AWARD NUMBER: W81XWH-16-2-0038

TITLE: The Genomic Epigenomic, and Quality-of-Life Charteristics of Long-Term Survivors of Ovarian Cancer

PRINCIPAL INVESTIGATOR: Michael Birrer

CONTRACTING ORGANIZATION: Massachusetts General Hospital
Boston, MA 02114

REPORT DATE: October 2018

TYPE OF REPORT: Annual Report

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE:	2. REPORT TYPE	3. DATES COVERED
October 2018	Annual	30 Sep 2017 - 29 Sep 2018

4. TITLE AND SUBTITLE

The Genomic Epigenomic, and Quality-of-Life Charteristics of Long-Term Survivors of Ovarian Cancer

5a. CONTRACT NUMBER

5b. GRANT NUMBER

W81XWH-16-2-0038

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHOR(S)

Michael Birrer: mbirrer@partners.org, now mbirrer@uab.edu
Lari Wenzel: lwenzel@uci.edu
Giulia Fulci: gfulci@partners.org, now gfulci@uab.edu
George Coukos: gcks@mail.med.upenn.edu
Heather Lankes: hlankes@gogstats.org
Samuel Mok: scmok@mdanderson.org
Ken Nephew: Knephew@indiana.edu
Giovanni Parmigiani: gp@jimmy.harvard.edu
Mary Jackson Scroggins: nekima@aol.com
Nilsa Ramirez: Nilsa.Ramirez@nationwidechildrens.org

E-Mail:

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Massachusetts General Hospital
55 Fruit St
Boston, MA 02114

University of California, Irvine
Irvine, CA 92697

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Ovarian cancer (OC) remains a major health problem in the United Sates (US). In 2012, there will be an estimated 22,280 cases of epithelial OC (EOC) resulting in 15,500 deaths. While the median survival of OC patients has improved over the last two decades, the vast majority of patients suffer relapse and develop chemo-resistant disease. The overall survival of patients suffering from OC has not changed appreciably over the last three decades. Despite these dismal statistics, there is a minority of OC patients who are long-term (LT) survivors (>10 years). This includes a subset of advanced stage (~15%) and a higher proportion of early-stage disease (75%). Unfortunately, there is little genomic or biologic characterization of these tumors, or patient reported outcomes that characterize LT survivors. The clinical importance of identifying subsets of patients who may or may not benefit from therapy, and understanding the biology of their tumors, is significant both from a patient survival and quality of life (QOL) standpoint. The characterization of LT survivors of advanced stage OC will potentially identify molecular and clinical pathways that can be targeted to help women who have shorter survivals. Further, careful characterization of these patients, including their initial and longitudinal health-related QOL reports, their response to treatments, and their tumors will provide significant measures of prognostic factors. Accurate identification of women with high-grade, early stage OC who will recur will allow for tailoring therapy to only those who will benefit. Thus, the systematic molecular and patient-reported outcomes evaluation of LT survivors of OC (both early and advanced stage) will yield data, which can significantly impact the management of OC patients.

15. SUBJECT TERMS					
Ovarian cancer, long-term survival, consortium development, genomics, epigenomics, quality of life, psychosocial					
16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** USAMRMC
a. REPORT U Unclassified	**b. ABSTRACT** U Unclassified	**c. THIS PAGE** U Unclassified	UU Unclassified	**18**	**19b. TELEPHONE NUMBER** *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

1. INTRODUCTION

The consortium will encompass the scientific expertise and historic interest in ovarian cancer (OC) research of the Gynecologic Oncology Group (GOG) and its participating institutions. It will consist of a multi-disciplinary team of investigators from major institutions across the United States (US), each dedicated to the health of women with OC, and a broad-based team of consumer advocates with extensive experience and reach. The consortium will *address key questions related to outcomes of long-term (LT) survivors with OC.* The overarching goal of this multi-institutional proposal is to enhance understanding of molecular, biologic, and patient-reported outcome (PRO) mechanisms, thereby further characterizing predictive markers for LT OC survival and survivorship. This effort will therefore develop predictive biomarkers by (1) discovery, identification, and validation of molecular markers, (2) identification of critical biochemical and immunologic pathways, and (3) identification of PROs that differentiate short-term (ST) and LT survivors and potentially link to predictive biomarkers. The anticipated outcome of this effort will be better stratification and prognostication of OC patients, more effective therapies, and ultimately, increased numbers of LT survivors.

2. KEYWORDS:

Ovarian cancer, methylation, transcriptome, micro RNA, proteomics, quality of life

3. ACCOMPLISHMENTS

What were the major goals of the project?

The major goals of the project include the following:

1.) The determination of the genomic (RNAseq miRNAseq, methylation patterns) and proteomic characteristics of LT versus ST survivors.
2.) The characterization and quantification of immune infiltrates and angiogenesis in LT versus ST survivors.
3.) Validation of a genetic signature that predicts for recurrence of early-stage, high-grade EOC.
4.) Understanding the extent to which health-related QOL measures, additional PROs, and key CTCAE criteria predict LT OC survival
5.) Examination of the potential relationship between health-related QOL, PROs, and key CTCAE criteria and genomic features predicting disease recurrence.

What was accomplished under these goals?

The consortium has dedicated itself to analyzing the tumor specimens from long term survivors along with the appropriate control cases. This analysis has included multiple specific platforms: 1.) microRNAs, 2.) proteomics, 3.) transcriptome, 4.) immune profiling, 5.) methylation, and 6.) quality of life. The platforms have evolved at different rates with some still analyzing the test set of tumors and others moving on to the validation set. Specifically:

MicroRNA Platform

MicroRNAs provide a novel master layer of regulation for gene expression and are expected to be a powerful tool to obtain more representative molecular portraits of specific characteristics and behaviors of tumors at diagnosis. Recent studies identified microRNA (miRNA) signatures for different subtypes of HGSC, and for chemoresistance in HGSC. However, most of these studies used bulk ovarian tumor tissue with various amounts of stromal tissue contamination, which can skew the miRNA profiles that are specific for the cancer cells. In addition, the molecular mechanisms by which these prognostic and predictive miRNAs confer a chemoresistance phenotype, either intrinsic or induced, have not been elucidated. The clinical relevance and usefulness of most of these miRNAs for the prediction of chemoresistance before treatment and patient survival rates have not been defined.

The HTG EdgeSeq miRNA Whole Transcriptome Assay was used to generate microRNA profiles from a total of **254** microdissected formalin fixed paraffin embedded (FFPE) specimens from **GOG 172, 182, and 218**, which include advanced stage high-grade serous ovarian cancer samples. The training data included **178** serous samples, with the remaining **76** samples reserved for validation. In brief, 10 µm sections were cut from FFPE tumor tissue samples and mounted on PEN slides (Leica). After deparaffinization, sections were stained with 1% methyl green. Laser microdissection was then performed to procure 250-500 cells. RNAs were extracted and libraries were prepared using the HTG EdgeSeq system (HTG Molecular), which allows the quantitation of a collection of 2,084 human miRNA transcripts (from miRbase v20) together with 12 housekeepers using the next generation sequencing platform. Two general approaches were first used to normalize the data. One was to use ComBat batch-corrected data without further adjusting for batch in the analysis. The other was to use TMM within-batch normalized data and adjusting for batch as a covariate in the model. For each of these two approaches, several methods were used for analysis, including limma (after voom transformation), linear models (on ComBat adjusted or tmm-normalized data on the log2 scale), Cox model considering miRNA expression level as a covariate, and LASSO as implemented in R package glmnet (both binomial and time to event outcome). The third approach was to use DESeq2, which integrates normalization and differential expression analysis. Six different models were fit, adjusting for batch only, batch and matched strata, or batch, matched strata and additional covariates including age and primary tumor vs metastases, each comparing either those surviving > 8 years to < 8 years, or those surviving > 8 years to < 5 years. The ComBat batch-corrected data and the TMM within-batch normalized data did not identify any significantly differentially expressed mirs after multiple testing adjustments using Benjamini-Hochberg using 0.1 as the threshold. However, using the six models generated from DESeq2, a total of 80 differentially expressed miRNAs were identified at the false discovery rate of 10%.

Proteomics Platform

Archival formalin-fixed, paraffin-embedded (FFPE) ovarian cancer tissue specimens (n=290) from the Gynecologic Oncology Group **(GOG) trials 172, 182,** and **218** were thin sectioned and pathologically assessed to confirm diagnosis, tumor cellularity, and necrosis. Whole tumor tissue was harvested from each of the 290 cases for downstream quantitative proteomic analysis. The analytically validated proteomics discovery platform we employ relies on state-of-the-art high resolution liquid chromatography-tandem mass spectrometry (LC-MS/MS) and highly multiplexed isobaric mass-tag labeling (TMT 11-plex) combined with basic reversed-phase liquid chromatography (bRPLC) fractionation at the peptide level. This allows deep characterization and precise relative proteome-scale quantification from clinically-derived specimens, such as biopsies and surgical specimens (Fig. 1). Proteins were extracted from each ovarian cancer tissue specimen and digested in a 10% buffered acetonitrile solution at 50 °C with a heat stable form of porcine trypsin. Of the 290 specimens received, we recovered a sufficient input quantity (10 microgram) of peptide from 288 (0.7% drop-out rate), which were qualified for TMT labeling. A reference sample was generated by pooling small quantities of the digested proteome from each clinical specimen in the study set. The remaining 10 channels in each 11-plex TMT set comprised independent clinical specimens that were pooled in a randomized fashion to avoid batch effects. In this study, specimens from GOG-172/182 and GOG-218 were not co-mingled in the multiplexes so as to

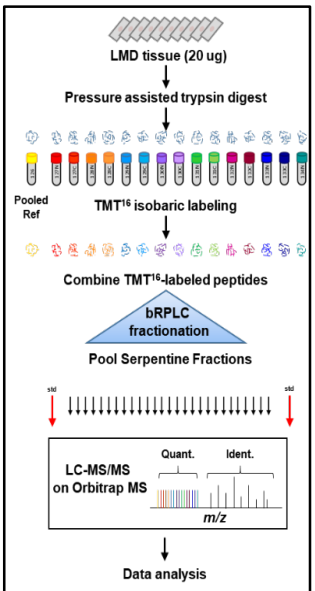

Figure 1: Quantitative tissue proteomics with tandem mass tags and high-resolution mass spectrometry.

serve as independent “discovery” (GOG172/182) and “validation” (GOG-281) cohorts. Each TMT-11 multiplex set was fractionated by basic reversed phase liquid chromatography (LC) and analyzed by high resolution LC-tandem mass spectrometry with a state-of-the-art Orbitrap mass spectrometer (Q-Exactive HF-X, ThermoFisher Scientific Inc.). Standard cell lysate analyses were performed prior to and after each TMT multiplex throughout the project to evaluate LC retention time and MS drift, which was assessed to be less than 6% across the analysis of these 348 peptide fractions, which required 1,218 hours of instrument time.

The quantitative data matrices (log2 transformed protein-level quantitation) were transferred to Dr. Victoria Wang representing the bioinformatics lead agent to identify and validate a protein-based signature of ovarian cancer survival. In total, 5,330 proteins were co-quantified across the 288 specimens. Unsupervised hierarchical clustering was performed with the co-quantified proteins to assess for the presence of batch effect (Fig. 2). No apparent batch effects are evident as a function of TMT multiplex or sample origin (GOG-172/182 or 218). This result indicates that the protein-level data from these formalin-fixed, paraffin-embedded archival ovarian tissue specimens is of high quality.

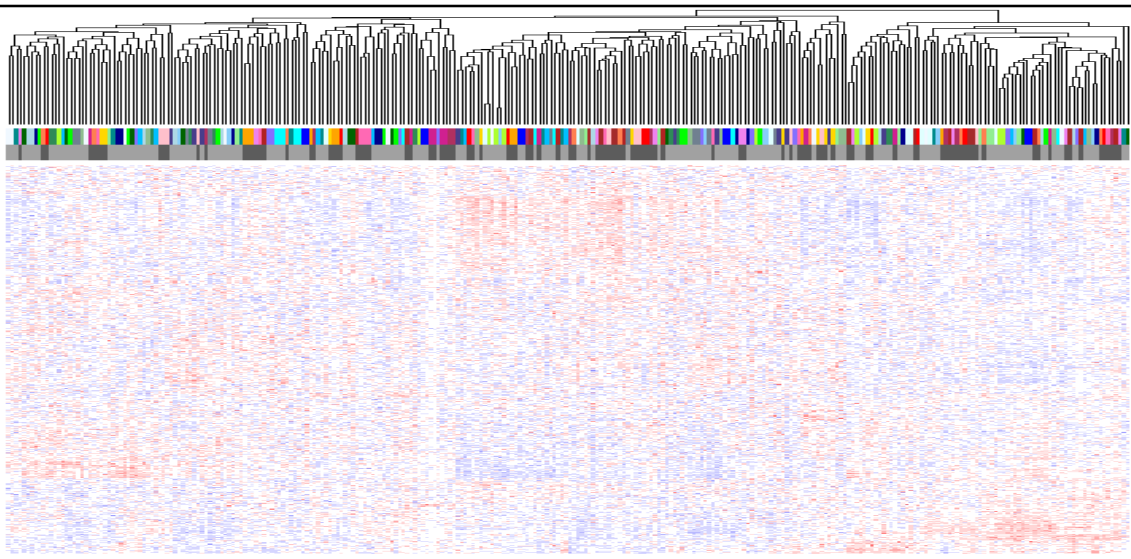

Figure 2: Unsupervised hierarchical cluster of 5,330 co-quantified proteins from 288 specimens. Top colored bars represent TMT multiplexes and gray bars represent sample clinical trial origin (light, GOG-172/182; dark, GOG-218).

Although limited data exists regarding proteomic signatures associated with progression-free or overall (PFS or OS) survival in ovarian cancer, multiple studies have identified prognostic signatures of outcome using RNA-level data. In one study, Konecny et al. validated the prognostic and therapeutic relevance of 4 molecular subtypes (immunoreactive, proliferative, differentiated, and mesenchymal) with a 667 transcript-based signature. Approximately 300 of the 667 transcripts in the Konecny et al. molecular sub-type signature were quantified at the protein level in these GOG-172/182/218 samples. We used our quantitative protein-level data to assess whether there were any significant associations with any of the four molecular sub-types in short term or long term ovarian cancer survivors. Interestingly, we found a significant association of short term survivor proteomes with the mesenchymal molecular subtype (p=0.031).

Diseases or Functions Annotation	p-Value	Activation z-score	Proteins (#)
Cell spreading	5.21E-06	3.08	10
Cell movement of leukocytes	1.88E-03	2.52	13
Adhesion of lymphatic system cells	5.41E-05	2.40	6
Binding of lymphatic system cells	5.54E-05	2.40	7
Binding of lymphocytes	2.64E-04	2.40	6
Cell movement	1.22E-08	2.37	39
Cell spreading of connective tissue cells	4.55E-07	2.22	5
Neoplasia of carcinoma cell lines	1.16E-04	2.21	5
Movement of vascular endothelial cells	1.21E-04	2.21	7
Adhesion of lymphocytes	4.05E-04	2.19	5
Glucose metabolism disorder	1.40E-04	-1.15	21
Apoptosis of endothelial cell lines	1.57E-04	-1.17	4
Tumorigenesis of tissue	6.41E-07	-1.31	92
Cardiac lesion	1.33E-03	-1.53	6
Fibrosis	2.69E-06	-1.65	15
Bleeding	1.39E-03	-1.86	9
Pulmonary emphysema	7.27E-04	-1.98	4
Kyphosis	1.42E-03	-2.00	4
Morbidity or mortality	8.06E-06	-2.36	35
Organismal death	1.65E-05	-2.52	34

Figure 3: Top ten disease function-related pathways predicted to be differentially activated (green) and inhibited disease in short term ovarian cancer survivors.

Among the 5,330 co-quantified proteins, 556 were identified to be significantly altered between short term (<5 yr) and long term (>8 yr) ovarian cancer survivors (LIMMA $p < 0.05$), of which 97 have a log2 fold change of ±1.25. These 97 proteins were assessed by functional gene set enrichment analyses (GSEA) that identified a significant enrichment in proteins predicted to be from the extracellular space in short term survivors. Functional pathway analysis (Ingenuity Pathway Analysis) predicted significant activation of pathways related to cell migration, spreading and invasion, and inhibition of pathways related to cell death, apoptosis, and glucose metabolism in tumors from short term ovarian cancer survivors (Fig. 3).

These data prompted us to hypothesize that there might be significant molecular alterations in the tumor microenvironment that may influence survival. To evaluate this hypothesis, we selected 20 extreme short term survivors and 20 extreme long term survivors. We thin sectioned their tumors onto polyethylene napthalate (PEN) membrane slides, and stained these with hematoxylin and eosin. The tumor epithelium and stromal compartments from each of these cases were selectively harvested by laser microdissection and independently analyzed using our quantitative TMT-based proteomics workflow. This analysis has been completed for 32 cases (17 STS and 15 LTS) and identified 7,582 and 6,337 proteins from the LMD harvested tumor epithelium and stroma, respectively. A differential analysis identified 274 and 195 significantly (LIMMA $p<0.05$) altered proteins in LTS vs STS in the LMD enriched tumor epithelium and stroma, respectively, of which 22 were co-significantly altered in tumor epithelium and stroma (Fig. 4). Semi-supervised hierarchical clustering of the significantly altered proteins from the LMD enriched tumor epithelium (Fig. 5A) and stroma (Fig. 5B) illustrates good classification of LTS from STS.

Figure 4: Venn diagram illustrating the overlap of significantly (LIMMA p<0.05) altered proteins between tumor epithelium and stroma from long term (n=15) and short term (n=17) ovarian cancer survivors.

Methylation

Epigenetic alterations are well recognized as important events in tumor development, growth, and progression. Global hypomethylation and localized hypermethylation have been shown to alter gene expression and disrupt transcription regulatory networks in cancer. Improved characterization of these alterations may contribute to the development of novel biomarkers for prognosis and response to anti-cancer therapies and new therapeutic targets. DNA methylation-mediated silencing of tumor suppressors has been shown to alter gene expression and disrupt

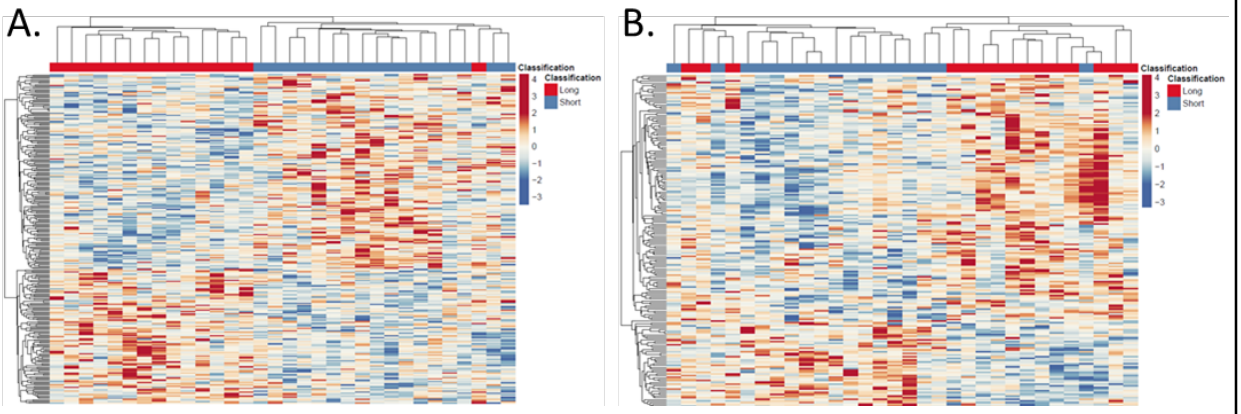

Figure 5: Heat maps illustrating the results of a semi-supervised hierarchical cluster analysis of significantly (LIMMA p<0.05) altered proteins from laser microdissection-enriched tumor epithelium (A) and stroma (B) from long (n=15) vs short (n=17) term ovarian cancer survivors.

transcription regulatory networks in cancer, including ovarian cancer. Improved characterization of these alterations in long- and short-term survivors may contribute to the development of therapies and new therapeutic targets in ovarian cancer.

Methyl-capture was used to profile 315 patient samples genome-wide in two phases. An initial analysis was performed on 100 serous samples profiled in the first phase. A methylation-specific method called csaw was used to identify differentially methylated regions. Window-sizes of 100, 200, 500 and 1000 base pairs were used for analysis. Nearby windows were then further merged into larger windows if found to be homogeneous.

Samples from **GOG 172, 182 and 218** were utilized. DNA was extracted from FFPE embedded tissues on slides and used for epigenomic (DNA methylation) analysis. As described in the original application, samples were analyzed in duplicate using MethylCap-sequencing (*in vitro* capture of methylated DNA by the recombinant methyl-CpG binding domain of MBD2 protein and subsequent analysis of enriched fragments by massively parallel sequencing). Briefly, MethylCap-sequenced genomic DNA (0.2-1 µg) was sonicated before capturing methylated DNA with MBD protein coupled to streptavidin beads. Following capture, the bound methylated DNA was eluted as a single fraction, subjected to library generation and 50-bp single-ended sequencing was performed on the Illumina NextSeq500 (Center for Genomics and Bioinformatics, Indiana University). Sequence files were processed to generate a binary file containing bin counts and scaled count values (a bin size of 500 bp was used to provide sufficient analysis resolution while smoothing the data statistically). Binary count files were interrogated by genomic feature (*e.g.,* CpG islands) to generate feature-specific count files. Samples were grouped according to duration of survival, i.e., LTS vs. STS (long term survivors vs. short term survivors, as defined in the original grant). Significant differences in methylation counts were determined. The Wilcoxon rank-sum test was used to measure the distribution of methylation counts for each locus across the two groups. Significant differentially methylated loci were selected by applying a multiple test corrected p-value cutoff (false discovery rate or FDR). Ingenuity Pathway Analysis (IPA) was used to identify functional interactions of genes differentially methylated between groups (LTS and STS). Methylated CpG islands were hierarchically clustered with Pearson dissimilarity and average linkage as clustering parameters. Boxplots and discretized methylation heat maps were created to visualize differential methylation in LTS and STS groups. Differentially methylated regions were identified at a FDR of 5%.

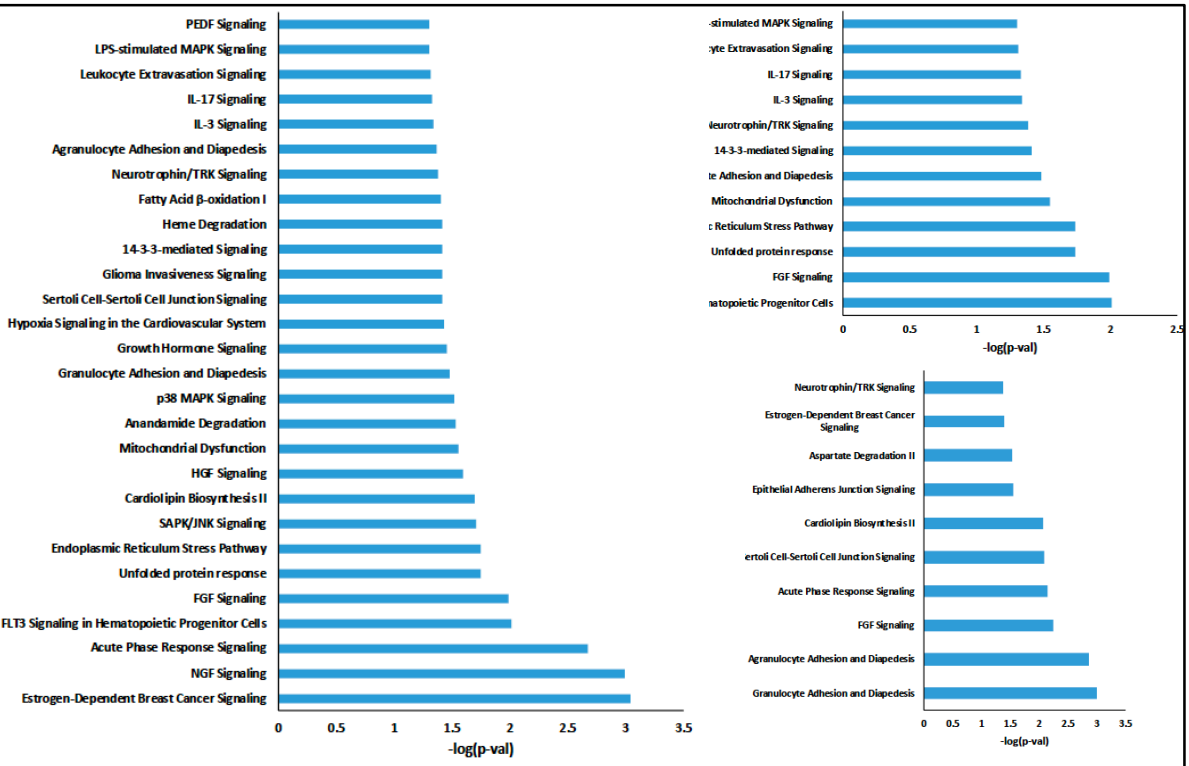

Figure 6. Left, Pathways enriched by 247 differentially methylated genes in LTS vs. STS (hyper- and hypo-methylated genes, left). Right, top panel: Pathways enriched in hypermethylated genes only (135 genes) in LTS vs. STS. Right, lower panel: Pathways enriched in hypomethylated genes (112 genes), LTS vs. STS.

Pathways and networks identified by IPA analysis of the differentially methylated genes in the two groups are shown below in Figure 6. A total of 247 genes with 135 hypermethylated genes in LTS and 112 hypomethylated genes in LTS. Representative results are shown.

Networks identified by IPA analysis of the differentially methylated genes in the two groups are shown below (total of 247 genes; 135 hypermethylated genes in LTS, and 112 hypomethylated genes in LTS). The significant networks (Score 20 or greater) and Representative networks are listed below in Figure 7. Preliminary cross-validation of the high-dimensional data is shown below in Figure 8 (heat map, survival curve, and risk score).

ID	Score	Focus Molecules	Top Diseases and Functions
1	39	22	[Cell-mediated Immune Response, Cellular Movement, Hematological System Development and Function]
2	34	20	[Developmental Disorder, Hereditary Disorder, Metabolic Disease]
3	34	20	[Cellular Assembly and Organization, Cellular Function and Maintenance, Endocrine System Development and Function]
4	30	18	[Auditory Disease, Cardiovascular System Development and Function, Connective Tissue Disorders]
5	28	17	[Cellular Movement, Connective Tissue Disorders, Developmental Disorder]
6	25	16	[Cell Signaling, Cellular Function and Maintenance, Inflammatory Response]
7	23	15	[Connective Tissue Development and Function, Embryonic Development, Organ Development]
8	23	15	[Developmental Disorder, Neurological Disease, Ophthalmic Disease]
9	21	14	[Energy Production, Lipid Metabolism, Small Molecule Biochemistry]

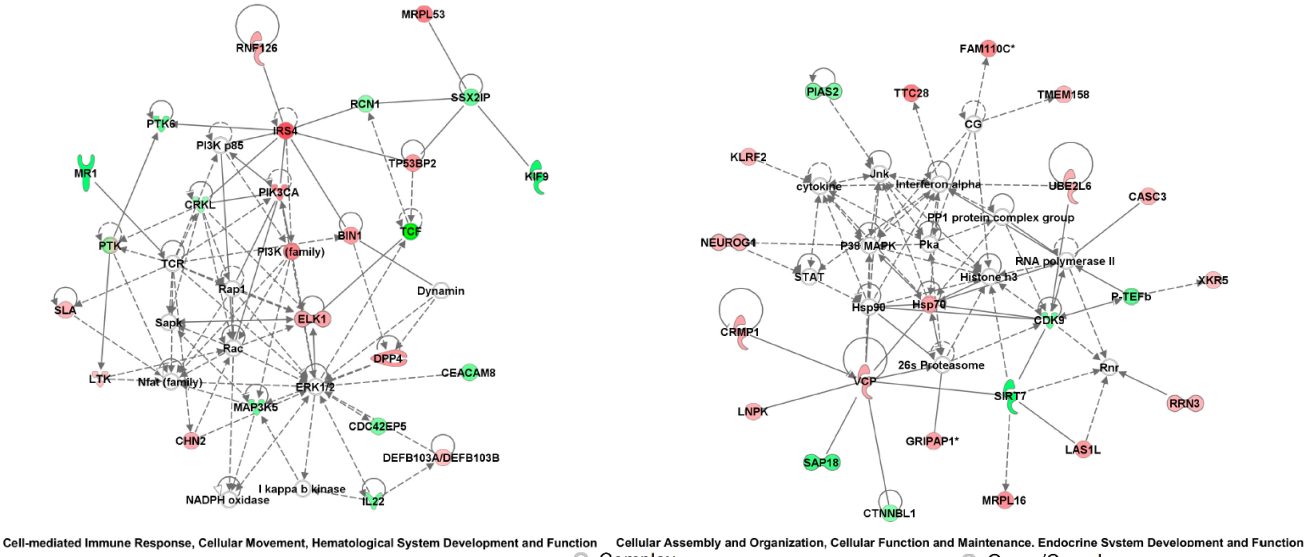

Figure 7. Representative networks enriched by 247 differentially methylated genes in LTS vs. STS (hyper- and hypo-methylated genes) from table. Left: Network ID1, Cell-mediated immune response. Right: Network ID3, Cellular Assembly and Organization. Key is to the right.

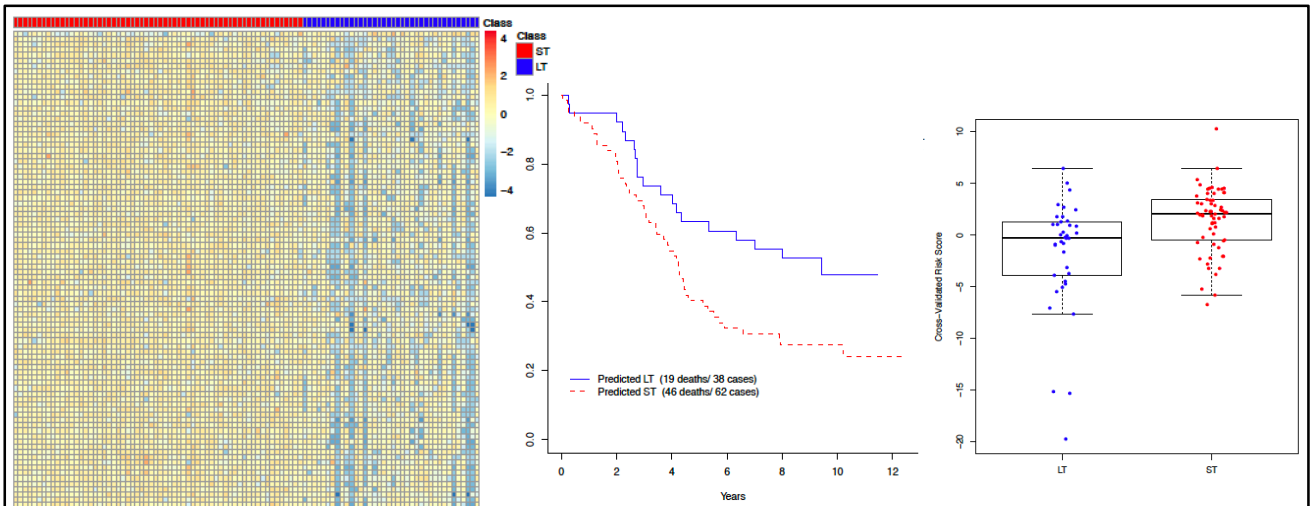

Figure 8. Cross-validation analysis. Representative heat map (left), survival probability (middle) and risk score (right) of differentially methylated genes in

Quality of Life

There is a critical need to identify and understand differences and similarities in long-term and short-term ovarian cancer survivor characteristics in order to improve upon long-term survival. Tumor biology (e.g., cancer stage, grade, histology, cytology) and age at diagnosis have clear associations with overall survival. In addition, there is considerable evidence that quality of life (QOL) at study initiation, as well as QOL changes over time are also significantly associated with overall survival. While these relationships are not well understood, it is known that a cancer patient's responsiveness to cancer therapy may positively affect QOL by decreasing disease burden, thereby influencing survival. By extension, it is reasonable to expect that QOL may be related to adverse events (AEs), as adherence to cancer treatment and maintenance regimens are clearly affected by experiencing AEs.

Examining both QOL and AEs across a clinical trial trajectory could assist in identifying the point in time at which risk for becoming a short versus long-term survivor can be detected and potentially addressed. Moreover, the extent to which QOL and AEs might predict long-term survival is unknown. The purpose of this component of our DOD study, and progress reported herein, is to determine the relationship between QOL and AEs on long-term vs. short-term survival of advanced ovarian cancer patients receiving frontline treatment.

We utilized data from Gynecologic Oncology Group 218 (GOG-218), a phase III randomized clinical trial that tested the efficacy of bevacizumab incorporated into standard frontline treatment of patients with stage III or stage IV ovarian epithelial, primary peritoneal, or fallopian tube cancer. QOL was measured using the Functional Assessment of Cancer Therapy – Ovary Trial Outcome Index (FACT-O-TOI), which captures the FACT-General QOL, dimensions of Physical Well-Being, Functional Well-Being, and an Ovarian Cancer Subscale. Measures were assessed at baseline, before cycles 4, 7, 13, 21, and 6 months after completing protocol-directed therapy. QOL and AEs were compared between short-term (≤5 years; STS) and long-term survivors (8+ years; LTS) at baseline, before cycles 4, 7, 13, 21, and 6 months post-treatment using general linear models and mixed hierarchical regression. Adjustments were made for age, stage, grade, performance status, residual disease, and treatment. We also examined change in QOL between adjacent time points to determine the potential point at which the short-term survivors' QOL decreases, utilizing a general linear model to include covariates of age, stage, grade, performance status, residual disease, treatment, number of treatment cycles, with baseline QOL. AEs were compared between survivor groups at each assessment using general linear models adjusting for patient covariates.

Significant QOL differences between short-term (<5 years; N=1,115) and long-term (8+ years; N=260) survivors first emerged at baseline, demonstrating a significantly increased probability of being a LTS, OR=1.012 (95% CI 1.000-1.025, p=0.046). Further, LTS have a greater improvement in QOL

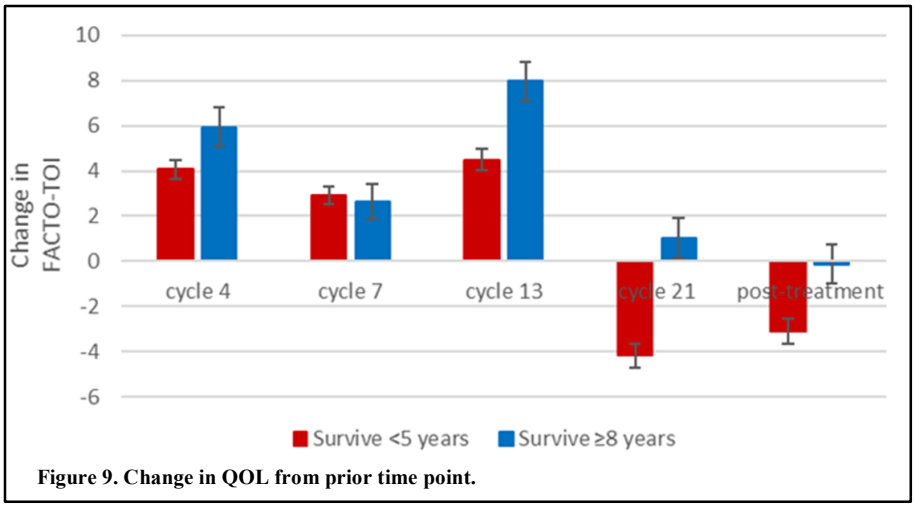

Figure 9. Change in QOL from prior time point.

than STS from baseline to cycle 4 (p=0.052), with no difference in change in QOL from cycle 4 to cycle 7 (p=0.753). Larger improvements in QOL changes occurred between cycles 7 and 13 (p=0.001), with continued greater improvement in QOL from cycle 13 to cycle 21 (p<0.001), and from cycle 21 to post-Rx (p=0.004). As evident in Figure 9, QOL declined for the STS group after cycle 13.

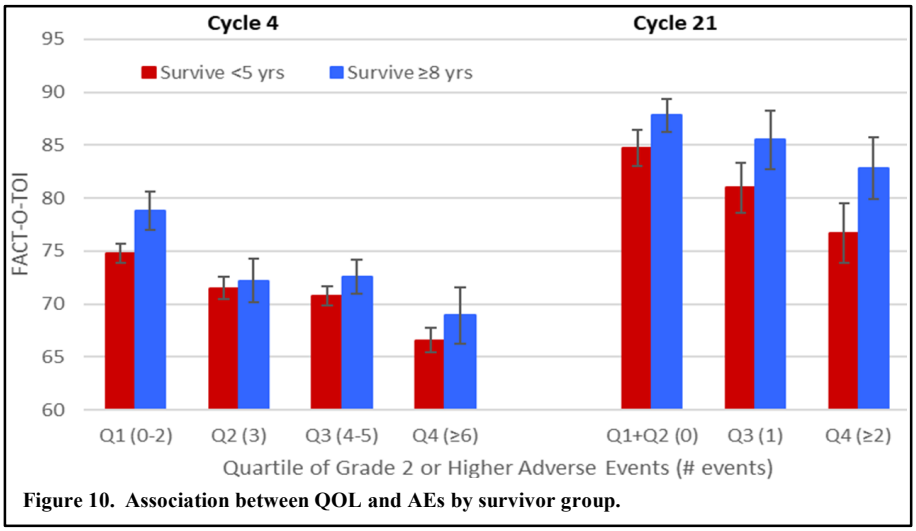

Figure 10. Association between QOL and AEs by survivor group.

With respect to QOL and AEs, QOL at cycle 4 decreased significantly with increasing AEs (p=0.042) independent of survivor group. Further, QOL decreased with increasing number of grade 2 or higher AEs at cycle 4 (p<0.001) and cycle 21 (p=0.017). The number of adverse events did not differ significantly between survivor groups at cycles 4, 7, or 21 (Fig. 10).

Our preliminary data from GOG-218 demonstrate that relationships between QOL at baseline, as well as later assessment time points, can distinguish short versus long-term survivors. Further, methylome analyses utilizing GOG-218 and other samples indicate that a cross-validated risk score can distinguish differentially methylated genes in LTS vs. STS. We anticipate that, in a pilot test of a subsample of 218 QOL and biomarker data, differential methylation in LTS and STS groups may be correlated with baseline quality of life. In addition, we hypothesize that advanced quantitative mass spectrometry analysis of baseline serum samples collected from GOG-218 patients at the initial surgery (prior to chemotherapy cycle) will enable identification of a protein biomarker panel that correlates with our baseline QOL data. Proteins that are identified to be associated with baseline QOL will be prioritized for validation in an independent set of serum samples from GOG-218.

Bioinformatic Analysis

There are approximately 315 ovarian samples from three GOG trials (172, 182 and 218), profiled by RNA-seq, miRNA, proteomics, methylation and tumor infiltrating lymphocytes. The goal is to identify molecular features that can predict long-term survival, currently defined as 8 years or longer from randomization. About 1/3 of the samples are from long-term survivors, and the majority (90%) are of serous type. A stratified sampling approach was used to select samples, where the strata were defined by trial, treatment arm, stage and residual disease status.

Samples are divided into training and validation at an approximately 2:1 ratio. Using the training data, molecular features showing a difference between long-term and short-term survivors are identified, and classifiers are developed that can predict long-term survival status. Cross-validation is used to tune model parameters as well as assess classifier performance. The best performing classifier by cross-validation is then tested in the validation data.

Summary of progress

We have completed the profiling of samples for miRNA, methylation and proteomics. Gene expression by RNA-seq and tumor infiltrating lymphocytes are complete for about half of the samples.

miRNAs

The HTG EdgeSeq system was used to profile about 2000 miRNAs by sequencing for 254 samples in 8 batches. The training data included 178 serous samples, with the remaining serous samples reserved for validation. Two general approaches were first used because of difficulty in normalizing the data. One was to use ComBat batch-corrected data without further adjusting for batch in the analysis. The other was to use TMM within-batch normalized data and adjusting for batch as a covariate in the model. For each of these two approaches, several methods were used for analysis, including limma (after voom transformation), linear models (on ComBat adjusted or tmm-normalized data on the log2 scale), Cox model considering mir expression level as a covariate, and LASSO as implemented in R package glmnet (both binomial and time to event outcome). None of the above identified any significantly differentially expressed mirs after multiple testing adjustments using Benjamini-Hochberg with 0.1 as the threshold. The third approach was to use DESeq2, which integrates normalization and differential expression analysis. Six different models were fit, adjusting for batch only, batch and matched strata, or batch, matched strata and additional covariates including age and primary tumor vs metastases, each comparing either those surviving > 8 years to < 8 years, or those surviving > 8 years to < 5 years. A total of 80 differentially expressed mirs were identified from these six models at the false discovery rate of 10%.

Proteins

An LC/MS system was used to profile 288 patient samples in two phases. There were 4536 proteins measured in both phases. Protein data was processed and normalized by INOVA, where the data was generated. The training data included 190 serous samples, with the remaining serous samples reserved for validation. Several methods were used to identify proteins associated with long-term survival, including (1) linear models (Limma or ordinary linear models for each protein), adjusting for matched strata, matched strata and age, or no adjustments, (2) Wilcoxon test, each for a >8yrs vs <8 yrs or >8rs vs < 5 yrs comparison, (3) Cox regression adjusting for matched strata and age, and (4) Lasso and coxnet for model selection. A total of 44 proteins were identified to be associated with long-term survival at the false discovery rate of 10%.

Methylation

Methyl-capture was used to profile 315 patient samples genome-wide in two phases. An initial analysis was performed on 100 serous samples profiled in the first phase. A methylation-specific method called csaw was used to identify differentially methylated regions. Window-sizes of 100, 200, 500 and 1000 base pairs were used for analysis. Nearby windows were then further merged into larger windows if found to be homogeneous. Approximately 900 differentially methylated regions were identified at a false discovery rate of 5%.

Transcriptome

Gene expression was profiled by RNA-seq for 95 serous samples. Reads were mapped to the transcriptome using STAR, and counts were summarized for about 20,000 protein-coding genes. Data were normalized using TMM and followed by batch correction using ComBat. DESeq2, limma, and Cox regression were used to identify genes whose expression levels have an effect on survival status, each adjusting for matched strata, matched strata and age, or no adjustment. A total of 50 genes were identified at a false discovery rate of 10%.

Tumor infiltrating lymphocytes
TILs have been profiled for the tumor and stroma components of 102 serous samples. Initial analysis was performed using Cox and logistic regression to assess the effect of immune markers on long-term survival, identifying a small number of potential candidates. The analysis will be repeated when more samples are profiled.

Classifier development
Classifiers to predict long-term survival were developed from each platform separately, for patients with or without residual disease. These were not initially considered for the stratified sampling design. Estimated areas under the curve (AUCs) ranged between 0.71 and 0.78 as assessed by cross-validation. While molecular profiling was completed for all platforms, performance assessment for classifiers is being updated to reflect the stratified sampling design so that the AUCs reflect how well the classifiers would do in the patient population from the samples selected, and not just how well they do among the selected samples. In addition, the split of training and validation data are also being updated so that distribution of patient populations is well represented in the validation set. Finally, an integrated classifier will be developed that incorporates information from all molecular platforms as well as clinical variables.

What opportunities for training and professional development has the project provided?

The project has provided multiple opportunities for training both at the graduate and post-doctoral level:
Allison Hunt, PhD, Postdoc, Inova Women's Health Integrated Research Center
Amanda Jackson, MD, GYN-ONC Fellow, National Capital Consortium Fellowship, Walter Reed National Military Medical Center
Jeremy Loffredo, Graduate (PhD) Student, George Mason University
Tomasz Szul, Research Associate UAB
Brandon Roane GYN-ONC Fellow UAB
Justin Wilford, PhD – Post-doc UCI
La Keisha Jeanmarie –Grad student in Public Health, UCI

How were the results disseminated to communities of interest?

The consortium has attempted to disseminate the results by publishing relevant papers (see above) and presenting the data at national meetings (see above).

What do you plan to do during the next reporting period to accomplish the goals?

Nothing to report

4. IMPACT

Impact on the development of the principal discipline(s) of the project:

The results from the individual technology platforms will undoubtedly provide significant information on the proteogenomic basis for long-term survivors of ovarian cancer. This data should have a large impact on providing important information on identifying and triaging patients who would be predicted to be long term survivors. In addition, the quality of life data will provide a complete characterization of the predictive value of QOL for these patients and important information related to their QOL later in life.

Impact on other disciplines:

Nothing to report

Impact on technology transfer:

Nothing to report

Impact on society beyond science and technology:

Nothing to report

5. CHANGES/PROBLEMS

Changes in approach and reasons for change:

Nothing to report

Actual or anticipated problems or delays and actions or plans to resolve them:

Nothing to report

Changes that had a significant impact on expenditures:

None

Significant changes in use or care of human subjects:

None

Significant changes in use or care of vertebrate animals:

None

Significant changes in use of biohazards and/or select agents:

None

6. PRODUCTS

Journal Publications
Acknowledgment of support is included in the following published and submitted manuscripts

Yeung TL, Sheng J, Leung CS, Li F, Kim J, Ho SY, Matzuk MM, Lu KH, Wong STC, Mok SC. Systematic Identification of Druggable Epithelial-Stromal Crosstalk Signaling Networks in Ovarian Cancer. J Natl Cancer Inst. e-Pub 5/2018. PMID: 29860390.

Hinchcliff E, Paquette C, Roszik J, Kelting S, Stoler MH, Mok SC, Yeung TL, Zhang Q, Yates M, Peng W, Hwu P, Jazaeri A. Lymphocyte-specific kinase expression is a prognostic indicator

in ovarian cancer and correlates with a prominent B cell transcriptional signature. Cancer Immunol Immunother 68(9):1515-1526, 9/2019. e-Pub /2019. PMID: 31515669.

Zhu Y, Ferri-Borgogno S, Sheng J, Yeung TL, Burks JK, Jazaeri AA, Capello P, Birrer MJ, Mok SC, Wong STC. A deep machine learning pipeline to identify prognostic immune biomarker in advanced high grade serous ovarian cancer. Cancer Cell, submitted.

Books or other non-periodical, one time publications

None

Other publications, conference papers, and presentations.

AACR Special Meeting: ADVANCES IN OVARIAN CANCER RESEARCH SEPT 2019 Atlanta Georgia Plenary Presentation
Long term survivors of ovarian cancer: a Department of Defense Initiative

O'Neal Cancer Canter Scientific Research Series, UAB Birmingham AL
Jan 2019
New Advances in Ovarian Cancer

Technologies or techniques

None

Interventions, patent applications, and/or licenses

None

Other Products

None

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?

Name	Michael Birrer
Project Role	PI
Researcher Identifier (ORCID ID)	0000-0001-6464-4225
Nearest Person Month Worked	2.4 calendar Months
Contribution to Project	Contributed to conception and design of study. Supervises all research procedures and analyses at UCI. Collaborates closely with each platform's teams on results interpretation, presentations, manuscripts, and grant proposals derived through the study.
Funding Support	1U01CA152990-01 1U01CA214114-02

Name	Lari Wenzel, PhD
Project Role	Co-PI
Researcher Identifier (ORCID ID)	0000-0002-3393-0349
Nearest Person Month Worked	2.4 calendar months
Contribution to Project	Contributed to conception and design of study. Supervises all research procedures and analyses at UCI. Collaborates closely with each platform's teams on results interpretation, presentations, manuscripts, and grant proposals derived through the study.
Funding Support	DOD W81XWH-16-2-0038

Name	Samuel C. Mok
Project Role	Co-investigator
Researcher Identifier (ORCID ID)	https://orcid.org/0000-0001-7013-1805
Nearest Person Month Worked	0.6 calendar months
Contribution to Project	Perform microdissection and generate miRNA transcriptome data from GOG tissue samples
Funding Support	R01CA239519 (NIH) W81XWH-17-1-0146 (DOD) W81XWH-17-1-0126 (DOD) W81XWH-18-1-0177 (DOD) Tina Brozman Foundation for Ovarian Cancer Research R01CA184918 (NIH) P50CA098258 (NIH)

Name	Thomas P. Conrads, PhD
Project Role	Co-Investigator
Researcher Identifier (ORCID ID)	https://orcid.org/0000-0003-4742-3281
Nearest Person Month Worked	0.6 calendar months
Contribution to Project	Oversight and execution of mass spectrometry-based tissue proteomics.
Funding Support	$136, 306

Name	Kathryn Osann, PhD
Project Role	Sub-award Biostatistician
Researcher Identifier (ORCID ID)	0000-0003-1364-7367
Nearest Person Month Worked	1.8 calendar months
Contribution to Project	Conducts statistical analyses to execute study aims they relate to patient-reported outcome endpoints. Collaborates with other

	biostatisticians on the study to integrate analyses of clinical and patient-reported outcomes data with biological specimens to predictor long-term survival of OC patients.
Funding Support	DOD W81XWH-16-2-0038

Name	Chelsea McKinney, PhD, MPH
Project Role	Sub-award Project Manager
Researcher Identifier (ORCID ID)	0000-0002-0873-8582
Nearest Person Month Worked	1.8 calendar months
Contribution to Project	Manages data acquisition and quality. Tracks study progress towards milestones and assists with the preparation of presentations, manuscripts, and subsequent grant proposals.
Funding Support	DOD W81XWH-16-2-0038

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

No

What other organizations were involved as partners?

Not applicable

8. SPECIAL REPORTING REQUIREMENTS

Not Applicable

9. APPENDICES

Not applicable