

# The HE4 (WFDC2) Protein Is a Biomarker for Ovarian Carcinoma<sup>1</sup>

Ingegerd Hellström,<sup>2</sup> John Raycraft, Martha Hayden-Ledbetter, Jeffrey A. Ledbetter, Michèle Schummer,<sup>3</sup> Martin McIntosh, Charles Drescher, Nicole Urban, and Karl Erik Hellström

Pacific Northwest Research Institute [I. H., J. R., M. H-L., J. A. L., K. E. H.], Institute for Systems Biology [M. S.], Fred Hutchinson Cancer Research Center [M. M., C. D., N. U.], and Swedish Hospital and Medical Center [C. D.], Seattle, Washington 98122

## ABSTRACT

The *WFDC2* (*HE4*) gene is amplified in ovarian carcinomas, whereas its expression in normal tissues, including ovary, is low. Although the function of the HE4 protein is unknown, it is a member of a family of stable 4-disulfide core proteins that are secreted at high levels. We therefore performed experiments to explore whether quantitation of HE4 protein levels in serum can be used as a biomarker for ovarian carcinoma. A fusion gene was constructed encoding the HE4 protein fused to a gene encoding the murine IgG2a Fc domain. Subsequently, protein produced in mammalian cells was purified by affinity chromatography and used to immunize mice to generate hybridomas specific for HE4. Hybridoma supernatants were screened for binding to a similar fusion protein that, instead, had a human immunoglobulin tail. Two hybridomas, 2H5 and 3D8, were selected that produce monoclonal antibodies to different HE4 epitopes, and a double determinant (“Sandwich”) ELISA was constructed and shown to detect a signal at the 160-pg level. Blinded studies on sera from postmenopausal patients with ovarian carcinoma and controls indicate that the specificity and sensitivity of the HE4-based ELISA is equivalent to that of the CA125 assay. However, the HE4 assay may have an advantage over the CA125 assay in that it is less frequently positive in patients with nonmalignant disease.

## INTRODUCTION

Ovarian cancer is potentially curable if diagnosed when localized (1), but most ovarian cancers are unfortunately diagnosed at an advanced stage when the survival rate is <20% despite aggressive surgery and chemotherapy (1). Currently, the only clinically accepted serum marker for ovarian cancer is CA125 (2, 3), which is found elevated in the majority of all ovarian cancers but in only half of those with early stage disease. Although overall CA125 has high sensitivity, its clinical use as an early detection marker for screening is severely limited, because it is also frequently elevated in women with more highly prevalent benign disease (4). Because these false positives are not easily eliminated by available imaging technologies, CA125-based screening would result in a large number of unnecessary surgeries. There is a need for novel markers that are sensitive and specific and/or improve diagnosis when used in combination with CA125 or that can replace it.

A time-honored approach to detect tumor markers is to immunize mice with human tumors, in this case, ovarian carcinoma, to obtain tumor-selective MAbs<sup>4</sup> (5–7). Recently, MAbs obtained this way were used to construct an ELISA that measures antigen(s) of the MPF (megakaryocyte potentiating factor) mesothelin family (referred to as

MSLN) in patient sera (7). Initial studies performed with this assay demonstrated antigen in sera from many patients with ovarian carcinoma as well as in sera from some patients with other malignancies (7). A subsequent “blinded” study indicated that this ELISA can complement the classical CA125 assay to improve the detection of patients with ovarian cancer.<sup>5</sup>

Another approach is based on the identification of overexpressed genes using techniques such as serial analysis of gene expression or cDNA microarrays (8–14). The latter was applied to identify two genes, *WFDC2* and *MSLN*, which are amplified in ovarian carcinomas but not in normal tissues (13, 14). The *WFDC2* gene was initially identified in epithelial cells of human epididymis and referred to as an epididymis-specific, fertility-related protein, HE4 (Homo sapiens epididymis specific; Refs. 15 and 16). HE4 belongs to a “four-disulfide core” family that comprises a heterogeneous group of small acid- and heat-stable proteins of divergent function. One of these proteins, the Wp protein (17), is expressed in milk, and secretory leukocyte protease inhibitor is present in seminal plasma and parotid secretions (18). Some members of the four-disulfide core family of proteins are protease inhibitors. These include secretory leukocyte protease inhibitor, which is a potent inhibitor of trypsin, chymotrypsin, elastase, and cathepsin G, as well as elafin (19), which is a potent inhibitor of elastase but not of other proteases, such as trypsin. No protease inhibitory activity has been identified for HE4, whose function remains unknown.

The initial microarray studies to search for genes that are overexpressed in ovarian carcinoma were extended by using membrane-based cDNA arrays, glass-based cDNA microarrays, and complementary statistical approaches to examine the expression of >100,000 transcripts in 13 ovarian cancers, 2 benign ovarian tumors, and 17 normal tissues.<sup>6</sup> Altogether, 101 transcripts were identified that were overexpressed in ovarian cancers as compared with normal tissues, including normal ovary. Real-time PCR of an independent set of benign and malignant tissues was then performed to characterize a subset of 64 genes of greatest interest, 12 of which were confirmed as overexpressed in ovarian cancers. Two of them, *WFDC2* (HE4) and *MSLN*, appeared to have the highest selectivity for ovarian cancer. Importantly, the former gene, in this study referred to as *HE4*, was not amplified in any of 19 tissues from women with benign ovarian masses who had elevated serum levels of CA125 (>30 units/ml).

Bingle *et al.* (20) have recently reported that the *HE4* gene is expressed also in some normal tissues and undergoes complex alternative splicing to yield multiple protein isoforms. Although this implies that HE4 proteins are not tumor specific, they may still have sufficient levels of tumor selectivity to serve as putative biomarkers for ovarian carcinoma. We have therefore generated MAbs to HE4 epitopes using a method that has been successfully applied in other systems (21). Constructs were made that encoded fusion proteins

Received 1/20/03; accepted 4/14/03.

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

<sup>1</sup> Supported by NIH/National Cancer Institute Specialized Program of Research Excellence Grant P50 CA83636 (Fred Hutchinson Cancer Research Center) and NIH Grant R01 CA 85780 (Pacific Northwest Research Institute).

<sup>2</sup> To whom requests for reprints should be addressed, at Pacific Northwest Research Institute, 720 Broadway, Seattle, WA 98122. Phone: (206) 726-1209; Fax: (206) 726-0302; E-mail: ihellstrom@pnri.org.

<sup>3</sup> Present address: MacroGenics, Inc., 1441 N. 34th St., Seattle, WA 98103-8914.

<sup>4</sup> The abbreviations used are: MAb, monoclonal antibody; MSLN, megakaryocyte potentiating factor/mesothelin family; AUC, area under curve; CHO, Chinese hamster ovary; Ig, immunoglobulin; ROC, receiver operating characteristic.

<sup>5</sup> M. W. McIntosh, C. Drescher, N. Urban, K. E. Hellstrom, and I. Hellstrom. Mesothelin/MPF antigen(s) as a diagnostic and screening marker for ovarian carcinoma when used alone or in combination with CA 125, submitted for publication.

<sup>6</sup> M. Schummer, N. Kiviat, L. Hood, C. Drescher, A. Ben-Dor, Z. Yakhini, M. W. McIntosh, A. F. Siegel, I. Podoflsky, I. Hellstrom, K. E. Hellstrom, and N. Urban. Microarray-based identification of potential ovarian cancer biomarkers, submitted for publication.

incorporating the *HE4* gene fused to a gene encoding Ig Fc domains from (IgG) mouse or humans. Mice were immunized with HE4 fusion protein that had a mouse Ig tail, after which hybridomas were generated and their supernatants screened against the human HE4 fusion protein. Two MAbs were derived, 2H5 and 3D8, which recognize different HE4 epitopes.

MAbs 2H5 and 3D8 were used to construct a double determinant ("Sandwich") ELISA. The ELISA was applied in a blinded study to test sera from a group of postmenopausal ovarian carcinoma patients and postmenopausal controls applying common processing and storage protocols. Our findings indicate that the HE4-based ELISA can aid the diagnosis of ovarian carcinoma and that it has an advantage over the classical CA125 test by being less frequently positive in women who have benign disease.

## MATERIALS AND METHODS

**Amplification of HE4 (WDFC2) cDNA from a High-throughput HE4 cDNA Clone toward Construction of a Fusion Construct.** Our strategy was to make constructs in which the *HE4* (WDFC2) gene was combined with genes encoding IgG to construct fusion proteins to immunize mice and obtain MAbs. The mice were immunized against fusion proteins with a mouse Ig tail, and the hybridomas were screened against fusion proteins with a human Ig tail. Subsequently, a double determinant (Sandwich) ELISA was constructed.

The mRNA sequence for *HE4* as originally published by Kirchoff *et al.* (15, 22) and deposited in GenBank (accession no. X63 187) provided the basis for oligonucleotide primer design to clone cDNA that encodes HE4. The cDNA for HE4, identified and isolated as a differentially expressed gene product using high-throughput cDNA arrays, was cloned in pSPORT as a 526-bp fragment. This cDNA was used as a template in PCR to amplify HE4 in a form appropriate for creating synthetic fusion protein genes.

A fusion protein was designed that incorporated the complete HE4 gene product fused to the human IgG1 Fc domain. Primers were designed that encoded appropriate restriction sites for cloning and created the necessary in-frame fusions of protein domains for the final construct. The 5' primer included a *HindIII* site, a Kozak sequence to improve expression adjacent to the first ATG, and a portion of the HE4 leader peptide based on the HE4 sequence published previously. The 3' primer included an in-frame *BamHI* site for fusion to the human-Ig tail cDNA, with the 3' end of the HE4 coding sequence truncated just before the STOP codon. PCR amplification reactions were performed according to manufacturer's instructions (ExTaq; Takara Bio, Inc., Otsu, Shiga, Japan) using 100 ng of *HE4*/pSPORT plasmid as a template and 30 cycles of amplification (1 min at 94°C, 1 min at 55°C, and 30 s at 72°C). Primer sequences are available on request. PCR products of the expected size (~400 bp) for the full-length HE4 were obtained and then purified using the QIAquick PCR Purification Kit (Qiagen, Valencia, CA). The purified PCR fragments were restriction digested, purified using the QIAex II Gel Extraction Kit (Qiagen), and ligated into the appropriately digested mammalian expression plasmid already containing the human IgG1 insert.

Ligation products were transformed into DH5 $\alpha$  bacterial cells, and transformants were screened for the presence of *HE4*-hIgG1 fusion gene inserts and verified by sequence analysis. In addition, protein expression was confirmed using plasmid DNA from these isolates to transiently transfect COS7 cells by the DEAE-Dextran technique as described (23). Culture supernatants were harvested after 72 h and screened by immunoprecipitation with protein A-agarose (Repligen, Cambridge, MA), reducing SDS-PAGE electrophoresis, and Western blotting. Western blots were probed using a goat antihuman IgG horseradish peroxidase conjugate (Caltag, Burlingame, CA) at 1:5000, followed by enhanced chemiluminescence development (Amersham, Little Chalfont, United Kingdom).

To clone HE4 cDNA, RNA was prepared from ovarian tumors and several ovarian tumor cell lines, including 4007 and OVCAR3 (24), using TRIzol (Life Technologies, Inc., Gaithersburg, MD) according to the manufacturer's instructions. cDNA was prepared using 1–3  $\mu$ g of RNA, random hexamers, and Superscript II Reverse Transcriptase (Life Technologies, Inc.) according to the manufacturer's directions. HE4 cDNA was PCR amplified from the random primed cDNA using standard conditions. PCR products of the expected

size for the full-length *HE4* were obtained, then cloned and confirmed by sequence analysis. As described in "Results," sequence discrepancies were detected between the originally published *HE4* sequence (16) and the sequence of HE4 obtained from 4007 and other ovarian carcinoma cell lines, which led us also to sequence *HE4* isolated from epididymis. The corrected sequence was used for construction of fusion proteins.

**Production of HE4Ig Fusion Protein.** A technique introduced by Capon *et al.* (25) was applied to construct fusion proteins with an Ig tail, which were of either mouse (HE4-mIgG) or human (HE4-hIgG) origin and used to immunize mice. The HE4-mIgG 2a and HE4-hIgG1 cDNA constructs were inserted as a *HindIII*-*XbaI* fragment (Fig. 1) into the multiple cloning site of the mammalian expression vector pD18, a derivative of pCDNA 3 as described previously (26). Constructs initially were transfected into COS cells by DEAE-Dextran transient transfections as described (23). Culture supernatants were harvested after 72 h and screened by immunoprecipitation with protein A-agarose, reducing SDS-PAGE electrophoresis, and Western blotting (Fig. 1).

CHO-DG44 cells (27) were used to construct stable lines expressing high levels of the fusion proteins of interest. Stable CHO lines expressing HE4Ig were created by high copy electroporation in the pD18 vector (26, 28) and selection of methotrexate-resistant clones by limiting dilution in Excell 302 CHO media (JRH Biosciences, Denver, PA) containing recombinant insulin (Life Technologies, Inc.), sodium pyruvate (Invitrogen Corp., Carlsbad, CA), L-glutamine (Invitrogen Corp.), 2  $\times$  nonessential amino acids (Invitrogen Corp.), and 100 nM methotrexate (Sigma, St. Louis, MO). Culture supernatants from resistant clones were then assayed by IgG sandwich ELISA to screen for high producing lines. Spent supernatants were harvested from large-scale cultures, and HE4Ig was purified by protein A affinity chromatography, after which the fusion proteins were checked by Western blotting (data not shown). The HE4-hIgG1 fusion protein migrated at an apparent molecular weight of  $M_r$  ~48,000 on reduced gels or Western blots, larger than the  $M_r$  36,000 expected based on the predicted amino acid sequence, suggesting that the molecule was glycosylated. Stable transfectants were used to produce enough protein for immunization of BALB/c mice.

**Generation of Anti-HE4 MAbs.** BALB/c mice were immunized six times against HE4-mIgG, the first time together. Three days after the last immunization (day 62), when the mice had high antibody titers to HE4-hIgG, they were sacrificed, and hybridomas were made by fusing spleen cells with the myeloma partner P3-X63-Ag8-653 (7). Hybridomas 2H5 and 3D8 were obtained and found to identify different epitopes according to competition assays.

**Construction and Application of an ELISA for Tumor Diagnosis.** A double determinant (Sandwich) ELISA was constructed, using a similar approach as that used to make an ELISA that measures mesothelin/*MPF* and related antigens in serum (7) using MAbs 2H5 and 3D8. MAb 2H5 was biotinylated, whereas MAb 3D8 was immobilized by binding to the bottom of the test plates; the respective concentrations of the two MAbs for the assay were 2.5 and 40  $\mu$ g/ml. A standard curve was constructed by using, as antigen, HE4-hIgG diluted in rLAV EIA specimen diluent buffer (Bio-Rad, Hercules, CA). Serial dilution of the samples was performed to measure sensitivity and specificity of the ELISA. Fig. 2 illustrates data obtained by the HE4 capture assay. Samples included serum from an ovarian carcinoma patient (200889), which was used as a positive control; serum from a healthy donor (JR7721), which was used as a negative control; HE4-hIgG fusion protein; serum from the healthy donor spiked with HE4-hIgG fusion protein; and MSLN-hIgG fusion protein, which was used as a second negative control. All samples were diluted in the same manner as the HE4-hIgG. As shown in Fig. 2, a signal was detected at the 160-pg level when HE4-Ig was used as antigen. The ELISA was subsequently used in a blinded study to measure HE4 protein in serum (Fig. 2).

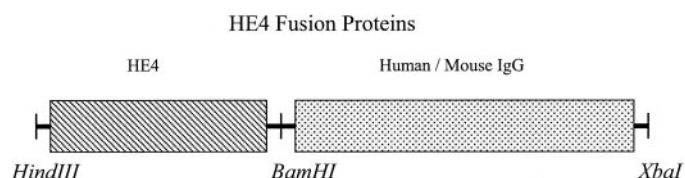


Fig. 1. Fusion protein constructed for immunization of mice.

## HE4 Capture ELISA

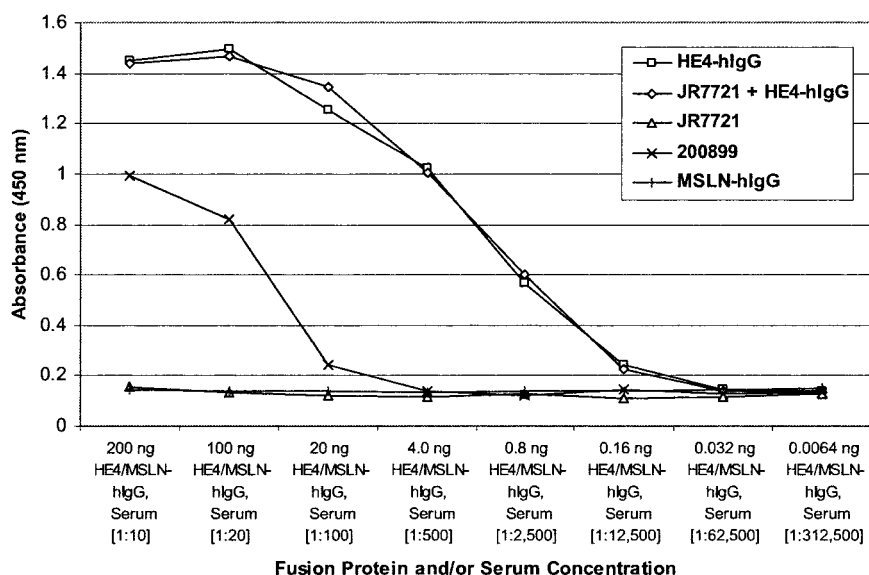


Fig. 2. Illustration of data obtained by the HE4 capture ELISA.

## RESULTS

**Correction of the Originally Deposited *HE4* (*WDCF2*) Gene Sequence.** Results from sequence analysis of the *HE4* gene isolated from ovarian carcinoma line 4007 indicated that the *HE4* coding sequence and deduced amino acid sequence differed at several positions from the published *HE4* coding and translated sequences as deposited in GenBank (15). Sequences were therefore also obtained from cDNAs derived from several ovarian carcinoma cell lines. Clones with inserts were sequenced as described above, and the PCR fragments from the tumor RNA samples were found to encode *HE4* sequence identical to the original clones from the 4007 tumor. All of these *HE4* sequences differed from the published sequence for *HE4* (15). Similarly, *HE4* coding sequence was obtained from normal epididymis and primary tumor tissue cDNAs and matched the new *HE4* sequence described above but differed from the *HE4* sequence of Kirchoff *et al.* (15). The corrected sequence for *HE4*, which has been deposited in GenBank (accession no. AY212888), was used to construct fusion proteins.

**Measurement of HE4 Antigen in Sera from Ovarian Cancer Patients and Controls.** Pilot tests using a reference group of sera from patients with ovarian carcinoma and matched controls indicated that the HE4 protein was elevated in a significant fraction of patients with ovarian carcinoma. Therefore, we performed blinded tests on sera from 37 ovarian cancer patients (7 early stages and 30 late stages), 65 healthy asymptomatic controls, and 19 with benign ovarian disease. All specimens were selected at random from among all postmenopausal subjects contained in the serum repository of an National Cancer Institute-funded ovarian cancer Specialized Programs of Research Excellence program at the FHCRC. Specimens from ovarian carcinoma cases and benign disease were collected before surgery and treatment. Specimens from healthy asymptomatic controls were collected as part of a research trial on ovarian cancer screening of normal risk women. All specimens were processed using the same protocol. Sample sizes were determined to have the power to detect sensitivities as low as 10% at high specificity with the Wilcoxon rank-sum test. Blinding of the status of the specimens was enforced by its Bioinformatics Core. For comparison, all sera were also tested for CA125 in the FHCRC core laboratory using the FDI (formerly Centocor) CA125 II assay.

Of the 37 sera tested, the most common histology was serous ovarian carcinoma ( $n = 21$ ), and the most common stage was stage III ( $n = 24$ ). Sixteen patients had the most common histology and most common stage (stage III serous carcinoma). Other histological types represented are poorly differentiated ( $n = 4$ ), undifferentiated ( $n = 2$ ), adenocarcinoma ( $n = 5$ ), mucinous ( $n = 1$ ), clear cell ( $n = 1$ ), and “mixed histology” ( $n = 3$ ).

Neither CA125 nor HE4 was a perfect predictor of ovarian cancer. Using a 95% specificity threshold for positivity, HE4 failed to identify 7 cases, and CA125 failed to identify 8 cases. For HE4, the false negative cases included serous (3 of 21), undifferentiated (1 of 4), clear cell (1 of 1), undifferentiated (1 of 2), and mixed histology (1 of 3). Cases that were falsely negative when tested with the CA125 assay include serous ovarian carcinoma (3 of 21), poorly differentiated (2 of 4), and undifferentiated carcinoma (3 of 4). Two of 21 serous ovarian carcinomas were missed by both the HE4 and CA 125 assays, as was 1 of 2 undifferentiated carcinomas and 1 of 4 poorly differentiated carcinomas.

Table 1, A–C summarizes the sensitivity achieved by the assay at cutoff levels associated with various correct classification rates, or specificity, among control subjects of different types. The left part of the table shows the specificity achieved, whereas the center and right columns compare the sensitivities of HE4 (center) and CA 125 (right) achieved at these thresholds. Table 1A compares early stage cases to healthy controls, Table 1B compares late stage cases to healthy controls, and Table 1C compares all cases to benign controls, *e.g.*, at a threshold with no misclassification of healthy subjects, or perfect specificity (1.0), the HE4 assay was able to detect 16 of the 30 late stage cases for 53.3% sensitivity (Table 1B). At a threshold that misclassified 3 of 65 controls (96% specificity), HE4 correctly classified 24 of 30 late stage cases (80% sensitivity). Data from the CA125 assay are shown on the right side of the table. For the specificities represented in Table 1A (early stage cases) and Table 1B (late stage cases), CA125 and HE4 have comparable, and not statistically different, sensitivity (Table 1, A–C).

However, when compared with controls having benign ovarian disease, the HE4 assay was slightly better than the CA125 assay in detecting early cases with high specificity (Table 1C). The 17 cases with benign ovarian disease were histologically classified as “muci-

Table 1 Summary of data from blinded assays of sera from patients with ovarian carcinoma and from controls

| A. Early stage cancer versus normal |  |        |  |        |        |        |
|-------------------------------------|--|--------|--|--------|--------|--------|
|                                     | No. of 65 healthy controls exceeding threshold (specificity) |        | No. of 7 early stage cases exceeding threshold (sensitivity) |        |        |        |
|                                     |  |        | HE4  |        | CA 125 |        |
| 1                                   | 24   | (0.60) | 6  | (0.86) | 7      | (1.00) |
| 2                                   | 13   | (0.80) | 6  | (0.86) | 7      | (1.00) |
| 3                                   | 8  | (0.88) | 6  | (0.86) | 7      | (1.00) |
| 4                                   | 6  | (0.92) | 6  | (0.86) | 7      | (1.00) |
| 5                                   | 3  | (0.96) | 6  | (0.86) | 5      | (0.71) |
| 6                                   | 0  | (1.00) | 3  | (0.43) | 2      | (0.29) |

| B. Late stage cancer versus normal |  |        |  |        |        |        |
|------------------------------------|--|--------|--|--------|--------|--------|
|                                    | No. of 65 healthy controls exceeding threshold (specificity) |        | No. of 30 late stage cases exceeding threshold (sensitivity) |        |        |        |
|                                    |  |        | HE4  |        | CA 125 |        |
| 1                                  | 24   | (0.60) | 29   | (0.97) | 28     | (0.93) |
| 2                                  | 13   | (0.80) | 29   | (0.97) | 26     | (0.87) |
| 3                                  | 8  | (0.88) | 28   | (0.93) | 26     | (0.87) |
| 4                                  | 6  | (0.92) | 25   | (0.83) | 25     | (0.83) |
| 5                                  | 3  | (0.96) | 24   | (0.80) | 24     | (0.80) |
| 6                                  | 0  | (1.00) | 16   | (0.53) | 14     | (0.47) |

| C. All cases versus benign disease |   |        |   |        |        |        |
|------------------------------------|---|--------|---|--------|--------|--------|
|                                    | No. of 19 benign controls exceeding threshold (specificity) |        | No. of 37 cases exceeding threshold (sensitivity) |        |        |        |
|                                    |   |        | HE4   |        | CA 125 |        |
| 1                                  | 8   | (0.60) | 29  | (0.97) | 26     | (0.87) |
| 2                                  | 4   | (0.80) | 29  | (0.97) | 24     | (0.80) |
| 3                                  | 3   | (0.88) | 23  | (0.77) | 24     | (0.80) |
| 4                                  | 2   | (0.92) | 21  | (0.70) | 22     | (0.73) |
| 5                                  | 1   | (0.96) | 20  | (0.67) | 12     | (0.40) |
| 6                                  | 0   | (1.00) | 18  | (0.60) | 4      | (0.13) |

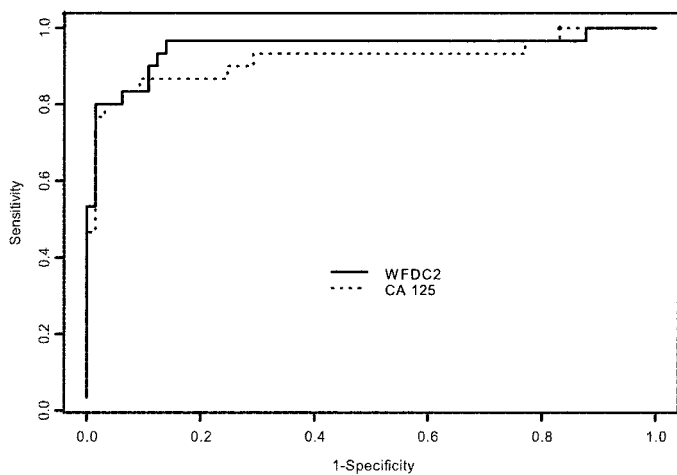


Fig. 3. ROC curves for identifying late stage ovarian cancer from healthy women.

nous” (1), “serous” (8), “non-neoplastic” (3), and other benign histologies (4). Note that endometrioid histology is absent from our benign disease cases, which is the benign histology most often associated with false elevations of CA 125. However, this benign condition can be diagnosed by other means, such as medical history and specific symptomatology, and so it is not as commonly mistaken for ovarian cancer as other benign histologies. Thus, the benign conditions that are most relevant for screening are those nonendometrioid histologies represented among the subjects we tested.

Table 1, A–C gives the sensitivities of each marker at selected specificities, and Figs. 3–5, respectively, summarize the sensitivities at all specificities in ROC curves. These curves display the ability of

CA125 and HE4 to distinguish late stage cancer (Fig. 3) and early stage cancer (Fig. 4) from healthy controls and patients with ovarian cancer from those with benign ovarian disease (Fig. 5). The horizontal axis represents the misclassification rate in the controls (equal to one minus specificity), and the vertical axis represents the true classification rate in the cases (the sensitivity). In this way, the complete tradeoff between sensitivity and specificity is displayed. When comparing two ROC curves, the marker achieving the highest curve (highest sensitivity) can be considered superior at that misclassification rate (Figs. 3–5).

It is also common to summarize an ROC curve by the AUC (29, 30). A perfect ROC curve has AUC = 1 (or 0), and a useless marker has AUC = 0.5. We performed statistical tests of markers using the Wilcoxon rank-sum test, because it is equivalent to testing whether its ROC curve has AUC = 0 and so is appropriate for testing diagnostic markers (31). Table 2 reports the AUC summarizing the ability of HE4 and CA 125 to distinguish cases from healthy controls and cases from benign controls. Each AUC reported in the table is statistically significant as measured by the Wilcoxon rank-sum test, meaning that both markers have the ability to discriminate cases from the respective controls (Table 2).

When using normal subjects as the comparison group for all of the cases, the AUC values for HE4 and CA125 are very similar (Figs. 4 and 5 and Table 2). The ROC curves displayed have the following AUC: (a) HE4 = 0.96, CA125 = 0.95 (Fig. 4) and (b) HE4 = 0.85,

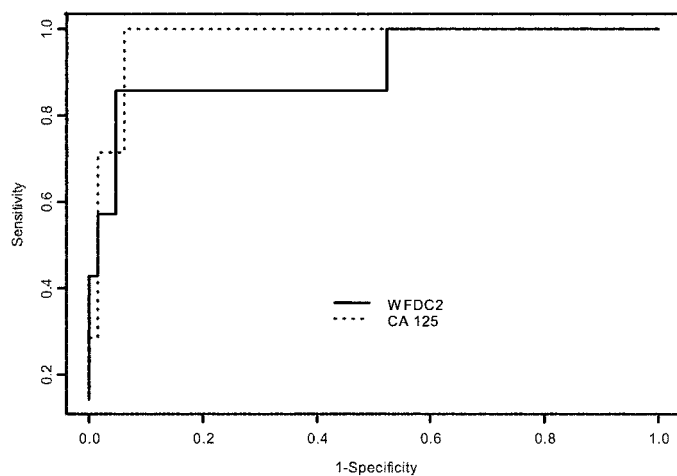


Fig. 4. ROC curves for identifying early stage ovarian cancer from healthy women.

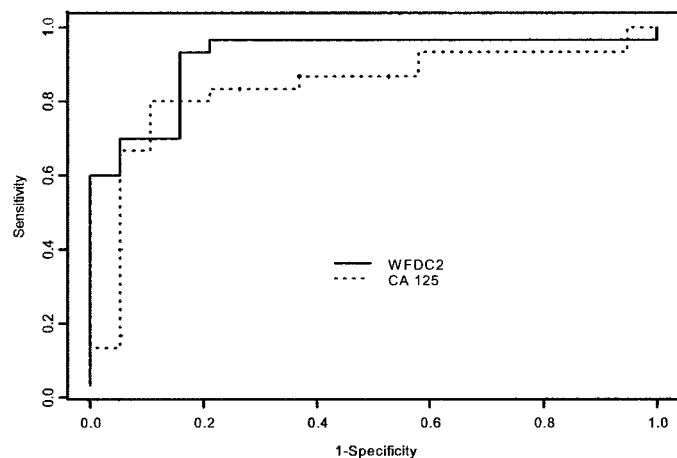


Fig. 5. ROC curves for identifying all cases of ovarian cancer from women with benign disease.

Table 2 Summary of AUC values for HE4 and CA 125 for discriminating between cases and various controls<sup>a</sup>

| Cases            | HE4 Assay       |                | CA 125 Assay    |                |
|------------------|-----------------|----------------|-----------------|----------------|
|                  | Healthy control | Benign control | Healthy control | Benign control |
| All              | 0.95            | 0.90           | 0.94            | 0.84           |
| Early stage only | 0.85            | 0.86*          | 0.90            | 0.80*          |
| Late stage only  | 0.96            | 0.91           | 0.95            | 0.85           |

<sup>a</sup> All AUC values except \* (with  $P < 0.01$ ) are statistically significant and have  $P < 0.001$ , by the Wilcoxon rank-sum test.

CA125 = 0.9 (Fig. 5). Although the AUC for HE4 exceeds that for CA125, the difference between them is not statistically significant ( $P = 0.76$ ). Thus, by the ROC curves and statistical test, we conclude that HE4 and CA125 have overall comparable ability to discriminate ovarian cancer cases from normal subjects.

When using benign controls as the comparison group for all cases, the overall ROC curves for HE4 and CA125 are similar along most of the ROC curves (Fig. 5 and Table 2). The AUC for the curves in Fig. 5 are 0.9 for HE4 and 0.84 for CA125, but the greater overall AUC for HE4 is not statistically significant ( $P = 0.27$ ). The AUC summarizes the overall ROC curve, but the performance at very high specificity, especially for benign ovarian disease, is most relevant for cancer screening. Indeed, Jacobs (32) reports that for ovarian cancer screening, a marker will need to achieve a specificity >99% to achieve clinical relevance. Therefore, in addition to comparing the overall ROC curves, it is also particularly relevant to compare markers at highest specificity. Of the 37 cases, only 4 (10.8% sensitivity) have CA125 levels exceeding the maximum found among the benign controls (100% specificity), whereas HE4 identified these 4 and 17 others, for a total of 21 of 37 cases (56.7% sensitivity). This difference is statistically significant ( $P = 0.001$  by McNamara test) and is indicative of HE4 having a better ability to discriminating between women with malignant and benign disease.

## DISCUSSION

Most existing antitumor MAbs have been obtained from mice that were immunized with human tumors, including MAbs to MSLN (5, 6), which is a promising new marker for ovarian carcinoma (7). However, the probing of large numbers of genes with microarray technology for amplification in neoplastic *versus* normal cells and generation of MAbs to the products of selectively amplified genes can dramatically facilitate the identification of novel tumor markers.

Although the *HE4* (WDFC2) gene was cloned already in 1991 (15), it was not until microarray analyses had been performed that its association with ovarian carcinoma became known (8, 9).<sup>6</sup> Further support for the microarray approach comes from the demonstration that the MSLN genes were also amplified in ovarian carcinoma. This agrees with the finding that MAbs to MSLN have high and relatively selective binding to cells from ovarian carcinoma (7).<sup>5</sup>

Rapid and reliable methods to generate and characterize MAbs to the products of newly discovered tumor-selective genes will facilitate their evaluation as diagnostic markers and therapeutic targets. Construction of fusion proteins with Ig tails has two advantages to facilitate the generation of MAbs: (a) proteins with a mouse Ig tail are likely to be taken up and processed by antigen-presenting cells to induce an antibody response; and (b) screening of hybridoma supernatants for binding to the same fusion proteins that instead have a human Ig tail facilitates the identification of hybridomas that make antibodies to the respective antigen.

We performed a blinded study on sera from patients with ovarian carcinoma (late or early) or who have benign ovarian disease with sera

from matched healthy controls. In parallel, we also determined the CA125 levels in the sera. Although the number of sera from patients with early stage disease was small, the findings indicate that HE4 is a potentially useful biomarker for ovarian carcinoma, comparable with CA125 in distinguishing women with ovarian cancer, both localized and advanced, from healthy women. Results indicate that HE4 is better than CA125 in distinguishing patients with malignant ovarian disease from those with benign ovarian disease at high specificity.

Future studies will be performed to evaluate HE4 as a diagnostic/prognostic marker for ovarian carcinoma. These studies will include larger samples of sera from patients with early ovarian carcinoma, benign disease, and high-risk women participating in ovarian cancer screening trials.

## REFERENCES

- Ries, L. A. G., Kosary, C. L., Hankey, B. F., Miller, B. A., Clegg, L., and Edwards, B. K. SEER Cancer Statistics Review. National Cancer Institute.
- Jacobs, I., and Bast, R. C. The CA-125 tumour associated antigen: a review of the literature. *Hum. Reprod.*, 4: 1-12, 1989.
- Berek, J. S., and Bast, R. C., Jr. Ovarian cancer screening. The use of serial complementary tumor markers to improve sensitivity and specificity for early detection. *Cancer*, 76: 2092-2096, 1995.
- Maggino, T., Gadducci, A., D'Addario, V., Pecorelli, S., Lissoni, A., Stella, M., Romagnolo, C., Federghini, M., Zucca, S., Trio, D., *et al.* Prospective multicenter study on CA 125 in postmenopausal pelvic masses. *Gynecol. Oncol.*, 54: 117-123, 1994.
- Chang, K., Pai, L. H., Batra, J. K., Pastan, I., and Willingham, M. C. Characterization of the antigen (CAK1) recognized by monoclonal antibody K1 present on ovarian cancers and normal mesothelium. *Cancer Res.*, 52: 181-186, 1992.
- Kojima, T., Oh-eda, M., Hattori, K., Taniguchi, Y., Tamura, M., Ochi, N., and Yamaguchi, N. Molecular cloning and expression of megakaryocyte potentiating factor cDNA. *J. Biol. Chem.*, 270: 21984-21990, 1995.
- Scholler, N., Fu, N., Yang, Y., Ye, Z., Goodman, G., Hellstrom, K., and Hellstrom, I. Soluble member(s) of the mesothelin/megakaryocyte potentiating factor family are detectable in sera from patients with ovarian carcinoma. *Proc. Natl. Acad. Sci. USA*, 96: 11531-11536, 1999.
- Schummer, M., Ng, W., Bumgarner, R., Nelson, P., Schummer, B., Bednarski, D., Hassell, L., Baldwin, R., Karlan, B., and Hood, L. Comparative hybridization of an array of 21 500 ovarian cDNAs for the discovery of genes overexpressed in ovarian carcinomas. *Gene*, 238: 375-385, 1999.
- Wang, K., Gan, L., Jeffrey, E., Gayle, M., Gown, A. M., Skelly, M., Nelson, P. S., Ng, W., Schummer, M., Hood, L., and Mulligan, J. Monitoring gene expression profile changes in ovarian carcinomas using cDNA microarray. *Gene*, 229: 101-108, 1999.
- Hough, C. D., Sherman-Baust, C. A., Pizer, E. S., Montz, F. J., Im, D. D., Ronsenshein, N. B., Cho, K. R., Riggins, G. J., and Morin, P. J. Large-scale serial analysis of gene expression reveals genes differentially expressed in ovarian cancer. *Cancer Res.*, 60: 6281-6287, 2000.
- Ismail, R. S., Baldwin, R. L., Fang, J., Browning, D., Karlan, B. Y., Gasson, J. C., and Chang, D. D. Differential gene expression between normal and tumor-derived ovarian epithelial cells. *Cancer Res.*, 60: 6744-6749, 2000.
- Martoglio, A. M., Tom, B. D., Starkey, M., Corps, A. N., Charnock-Jones, D. S., and Smith, S. K. Changes in tumorigenesis- and angiogenesis-related gene transcript abundance profiles in ovarian cancer detected by tailored high density cDNA arrays. *Mol. Med.*, 6: 750-765, 2000.
- Ono, K., Tanaka, T., Tsunoda, T., Kitahara, O., Kihara, C., Okamoto, A., Ochiai, K., Takagi, T., and Nakamura, Y. Identification by cDNA microarray of genes involved in ovarian carcinogenesis. *Cancer Res.*, 60: 5007-5011, 2000.
- Welsh, J. B., Zarrinkar, P. P., Sapinoso, L. M., Kern, S. G., Behling, C. A., Monk, B. J., Lockhart, D. J., Burger, R. A., and Hampton, G. M. Analysis of gene expression profiles in normal and neoplastic ovarian tissue samples identifies candidate molecular markers of epithelial ovarian cancer. *Proc. Natl. Acad. Sci. USA*, 98: 1176-1181, 2001.
- Kirchhoff, C., Habben, I., Ivell, R., and Krull, N. A major human epididymis-specific cDNA encodes a protein with sequence homology to extracellular proteinase inhibitors. *Biol. Reprod.*, 45: 350-357, 1991.
- Kirchhoff, C. Molecular characterization of epididymal proteins. *Rev. Reprod.*, 3: 86-95, 1998.
- Dandekar, A. M., Robinson, E. A., Appella, E., and Quasba, P. K. Complete sequence analysis of cDNA clones encoding rat whey phosphoprotein: homology to a protease inhibitor. *Proc. Natl. Acad. Sci. USA*, 79: 3987-3991, 1982.
- Heinzel, R., Appelhaus, H., Gassen, G., Seemuler, U., Machleidt, W., Fritz, H., and Steffens, G. Molecular cloning and expression of cDNA for human antileukoprotease from cervix uterus. *Eur. J. Biochem.*, 160: 61-67, 1986.
- Wiedow, O., Schroder, J. M., Gregory, H., Young, J. A., and Christophers, E. Elafin: an elastase-specific inhibitor of human skin. Purification, characterization, and complete amino acid sequence. *J. Biol. Chem.*, 265: 14791-14795, 1990.
- Bingle, L., Singleton, V., and Bingle, C. D. The putative ovarian tumour marker gene HE4 (WDFC2), is expressed in normal tissues and undergoes complex alternative splicing to yield multiple protein isoforms. *Oncogene*, 21: 2768-2773, 2002.

21. Hayden, M. S., Gilliland, L. K., and Ledbetter, J. A. Antibody engineering. *Curr. Opin. Immunol.*, *9*: 201–212, 1997.
22. Kirchoff, C., and Willison, K. Nucleotide and amino-acid sequence of human testis-derived TCP1. *Nucleic Acids Res.*, *18*: 4247, 1990.
23. Hayden, M. S., Linsley, P. S., Gayle, M. A., Bajorath, J., Brady, W. A., Norris, N. A., Fell, H. P., Ledbetter, J. A., and Gilliland, L. K. Single-chain mono- and bispecific antibody derivatives with novel biological properties and antitumour activity from a COS cell transient expression system. *Ther. Immunol.*, *1*: 3–15, 1994.
24. Hellstrom, I., Goodman, G., Pullman, J., Yang, Y., and Hellstrom, K. E. Overexpression of HER-2 in ovarian carcinomas. *Cancer Res.*, *61*: 2420–2423, 2001.
25. Capon, D. J., Chamow, S. M., Mordenti, J., Marsters, S. A., Gregory, T., Mitsuya, H., Byrn, R. A., Lucas, C., Wurm, F. M., Groopman, J. E., *et al.* Designing CD4 immunoadhesins for AIDS therapy. *Nature*, *337*: 525–531, 1989.
26. Hayden, M. S., Grosmaire, L. S., Norris, N. A., Gilliland, L. K., Winberg, G., Tritschler, D., Tsu, T. T., Linsley, P. S., Mittler, R. S., Senter, P. D., Fell, H. P., and Ledbetter, J. A. Costimulation by CD28 sFv expressed on the tumor cell surface or as a soluble bispecific molecule targeted to the L6 carcinoma antigen. *Tissue Antigens*, *48*: 242–254, 1996.
27. Urlaub, G., Mitchell, P. J., Kas, E., Chasin, L. A., Funanage, V. L., Myoda, T. T., and Hamlin, J. Effect of gamma rays at the dihydrofolate reductase locus: deletions and inversions. *Somatic Cell Mol. Genet.*, *12*: 555–566, 1986.
28. Barsoum, J. Introduction of stable high-copy-number DNA into Chinese hamster ovary cells by electroporation. *DNA Cell Biol.*, *9*: 293–300, 1990.
29. Hanley, J. A., and McNeil, B. J. The meaning and use of the area under the receiver operating characteristic (roc) curve. *Radiology*, *143*: 23–29, 1982.
30. Pepe, M. S. An interpretation for the ROC curve and inference using GLM procedures. *Biometrics*, *56*: 352–359, 2000.
31. Pepe, M. S. Receiver operating characteristics methodology. *J. Am. Stat. Assoc.*, *95*: 308–311, 2000.
32. Jacobs, I. Overview-progress in screening for ovarian cancer. *In*: F. Sharp, T. Blackett, J. Berek, and R. Bast (eds.), *Ovarian Cancer*, Vol. 5, pp. 173–185. Oxford: ISIS Medical Media, 1998.

# Cancer Research

The Journal of Cancer Research (1916–1930) | The American Journal of Cancer (1931–1940)

## The HE4 (WFDC2) Protein Is a Biomarker for Ovarian Carcinoma

Ingegerd Hellström, John Raycraft, Martha Hayden-Ledbetter, et al.

*Cancer Res* 2003;63:3695-3700.

**Updated version** Access the most recent version of this article at:  
<http://cancerres.aacrjournals.org/content/63/13/3695>

**Cited articles** This article cites 31 articles, 11 of which you can access for free at:  
<http://cancerres.aacrjournals.org/content/63/13/3695.full#ref-list-1>

**Citing articles** This article has been cited by 51 HighWire-hosted articles. Access the articles at:  
<http://cancerres.aacrjournals.org/content/63/13/3695.full#related-urls>

**E-mail alerts** [Sign up to receive free email-alerts](#) related to this article or journal.

**Reprints and Subscriptions** To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at [pubs@aacr.org](mailto:pubs@aacr.org).

**Permissions** To request permission to re-use all or part of this article, use this link  
<http://cancerres.aacrjournals.org/content/63/13/3695>.  
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.