Differential expression of collagen type IV alpha 6 chain in human epithelial ovarian cancer.

Shahan Mamoor, MS¹

¹shahanmamoor@gmail.com

East Islip, NY USA

Epithelial ovarian cancer (EOC) is the most lethal gynecologic cancer (1). We performed discovery of genes associated with epithelial ovarian cancer and of the high-grade serous ovarian cancer (HGSC) subtype, using published microarray data (2, 3) to compare global gene expression profiles of normal ovary or fallopian tube with that of primary tumors from women diagnosed with epithelial ovarian cancer or HGSC. We identified the gene encoding collagen type IV alpha 6 chain, COL4A6, as among the genes whose expression was most different in epithelial ovarian cancer as compared to the normal fallopian tube. COL4A6 expression was significantly lower in high-grade serous ovarian tumors relative to normal fallopian tube. COL4A6 expression correlated with progression-free survival in patients with ovarian cancer. These data indicate that expression of COL4A6 is perturbed in epithelial ovarian cancers broadly and in ovarian cancers of the HGSC subtype. COL4A6 may be relevant to pathways underlying ovarian cancer initiation (transformation) or progression.

Keywords: ovarian cancer, epithelial ovarian cancer, HGSC, high-grade serous ovarian cancer, systems biology of ovarian cancer, targeted therapeutics in ovarian cancer.

The five-year survival rate for women diagnosed with high-grade serous ovarian cancer is between 30-40% and has not changed significantly in decades (4, 5). The development of novel, targeted therapeutics to treat HGSC can be facilitated by an enhanced understanding of the transcriptional behavior of ovarian tumors relative to that of the normal ovary. We mined published microarray data (2, 3) to compare global gene expression profiles between human ovarian tumors, including that of the HGSC subtype, and that of normal ovarian and fallopian tissue. We identified the gene encoding collagen type IV alpha 6 chain, COL4A6, as among the most differentially expressed in HGSC tumors of the ovary and in epithelial ovarian cancer broadly. COL4A6 may be a gene of interest when prioritizing the study of target genes and pathways for the development of novel therapeutic interventions in epithelial ovarian cancer and specifically in high-grade serous ovarian cancers.

Methods

We used microarray data from datasets GSE124766 (2) and GSE146556 (3) for this differential gene expression analysis of epithelial ovarian cancer in conjunction with GEO2R. The Benjamini and Hochberg method of p-value adjustment was used for ranking of differential expression but raw p-values were used for assessment of statistical significance of global differential expression. Log transformation of data was auto-detected, and the NCBI generated category of platform annotation was used. GSE124766 (2) was generated using Agilent-014850 Whole Human Genome Microarray 4x44K G4112F with $n=3$ normal fallopian tube tissue (control) and $n=8$ for tumors from patients with high-grade serous ovarian cancer. GSE146556 (3) was generated using Affymetrix Human Gene 1.0 ST Array with $n=3$ for normal ovarian tissue and $n=40$ for tumors from patients with epithelial ovarian cancer: $n=26$ high-grade serous ovarian cancer; $n=3$ carcinosarcoma, $n=2$ clear cell, $n=1$ endometrioid, $n=1$ granulosa cell, $n=1$ immature teratoma, $n=4$ mixed, and $n=2$ mucinous; analysis was performed using platform GPL6244. GEO2R provides mRNA expression levels only for the top 250 most differentially expressed genes transcriptome-wide.

A statistical test was performed to evaluate whether COL4A6 expression was significantly different when comparing normal ovarian tissue from control subjects and primary tumors from women diagnosed with HGSC using a two-tailed, unpaired t-test with Welch's correction. We used PRISM for all statistical analyses (Version 8.4.0)(455). For Kaplan-Meier survival analysis, we used the Kaplan-Meier plotter online tool (6) for correlation of COL4A6 mRNA expression levels with progression-free survival (PFS) in $n=483$ ovarian cancer patients.

Results

We mined published microarray data (2, 3) to identify differentially expressed genes in epithelial ovarian cancer and specifically in high-grade serous ovarian cancer (HGSC), the type of ovarian cancer responsible for 70-80% of deaths resulting from the most lethal gynecologic malignancy.

COL4A6 is differentially expressed in ovarian tumors from women diagnosed with HGSC

We identified COL4A6 as among the genes whose expression was most different when comparing primary HGSC tumors to normal fallopian tube tissue (2) (Chart 1). When sorting all of the transcripts measured by microarray based on change in expression between HGSC and the normal fallopian tube, in this data set, COL4A6 ranked 12 out of 40481 total transcripts (Chart 1), equating to 99.97% differential expression. Differential expression of COL4A6 in HGSC tumors was statistically significant (Chart 1; $p=5.95E-09$).

We analyzed a second microarray dataset (3) generated using normal ovarian tissues and tumors

from women diagnosed with epithelial ovarian cancer (EOC) to determine whether differential expression of COL4A6 could be observed in tumors from a different group of patients, which included HGSC tumors.

We found that COL4A6 was differentially expressed in epithelial ovarian cancers broadly (Chart 2). When sorting all of the transcripts measured by microarray based on change in expression between tumors from patients with epithelial ovarian cancers and the normal ovary, COL4A6 ranked 5216 of 29088 total transcripts (Chart 2), equating to 82.1% differential expression. Differential expression of COL4A6 in EOC was statistically significant (Chart 2; $p=1.76E-02$).

COL4A6 is expressed at significantly lower levels in HGSC when compared to the normal fallopian tubes.

We then extracted exact mRNA expression levels for the differentially expressed COL4A6 transcript from both normal fallopian tube tissue and from epithelial ovarian tumors. COL4A6 was expressed at significantly lower levels in high-grade serous ovarian cancers when compared to the normal fallopian tube (Figure 1; $p<0.0001$). We calculated a mean fold change of 0.59 in COL4A6 expression when comparing HGSC tumors to normal fallopian tissue, as COL4A6 was expressed at 12.31 ± 0.40 arbitrary units (AU) in the fallopian tubes but at 7.22 ± 0.55 AU in tumors of the ovary.

COL4A6 tumor expression is correlated with survival outcomes in patients with ovarian cancer.

We performed Kaplan Meier survival analysis (6) to evaluate whether COL4A6 tumor expression correlated with survival outcomes in patients with ovarian cancer. COL4A6 primary tumor expression was correlated with progression-free survival (PFS) in patients with p53 mutant ($p53^{mut}$) ovarian cancer, in the lower quartile (Figure 2). COL4A6 expression was a negative prognostic indicator in patients with ovarian cancer. While median PFS was 22.57 months for ovarian cancer patients with low tumor expression of COL4A6, median PFS was 16.63 months for ovarian cancer patients with high tumor expression of COL4A6 (Chart 3). Correlation of PFS with COL4A6 tumor expression in patients with $p53^{mut}$ ovarian cancer, in the lower quartile, was statistically significant (Figure 2: log rank p-value: 0.0034; hazard ratio: 1.47 (1.13-1.9)).

COL4A6 primary tumor expression was also correlated with progression-free survival (PFS) in patients with $p53^{mut}$ ovarian cancer, in the lower tertile (Figure 2). COL4A6 expression was a negative prognostic indicator in patients with ovarian cancer. While median PFS was 19.23 months for ovarian cancer patients with low tumor expression of COL4A6, median PFS was 17.07 months for ovarian cancer patients with high tumor expression of COL4A6 (Chart 3). Correlation of PFS with COL4A6 tumor expression in patients with $p53^{mut}$ ovarian cancer, in the lower tertile, trended towards statistical significance (Figure 2: log rank p-value: 0.097; hazard ratio: 1.22 (0.96-1.56)).

COL4A6 primary tumor expression was also correlated with progression-free survival (PFS) in patients with $p53^{mut}$ ovarian cancer, in the upper tertile (Figure 2). COL4A6 expression was a negative prognostic indicator in patients with ovarian cancer. While median PFS was 18.79 months for ovarian cancer patients with low tumor expression of COL4A6, median PFS was 16.85 months for ovarian cancer patients with high tumor expression of COL4A6 (Chart 3). Correlation of PFS with COL4A6 tumor expression in patients with $p53^{mut}$ ovarian cancer, in the upper tertile, trended towards statistical significance (Figure 2: log rank p-value: 0.079; hazard ratio: 1.23 (0.98-1.55)).

COL4A6 primary tumor expression was not correlated with progression-free survival in the upper quartile (Figure 2: log rank p-value: 0.17; hazard ratio: 1.19 (0.93-1.53)), in $p53^{mut}$ ovarian cancer patients.

Thus, we found using published microarray data (2, 3) that COL4A6 was among the genes most differentially expressed in tumors from women with high-grade serous carcinomas, and in epithelial ovarian cancer broadly. COL4A6 expression was significantly lower in HGSC tumors when compared to the normal fallopian tubes, and COL4A6 tumor expression was correlated with progression-free survival in patients with ovarian cancer.

Discussion

Transcriptional profiling of HGSC in a Cancer Genome Atlas integrated genomic analysis classified HGSC into four subtypes based on gene clustering: immunoreactive, differentiated, proliferative, and mesenchymal (7). We sought to describe the transcriptional landscape of epithelial ovarian cancers in an unbiased fashion by using published microarray data from primary tumors of women diagnosed with HGSC compared to transcriptome data from normal ovarian and fallopian tube tissue (2, 3). In both datasets analyzed (2, 3), COL4A6 was among the genes whose expression changed most significantly when comparing normal gynecological tissue to ovarian tumors, from women with epithelial ovarian cancer as well as specifically from women with HGSC. COL4A6 may be a molecule of relevance in understanding transformation or disease progression in human epithelial ovarian cancer.

References

1. Guppy, A.E., Nathan, P.D. and Rustin, G.J., 2005. Epithelial ovarian cancer: a review of current management. *Clinical Oncology*, 17(6), pp.399-411.
2. Zhang, W., Klinebiel, D., Barger, C.J., Pandey, S., Guda, C., Miller, A., Akers, S.N., Odunsi, K. and Karpf, A.R., 2020. Global DNA hypomethylation in epithelial ovarian cancer: passive demethylation and association with genomic instability. *Cancers*, 12(3), p.764.
3. Hoffmann, K., Berger, H., Kulbe, H., Thillainadarasan, S., Mollenkopf, H.J., Zemojtel, T., Taube, E., Darb-Esfahani, S., Mangler, M., Sehouli, J. and Chekerov, R., 2020. Stable expansion of high-grade serous ovarian cancer organoids requires a low-Wnt environment. *The EMBO journal*, 39(6), p.e104013.
4. Bowtell, D.D., Böhm, S., Ahmed, A.A., Aspuria, P.J., Bast Jr, R.C., Beral, V., Berek, J.S., Birrer, M.J., Blagden, S., Bookman, M.A. and Brenton, J.D., 2015. Rethinking ovarian cancer II: reducing mortality from high-grade serous ovarian cancer. *Nature reviews Cancer*, 15(11), pp.668-679.
5. Vaughan, S., Coward, J.I., Bast, R.C., Berchuck, A., Berek, J.S., Brenton, J.D., Coukos, G., Crum, C.C., Drapkin, R., Etemadmoghadam, D. and Friedlander, M., 2011. Rethinking ovarian cancer: recommendations for improving outcomes. *Nature Reviews Cancer*, 11(10), pp.719-725.
6. Györfy, B., Lánckzy, A. and Szállási, Z., 2012. Implementing an online tool for genomewide validation of survival-associated biomarkers in ovarian-cancer using microarray data from 1287 patients. *Endocrine-related cancer*, 19(2), pp.197-208.
7. Cancer Genome Atlas Research Network, 2011. Integrated genomic analyses of ovarian carcinoma. *Nature*, 474(7353), p.609.

Probe ID: A_23_P217379

p-value: 5.95E-09

t: 14.7765203

B: 10.52306

Gene: COL4A6

Gene name: collagen type IV alpha 6 chain

Chart 1: COL4A6 is differentially expressed in high-grade serous ovarian carcinomas (HGSC).

The rank of differential expression, the probe/transcript ID, the p-value with respect to global differential expression, t, a moderated t-statistic, B, the log-odds of differential expression between the two groups compared, the gene and gene name are listed in this chart.

1 Probe ID: 8174389
2 p-value: 1.76E-02
3 t: 2.465982
4 B: -3.6856398
5 Gene: COL4A6
6 Gene name: collagen type IV alpha 6 chain

7 Chart 2: COL4A6 is differentially expressed in epithelial ovarian cancers, including high-grade serous
8 ovarian carcinomas (HGSC).

9 The rank of differential expression, the probe/transcript ID, the p-value with respect to global differential
10 expression, t, a moderated t-statistic, B, the log-odds of differential expression between the two groups
11 compared, the gene and gene name are listed in this chart.
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

COL4A6

Figure 1: COL4A6 is expressed at significantly lower levels in HGSC tumors when compared to normal fallopian tube.

The mRNA expression of COL4A6 in normal fallopian tube tissues (left) and in the primary tumors of patients with HGSC (right) is represented with mean mRNA expression level marked and the result of a statistical test evaluating the significance of difference in COL4A6 expression between the normal fallopian tube and primary tumors from patients with HGSC, a *p*-value, listed above.

Figure 2: COL4A6 tumor expression correlates with progression-free survival in patients with ovarian cancer.

Depicted in this Kaplan-Meier plot is the probability of progression-free survival (OS) for n=483 patients, in the lower quartile (top left), in the lower tertile (top right), in the upper tertile (bottom left), and in the upper quartile (bottom right), based on primary tumor expression of COL4A6 (low and high). The log rank *p*-value denoting statistical significance of difference in progression-free survival when comparing the two groups, as well as hazard ratio for this comparison is listed above. Listed below is the number of patients at risk (number of patients alive) per interval, after stratification based on COL4A6 expression; in the first interval, number at risk is number of patients alive; in each subsequent interval, number at risk is the number at risk less those who have expired or are censored.

Lower quartile

1 Low COL4A6 expression: 22.57 months

2 High COL4A6 expression: 16.63 months

3 **Lower tertile**

4 Low COL4A6 expression: 19.23 months

5 High COL4A6 expression: 17.07 months

6 **Upper tertile**

7 Low COL4A6 expression: 18.79 months

8 High COL4A6 expression: 16.85 months

9 Chart 3: Progression-free survival is superior in ovarian cancer patients with low tumor expression of
10 COL4A6.

11 The median progression-free survival of $n=483$ ovarian cancer patients based on stratification into low or
12 high tumor expression of COL4A6 is listed in this chart, in the lower quartile, in the lower tertile, or in the
13 upper tertile.