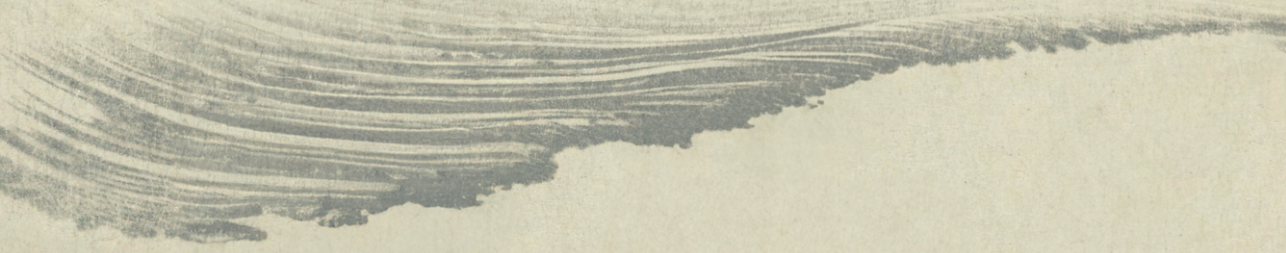
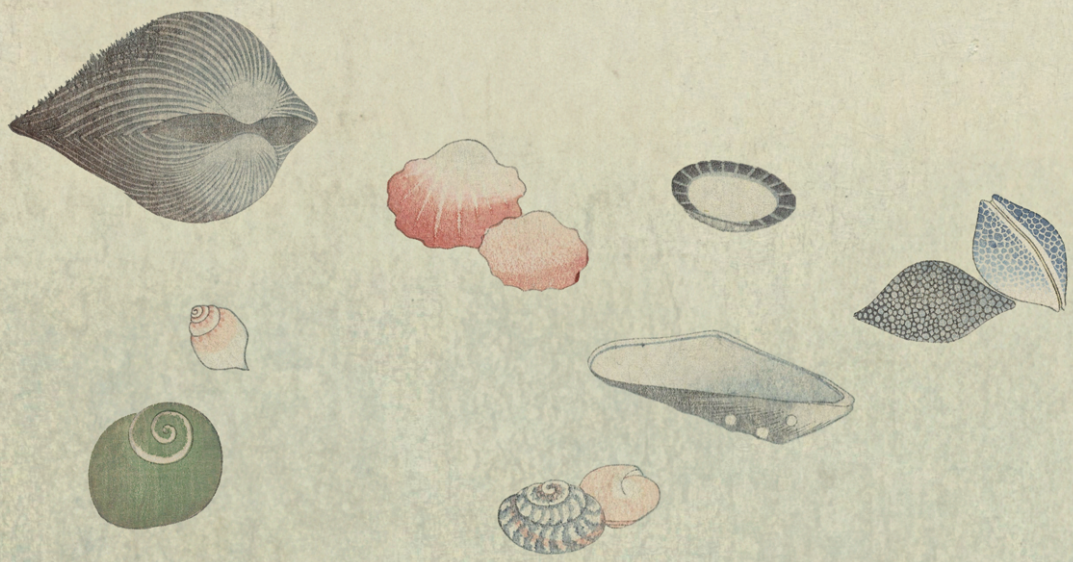
# Lymphovascular space invasion in endometrial carcinoma

Peters, E.E.M.





# CHAPTER 8

## GENE EXPRESSION ANALYSIS OF MISMATCH REPAIR DEFICIENT ENDOMETRIAL CANCERS WITH LYMPHOVASCULAR SPACE INVASION

Elke E.M. Peters, Lisa Vermij, Jan Oosting, Dana A.M. Mustafa,
Remi Nout, Vincent T.H.B.M. Smit and Tjalling Bosse

*In preparation*

## Abstract

Background: the prevalence of lymphovascular space invasion (LVSI) in endometrial cancer (EC) is not evenly distributed among molecular subgroups and the molecular events leading to LVSI are largely undefined. The aim of this study was to find a vascular invasion profile (VIP) by comparing DNA expression of MMRd ECs with substantial LVSI and without LVSI.

Methods: first, differential DNA expression analysis was performed on TCGA data. Next, DNA expression of archived MMRd ECs with substantial LVSI was compared to MMRd ECs without LVSI using NanoString technologies. Lastly, literature was reviewed for VIPs among and beyond ECs (3) Results: TCGA analysis revealed 94 differentially expressed genes between ECs without LVSI and those with extensive LVSI. No differentially expressed genes were found in the NanoString analysis of MMRd ECs. Among 13 published VIPs containing 345 genes, 12 overlapping genes recurred in two or three VIPs.

Conclusions: although this study did not reveal a VIP for MMRd ECs, there may be indications LVSI is associated with differential gene expression. Further research using a wider panel and larger sample may be helpful.

## Introduction

Lymphovascular space invasion (LVSI) in endometrial cancer (EC) is associated with an increased risk of recurrent disease, lymph node (LN) metastases, distant metastases and reduced survival [1-3]. Especially substantial LVSI, rather than focal LVSI, has proven to be the strongest prognostic factor in early-stage EC [4, 5]. LVSI is more prevalent in high grade EC and in higher stages of disease, compared to low grade and early-stage EC [6, 7]. While the clinical relevance of LVSI has been studied well, the underlying molecular biology leading to LVSI and potential markers to predict LVSI are largely undefined.

The TCGA molecular classification of EC describes four molecular subgroups, including 1) the polymerase epsilon (*POLE*) mutated (*POLE*-mt) tumors, characterized by an ultra-high mutated phenotype; 2) mismatch repair deficient (MMRd), characterized by a hypermutated phenotype; 3) *p53-abnormal* (p53abn), characterized by high copy number alterations and 4) tumors without a specific molecular profile (NSMP), lacking any of these characteristics [46]. The incidence of LVSI is highest in p53abn tumors, followed by MMRd tumors, but has a low incidence in *POLE* mt and NSMP tumors. In a pooled histopathological characterization of molecular groups, LVSI was present in 48,8%; 41,3%; 32,7% and 13,8% for p53abn, MMRd, *POLE*-mt and NSMP EC respectively [154]. In a combined early stage, high-intermediate risk EC study population, substantial LVSI was present in 5,4%; 8,9%, 0% and 2,4% in p53abn, MMRd, *POLE*-mt and NSMP EC respectively [10]. Potentially, these differences in the incidence of LVSI between the molecular subgroups may suggest molecular mechanisms resulting in LVSI.

Few studies have described an EC-specific vascular invasion profile (VIP) for LVSI based on differential DNA expression analyses. Watanabe et al. designed a 55-gene profile to predict LVSI in EC [11]. In addition, Mannelqvist et al. found an 18-gene profile for vascular invasion in EC [12]. Although both gene panels were designed for EC, no overlapping genes are listed. A 55-gene expression profile to predict lymph node metastases in EC, which is strongly associated with vascular invasion, included one overlapping gene (*CLCN2*) with one of the previous panels [13]. Differential expression analyses seeking vascular invasion profiles in other tumor types like breast [14-17], ovarian [18], hepatocellular [19-21], non-small cell lung carcinoma (NSCLC) [22] as well as gastric adenocarcinoma [23, 24] have all resulted in signatures but show little overlap between, as well as within tumor types.

Previous studies exploring a vascular invasion expression profile for EC have not incorporated the molecular EC classification. Given differences in LVSI incidence among the four molecular subgroups, we hypothesized there are molecular alterations associated with LVSI which may not be similar between molecular subgroups. Since LVSI is associated with MMRd, gene expression analysis was performed on archived MMRd EC, stratified by LVSI to evaluate altered gene expression associated with LVSI.

## Materials and Methods

*TCGA*

Publicly available whole genome expression data from the TCGA Uterine Corpus Endometrial Carcinomas (UCEC) cohort was used to explore differentially expressed genes (DEG) in LVSI positive EC [25]. LVSI status was extracted from the accompanying pathology report. LVSI was reported by local pathologists as: 'no', 'yes', 'focal', 'extensive' or 'indeterminate'. DEG analysis using the genes in the MsigDB hallmark gene sets [170] was carried out on the following subsets: all EC without LVSI versus EC with extensive LVSI; all EC without LVSI versus EC with any extent of LVSI; MMRd EC without LVSI versus MMRd EC with any extent of LVSI.

*Patient and tissue selection*

For this study 48 patients were selected from molecularly classified archived EC that were previously collected by our research group, including the international PORTEC-1/2/4a clinical trials [27-29]. Pathology review was performed to asses histological subtype, stage, grade and extent of LVSI. For the purpose of this study, all cases with documented (substantial) LVSI were re-reviewed by the study pathologists (E.E.M.P. and T.B.) to confirm LVSI status. Substantial LVSI was defined by the presence of at least four LVSI positive vessels in at least one hematoxylin and eosin (H&E) slide [30]. MMRd EC with substantial LVSI and MMRd EC without LVSI, matched for International Federation of Gynecology and Obstetrics (FIGO) stage and grade were selected for gene expression analysis. MMRd status was confirmed by immunohistochemistry [31].

*RNA isolation*

RNA was isolated from formalin-fixed paraffin-embedded (FFPE) tissue blocks using the automated Tissue Preparation System (Siemens Healthcare Diagnostics). RNA was extracted by taking four 0.8mm cores from the invasive border of the tumor, annotated by the study pathologists. The quality and quantity of the RNA was measured using the 2100 Bioanalyser (Agilent, Santa Clara, California USA). Samples were considered of sufficient quality for further analysis when >20% of RNA fragment length was at least 300 bases, and the corrected concentration remained >42 µg/µL.

*Panel selection for NanoString*

In order to select the best performing gene panel, we applied three available NanoString panels (PanCancer Progression Panel, PanCancer Pathways Panel (Human) and PanCancer IO 360 Panel) to TCGA cases and gene expression was performed, comparing tumors without LVSI to tumors with extensive LVSI. The panels contained 770 key genes of pathways involved in angiogenesis, extracellular matrix, epithelial-mesenchymal transition, signaling, and/or immune response and 30 to 40 housekeeping genes. Selection of the panel was based on the highest number of differentially expressed genes. The available nCounter gene panels allowed customization by adding up to 30 genes of choice. We choose to add the 15 most up-regulated and 15 most down-regulated genes from the TCGA comparison.

*NanoString gene expression analysis*

A total of 300 ng RNA was hybridized to the PanCancer IO 360 Panel for 17 hours, following the manufacturer's instruction (NanoString Technologies, Seattle, Washington, USA). The nCounter FLEX stsytem (Nanostring Technologies) was used to scan 490 Field of Views and count the number of copies per gene in each sample. Gene expression data was analyzed using the Advanced Analysis Module in the nSolver Analysis Software version 4.0 (NanoString Technologies, Washington, USA). The nSolver Analysis Software enables for quality control, normalization, cluster analysis and DEG analysis. Raw data was normalized by subtracting the mean plus one standard deviation of eight negative controls and technical variation was normalized through internal positive controls. Data was corrected for input volume via internal housekeeping genes.

*Statistical analyses*

Gene expression analyses on TCGA data were done using R version 4.0.2. MsigDB datasets were retrieved using the EGSEAdata package version 1.16 [173], and the genes from the NanoString panels were retrieved from the manufacturer documentation [33]. In order to limit the analysis to biologically relevant genes for the TCGA analysis, only genes belonging to the hallmark gene sets (category = "H") were selected. TCGA data of the selected gene sets and NanoString panels was retrieved using the cgdsr package version 1.3.0 [34]. Genes were filtered out when less than 50% of the samples had a valid data point. The statistical analysis was performed using voom from the limma package version 3.44.3 [35]. Genes were considered differentially expressed when the adjusted p-value according to Benjamini and Hochberg was below 0.05. Expression data was incomplete for one LVSI negative tumor and therefore excluded from further analyses.

*Literature review for vascular invasion profiles*

Literature on vascular invasion profiles in epithelial cancers was searched using the search terms listed in box 1.

**Box 1. Entries used to search PubMed for vascular invasion profiles in cancer.**

("Transcriptome"[Mesh] OR "gene expression profile"[all fields] OR "gene expression profiles"[all fields] OR "Transcriptome profile"[all fields] OR "Transcriptome profiles" OR "Gene expression signature"[all fields] OR "gene expression signatures"[all fields] OR "gene expression signature"[all fields] OR "gene expression signatures"[all fields]) AND ("lymphovascular space invasion"[all fields] OR "lymphvascular space invasion"[all fields] OR "lymphovascular space involvement"[all fields] OR "lymphvascular space involvement"[all fields] OR "vascular space invasion"[all fields] OR "vascular space involvement"[all fields] OR "vascular invasion"[all fields] OR "vascular involvement"[all fields] OR "LVI"[all fields] OR "LVSI"[all fields])

## Results

*TCGA*

Gene expression data was available for 176/529 (33,3%) EC in the TCGA dataset. According to the accompanying pathology reports, 88/176 (50,0%) tumors had no LVSI; 78/176 (44,3%) were positive for LVSI and LVSI status was missing in 10/176 (5,6%) tumors. The extent of LVSI was specified in 28/78 (35,9%) of LVSI positive tumors and reported as extensive (n=19/78; 24,4% ),

focal (n=8/78; 10,3%) and intermediate (n=1/78; 1,3%). Extensive LVSI was most prevalent in p53abn EC (n=14/19; 73,7%) as shown in table 1. Molecular subtype was unknown in 9 (9/176; 5,1%).

**Table 1. LVSI status or extent according to molecular classifier for TCGA cases with gene expression data.**

| Molecular classifier | No LVSI | LVSI positive | Focal LVSI | Extensive LVSI | LVSI missing |
|---|---|---|---|---|---|
| *POLE* mutated | 8 (62%) | 3 (23%) | 0 (0%) | 2 (15%) | 2 |
| MMR deficient | 19 (48%) | 15 (38%) | 3 (7.5%) | 3 (7.5%) | 2 |
| *TP53* mutated | 38 (49%) | 24 (31%) | 2 (2,6%) | 14 (18%) | 4 |
| NSMP | 19 (70%) | 5 (19%) | 3 (11%) | 0 (0%) | 1 |

* Information about LVSI status of TCGA cases with gene expression data (n = 176) was retrieved from pathology reports. If LVSI was reported to be present, the extent of LVSI (focal or extensive) as stated in the pathology report was adopted. The extent of 'LVSI positive' cases is unknown. LVSI status in TCGA is not verified by pathology review, nor definitions are given of the extent of LVSI. LVSI status was reported 'indeterminate' in one NSMP case and left out from this table. Nine cases were not assigned to a molecular classifier, including four without LVSI, three were LVSI positive, LVSI was 'indeterminate' in one and LVSI status was missing in another one. Abbreviations: *POLE*: polymerase ε; MMR: mismatch repair, NSMP: no specific molecular profile.

Analysis of differential expression of genes between 19 tumors with extensive LVSI and 87 without LVSI revealed 94 DEGs (table 2, supplementary table 1 for details). In one of 88 LVSI negative tumors DNA expression data was incomplete and excluded from analysis. No DEGs were found in comparison between LVSI positive (any extent) versus no LVSI regardless of molecular group; any LVSI versus no LVSI among MMRd (table 3).

**Table 2. Differentially expressed genes in alphabetical order**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *ABCD1* | *CD207* | *DES* | *GLS* | *MAST4* | *NID2* | *PSMD3* | *SLC7A2* | *TOMM40* |
| *AMOT* | *CDC25B* | *DHRS2* | *HADH* | *MPPED2* | *NOP56* | *PVR* | *SLN* | *TRIB3* |
| *AMPH* | *CDC42* | *DUSP5* | *HSPB2* | *MSX1* | *NPHP1* | *RBP2* | *SMAD3* | *TUBA4A* |
| *ATP1A3* | *CDH16* | *EHD1* | *ICA1* | *MT1E* | *OASL* | *PHOF* | *SNX10* | *UPP1* |
| *ATP2B4* | *CLCF1* | *EI24* | *IFI30* | *MX2* | *OSTC* | *SCHIP1* | *SORBS2* | *WFS1* |
| *AURKA* | *CLCN2* | *EPHX1* | *IL6* | *MYL4* | *OVOL2* | *SGCB* | *SPHK1* | |
| *BCL2L1* | *CLDN19* | *ESR1* | *IRF6* | *NDP* | *PCDH7* | *SKP1* | *SSH2* | |
| *CANT1* | *CYB5A* | *FABP3* | *KCNQ2* | *NDRG2* | *PDCD4* | *SLC1A1* | *SST* | |
| *CANX* | *CYB5R1* | *FGFR1* | *KDM4B* | *NEFH* | *PERP* | *SLC27A1* | *TEAD4* | |
| *CASP10* | *CYP2C18* | *FUT1* | *L1CAM* | *NEXN* | *PEX2* | *SLC3A2* | *THY1* | |
| *CASP6* | *DDC* | *GAL* | *LRIG1* | *NFKB2* | *PLA2G12A* | *SLC6A12* | *TMED10* | |

*Panel selection for NanoString analysis*

Selection of the NanoString gene panel was done in a two-way comparison. First, the genes listed in the PanCancer IO360, PanCancer Progression and PanCancer Pathways panels were

**Table 3. Differential gene expression analysis in several comparisons.**

| Case selection | n | Reference | n | Number of DEG |
|---|---|---|---|---|
| Extensive LVSI* | 19 | No LVSI* | 87 | 94 |
| Extensive LVSI or LVSI positive* | 68 | No LVSI* | 87 | 0 |
| Any LVSI (MMRd) | 21 | No LVSI (MMRd) | 19 | 0 |

*Any molecular classifier. Abbreviations: DEG: differentially expressed genes; MMRd: mismatch repair deficient.

compared to the list of 94 DEGs in extensive LVSI in the TCGA for overlapping genes and revealed 12, 10 and 9 overlapping genes, respectively. Additionally, a gene expression analysis

using the genes listed in these panels was performed for the TCGA selection of 19 tumors with extensive LVSI versus 87 without LVSI and resulted in 12, 44 and 31 DEGs respectively (table 4). Based on these results, the PanCancer Progression panel was selected for further analysis and enriched with the top 30 DEGs from the 94 DEGs from the TCGA analysis (table 5).

**Table 4. Performance of different NanoString gene panels**

| NanoString panel | DEG in TCGA dataset[1] | Overlapping genes[2] |
|---|---|---|
| PanCancer IO 360 | 12 | 12 |
| PanCancer Progression | 44 | 10 |
| PanCancer Pathways | 31 | 9 |

[1] Differential expression analysis of TCGA data with the genes in the panels listed in the comparison of "no LVSI" versus "extensive LVSI". [2] Number of genes listed in the NanoString panel which were also among the 94 differentially expressed genes in the initial TCGA analysis ("no LVSI" versus "extensive LVSI")

**Table 5. List of genes customized to the NanoString PanCancer progression panel.**

| | | | | | |
|---|---|---|---|---|---|
| *AMOT* | *CLDN19* | *ESR1* | *LRIG1* | *NDRG2* | *SLC6A12* |
| *AMPH* | *CYP2C18* | *FABP3* | *MPPED2* | *NEFH* | *SLC7A2* |
| *ATP1A3* | *DDC* | *GAL* | *MSX1* | *PCDH7* | *SLN* |
| *CD207* | *DES* | *KCNQ2* | *MT1E* | *RBP1* | *SORBS2* |
| *CDH16* | *DHRS2* | *L1CAM* | *MYL4* | *SLC1A1* | *SST* |

*NanoString gene expression analysis*

Sufficient RNA for analyses was extracted from 32 MMRd EC without LVSI and 16 MMRd EC with extensive LVSI. None of the pathways or gene clusters involved in angiogenesis, extracellular matrix, epithelial-mesenchymal transition, signaling or immune response was associated with LVSI status. Analysis of differential expression of genes revealed no significant DEGs. In particular, the 30 genes customized to the panel which were differentially expressed in the TCGA analysis, showed no significant differential expression.

*Literature review for vascular invasion profiles (VIPs)*

Using the entries listed in Box 1, a PubMed search was performed which resulted in 13 articles with VIPs on eight epithelial tumor types, including endometrial, lung, breast, hepatocellular, serous ovarian, colon, urothelial cell and esophageal squamous cell carcinomas. A summary of these studies is listed in Table 6. DEGs per study are listed in supplementary table 2. There were three VIPs for EC, these included 17, 12 and 55 genes respectively with one overlapping gene (*CLCN2*) in two of the three VIPs. In the comparison of the three VIPs for EC with our TCGA analysis, three overlapping genes were found: *CLCN2*, *AURKA* and *NOP56*. In the comparison of the 13 published VIPs, 8 genes appeared in two VIPs (*AURKA, ATP2B4, CCNB2, HMGA1, MED1, MT1E, NDP* and *TM7SF2*) and 4 genes were listed in 3 VIPs *CLCN2, ESR1, NOP56* and *UBE2C*). 7 of these were also differentially expressed in our TCGA analysis (*AURKA, ATP2B4, CLCN2, ESR1, MT1E, NDP* and *NOP56,* figure 1). Overlap was found in VIPs for EC (*AURKA, CLCN2, ESR1* and *NOP56*) breast cancer (*MT1E, NDP* and *NOP56*), NSCLC (*ATP2B4*) and colon cancer (*ESR1*).

Table 6. Overview of reported vascular invasion profiles.

| Author | Tumor | Input[1] | Cases (n) | DEG[2] |
|---|---|---|---|---|
| Mannelqvist[12] | Endometrial | FF | 57 | 17 |
| Kang[13] | Endometrial | FF | 330 | 12 |
| Watanabe[11] | Endometrial | FF | 88 | 55 |
| Regan[22] | NSCLC | Cell lines | 9 | 17 |
| Fidalgo[17] | Breast | FF | 57 | 22 |
| Kurozumi[14] | Breast | FF, data set | 1565 + 854 | 99 |
| Asaoka[15] | Breast | Data set | 835 | 3 |
| Minguez[19] | Hepatocellular | FF | 214 | 35 |
| Ho[21] | Hepatocellular | FF | 53 | 14 |
| Yue[18] | Serous ovarian | Data set | 192 | 43 |
| Jiang[36] | Colon | FF | 47 | 20 |
| Poyet[37] | Urothelial | FF | 23 | 3 |
| Sonohara[38] | Esophageal squamous | FF, data set | 267 + 96 | 5 |

[1] FF: fresh frozen tissue. [2] DEG: differentially expressed genes; number of genes in the vascular invasion profile.

## Discussion

This study aimed to identify differentially expressed genes associated with LVSI in EC. The initial analysis of TCGA data confirmed LVSI is more frequent in MMRd and p53abn tumors compared to NSMP and POLE-mutated tumors. Extensive LVSI was associated with differential expression of 94 genes. In the subsequent gene expression analysis of 48 MMRd EC on the NanoString platform we were not able to establish a gene expression profile associated with substantial LVSI. Literature review of differential expression of genes associated with vascular invasion among several tumor types resulted in a longlist of genes with limited overlap between and within tumor types.

The motivation for this study was the hypothesis that LVSI is associated with specific molecular differences between tumors which is manifested by altered gene expression. The analysis of TCGA data yielded 94 DEG supporting this hypothesis. The candidate genes are ideally confirmed in an independent study set. Due to the coordination of the various PORTEC RCTs, our research group has access to a large amount of data and FFPE EC from which this study set could be compiled.

NanoStrings direct hybridization technique without amplification work up, enables to perform expression analysis on FFPE tissue, instead of fresh frozen tissue only. The downside for choosing this platform is that expression analysis is limited to genes included in a fixed panel with limited options for customization. Despite of a well-considered selection and customization of the panel, no DEGs were found. The NanoString panel we choose was directed towards genes involved in angiogenic, tumor progression and metastatic processes. However, genes included in other VIPs are only involved in these processes to a limited extent.

Eight of the 94 differentially expressed genes associated with substantial LVSI in TCGA were also listed in other VIPs. These genes are involved in divergent cellular processes, like homeostatic processes involving *ATP2B4* (intracellular calcium homeostasis), *CLCN2* (voltage gated chloride channel) and *MTE1* (intracellular Acyl-CoA regulation), and not directly associated with vascular invasion. Other genes like *ESR1* encode for the estrogen receptor and like *MED1* contributes to transcription. *AURKA* contributes to the regulation of cell cycle progression and *NDP* activates the Wnt signaling pathway, associated with carcinogenic processes, but not directly linked to LVSI. Direct association with LVSI is even less evident for *NOP56,* which plays a role in pre-ribosomal RNA processing.

There are several explanations for the lack of concordance between our results and previous VIPs for EC. First, the latter were designed without knowledge of the molecular subtype. Second, stratification for LVSI was done using a binary system for LVSI (present or absent) without quantification [11, 12], presumably including focal LVSI we know have no clinical impact [4]. Third, the profile by Kang et al. was designed to predict nodal metastases which is strongly associated with LVSI, but is not the same [13]. Last, but not least, different methods were used.

Multiple studies have shown substantial LVSI is a relevant prognostic factor, in contrast to focal LVSI [4, 39-41]. Differences in prognosis and prevalence of LVSI are observed among the four molecular subtypes of EC and this study was designed to find a VIP within one molecular subtype. There were two reasons to limit this study to MMRd EC. First, we expected the mutations resulting from the MMRd nature, to appear randomly which would not impede the detection of specific, LVSI associated altered gene expression. Second, MMRd ECs have a relatively good prognosis, but substantial LVSI is the strongest prognostic factor for adverse prognosis for EC [4]. However, it is unknown how substantial LVSI affects the prognosis of patients with MMRd EC, a VIP for MMRd EC could be a predictive factor to identify patients with a high risk for recurrence among a group of patients with a relatively good prognosis. Third, by limiting the study to one molecular subtype, possible bias relating to the molecular subgroup was excluded.



Despite the well-considered study design, sample analysis revealed no VIP. This was in line with our TCGA subanalysis of MMRd ECs which did not result in DEGs associated with LVSI. A possible explanation is the sample size was inadequate to detect altered gene expression related to LVSI among hypermutated tumors in which genetic heterogeneity possibly overturns the specific molecular alterations associated with LVSI. Another possibility is the lack of relevant genes in the panels used.

There are some limitations to this study. First, TCGA analysis was used as input for panel selection and customization, however the majority of tumors with extensive LVSI in TCGA had a p53abn phenotype in contrast to the MMRd study set. TCGA analysis among MMRd only resulted

in a small sample size without DEGs and therefore could not be used for panel selection and customization. Second, although substantial LVSI occurs more often in MMRd and p53abn tumors [42], it is still a rather uncommon event, limiting the number of cases we could select for this study. Third, whole exome sequencing (WES) is the golden standard for gene expression studies. The archived tumor material was FFPE only and however the quality of extracted RNA from FFPE tissue was adequate for the NanoString platform, the quality was insufficient for WES.

## Conclusions

In this study we were not able to establish a VIP for MMRd EC. Nevertheless, based on differences in the prevalence of LVSI between the molecular subgroups in EC, TCGA data analysis and literature search, there are sufficient indications there are molecular alterations underlying LVSI. The molecular mechanisms resulting in LVSI and the key genes involved are, however, still largely unknown. As a result, the search for a VIP is still unfocused, but advancing insights into tumor biology and adequate sample size will increase the likelihood of finding a VIP in future studies.

## References


1. Gadducci, A.; Cavazzana, A.; Cosio, S.; et al. *Lymph-vascular space involvement and outer one-third myometrial invasion are strong predictors of distant haematogeneous failures in patients with stage I-II endometrioid-type endometrial cancer.* Anticancer Res 2009, **29**, 1715-1720.
2. Guntupalli, S.R.; Zighelboim, I.; Kizer, N.T.; et al. *Lymphovascular space invasion is an independent risk factor for nodal disease and poor outcomes in endometrioid endometrial cancer.* Gynecologic oncology 2012, **124**, 31-35,
3. O'Brien, D.J.; Flannelly, G.; Mooney, E.E.; et al. *Lymphovascular space involvement in early stage well-differentiated endometrial cancer is associated with increased mortality.* BJOG 2009, **116**, 991-994.
4. Bosse, T.; Peters, E.E.; Creutzberg, C.L.; et al. *Substantial lymph-vascular space invasion (LVSI) is a significant risk factor for recurrence in endometrial cancer--A pooled analysis of PORTEC 1 and 2 trials.* Eur J Cancer 2015, **51**, 1742-1750.
5. Peters, E.E.M.; Leon-Castillo, A.; Hogdall, E.; et al. *Substantial Lymphovascular Space Invasion Is an Adverse Prognostic Factor in High-Risk Endometrial Cancer.* Int J Gynecol Pathol 2022, **41**, 227-234.
6. Boothe, D.; Wolfson, A.; Christensen, M.; et al. *Lymphovascular Invasion in Endometrial Cancer: Prognostic Value and Implications on Adjuvant Radiation Therapy Use.* Am J Clin Oncol 2019, **42**, 549-554.
7. Matsuo, K.; Garcia-Sayre, J.; Medeiros, F.; et al. *Impact of depth and extent of lymphovascular space invasion on lymph node metastasis and recurrence patterns in endometrial cancer.* J Surg Oncol 2015, **112**, 669-676.
8. Cancer Genome Atlas Research, N.; Kandoth, C.; Schultz, N.; Cherniack, A.D.; et al. *Integrated genomic characterization of endometrial carcinoma.* Nature 2013, **497**, 67-73.
9. Raffone, A.; Travaglino, A.; Mascolo, M.; et al. *Histopathological characterization of ProMisE molecular groups of endometrial cancer.* Gynecologic oncology 2020, **157**, 252-259.
10. Stelloo, E.; Nout, R.A.; Osse, E.M.; et al. *Improved Risk Assessment by Integrating Molecular and Clinicopathological Factors in Early-stage Endometrial Cancer-Combined Analysis of the PORTEC Cohorts.* Clin Cancer Res 2016, **22**, 4215-4224.
11. Watanabe, T.; Honma, R.; Kojima, et al. *Prediction of lymphovascular space invasion in endometrial cancer using the 55-gene signature selected by DNA microarray analysis.* PLoS One 2019, **14**, e0223178.
12. Mannelqvist, M.; Stefansson, I.M.; Bredholt,et al. *Gene expression patterns related to vascular invasion and aggressive features in endometrial cancer.* Am J Pathol 2011, **178**, 861-871.
13. Kang, S.; Thompson, Z.; McClung,et al. *Gene Expression Signature-Based Prediction of Lymph Node Metastasis in Patients With Endometrioid Endometrial Cancer*. Int J Gynecol Cancer 2018, **28**, 260-266.
14. Kurozumi, S.; Joseph, C.; Sonbul, S.; et al. *A key genomic subtype associated with lymphovascular invasion in invasive breast cancer.* Br J Cancer 2019, **120**, 1129-1136.
15. Asaoka, M.; Patnaik, S.K.; Zhang, F.; et al. *Lymphovascular invasion in breast cancer is associated with gene expression signatures of cell proliferation but not lymphangiogenesis or immune response.* Breast Cancer Res Treat 2020, **181**, 309-322.
16. Mannelqvist, M.; Wik, E.; Stefansson, I.M.;et al. *An 18-gene signature for vascular invasion is associated with aggressive features and reduced survival in breast cancer*. PLoS One 2014, **9**, e98787.
17. Fidalgo, F.; Rodrigues, T.C.; Pinilla, et al. *Lymphovascular invasion and histologic grade are associated with specific genomic profiles in invasive carcinomas of the breast.* Tumour Biol 2015, **36**, 1835-1848.
18. Yue, H.; Wang, J.; Chen, R.; et al. *Gene signature characteristic of elevated stromal infiltration and activation is associated with increased risk of hematogenous and lymphatic metastasis in serous ovarian cancer*. BMC Cancer 2019, **19**, 1266.

19. Minguez, B.; Hoshida, Y.; Villanueva, A.; et al. *Gene-expression signature of vascular invasion in hepatocellular carcinoma.* J Hepatol 2011, **55**, 1325-1331.
20. Hou, Y.; Zou, Q.; Ge, R.; et al. *The critical role of CD133(+)CD44(+/high) tumor cells in hematogenous metastasis of liver cancers.* Cell Res 2012, **22**, 259-272.
21. Ho, M.C.; Lin, J.J.; Chen, C.N.; et al. *A gene expression profile for vascular invasion can predict the recurrence after resection of hepatocellular carcinoma: a microarray approach.* Ann Surg Oncol 2006, **13**, 1474-1484.
22. Regan, E.; Sibley, R.C.; Cenik, et al. *Identification of Gene Expression Differences between Lymphangiogenic and Non-Lymphangiogenic Non-Small Cell Lung Cancer Cell Lines.* PLoS One 2016, **11**, e0150963.
23. Dicken, B.J.; Graham, K.; Hamilton, S.M.; et al. *Lymphovascular invasion is associated with poor survival in gastric cancer: an application of gene-expression and tissue array techniques.* Ann Surg 2006, **243**, 64-73.
24. Korenberg, M.J.; Dicken, B.J.; Damaraju, S.; et al. *Predicting node positivity in gastric cancer from gene expression profiles*. Biotechnol Lett 2009, **31**, 1381-1388.
25. TCGA. Consortium Uterine Corpus Endometrial Carcinoma (TCGA, PanCancer Atlas). 2018.
26. Liberzon, A.; Birger, C.; Thorvaldsdottir, H.; et al. *The Molecular Signatures Database (MSigDB) hallmark gene set collection.* Cell Syst 2015, **1**, 417-425.
27. Creutzberg, C.L.; van Putten, W.L.; Koper, P.C.; et al. *Surgery and postoperative radiotherapy versus surgery alone for patients with stage-1 endometrial carcinoma: multicentre randomised trial. PORTEC Study Group. Post Operative Radiation Therapy in Endometrial Carcinoma.* Lancet 2000, **355**, 1404-1411
28. Nout, R.A.; Smit, V.T.; Putter, H.;et al. *Vaginal brachytherapy versus pelvic external beam radiotherapy for patients with endometrial cancer of high-intermediate risk (PORTEC-2): an open-label, non-inferiority, randomised trial.* Lancet 2010, **375**, 816-823
29. van den Heerik, A.; Horeweg, N.; Nout, R.A.; et al. *PORTEC-4a: international randomized trial of molecular profile-based adjuvant treatment for women with high-intermediate risk endometrial cancer.* Int J Gynecol Cancer 2020, **30**, 2002-2007.
30. Peters, E.E.M.; Leon-Castillo, A.; Smit, V.; et al. *Defining Substantial Lymphovascular Space Invasion in Endometrial Cancer.* Int J Gynecol Pathol 2022, **41**, 220-226.
31. Stelloo, E.; Jansen, A.M.L.; Osse, E.M.; et al. *Practical guidance for mismatch repair-deficiency testing in endometrial cancer.* Ann Oncol 2017, **28**, 96-102.
32. Alhamdoosh M, H.Y., Smyth GK. EGSEAdata: Gene set collections for the EGSEA package. 2021, **10**.18129/B9.bioc.EGSEAdata.
33. NanoString. nCounter Panels & Assys | Oncology. Available online: https://nanostring.com/products/ncounter-assays-panels/oncology/
34. Jacobsen A, L.A. cgdsr: R-Based API for Accessing the MSKCC Cancer Genomics Data Server (CGDS). Available online: https://github.com/cBioPortal/cgdsr
35. Law, C.W.; Chen, Y.; Shi, W.; et al: *Precision weights unlock linear model analysis tools for RNA-seq read counts.* Genome Biol 2014, **15**, R29.
36. Jiang, H.H.; Zhang, Z.Y.; Wang, X.Y.; Tang, X.; Liu, H.L.; Wang, A.L.; Li, H.G.; Tang, E.J.; Lin, M.B. Prognostic significance of lymphovascular invasion in colorectal cancer and its association with genomic alterations. World J Gastroenterol 2019, 25, 2489-2502, doi:10.3748/wjg.v25.i20.2489.
37. Poyet, C.; Thomas, L.; Benoit, T.M.; et al. *Implication of vascular endothelial growth factor A and C in revealing diagnostic lymphangiogenic markers in node-positive bladder cancer.* Oncotarget 2017, **8**, 21871-21883.
38. Sonohara, F.; Gao, F.; Iwata, N.; et al. *Genome-wide Discovery of a Novel Gene-expression Signature for the Identification of Lymph Node Metastasis in Esophageal Squamous Cell Carcinoma.* Ann Surg 2019, **269**, 879-886.

39. Tortorella, L.; Restaino, S.; Zannoni, G.F.; et al. *Substantial lymph-vascular space invasion (LVSI) as predictor of distant relapse and poor prognosis in low-risk early-stage endometrial cancer.* J Gynecol Oncol 2021, **32**, e11.
40. Barnes, E.A.; Martell, K.; Parra-Herran, C.; et al. *Substantial lymphovascular space invasion predicts worse outcomes in early-stage endometrioid endometrial cancer.* Brachytherapy 2021, **20**, 527-535.
41. Restaino, S.; Tortorella, L.; Dinoi, G.; et al. *Semiquantitative evaluation of lymph-vascular space invasion in patients affected by endometrial cancer: Prognostic and clinical implications.* Eur J Cancer 2021, **142**, 29-37.
42. Stelloo, E.; Nout, R.A.; Naves, L.C.; et al *High concordance of molecular tumor alterations between pre-operative curettage and hysterectomy specimens in patients with endometrial carcinoma.* Gynecologic oncology 2014, **133**, 197-204.

Table S1: Differential expression analysis results of TGCA DNA expression data in a comparison of tumors with extensive LVSI (n=19) versus tumors without LVSI (n=87)

| Gene | logFC | AveExpr | t | P value | adj P value | B |
|---|---|---|---|---|---|---|
| FABP3 | -2,0299 | 4,1149 | -0,5937 | 0,0000 | 0,0001 | 8,4812 |
| MYL4 | -2,0264 | -1,5864 | -0,5750 | 0,0000 | 0,0002 | 6,7377 |
| MT1E | -2,2278 | 5,1918 | -0,5330 | 0,0000 | 0,0006 | 5,9762 |
| SCHIP1 | -1,5680 | 4,3878 | -0,5343 | 0,0000 | 0,0006 | 6,0127 |
| CDC25B | -1,1029 | 7,5442 | -0,5080 | 0,0000 | 0,0013 | 4,9084 |
| CDH16 | -3,6105 | 1,6080 | -0,5001 | 0,0000 | 0,0015 | 4,5424 |
| NDP | 2,9624 | 2,6835 | 0,4824 | 0,0000 | 0,0027 | 3,5495 |
| BCL2L1 | -0,6438 | 8,4508 | -0,4754 | 0,0000 | 0,0032 | 3,6023 |
| IL6 | -2,3317 | 1,4848 | -0,4719 | 0,0000 | 0,0033 | 3,3896 |
| SLC6A12 | -2,2233 | 3,1225 | -0,4664 | 0,0000 | 0,0037 | 3,3489 |
| ESR1 | 2,6802 | 8,3879 | 0,4635 | 0,0000 | 0,0038 | 3,2320 |
| KCNQ2 | -3,4968 | -1,4394 | -0,4554 | 0,0000 | 0,0048 | 2,6266 |
| SNX10 | -1,3123 | 4,4351 | -0,4501 | 0,0000 | 0,0055 | 2,7796 |
| WFS1 | 1,0287 | 7,7131 | 0,4471 | 0,0000 | 0,0057 | 2,5860 |
| NOP56 | -0,6416 | 8,5424 | -0,4378 | 0,0000 | 0,0077 | 2,1839 |
| CANT1 | 0,6326 | 7,5680 | 0,4298 | 0,0000 | 0,0096 | 1,9460 |
| CLCF1 | -1,0251 | 5,1740 | -0,4289 | 0,0000 | 0,0096 | 2,0048 |
| KIFAP3 | 0,9217 | 6,3257 | 0,4266 | 0,0000 | 0,0099 | 1,9244 |
| PCDH7 | 2,5788 | 6,3626 | 0,4229 | 0,0000 | 0,0108 | 1,8177 |
| NEXN | -1,3893 | 2,9305 | -0,4210 | 0,0001 | 0,0110 | 1,7247 |
| SLC3A2 | -0,5464 | 8,8881 | -0,4191 | 0,0001 | 0,0113 | 1,5014 |
| CASP10 | -1,2263 | 4,6060 | -0,4108 | 0,0001 | 0,0120 | 1,3971 |
| CYB5R1 | 0,9276 | 7,1347 | 0,4119 | 0,0001 | 0,0120 | 1,3558 |
| KDM4B | 0,7460 | 7,3263 | 0,4126 | 0,0001 | 0,0120 | 1,3586 |
| LRIG1 | 1,6564 | 8,7247 | 0,4115 | 0,0001 | 0,0120 | 1,2747 |
| NDRG2 | 1,6820 | 6,8336 | 0,4124 | 0,0001 | 0,0120 | 1,4313 |
| SST | -3,4973 | 1,9205 | -0,4138 | 0,0001 | 0,0120 | 1,5022 |
| SGCB | 1,1724 | 6,7380 | 0,4081 | 0,0001 | 0,0129 | 1,2666 |
| NEFH | -2,5620 | 2,2822 | -0,4044 | 0,0001 | 0,0137 | 1,1867 |
| NID2 | -1,2089 | 4,1919 | -0,4037 | 0,0001 | 0,0137 | 1,1711 |
| SLN | -1,7418 | -1,9876 | -0,4043 | 0,0001 | 0,0137 | 0,9518 |
| TMED10 | 0,5535 | 9,8296 | 0,4025 | 0,0001 | 0,0138 | 0,9190 |
| RHOF | -1,4629 | 5,2830 | -0,4010 | 0,0001 | 0,0142 | 1,0137 |
| PSMD3 | -0,8079 | 8,6224 | -0,3857 | 0,0002 | 0,0215 | 0,3505 |
| PVR | -0,6556 | 7,1363 | -0,3869 | 0,0002 | 0,0215 | 0,4379 |
| SPHK1 | -1,0033 | 3,9545 | -0,3862 | 0,0002 | 0,0215 | 0,6065 |
| TOMM40 | -0,6261 | 7,8381 | -0,3862 | 0,0002 | 0,0215 | 0,3870 |
| TRIB3 | -1,1477 | 5,2422 | -0,3888 | 0,0002 | 0,0215 | 0,6185 |
| PERP | 0,9467 | 9,1857 | 0,3844 | 0,0002 | 0,0220 | 0,3158 |
| AURKA | -0,8133 | 5,9878 | -0,3826 | 0,0002 | 0,0227 | 0,3667 |
| NFKB2 | -0,6685 | 7,3977 | -0,3822 | 0,0002 | 0,0227 | 0,2665 |
| L1CAM | -2,5584 | 3,9183 | -0,3771 | 0,0003 | 0,0259 | 0,2816 |

Table S1. *Continued*

| Gene | logFC | AveExpr | t | P value | adj P value | B |
|---|---|---|---|---|---|---|
| PLA2G12A | 0,6663 | 6,5599 | 0,3776 | 0,0003 | 0,0259 | 0,2357 |
| AMOT | 1,7704 | 5,8448 | 0,3762 | 0,0003 | 0,0261 | 0,2871 |
| SSH2 | -0,7353 | 5,6436 | -0,3756 | 0,0003 | 0,0261 | 0,1755 |
| TEAD4 | -0,7139 | 5,5822 | -0,3747 | 0,0003 | 0,0263 | 0,1510 |
| MPPED2 | 2,8872 | 4,4531 | 0,3715 | 0,0003 | 0,0288 | 0,1109 |
| CANX | 0,6814 | 10,6966 | 0,3696 | 0,0003 | 0,0291 | -0,1886 |
| DHRS2 | -2,4084 | 0,7854 | -0,3703 | 0,0003 | 0,0291 | 0,0682 |
| RBP1 | 1,6826 | 8,3220 | 0,3695 | 0,0003 | 0,0291 | -0,1129 |
| NPHP1 | 1,1209 | 4,7290 | 0,3657 | 0,0004 | 0,0325 | -0,0257 |
| CYB5A | 0,8665 | 6,6845 | 0,3637 | 0,0004 | 0,0340 | -0,2086 |
| DDC | -2,0983 | -1,5586 | -0,3604 | 0,0005 | 0,0340 | -0,2889 |
| EI24 | 0,4955 | 8,3705 | 0,3603 | 0,0005 | 0,0340 | -0,4519 |
| IFI30 | -0,8229 | 9,0822 | -0,3612 | 0,0005 | 0,0340 | -0,4545 |
| MAST4 | 1,5399 | 3,6844 | 0,3601 | 0,0005 | 0,0340 | -0,2195 |
| MX2 | -1,2863 | 5,4121 | -0,3610 | 0,0005 | 0,0340 | -0,2993 |
| SLC1A1 | 1,6747 | 3,5042 | 0,3629 | 0,0004 | 0,0340 | -0,1484 |
| SLC7A2 | -1,8538 | 4,2930 | -0,3615 | 0,0005 | 0,0340 | -0,2084 |
| SMAD3 | -0,7800 | 6,2815 | -0,3597 | 0,0005 | 0,0340 | -0,3915 |
| HADH | 0,8239 | 7,4176 | 0,3591 | 0,0005 | 0,0341 | -0,4219 |
| CD207 | 1,7645 | 1,0437 | 0,3578 | 0,0005 | 0,0351 | -0,3566 |
| AMPH | 2,0462 | 0,7544 | 0,3562 | 0,0005 | 0,0364 | -0,4139 |
| CDC42 | 0,4488 | 9,2736 | 0,3559 | 0,0006 | 0,0364 | -0,6185 |
| PDCD4 | 0,7863 | 7,8440 | 0,3548 | 0,0006 | 0,0372 | -0,5914 |
| ABCD1 | -0,6411 | 5,7953 | -0,3540 | 0,0006 | 0,0376 | -0,5197 |
| ATP1A3 | -1,6121 | -0,1840 | -0,3525 | 0,0006 | 0,0390 | -0,4837 |
| IRF6 | 1,2689 | 5,8938 | 0,3519 | 0,0006 | 0,0392 | -0,4794 |
| GLS | -1,1404 | 6,3986 | -0,3510 | 0,0007 | 0,0398 | -0,6864 |
| DES | -1,9365 | 4,4076 | -0,3458 | 0,0008 | 0,0434 | -0,7041 |
| FGFR1 | 1,0630 | 8,3473 | 0,3453 | 0,0008 | 0,0434 | -0,9019 |
| HSPB2 | -1,2239 | 2,1203 | -0,3458 | 0,0008 | 0,0434 | -0,6199 |
| ICA1 | 0,8456 | 6,5131 | 0,3452 | 0,0008 | 0,0434 | -0,7645 |
| OASL | -1,2933 | 4,2848 | -0,3465 | 0,0008 | 0,0434 | -0,6402 |
| OVOL2 | 0,9084 | 4,9861 | 0,3471 | 0,0007 | 0,0434 | -0,5845 |
| SKP1 | 0,4125 | 9,0654 | 0,3471 | 0,0007 | 0,0434 | -0,8871 |
| SLC27A1 | 0,9219 | 6,4060 | 0,3460 | 0,0008 | 0,0434 | -0,7228 |
| TUBA4A | -0,9864 | 6,5425 | -0,3447 | 0,0008 | 0,0435 | -0,8871 |
| CLDN19 | -2,2421 | -1,8508 | -0,3441 | 0,0008 | 0,0436 | -0,7328 |
| FUT4 | -0,9036 | 4,2550 | -0,3439 | 0,0008 | 0,0436 | -0,7011 |
| EHD1 | -0,3938 | 7,4339 | -0,3433 | 0,0008 | 0,0438 | -0,9662 |
| THY1 | -0,9798 | 7,3973 | -0,3430 | 0,0009 | 0,0438 | -0,9869 |
| DUSP5 | -1,1731 | 5,2921 | -0,3425 | 0,0009 | 0,0441 | -0,8470 |
| OSTC | 0,5092 | 8,1118 | 0,3419 | 0,0009 | 0,0444 | -1,0146 |
| ATP2B4 | 0,8375 | 8,7907 | 0,3403 | 0,0009 | 0,0457 | -1,0792 |

Table S1. *Continued*

| Gene | logFC | AveExpr | t | P value | adj P value | B |
|---|---|---|---|---|---|---|
| EPHX1 | 1,0299 | 9,4952 | 0,3404 | 0,0009 | 0,0457 | -1,0931 |
| CASP6 | 0,6181 | 6,2815 | 0,3395 | 0,0010 | 0,0463 | -0,9249 |
| CLCN2 | -0,8221 | 4,3454 | -0,3388 | 0,0010 | 0,0463 | -0,8522 |
| GAL | -2,0534 | 0,8143 | -0,3390 | 0,0010 | 0,0463 | -0,8226 |
| PEX2 | 0,5866 | 6,8017 | 0,3384 | 0,0010 | 0,0464 | -1,0161 |
| MSX1 | 2,8680 | 8,2560 | 0,3369 | 0,0010 | 0,0478 | -1,0358 |
| CYP2C18 | -2,1826 | -0,8609 | -0,3361 | 0,0011 | 0,0480 | -0,9296 |
| SORBS2 | 2,0008 | 6,0159 | 0,3341 | 0,0011 | 0,0500 | -0,9752 |
| UPP1 | -0,8969 | 5,4543 | -0,3339 | 0,0011 | 0,0500 | -1,1048 |

logFC: fold change of gene expression; AveExpr: average expression; t: P-value; adj P value: *p*-value after Benjamini Hochberg correction for multiple testing.

Table S2. Overview of differentially expressed genes in vascular invasion profiles available in literature.

| | |
|---|---|
| Mannelqvist (2011)<br>Endometrial carcinoma | *Upregulated:* ANGPTL4, COL8A1, FPR2, IL8, MMP, SERPINE1, TNFAIP6<br>*Downregulated:* ALDH1A2, ATCAY, C1orf114, COL4A6, FGFR2, ITIH5, KLHL13, MAMDC2, OGN, OSR2, SEMA5A |
| Kang (2018)<br>Endometrial carcinoma | CLCN2, CPB1, ESR1, FMO2, GREM2, PKHDIL1, PRR9, RPTN, SLC9C2, TCHHL1, TMEM212 |
| Watanabe (2019)<br>Endometrial carcinoma | AUNIP, AURKA, BUB1, BUB1B, C5orf34, CAD, CCNA2, CCNB2, CDC20, CDCA8, CENPA, CHEK1, CKAP2, CKS1B, CLCN2, CSE1L, EIF2AK2, ESPL1, FAM38B, FANCA, FTL, GEN1, GINS1, HJURP, HMGA1, ISYNA1, KIF14, KIF20A, KIFC1, LAGE3, LSM5, LY6K, MCM4, MSH6, NDUFB9, NMU, NOP56, NT5DC2, PHGDH, PLCXD1, PLP2, PRR11, RACGAP1, RAD51AP1, RECQL4, SOX12, SRM, STON2, TACC3, TM7SF2, TNNT1, TONSL, TOP2A, TPX2, UBE2C |
| Regan (2016)<br>NSCLC | *Upregulated:* VEGFC, LYPD6B, TSC22D1, ERGIC2, RPL7A, FTH1, PTPLAD2, ATP2B4, MTRNR2L1<br>*Downregulated:* ZNF280B, LYPD5, NF2, C10orf125, UNC5A |
| Fidalgo (2015)<br>Breast carcinoma | *Upregulated:* C19orf33, LGALS7B, CPE, AGBL2, ARSG, UMOD, C13orf31, MYCBPAP, CXXC4, MED1, SHISA5<br>*Downregulated:* GFRA1, MF12, TBX21, KRT15, KFASC, FZD5, TCF7L1, MYBPC2, CRABP1, CIB2, DUSP3 |
| Kurozumi (2019)<br>Breast carcinoma | *Upregulated:* APOC1, APOE, CALML5, CCNB2, CDCA5, COX6C, DNAJA4, EEF1A2, ELF3, ERBB2, GNAS, HMGA1, HMGB2, HSPB1, IDH2, IFI27, ISG15, KRT18, KRT18P55, KRT19, KRT7, KRT8, LAPTM4B, LRRC26, LY6E, MMP11, MX1, NME1, NOP56, PGAP3, PITX1, PTTG1, S100P, SCD, SLC52A2, SLC9A3R1, SPDEF, TM7SF2, UBE2C, UBE2S, UCP2, YWHAZ<br>*Downregulated:* ACTG2, ANG, ANXA1, C15, CDC42EP4, CEBPD, CFB, CFD, CLIC6, CXCL12, CXCL14, CYBRD1, CYP4X1, DCN, DKK3, DPYSL2, DUSP1, EEF1B2, FBLN1, FCER1A, FCGBP, FGD3, FOS, FST, GAS1, GSTP1, HBA2, HBB, HLA-DQA1, IL17RB, MAOA, MFAP4, MGP,MT1E, NDP, NINJ1, PDGFRL, PLGRKT, PYCARD, RPL3, S100A4, SELENOM, SERPINA3, SERPINE2, SGCE, SLC40A1, SLC44A1, SRPX, STC2, SUSD3, TNS3, TPM2, TXNIP, UBD, VIM, VTCN1, ZBTB20 |
| Asaoka (2020)<br>Breast carcinoma | *Upregulated:* STK26, CDH1, MKI67<br>*Downregulated:* - |
| Minguez (2011)<br>Hepatocellular carcinoma | *Upregulated:* CD24, PGLS, HDLBP, GORASP2, TYMS, UBE2C, CPD, XPOT, YY1AP1, CDKN3, NARF, KDELR1, NOMO2, NDUFS8<br>*Downregulated:* PAH, PPARGC1A, MASP2, DEPDC7, GLYAT, UGT2B15, ZFAND5, KLF9, CYP3A4, SLC38A4, PIK3R1, PON1, DPYS, SLC38A2, GLYATL1, PCK1, MYLK, AASS, MAT1A, ADH4, RCL1 |

Table S2. *Continued*

| | |
|---|---|
| Ho (2006)<br>Hepatocellular carcinoma | AMDP3, TAF4B, SLC4A7, RAB38, RYR1, KIAA0010, DKFZP727M111, KIAA1441, TRIM8, THIN, OGG, MAFA, CSF3R |
| Yue (2019)<br>Serous ovarian carcinoma | *Upregulated:* POSTN, LUM, THBS2, COL3A1, COL5A1, COL5A2, FAP, FBN1<br>*Downregulated:* - |
| Jiang (2019)<br>Colon carcinoma | *DNA copy number alterations:* EP300, NOTCH1, ESR1, AKT1, DVL1, STAT5A, MED1, STAT3, STAT5B, SIRT1, PPARA, HDAC2, ARNT, CSNK1E, HDAc10, KAT2A, MYB, CITED2, IGF1R, MAPK1 |
| Poyet (2017)<br>Urothelial cell carcinoma | *Upregulated:* PDPN, LYVE1, SLP76<br>*Downregulated:* - |
| Sonohara (2019)<br>Oesophageal squamous cell carcinoma | PLAC8, SLC12A8, CSPG4, TFPI, TNFSF10 |
| Peters (2023)<br>Endometrial carcinoma | *Upregulated:* CDH16, SST, KCNQ2, NEFH, L1CAM, DHRS2, IL6, CLDN19, MT1E, SLC6A12, CYP2C18, DDC, GAL, FABP3, MYL4, DES, SLC7A2, SLN, ATP1A3, SCHIP1, RHOF, NEXN, SNX10, OASL, MX2, CASP10, HSPB2, NID2, DUSP5, TRIB3, GLS, CDC25B, CLCF1, SPHK1, TUBA4A, THY1, FUT4, UPP1, IFI30, CLCN2, AURKA, PSMD3, SMAD3, SSH2, TEAD4, NFKB2, PVR, BCL2L1, NOP56, ABCD1, TOMM40, SLC3A2, EDH1<br>*Downregulated:* SKP1, CDC42, EI24, OSTC, TMED10, PEX2, CASP6, CANT1, PLA2G12A, CANX, KDM4B, PDCD4, HADH, ATP2B4, ICA1, CYB5A, OVOL2, KIFAP3, SLC27A1, CYB5R1, PERP, WFS1, EPHX1, FGFR1, NPHP1, SGCB, IRF6, MAST4, LRIG1, SLC1A1, NDRG2, RBP1, CD207, AMOT, SORBS2, AMPH, PCDH7, ESR1, MSX1, MPPED2, NDP |