

Article

## **HE4 Transcription- and Splice Variants-Specific Expression in Endometrial Cancer and Correlation with Patient Survival**

**Shi-Wen Jiang**<sup>1,2,3,†</sup>, **Haibin Chen**<sup>4,†</sup>, **Sean Dowdy**<sup>2</sup>, **Alex Fu**<sup>1</sup>, **John Attewell**<sup>5</sup>,  
**Eleftheria Kalogera**<sup>2</sup>, **Ronny Drapkin**<sup>6</sup>, **Karl Podratz**<sup>2</sup>, **Russell Broaddus**<sup>7</sup> and **Jinping Li**<sup>1,2,3,\*</sup>

<sup>1</sup> Department of Biomedical Science, School of Medicine, Mercer University, Savannah, GA 31404, USA; E-Mails: jiang\_s@mercer.edu (S.-W.J.); alex.a.fu@gmail.com (A.F.)

<sup>2</sup> Division of Gynecologic Surgery, Mayo Clinic, Rochester, MN 55902, USA; E-Mails: dowdy.sean@mayo.edu (S.D.); kalogera.eleftheria@mayo.edu (E.K.); podratz.karl@mayo.edu (K.P.)

<sup>3</sup> Department of Laboratory Oncology Research, Curtis and Elizabeth Anderson Cancer Institute, Memorial University Medical Center, Savannah, GA 31404, USA

<sup>4</sup> Department of Histology and Embryology, Shantou University Medical College, Shantou 515041, Guangdong, China; E-Mail: chenhb@stu.edu.cn

<sup>5</sup> Healthcare Analytics, Mayo Clinic, Rochester, MN 55902, USA; E-Mail: john.attewell@charter.net

<sup>6</sup> Department of Pathology, Harvard Medical School, Boston, MA 02115, USA; E-Mail: ronny\_drapkin@dfci.harvard.edu

<sup>7</sup> Department of Pathology, the University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA; E-Mail: rbroadus@mdanderson.org

† These authors contributed equally to this work.

\* Author to whom correspondence should be addressed; E-Mail: li\_j@mercer.edu; Tel.: +1-912-350-1732; Fax: +1-921-350-1281.

Received: 16 June 2013; in revised form: 14 October 2013 / Accepted: 15 October 2013 /

Published: 18 November 2013

---

**Abstract:** We investigated the *HE4* variant-specific expression patterns in various normal tissues as well as in normal and malignant endometrial tissues. The relationships between mRNA variants and age, body weight, or survival are analyzed. ICAT-labeled normal and endometrial cancer (EC) tissues were analyzed with multidimensional liquid chromatography followed by tandem mass spectrometry. Levels of *HE4* mRNA variants were measured by real-time PCR. Mean mRNA levels were compared among 16 normal endometrial samples, 14 grade 1 and 14 grade 3 endometrioid EC, 15 papillary serous EC,

and 14 normal human tissue samples. The relationship between levels of *HE4* variants and EC patient characteristics was analyzed with the use of Pearson correlation test. We found that, although all five *HE4* mRNA variants are detectable in normal tissue samples, their expression is highly tissue-specific, with epididymis, trachea, breast and endometrium containing the highest levels. *HE4*-V0, -V1, and -V3 are the most abundant variants in both normal and malignant tissues. All variants are significantly increased in both endometrioid and papillary serous EC, with higher levels observed in grade 3 endometrioid EC. In the EC group, *HE4*-V1, -V3, and -V4 levels inversely correlate with EC patient survival, whereas *HE4*-V0 levels positively correlate with age. *HE4* variants exhibit tissue-specific expression, suggesting that each variant may exert distinct functions in normal and malignant cells. *HE4* levels appear to correlate with EC patient survival in a variant-specific manner. When using *HE4* as a biomarker for EC management, the effects of age should be considered.

**Keywords:** *HE4*; transcription; splice variants; gene expression; endometrial cancer

---

## 1. Introduction

Endometrial cancer (EC) is the most common gynecologic malignancy and the fourth most common cancer in females, ranking behind breast, lung and colorectal cancers in the United States. While early stage EC has a favorable prognosis, late stage EC is associated with a relatively high rate of post-treatment recurrence and mortality. The importance of early diagnosis is underscored by the dramatic reduction of five-year survival rates from 96% to 67% and further to 17% for localized, regional, and metastatic disease, respectively [1]. Several studies have evaluated the efficacy of various tumor markers, including CA125, for EC early detection and management, but only limited specificities and sensitivities was observed [2–4].

*HE4* was first cloned as one of four proteins highly expressed in human epididymis [5]. Amino acid sequence analysis suggested that *HE4* belongs to a whey-acidic-protein (WAP) family and contains two four-disulfide core domains (WFDC2). *HE4* gene is located on chromosome 20q13.12, a region harboring a cluster of genes for WAP domain proteins [6–8]. Although it is well known that at least some members of this family serve as endogenous peptide protease inhibitors, the function of *HE4* remains obscure. Using the cDNA microarray technique Schummer *et al.* found that *HE4* is overexpressed in ovarian cancer (OC) [9]. Bingle *et al.* used Northern blot hybridization in order to characterize its tissue-specific expression and found that *HE4* mRNA is abundant in lung, kidney, and the salivary gland [8]. Lu *et al.* performed oligo-array analysis and observed that *HE4* is overexpressed in OC in comparison to normal ovarian epithelium [10]. Drs. Drapkin and Galgano *et al.* analyzed the expression of *HE4* in a series of normal and malignant tissues and observed high levels of *HE4* mRNA in the trachea and salivary gland [11,12]. Tissue microarray studies indicate that *HE4* protein is present in both normal and malignant ovarian and endometrial cells [10–13].

Bingle *et al.* compared the sequences of two lung-derived *HE4* ESTs (expressed sequence tags) with those of the prototypic *HE4* cDNA, leading to the identification of five distinct *HE4* mRNA variants [8]. Four of these variants (*HE4*-V0, -V1, -V2, -V4) are generated by differential splicing

mechanisms whereas one (*HE4-V3*) utilizes an alternative promoter as suggested by the presence of an extended 5' prime region. The disparate structures and domain arrangements among the *HE4* variants raised questions on their tissue-specific expression and potentially distinct cellular functions.

Being a small, secretory glycoprotein [11], HE4 is readily detectable in plasma and the serum levels of HE4 are considered a potential biomarker for ovarian cancer diagnosis and management [13–16]. An antibody-based ELISA has been developed to measure HE4 serum levels in various clinical settings [17–19]. It has been reported that HE4 alone or in combination with CA125 has higher sensitivity than any other single marker or combination of markers, in early detection [20–22], diagnosis [15,23–26], risk stratification [27] and prediction of postsurgical outcome [10,20,24,28] for patients with OC. With regards to EC, Moore *et al.* found that HE4 serum levels are consistently elevated in all stages of EC and that it is more sensitive in detecting early-stage disease than CA125 [29]. Bignotti *et al.* assessed the diagnostic and prognostic value of HE4 in EC. In their study, a significant correlation was found between increased HE4 serum levels and aggressive EC phenotype [30]. Based on these findings, HE4 test has been proposed as a valuable tool for the triage of patients presenting with a pelvic mass [31–33]. It is noteworthy that, despite the data in favor of HE4 protein levels as a useful biomarker, negative results from several studies challenged the predictive value of the ROMA algorithm when HE4 was used alone or in combination with CA125 [34,35].

While most studies suggest that HE4 is a clinically useful biomarker in EC and OC, the assays used to measure mRNA or protein levels were designed based on the structure of the prototype HE4, designated as HE4-V0 in this report. In previous studies, no distinction was made among the multiple structurally and potentially functionally diverged HE4 variants. The goal of this study is to compare HE4 mRNA variants-specific expression in normal and malignant endometrial tissues, and to analyze the relationship between HE4 variant-specific mRNA levels and clinical features including age, surgical stage and survival.

## 2. Results

### 2.1. HE4 Protein Is Overexpressed in Endometrial Cancer

Six pairs of matched normal and endometrial cancer tissue samples were prepared and analyzed on mass spectrometry. Figure S1 shows representative chromatograms for LC and mass spectrometry identification and comparison of HE4 levels between EC and normal endometrium samples. Thousands of proteins were identified from each of these pairs (3006, 2826, 2434, 2389, 1925, and 1803 for pairs 1 to 6, respectively; 75% confidence level). A more in-depth analysis of the data resulted in the identification of a group of 77 proteins which presented a greater than 1.5-fold changes in their expression levels between normal and cancer tissue in at least two out of six pairs of samples [36]. These proteins were found to be involved in a broad spectrum of cell functions including cellular structure determination, cell signaling, transcription control, as well as cell cycle and apoptosis regulation (Figure S1). Table 1 presents data on 8 representative proteins out of the 77 that exhibiting altered expression in endometrioid cancers. In accordance with previous studies [24,30], HE4 was found to be one of the 27 proteins significantly up-regulated in EC (Figure S2).

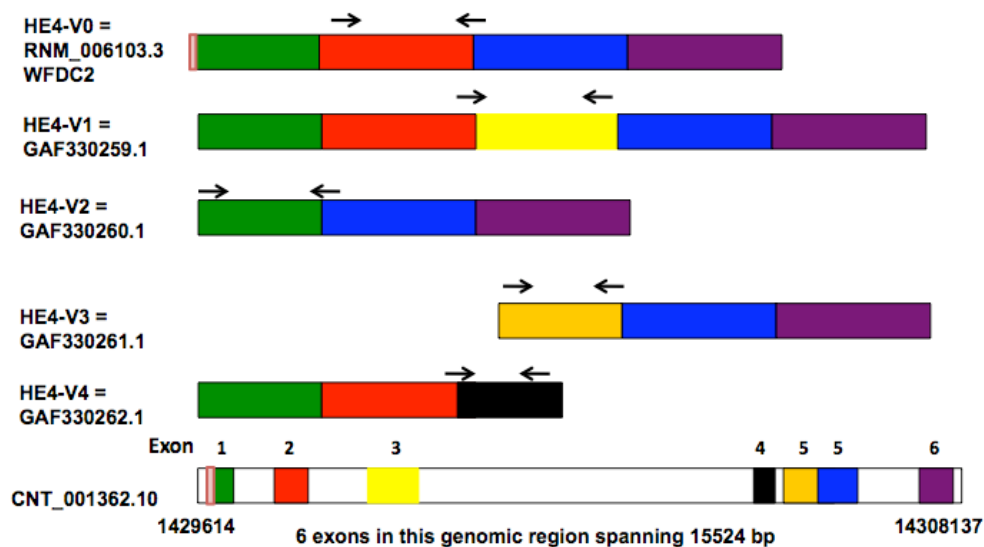
**Table 1.** Representative proteins identified by Isotope-Coded Affinity Tag (ICAT)-mass spectrometry. The table provides basic information on accession number (GI number is NCBI/PubMed protein ID number), protein identity, cellular function, numbers of paired-samples in which the protein was detected (Presence), the average ratio between endometrioid cancers and normal controls (Mean), and the standard errors (STE).

| Accession ID number | Protein identity   | Function   | Presence | Peptide | Mean | STE  |
|---------------------|--|--|----------|---------|------|------|
| GI: 32307109        | HE4, WAP four-disulfide core domain protein 3 precursor, WFDC2 | Male fertility function. Putative ovarian tumor marker       | 2        | 1       | 2.45 | 1.23 |
| GI: 3859849         | LIM protein SLIMMER  | Intracellular signaling pathway                              | 3        | 1       | 2.39 | 1.11 |
| GI: 15779184        | FLNA protein, Filamin  | Actin binding, cell structure and mobility                   | 4        | 1       | 2.00 | 0.57 |
| GI: 4504165         | Gelsolin   | Calcium-regulated, actin modulation; Silenced in tumor cells | 2        | 2       | 0.52 | 0.07 |
| GI: 86947           | Galectin 1, beta-galactoside-binding lectin, placental         | Cell differentiation and apoptosis                           | 5        | 2       | 0.25 | 0.05 |
| GI: 3193336         | DBI-related protein  | Fatty acid metabolism  | 2        | 1       | 0.54 | 0.20 |
| GI: 37544577        | Hypothetical protein XP_294867                                 | Unknown  | 5        | 4       | 1.60 | 0.13 |
| GI: 27545315        | Translational activator of cytochrome c oxidase 1              | Translational activator                                      | 5        | 10      | 1.58 | 0.21 |

## 2.2. Design of *HE4* Variant-Specific PCR Primers

We took advantage of the distinct intra-exon sequences and exon-exon boundaries within the *HE4* variants to design variant-specific PCR primers (Figure 1). The designations and sequences of these PCR primers are listed in Table 2. PCR conditions were extensively optimized to ensure specific and linear amplification. Aliquots of the real-time PCR products were resolved electrophoretically in agarose gels and subsequently visualized with ethidium bromide staining. A single band pattern with a predicted size indicated that specific amplification had been achieved (Figure 2).

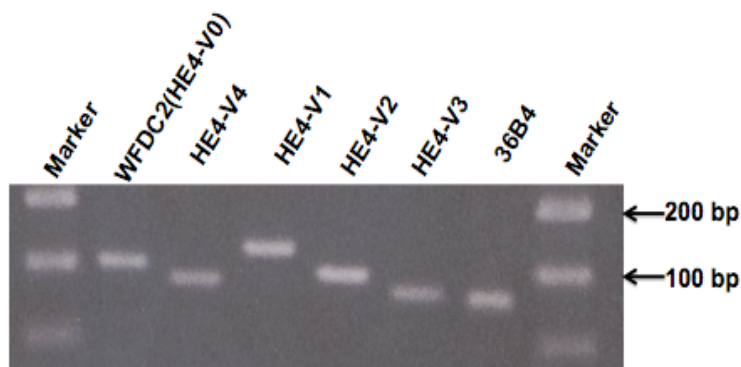
**Figure 1.** Schematic outline of real-time PCR primer design for each *HE4* variant. The structures of *HE4* variants were illustrated with the six exons marked in different colors. The locations of real-time PCR primers were indicated by arrows. On the bottom the *HE4* genomic structure of *HE4* prototype (*HE4*-V0) (NM\_006103.3, GeneID: 10406) was shown based on the information from NCBI website [37]. The *HE4* gene is located at 20q13.12, spanning 1429614nt to 14308137nt in NM\_006103.3.



**Table 2.** Real-time PCR primers for *HE4* transcript variants. The sequences and designation of PCR primers for real-time PCR measurement of *HE4* transcript variant were listed.

| Primer name                    | Primer sequence                       |
|--------------------------------|---------------------------------------|
| WFDC2 = <i>HE4</i> -V0 Forward | 5'-CCG ACA ACC TCA AGT GCT G-3'       |
| WFDC2 = <i>HE4</i> -V0 Reverse | 5'-CGA GCT GGG GAA AGT TAA TG-3'      |
| <i>HE4</i> -V1 Forward         | 5'-AAT GGC CAA CTG GCT GAG-3'         |
| <i>HE4</i> -V1 Reverse         | 5'-TTT GAG AGA GTC CCC AGC TC-3'      |
| <i>HE4</i> -V2 Forward         | 5'-CCA TGC CTG CTT GTC GCC-3'         |
| <i>HE4</i> -V2 Reverse         | 5'-CAG GAA CCC TCC TTA TCT GA-3'      |
| <i>HE4</i> -V3 Forward         | 5'-GCC ATG CTG CAG GTA CAA GT-3'      |
| <i>HE4</i> -V3 Reverse         | 5'-ATC TGG GTA GAA AAA GGA GTA AGG-3' |
| <i>HE4</i> -V4 Forward         | 5'-CCC AAT GCA CTG TTC CAC T-3'       |
| <i>HE4</i> -V4 Reverse         | 5'-AGT CCC AAG TGG GCC TTC-3'         |
| 36B4 Forward                   | 5'-ATG CAG CAG ATC CGC ATG-3'         |
| 36B4 Reverse                   | 5'-TCA TGG TGT TCT TGC CCA TCA-3'     |

**Figure 2.** Gel electrophoresis of real-time PCR products. Real-time PCR products of the five *HE4* transcript variants are obtained from normal endometrial specimen, and separated in 2% agarose gel. The real-time PCR product size for each variant was: for *HE4-V0*, 112 bp; *HE4-V1*, 131 bp; *HE4-V2* 96 bp; *HE4-V3*, 74 bp; *HE4-V4*, 84 bp. *36B4* (63 bp) was used as an internal control. The single band pattern indicates specific amplification of cDNA by real-time PCR.

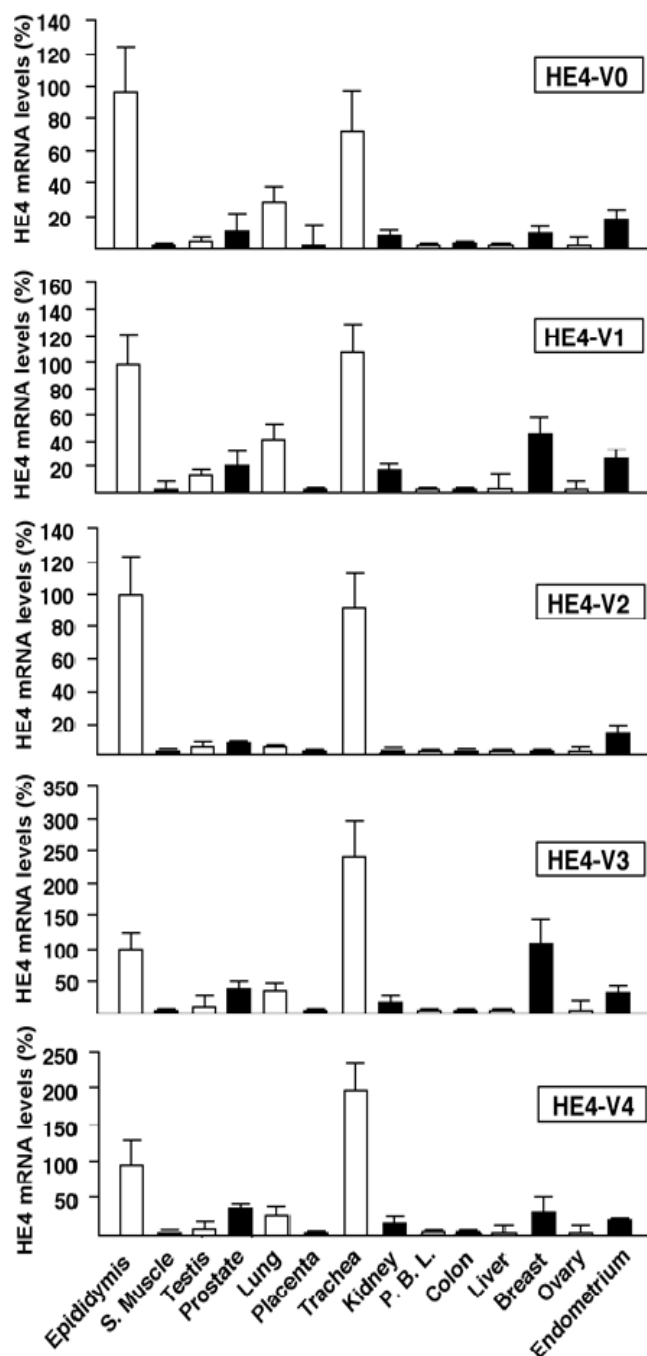


### 2.3. Expression of *HE4* mRNA Variants in Normal Tissues

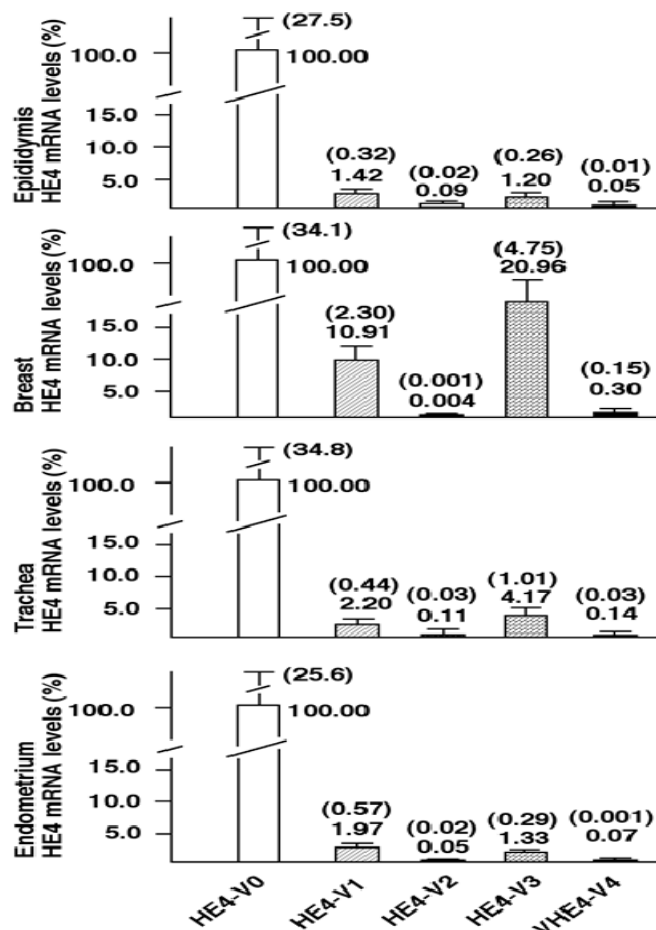
To obtain general knowledge on the *HE4* mRNA variants distribution, we first compared by real-time PCR the levels of their expression in 14 normal human tissues, including epididymis, skeletal muscle (s. muscle), testis, prostate, lung, placenta, trachea, kidney, peripheral blood lymphocyte (PBL), colon, liver, breast, ovary and endometrium. Relative mRNA levels of *HE4* variants are shown in Figure 3, which indicates a high degree of variability across different tissues. Overall, epididymis and trachea express the highest levels of *HE4* mRNA variants followed by breast, endometrium, lung and prostate, which expressed intermediate levels. Interestingly, this is consistent with Dr. Drapkin *et al.*'s previous finding [11]. The remaining tissues tested express only low to moderate levels of *HE4* mRNA variants. Different tissues express diverged levels of *HE4* variants. For example, the highest levels of *HE4-V0* and -*V2* were found in the epididymis, whereas the highest levels of *HE4-V1*, -*V3* and -*V4* were in the trachea. *HE4-V3* is also expressed at a relatively high level in breast tissue.

In order to compare horizontally the levels of each variant across various tissues, the data was rearranged and presented by tissues. Figure 4 shows that the *HE4-V0* to -*V4* distribution in normal endometrium, breast, epididymis, and trachea. The distribution in the rest 10 types of tissues was shown in the supplemental data Figure S3. *HE4-V0* represents the most abundant variant, followed by *HE4-V1* and -*V3* at approximately 10- to 100-fold lower levels than *HE4-V0*. Diminished levels of *HE4-V2* and -*V4*, at approximately 100- to 1000-fold lower than *HE4-V0*, were detected in all the tissues.

**Figure 3.** HE4 variant expression levels in normal tissues. Total RNA was extracted from 14 types of normal human tissues: epididymis, skeletal muscle (S. Muscle), testis, prostate, lung, placenta, trachea, kidney, peripheral blood lymphocyte (PBL), colon, liver, breast, ovary and endometrium. Reverse transcription and real-time PCR were performed on these samples. The expression levels of each variant in different tissues were normalized by the level of epididymis (set as 100). The mRNA levels of HE4 were ranked in three groups: <5% was ranked as low, 5%–50% was ranked as moderate, >50% was ranked as high. In general, all the variants were expressed at the highest levels in epididymis or trachea, moderate levels in prostate, lung, breast and endometrium, and very low levels in skeletal muscle, testis, placenta, kidney, PBL, colon and liver.



**Figure 4.** Comparison of *HE4* variant expression levels in normal tissues. The data from Figure 3 was re-set to best describe the *HE4* variant expression levels in epididymis, breast, trachea and endometrium. The relative expression levels were converted based on *HE4-V0* (set as 100). Average and standard errors were indicated by the number on the top of bars.



To further analyze the tissue/organ-specific expression patterns of the 5 *HE4* variants, we performed regression analysis in the 13 normal tissues/organs. Similar expression patterns of *HE4* variants were found among the majority of the tissue types, with correlation coefficients all exceeding 0.88 and *p* values small than 0.05 (Table 3). Thus, the results indicated a significant inter-correlation among the different *HE4* variants (*N* = 5) in those normal tissues. Interestingly, no significant correlation in the *HE4* variant expression patterns could be established between the colon and several tissues such as ovary, PBL, placenta, kidney and prostate. Consequently, the colon appears to represent a special tissue with a unique *HE4* variant expression pattern when compared to other human tissues. It is intriguing to observe this highly consistent *HE4* variant expression patterns among a variety of human tissues, and at the same time, a dramatic divergence between colon and other tissues. The regulatory mechanism and functional significance of this phenomenon is not clear at this time.

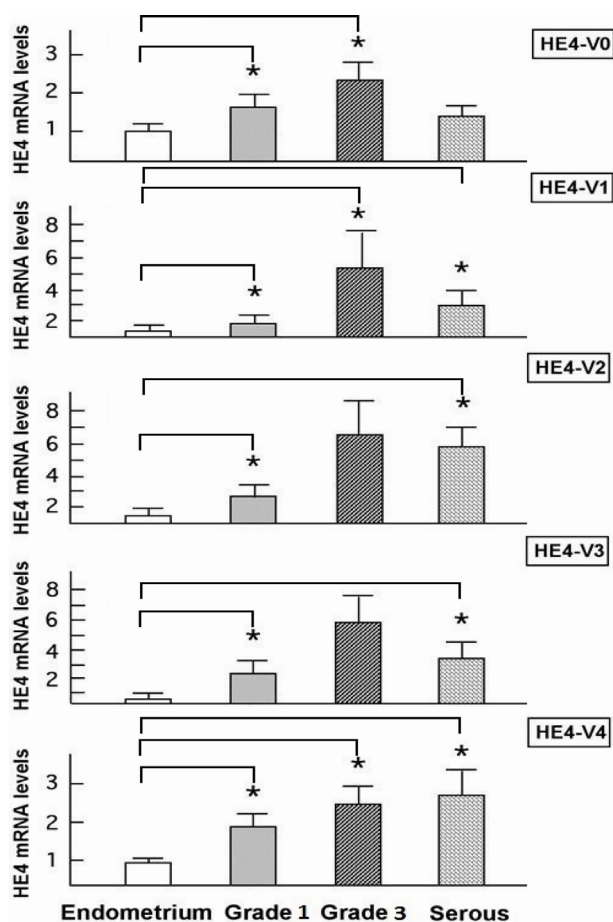




#### 2.4. Increased Levels of HE4 mRNA Variants in Endometrial Cancer

Using the same set of primers we performed real-time PCR to compare the *HE4* variant-specific expression between EC and normal endometrium tissues (Figure 5). The mRNA levels of HE4-V0, -V1, and -V4 were significantly higher in grade 1 and grade 3 endometrioid EC than in normal endometrium ( $p < 0.05$ ). mRNA of HE4-V2 and -V3 were increased in grade 1 compared to normal endometrium ( $p < 0.05$ ). The difference in the expression levels of *HE4*-V2 and -V3 in grade 3 EC compared to normal endometrium were of borderline significance ( $p = 0.057$ ,  $p = 0.067$ , respectively). Likewise, the levels of all variants but *HE4*-V0 ( $p = 0.17$ ) were significantly increased in papillary serous EC compared to normal endometrium ( $p < 0.05$ ). The individual expression levels of each *HE4* variant compared to housekeeping gene measured in each clinical case were shown in the Figure S4 (supplemental data section). As it was observed in all normal tissues, *HE4*-V0 is the most abundant variant among the five counterparts in malignant as well as benign endometrial tissues.

**Figure 5.** Increased *HE4* variant expression levels in EC. mRNA levels of the five *HE4* variants were measured in normal endometrium, grade 1 and 3 endometrioid cancer, and serous cancers by real-time PCR. *HE4* mRNA levels in normal endometrium were set as 1. Statistical significance ( $p < 0.05$ ) was indicated by asterisks.



While *HE4* expression patterns in grade 1 endometrioid, grade 3 endometrioid and papillary serous EC were significantly inter-correlated with each other ( $p < 0.05$ ) (Table S1), they were not statistically correlated with the expression patterns in normal endometrium ( $p = 0.32$ ,  $p = 0.18$ ,  $p = 0.23$

respectively), drawing a clear line that separate the malignant and normal endometrial tissues as two distinct entities.

### 2.5. Increased HE4 Expression Levels Are Negatively Associated with Patient Outcome

The relationship between *HE4* variant expression levels and clinical outcomes were analyzed. Clinical data on survival, BMI and age are retrieved and summarized in supplemental Table S2. No statistically significant correlation was observed between *HE4* variants levels and age or BMI in the normal control group (Table 4). In contrast, a significant inverse correlation was found between *HE4*-V1, -V3, -V4 mRNA levels and EC patient survival time in grade 1 and grade 3 endometrioid cancer sub-cohort (Table 5) ( $p = 0.031$ ,  $p = 0.048$ ,  $p = 0.01$ , respectively). The detailed one-to-one correlation is documented in Figure 6. Interestingly, in the same endometrioid cancer sub-cohort, *HE4*-V0 levels were found to be positively correlated with patients' age ( $p = 0.033$ ), but not with body mass index (BMI) ( $p = 0.79$ ) or survival ( $p = 0.23$ ) (Table 5). Multivariate analysis adjusted by patient age and survival time confirmed that age was independently correlated with *HE4* levels. It is surprising to see a lack of significant correlation between mRNA levels of the five variants of *HE4* and patient survival, BMI or age in the serous cancer group (Table 6). The mean levels of Age, BMI, and number of study subjects are presented in Table 7. The different relationships between *HE4* variant expression levels and clinical parameters in the endometrioid and serous tumors underscore a divergence between the type I and II EC cases, as expected by their differing etiology and molecular make-up.

**Table 4.** Correlation of *HE4* variants expression levels with endometrial cancer (EC) patients' clinical data. The Pearson correlation coefficient analysis is performed between *HE4* variant levels and patients' age, BMI, and survival time. No significant correlation between *HE4* variant levels and clinical data was found in normal endometrium group (Table 4).

|                             | BMI           | Age           |
|-----------------------------|---------------|---------------|
| <b>Log<sub>10</sub>(V0)</b> | <b>-0.172</b> | <b>-0.184</b> |
|                             | $p = 0.524$   | 0.495         |
|                             | $n = 16$      | 16            |
| <b>Log<sub>10</sub>(V1)</b> | <b>0.086</b>  | <b>-0.076</b> |
|                             | $p = 0.752$   | 0.779         |
|                             | $n = 16$      | 16            |
| <b>Log<sub>10</sub>(V2)</b> | <b>0.046</b>  | <b>-0.075</b> |
|                             | $p = 0.865$   | 0.782         |
|                             | $n = 16$      | 16            |
| <b>Log<sub>10</sub>(V3)</b> | <b>-0.087</b> | <b>0.377</b>  |
|                             | $p = 0.749$   | 0.150         |
|                             | $n = 16$      | 16            |
| <b>Log<sub>10</sub>(V4)</b> | <b>0.229</b>  | <b>0.134</b>  |
|                             | $p = 0.394$   | 0.620         |
|                             | $n = 16$      | 16            |

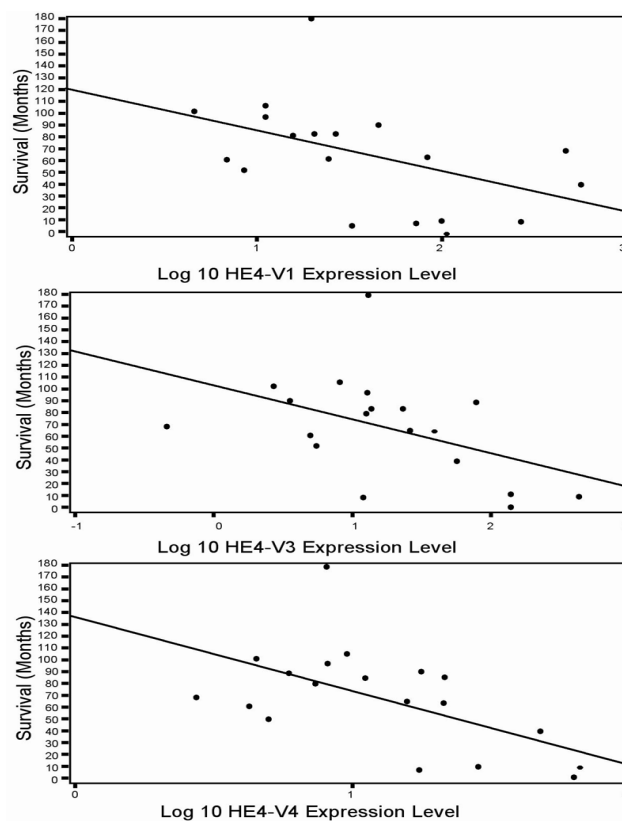
Prob >  $|r|$  under H0: Rho = 0.

**Table 5.** Correlation of *HE4* variants expression levels with endometriod EC patients’ clinical data. A significant inverse correlation between three of the variants (*HE4*-V1, -V3, -V4) and patients’ survival were found in endometrioid EC group. There was also a positive correlation found between variants *HE4*-V0 and patients’ age. No significant correlation was found between *HE4* expression levels and patients’ BMI in this group.

|                             | <b>Survival</b>  | <b>BMI</b>    | <b>Age</b>   |
|-----------------------------|------------------|---------------|--------------|
| <b>Log<sub>10</sub>(V0)</b> | <b>-0.285</b>    | <b>-0.057</b> | <b>0.419</b> |
|                             | <i>p</i> = 0.236 | 0.791         | 0.033        |
|                             | <i>n</i> = 19    | 24            | 26           |
| <b>Log<sub>10</sub>(V1)</b> | <b>-0.497</b>    | <b>-0.114</b> | <b>0.226</b> |
|                             | <i>p</i> = 0.031 | 0.595         | 0.267        |
|                             | <i>n</i> = 19    | 24            | 26           |
| <b>Log<sub>10</sub>(V2)</b> | <b>-0.338</b>    | <b>-0.085</b> | <b>0.179</b> |
|                             | <i>p</i> = 0.157 | 0.692         | 0.380        |
|                             | <i>n</i> = 19    | 24            | 26           |
| <b>Log<sub>10</sub>(V3)</b> | <b>-0.459</b>    | <b>-0.038</b> | <b>0.262</b> |
|                             | <i>p</i> = 0.048 | 0.858         | 0.196        |
|                             | <i>n</i> = 19    | 24            | 26           |
| <b>Log<sub>10</sub>(V4)</b> | <b>-0.572</b>    | <b>-0.056</b> | <b>0.299</b> |
|                             | <i>p</i> = 0.011 | 0.796         | 0.138        |
|                             | <i>n</i> = 19    | 24            | 26           |

Prob > |*r*| under H0:Rho = 0.

**Figure 6.** Correlation between *HE4* variants mRNA levels and EC patient survival time. A significant inverse correlation was found between *HE4*-V1, -V3, -V4 mRNA levels and EC patient survival time in this endometrioid cancer sub-cohort (*p* = 0.031, *p* = 0.048, *p* = 0.01, respectively, *n* = 19).



**Table 6.** Correlation of *HE4* variants expression levels with Serous EC patients' clinical data. In serous cancer group, there was no significant correlation between the variants expression and patients' survival, BMI and Age.

|                             | <b>Survival</b>  | <b>BMI</b>   | <b>Age</b>    |
|-----------------------------|------------------|--------------|---------------|
| <b>Log<sub>10</sub>(V0)</b> | <b>0.088</b>     | <b>0.253</b> | <b>-0.075</b> |
|                             | <i>p</i> = 0.774 | 0.364        | 0.790         |
|                             | <i>n</i> = 13    | 15           | 15            |
| <b>Log<sub>10</sub>(V1)</b> | <b>0.304</b>     | <b>0.145</b> | <b>-0.077</b> |
|                             | <i>p</i> = 0.313 | 0.605        | 0.786         |
|                             | <i>n</i> = 13    | 15           | 15            |
| <b>Log<sub>10</sub>(V2)</b> | <b>0.247</b>     | <b>0.454</b> | <b>-0.254</b> |
|                             | <i>p</i> = 0.416 | 0.089        | 0.361         |
|                             | <i>n</i> = 13    | 15           | 15            |
| <b>Log<sub>10</sub>(V3)</b> | <b>0.32</b>      | <b>0.112</b> | <b>-0.085</b> |
|                             | <i>p</i> = 0.287 | 0.691        | 0.763         |
|                             | <i>n</i> = 13    | 15           | 15            |
| <b>Log<sub>10</sub>(V4)</b> | <b>0.332</b>     | <b>0.482</b> | <b>-0.074</b> |
|                             | <i>p</i> = 0.268 | 0.069        | 0.794         |
|                             | <i>n</i> = 13    | 15           | 15            |

Prob > |*r*| under H<sub>0</sub>:Rho = 0.

**Table 7.** In endometrioid cancer group, the cases numbers (*N*), mean value (mean), the maximum value and minimum value on patients' survival, BMI and age are summarized.

|          | <i>N</i> | <b>Mean</b> | <b>σ</b> | <b>Minimum</b> | <b>Maximum</b> |
|----------|----------|-------------|----------|----------------|----------------|
| Survival | 19       | 68.7        | 43.1     | 1              | 180            |
| BMI      | 24       | 34.4        | 9.4      | 19.6           | 63.5           |
| Age      | 26       | 59.7        | 12.8     | 39             | 91             |

### 3. Discussion

Utilization of alternative splicing or alternative promoters allows a single gene to encode multiple structurally and functionally related proteins. These processes have greatly increased the diversity of proteome, and may constitute an effective and economical mechanism for fine-tuned regulations of cell functions. Indeed, complex posttranscriptional regulatory mechanisms appear to be a common feature of the WAP domain genes. *Elafin* and *SLPI* genes, the two best-studied WAP domain family members, are both able to generate multiple splicing products [38–41]. Eppin, another recently identified WAP protein, is encoded by a single-copy gene as evidenced by the presence of single band on the Southern blotting of human genomic DNA [42]. Computer-aided analysis of *Eppin* sequence predicts the existence of three splice variants, all of which conform to the AG/GT splicing rule [42].

*HE4* gene generates five mRNA variants as a result of alternative splicing and utilization of alternative promoters [8]. Amino acid sequence and domain analysis of the deduced peptides predicts that the *HE4* N-terminal and C-terminal WAP domains are encoded by exon 2 and exon 5, respectively. Based on the arrangement of WAP domains, the five *HE4* transcript variants can be categorized into three groups: Group 1 (*HE4*-V0 and -V1) contains both the N-terminal and C-terminal

WAP domains; Group 2 (*HE4-V4*) contains only the *N*-terminal WAP domain; Group 3 (*HE4-V2* and *-V3*) contains only the *C*-terminal WAP domain. In addition, *HE4-V0*, *-V1*, and *-V4* share the same secretory signal peptide and are therefore likely to be present in the plasma. For the rest of *HE4* variants, *HE4-V2* signal peptide is homologous to that of *HE4-V0* with the exception of a deletion in the last 10 amino acids, suggesting differential regulation of secretion; *HE4-V3* does not contain a consensus signal sequence, raising questions regarding its cellular location and secretory potential [8]. In this study, we provide experimental data showing that these structurally unique *HE4* variants are concurrently expressed in various human tissues. Although all variants were readily detectable by real-time PCR, their expression levels differ dramatically, by as many as 10,000 fold. Taken together, these data suggest that the production of *HE4* variants, either through differential splicing or use of alternative promoters, are subject to tight and distinct regulation. Furthermore, the structural divergence, differential locations, e.g., secreted form versus intracellular form, and markedly different expression patterns suggest that they may be implicated in tissue-specific functions.

Although total *HE4* expressions have been characterized in a variety of normal and cancerous tissues by many laboratories [8,11,12], *HE4* variant-specific expression patterns has not been comprehensively investigated. We showed that the expression patterns of the five *HE4* variants share certain similarity among normal tissues. On the contrary, *HE4* variant expression patterns diverge significantly between malignant and normal endometrial tissues. This observation in conjunction with the fact that these variants are significantly overexpressed in endometrial cancer cells points to a potential role(s) for *HE4* in EC tumorigenesis or progression. In a parallel study, we have tested this hypothesis. We demonstrated that at least the prototype *HE4-V0* exhibits a potent tumor-promoting activity both *in vitro* and *in vivo* [43]. Given the fact that the deduced peptides from these mRNA variants contain a different number of WAP domains, and that the *N*- and *C*-terminal WAP domains share low homology, it is unlikely that these peptides carry out the same functions. Comparative variant-specific functional studies are required to delineate the roles of each variant in EC development.

In addition to its potential pathological role(s), *HE4* levels may also reflect the status of EC progression from a clinical point of view. Yamashita *et al.* performed immunohistochemistry on lung adenocarcinoma tissue samples to investigate the predictive value of *HE4* for patient outcome and found that the five-year disease-free survival in the *HE4*-positive group (44.6%) was significantly lower than that in the *HE4*-negative group (82.3%,  $p = 0.001$ ). Moreover, *HE4* expression and the nodal status were noted to be independent prognostic factors for disease-free and overall survival [44]. Kamei *et al.* examined the *HE4* expression in breast cancer tissues and confirmed that *HE4* protein level was associated with lymph node invasion. Furthermore, the five-year survival in the *HE4*-positive group (58.6%) was significantly worse than that in the *HE4*-negative group (85.6%,  $p = 0.04$ ) [45]. Bignotti, *et al.* found that high serum *HE4* levels correlated with reduced survival in the poorly differentiated EC cohort [30]. A study from our group recently indicated a correlation between the high serum *HE4* levels and high surgical stage, increased myometrial invasion, and large primary tumor diameters in EC patients [46]. In the current study although all five *HE4* mRNA variants were significantly increased in malignant compared to normal endometrium, only the levels of three variants (*HE4-V1*, *-V3*, and *-V4*) were found to be correlated with the patient survival data. This finding suggests that not all *HE4* variants are predictive of patient outcomes, which underscores the importance of a variant-specific approach in the measurement of *HE4* levels. It is noteworthy that the

variants associated with survival (*HE4-V1*, *-V3*, and *-V4*) are those expressed at the highest level in the trachea but not the ones (*HE4-V0* and *-V2*) expressed at the highest levels in the epididymis. Neither the interpretation nor the significance of this observation is clear at this time.

Despite the ambiguity regarding each variant's distinct function, correlation between the expression of certain *HE4* variants and patient survival may prove important from a clinical perspective. Numerous studies have shown that mRNA detection in serum as well as in other body fluids may provide a promising approach to cancer assessment and management. Tumor cell mRNA in plasma may originate from necrosis or apoptosis of circulating cancer cells, or as the result of an active release mechanism in which mRNA transportation occurs through vesicle-like structures [47,48]. Based on the high expression levels of the *HE4* mRNA variants in EC tissue samples, the same mRNA variants may be detectable in patients' serum as well. Thus, the variant-specific approach may offer insights into EC detection, prognostication, and disease monitoring. The optimal technical parameters we outline here could serve as a foundation for future studies along this direction.

Because both ovary and endometrium develop from the Müllerian system [49–51], carcinomas arising from these sites tend to have similar etiological factors, gene expression profiles, and tumorigenic mechanisms. Consequently, defining the expression patterns of *HE4* variants in normal and malignant ovarian tissues may demonstrate similar findings. However, early detection of *HE4* variants is more convenient in EC compared to ovarian or other types of cancer given that tumor cells can be collected non-invasively in menses or vaginal secretions by the use of pap smear-like devices or tampons [52]. RNA samples extracted from pap smear specimens have been successfully utilized for transcriptome profiling in the past [53]. The currently available *HE4* ELISA assay measures the total *HE4* levels and does not distinguish the *HE4* protein species produced by different *HE4* mRNA variants. One could envision using either antibody-based ELISA or mRNA-based PCR assays to detect *HE4* mRNA variants in these body fluids for the purpose of early EC detection.

Anastasi *et al.* observed that serum *HE4* levels vary substantially among the follicular, ovulatory and luteal phases in healthy young women, suggesting hormonal regulation of *HE4* production in the ovary and other somatic tissues [28]; recently, Levanon *et al.* have also shown that normal fallopian tube secretes *HE4*, these data indicate the characteristic of the *HE4* origin from the Mullerian epithelium [54]. Malignant endometrial cells express varied levels of steroid hormone receptors due to differential epigenetic modification. This may have a significant effect on *HE4* variant expression patterns and at the same time, cancers' response to hormonal therapy. Since most of the patients included in our study were postmenopausal women, we could not assess the influence of menstrual cycle on *HE4* expression. Interestingly, *HE4-V0*, the most abundant *HE4* variant, was found to be correlated with age in the endometrioid cancer cohort. This result is consistent with another study showing that serum *HE4* levels are closely correlated with age in EC patients [46]. Given the fact that the relationship between *HE4* mRNA levels and age is not evident in the normal control group, this age-related change of *HE4* levels may reflect tumor characteristics rather than varying hormonal status. This hypothesis is further supported by the data from our recent study in which *HE4* serum levels positively correlated with high-risk tumor characteristics in EC [46]. The association between *HE4* /*HE4* levels and age also suggest that the age-adjusted *HE4* levels may be more accurate criteria for cancer diagnosis and prognostication. We should point out that this observation is in agreement with the data recently published by Bolstad *et al.* [55]. In that study, it was found that higher *HE4*

levels are associated with age and smoking status, as well as creatinine levels and BMI in healthy subjects, while the two later factors are also associated with age. These new findings are very important in the way that when *HE4* variants are applied as a biomarker for EC diagnosis, these factors need to be taken into account to rule out the possibility that the elevated *HE4* levels may be caused by the patients' age, smoking status and BMI in addition to the tumor itself. As the higher creatinine levels and smoking status are associated with age, the relationship between *HE4* levels and creatinine levels would be an interesting issue for future studies.

This study is to our knowledge the first one to demonstrate a tissue-specific distribution of *HE4* variants and establish an *HE4* variant-specific correlation with survival in EC. Furthermore, we suggest that a variant-specific approach may prove useful for improving the current use of *HE4* assays in the diagnosis, triage and postsurgical surveillance for the patient with EC.

## 4. Experimental Section

### 4.1. Collection of Tissue Samples

Both normal and malignant endometrial tissue samples were collected randomly from patients treated at Mayo Clinic for a variety of gynecologic conditions. Central pathology review was conducted to verify institutional diagnoses and eliminate variability among experts. Samples obtained from patients undergoing hysterectomy for benign indications were used as normal endometrium controls. Endometrial tissue specimens were obtained from 14 patients with grade 1 EC, 14 patients with grade 3 endometrioid EC, and 15 patients with papillary serous EC, and 16 patients with normal endometrium. For the six pairs of tissues analyzed by mass spectrometry study, EC and normal samples were selected based on match criteria of age ( $\pm 5$  years), body weight ( $\pm 10$  kg) and menopausal status, which is a separate pool of the above described cohort. Clinical data on survival, BMI and age were retrieved and confirmed for most of the patients. All samples were snap frozen, stored at  $-80$  °C. For total RNA extraction, tissues were cut into 10  $\mu\text{m}$ -thick series sections; one section was used for H&E staining to define the tumor tissues (circled with marker) vs. myometrium. 10 series sections were matched with H&E stained one, only tumor tissues were macro-dissected to ensure that at least 50% of the samples were tumor tissue. These studies were approved by the Institutional Review Board of Mayo Foundation (IRB# 07-004290). In accordance with the Helsinki Declaration and Minnesota Statute for Use of Medical Information in Research, only patients who consented to the use of their medical records were included in the study.

### 4.2. Tissue Preparation and ICAT Labeling Procedure

Specimens were mixed with 0.5 ml of lysis buffer (0.5 mM pH 8.3 Tris plus 0.1% SDS), and a glass grinder was used to homogenize the tissues. After separating the supernatant from the tissue debris by centrifugation at 13523 RCF (or 12,000 rpm), the supernatant was used for Isotope-Coded Affinity Tag (ICAT) labeling for a high throughput comparison of normal and malignant endometrial tissues. Protein concentrations were measured using the Bio-Rad Protein Assay Kit (Bio-Rad, Hercules, CA, USA). 100  $\mu\text{g}$  of protein was labeled with Cleavable ICAT (cICAT) purchased from Applied Biosystems (Applied Biosystems, Foster City, CA, USA). We labeled the control samples with light



isotopes (containing 8 hydrogens, D<sub>0</sub>) and the EC samples with heavy isotopes (containing 8 deuteriums, D<sub>8</sub>). Briefly, protein extracts were mixed with 2 µL of reducing agent. After boiling for 10 min, the labeling reaction was carried out overnight at 37 °C. Each labeled EC sample was combined with its labeled control sample and the mixture was incubated overnight at 37 °C in digestion buffer solution containing 2 µL of CaCl<sub>2</sub> and 1 vial of trypsin as provided in the kit. Following digestion, the peptide mixture was subject to cation ion exchange purification and affinity purification using an Avidin column to collect peptides containing ICAT-labeled cysteine. 100 µL of cleavage reagent was added and the mixture was kept at 37 °C for 2 h to cleave the biotin portion. Cleavage mixture was dried under vacuum and stored at -80 °C until mass spectrometric analysis.

#### 4.3. Multidimensional Liquid Chromatography Coupled to Tandem Mass Spectrometry (LC/MS/MS)

The cleaved samples were separated into 8 fractions on a cation exchange column (BioX SCX 300 mm × 5 cm, Dionex, Sunnyvale, CA, USA) using an off-line Agilent 1100 series capillary liquid chromatography system (Wilmington, DE, USA). The API QSTAR Pulsar I, a hybrid quadrupole time-of-flight mass spectrometer (Applied Biosystems, Foster City, CA, USA) was used to perform the LC/MS/MS analysis of the peptides from each fraction with the help of the Analyst QS 1.0 software (Applied Biosystems, Foster City, CA, USA). The QSTAR was configured with the Protana Nanospray Source (Proxeon, Denmark), and coupled to an Ultimate Nano liquid chromatography system (Dionex, Sunnyvale, CA, USA) equipped with a Zorbax C18 100 mm × 150 mm microbore column (Agilent, Wilmington, DE, USA). Mass spectrometry analysis consists of a one-second survey scan from 400 to 1600 mass-to-charge ratio ( $m/z$ ) followed by two-second MS/MS fragmentation ion scan with a threshold of 10 counts per second. Once the three most intense ions were fragmented in each survey scan, they were excluded from repeated fragmentation for 60 s. The collision energy applied varied automatically depending on the precursor  $m/z$  and charge state. The peptide mass tolerance was set at 0.2 Da.

#### 4.4. Data Analysis of LC/ICAT

Acquired data were searched against the CDS FASTA database and quantified using ProICAT 1.0 SP3 software (Applied Biosystems, Foster City, CA, USA). Relative quantification of proteins was performed on the TOF (time of flight) MS scan by calculating the ratio between the areas representing the light- and heavy-labeled peptide peaks. Two different search engines were used for protein identification, the ProICAT and the Mascot (Matrix Science, Boston, MA, USA) from the human non-redundant database. Only the proteins with a confidence level of 90% were used for further identification and quantification. The protein function annotation was supplemented using an online bioinformatics program, the Bioinformatics Harvester [56]. All proteins identified (Table 1) were individually verified by manually inspecting its precursor on chromatogram and mass spectrum. The protein accession numbers (GI number) are from NCBI protein database.

#### 4.5. RNA Isolation, cDNA Synthesis, and Quantitative PCR

Total RNA of endometrial tissues was isolated from 20 µm-thick sections of frozen tissue using TRIzol<sup>®</sup> reagent (Invitrogen, Carlsbad, CA, USA) followed the protocol described previously [57]. All

the RNA samples were treated with DNase (RNase-Free DNase Set, cat No. 79254, Qiagen, Valencia, CA, USA) before reverse transcription. The purity of RNA has been determined by OD260/280 ratio. The quality of RNA and cDNA was ensured by the results of pilot PCR in which clear specific DNA band of the target gene was observed in agarose gel electrophoresis. Poly-A RNA from human epididymis, skeletal muscle (s. muscle), testis, prostate, lung, placenta, trachea, kidney, peripheral blood lymphocyte (PBL), colon, liver, breast, and ovary were purchased from Clontech Laboratories (Clontech Laboratories, Inc., Mountain View, CA, USA). cDNA was synthesized from 1 µg of total RNA using the SuperScript™ kit (Invitrogen, Carlsbad, CA, USA). 20 µL of reverse transcription product was diluted to 100 µL and 2 µL used for each real-time PCR reaction. The cDNA sequences for human *HE4* transcript variants (*HE4-V0* to *-V4*) and human *36B4* (internal control) and *GAPDH* (internal control) were obtained from the PubMed gene bank and primers were designed by using primer 3 software [58]. All the primers were synthesized from IDT (Integrated DNA Technology). *36B4* is a ribosomal RNA and ubiquitously expressed in all the cells. Owing to its housekeeping function, and the fact that CT value of *36B4* is lower than the CT value of *GAPDH*, and *36B4* has lower variance, *36B4* was used as a control experiment for real-time PCR. We have also used *36B4* as housekeeping gene in previously works [59–61]. Negative control has been performed on the same 96-well PCR plate to test the PCR specificity. Real-time PCR was performed following previously described protocols [57,61] in a 25 µL reaction containing 2 µL of cDNA, 12.5 µL of SYBR® Green PCR Master Mix (Stratagene, Cedar Creek, TX, USA) and 50 nM of forward and reverse primers, respectively. The ABI 7900HT Fast Real-time PCR® System (Applied Biosystems, Foster City, CA, USA) was used with the following regimen of thermal cycling: Stage 1: 1 cycle, 2 min at 50 °C; Stage 2: 1 cycle, 5 min at 95 °C; Stage 3: 40 cycles, 15 s at 95 °C, 1 min at 60 °C.

Relative mRNA concentrations were calculated as the following: The threshold cycle number (CT) at which PCR products reached a preset threshold value, which was defined as the value where all products were undergoing exponential amplification, was determined by the fluorescence detector. Real-time PCR for the housekeeping gene *36B4* was performed in the same plates as for *HE4* experiments. The *HE4* real-time PCR results were standardized by those of *36B4* using the formula  $\Delta CT = CT_{HE4} - CT_{36B4}$  ( $CT_{HE4}$  and  $CT_{36B4}$ , threshold values for *HE4* and *36B4*, respectively). Relative *HE4* mRNA levels were expressed as fold over *36B4* mRNA levels ( $F = 2^{\Delta CT}$ ). The experiments were repeated at least twice and the CT for each sample was determined in duplicate or triplicate. For practical reasons, relatively *HE4* mRNA levels were arbitrarily amplified by a factor of 100 or 10,000 as indicated in figure legends. All the real-time PCR data were normalized with housekeeping gene.

#### 4.6. Statistical Analysis

Mean values among different groups were compared (grade 1 or grade 3 endometrioid EC or papillary serous EC *versus* normal endometrium) using a *post-hoc* Bonferroni *t* test with the assumption that *HE4* mRNA levels exhibited bimodal distribution and that the variances of the two samples were equal. Pearson's correlation coefficient test was used to examine the relationship between the expression patterns of the five *HE4* variants among normal tissues as well as normal and malignant endometrium. Pearson's correlation coefficient test was also applied to study the relationship between *HE4* mRNA levels and patient characteristics including age, surgical stage, and

disease free survival. To avoid overly influential observations, our data were log-transformed before statistical analysis. Correlations with a  $p$  value smaller than 0.05 and a coefficient close to 1.0 were considered statistically significant.

## 5. Conclusions

This study is to our knowledge the first one to demonstrate a tissue-specific distribution of *HE4* variants and establish an *HE4* variant-specific correlation with survival in EC. Furthermore, we suggest that a variant-specific approach may prove useful for improving the current use of HE4 assays in the diagnosis, triage and postsurgical surveillance for the patient with EC.

## Acknowledgments

We thank Ying Zhao and Run Zhang for their technical support, and Clare Zhou for her efforts in retrieving part of the clinical data.

This study is supported by the Career Development Program of the NIH/NCI MD Anderson Uterine Cancer SPORE (J. Li, S-W. Jiang, K. Podratz); Mercer University Seed Grant (J. Li, S-W. Jiang); Distinguished Cancer Scholar Program of Georgia Cancer Coalition (S-W. Jiang); Research Supplement from Mercer University School of Medicine (S-W. Jiang); Fraternal Order Of Eagles Cancer Research Fellow, Mayo Cancer Center (S-W. Jiang.); Georgia Research Alliance Cancer Research Reward (S-W. Jiang); Research Supplement from the Mercer University School of Medicine (J. Li); Robert and Debra First Fund, Mary Kay Foundation, and Sandy Rollman Ovarian Cancer Foundation (R. Drapkin) and NIH 2P50 CA098258-06 SPORE in Uterine Cancer (R. Broaddus).

## Conflicts of Interest

The authors declare no conflict of interest.

## References

1. SEER Stat Fact Sheets: Endometrial Cancer. Available online: <http://seer.cancer.gov/statfacts/html/corp.html> (accessed on 14 June 2013).
2. Takeshima, N.; Shimizu, Y.; Umezawa, S.; Hirai, Y.; Chen, J.T.; Fujimoto, I.; Yamauchi, K.; Hasumi, K. Combined assay of serum levels of CA125 and CA19-9 in endometrial carcinoma. *Gynecol. Oncol.* **1994**, *54*, 321–326.
3. Yildiz, A.; Yetimalar, H.; Kasap, B.; Aydin, C.; Tatar, S.; Soyulu, F.; Yildiz, F.S. Preoperative serum CA 125 level in the prediction of the stage of disease in endometrial carcinoma. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **2012**, *164*, 191–195.
4. Sebastianelli, A.; Renaud, M.C.; Gregoire, J.; Roy, M.; Plante, M. Preoperative CA 125 tumour marker in endometrial cancer: Correlation with advanced stage disease. *J. Obstet. Gynaecol. Can.* **2010**, *32*, 856–860.
5. Kirchhoff, C.; Osterhoff, C.; Habben, I.; Ivell, R. Cloning and analysis of mRNAs expressed specifically in the human epididymis. *Int. J. Androl.* **1990**, *13*, 155–167.

6. Kirchoff, C.; Habben, I.; Ivell, R.; Krull, N. A major human epididymis-specific cDNA encodes a protein with sequence homology to extracellular proteinase inhibitors. *Biol. Reprod.* **1991**, *45*, 350–357.
7. Kirchoff, C. Molecular characterization of epididymal proteins. *Rev. Reprod.* **1998**, *3*, 86–95.
8. Bingle, L.; Singleton, V.; Bingle, C.D. The putative ovarian tumour marker gene HE4 (WFDC2), is expressed in normal tissues and undergoes complex alternative splicing to yield multiple protein isoforms. *Oncogene* **2002**, *21*, 2768–2773.
9. Schummer, M.; Ng, W.V.; Bumgarner, R.E.; Nelson, P.S.; Schummer, B.; Bednarski, D.W.; Hassell, L.; Baldwin, R.L.; Karlan, B.Y.; Hood, L. Comparative hybridization of an array of 21,500 ovarian cDNAs for the discovery of genes overexpressed in ovarian carcinomas. *Gene* **1999**, *238*, 375–385.
10. Anderson, G.L.; McIntosh, M.; Wu, L.; Barnett, M.; Goodman, G.; Thorpe, J.D.; Bergan, L.; Thornquist, M.D.; Scholler, N.; Kim, N.; *et al.* Assessing lead time of selected ovarian cancer biomarkers: A nested case-control study. *J. Natl. Cancer Inst.* **2010**, *102*, 26–38.
11. Drapkin, R.; von Horsten, H.H.; Lin, Y.; Mok, S.C.; Crum, C.P.; Welch, W.R.; Hecht, J.L. Human epididymis protein 4 (HE4) is a secreted glycoprotein that is overexpressed by serous and endometrioid ovarian carcinomas. *Cancer Res.* **2005**, *65*, 2162–2169.
12. Galgano, M.T.; Hampton, G.M.; Frierson, H.F., Jr. Comprehensive analysis of HE4 expression in normal and malignant human tissues. *Mod. Pathol.* **2006**, *19*, 847–853.
13. Hellstrom, I.; Raycraft, J.; Hayden-Ledbetter, M.; Ledbetter, J.A.; Schummer, M.; McIntosh, M.; Drescher, C.; Urban, N.; Hellstrom, K.E. The HE4 (WFDC2) protein is a biomarker for ovarian carcinoma. *Cancer Res.* **2003**, *63*, 3695–3700.
14. Bast, R.C., Jr.; Brewer, M.; Zou, C.; Hernandez, M.A.; Daley, M.; Ozols, R.; Lu, K.; Lu, Z.; Badgwell, D.; Mills, G.B.; *et al.* Prevention and early detection of ovarian cancer: Mission impossible? *Recent Results Cancer Res.* **2007**, *174*, 91–100.
15. Hellstrom, I.; Hellstrom, K.E. SMRP and HE4 as biomarkers for ovarian carcinoma when used alone and in combination with CA125 and/or each other. *Adv. Exp. Med. Biol.* **2008**, *622*, 15–21.
16. Li, J.; Dowdy, S.; Tipton, T.; Podratz, K.; Lu, W.G.; Xie, X.; Jiang, S.W. HE4 as a biomarker for ovarian and endometrial cancer management. *Expert Rev. Mol. Diagn.* **2009**, *9*, 555–566.
17. Scholler, N.; Garvik, B.; Quarles, T.; Jiang, S.; Urban, N. Method for generation of *in vivo* biotinylated recombinant antibodies by yeast mating. *J. Immunol. Methods* **2006**, *317*, 132–143.
18. Scholler, N.; Crawford, M.; Sato, A.; Drescher, C.W.; O'Briant, K.C.; Kiviat, N.; Anderson, G.L.; Urban, N. Bead-based ELISA for validation of ovarian cancer early detection markers. *Clin. Cancer Res.* **2006**, *12*, 2117–2124.
19. Scholler, N.; Lowe, K.A.; Bergan, L.A.; Kampani, A.V.; Ng, V.; Forrest, R.M.; Thorpe, J.D.; Gross, J.A.; Garvik, B.M.; Drapkin, R.; *et al.* Use of yeast-secreted *in vivo* biotinylated recombinant antibodies (Biobodies) in bead-based ELISA. *Clin. Cancer Res.* **2008**, *14*, 2647–2655.
20. Havrilesky, L.J.; Whitehead, C.M.; Rubatt, J.M.; Cheek, R.L.; Groelke, J.; He, Q.; Malinowski, D.P.; Fischer, T.J.; Berchuck, A. Evaluation of biomarker panels for early stage ovarian cancer detection and monitoring for disease recurrence. *Gynecol. Oncol.* **2008**, *110*, 374–382.

21. Hellstrom, I.; Heagerty, P.J.; Swisher, E.M.; Liu, P.; Jaffar, J.; Agnew, K.; Hellstrom, K.E. Detection of the HE4 protein in urine as a biomarker for ovarian neoplasms. *Cancer Lett.* **2010**, *296*, 43–48.
22. Yurkovetsky, Z.; Skates, S.; Lomakin, A.; Nolen, B.; Pulsipher, T.; Modugno, F.; Marks, J.; Godwin, A.; Gorelik, E.; Jacobs, I.; *et al.* Development of a multimarker assay for early detection of ovarian cancer. *J. Clin. Oncol.* **2010**, *28*, 2159–2166.
23. Huhtinen, K.; Suvitie, P.; Hiissa, J.; Junnila, J.; Huvila, J.; Kujari, H.; Setala, M.; Harkki, P.; Jalkanen, J.; Fraser, J.; *et al.* Serum HE4 concentration differentiates malignant ovarian tumours from ovarian endometriotic cysts. *Br. J. Cancer* **2009**, *100*, 1315–1319.
24. Moore, R.G.; Brown, A.K.; Miller, M.C.; Badgwell, D.; Lu, Z.; Allard, W.J.; Granai, C.O.; Bast, R.C., Jr.; Lu, K. Utility of a novel serum tumor biomarker HE4 in patients with endometrioid adenocarcinoma of the uterus. *Gynecol. Oncol.* **2008**, *110*, 196–201.
25. Park, Y.; Kim, Y.; Lee, E.Y.; Lee, J.H.; Kim, H.S. Reference ranges for HE4 and CA125 in a large Asian population by automated assays and diagnostic performances for ovarian cancer. *Int. J. Cancer* **2012**, *130*, 1136–1144.
26. Ruggeri, G.; Bandiera, E.; Zanotti, L.; Belloli, S.; Ravaggi, A.; Romani, C.; Bignotti, E.; Tassi, R.A.; Tognon, G.; Galli, C.; *et al.* HE4 and epithelial ovarian cancer: Comparison and clinical evaluation of two immunoassays and a combination algorithm. *Clin. Chim. Acta* **2011**, *412*, 1447–1453.
27. Shah, C.A.; Lowe, K.A.; Paley, P.; Wallace, E.; Anderson, G.L.; McIntosh, M.W.; Andersen, M.R.; Scholler, N.; Bergan, L.A.; Thorpe, J.D.; *et al.* Influence of ovarian cancer risk status on the diagnostic performance of the serum biomarkers mesothelin, HE4, and CA125. *Cancer Epidemiol. Biomark. Prev.* **2009**, *18*, 1365–1372.
28. Anastasi, E.; Granato, T.; Marchei, G.G.; Viggiani, V.; Colaprisca, B.; Comploj, S.; Reale, M.G.; Frati, L.; Midulla, C. Ovarian tumor marker HE4 is differently expressed during the phases of the menstrual cycle in healthy young women. *Tumour. Biol.* **2010**, *31*, 411–415.
29. Moore, R.G.; Brown, A.K.; Miller, M.C.; Skates, S.; Allard, W.J.; Verch, T.; Steinhoff, M.; Messerlian, G.; DiSilvestro, P.; Granai, C.O.; *et al.* The use of multiple novel tumor biomarkers for the detection of ovarian carcinoma in patients with a pelvic mass. *Gynecol. Oncol.* **2008**, *108*, 402–408.
30. Bignotti, E.; Ragnoli, M.; Zanotti, L.; Calza, S.; Falchetti, M.; Lonardi, S.; Bergamelli, S.; Bandiera, E.; Tassi, R.A.; Romani, C.; *et al.* Diagnostic and prognostic impact of serum HE4 detection in endometrial carcinoma patients. *Br. J. Cancer* **2011**, *104*, 1418–1425.
31. Lowe, K.A.; Shah, C.; Wallace, E.; Anderson, G.; Paley, P.; McIntosh, M.; Andersen, M.R.; Scholler, N.; Bergan, L.; Thorpe, J.; *et al.* Effects of personal characteristics on serum CA125, mesothelin, and HE4 levels in healthy postmenopausal women at high-risk for ovarian cancer. *Cancer Epidemiol. Biomark. Prev.* **2008**, *17*, 2480–2487.
32. Montagnana, M.; Lippi, G.; Ruzzenente, O.; Bresciani, V.; Danese, E.; Scevarolli, S.; Salvagno, G.L.; Giudici, S.; Franchi, M.; Guidi, G.C. The utility of serum human epididymis protein 4 (HE4) in patients with a pelvic mass. *J. Clin. Lab Anal.* **2009**, *23*, 331–335.
33. Moore, R.G.; McMeekin, D.S.; Brown, A.K.; DiSilvestro, P.; Miller, M.C.; Allard, W.J.; Gajewski, W.; Kurman, R.; Bast, R.C., Jr.; Skates, S.J. A novel multiple marker bioassay utilizing HE4 and CA125 for the prediction of ovarian cancer in patients with a pelvic mass. *Gynecol. Oncol.* **2009**, *112*, 40–46.

34. Jacob, F.; Meier, M.; Caduff, R.; Goldstein, D.; Pochechueva, T.; Hacker, N.; Fink, D.; Heinzelmann-Schwarz, V. No benefit from combining HE4 and CA125 as ovarian tumor markers in a clinical setting. *Gynecol. Oncol.* **2011**, *121*, 487–491.
35. Montagnana, M.; Danese, E.; Ruzzenente, O.; Bresciani, V.; Nuzzo, T.; Gelati, M.; Salvagno, G.L.; Franchi, M.; Lippi, G.; Guidi, G.C. The ROMA (Risk of Ovarian Malignancy Algorithm) for estimating the risk of epithelial ovarian cancer in women presenting with pelvic mass: Is it really useful? *Clin. Chem. Lab. Med.* **2011**, *49*, 521–525.
36. Li, J.; Fu, A.; Huang, T.H.; Dowdy, S.C.; Podratz, K.C.; Jiang, S.W. Proteomic Identification of HE4 as a potential molecular marker for endometrioid cancer of uterus. *J. Obstet Gynaecol Can.* **2013**, in preparation.
37. Genomic Structure of *HE4*. Available online: [http://www.ncbi.nlm.nih.gov/sutils/evv.cgi?taxid=9606&contig=NT\\_011362.10&gene=WFDC2&lid=10406](http://www.ncbi.nlm.nih.gov/sutils/evv.cgi?taxid=9606&contig=NT_011362.10&gene=WFDC2&lid=10406) (accessed on 19 August 2004).
38. Kluger, H.M.; Chelouche Lev, D.; Kluger, Y.; McCarthy, M.M.; Kiriakova, G.; Camp, R.L.; Rimm, D.L.; Price, J.E. Using a xenograft model of human breast cancer metastasis to find genes associated with clinically aggressive disease. *Cancer Res.* **2005**, *65*, 5578–5587.
39. Westin, U.; Nystrom, M.; Ljungcrantz, I.; Eriksson, B.; Ohlsson, K. The presence of elafin, SLPI, IL1-RA and STNFalpha RI in head and neck squamous cell carcinomas and their relation to the degree of tumour differentiation. *Mediat. Inflamm.* **2002**, *11*, 7–12.
40. Yoshida, N.; Egami, H.; Yamashita, J.; Takai, E.; Tamori, Y.; Fujino, N.; Kitaoka, M.; Schalkwijk, J.; Ogawa, M. Immunohistochemical expression of SKALP/elafin in squamous cell carcinoma of human lung. *Oncol. Rep.* **2002**, *9*, 495–501.
41. Devoogdt, N.; Hassanzadeh Ghassabeh, G.; Zhang, J.; Brys, L.; de Baetselier, P.; Revets, H. Secretory leukocyte protease inhibitor promotes the tumorigenic and metastatic potential of cancer cells. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 5778–5782.
42. Richardson, R.T.; Sivashanmugam, P.; Hall, S.H.; Hamil, K.G.; Moore, P.A.; Ruben, S.M.; French, F.S.; O’Rand, M. Cloning and sequencing of human Eppin: A novel family of protease inhibitors expressed in the epididymis and testis. *Gene* **2001**, *270*, 93–102.
43. Li, J.; Chen, H.; Mariani, A.; Chen, D.; Klatt, E.; Podratz, K.; Drapkin, R.; Broaddus, R.; Dowdy, S.; Jiang, S.W. HE4 (WFDC2) Promotes Tumor Growth in Endometrial Cancer Cell Lines. *Int. J. Mol. Sci.* **2013**, *14*, 6026–6043.
44. Idoji, Y.; Watanabe, Y.; Yamashita, A.; Yamanishi, K.; Nishiguchi, S.; Shimada, K.; Yasunaga, T.; Yamanishi, H. *In silico* study of whey-acidic-protein domain containing oral protease inhibitors. *Int. J. Mol. Med.* **2008**, *21*, 461–468.
45. Kamei, M.; Yamashita, S.; Tokuishi, K.; Hashimoto, T.; Moroga, T.; Suehiro, S.; Ono, K.; Miyawaki, M.; Takeno, S.; Yamamoto, S.; *et al.* HE4 expression can be associated with lymph node metastases and disease-free survival in breast cancer. *Anticancer Res.* **2010**, *30*, 4779–4783.
46. Kalogera, E.; Scholler, N.; Powless, C.; Weaver, A.; Drapkin, R.; Li, J.; Jiang, S.W.; Podratz, K.; Urban, N.; Dowdy, S.C. Correlation of serum HE4 with tumor size and myometrial invasion in endometrial cancer. *Gynecol. Oncol.* **2012**, *124*, 270–275.
47. Garcia, J.M.; Garcia, V.; Pena, C.; Dominguez, G.; Silva, J.; Diaz, R.; Espinosa, P.; Citores, M.J.; Collado, M.; Bonilla, F. Extracellular plasma RNA from colon cancer patients is confined in a vesicle-like structure and is mRNA-enriched. *RNA* **2008**, *14*, 1424–1432.

48. Skog, J.; Wurdinger, T.; van Rijn, S.; Meijer, D.H.; Gainche, L.; Sena-Esteves, M.; Curry, W.T., Jr.; Carter, B.S.; Krichevsky, A.M.; Breakefield, X.O. Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. *Nat. Cell Biol.* **2008**, *10*, 1470–1476.
49. Leitzmann, M.F.; Koebnick, C.; Danforth, K.N.; Brinton, L.A.; Moore, S.C.; Hollenbeck, A.R.; Schatzkin, A.; Lacey, J.V., Jr. Body mass index and risk of ovarian cancer. *Cancer* **2009**, *115*, 812–822.
50. Parkin, D.M. Cancers of the breast, endometrium and ovary: Geographic correlations. *Eur. J. Cancer Clin. Oncol.* **1989**, *25*, 1917–1925.
51. Ben-Arie, A.; Huszar, M.; Ben-Zvi, N.; Smirnov, A.; Altevogt, P.; Fogel, M. The role of L1-CAM immunohistochemical staining in the diagnosis of abdominal-pelvic cancer of uncertain primary site in women. *Eur. J. Surg. Oncol.* **2008**, *34*, 795–799.
52. Jiang, S.W.; Li, J.; Podratz, K.; Dowdy, S. Application of DNA methylation biomarkers for endometrial cancer management. *Expert Rev. Mol. Diagn.* **2008**, *8*, 607–616.
53. Cricca, M.; Venturoli, S.; Leo, E.; Costa, S.; Musiani, M.; Zerbini, M. Molecular analysis of HPV 16 E6I/E6II spliced mRNAs and correlation with the viral physical state and the grade of the cervical lesion. *J. Med. Virol.* **2009**, *81*, 1276–1282.
54. Levanon, K.; Ng, V.; Piao, H.Y.; Zhang, Y.; Chang, M.C.; Roh, M.H.; Kindelberger, D.W.; Hirsch, M.S.; Crum, C.P.; Marto, J.A.; *et al.* Primary *ex vivo* cultures of human fallopian tube epithelium as a model for serous ovarian carcinogenesis. *Oncogene* **2010**, *29*, 1103–1113.
55. Bolstad, N.; Oijordsbakken, M.; Nustad, K.; Bjerner, J. Human epididymis protein 4 reference limits and natural variation in a Nordic reference population. *Tumour. Biol.* **2012**, *33*, 141–148.
56. Bioinformatics Harvester Program. Available online: <http://www.embl.de/services/bioinformatics/> (accessed on 26 October 2011).
57. Jin, F.; Dowdy, S.C.; Xiong, Y.; Eberhardt, N.L.; Podratz, K.C.; Jiang, S.W. Up-regulation of DNA methyltransferase 3B expression in endometrial cancers. *Gynecol. Oncol.* **2005**, *96*, 531–538.
58. Primer3 tool. Available online: [http://biotools.umassmed.edu/bioapps/primer3\\_www.cgi](http://biotools.umassmed.edu/bioapps/primer3_www.cgi) (accessed on 14 January 2006).
59. Li, J.; Wang, E.; Rinaldo, F.; Datta, K. Upregulation of VEGF-C by androgen depletion: The involvement of IGF-IR-FOXO pathway. *Oncogene* **2005**, *24*, 5510–5520.
60. Rinaldo, F.; Li, J.; Wang, E.; Muders, M.; Datta, K. RalA regulates vascular endothelial growth factor-C (VEGF-C) synthesis in prostate cancer cells during androgen ablation. *Oncogene* **2007**, *26*, 1731–1738.
61. Zhang, H.; Muders, M.H.; Li, J.; Rinaldo, F.; Tindall, D.J.; Datta, K. Loss of NKX3.1 favors vascular endothelial growth factor-C expression in prostate cancer. *Cancer Res.* **2008**, *68*, 8770–8778.