AWARD NUMBER: W81XWH-18-1-0751

TITLE: "Identification of novel genetic causes of lung cancer by sequencing the germline exome of never smoker young lung cancer patients".

PRINCIPAL INVESTIGATOR:

PI: Dr. Zoltan Szallasi,

CONTRACTING ORGANIZATION:

Children's Hospital, Boston

REPORT DATE: October 2019

TYPE OF REPORT: Annual progress 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
Oct 2019	Annual	09/30/18-09/29/19

4. TITLE AND SUBTITLE
"Identification of novel genetic causes of lung cancer by sequencing the germline exome of never smoker young lung cancer patients".

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-18-1-0751

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Zoltan Szallasi MD

E-Mail:Zoltan.szallasi@childrens.harvard.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

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

Children's Hospital, Boston
Massachusetts, 02215

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 Treatment efficacy and survival in cancer has been constantly improving over the past several decades for children and for adults over forty years of age. For the young adult population, however, no progress has been made. Identification of genetic factors contributing to the development of cancer may help to identify individuals at high risk and thus provide rational strategy for prevention and through understanding the underlying biology it can also help to devise novel therapeutic strategies. We will sequence the germline DNA of lung cancer patients that are under 35 years of age, who never smoked and had no apparent exposure to toxic chemicals. In such a population, genetic factors will play a more prominent role than in an older population more affected by environmental factors, therefore, those genetic factors could be identified more easily.

15. SUBJECT TERMS
Germline sequencing, oncogenes, lung cancer in young non-smoker adults, genetic causes of lung cancer

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		19b. TELEPHONE NUMBER *(include area code)*

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

TABLE OF CONTENTS

Introduction:

Treatment efficacy and survival in cancer has been constantly improving over the past several decades for children and for adults over forty years of age. For the young adult population, however, no progress has been made. Identification of genetic factors contributing to the development of cancer may help to identify individuals at high risk and thus provide rational strategy for prevention and through understanding the underlying biology it can also help to devise novel therapeutic strategies. Earlier efforts to identify genetic factors contributing to lung cancer probably missed several relevant mutations since they have been limited in many ways that we are aiming to overcome in this proposal. We will sequence the germline DNA of lung cancer patients that are under 35 years of age, who never smoked and had no apparent exposure to toxic chemicals. In such a population, genetic factors will play a more prominent role than in an older population more affected by environmental factors, therefore, those genetic factors could be identified more easily.

Keywords:

Germline sequencing, oncogenes, lung cancer in young non-smoker adults, genetic causes of lung cancer

Accomplishments:

Specific Aim 1: Whole exome sequencing of the germline DNA of young, (<35 years of age) non-smoker lung cancer patients

Major Task 1: Whole exome sequencing of the germline DNA of 40 young non-smoker lung cancer patients

Subtask 1: Finalizing and obtaining all relevant IRB approvals

Subtask 2: Whole exome sequencing of the germline DNA young non-smoker lung cancer patients – 50 cases

1) **Major activities:**
   - We have finalized and obtained all necessary IRB approvals for the project.
   - We have characterized the cohort for the most relevant biomarkers such as ras mutations, EGFR mutation, ALK translocation.
   - We have isolated germline DNA from the first 30 cases, those were processed by next generation sequencing.
   - We received the first batch of sequencing data. Bioinformatics processing of the data is currently underway by standard tools. The quality control of the raw sequencing data has been completed. Mapping the quality-controlled sequencing reads to the reference genome is underway by using the Burrows-Wheeler Aligner (BWA).

2) **Specific objectives:** The main objective of the first part of the project is to perform a next generation sequencing on the germline DNA and identify germline genomic variants enriched in this cohort.

3) **Significant results**: Histological analysis showed that with one exception, all 35 cases sequenced so far were lung adenocarcinoma cases. We have obtained somatic molecular biology markers for 22 cases of the 35 germline cases sequenced so far. Of the 22 cases 13 had ALK translocation and 4 had EGFR mutation. This represents a higher than 50% frequency of ALK positivity which is much higher than the ~5% frequency seen in the average population. This strongly suggests that our assembled cohort has unusual enrichment for certain molecular mechanism the germline cause of which we aim to discover in this work.

Summary of tumor properties from patients whose germline DNA has been sequenced.

ID	histology	stage	Molecular drivers	brain metastasis
YLB17	ADC			
YLB18	ADC	T2 N1 M0		
YLB01	ADC	T1a N0 M0	no ras, ALK or EGFR mutation	
YLB39	ADC	T2 N2 M1b (brain, bone, metastasis)	ALK translocation	YES
YLB02	ADC	T1a N0 M0		
SYLB01	ADC	T1 N3 M1 (brain met)	ALK translocation	YES
YLB23	ADC	T4 N0 M0	no ras, ALK or EGFR mutation	
YLB20	ADC			
YLB02	ADC	M1 (bone)	ALK translocation	
YLB03	ADC			
YLB22	ADC		ALK translocation	
YLB04	ADC	T3NxM1 (pleura)	EGFR (deletion in Exon 19)	YES
YLB19	ADC	T4 N2 M1a (lung metastasis)		
YLB28	ADC			
YLB38	ADC		ALK translocation	
YLB05	ADC	M1 (intrapulm.met.)	no ras, ALK or EGFR mutation	
YLB29	ADC		ALK translocation	
YLB08	ADC			
YLB21	ADC	T1bN0M0		
YLB35	ADC		ALK translocation	
YLB09	ADC	T2N3M1	ALK translocation	
YLB14	ADC		EGFR (deletion in Exon 19)	
YLB10	ADC	T4-N3-M1b (bone, liver metastasis)	EGFR (deletion in Exon 19)	
YLB07	ADC	pT2aN0M0		
YLB11	ADC	M1 (pleural metastasis)	ALK translocation	

YLB16	ADC	T2a-N0-M0	no ras, ALK or EGFR mutation	
YLB15	ADC		ALK translocation	
YLB12	ADC			
YLB13	ADC	M1 (pleural metastasis)	ALK translocation	
YLB24	ADC	T4N2Mab	ALK translocation	
YLB25	ADC	T4NxM1b	EGFR (deletion in Exon 19)	
YLB26	ADC	T2N2M0	ALK translocation	
YLB27	ADC	T3N0M1a		
YLB36	squamous	T4N2M0		
YLB37	ADC	T4NxM1		

What opportunities for training and professional development has the project provided? Nothing to report.

How were the results disseminated to communities of interest? Nothing to report.

What do you plan to do during the next reporting period to accomplish the goals? We will complete the bioinformatic analysis of the first thirty cases and we will start processing the next twenty cases for next generation sequencing during next few months. Once we complete the analysis and identify germline genomics variants enriched in this cohort, we will test those on our FFPE validation cohort (fifty cases) with targeted sequencing and we will also test the associated somatic genomic aberrations (loss of heterozygosity etc.).

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

What was the impact on other disciplines? Nothing to report.
What was the impact on technology transfer? Nothing to report.
What was the impact on society beyond science and technology? Nothing to report

CHANGES/PROBLEMS:
Changes in approach and reasons for change: There are no changes.

Actual or anticipated problems or delays and actions or plans to resolve them:
We have encountered several issues that caused delays with the project.
It took about 6 weeks longer to obtain IRB approvals than anticipated. We initiated DNA extraction afterwards. We sent out the first test batch for sequencing, but the quality of the DNA in the first batch did not meet the quality criteria of the sequencing facility for all cases and we needed to isolate another batch of DNA from the

less than appropriate samples. We also had some issues with personnel (one of the associates helping with handling of the samples delivered her child and we needed to find a replacement). All these together caused and approximately 6-7 months delay with the project. Therefore, we requested and obtained a 1-year, no cost extension. With this 6-months delay accounted for, the project is on schedule and we do not anticipate any further delays or changes.

Changes that had a significant impact on expenditures: Nothing to report.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents: Nothing to report.

PRODUCTS:

Journal publications: Nothing to report

Books or other non-periodical, one-time publications: Nothing to report

Other publications, conference papers and presentations: Nothing to report

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

PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

Name: Zoltan Szallasi, MD

Project Role: PI

Researcher Identifier (e.g. ORCID ID): 0000-0001-5395-7509

Nearest person month worked: 2 months

Contribution to Project: Dr. Szallasi provided directions for the project and supervise and perform some of the bioinformatics analysis

Funding Support:

Name: Viktoria Tisza, PhD

Project Role: research associate

Researcher Identifier (e.g. ORCID ID):

Nearest person month worked: 4 months

Contribution to Project: Dr Tisza is helping with processing the samples for next generation sequencing and she also set up the validation experiments.

Funding Support: