



Pilot Study to Establish a Novel Five-Gene Biomarker Panel for Predicting Lymph Node Metastasis in Patients With Early Stage Endometrial Cancer

Chia-Yen Huang^{1,2,3}, Kuang-Wen Liao^{1,4}, Chih-Hung Chou^{1,5,6}, Sirjana Shrestha¹, Chi-Dung Yang^{1,7,8}, Men-Yee Chiew¹, Hsin-Tzu Huang¹, Hsiao-Chin Hong^{1,6,7,8}, Shih-Hung Huang⁹, Tzu-Hao Chang^{10,11*} and Hsien-Da Huang^{7,8*}

¹ Department of Biological Science and Technology, National Chiao Tung University, Hsinchu, Taiwan, ² Department of Obstetrics and Gynecology, Gynecologic Cancer Center, Cathay General Hospital, Taipei, Taiwan, ³ School of Medicine, Fu Jen Catholic University, New Taipei City, Taiwan, ⁴ Institute of Molecular Medicine and Bioengineering, National Chiao Tung University, Hsinchu, Taiwan, ⁵ Center for Intelligent Drug Systems and Smart Bio-devices (IDS2B), National Chiao Tung University, Hsinchu, Taiwan, ⁶ Institute of Bioinformatics and Systems Biology, National Chiao Tung University, Hsinchu, Taiwan, ⁷ School of Life and Health Sciences, Chinese University of Hong Kong, Shenzhen, China, ⁸ Warshel Institute for Computational Biology, Chinese University of Hong Kong, Shenzhen, China, ⁹ Department of Pathology, Cathay General Hospital, Taipei, Taiwan, ¹⁰ International Center for Health Information Technology, Taipei Medical University, Taipei, Taiwan, ¹¹ Graduate Institute of Biomedical Informatics, College of Medical Science and Technology, Taipei Medical University, Taipei, Taiwan

OPEN ACCESS

Edited by:

San-Gang Wu,
First Affiliated Hospital of Xiamen
University, China

Reviewed by:

Emanuele Perrone,
A. Gemelli University Hospital
Foundation, Italy
Khalid El Bairi,
Mohamed Premier
University, Morocco

*Correspondence:

Tzu-Hao Chang
kevinchang@tmu.edu.tw
Hsien-Da Huang
huanghsienda@cuhk.edu.cn

Specialty section:

This article was submitted to
Women's Cancer,
a section of the journal
Frontiers in Oncology

Received: 24 August 2019

Accepted: 16 December 2019

Published: 21 January 2020

Citation:

Huang C-Y, Liao K-W, Chou C-H,
Shrestha S, Yang C-D, Chiew M-Y,
Huang H-T, Hong H-C, Huang S-H,
Chang T-H and Huang H-D (2020)
Pilot Study to Establish a Novel
Five-Gene Biomarker Panel for
Predicting Lymph Node Metastasis in
Patients With Early Stage Endometrial
Cancer. *Front. Oncol.* 9:1508.
doi: 10.3389/fonc.2019.01508

Introduction: In the United States and Europe, endometrial endometrioid carcinoma (EEC) is the most prevalent gynecologic malignancy. Lymph node metastasis (LNM) is the key determinant of the prognosis and treatment of EEC. A biomarker that predicts LNM in patients with EEC would be beneficial, enabling individualized treatment. Current preoperative assessment of LNM in EEC is not sufficiently accurate to predict LNM and prevent overtreatment. This pilot study established a biomarker signature for the prediction of LNM in early stage EEC.

Methods: We performed RNA sequencing in 24 clinically early stage (T1) EEC tumors (lymph nodes positive and negative in 6 and 18, respectively) from Cathay General Hospital and analyzed the RNA sequencing data of 289 patients with EEC from The Cancer Genome Atlas (lymph node positive and negative in 33 and 256, respectively). We analyzed clinical data including tumor grade, depth of tumor invasion, and age to construct a sequencing-based prediction model using machine learning. For validation, we used another independent cohort of early stage EEC samples ($n = 72$) and performed quantitative real-time polymerase chain reaction (qRT-PCR). Finally, a PCR-based prediction model and risk score formula were established.

Results: Eight genes (*ASRGL1*, *ESR1*, *EYA2*, *MSX1*, *RHEX*, *SCGB2A1*, *SOX17*, and *STX18*) plus one clinical parameter (depth of myometrial invasion) were identified for use in a sequencing-based prediction model. After qRT-PCR validation, five genes (*ASRGL1*, *RHEX*, *SCGB2A1*, *SOX17*, and *STX18*) were identified as predictive biomarkers. Receiver operating characteristic curve analysis revealed that these five genes can predict LNM. Combined use of these five genes resulted in higher diagnostic accuracy than use of any

single gene, with an area under the curve of 0.898, sensitivity of 88.9%, and specificity of 84.1%. The accuracy, negative, and positive predictive values were 84.7, 98.1, and 44.4%, respectively.

Conclusion: We developed a five-gene biomarker panel associated with LNM in early stage EEC. These five genes may represent novel targets for further mechanistic study. Our results, after corroboration by a prospective study, may have useful clinical implications and prevent unnecessary elective lymph node dissection while not adversely affecting the outcome of treatment for early stage EEC.

Keywords: endometrial cancer, lymph node metastasis, RNA sequencing, TCGA, prediction model

INTRODUCTION

As the gynecologic malignancy with the highest prevalence in Western countries (1) and Taiwan (2), endometrial cancer has become more common in recent years (3, 4). Endometrial endometrioid carcinoma (EEC) is the most prevalent histological type of endometrial cancer (5). The primary treatment for patients with medically operable EEC is surgical staging, which involves collecting peritoneal fluid for cytologic examination; inspecting the whole abdominal and pelvic cavities, after which any suspicious lesions are biopsied or excised; total extrafascial hysterectomy along with bilateral salpingo-oophorectomy; and lymph node evaluation (6). This procedure obtains complete pathologic and prognostic information for tailoring further adjuvant therapy.

The surgical treatment employed in patients with endometrial cancer has changed. Previously, lymph node assessment in EEC comprised dissection of the para-aortic and pelvic nodes in all patients (6). However, EEC is often diagnosed early, when the risk of lymph node metastasis (LNM) is low (5). In clinical stage I EEC, the incidence of pelvic and para-aortic nodal metastases is only 10 and 6%, respectively (7). Most patients with early stage EEC can avoid lymph node dissection to prevent morbidities. Nerve injury, prolonged operation time, lymphedema, blood loss, and lymphocyst formation are the principal morbidities associated with lymph node dissection (7–10). Thus, a more selective lymphadenectomy is suggested for patients with EEC to prevent overtreatment (11).

Sentinel lymph node (SLN) mapping can be an alternative to full lymphadenectomy in patients with EEC. According to the National Comprehensive Cancer Network guidelines (12), SLN mapping can identify metastatic lymph nodes in diseases confined to the uterus (13–16). SLN mapping has improved surgeons' detection of LNM in small-volume disease and has considerably reduced intraoperative and postoperative morbidity. However, the skills, knowledge, and attention to technical detail of the surgeon are crucial. Only institutions with the relevant expertise in SLN mapping should perform the procedure to minimize false negatives, which may compromise treatment outcomes (12). Most critically, SLN mapping is performed during surgery and is not a preoperative setting for the assessment of lymph node status.

At present, lymph node status is preoperatively assessed using clinical risk factors including tumor invasion depth, tumor grade, or tumor size, but in clinical practice, such assessment has a low positive predictive value (17–19). Scholars have concluded that image-based diagnostic tools—including computed tomography (20), magnetic resonance imaging (21, 22), and positron emission tomography (23)—are inadequate for detection of micrometastatic nodal disease because they have low sensitivity. Recent advances in molecular profiling have provided important insights into the biological nature of tumors. These advances have enabled researchers to investigate gene expression signatures as diagnostic tools for clinical decision-making (24) and prediction of outcomes, including tumor metastases (25–27).

In the current study, we investigated the clinicopathological parameters and molecular markers of LNM in patients with early clinical stage EEC by analyzing gene expression in surgical specimens from such patients. Such genetic signatures may help guide individualized therapies and reduce the surgical morbidities of patients with EEC.

METHODS

Cathay General Hospital Dataset Patient Enrollment and Sampling

The Institutional Review Board of Cathay General Hospital (CGH) approved the experimental protocols employed in this study. Prior to surgery, informed consent was obtained from all participants (IRB: CGH-P104102). Between March 2011 and December 2017, 113 patients with EEC who underwent staging surgery through either laparotomy or a minimally invasive procedure were enrolled. We included patients with (i) unambiguous histologic diagnosis of endometrioid carcinoma and absence of a mixed tumor type, (ii) tissue that was sufficient to enable RNA extraction, and (iii) no prior treatment history. Clinical TMN staging before surgery was evaluated according to the *American Joint Committee on Cancer Staging Manual, Seventh Edition* (28). Only patients receiving complete surgical staging surgery—that is, undergoing total hysterectomy and/or bilateral salpingo-oophorectomy, para-aortic lymphadenectomy, and bilateral pelvic lymph node dissection—were included. Our aim was to identify the molecular signature of early clinical

stage EEC. Therefore, only patients with clinical T1 tumors documented through preoperative magnetic resonance imaging were included.

Following hysterectomy, tissue specimens were collected and preserved in RNAlater[®] (Qiagen, Valencia, CA, USA) immediately for storage at -80°C until they were analyzed. The 2009 International Federation of Obstetrics and Gynecologic staging system was the basis of our EEC tumor staging and grading criteria (29). A gynecological pathologist (S-HH) reviewed the histopathology. Lymph node status was considered positive if metastasis was histologically proven and negative if no metastasis was detected in the surgical specimen and during at least the subsequent 1 year. Clinical data collected prior to hysterectomy regarding age, menopausal status, and body mass index were recorded. We examined the patients' detailed medical records for the period up to December 31, 2018. A multistep case-control study was designed to identify mRNA markers for predicting LNM in patients with EEC. **Figure 1** presents a study workflow schematic. Additionally, we applied the REMARK reporting recommendations for prognostic biomarkers (**Supplement Table 1**) (30).

RNA Sequencing and Data Processing

In the initial screening phase, we selected 6 patients with and 18 patients without LNM for RNA sequencing (**Table 1**). For RNA preparation, ~ 25 mg of tissue from an EEC tumor specimen underwent homogenization in a TissueLyser II bead mill (Qiagen) running at 30 Hz for 10–20 s and chilled lysis buffer from the AllPrep RNA/DNA/Protein Mini Kit (Qiagen) in accordance with the guidelines of the manufacturer. Quantification of the RNA elution was performed using a Nanodrop spectrophotometer (Thermo Fisher Scientific, MA, USA), and the elution was further examined using a Bioanalyzer 2200 (Agilent Technologies, Santa Clara, CA, USA), which

output an RNA integrity number between 1 (indicating lowest quality and most degradation) and 10 (indicating highest quality and least degradation) (31). Samples with an RNA integrity number < 7 —which may have substantially affected the results of sequencing (e.g., 3'-5' transcript bias and uneven gene coverage) and led to inaccurate conclusions—were not included in the analysis. A TruSeq mRNA Sample Preparation Kit V2 (Illumina, San Diego, CA, USA) used according to the manufacturer protocol was employed to construct RNA sequencing libraries. Briefly, the protocol for preparing these RNA sequencing libraries involved poly-A RNA isolation, RNA fragmentation, reverse transcription to cDNA using random primers, 3'-end adenylation, adapter ligation, and polymerase chain reaction (PCR) enrichment. The resultant cDNA libraries quantified through quantitative real-time (qRT)-PCR (using the Roche LightCycler[®] 480 System) and Qubit fluorometry (Invitrogen, Carlsbad, CA, USA). After they were purified, quantified, and validated, the libraries were sequenced on the Illumina Sequencing System (Next-seq 500 sequencing system, 75 single-end) according to the manufacturer's standard workflow. The RNA sequencing data were subject to quadripartite analysis involving (i) quality control, (ii) alignment to the reference genome, (iii) splice junction discovery and transcript assembly, and (iv) abundance estimation and normalization. We deposited the raw sequences of the EEC samples in the National Center for Biotechnology Information Sequence Read Archive (access number PRJNA554534).

Our ultimate goal was to identify clinically predictive markers by using RT-PCR to measure specific gene expression in surgical specimens of patients with EEC. We validated the differentially expressed genes (DEGs) identified from next-generation sequencing data by using RT-PCR. Therefore, the selection of the reference gene for comparison was critical. The interpretation of RNA sequencing results is highly dependent

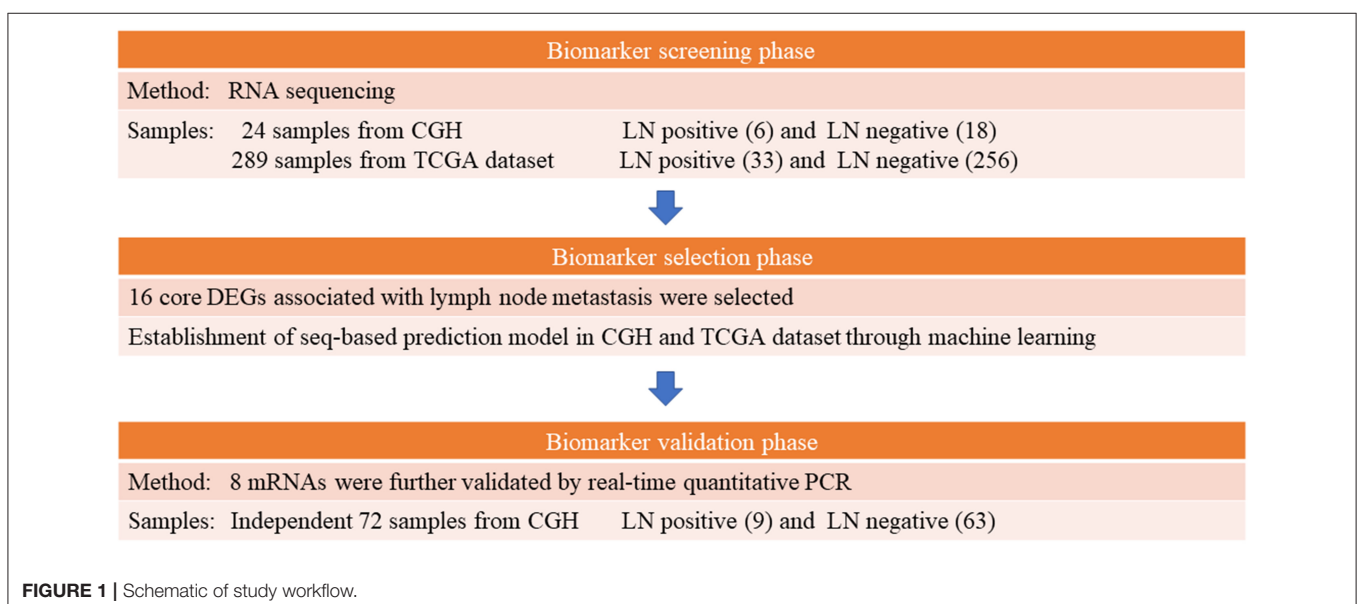


TABLE 1 | Basic characteristics of patients from CGH for RNA sequencing (CGH dataset, $n = 24$).

| Variable | Lymph node positive ($n = 6$) | Lymph node negative ($n = 18$) | p |
|-------------------------------|------------------------------------|-------------------------------------|----------------------|
| Age, years (median) | 61.8 | 54.8 | 0.73 ^a |
| Body mass index (mean) | 23.7 | 24.2 | 0.55 ^b |
| Menopausal status, n (%) | | | 1.00 ^c |
| Premenopausal | 1 (16.7%) | 3 (16.7%) | |
| Perimenopausal | 5 (83.3%) | 15 (83.3%) | |
| Surgical approach, n (%) | | | 1.00 ^c |
| Open | 5 (83.3%) | 13 (72.2%) | |
| Minimally invasive | 1 (16.6%) | 5 (27.8%) | |
| Tumor grade | | | 0.007 ^c |
| 1 | 1 (16.7%) | 15 (83.3%) | |
| 2 | 3 (50%) | 2 (11.1%) | |
| 3 | 2 (33.3%) | 1 (5.6%) | |
| Myometrial invasion | | | 0.061 ^c |
| <1/2 | 1 (16.7%) | 12 (66.7%) | |
| $\geq 1/2$ | 5 (83.3%) | 6 (33.3%) | |
| Lymphovascular space invasion | | | <0.0001 ^c |
| Present | 6 (100%) | 5 (27.8%) | |
| Absent | 0 (0%) | 13 (72.2%) | |

^a*t*-test.^bMann–Whitney *U*-test.^cFisher-exact test.**TABLE 2** | Basic characteristics of TCGA patients with uterine corpus endometrial carcinoma (TCGA dataset, $n = 289$).

| Variable | Lymph node positive ($n = 33$) | Lymph node negative ($n = 256$) | p |
|------------------------------|-------------------------------------|--------------------------------------|----------------------|
| Age, years (mean) | 59.4 | 63.0 | 0.69 ^a |
| Body mass index (mean) | 33.1 | 34.8 | 0.46 ^b |
| Menopausal status, n (%) | | | 0.58 ^c |
| Premenopausal | 4 (12.1%) | 19 (7.4%) | |
| Perimenopausal | 2 (6.1%) | 10 (3.9%) | |
| Postmenopausal | 26 (78.8%) | 210 (82.0%) | |
| Indeterminate & NA | 1 (3.0%) | 17 (6.6%) | |
| Surgical approach, n (%) | | | 0.005 ^c |
| Open | 27 (82.8%) | 142 (55.5%) | |
| Minimally invasive | 6 (18.2%) | 109 (42.6%) | |
| NA | | 5 (1.9%) | |
| Tumor grade, n (%) | | | 0.001 ^c |
| 1 | 2 (6.1%) | 78 (30.5%) | |
| 2 | 8 (24.2%) | 79 (30.9%) | |
| 3 | 23 (69.7%) | 99 (38.7%) | |
| Myometrial invasion, n (%) | | | <0.0001 ^c |
| <1/2 | 9 (27.3%) | 165 (64.5%) | |
| $\geq 1/2$ | 24 (72.7%) | 91 (35.5%) | |

^a*t*-test.^bMann–Whitney *U*-test.^cChi-square test.

upon the reference genes employed for normalization (32). To identify a gene that was highly expressed in both The Cancer Genome Atlas (TCGA; www.cancergenome.nih.gov) and CGH datasets with minimal variability between samples, NormFinder (<https://moma.dk/normfinder-software>) was used for identification of the appropriate normalization gene (33). Using NormFinder, candidate normalization genes could be ranked by the stability of their expression in the CGH and TCGA datasets. *RPL19* was identified as a suitable reference gene in this study.

TCGA Dataset Databases

We analyzed mRNA expression profiles among patients present in TCGA uterine corpus endometrial carcinoma series. Data were collected in August 2016. Only cases involving a histologic EEC diagnosis and with complete clinical information regarding tumor grade, lymph node status, and myometrial invasion depth were selected for analysis. Additionally, we selected only patients with clinical stage I (negative lymph nodes) or IIIC (positive lymph nodes) disease for comparison. After filtering out excluded cases, 289 patients with endometrioid carcinoma were selected for analyses (Table 2). TCGA comprises three levels of data. We downloaded level 3 RNAseqV2 normalized data, containing gene expression information, for the aforementioned 289 patients. The patients and their clinical data were merged into a single dataset to eliminate duplicate entries.

RNA Sequence Data Processing and Differential Expression Analysis

We corrected the raw data downloaded from TCGA for background and normalization. The expression level was calculated using the Bioconductor package. After the raw data were preprocessed and normalized to the *RPL19* reference gene, the DEGs between patients with and without LNM were identified. The DEGs were screened using the cut-off criteria, which were set as a fold change (FC) ≥ 1.5 and an adjusted Wilcoxon rank-sum *p*-value of <0.05.

Sequencing-Based Model Construction and Feature Selection Through Machine Learning

The DEGs in both the CGH and TCGA datasets were selected as core DEGs. The prediction model was tuned using variables including tumor grade, myometrial invasion depth, and age along with gene expression data. Although, lymphovascular space invasion (LVSI) was previously considered a crucial determinant of LNM (34), LVSI information is unavailable in TCGA. In addition, the LVSI status is not routinely reported in small presurgical biopsy specimens collected through dilatation and curettage or endometrial biopsy; therefore, we did not include LVSI in our model. Four machine learning methods—namely, the random forest (35), sequential minimal optimization, J48, and Naïve Bayes methods—were employed for predictions. Threefold cross validation with 10 repetitions was used to assess these methods' performance.

High sensitivity was critical to this study's model. However, the CGH and TCGA datasets have imbalanced numbers of patients with and without LNM. Therefore, we employed the WEKA cost-sensitive classifier to enhance the penalty for false-negative samples. Finally, the predictive performance of the sequencing-based model with the CGH and TCGA datasets was examined.

qRT-PCR Validation and PCR-Based Prediction Model Construction Through Machine Learning

To validate our findings, we performed qRT-PCR on clinical EEC tissues from another independent set of patients with early stage EEC (Table 3). The mRNA expression levels of eight genes—*ASRGL1*, *ESR1*, *EYA2*, *MSX1*, *RHEX*, *SCGB2A1*, *SOX17*, and *STX18*—were evaluated using TaqMan qRT-PCR in triplicate. We performed qRT-PCR following a previous study (36). Briefly, for amplification, we used 2 μ L of cDNA at 1:10 dilution and TaqMan real-time probes and primers (Applied Biosystems, Foster City, CA, USA). Supplement Table 2 lists the primer sequences. TaqMan Universal PCR Master Mix (ABgene, Rochester, NY, USA) and an ABI Prism 7900-HT Sequence Detection System (Applied Biosystems) were employed for all PCR. For each sample, individual genes' relative mRNA expression levels were normalized to *RPL19* expression levels through the comparative Ct ($2^{-\Delta\Delta Ct}$) method (36). Five DEGs in the validation cohort were used for further machine learning and prediction model construction. The four machine learning methods mentioned in the previous subsection were employed to construct a PCR-based prediction model.

Risk Score Calculation and Predictive Model Construction

For clinical application of the mRNA signature, we used binary logistic regression analysis to generate a comprehensive set of multi-mRNA markers (combination of five mRNA markers) for the prediction of lymph node status. Subsequently, the expression levels of the mRNA markers were linearly combined into an LNM risk score as follows. Risk score = $0.249 \times (\text{Ct of } ASRGL1 - \text{Ct of } RPL19) + 0.166 \times (\text{Ct of } RHEX - \text{Ct of } RPL19) + 0.434 \times (\text{Ct of } SCGB2A1 - \text{Ct of } RPL19) + 0.02 \times (\text{Ct of } SOX17 - \text{Ct of } RPL19) - 0.51 \times (\text{Ct of } STX18 - \text{Ct of } RPL19) - 4.06$, where Ct is the threshold cycle of each gene. Area under the receiver operating characteristic (ROC) curve (AUC) for each (or the combined) mRNA marker was calculated using a 95% confidence interval to evaluate the performance of the mRNA expression markers in identifying patients with LNM. The optimal cutoffs and related specificity, sensitivity, and positive and negative predictive value (PPV and NPV) from ROC curves were determined using a common method. We considered $p < 0.05$ statistically significant.

Statistical Analysis

We performed statistical analysis in SPSS version 15.0 (SPSS Inc., Chicago, IL, USA) for Windows. Intergroup differences in continuous variables were evaluated through unpaired *t* or

TABLE 3 | Basic characteristics of validation cohort ($n = 72$).

| Variable | Lymph node positive ($n = 9$) | Lymph node negative ($n = 63$) | <i>p</i> |
|---------------------------------|---------------------------------|----------------------------------|----------------------|
| Age, years (mean) | 58.7 | 54.0 | 0.45 ^a |
| Body mass index (mean) | 24.3 | 25.6 | 0.18 ^b |
| Menopausal status, <i>n</i> (%) | | | |
| Premenopausal | 2 (22.2%) | 18 (28.6%) | 1.00 ^c |
| Perimenopausal | 7 (77.8%) | 45 (71.4%) | |
| Surgical approach, <i>n</i> (%) | | | 1.00 ^c |
| Open | 2 (22.2%) | 17 (27.0%) | |
| Minimally invasive | 7 (77.8%) | 46 (73.0%) | |
| Tumor grade | | | 0.0001 ^c |
| 1 | 1 (11.1%) | 50 (79.4%) | |
| 2 | 4 (44.4%) | 8 (12.7%) | |
| 3 | 4 (44.4%) | 5 (7.9%) | |
| Myometrial invasion | | | 0.023 ^c |
| Inner 1/2 | 2 (22.2%) | 42 (66.7%) | |
| Outer 1/2 | 7 (77.8%) | 21 (33.3%) | |
| Number of lymph nodes removed | | | 0.08 ^b |
| Median (range) | 26 (14–36) | 17 (11–57) | |
| Lymphovascular space invasion | | | <0.0001 ^c |
| Present | 9 (100%) | 17 (27.0%) | |
| Absent | 0 (0%) | 46 (73.0%) | |

^a*t*-test.

^bMann–Whitney *U*-test.

^cFisher-exact test.

Mann–Whitney *U* tests after data normality was confirmed using the Shapiro–Wilk test, and differences in categorical variables were evaluated through Fisher-exact or chi-square tests. For gene expression analysis, the Bioconductor package (version 3.6) in R (version 3.4.3) was used to identify DEGs between patients with and without LNM (37). The DEGs were screened using the cut-off criteria, which were set as $FC \geq 2$ and adjusted Wilcoxon rank-sum $p < 0.05$ in CGH dataset. We employed R packages to statistically analyze the RNA sequencing and clinical data.

RESULTS

Study Population's Basic Characteristics

The basic clinicopathological parameters of the patients in the CGH dataset ($n = 24$) are listed in Table 1. These patients were divided into those with negative [LN (-); 18 patients [75%]] or positive lymph nodes [LN (+); 6 patients [25%]]. These two groups were similar in body mass index, age, surgical route, and menopausal status. The LN (+) group had higher proportions of high-grade tumors ($p = 0.007$) and positive LVSI status ($p < 0.0001$).

From TCGA data, we identified 289 EEC patients with comprehensive lymph node status, genomic expression, and clinical data. Of these 289 patients, 256 (88.6%) were LN (-) and 33 (11.4%) were LN (+) (Table 2). Patients in TCGA with LNM were more likely to have undergone an open procedure (p

TABLE 4 | LN (+) versus LN (-) DEGs in CGH and TCGA datasets.

| | CGH dataset (n = 24) | | | TCGA dataset (n = 289) | | |
|----------------|----------------------|------------------|----------|------------------------|-------------------|----------|
| | LN (+) n = 6 | LN (-) n = 18 | p | LN (+) n = 33 | LN (-) n = 256 | p |
| <i>SCGB2A1</i> | -6.48 | -1.35 | 6.54E-04 | -1.75 | -0.26 | 5.35E-03 |
| <i>RHEX</i> | -9.44 | -4.35 | 5.94E-05 | -5.21 | -4.48 | 4.14E-02 |
| <i>EYA2</i> | -9.37 | -5.30 | 4.46E-04 | -4.25 | -3.25 | 2.74E-02 |
| <i>MSX1</i> | -6.96 | -3.59 | 1.18E-02 | -3.22 | -1.94 | 6.78E-04 |
| <i>SOX17</i> | -7.58 | -3.90 | 4.46E-04 | -3.21 | -2.39 | 4.44E-02 |
| <i>ASRGL1</i> | -7.56 | -3.69 | 2.97E-05 | -4.37 | -3.55 | 8.31E-03 |
| <i>STX18</i> | -7.15 | -4.09 | 9.51E-04 | -4.73 | -3.92 | 1.17E-03 |
| <i>ESR1</i> | -8.49 | -5.49 | 4.43E-03 | -5.40 | -4.58 | 8.36E-03 |
| <i>TMEM101</i> | -8.07 | -5.16 | 3.37E-03 | -5.38 | -4.43 | 1.83E-03 |
| <i>CFI</i> | -8.20 | -4.94 | 2.82E-04 | -5.32 | -4.65 | 1.81E-02 |
| <i>LRIG1</i> | -8.12 | -5.62 | 6.54E-04 | -4.85 | -4.18 | 2.82E-02 |
| <i>ELP3</i> | -8.03 | -5.97 | 4.46E-04 | -5.31 | -4.69 | 1.84E-03 |
| <i>BASP1</i> | -7.87 | -5.81 | 7.40E-03 | -3.53 | -4.14 | 7.33E-03 |
| <i>SERINC2</i> | -5.68 | -3.92 | 5.77E-03 | -2.87 | -2.13 | 1.83E-02 |
| <i>NOXA1</i> | -7.51 | -5.92 | 4.43E-03 | -5.26 | -4.47 | 3.27E-04 |
| <i>HOXB5</i> | -7.42 | -5.18 | 1.18E-02 | -5.93 | -4.83 | 6.64E-03 |

Data are expressed as \log_2 (intensity of corresponding gene/intensity of RPL19).

= 0.005), a high-grade tumor ($p = 0.001$), or deep myometrial invasion ($p < 0.0001$; **Table 2**).

DEGs Between Node-Positive and Node-Negative EEC in the CGH and TCGA Datasets

We conducted RNA sequencing for 24 patients in the CGH dataset. On average, we obtained 20.8 million reads from each of these 24 libraries. We assessed the quality of the reads by using FastQC, and all reads were discovered to have a Phred score higher than 30. **Supplement Table 3** presents alignment and mapping quality metrics produced by Qualimap over the output BAM files through STAR. We discovered 779 DEGs between the LN (+) and LN (-) cancer tissues in the CGH dataset ($FC \geq 2$ and $p < 0.05$ after adjustment for false discovery rate). Additionally, data on mRNA expression in patients with EEC were obtained from TCGA and included 33 LN (+) and 256 LN (-) EEC tissues. Between these tissues, 41 DEGs were identified ($FC \geq 1.5$ and $p < 0.05$ after adjustment for false discovery rate).

Sequencing-Based Model Construction and Feature Selection

Sixteen DEGs were identified in both the CGH and TCGA datasets (**Table 4**). A heat map of these 16 DEGs in the CGH dataset is presented in **Figure 2**. As stated previously, gene expression data and age, grade and depth of myometrial invasion, and other clinical information were used to tune the predictive model. Machine learning was performed to select features. One clinical feature (depth of myometrial invasion) and eight DEGs (*ASRGL1*, *ESR1*, *EYA2*, *MSX1*, *RHEX*, *SCGB2A1*, *SOX17*, and *STX18*) were identified. For the CGH dataset, the model achieved 100% sensitivity and 88.9% specificity, whereas for TCGA data,

these were 97% sensitivity and 38.7% specificity, respectively. The contingency tables of the agreement between the datasets along with the predictive model are displayed in **Table 5**. Only one patient with LNM in TCGA (**Table 5A**) was incorrectly classified; other LNM predictions for both datasets (**Tables 5A,B**) were correct.

PCR Validation and PCR-Based Prediction Model Construction

We used another independent cohort of 72 patients with early stage EEC for qRT-PCR validation of the eight identified DEGs (*ASRGL1*, *ESR1*, *EYA2*, *MSX1*, *RHEX*, *SCGB2A1*, *SOX17*, and *STX18*). Other 17 patients without complete pelvic and para-aortic lymphadenectomy were excluded. The cohort's ($n = 72$) clinical characteristics are listed in **Table 3**. All 72 patients had clinical T1 tumors (confined to the uterus), with 63 patients (87.5%) being LN (-) and 9 (12.5%) being LN (+). The medians of lymph nodes harvested were 17 in the LN (-) group and 26 in the LN (+) group. Similar to those in TCGA data, these patients with LNM were more likely to have high-grade tumors ($p = 0.0001$) and deep myometrial invasion ($p = 0.023$). The LN (+) patients also had a higher proportion of LVSI ($p < 0.0001$).

The relative expression of the eight DEGs is illustrated in **Figure 3A**. Five DEGs were observed between the LN (+) and LN (-) groups. Significantly lower relative expression of *ASRGL1* ($p = 0.002$), *RHEX* ($p = 0.005$), *SCGB2A1* ($p = 0.002$), *SOX17* ($p = 0.003$), and *STX18* ($p = 0.048$) was discovered in the LN (+) group. After the expression of these five genes and the one clinical parameter (depth of myometrial invasion) were incorporated, the performance of the various machine learning model was determined, as shown in **Table 6**. The random forest model had the highest sensitivity and specificity: 92 and 31%, respectively.

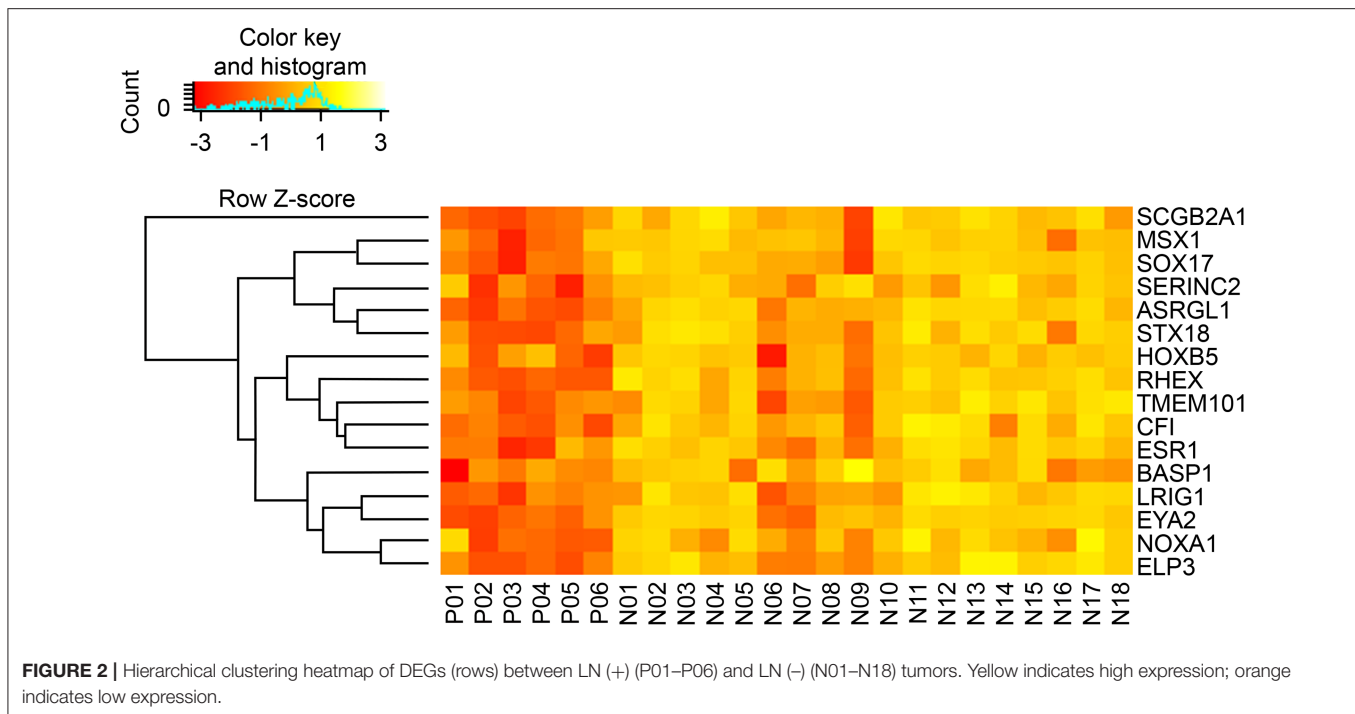


FIGURE 2 | Hierarchical clustering heatmap of DEGs (rows) between LN (+) (P01–P06) and LN (-) (N01–N18) tumors. Yellow indicates high expression; orange indicates low expression.

TABLE 5 | Contingency table of level of agreement between the two datasets (A. TCGA and B. CGH) and sequencing-based prediction model.

| | | Predicted as | |
|----------|---------------------|--------------|----------|
| | | Positive | Negative |
| A | | | |
| | TCGA dataset | | |
| | Positive | 32 | 1 |
| | Negative | 157 | 99 |
| B | | | |
| | CGH dataset | | |
| | Positive | 6 | 0 |
| | Negative | 2 | 16 |

Patients classified correctly (Pos-Pos and Neg-Neg) are highlighted in yellow. Negative, patients with negative lymph node status; positive, patients with positive lymph node status.

The accuracy, PPV, and NPV of the PCR-based prediction model were 40, 16, and 97%, respectively.

Construction of mRNA Signature and Risk Score Formulation

To evaluate the performance of these mRNA markers for predicting LNM in EEC, ROC analysis was performed (Figures 3B,C). The AUC, sensitivity, and specificity of each mRNA marker are listed in Table 7. The AUC of the five mRNAs ranged from 0.705 to 0.829, indicating that all five markers could distinguish LN (+) patients from LN (-) patients (Figure 3B). The sensitivity and specificity of the mRNA markers for detecting patients with LN (+) ranged from 77.8 to 88.9% and 54.0 to 74.6%, respectively. Among the five mRNAs studied, *SCGB2A1* was identified as the best marker for identifying LN (+) patients (AUC = 0.829) because it had the highest sensitivity and specificity.

To enable these mRNA signatures to be applied in clinical practice and to improve the mRNA prediction efficacy, for each patient in the validation cohort a risk score was calculated on the basis of the five-mRNA combination. On the basis of the optimal risk score cut-off of -1.77 , the 72 patients with EEC in the validation cohort were classified as being at high risk ($n = 18$; risk score > -1.77) or low risk ($n = 54$; risk score < -1.77). The distribution of clinicopathological characteristics in relation to risk score is displayed in Supplement Table 4. The risk score was significantly higher for patients with grade 3 tumors than those with grade 1 or 2 tumors (-3.06 , -2.17 , and -1.65 for grade 1–3 tumors, respectively, $p = 0.034$). The LN (+) group also had significantly higher risk scores than the LN (-) group (-0.36 vs. -3.08 , respectively, $p < 0.0001$). Additionally, the predictive performance of the combined set of five mRNAs was significantly higher than that of any single mRNA marker and achieved an excellent AUC of 0.898, sensitivity of 88.9%, specificity of 84.1%, accuracy of 84.7%, PPV of 44.4%, and NPV of 98.1% (Figure 3C and Table 7).

DISCUSSION

LNM is a critical prognostic factor for EEC, and the status of lymph nodes is an essential consideration when making decisions about personalized treatment for EEC, especially in patients with the early stage of the disease. Accurately diagnosing LNM before surgical staging remains difficult. A test that can accurately identify LNM in patients with early stage EEC before staging surgery would have substantial clinical value.

Most previous risk models for LNM prediction have been constructed from intraoperative or final pathological findings

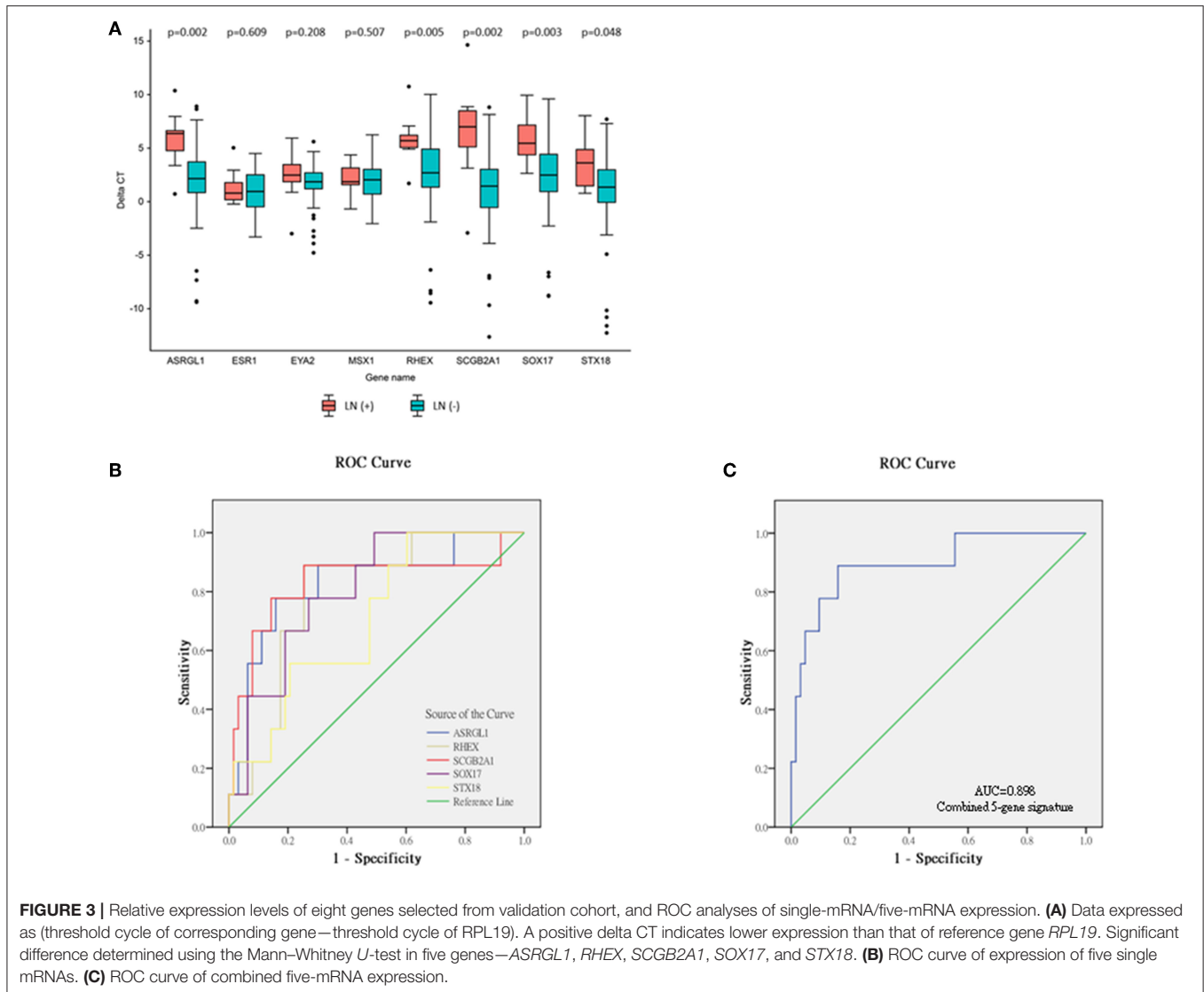


TABLE 6 | Prediction performance of PCR-based model using five-gene panel with various machine learning approaches (SMO, sequential minimal optimization).

| | Naive Bayes (%) | J48 (%) | SMO (%) | Random forest (%) |
|-------------|-----------------|---------|---------|-------------------|
| Accuracy | 62 | 53 | 36 | 40 |
| Sensitivity | 83 | 83 | 89 | 92 |
| Specificity | 61 | 49 | 29 | 31 |
| PPV | 23 | 19 | 15 | 16 |
| NPV | 96 | 95 | 95 | 97 |

TABLE 7 | Performance of mRNA expression markers in predicting lymph node status in EEC.

| Marker | Sensitivity(%) | Specificity(%) | AUC (95% CI) | PPV(%) | NPV(%) |
|------------------------------|----------------|----------------|---------------------|--------|--------|
| <i>ASRGL1</i> | 88.9 | 69.8 | 0.827 (0.671–0.983) | | |
| <i>RHEX</i> | 88.9 | 74.6 | 0.792 (0.662–0.922) | | |
| <i>SCGB2A1</i> | 88.9 | 74.6 | 0.829 (0.644–1.014) | | |
| <i>SOX17</i> | 77.8 | 73.0 | 0.804 (0.678–0.930) | | |
| <i>STX18</i> | 88.9 | 54.0 | 0.705 (0.544–0.867) | | |
| Combined five mRNA signature | 88.9 | 84.1 | 0.898 (0.781–1.014) | 44.4 | 98.1 |

(38–40). In a multicenter retrospective study, the Korean Gynecologic Oncology Group developed a preoperative model for predicting LNM in endometrial cancer (KGOG-2014) and identified serum CA-125 level plus magnetic resonance imaging as a combination achieving accurate identification of low LNM risk among patients with endometrial cancer (17). However, that study included patients with non-endometrioid histology.

The molecular mechanisms underlying the development, invasion, and metastasis of EEC have not yet been fully characterized and consequently are poorly understood (41). The clinical heterogeneity of patients with EEC probably reflects variation at the molecular level (42); thus, the metastasis mechanisms of EEC tumors likely vary as well.

We employed a genome-wide strategy for characterization of the molecular signatures associated with LNM in early stage EEC; mRNA sequencing data were comprehensively analyzed to identify five mRNAs closely associated with LNM in patients with early stage EEC. Assuming that a combination of several mRNAs has higher sensitivity and specificity than a single mRNA for prediction of LNM, we formulated a risk score on the basis of these five mRNAs. Using the five-mRNA signature, we could effectively categorize patients in our validation cohort as high or low risk, and favorable reliability was achieved in our prediction of LNM in EEC. Logically, our next step is to prospectively validate the use of this gene expression signature in presurgical biopsy specimens of patients with EEC.

The most vital contribution of our study is preoperative LNM risk determination. In most cases, endometrial cancer is diagnosed through endometrial biopsy (43, 44) or fractional dilatation and curettage (43, 45). Once EEC has been diagnosed, presurgical biopsy specimens may be used for molecular testing for our proposed five-gene PCR-based markers. Thus, patients at high risk can be identified before staging surgery. Additionally, low-LNM-risk patients, for whom systemic lymph node dissection may be unnecessary, can be identified. In addition, SLN mapping can be included in prospective studies to reduce surgical morbidities.

Another advantage of our study is its homogeneity in the following respects. First, all patients had pure endometrioid carcinoma, which was confirmed after staging surgery. Second, all patients received complete surgical staging operations, which included para-aortic and pelvic lymph node dissection. Third, gynecologic oncologists performed all treatments to an adequate extent of lymphadenectomy. In the LN (+) and LN (-) groups, the median numbers of lymph nodes removed were 26 and 17, respectively. One study concluded removal of 11 or more pelvic lymph nodes to be adequate (46).

We identified five PCR-based gene markers (*ASRGL1*, *RHEX*, *SCGB2A1*, *SOX17*, and *STX18*) for predicting LNM in patients with early stage EEC. These five markers and their associations with lymph node status in patients with EEC may reveal novel targets for therapy and further investigation of endometrial cancer.

L-asparaginase is an enzyme encoded by the *ASRGL1* gene in humans. Asparagine and glutamine is hydrolyzed by this enzyme into aspartic and glutamic acid, respectively (47). The enzyme has been employed in non-Hodgkin's lymphoma and acute lymphoblastic leukemia treatment (48–50). Fonnes et al. (51) analyzed *ASRGL1* expression by using immunohistochemistry in curettage specimens from 1,144 women with endometrial carcinoma and discovered LNM to be more common in patients with low *ASRGL1* expression than in those with high *ASRGL1* expression. Low *ASRGL1* expression could also predict LNM independently. This finding is consistent with ours. *RHEX*, also known as *C1orf186*, acts as a transduction factor of the EPO–EPOR signaling pathway and promotes erythroid cell differentiation (52). In humans, uterine *RHEX* expression is high, and the abundance of *RHEX* is determined by the stage of a woman's menstrual cycle (53).

SCGB2A1, also known as mammaglobin B (*MGB-2*), is a secretoglobin involved in adenocarcinoma development when

the tumor originates in organs such as the ovary and endometrium (54, 55). Tassi et al. (54) employed RT-PCR to analyze *MGB-2* expression in EEC tumors with different grades of differentiation and found this gene was expressed in 10 of 11 grade 1 (91%), 16 of 17 in grade 2 (94%), and 6 of 22 grade 3 (27%) EEC cases. Pairwise comparisons of *MGB-2* gene expression levels between the grades 1 or 2 cases and grade 3 cases revealed significantly higher levels in the grades 1 and 2 cases (grade 2 vs. grade 3: $p < 0.001$; grade 1 vs. grade 3: $p = 0.016$). In our series, we discovered that patients with LNM have lower *SCGB2A1* expression levels. This can be attributed to the higher tumor grade of patients with LNM. *SOX17* is a critical antagonist and inhibitor of the Wnt signaling pathway (56). Endometrial tumors expressing higher *SOX17* levels are associated with higher recurrence-free survival rates; conversely, progression of endometrial cancer may be promoted by low *SOX17* expression through downregulation of *MAML3* expression and Wnt signaling (57). A soluble N-ethylmaleimide-sensitive factor attachment protein receptor, *STX18* is involved in several cellular activities, such as the cell cycle and organelle assembly (58). Downregulation of *STX18* was reported as significantly increasing growth of the MCF-7 human breast cancer cells (59).

These reports combined with our present analysis support the appropriateness of using our five-gene signature for predicting LNM in EEC. Nonetheless, our study may have some shortcomings. The primary shortcoming is the small sample size. Only 72 patients were included in the validation cohort. Prior to clinical application of the mRNA signature, larger independent cohorts recruited from multiple centers should be used to confirm the signature's predictive performance. Another shortcoming is that we analyzed only Asian patients. Whether our finding is applicable to non-Asian populations is unknown. Compared with Caucasian populations, other populations with endometrial cancer are at higher risk for LNM and have poorer prognoses (60, 61). Whether ethnic characteristics affect the predictive results is unclear, but we recommend testing our model in non-Asian populations.

In summary, after comprehensive analysis, we established a model with a five-gene expression profile that can serve as a useful tool for predicting LNM in patients with early stage EEC. The predictive model should be tested prospectively and used in conjunction with the present clinical assessments to optimize individual treatment plans for patients with EEC before the first treatment.

DATA AVAILABILITY STATEMENT

The data used in this study can be obtained from the National Center for Biotechnology Information Sequence Read Archive (bioproject access number: PRJNA554534).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by institutional review board of Cathay General

Hospital (CGH). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

C-YH, T-HC, and H-DH designed the study and provided the ideas. C-YH, T-HC, K-WL, and H-DH designed the experiments. C-YH and S-HH provided materials. SS, C-DY, M-YC, H-TH, and H-CH conducted the experiments. C-YH, C-HC, and T-HC performed the analysis and created and adjusted the figures and tables. C-YH and T-HC wrote the manuscript. All authors read and approved the final manuscript.

REFERENCES

- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin.* (2018) 68:7–30. doi: 10.3322/caac.21442
- Health Promotion Administration, Ministry of Health and Welfare. *Cancer Registry Annual Report*. Ministry of Health and Welfare, Taipei (2015).
- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* (2018) 68:394–424. doi: 10.3322/caac.21492
- Lee J-Y, Kim E-Y, Jung K-W, Shin A, Chan KKL, Aoki D, et al. Trends in gynecologic cancer mortality in East Asian regions. *J Gynecol Oncol.* (2014) 25:174–82. doi: 10.3802/jgo.2014.25.3.174
- Creasman WT, Odicino F, Maisonneuve P, Quinn MA, Beller U, Benedet JL, et al. Carcinoma of the corpus uteri. FIGO 26th annual report on the results of treatment in gynecological cancer. *Int J Gynaecol Obstet.* (2006) 95:S105–S43. doi: 10.1016/S0020-7292(06)60031-3
- American College of Obstetricians and Gynecologists. ACOG practice bulletin, clinical management guidelines for obstetrician-gynecologists, number 65, august 2005: management of endometrial cancer. *Obstet Gynecol.* (2005) 106:413–25. doi: 10.1097/00006250-200508000-00050
- Morrow CP, Bundy BN, Kurman RJ, Creasman WT, Heller P, Homesley HD, et al. Relationship between surgical-pathological risk factors and outcome in clinical stage I and II carcinoma of the endometrium: a gynecologic oncology group study. *Gynecol Oncol.* (1991) 40:55–65. doi: 10.1016/0090-8258(91)90086-K
- Abu-Rustum NR, Alektiar K, Iasonos A, Lev G, Sonoda Y, Aghajanian C, et al. The incidence of symptomatic lower-extremity lymphedema following treatment of uterine corpus malignancies: a 12-year experience at memorial sloan-kettering cancer center. *Gynecol Oncol.* (2006) 103:714–8. doi: 10.1016/j.ygyno.2006.03.055
- Homesley HD, Kadar N, Barrett RJ, Lentz SS. Selective pelvic and periaortic lymphadenectomy does not increase morbidity in surgical staging of endometrial carcinoma. *Am J Obstet Gynecol.* (1992) 167:1225–30. doi: 10.1016/S0002-9378(11)91693-7
- Orr JW, Holimon JL, Orr PF. Stage I corpus cancer: is teletherapy necessary? *Am J Obstet Gynecol.* (1997) 176:777–88. doi: 10.1016/S0002-9378(97)70601-X
- Soliman PT, Frumovitz M, Spannuth W, Greer MJ, Sharma S, Schmeler KM, et al. Lymphadenectomy during endometrial cancer staging: practice patterns among gynecologic oncologists. *Gynecol Oncol.* (2010) 119:291–4. doi: 10.1016/j.ygyno.2010.07.011
- NCCN Guidelines, *Uterine Neoplasms* (2018). Available online at: <https://www2.tri-kobe.org/nccn/guideline/gynecological/english/uterine.pdf>
- Ballester M, Koskas M, Coutant C, Chereau E, Seror J, Rouzier R, et al. Does the use of the 2009 FIGO classification of endometrial cancer impact on indications of the sentinel node biopsy? *BMC Cancer.* (2010) 10:465. doi: 10.1186/1471-2407-10-465
- Barlin JN, Khoury-Collado F, Kim CH, Leitao MM, Chi DS, Sonoda Y, et al. The importance of applying a sentinel lymph node mapping algorithm in endometrial cancer staging: beyond removal of blue nodes. *Gynecol Oncol.* (2012) 125:531–5. doi: 10.1016/j.ygyno.2012.02.021
- Holloway RW, Abu-Rustum NR, Backes FJ, Boggess JF, Gotlieb WH, Lowery WJ, et al. Sentinel lymph node mapping and staging in endometrial cancer: a society of gynecologic oncology literature review with consensus recommendations. *Gynecol Oncol.* (2017) 146:405–15. doi: 10.1016/j.ygyno.2017.05.027
- How J, Lau S, Press J, Ferenczy A, Pelmus M, Stern J, et al. Accuracy of sentinel lymph node detection following intra-operative cervical injection for endometrial cancer: a prospective study. *Gynecol Oncol.* (2012) 127:332–7. doi: 10.1016/j.ygyno.2012.08.018
- Kang S, Kang WD, Chung HH, Jeong DH, Seo SS, Lee JM, et al. Preoperative identification of a low-risk group for lymph node metastasis in endometrial cancer: a Korean gynecologic oncology group study. *J Clin Oncol.* (2012) 30:1329–34. doi: 10.1200/JCO.2011.38.2416
- Mariani A, Dowdy SC, Cliby WA, Gostout BS, Jones MB, Wilson TO, et al. Prospective assessment of lymphatic dissemination in endometrial cancer: a paradigm shift in surgical staging. *Gynecol Oncol.* (2008) 109:11–8. doi: 10.1016/j.ygyno.2008.01.023
- Todo Y, Okamoto K, Hayashi M, Minobe S, Nomura E, Hareyama H, et al. A validation study of a scoring system to estimate the risk of lymph node metastasis for patients with endometrial cancer for tailoring the indication of lymphadenectomy. *Gynecol Oncol.* (2007) 104:623–8. doi: 10.1016/j.ygyno.2006.10.002
- La Fianza A, Di Maggio EM, Preda L, Coscia D, Tateo S, Campani R. Clinical usefulness of CT in the treatment of stage I endometrial carcinoma. *Radiol Med.* (1997) 93:567–71.
- Frei KA, Kinkel K, Bonel HM, Lu Y, Zaludek C, Hricak H. Prediction of deep myometrial invasion in patients with endometrial cancer: clinical utility of contrast-enhanced MR imaging—a meta-analysis and bayesian analysis. *Radiology.* (2000) 216:444–9. doi: 10.1148/radiology.216.2.r00au17444
- Rockall AG, Sohaib SA, Harisinghani MG, Babar SA, Singh N, Jeyarajah AR, et al. Diagnostic performance of nanoparticle-enhanced magnetic resonance imaging in the diagnosis of lymph node metastases in patients with endometrial and cervical cancer. *J Clin Oncol.* (2005) 23:2813–21. doi: 10.1200/JCO.2005.07.166
- Nogami Y, Banno K, Irie H, Iida M, Kisu I, Masugi Y, et al. The efficacy of preoperative positron emission tomography-computed tomography (PET-CT) for detection of lymph node metastasis in cervical and endometrial cancer: clinical and pathological factors influencing it. *Jpn J Clin Oncol.* (2015) 45:26–34. doi: 10.1093/jjco/hyu161
- Wu SG, Zhang WW, Wang J, Dong Y, Chen YX, He ZY. Effect of 21-gene recurrence score in decision-making for surgery in early stage breast cancer. *Onco Targets Ther.* (2019) 12:2071–8. doi: 10.2147/OTT.S199625
- van Hooff SR, Leusink FK, Roepman P, de Jong RJB, Speel EJ, van den Brekel MW, et al. Validation of a gene expression signature for assessment of lymph node metastasis in oral squamous cell carcinoma. *J Clin Oncol.* (2012) 30:4104–10. doi: 10.1200/JCO.2011.40.4509
- O'Donnell RN, Kupferman M, Wei SJ, Singhal S, Weber R, O'Malley B, et al. Gene expression signature predicts lymphatic metastasis in

FUNDING

This study was financially supported by the National Science Council, Ministry of Science and Technology of Taiwan, and CGH in Taipei, Taiwan (NSC 102-2314-B-281-004, MOST 104-2314-B-281-005-MY2, 105CGH-TMU-06, and 107CGH-TMU-04).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2019.01508/full#supplementary-material>

- squamous cell carcinoma of the oral cavity. *Oncogene*. (2005) 24:1244–51. doi: 10.1038/sj.onc.1208285
27. Zevallos JP, Mazul AL, Walter V, Hayes DN. Gene expression subtype predicts nodal metastasis and survival in human papillomavirus-negative head and neck cancer. *Laryngoscope*. (2019) 129:154–61. doi: 10.1002/lary.27340
 28. Edge SB, Compton CC. The American joint committee on cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol*. (2010) 17:1471–4. doi: 10.1245/s10434-010-0985-4
 29. Creasman W. Revised FIGO staging for carcinoma of the endometrium. *Int J Gynaecol Obstet*. (2009) 105:109. doi: 10.1016/j.ijgo.2009.02.010
 30. Altman DG, McShane LM, Sauerbrei W, Taube SE. Reporting recommendations for tumor marker prognostic studies (REMARK): explanation and elaboration. *PLoS Med*. (2012) 9:e1001216. doi: 10.1371/journal.pmed.1001216
 31. Kukurba KR, Montgomery SB. RNA sequencing and analysis. *Cold Spring Harb Protoc*. (2015) 2015:951–69. doi: 10.1101/pdb.top084970
 32. Chari R, Lonergan KM, Pikor LA, Coe BP, Zhu CQ, Chan TH, et al. A sequence-based approach to identify reference genes for gene expression analysis. *BMC Med Genom*. (2010) 3:32. doi: 10.1186/1755-8794-3-32
 33. Andersen CL, Jensen JL, Orntoft TF. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res*. (2004) 64:5245–50. doi: 10.1158/0008-5472.CAN-04-0496
 34. Jorge S, Hou JY, Tergas AI, Burke WM, Huang Y, Hu JC, et al. Magnitude of risk for nodal metastasis associated with lymphovascular space invasion for endometrial cancer. *Gynecol Oncol*. (2016) 140:387–93. doi: 10.1016/j.ygyno.2016.01.002
 35. Breiman L. Random forests. *Mach Learn*. (2001) 45:5–32. doi: 10.1023/A:1010933404324
 36. Huang CY, Chang MC, Huang WY, Huang CT, Tang YC, Huang HD, et al. Urokinase-type plasminogen activator resulting from endometrial carcinogenesis enhances tumor invasion and correlates with poor outcome of endometrial carcinoma patients. *Sci Rep*. (2015) 5:10680. doi: 10.1038/srep10680
 37. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol*. (2010) 11:R106. doi: 10.1186/gb-2010-11-10-r106
 38. Boronow RC, Morrow CP, Creasman WT, Disaia PJ, Silverberg SG, Miller A, et al. Surgical staging in endometrial cancer: clinical-pathologic findings of a prospective study. *Obstet Gynecol*. (1984) 63:825–32.
 39. Creasman WT, Morrow CP, Bundy BN, Homesley HD, Graham JE, Heller PB. Surgical pathologic spread patterns of endometrial cancer. A gynecologic oncology group study. *Cancer*. (1987) 60(8 Suppl):2035–41. doi: 10.1002/1097-0142(19901015)60:8+<2035::aid-cnrcr2820601515>3.0.co;2-8
 40. Schink JC, Lurain JR, Wallemark CB, Chmiel JS. Tumor size in endometrial cancer: a prognostic factor for lymph node metastasis. *Obstet Gynecol*. (1987) 70:216–9. doi: 10.1016/0090-8258(87)90055-2
 41. Abal M, Llaurodo M, Doll A, Monge M, Colas E, Gonzalez M, et al. Molecular determinants of invasion in endometrial cancer. *Clin Transl Oncol*. (2007) 9:272–7. doi: 10.1007/s12094-007-0054-z
 42. Zgheib NB, Marchion DC, Bush SH, Judson PL, Wenham RM, Apte SM, et al. Molecular determinants for lymph node metastasis in clinically early-stage endometrial cancer. *Oncol Lett*. (2016) 11:323–9. doi: 10.3892/ol.2015.3883
 43. McCluggage WG. My approach to the interpretation of endometrial biopsies and curettings. *J Clin Pathol*. (2006) 59:801–12. doi: 10.1136/jcp.2005.029702
 44. McKenney JK, Longacre TA. Low-grade endometrial adenocarcinoma: a diagnostic algorithm for distinguishing atypical endometrial hyperplasia and other benign (and malignant) mimics. *Adv Anat Pathol*. (2009) 16:1–22. doi: 10.1097/PAP.0b013e3181919e15
 45. Leitao MM, Kehoe S, Barakat RR, Alektiar K, Gattoc LP, Rabbitt C, et al. Comparison of D&C and office endometrial biopsy accuracy in patients with FIGO grade I endometrial adenocarcinoma. *Gynecol Oncol*. (2009) 113:105–8. doi: 10.1016/j.ygyno.2008.12.017
 46. Mariani A, Dowdy SC, Cliby WA, Haddock MG, Keeney GL, Lesnick TG, et al. Efficacy of systematic lymphadenectomy and adjuvant radiotherapy in node-positive endometrial cancer patients. *Gynecol Oncol*. (2006) 101:200–8. doi: 10.1016/j.ygyno.2006.01.032
 47. Becker FF, Broome JD. L-asparaginase: inhibition of early mitosis in regenerating rat liver. *Science*. (1967) 156:1602–3. doi: 10.1126/science.156.3782.1602
 48. Den Boer ML, van Slegtenhorst M, De Menezes RX, Cheok MH, Buijs-Gladdines JG, Peters ST, et al. A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genome-wide classification study. *Lancet Oncol*. (2009) 10:125–34. doi: 10.1016/S1470-2045(08)70339-5
 49. Rigouin C, Nguyen HA, Schalk AM, Lavie A. Discovery of human-like L-asparaginases with potential clinical use by directed evolution. *Sci Rep*. (2017) 7:10224. doi: 10.1038/s41598-017-10758-4
 50. Shrivastava A, Khan AA, Khurshid M, Kalam MA, Jain SK, Singhal PK. Recent developments in L-asparaginase discovery and its potential as anticancer agent. *Crit Rev Oncol Hematol*. (2016) 100:1–10. doi: 10.1016/j.critrevonc.2015.01.002
 51. Fonnes T, Trovik J, Edqvist PH, Fasmer KE, Marcickiewicz J, Tingulstad S, et al. Asparaginase-like protein 1 expression in curettage independently predicts lymph node metastasis in endometrial carcinoma: a multicentre study. *BJOG*. (2018) 125:1695–703. doi: 10.1111/1471-0528.15403
 52. Verma R, Su S, McCrann DJ, Green JM, Leu K, Young PR, et al. RHEX, a novel regulator of human erythroid progenitor cell expansion and erythroblast development. *J Exp Med*. (2014) 211:1715–22. doi: 10.1084/jem.20130624
 53. Simmen FA, Su Y, Xiao R, Zeng Z, Simmen RC. The Kruppel-like factor 9 (KLF9) network in HEC-1-A endometrial carcinoma cells suggests the carcinogenic potential of dys-regulated KLF9 expression. *Reprod Biol Endocrinol*. (2008) 6:41. doi: 10.1186/1477-7827-6-41
 54. Tassi RA, Bignotti E, Falchetti M, Calza S, Ravaggi A, Rossi E, et al. Mammaglobin B expression in human endometrial cancer. *Int J Gynecol Cancer*. (2008) 18:1090–6. doi: 10.1111/j.1525-1438.2007.01137.x
 55. Tassi RA, Bignotti E, Rossi E, Falchetti M, Donzelli C, Calza S, et al. Overexpression of mammaglobin B in epithelial ovarian carcinomas. *Gynecol Oncol*. (2007) 105:578–85. doi: 10.1016/j.ygyno.2007.01.043
 56. Jia Y, Yang Y, Brock MV, Zhan Q, Herman JG, Guo M. Epigenetic regulation of DACT2, a key component of the Wnt signalling pathway in human lung cancer. *J Pathol*. (2013) 230:194–204. doi: 10.1002/path.4073
 57. Zhang Y, Bao W, Wang K, Lu W, Wang H, Tong H, et al. SOX17 is a tumor suppressor in endometrial cancer. *Oncotarget*. (2016) 7:76036–46. doi: 10.18632/oncotarget.12582
 58. Li X, Shi S, Li FF, Cheng R, Han Y, Diao LW, et al. Characterization of soluble N-ethylmaleimide-sensitive factor attachment protein receptor gene STX18 variations for possible roles in congenital heart diseases. *Gene*. (2017) 598:79–83. doi: 10.1016/j.gene.2016.10.043
 59. Bassett T, Harpur B, Poon HY, Kuo KH, Lee CH. Effective stimulation of growth in MCF-7 human breast cancer cells by inhibition of syntaxin18 by external guide sequence and ribonuclease P. *Cancer Lett*. (2008) 272:167–75. doi: 10.1016/j.canlet.2008.07.014
 60. Tangjitgamol S, Anderson BO, See HT, Lertbutsayanukul C, Sirisabya N, Manchana T, et al. Management of endometrial cancer in Asia: consensus statement from the Asian oncology summit 2009. *Lancet Oncol*. (2009) 10:1119–27. doi: 10.1016/S1470-2045(09)70290-6
 61. Kost ER, Hall KL, Hines JF, Farley JH, Nycum LR, Rose GS, et al. Asian-Pacific Islander race independently predicts poor outcome in patients with endometrial cancer. *Gynecol Oncol*. (2003) 89:218–26. doi: 10.1016/S0090-8258(03)00050-7

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Huang, Liao, Chou, Shrestha, Yang, Chiew, Huang, Hong, Huang, Chang and Huang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.