AWARD NUMBER: W81XWH-20-1-0849

TITLE: Plasma Cell-Free RNA as Non-Invasive Biomarker for Parkinson's Disease

PRINCIPAL INVESTIGATOR: Laura Ibanez

CONTRACTING ORGANIZATION: Washington University in Saint Louis

REPORT DATE: Oct 2021

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**
Oct 2021	Annual Technical Report	09/15/2020 - 09/14/2021

4. TITLE AND SUBTITLE
Plasma Cell-Free RNA as Noninvasive Biomarker for Parkinson's Disease

5a. CONTRACT NUMBER
W81XWH-20-1-0849

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Laura Ibanez

E-Mail:ibanezl@wustl.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Washington University
660 S. Euclid Ave. St. Louis
63110-1010, Missouri - US

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

Parkinson disease (PD) is the most common neurodegenerative disorder, after Alzheimer disease (AD). Many attempts have been made to find a good biomarker, including alpha-synuclein protein levels in the cerebrospinal fluid (CSF). Cell-free nucleic acids-based diagnostic tests have revolutionized prenatal screening. They have also been investigated in cancer and fetal development among other traits, including neurodegenerative diseases. We have successfully developed a preliminary predictive model for AD using cell-free plasma RNA sequencing (cfRNASeq) and machine learning techniques. We used an exploratory dataset (10 AD cases and 10 controls) to train a predictive model. We obtained an area under the ROC (AUC) of 0.84 in an independent replication dataset (10 independent AD cases and 10 controls). Moreover, this model provided similar accuracy (AUC=0.86) when tested in four preclinical AD. Using state-of-art deep neural network approaches, the accuracy increased up to 0.94. Overall, these results indicate that we can identify individuals that will progress to dementia. We think this technique can be applied to PD to generate disease-specific predictive model.

We ***hypothesize*** that there are detectable changes in the plasma free nucleic acid composition due to PD pathogenesis, even in early stages. We will use bioinformatics tools to construct a predictive model for PD, leveraging longitudinal plasma data that will allow the modeling of plasma cfRNA composition changes over the curse of the disease, thus maximizing the power of selecting informative transcripts to construct the predictive model.

We will firstly **accurately predict preclinical PD using cell-free nucleic species** in three steps: *A. Create prediction models for PD using cfRNA* generating cfRNASeq data from 200 plasma samples (50 PD individuals at 3 time-points – early pre-clinical (5-10 years before symptoms), pre-clinical (2-5 years before symptoms) and symptomatic (5-8 years after diagnostic) and 50 controls). We will use multiple analytical approaches including digital deconvolution and machine learning, feature selection and deep neural networks (similar to what we have used to generate our model for AD) to build a robust predictive model that includes the optimal number of transcripts. *B. Replication:* We will quantify the transcripts selected in A to be part of the predictive model using a more scalable and cost-effective technology such as Nanostring, Sequenom or custom transcript array to replicate the predictive model in an independent dataset (50 preclinical PD cases and 50 controls). We will also include subjects from African American and Latin ancestries (20 cases and 20 controls from each) and carriers of PD-causing mutations in *PARK1* and *PARK2* (n=20) to test the performance of the model in non-European ethnicities and in mutation carriers. *C. Specificity:* We will quantify the transcripts selected in A and replicated in B using the same scalable and cost-effective technology in 80 cases of other neurodegenerative diseases (AD, Lewy body dementia, progressive supranuclear palsy, amyotrophic lateral sclerosis and frontotemporal dementia) and additional 40 controls to test whether the predictive model is specific for PD or neurodegeneration. We expect the predictive model to be specific for PD; however, some overlap is expected due the commonalities of neurodegenerative diseases. We are currently generating longitudinal cfRNASeq data on AD individuals. With the data generated in the first aim, we will perform **integrative analyses of AD and PD** to describe biological differences and commonalities across the two most common neurodegenerative diseases. This will allow the description of biological mechanisms such as differences in genes or pathways across diseases, differences in the timeline of the disease for common genes/pathways and the improvement of the differential diagnosis.

If successful, this method could improve the cost-effectiveness of the currently available tools to diagnose and monitor PD, and provide a scalable blood-based early diagnostic screening tool.

Dr. Ibanez research interest is focused on using genetics to improve the management of individuals that suffer from neurodegenerative diseases, specially the early management by using high-throughput technologies and bioinformatics. Currently there is a great potential on using multi-omic approaches that integrate all levels of biological information. By using powerful bioinformatic tools we can combine genetic variance data with the RNA translation and the final protein production. This can be linked to diseases and be used to create predictive models that can easily be used in clinical settings to improve the management and quality of life of patients. In the last year and a half, Dr. Ibanez efforts have been focused on creating predictive models for Alzheimer Disease using different approaches. Even this model is on preliminary stages, she has proven that she can successfully create predictive models using cell-free RNA. This proposal will allow her to leverage the biology from the data generated, which could add some biological understanding to the biology of these neurodegenerative diseases.

This proposal has the potential of leading to a biomarker for early diagnosis and prognosis of PD. Moreover, the transcripts included in the prediction model will probably have biological relevance which, together with the integrative analyses with AD, might lead to potential drug targets in future studies.

15. SUBJECT TERMS

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

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

TABLE OF CONTENTS

Introduction

Parkinson's disease (PD) is the most common neurodegenerative disorder, after Alzheimer disease (AD). Many attempts have been made to find a good biomarker, including alpha-synuclein protein levels in the cerebrospinal fluid (CSF). Cell-free nucleic acids-based diagnostic tests have revolutionized prenatal screening. They have also been investigated in cancer and fetal development among other traits, including neurodegenerative diseases. We will use bioinformatics tools to construct a predictive model for PD, leveraging longitudinal plasma data that will allow the modeling of plasma cfRNA composition changes over the curse of the disease, thus maximizing the power of selecting informative transcripts to construct the predictive model.

Keywords

Parkinson's Disease, Biomarkers, cell-free RNA, Machine Learning

Accomplishments

What are the major goals of the project?

This project has four major goals:

1. Create a Predictive Model for Parkinson's Disease
2. Replicate the Predictive Model in an independent Dataset using a cost-effective Platform
3. Test the Specificity & Sensitivity of the Predictive Model
4. Conduct Integrative Analyses

What was accomplished under these goals?

1) Major activities: We have successfully generated data on 189 plasma samples.

2) Specific objectives: The goals for the first 12 months of the award were to perform Sample Selection, Protocol optimization, Sample Extraction, RNA QC and Sequencing, Sequencing Processing and QC and, create the predictive model. Additionally, we also planned to select an alternative platform to quantify the genes included in the predictive model and start to design the assay.

3) Significant results or key outcomes:

Sample Selection: We have selected 99 PD cases from the Washington University Movement Disorders Clinic and 90 control individuals from Washington University Movement Disorder Clinic and the Knight-ADRC (Table 1). The PD individuals included have been divided into two categories, early PD (less than five years after onset – N=45) and advanced PD (more than five years after onset – N=54).

Protocol Optimization: To generate the preliminary data we used a labor-intensive extraction. To improve the yields of RNA, reduce the original input volume from 1ml to 500µL and, the reproducibility of the extractions we optimized the use of the Maxwell RSC Instrument. The Maxwell equipment function with preloaded protocols and kits that are specific to sample types and nucleic acid. During the preliminary data generation, we extracted the cfRNA using a kit that extracts nucleic acids (both DNA and RNA) followed by DNAseI digestion to eliminate the genomic DNA contamination (QIAamp Circulating Nucleic Acid Kit – Qiagen – 1017647). This kit allows for

large volumes of plasma, but is time consuming and its scalability is limited. After conversations with the support consultants at Promega (Maxwell provider), we were suggested two kits; the viral RSC total nucleic acid purification kit (AS1330) followed by a DNAseI digestion and the Maxwell RSC miRNA from tissue and plasma or serum (AS1460).

- *Viral RSC total nucleic acid purification kit with protocol modification*: This protocol has a first incubation step (10min at 60°C) that aims to lyse the viruses. Since the potential viruses present in the plasma samples are of no interest to this project, the first incubation was skipped to preserve the integrity of the RNA. We tested three samples from healthy donors that were stored in the freezer for two years. We try to avoid using fresh samples for optimization tests, when possible, because they do not reflect the reality of the biobank samples that could have been stored for up to 24 years. The results show that this kit extracts both DNA and RNA, however, it has a clear preference for DNA. The yield of DNA is ten times greater to that of RNA. The use of DNAseI removes the DNA population, but not completely. It probably need longer incubations or higher concentrations, which might affect the integrity of the RNA. On top of that, this protocol did not improve the yield of RNA when we compare it to the extraction used to generate the preliminary data (Qiagen).

- *Maxwell RSC miRNA from plasma or serum*: This kit extracts all the RNA present in the plasma, from miRNA to mRNA. We extracted the RNA using the Maxwell kit and the Qiagen kit in 12 samples form healthy donors that were stored in the freezer for two months and compared the results. The yield were similar but not correlated, low yield samples were not consistent. After library preparation and sequencing I observed that the samples that the alignment parameters were very similar in both extraction methods. However, the proportion of ribosomal RNA was significantly higher in the Maxwell samples. We proceeded to plan another experiment with six samples that were extracted with the Maxwell equipment only. Half of the elution volume was ribodepleted before library preparation, and the other half went straight to library preparation. After sequencing, we compared the alignment. The ribosomal proportion of the ribodepleted samples was decreased to almost zero, but also the sequences that aligned to the coding regions of the genome (Figure 1). All samples had a large proportion of adapters dimers (Figure 1 – pink box), which might indicate that the performance of library preparation was not optimal. On our last experiment, we extracted RNA from six samples and ribodepleted half of the elution volume. We then proceeded to generate libraries for the 12 samples (6 non-ribodepleted and 6 ribodepleted) and cleaned them from the adaptor dimer. After sequencing, we observed a highly significant improvement regarding the proportion of coding RNA (Figure 2 – blue box). Coding RNA proportions increased in all samples with adaptor dimer removal, but especially on those that were ribodepleted. Even though the removal of the adapters dimers was not complete, we improved the overall quality of the data. In conclusion, we have optimized the protocol to extract and sequence cfRNA from plasma samples minimizing the hands-on time and maximizing the number of coding RNA sequenced.

Table 1. Summary Demographics of the processed samples

Group	N	Sex (Female %)	Age at draw (mean age)	Age at Onset (mean age)	Disease Duration (mean years)
Controls	90	61%	74.43	-	-
Early PD	45	43%	69.58	66.29	3.29
Advanced PD	54	33%	73.91	63.13	10.71

RNA Extraction, QC and Sequencing: Once the protocols for plasma cfRNA extraction and sequencing were optimized, we proceeded directly to process all the samples selected. We have extracted, ribodepleted, and generated the libraries for a total of 189 plasma samples (Table 1).

We targeted 40 million of single end reads from each library. The Genome Technology Access Center (GTAC), the sequencing facility at Washington University performed the final quality control for each library and generated the reads using a NovaSeq6000. They distributed the samples across 10 pools with mixed samples from other projects. This is normal practice.

Figure 1. Library sequencing quality control. RNA was obtained using the Maxwell RSC miRNA from plasma or serum kit. Each line represents a sample. Samples labeled with the "ribo" tag underwent the ribodepletion protocol.

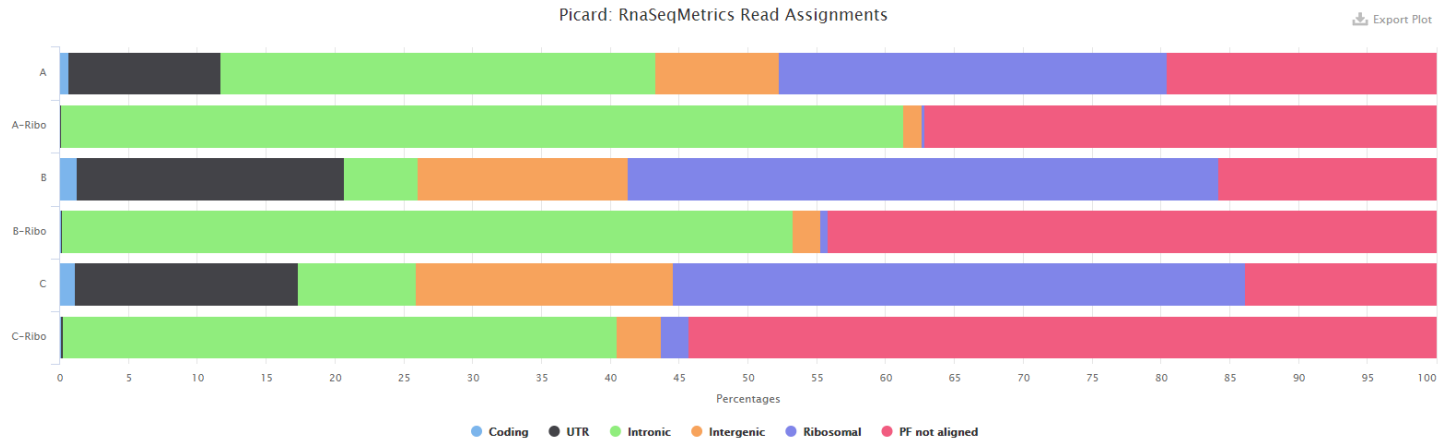
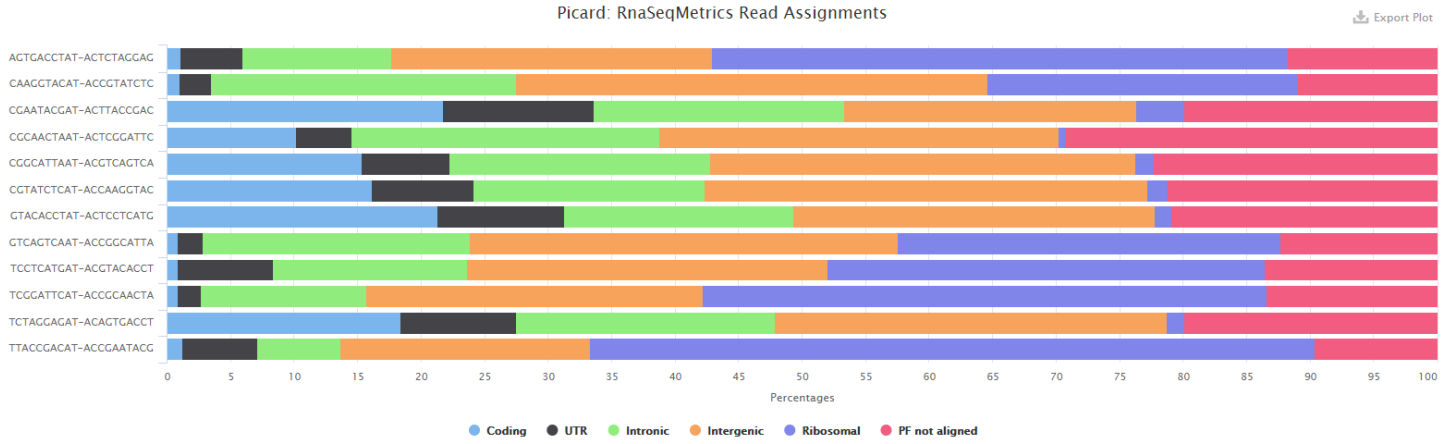


Figure 2. Library sequencing quality control. RNA was obtained using the Maxwell RSC miRNA from plasma or serum kit. Each of the rows represents a sample. Half of the samples were ribodepleted (rows: 3,4,5,6,7 and 11). After pooling and prior to sequencing, adapter dimers were removed with magnetic beads in all of the samples.



<u>Sequencing Processing and Quality Control:</u> We then proceeded to perform quality control of the raw reads with FastQC, STAR and Picard tools. We first evaluated differences in quality of the reads based on pools. We detected that the sample located in Pool 1 had a very low number of reads compared to the rest of the pools (Figure 3; red box). However, when we explored the number of usable reads for the analyses (corresponding to the sum of coding and UTR boxes from Figures 1 and 2); no significant differences were observed (Figure 4).

We then proceeded to the quantify the genes using salmon. All transcripts or genes with less than 10 reads in more than 90% of the individuals were removed and considered unreliable. We then normalized the raw counts for library complexity followed by log transformation using DeSeq2. Finally, we performed principal component analysis to identify outliers and main drivers of the variance within these samples (Figure 5). To identify the source of variation we performed correlation analyses between principal components number one and two and technical and phenotypical variables (Figure 6). Principal Component 1 is clearly associated to case/control

status and the number of reads. When we subset samples with similar number of reads, the association between case/control remains but not with the numbers of reads, suggesting that the association with status is true.

Figure 3. Number of reads distribution for each sequencing pool

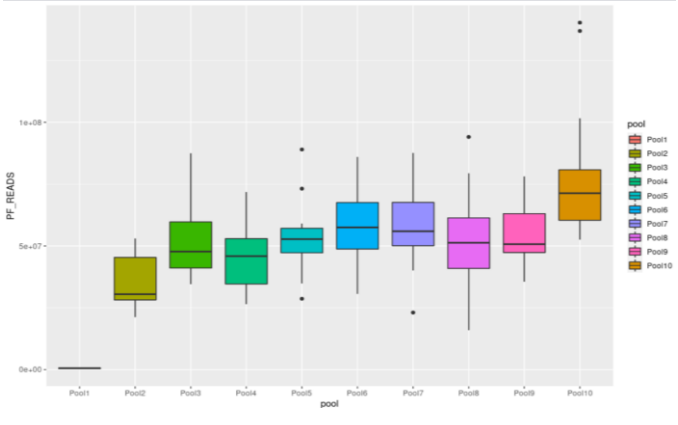


Figure 3. Number of reads distribution for each sequencing pool

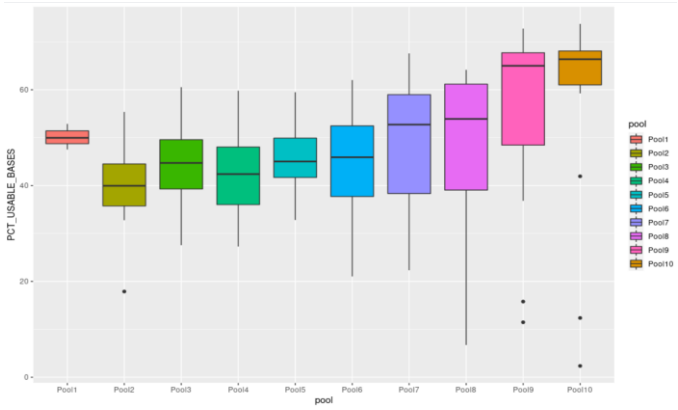


Figure 5. Principal Components from the top 500 more variable genes colored by case control status.

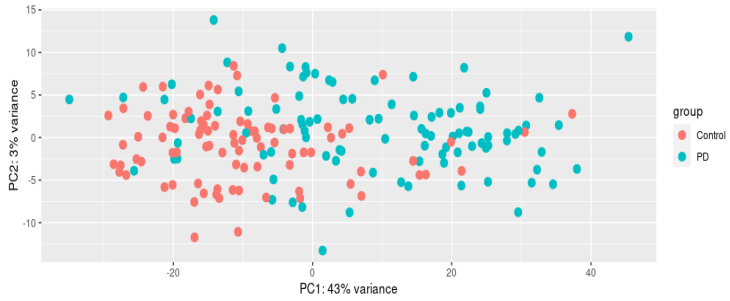


Figure 7. Number of reads distribution for each sequencing pool

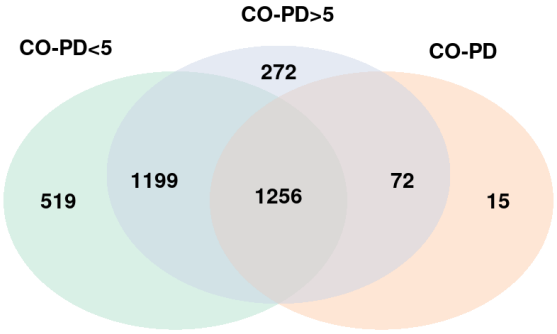
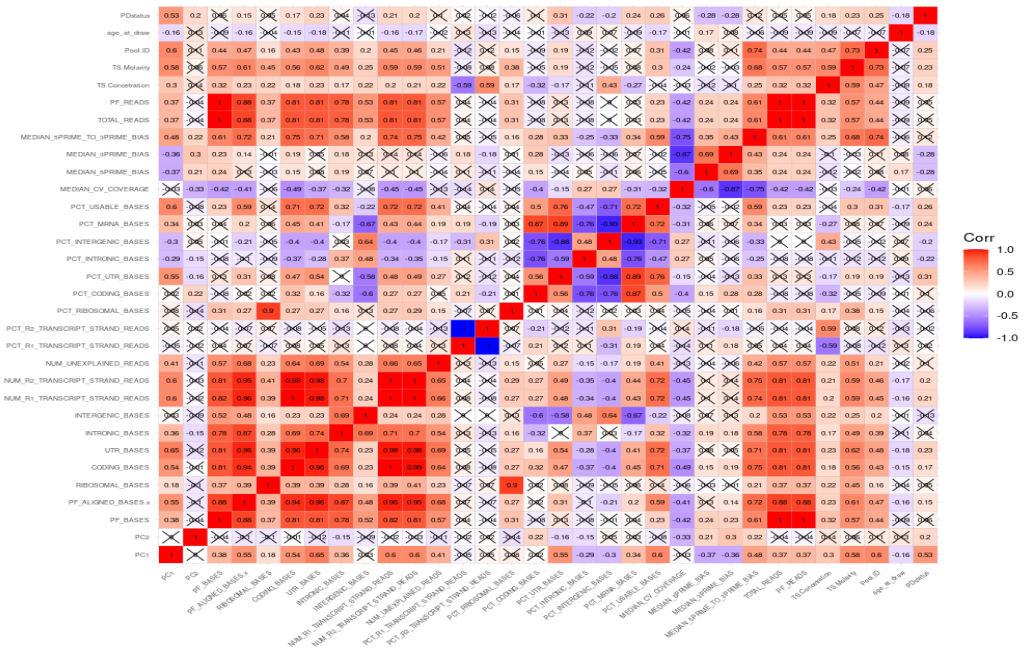


Figure 6. Correlation between technical and phenotypic variables to understand the drivers of PC1 and adjust the model accordingly.



Differential Expression Analyses: We have identified 1,749 upregulated genes and 1,050 down-regulated that pass Bonferroni correction when comparing all PD cases to controls without adjusting the model. When comparing early PD with controls, the number of up-regulated genes increased to 1,806 and the down-regulated to 1,168. When comparing late PD with Controls, the number of up-regulated genes was reduced to 893 and the down-regulated to 450. The overlap between these comparisons is highly significant; 1,256 genes are commonly dysregulated in early and late PD, whereas 519 and 272 seem to be unique to early and late PD respectively (Figure 7). We are currently testing logistic models to adjust by age, sex and other phenotypic and technical variables to minimize the number of false positive results.

Creation of the Predictive Model: Due to all the optimization of the RNA extraction, library preparation and sequencing combined with the limited access to the laboratory due to COVID-19, we are currently generating predictive models.

Alternative Platform Selection: We still do not have the predictive model, so we cannot design the final assay. From our project on Alzheimer's disease we are targeting about 25 genes, thus we will most likely use real-time PCR as alternative assay to quantify the genes that will be included in the predictive model. However, experts on the field have suggested to generate transcriptome data to leverage all the potential from the samples due to the high value of the samples that we are using.

4) Other achievements. Include a discussion of stated goals not met.

Other Achievements: We are performing similar analyses with Alzheimer's disease samples. We have generated three possible predictive models for Alzheimer's disease that we are going to test in an independent dataset. Together with the 189 plasma samples processed for the present award, we have generated data on 264 additional plasma samples distributed as follows:

- 123 European healthy controls
- 141 confirmed European Alzheimer disease cases (25 AD cases more than five years before symptom onset; 42 AD cases two to five years before symptom onset; 30 AD cases at the early stages of the disease (clinical dementia rating (CDR)=0.5) and, 44 to more advanced stages of the disease (CDR=1)).

Inclusion of samples from pre-symptomatic phases of the disease: In the original plan, we aimed to include two groups of pre-symptomatic PD patients (two to five years before onset, and five to eight years before onset). We identified them in the context of the Framingham Study, which has 79 participants with available plasma samples ten years prior to diagnosis and 92 participants with a plasma samples two to five years before PD diagnostic. We submitted the tissue request November 4th 2020 that was reviewed by the pertinent committees on December 2nd 2020. The committee found the proposal to have scientific merit and fills a clear need, but rejected the application due to concerns about the suitability of Framingham samples due to their long-term storage and lack of preliminary data on PD. They suggested a resubmission with preliminary data generated on PD individuals.

Creation of the Predictive Model: Due to all the optimization of the RNA extraction, library preparation and sequencing combined with the supply shortage, limited access to the laboratory and the extended waiting times due to COVID-19, we have not been able to analyze the data we have generated. Both will be discussed in the challenges section.

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

The present Early Investigator Research Award has facilitated the transition to independence of Dr. Ibanez. The award has allowed her to increase her team with a Postdoctoral Research Associated

(Dr. Chan) that replace the bioinformatics analyst (Dr. Farias) proposed in the original award. A side from the lab technician (Ms. Bergmann), the team has also gained a visiting researcher from Spain (Mr. Cisterna) that is not involved or financially cover by the present project. However, the award has facilitated the contact since Mr. Cisterna's mentor, Dr. Botia is part of the PD network.

As stated in the Researcher Development plan, Dr. Ibanez has followed the original training plan (funded by her K99/R00 award) as much as possible given the cancellations and transitions to virtual due to Covid-19. She has attended the following courses:

1- Understanding the Role of the Exposome in Brain Aging, Alzheimer's Disease and AD-Related Dementias. NIH 2-3 December 2020
2- 2021 Researcher Forum. Washington University in Saint Louis. 13-14 January 2021
3- Genomics in Medicine Seminar Series. Washington University in Saint Louis – Spring Semester 2021
4- Evidence-Based Diagnosis and Screening. University of Oxford (UK) – 8-12 February 2021
5- Clinical Research Training Center Career Development Virtual Seminar. Washington University in Saint Louis. 9 February 2021
6- Skandalaria Startup Webminar: Women in Stem. Skandalaris Center. 11 February 2021
7- Machine Learning in Genomics: Tools, Resources, Clinical Applications, and Ethics Workshop. NIH. 13-14 April 2021
8- PERCSS – RCR Ethics Workshop. Washington University in Saint Louis. 22 April 2021
9- The Biology of Genomes. Cold Spring harbor Laboratory. 11-14 May 2021
10- University of Washington Summer Institute in Statistics for Big Data – Supervised Methods for Statistical Machine Learning. 14-16 July 2021
11- University of Washington Summer Institute in Statistics for Big Data – Unsupervised Methods for Statistical Machine Learning. 19-21 July 2021

Dr. Ibanez has participated in several Conferences, seminars and invited communications:

A. Conferences:
    a. Alzheimer Drug Discovery Foundation Conference – On line October 2020
    b. Alzheimer Drug Discovery Foundation Investigators Meeting – On line October 2020
    c. Funding Opportunities in Neurodegenerative Disease Panel. Alzheimer's Drug Discovery Foundation. 17 February 2021
    d. The 15th International Conference on Alzheimer & Parkinson's Disease – Online March 9-14 2021
    e. Midwest Parkinson Congress 2021 – Online April 8-9 2021
    f. CZI Meeting 2021 – Online June 22-24 2021
    g. Alzheimer's Association International Conference – On line 26-30 July 2021
B. Seminars:
    a. Hope Center Weekly Seminars
    b. ADRC Weekly Seminars
    c. Neurogenetics and Transcriptomics Monthly Seminars
    d. Genetics Department Weekly Seminar
    e. CHARGE monthly Virtual Meetings
    f. ADSP monthly Functional Genomics Meeting
C. Invited Communications:
    a. Alzheimer Drug Discovery Foundation Investigators Meeting – On line October 9 2020
    b. Join European/World Stroke Organization Conference – On line November 7 2020 → Dr. Ibanez has received the **Young Investigator Award** at the Join European and World Stroke Organization Conference for her work in early outcomes after ischemic stroke genetics done during her time as post-doctoral research associate.
    c. Neurogenetics and Transcriptomics Seminar at Washington University – On line December 4 2020
    d. Fall 2020 LEAP program final evaluation – On line December 9 2020
    e. The 15th International Conference on Alzheimer & Parkinson's Disease – Online March 9-14 2021

f. Invited Seminar at Texas Tech University Health Sciences Center, School of Medicine – Online March 8 2021
g. Innovation and Science Seminar Series at Rosalind Franklin University School of Medicine – Online April 2 2021
h. Master in Bioinformatics Invited Talk at University of Murcia (Spain) – Online April 29 2021
i. Alzheimer Drug Discovery Foundation Investigators Meeting – On line October 12 2021

D. Upcoming Invited Communications
   a. Neurogenetics and Transcriptomics Seminar at Washington University – On line October 15 2021

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?

The main goals for the next reporting period are:

1. Finalize Aim 1: We will finalize the generation of the predictive model and the differential expression analyses. Then we will proceed with the replication in an independent dataset of pre-symptomatic individuals (Framingham Study Samples). We will also evaluate the performance of the model in an ancestry diverse population and other neurodegenerative diseases. We have already generated cfRNAseq data for several individuals with a diverse ancestry and from other diseases.
2. Aim 2: We have generated cfRNAseq data on samples from Alzheimer's Disease. We will integrate it with the data we have generated for Parkinson's disease. We will perform differential expression analyses to describe biological differences and commonalities across the two most common neurodegenerative diseases. We will also apply supervised and unsupervised machine learning methods to identify the main drivers of the differences between Alzheimer's Disease and PD, we will also use those to improve the accuracy of the predictive models.
3.

Impact

What was the impact of the development of the principal discipline of the project?

Nothing to Report

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

Challenges / Problems

Changes in approach and reasons for change?

In the original plan, we aimed to include two groups of pre-symptomatic PD patients (two to five years before onset, and five to eight years before onset). We identified them in the context of the Framingham Study, which has 79 participants with available plasma samples ten years prior to diagnosis and 92 participants with a plasma samples two to five years before PD diagnostic. We submitted the tissue request November 4th 2020 that was reviewed by the pertinent committees on December 2nd 2020. The committee found the proposal to have scientific merit and fills a clear need, but rejected the application due to concerns about the suitability of Framingham samples due to their long-term storage and lack of preliminary data on PD. They suggested a resubmission with preliminary data generated on PD individuals.

To overcome the rejection, we have included samples from early and late PD to generate the predictive model. Once this data is available, we plan to resubmit the tissue request to the Framingham Study to access the pre-symptomatic samples and then replicate the predictive model there. Due to the high value of these samples, we will generate RNAseq data instead of only quantifying the genes included in the predictive model.

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

The COVID-19 pandemic has affected all the research labs, and we are no exception. Due to lab closures and shortage of supplies, we estimate that we are between four and six months behind the initial schedule. We are already analyzing the data at fast pace to try to minimize the delay.

We anticipate some hurdles to access the Framingham Study samples. They requested a feasibility study to demonstrate that we can work with low volumes and aged samples. We are currently analyzing data from Alzheimer's disease individuals that have been more than 25 years in storage. We plan to use that data as the feasibility study.

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

Significant changes in use or care of human subjects

Nothing to Report

Significant changes in use or care of vertebrate animals.

Nothing to Report

Significant changes in use of biohazards and/or select agents

Nothing to Report

Products

Nothing to Report

Participants and other collaborating organizations

What individuals have worked on the project?

Name:	*Laura Ibanez*
Project Role:	*Principal Investigator*
Researcher Identifier (e.g. ORCID ID):	0000-0003-2381-7059
Nearest person month worked:	*3*
Contribution to Project:	*Dr. Ibanez has supervised all the steps for the correct development of the project. She performed half of the ribodepletion protocols, generated all the libraries and performed the quality control. She planned for the pooling and was in contact with the sequencing center during the sequencing project. She has supervised Dr. Chen.*
Funding Support:	*National Institute of Aging* *Alzhiemer's Drug Discovery Fundation*

Name:	*Hsiang-Han Chen*
Project Role:	*Post-Doctoral Research Associate*
Researcher Identifier (e.g. ORCID ID):	0000-0003-1329-3298
Nearest person month worked:	*5*
Contribution to Project:	*Dr. Chen has prepared all the bioinformatics pipelines needed to process the files provided by the sequencing center. He has processed all the files and performed the quality control. He is currently working on the analysis of the data (not included in this report)*
Funding Support:	*Alzheimer's Drug Discovery Foundation*

Name:	*Kristy Bergmann*
Project Role:	*Lab Technician*
Researcher Identifier (e.g. ORCID ID):	N/A

Nearest person month worked:	*3*
Contribution to Project:	*Ms. Bergmann has performed all the RNA extractions and quality control. She has also performed half of the ribodepletion protocols all the pooling and the adapter-primer cleaning.*
Funding Support:	*National Institute of Aging* *Alzheimer's Drug Discovery Fundation*

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

There is no pending support at the time of this report.

Dr. Ibanez was awarded a Bright Focus Foundation Standard Award Program in Alzheimer's Disease research for her grant application entitled "*Pathophysiology of sRNAs in Alzheimer's Disease"* in the amount of $299,181.00 for the period 07/01/2021 – 06/20/2024.

Title: *Pathophisiology of miRNAs in Alzheimer Disease*
Effort: *0.3 calendar*
Supporting Agency: *BrightFocus Foundation A2021033S (Ibanez)*
Address:
22512 Gateway Center Drive
Clarksburg, MD 20871
1-800-437-2423
Contracting/Grants Officer: *Diane Bovenkamp, PhD; dbovenkamp@brightfocus.org*
Performance Period: *7/1/21-6/30/24*
Level of Funding: *$299,181 (total award)*
Project Goals: *In this proposal, we will characterize the different populations of small RNAs in the brain, plasma, and cerebrospinal fluid of individuals with Alzheimer's disease.*
Specific Aims: *We aim to identify differentially expressed sRNAs to understand the circulation of small RNAs (sRNAs) among CSF, blood (plasma) and brain (Aim 1). After proving that there are sRNAs (different from microRNAs) that cross the blood brain barrier we plan keep investigating the biology of sRNAs and their involvement in the pathophysiology of Alzheimer disease. To have more power to understand the biology of sRNAs we want to explore more complex networks and pathways. Thus, we plan to increase the sample size of all three body compartments to identify novel sRNAs that cross the blood brain barrier. This will also allow us to replicate and improve the accuracy of the predictive model (proposed in Aim 1). If our predictive model has a good accuracy, we plan to generate sRNA data in other neurodegenerative diseases to perform specificity analyses. We plan to investigate if the model generated is predicting Alzheimer disease specifically or neurodegeneration in general. In Aim 2 we will investigate the involvement of small RNAs (sRNAs) in the dysregulation of amyloid-beta. We plan to follow-up the sRNAs that alter the production of amyloid-beta as potential drug targets.*
Overlap: *No scientific or budgetary overlap with the proposal*

What other organizations were involved as partners?

Nothing to Report

Special Reporting Requirements

N/A

Appendices