AWARD NUMBER: W81XWH-16-1-0403

TITLE: Transcriptomic Profiling and Functional Characterization of Fusion Genes in Recurrent Ovarian Cancer

PRINCIPAL INVESTIGATOR: Adrian V. Lee, Ph.D.

CONTRACTING ORGANIZATION: University of Pittsburgh
Pittsburgh, PA 15213

REPORT DATE: Sept 2018

TYPE OF REPORT: Annual

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
Sept 2018	Annual	15 Aug 2017-14 Aug 2018

4. TITLE AND SUBTITLE
Transcriptomic Profiling and Functional Characterization of Fusion Genes in Recurrent Ovarian Cancer

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-16-1-0403

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Adrian V. Lee, Ph.D.

E-Mail: leeav@upmc.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Pittsburgh
4200 Fifth Ave,
Pittsburgh,
PA 15260

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
High-grade serous ovarian cancer (HGSOC) is known for its lack of early detection, limited therapies, and high rate of recurrence. Recent advances in transcriptomic sequencing have identified drug-targetable, pathogenic fusion genes in solid cancers. We hypothesize that fusion genes are commonly acquired or enriched in relapsed HGSOC and contribute to the enhanced malignancy observed in recurrent disease. In the first year of this proposal we assembled a cohort of 18 patient matched pairs of chemotherapy naïve and resistant HGSOC and performed RNA sequencing. We noted transcriptional similarity between the patient-matched pairs of samples, but several recurrent transcriptional remodeling events were noted. Every case showed acquisition of RNA fusions in the recurrent disease. Some fusions acquired in the chemotherapy-resistant HGSOC are found in HGSOC cell lines. We will examine the expression of these fusions in patients treated with neo-adjuvant chemotherapy (pre and post therapy) and examine the biologic function of prioritized RNA fusion events.

15. SUBJECT TERMS
NONE LISTED

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	14	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

1) INTRODUCTION:

High-grade serous ovarian cancer (HGSOC) is known for its lack of early detection, limited therapies, and high rate of recurrence. Greater than 80% of patients with late-stage HGSOC recur after an initial response to chemotherapy, with the majority of relapsed tumors developing deadly resistance to subsequent chemotherapies. The generation of fusion mRNA transcripts is an oncogenic event in many cancer types. Recent advances in transcriptomic sequencing have identified drug-targetable, pathogenic fusion genes in solid cancers. We hypothesize that fusion genes are commonly acquired or enriched in relapsed HGSOC and contribute to the enhanced malignancy observed in recurrent disease. The goal of this proposal is to test this hypothesis with the following specific aims; 1) To define the presence and relative expression of fusion mRNA transcripts in primary and recurrent high grade serous ovarian cancer (HGSOC). 2) To establish the prevalence and clinical importance of identified pathogenic gene fusions 3) To determine the biological effect, and mechanistic action, of fusion candidates acquired in relapsed disease. This study will provide novel targets and biomarkers for a cancer with limited options. This pilot project will develop key preliminary data critical for further analysis of RNA fusions in recurrent HGSOC and may identify new prognostic markers and ultimately therapeutic targets for reversing HGSOC chemoresistance, reducing recurrence, and extending patient survival.

2) KEYWORDS:

High grade serous ovarian cancer, chemotherapy resistance, RNA fusions, prognosis, recurrence, sequencing

3) ACCOMPLISHMENTS:

What were the major goals of the project?

Specific Aim 1) Identify fusion transcripts in recurrent HGSOC	**Timeline**	**Progress**
Major Task 1 RNA-sequencing of recurrent HGSOC	Months	
Local IRB/IACUC Approval	0	Completed
Submission of institution's IRB approval and related material for DoD's HRPO approval	0-1	Completed
Receive HRPO approval or exempt finding before initiating relevant tasks	1-3	Completed
Subtask 1 Pathology analysis of 20 pairs of primary and recurrent HGSOC	3-4	Completed
Subtask 2 Isolation of RNA from 20 pairs	3-5	Completed

Subtask 3 Sequencing using NextSeq500 in sequencing core at Pitt	5-6	Completed
Milestone(s) Achieved – RNA-seq data from primary and recurrent Pitt tumors (n=23 total – 3 pairs performed for preliminary data)		Completed
Major Task 2 Bioinformatic analysis and validation of fusions		
Subtask 1 Analysis of RNA-seq by mapping with STAR and calling fusions using Fusion MetaCaller. Analysis of gene expression using STAR and Deseq	6-7	Completed
Subtask 2 Validation of candidate fusions using RT-PCR and Q-RT-PCR	7-8	Partial – in progress
Subtask 3 Validation of a select number of fusions by FISH	8-10	Not started
Milestone(s) Achieved: Validated fusion mRNAs present in recurrent HGSOC		Partial – in progress
Specific Aim 2) Establish the prevalence and clinical significance of identified fusion genes		
Major Task 3 Isolate RNA and measure fusion using Nanostring		
Subtask 1 Procure 60 FFPE recurrent samples and 200 primary HGSOC	4-8	Partial – in progress
Subtask 2 Isolate and measure RNA from 260 samples	8-10	Partial – in progress
Subtask 3 Develop NanoString codeset based upon fusions from Aim 1	10-14	Not started

What was accomplished under these goals?

For this reporting period describe:

1) major activities

The major activity of the second year of funding was to complete bioinformatics analysis of the RNA fusion dataset from year 1 and prioritize fusions for further analysis and validate expression in an independent cohort.

2) specific objectives

The major tasks of the second year were a) Bioinformatic analysis and validation of fusions, and b) Isolate RNA and measure fusion using Nanostring

3) significant results or key outcomes, including major findings, developments, or conclusions

a) Local IRB/IACUC Approval

Written and approved

b) Submission of institution's IRB approval and related material for DoD's HRPO approval

Written and approved

c) Receive HRPO approval or exempt finding before initiating relevant tasks

Written and approved

d) Subtask 1 Pathology analysis of 20 pairs of primary and recurrent HGSOC

The details of the unique cohort of patient-matched chemotherapy naïve and treated HGSOC are described in the Yr 1 report.

e) Subtask 2 Isolation of RNA from 20 pairs

Completed. All samples had a RIN score above 7.5 and this made them suitable for downstream sequencing.

f) Subtask 3 Sequencing using NextSeq500 in sequencing core at Pitt

Completed. All QC metrics were excellent.

Major Task 2 Bioinformatic analysis and validation of fusions

g) Subtask 1 Analysis of RNA-seq by mapping with STAR and calling fusions using Fusion MetaCaller

In the year 1 report, we detailed analysis of the data based upon FusionCatcher which has very high specificity. Fusion RNAs were called with FusionCatcher v0.99.7b. We identified fusions in the chemotherapy naïve (early, E) and recurrent (late, L) HGSOC. Nearly all recurrences harbored “preserved” fusions—fusion transcripts detected in both the early and late lesion (Figure 1A). Although no fusions acquired in the recurrent sample were present in more than one patient, fusions of particular interest were identified. Because preserved fusions (present in both early and late samples) were found to be common in ovarian cancer recurrences, we searched for preserved fusion genes that were shared in multiple samples. CCDC6-ANK3 was found to harbor distinct breakpoints in two different HGSCOC samples and in the OVCA3 cell line. These breakpoints were confirmed with RT-PCR.

Since the main goal of the proposal was to study recurrent fusions, in Year 2 we applied a more comprehensive approach to fusion detection, using two different algorithms StarFusion and FusionCatcher and using two different versions of each (as new versions came out while the analysis was underway and gave quite different results. The versions are FusionCatcher v0.99.7 and v1.00, and StarFusion 2.4 and 2.5.3a.

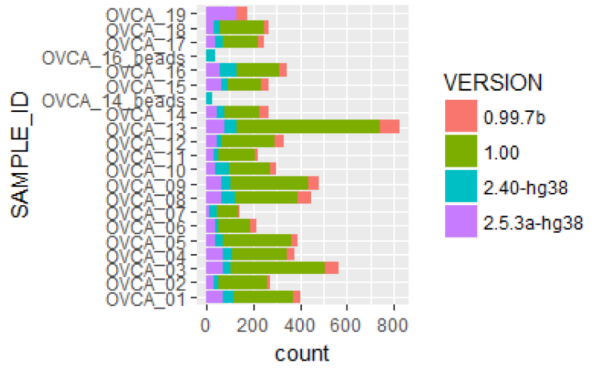


Figure 1: RNA fusions called by two different algorithms (StarFusion and FusionCatcher) using two different versions. Y-axis represents the number of fusions called per case indicated on the X-axis.

Figure 1 shows the results from fusions identified in the 19 cases of early and late HGSOC using the four different algorithms. Note the large variation even between different versions of the same algorithm. For example, we originally used FusionCatcher v 0.99.7b and found only a small numbers of fusions per case (Fig 1, red), however, using the newly released v 1.00 showed a much greater number of fusions were identified (Fig 1, green). This highlights the challenge of RNA fusion detection.

We examined the genomic distribution of these variants using visualization with CIRCOS plots. Figure 2 shows that fusions in case #1 are spread across most of the genome with a number of genomic regions which show an enrichment for a higher rate of fusion events

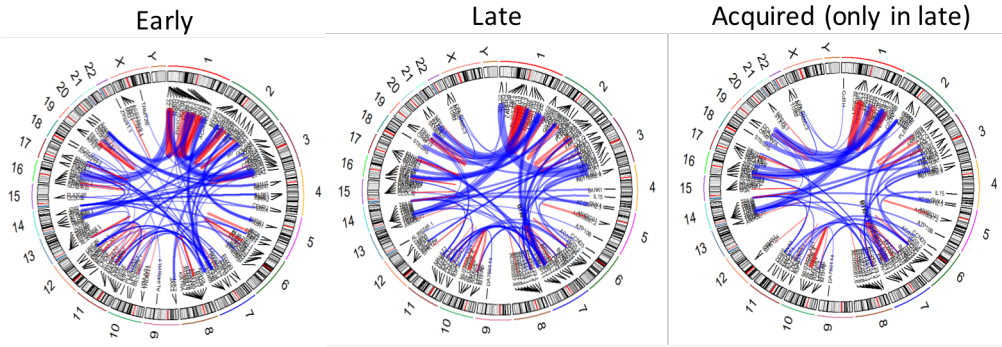
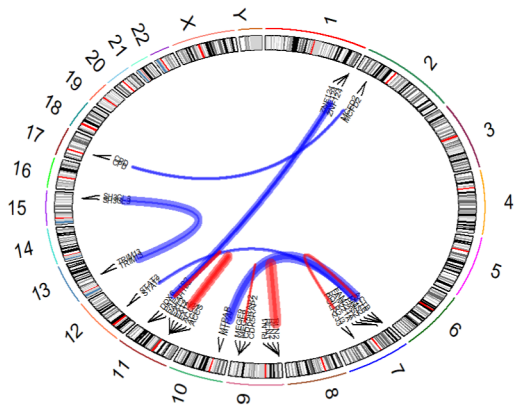


Figure 2: Union of RNA fusions called by two different algorithms (StarFusion and FusionCatcher) represented by CIRCOS plot in case #1. Early represents those found in the early HGSOC, Late those in the late recurrent HGSOC, and Acquired are those that are found only in the late recurrent HGSOC. Blue represents intergenic fusions, while red are intragenic.

We next developed an approach to identify the most appropriate fusions for further study. All of the acquired fusions i.e. those found only in the late recurrent HGSOC, were filtered for a

number of parameters including number of sequencing reads which span the fusion breakpoint. These were also examined for their potential to be a driver oncogenic fusion event using OncoFuse (score >0.75) and examined for their presence in ovarian cancer cell lines. Figure 3 shows a CIROCOS plot of all high confidence HGSOC recurrent fusions and the top 4 candidate fusions which have high confidence reads, a high OncoFuse score, and are present in one or more HGSOC cell lines. In the next year we will examine the expression and biological function of these fusions.



- **RAB9A – OFD1**
  - In OVSAHO, JHOC-5, JHOM-2B, JHO4
- **PRPF3—RPRD2**
  - In OVISE, SK-OV3
- **DBT – LRRC39**
  - In JHOM-1
- **NUDT21 – AMFR**
  - JHOM-1

Figure 3: High confidence RNA fusions in recurrent HGSOC. Four fusions are listed and the HGSOC cell lines which show the same fusion are listed.

h) Analysis of gene expression using STAR and Deseq

Completed. As detailed in the Year 1 report we performed unsupervised hierarchical clustering and found that nine patient-matched pairs of HGSOC clustered in the same doublet clade of their patient-matched primary, suggesting a profound transcriptional conservation between the recurrence and the early lesion. Differential expression analyses revealed heterogeneous expression between the patient-matched samples, only uncovering 39 differentially expressed genes. The most significantly upregulated gene in late ovarian cancer was NTRK2, showing upregulation in the majority of recurrences. Since resistance mechanisms in advanced cancers may be mutually exclusive, and thus would be missed by conventional differential expression analyses, we performed analysis focusing on outlier expression gains and losses—particularly in genes that are clinically actionable. Four clinically actionable genes showed outlier increases in at least one-third of late disease samples versus their matched early disease lesion—INHBA, IGF1 NTRK2 and EPHA3.

i) Subtask 2: Validation of candidate fusions using RT-PCR and Q-RT-PCR

In the year 1 report we showed validation of some initial fusions – however many of these didn't show high OnocFuse scores and thus may not be drivers of disease and were not in cell lines. We will thus now focus on the 4 fusions noted in Figure 3 and validation is underway.

J: Subtask 3 Validation of a select number of fusions by FISH

This has not yet been started due to the technical challenges of FISH and our desire to test fusions which have clinical importance.

Specific Aim 2) Establish the prevalence and clinical significance of identified fusion genes

Major Task 3 Isolate RNA and measure fusion using Nanostring

k) Subtask 1 Procure 60 FFPE recurrent samples and 200 primary HGSOC

This is in progress. The identification of the HGSOC for use in this project has proven challenging and taken more time than expected. We identified 1,006 cases of HGSOC treated at our hospital, and attempted to extract clinical information on them. However, while primary treatment and therapy is well documented in our cancer registry, this wasn't the case for recurrence.

As an alternative source of chemotherapy naïve and treated HGSOC we examined patients treated with neo-adjuvant chemotherapy. Importantly, we identified 174 cases which received neoadjuvant chemotherapy and thus would allow us to examine changes in fusion expression before and after neo-adjuvant chemotherapy. These samples were examined by a pathologist and scored with regards to the quality and size of the specimen. We are currently requesting these tissue for analysis.

l) Subtask 2 Isolate and measure RNA from 260 samples

This has not yet been started as we don't have the samples, however they have already been examined by a pathologist and we expect to have them shortly.

m) Subtask 3 Develop NanoString codeset based upon fusions from Aim 14) other achievements. Include a discussion of stated goals not met.

This has not yet been started as we are finalizing the prioritized list of fusions to analyze. We would like to note that since writing the original proposal, the cost of sequencing has dropped significantly, yet the costs for NanoString have remained constant. We may therefore use targeted RNA sequencing for this approach.

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

Nolan Priedigkeit participated in the year 1 study as part of his PhD thesis in the Lee laboratory. He graduated and received his PhD.

o) How were the results disseminated to communities of interest?

We presented our work as a poster at the AACR Addressing Critical Questions in Ovarian Cancer Research and Treatment, October 1 - 4, 2017, Wyndham Grand Pittsburgh. We received strong positive feedback.

p) Describe how the results were disseminated to communities of interest.

We presented our work as a poster at the AACR Addressing Critical Questions in Ovarian Cancer Research and Treatment, October 1 - 4, 2017, Wyndham Grand Pittsburgh.

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

We will obtain the next batch of clinical samples and continue study of the fusions and thus complete the goals and tasks for specific aim 2.

r) Describe briefly what you plan to do during the next reporting period to accomplish the goals and objectives.

Specific Aim 2) Establish the prevalence and clinical significance of identified fusion genes

Major Task 3 Isolate RNA and measure fusions

Major Task 4 Bioinformatic analysis of data

Major Task 5 Development of HGSOC cell lines with knockdown or overexpression of fusion genes

Major Task 6 Examine the phenotype of HGSOC cell lines with and without gene fusions

4) IMPACT:

What was the impact on the development of the principal discipline(s) of the project?

We are the first to identify RNA fusions in chemotherapy naïve and resistant HGSOC, and the finding of acquired fusions in all of the cases suggests there maybe biological drivers of recurrence. This is a novel and new finding. The next period of work will be key to decipher which of these RNA fusions is key to HGSOC. The gene expression findings are also novel and may highlight several new therapeutic targets which once reported will all be publicly available and can be studied by others.

Describe how findings, results, techniques that were developed or extended, or other products from the project made an impact or are likely to make an impact on the base of knowledge, theory, and research in the principal disciplinary field(s) of the project.

What was the impact on other disciplines?

Nothing to report

Describe how the findings, results, or techniques that were developed or improved, or other products from the project made an impact or are likely to make an impact on other disciplines.

We are using new methods for RNA fusion discovery, validation, and then driver prediction which will add to others work and help guide this field forward. For example, the use of RNA fusion driver prediction algorithms is nascent, and our prediction and then functional validation will deliver data and help improve these algorithms and approaches.

What was the impact on technology transfer?

Nothing to Report

Describe ways in which the project made an impact, or is likely to make an impact, on commercial technology or public use, including:

Nothing to Report

Describe how results from the project made an impact, or are likely to make an impact, beyond the bounds of science, engineering, and the academic world on areas such as:

Nothing to report

5) CHANGES/PROBLEMS:

Describe any changes in approach during the reporting period and reasons for these changes

We faced a challenge obtaining a validation set of naïve and chemotherapy resistant HGSOC samples. To overcome this we will focus on neoadjuvant treatment, where each patient is guaranteed to have a pre-biopsy and subsequent resection. Utilizing this approach we have identified a suitable cohort which has undergone pathology review and we are awaiting samples.

Since writing the proposal, the costs of sequencing have dropped considerably, whereas the costs for NanoString have remained constant (actually increased with inflation). Furthermore, the advantage of NanoString for FFPE samples is decreased as we have recently reported outstanding performance of capture exome RNA sequencing on breast cancer FFPE specimens (PMID: 28878133). We may therefore now use RNA sequencing instead of NanoString for our validation study.

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

We had a delay in the procurement of the second set of human HGSOC samples. The identification and procurement of human samples is always a challenge. Samples must be identified and then clinical characteristics identified to make sure that the correct samples are procured. However, we now have the neoadjuvant patient cohort identified, and pathology review has been completed. There is a possibility that some cases will not have available tissue, and if this is the case we may need to approach other collaborators to expand this cohort.

Describe problems or delays encountered during the reporting period and actions or plans to resolve them.

Delays in identifying and procuring tissue as noted above. We now have the patients identified and do not expect further issues.

Changes that had a significant impact on expenditures

Nothing to report

Describe changes during the reporting period that may have had a significant impact on expenditures, for example, delays in hiring staff or favorable developments that enable meeting objectives at less cost than anticipated.

Nothing to report

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents

Nothing to report

Describe significant deviations, unexpected outcomes, or changes in approved protocols for the use or care of human subjects, vertebrate animals, biohazards, and/or select agents during the reporting period.

Nothing to report

6) PRODUCTS:

Publications, conference papers, and presentations

We presented our work as a poster at the AACR Addressing Critical Questions in Ovarian Cancer Research and Treatment, October 1 - 4, 2017, Wyndham Grand Pittsburgh.

B55 Recurrent transcriptional remodeling events and acquired fusion RNAs in relapsed ovarian cancers. Nolan Priedigkeit, University of Pittsburgh, Pittsburgh, PA, United States.

https://www.aacr.org/Documents/OVA17_Poster_B_for_Web.pdf

Website(s) or other Internet site(s)

Nothing to report

Technologies or techniques

Nothing to report

Inventions, patent applications, and/or licenses

Nothing to report

Other Products

Nothing to report

7) PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?

Provide the following information for: (1) PDs/PIs; and (2) each person who has worked at least one person month per year on the project during the reporting period, regardless of the source of compensation (a person month equals approximately 160 hours of effort). If information is unchanged from a previous submission, provide the name only and indicate "no change."

Name: Adrian V. Lee, Ph.D. – no change

Name: Peter Lucas – no change

Name: Robert Edwards – no change

Name: George Tseng – no change

Name: Nick Smith left the laboratory and was replaced with Zeynep Erdogan. She is an equivalent trained employee who is performing computational analysis on fusion RNAs.

Name: Li Zhu - no change

Name: Nolan Priedigkeit graduated with his MD/PhD and left the project. He was replaced by John Willis who is a similar trained MD/PhD student who is working on the project. John's salary is currently supported by the graduate program of the University of Pittsburgh and no funds have been required from this grant.

What other organizations were involved as partners?

Organization Name: Roswell Park Cancer Institute (RPCI)

Location of Organization: Buffalo, NY

Partner's contribution to the project (identify one or more)

Financial support; N/A

In-kind support (e.g., partner makes software, computers, equipment, etc., available to project staff); N/A

Facilities (e.g., project staff use the partner's facilities for project activities); N/A

Collaboration (e.g., partner's staff work with project staff on the project); RPCI will analyze the RNAseq data in collaboration following transfer under MTA/DUA. This work is in-kind and no financial support is provided.

Personnel exchanges (e.g., project staff and/or partner's staff use each other's facilities, work at each other's site); N/A

8) SPECIAL REPORTING REQUIREMENTS

Nothing to report

9) APPENDICES

Nothing to report