Differential expression of collagen type I alpha 2 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 I alpha 2 chain, COL1A2, as among the genes whose expression was most different in epithelial ovarian cancer as compared to the normal fallopian tube. COL1A2 expression was significantly higher in high-grade serous ovarian tumors relative to normal fallopian tube. COL1A2 expression correlated with overall survival in patients with ovarian cancer. These data indicate that expression of COL1A2 is perturbed in epithelial ovarian cancers broadly and in ovarian cancers of the HGSC subtype. COL1A2 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 I alpha 2 chain, COL1A2, as among the most differentially expressed in HGSC tumors of the ovary and in epithelial ovarian cancer broadly. COL1A2 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 GSE126308 (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. GSE126308 (3) was generated using Affymetrix Human Transcriptome Array 2.0 with $n=20$ tumors from high-grade serous adenocarcinoma (HGSC) patients classified as early in disease progression and $n=30$ for tumors from HGSC patients classified as “late” in disease progression; analysis was performed using platform GPL17586.

A statistical test was performed to evaluate whether COL1A2 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 COL1A2 mRNA expression levels with overall survival (OS) in $n=1656$ 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.

COL1A2 is differentially expressed in ovarian tumors from women diagnosed with HGSC

We identified COL1A2 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, COL1A2 ranked 189 out of 40481 total transcripts (Chart 1), equating to 99.5% differential expression. Differential expression of COL1A2 in HGSC tumors was statistically significant (Chart 1; $p=1.19E-05$).

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 COL1A2 could be observed in tumors from a different group of patients, which included HGSC tumors.

We found that COL1A2 was differentially expressed in epithelial ovarian cancers when analyzing based on disease progression (Chart 2). When sorting all of the transcripts measured by microarray based on change in expression between tumors from patients early in disease progression and tumors from patients late in disease progression, COL1A2 ranked 18104 out of 70523 total transcripts (Chart 2), equating to 74.3% differential expression. Differential expression of COL1A2 in EOC based on disease progression trended towards statistical significance (Chart 2; $p=8.50E-02$).

COL1A2 is expressed at significantly higher levels in HGSC when compared to the normal fallopian tubes.

We then extracted exact mRNA expression levels for the differentially expressed COL1A2 transcript from both normal fallopian tube tissue and from epithelial ovarian tumors. COL1A2 was expressed at significantly higher 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 1.18 in COL1A2 expression when comparing HGSC tumors to normal fallopian tissue, as COL1A2 was expressed at 9.79 ± 0.11 arbitrary units (AU) in the fallopian tubes but at 11.58 ± 0.37 AU in tumors of the ovary.

COL1A2 tumor expression is correlated with survival outcomes in patients with ovarian cancer.

We performed Kaplan Meier survival analysis (6) to evaluate whether COL1A2 tumor expression correlated with survival outcomes in patients with ovarian cancer. COL1A2 primary tumor expression was correlated with overall survival (OS) in patients with ovarian cancer, in the upper tertile and in the upper quartile (Figure 2)

. COL1A2 expression was a negative prognostic indicator in patients with ovarian cancer, in the upper tertile. While median OS was 46.82 months for ovarian cancer patients with low tumor expression of COL1A2, median OS was 41.6 months for ovarian cancer patients with high tumor expression of COL1A2 (Chart 3). Correlation of OS with COL1A2 tumor expression in patients with ovarian cancer, in the upper tertile, was statistically significant (Figure 2: log rank p-value: 0.012; hazard ratio: 1.19 (1.04-1.36)).

COL1A2 expression was also a negative prognostic indicator in patients with ovarian cancer, in the upper quartile. While median OS was 46.52 months for ovarian cancer patients with low tumor expression of COL1A2, median OS was 38.3 months for ovarian cancer patients with high tumor expression of COL1A2 (Chart 3). Correlation of OS with COL1A2 tumor expression in patients with ovarian cancer, in the upper quartile, was statistically significant (Figure 2: log rank p-value: 0.0003; hazard ratio: 1.25 (1.08-1.44)).

COL1A2 primary tumor expression was not correlated with overall survival in the lower quartile (Figure 2: log rank p-value: 0.5; hazard ratio: 1.05 (0.91-1.22)), or in the lower tertile (Figure 2: log rank p-value: 0.55; hazard ratio: 1.04 (0.91-1.19)).

Thus, we found using published microarray data (2, 3) that COL1A2 was among the genes most differentially expressed in tumors from women with high-grade serous carcinomas, and in epithelial ovarian cancer broadly. COL1A2 expression was significantly higher in HGSC tumors when compared to the normal fallopian tubes, and COL1A2 tumor expression was correlated with overall 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), COL1A2 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. COL1A2 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. Biomarkers in older ovarian cancer patients. GSE126308.
<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE126308>. Sambavy Nadaraja. Odense University Hospital.
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_24_P277934

p-value: 1.19E-05

t: -7.6164318

B: 3.753558

Gene: COL1A2

Gene name: collagen type I alpha 2 chain

Chart 1: COL1A2 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.

Probe ID: TC07000559.hg.1
p-value: 8.50E-02
t: 1.76
B: -4.40831
Gene: COL1A2
Gene name: collagen type I alpha 2 chain

Chart 2: COL1A2 is differentially expressed in epithelial ovarian cancers when comparing tumors based on degree of disease progression - “early” versus “late”.

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.

COL1A2

Figure 1: COL1A2 is expressed at significantly higher levels in HGSC tumors when compared to normal fallopian tube.

The mRNA expression of COL1A2 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 COL1A2 expression between the normal fallopian tube and primary tumors from patients with HGSC, a *p*-value, listed above.

Figure 2: COL1A2 tumor expression correlates with overall survival in patients with ovarian cancer.

Depicted in this Kaplan-Meier plot is the probability of overall survival (OS) for n=1656 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 COL1A2 (low and high). The log rank *p*-value denoting statistical significance of difference in overall 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 COL1A2 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.

Upper tertile

Low COL1A2 expression: 46.82 months

High COL1A2 expression: 41.6 months

Upper quartile

Low COL1A2 expression: 46.52 months

High COL1A2 expression: 38.3 months

Chart 3: Overall survival is superior in ovarian cancer patients with low tumor expression of COL1A2.

The median overall survival of $n=1656$ ovarian cancer patients based on stratification into low or high tumor expression of COL1A2 is listed in this chart, in the upper tertile and upper quartile.