AWARD NUMBER: W81XWH-18-1-0499

TITLE: Rare Variants in Systemic Sclerosis

PRINCIPAL INVESTIGATOR: Brendan Lee

CONTRACTING ORGANIZATION: Baylor College of Medicine

REPORT DATE: September 2020

TYPE OF REPORT: Annual Technical Report

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

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
September 2020	ANNUAL	August 15, 2019 - August 14, 2020

4. TITLE AND SUBTITLE
RARE VARIANTS IN SYSTEMIC SCLEROSIS

5a. CONTRACT NUMBER
W81XWH-18-1-0499

5b. GRANT NUMBER
GRANT12507060

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Brendan Lee, MD, PhD

E-Mail: blee@bcm.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Baylor College of Medicine
One Baylor Plaza
Houston, Texas 77030-3411

8. PERFORMING ORGANIZATION REPORT NUMBER

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

U.S. Army Medical Research and Development 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

Systemic Sclerosis (SSc, scleroderma) has considerable morbidity and the highest mortality of all the systemic autoimmune diseases. The cause(s) of SSc is/are unknown and it is generally considered to result from a combination of external triggers operating in the context of genetic susceptibility, similar to other complex genetic autoimmune diseases. The genetic component has been estimated to contribute 30% to the risk of developing SSc. The overall goal of this project is to identify rare genetic variants, that increase susceptibility to SSc and that influence clinical outcomes. Metabolic pathways influenced by these variants will provide insight into pathogenetic mechanisms of this and other fibrotic diseases, leading to novel therapeutic approaches. This study is on track for completion of the largest whole genome sequencing (WGS) project in scleroderma. The samples have been sequenced and are currently been analyzed for genotype/phenotype correlations.

15. SUBJECT TERMS
Systemic Sclerosis, Scleroderma, Whole Genome Sequencing (WGS), RNA sequencing, genetic variants, rare variants.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRDC
Unclassified	Unclassified	Unclassified	Unclassified	12	19b. TELEPHONE NUMBER *(include area code)*

TABLE OF CONTENTS

1. **INTRODUCTION:** The subject/topic area of this research is Systemic Sclerosis (SSc, Scleroderma). The purpose of the research is to identify genetic variants that contribute to SSc disease susceptibility and influence outcome. The approach involves whole genome sequencing of 100 trios (300 individuals including affected case and both parents). Previous Genome-Wide-Association-Studies (GWAS) have identified gene regions that are associated with disease but the majority of these are in non-coding areas so the impact of these variants is unclear. This study will identify rare variants (both inherited and de novo mutations) and will analyze these mutations according to the role they likely plan in disease pathogenesis. The immediate outcome of this project will be identification of the causal variants in multiple pathways associated with SSc susceptibility with the long-range impact will be the identification of the role these variants plan in disease causation and severity/outcome.
2. **KEYWORDS:** Systemic Sclerosis, Scleroderma, Whole Genome Sequencing (WGS), genetic variants, rare variants.
3. **OVERALL PROJECT SUMMARY:** This study is on track for completion of the largest whole genome sequencing (WGS) project in scleroderma. The samples have completed sequencing and analysis is currently underway for genotype/phenotype correlations.

4. **ACCOMPLISHMENTS:**
    - **What were the major goals of the project?**
        - Task 1: Institutional Review Board (IRB) and DOD Human Research Protection Office (HRPO) approval. Timeline months 1-6 (15 Aug 2018 – 14 Jan 2019); actually HRPO approval was received later than expected on 11 Jul 2019 due to a miscommunication– but this has been resolved..
            - UT Houston IRB approval was obtained previously under the general approval for the genetics studies in SSc; a notice was sent to the IRB and accepted as a defined sub-study with the most recent annual reviews up-to-date as of 22 Jan 2019; so **completed**.
            - Baylor College of Medicine: Whole genome sequencing will catalog the largest amount of genetic variation in scleroderma cases and will provide information on noncoding and rare variation. The incorporation of whole genome generated information in analyses of complex traits is an emerging tool for defining underlying genetic underpinnings. The sequencing and processing will be completed at BGI and data analysis in the lab of Dr Brendan Lee. The generated data will be the basis of the proposed work.
            - HRPO approval was received 11 Jul 2019 (HRPO Log Number E00551.1a) so **completed**.
        - Task 2: Prioritization, preparation and distribution of samples to be sent to BCM for genotyping: Timeline months 7- 8 (15 Feb 2019 – 14 Apr 2019)

        - Samples were identified on the basis of protocol-defined criteria and prepared for shipping during this time and were subsequently sent to BCM. So **completed.**
    - Task 3: Sequencing of 300 samples with appropriate quality control measures: Timeline months 9-18 (15 Apr 2019 – 14 Feb 2020)
        - All samples were sent to BCM for whole genome sequencing; of the samples sent 2 failed quality control measures and new replacement aliquots were prepared and sent to BCM on 19 Aug 2019. Six more samples were sent to BCM to expand the cohort for whole genome sequencing**.** All the 306 samples were successfully sequenced. Data generated were uploaded into BCM computational cluster for analysis downstream. So **completed.**
    - Task 4: Processing, including imputation using the generated sequence data. Timeline months 19-21 (15 Feb 2020 – 14 May 2020).

      All the sequencing data of the 306 samples were processed through internal bioinformatics pipeline for variants calling according to given family structure. VCF files were generated for further downstream analysis. So **completed.**
    - Task 5: Data analysis and prioritization of candidates for validation in other, future cohorts. Timeline months 22-36 (15 May 2020 – 14 Aug 2021).

      Candidate genes/variants list now has been compiling and validating. Additional cohort has been adopted and under investigation. **Ongoing.**
    - Task 6: Association analysis of the most likely identified variants with clinical disease features. Timeline months 34-36 (15 May 2021 – 14 Aug 2021).

      Genome-wide Association Analysis (GWAS) of the non-trio Scleroderma Registry data is currently underway and will serve as a validation data for rare variant analysis. RV-TDT analysis will follow upon identification of candidate/rare variants predicted to effect protein structure (missense, nonsense and splicing variants). **Ongoing.**
    - Task 7: Preparation of manuscripts for publication. Timeline months 34-36 (15 May 2021– 14 Aug 2021). **Awaiting results** from Task 6.
    - Task 8: Quarterly meetings between the UT-H and BCM teams to coordinate all aspects of the project and review and interpret data as it becomes available. Timeline months 3-36 (15 Nov 2019 – 14 Aug 2021).**Ongoing.**
        - We have had 2 face-to-face meetings and 12 conference calls involving both teams for a total of 14 meetings since the study started. We are up-to-date on these and plan to continue to meet on a regular basis – at least quarterly or more frequently as the data become available.
- **What was accomplished under these goals?**
    - Tasks 1,2,3,4 have been completed. For Task 5 and Task 6, the analysis and interpretation phases are expected to be quite time-consuming so generating the sequence data ahead of schedule will be welcome. Task 8 (quarterly meetings of the UT-H and BCM teams) is ongoing as scheduled.

- BCM side of Task 2: Quality checking and trio family selecting on samples delivered by UT-H Timeline months 9- 10 (15 May 2019 – 14 June 2019)
    - Samples were checked for quality and quantity before final sequencing. Replacements were requested for failed samples. 100 trios were selected based on qualified criteria to initiate the sequencing. Status: **Complete**
- BCM side of Task 3: Sequencing of 300 samples. Timeline months 11-18 (15 June 2019 – 14 Feb 2020)
    - Replacement samples received to complete the 100 trios whole genome sequencing. Samples were sequenced in batches. 6 more samples were sequenced. Status: **Completed.**

- **What opportunities for training and professional development has the project provided?**
    - Nothing to report from either institution**.**
- **How were the results disseminated to communities of interest?**
    - Nothing to report from either institution**.**
- **What do you plan to do during the next reporting period to accomplish the goals?**
    - We fully anticipate that the candidate genes/variants analysis will be completed. (Tasks 5 and 6 anticipated completion 14 Aug 2021).
    - Task 5 evaluation in other cohorts will be completed.

5. **IMPACT:** *Describe distinctive contributions, major accomplishments, innovations, successes, or any change in practice or behavior that has come about as a result of the project relative to:*
    - **What was the impact on the development of the principal discipline(s) of the project?**
        - Nothing to Report.
        - Nothing to Report at this time.
    - **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.
6. **CHANGES/PROBLEMS:** *The Project Director/Principal Investigator (PD/PI) is reminded that the recipient organization is required to obtain prior written approval from the awarding agency Grants Officer whenever there are significant changes in the project or its direction. If not previously reported in writing, provide the following additional information or state, "Nothing to Report," if applicable:*

    - **Changes in approach and reasons for change**
        - In the upcoming grant cycle we have completed Whole Genome Sequencing ahead of schedule and will focus on completing Task 5 and Task 6, the analysis

and interpretation phases which are expected to be quite time-consuming. This will require additional personnel to be added to the project for analysis and interpretation. In addition, we plan to extend the spectrum of our analysis to include integration of multi-omic data.

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

In addition, due to the Covid-19 pandemic there were additional delays in sequencing and processing of the data; However, in spite of this we are back on track to meet our timeline goals.

- **Changes that had a significant impact on expenditures**

We are requesting approval of carryforward of more than 25% unobligated balance due to the Covid-19 pandemic. In response to the Covid-19 Pandemic, Baylor College of Medicine instituted college wide measures to help limit the spread of the virus and perform responsible conduct of research. Starting March 23, 2020 limited access to research facilities was implemented with phased increases of access as recovery efforts commenced. Following the OMB Flexibility guidelines, researchers were retained on grants during this period when they had both continuity support and direct activities in support of the grant.

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

7. **PRODUCTS:** *List any products resulting from the project during the reporting period. If there is nothing to report under a particular item, state "Nothing to Report."*
   - **Publications, conference papers, and presentations**
     - **Journal publications.** None to date.
     - **Books or other non-periodical, one-time publications.** None to date.
     - **Other publications, conference papers, and presentations.** None to date.
   - **Website(s) or other Internet site(s)** None to date.
   - **Technologies or techniques** None to date.
   - **Inventions, patent applications, and/or licenses** None to date.
   - **Other Products** None to date.
8. **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:	Maureen D. Mayes, MD, MPH (No change from submission)
Project Role:	Principal Investigator; Initiating PI
Researcher Identifier (e.g. ORCID ID):	
Nearest person month worked:	3.0 calendar months
Contribution to Project:	Dr. Mayes is responsible for the overall conduct of the study and for the timely completion of all aspects (Tasks 1 through 8); she supervises UT project personnel and organizes the weekly UT meetings, the quarterly UT-BCM meetings to review progress, potential problems, data collection and results.
Funding Support:	**New:** GSK: A multi-centre, randomized, double-blind (sponsor open), placebo-controlled, repeat dose, proof of mechanism study to evaluate the safety, tolerability, pharmacokinetics pharmacodynamics and explore efficacy of GSK2330811 in subjects with diffuse cutaneous systemic sclerosis. Galapagos: Orally administered GLPG1690 for 24 weeks in subjects with systemic sclerosis Eicos: A Multicenter, Double-Blind, Randomized, Placebo-Controlled, Phase 3 Study Evaluating the Efficacy and Safety of Intravenous Iloprost in Subjects With Raynaud's Phenomenon Secondary to Systemic Sclerosis University of Michigan: Long-term follow-up of Participants of the Phase II study to evaluate subcutaneous abatacept v. placebo in diffuse cutaneous systemic sclerosis-a double-blind, placebo-controlled, randomized trial (ASSET)

Name:	Dianna Milewicz, MD
Project Role:	Co-Investigator
Researcher Identifier (e.g. ORCID ID):	
Nearest person month worked:	0.6 calendar months
Contribution to Project:	Dr. Milewicz has a strong background in the genetic basis of vascular diseases. Her role to date is that she has provided guidance on study design thus far and in the upcoming years she will advise on the analysis and interpretation of the genetic data. She has served as an advisor on Dr. Mayes' previous scleroderma genetic studies and has worked with the Baylor College of Medicine Genetics group on multiple projects
Funding Support:	**New:** NIH/NHLBI: Novel genetic Insight into the molecular pathogenesis

	of atherosclerosis. NIH: Medical Scientist Training Program Texas Heart Institute: Fibromuscular Dysplasia Project Marfan Foundation: Asprosin's Role in Suppressed Appetite and Progeroid Appearance of Marfan Lipodystrophy Syndrome and Neonatal Marfan Syndrome Patients American Heart Association: Molecular Pathogenesis of Occlusive Cerebrovascular Disease Resulting from ACTA2 Mutations

Name:	Claudia Pedroza, PhD (no change from submission)
Project Role:	Co-Investigator, statistician
Researcher Identifier (e.g. ORCID ID):	
Nearest person month worked:	0.6 calendar months
Contribution to Project:	Dr. Pedroza has been involved in the planning and implementation of the project to date to ensure that data will be interpreted in light of available clinical outcomes.
Funding Support:	No changes

Name:	Patricia Gonzales, LVN (no change from submission)
Project Role:	Project Coordinator
Researcher Identifier (e.g. ORCID ID):	
Nearest person month worked:	3.6 calendar months
Contribution to Project:	Ms Gonzales is responsible for the day-to-day operations of the study, overseeing database queries and reporting to the investigators regarding progress, time lines and review of expenditures.
Funding Support:	

Name:	Julio Charles (no change from submission)
Project Role:	Laboratory Manager
Researcher Identifier (e.g. ORCID ID):	
Nearest person month worked:	3.0 calendar months
Contribution to Project:	Mr Charles is responsible for overseeing the selection, aliquoting and distribution of samples for genotyping and sequencing studies. He coordinates delivery to the Baylor research lab; he attends the weekly lab meetings as well as the project-specific quarterly meetings between the UT and Baylor research groups.
Funding Support:	

Name:	Hau Pham (no change from submission)
Project Role:	Research Assistant
Researcher Identifier (e.g. ORCID ID):	
Nearest person month worked:	3.0 calendar months
Contribution to Project:	Ms. Pham is responsible for the day-to-day work in the Rheumatology Research Lab, to implement sample selection, DNA quantification, DNA measurement and record keeping of samples distributed to Baylor as well as DNA extraction on new samples as needed, and autoantibody determination on these new samples.
Funding Support:	

- **Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?**
  - Dr. David Murdock, Assistant Professor, joined the project and leads the WGS sequencing interpretation team with Dr. Hongzheng Dai.
  - Dr. Monika Weisz Hubsman, Clinical Fellow, will join the project and lead the GWAS analysis of non-trio Scleroderma Registry data.
- **What other organizations were involved as partners?**

BAYLOR COLLEGE OF MEDICINE

Name:	Brendan Lee, MD, PhD (No change from submission)
Project Role:	Principal Investigator;

Researcher Identifier (e.g. ORCID ID):	
Nearest person month worked:	.30 calendar months
Contribution to Project:	Dr. Lee will be responsible for the overall implementation of the project as well as experiments in all aims. He will be responsible for the overall experimental design, data analysis, implementation, communication, and reporting of the experiments in the grant
Funding Support:	**No new support to report since submission**

Name:	Lindsay Burrage, MD, PhD (No change from submission)
Project Role:	Faculty
Researcher Identifier (e.g. ORCID ID):	
Nearest person month worked:	1.20 calendar months
Contribution to Project:	Dr. Burrage will assist with the overall implementation of the project as well as experiments in all aims. She will also assist with the overall experimental design, data analysis, implementation, communication, and reporting of the experiments in the grant
Funding Support:	**No new support to report since submission**

Name:	Hongzheng Dai, PhD
Project Role:	Research Associate
Researcher Identifier (e.g. ORCID ID):	
Nearest person month worked:	3.6 calendar months
Contribution to Project:	Dr. Dai will perform data analyses using existing publicly available software and, in addition, write programs to analyze the generated data. He will combine publicly available datasets with the generated data
Funding Support:	**No new support to report since submission**

Name:	Monika Weisz Hubsman, MD (New since last submission)
Project Role:	Clinical Fellow
Researcher Identifier (e.g. ORCID ID):	
Nearest person month worked:	6 calendar months
Contribution to Project:	Dr. Hubsman will be responsible for the Genome-wide Association Analysis (GWAS) of the non-trio Scleroderma Registry data and RV-TDT analysis.
Funding Support:	**No new support to report since submission**

Name:	David Murdock, MD (New since last submission)
Project Role:	Faculty
Researcher Identifier (e.g. ORCID ID):	
Nearest person month worked:	.60 calendar months
Contribution to Project:	Dr. Murdock will lead the sequencing interpretation team working closely with Dr. Hongzheng Dai and Dr. Lindsay Burrage. He will apply various analysis pipelines to the WGS data for variant and gene discovery.
Funding Support:	**No new support to report since submission**

9. **SPECIAL REPORTING**

REQUIREMENTS ○ **None**

APPENDICES: Not applicable