in mammalian cells.

The exact functions of FSA-encoded proteins remain to be determined. However, from the

expression patterns as determined by the immunochemical stainings, one must consider that FSA plays an important role in the growth and differentiation of epithelial cells. The levels of FSA expression were diminished in many proliferating epithelial compartments but high in postmitotic and well differentiated compartment, suggesting that FSA may have a growth suppression function. Incidentally, the *FSA* locus was encompassed in the hamster 1q31 fragile site, but whether its human homologue locus 4q24 is also a fragile site remains to be determined. Several growth (tumor) suppressor genes are associated with human fragile sites. These genes include the well-studied fragile histidine triad (FHIT) tumor suppressor gene located in *FRA3B* fragile site[19], candidate tumor suppressor genes including CAV1, CAV2, and TESTIN located at *FRA7G*[20-21], and the putative tumor suppressor gene WWOX located at FRA16D[22].

FSA predominantly expressed in many different epithelial types may have important physiological implications regarding epithelial differentiation. Epithelial tissues differ in capabilities according to their anatomical location. Morphologically, these cells may be cuboid (e.g., breast), columnar (colon) or squamous (skin). Some are monotonic whereas others are polarized (mammary gland). Some epithelia stratify into multiple layers of cells (bladder and skin) while others may retain simple monolayer structure (ovary and breast). From the physiological aspects, epithelial cells perform very diverse functions, e.g., absorption (colon), active transport (colon), secretion (mammary gland), and protective barriers (colon and bladder), depending upon cell-cell interaction and microenvironmental influences[23]. From the pathological point of view, adult epithelial carcinomas including basal cell carcinoma and squamous cell carcinomas, carcinomas of the breast, prostate, lung, and colon, and other epithelial tissues account for more than 80% of all human tumor types as compared with cancers of mesenchymal or haematopoietic origins. Pediatric epithelial carcinomas only account for less than 10%[24]. These findings suggest that epithelial cancers are age-related and rise exponentially in the final decades of life. FSA proteins are expressed in many if not all mature epithelial cell types, suggesting that these proteins may play an important role in maintaining the epithelial hierarchy. The identification of a putative growth suppressor gene whose expression is predominantly associated with mature epithelial compartments raise an important scenario regarding the role of this gene in epithelial carcinomas of human cancers.

At the genomic levels, because the *fas* was initially cloned from the CHO fragile site. This raises the following important issues. Is the human FSA genomic sequence also fragile? Do the common fragile sites contain unstable DNA sequences that are prone to damages under the challenges by carcinogens, antitumor agents and stress-inducing agents? How does the human FSA gene respond to the treatments of these cytotoxic agents? Because of their genomic instability, fragile sites are often associated with genomic rearrangements and chromosomal aberrations during neoplastic transformation[25]. It is therefore important to investigate further the structure and expression of FSA in human malignant cells.

In summary, cloning of the FSA gene and the preliminary characteristics as described here have introduced a new paradigm for investigating the role of FSA in cancer epithelial biology.

REFERENCES

1. Kuo M T, Sen S, Hittelman W N, *et al.* (1998). Chromosomal fragile sites and DNA amplification in drug-resistant cells. *Biochem Pharmacol* **56**, 7-13.
2. Kuo M T, Vyas R C, Jiang L X, *et al.*(1994). Chromosomal breakage at a major fragile site associated with P-glycoprotein gene amplification in multidrug-resistant CHO cells. *Mol Cell Biol* **14**, 5202-5211.
3. Coquelle A, Pipiras E, Toledo F, *et al.* (1997). Expression of fragile sites triggers intrachromosomal mammalian gene amplification and sets boundaries to early amplicons. *Cell* **89**, 215-225.
4. Hellman A, Zlotorynski E, Sherer S W, *et al.* (2002). A role for common fragile site induction in amplification of human oncogenes. *Cancer Cell* **1**, 89-97.
5. Rassool F V, McKeithan T W, Neilly M E, *et al.* (1991). Preferential integration of marker DNA into the chromosomal fragile site at 3p14: an approach to cloning fragile sites. *Proc Natl Acad Sci USA* **88,** 6657-6661.
6. Yamane Y, Furuichi M, Song R, *et al.* (1998). Expression of multidrug resistance protein/GS-X pump and gamma-glutamylcysteine synthetase genes is regulated by oxidative stress. *J Biol Chem* **273**, 31075-31085.
7. Ishikawa T, Bao J J, Yamane Y, *et al.* (1996). Coordinated up-regulation of MRP/GS-X pump and glutamylcystein synthetase expression by cisplatin and heavy metals. *J Biol Chem* **271**, 14981-14988.
8. Frank-Kamenetskii M D, Mirkin S M (1995). Triplex DNA structures. *Annu Rev Biochem* **64**, 65-69.
9. Dickinson L A, Joh T, Kohwi Y (1992). A tissue-specific MAR/SAR DNA-binding protein with unusual binding site recognition. *Cell* **70**, 631-645.
10. Mishmar D, Rahat A, Scherer S W, *et al.* (1998). Molecular characterization of a common fragile site (FRA7H) on human chromosome 7 by the cloning of a simian virus 40 integration site. *Proc Natl Acad Sci USA* **95**, 8141-8146.
11. Glover T W (1998). Instability at chromosomal fragile site. *Rec Results Cancer Res* **154**, 186-199.
12. Boldog F, Gemmill R M, West J, *et al.* (1997). Chromosome 3p14 homozygous deletions and sequence analysis of FRA3B. *Hum Mol Genet* **6**, 193-203.
13. Huebne K, Garrison P N, Barnes L D, *et al.* (1998). The role of the FHIT/FRA3B locus in cancer. *Annu Rev Genet* **32**, 7-31.
14. Corbin S, Neilly M E, Espinosa R, *et al.* (2002). Identification of unstable sequences within the common fragile site at 3p14.2:

implications for the mechanism of deletions within fragile histidine triad gene/common fragile site at 3p14.2 in tumors. *Cancer Res* **62**, 3477-3284.
15. Huang H, Qian J, Proffit J, *et al.* (1998). FRA7G extends over a broad region: coincidence of human endogenous retroviral sequences (HERV-H) and small polydispersed circular DNAs (spcDNA) and fragile sites. *Oncogene* **16**, 2311-2319.
16. Cheng H, Leblond C P (1974). Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. V. Unitarian theory of the origin of the four epithelial cell types. *Am J Anat* **141**, 537-561.
17. Mills J C, Gordon J I (2001). The intestinal stem cell niche: there grows the neighborhood. *Proc Natl Acad Sci USA* **98**, 12334-12336.
18. Owens D M, Watt F M (2002). Contribution of stems cells and differentiated cells to epidermal tumours. *Nature Rev Cancer* **3**, 444-451.
19. Huebner K, Croce C (2001). FRA3B and other common fragile sites: the weakest links. *Nature Rev Cancer* **1**, 214 -221.
20. Hurlstone A F, Reid G, Reeves J R, *et al.* (1999). Analysis of the CAVEOLIN-1 gene at human chromosome 7q31.1 in primary tumors and tumor-derived cell lines. *Oncogene* **18**, 1881-1890.
21. Tatarelli C, Linnenbach A, Mimori K, *et al.* (2000). Characterization of the human testin gene localized in the FRA7G region at 7q31.2. *Genomics* **68**, 1-12.
22. Paige A J W, Taylor K J, Taylor C, *et al.* (2001). A candidate tumor suppressor gene involved in multiple tumor types. *Proc Natl Acad Sci USA* **98**, 11417-11422.
23. Slack J M W (2000). Stem cells in epithelial tissues. *Science* **287**, 1431-1433.
24. DePinho R A (2000). The age of cancer. *Nature* **408**, 248-254.
25. Mimori K, Druck T, Inoue H, *et al.* (1999). Cancer-specific chromosome alterations in the constitutive fragile region FRA3B. *Proc Natl Acad Sci USA* **96**, 7456-7461.

(*Received March* 12, 2006 *Accepted August* 4, 2006)