National Surgical Adjuvant Breast and Bowel Project (NSABP) Foundation

Annual Progress Report: 2010 Formula Grant

Reporting Period

January 1, 2011 – June 30, 2011

Formula Grant Overview

The National Surgical Adjuvant Breast and Bowel Project (NSABP) Foundation, Inc. received $967,922 in formula funds for the grant award period January 1, 2011 through December 31, 2012. Accomplishments for the reporting period are described below.

Research Project 1: Project Title and Purpose

Next-generation Sequencing-based RNA Expression Profiling for Archived Tumor Tissue - The purpose of this project is to develop a protocol for next-generation (NexGen) sequencing of RNAs isolated from formalin-fixed, paraffin-embedded tissues.

Anticipated Duration of Project

1/1/2011 - 12/31/2012

Project Overview

The broad goals of the NSABP pathology laboratory are to find molecular markers that can be used for prognostic or predictive information that can improve the diagnosis and treatment of breast and colon cancers. The procedure currently used for the discovery of prognostic and predictive biomarkers involves several steps. First, whole genome expression analysis is done on a large discovery cohort consisting of approximately 1000 samples from NSABP's vast resources of archived tissues; second, prognostic and predictive genes are identified. Third, the expression profiles on the discovery cohort are repeated with a platform that is amenable as a clinically appropriate test. This step, referred to as platform validation, is important because genes discovered on the whole transcriptome platform are not always validated on a smaller clinically appropriate platform; also, this step adds a great deal of time, and cost to biomarker discovery. The final step is the clinical validation which interrogates a new cohort of patients. Next-generation (NexGen) RNA sequencing is a powerful new technology that allows for an unbiased and complete examination of the transcriptome. Furthermore, this technology is being developed as a clinical test and would simplify the process of discovery and validation.

To expedite the process of biomarker discovery and to lead to new biological insights into breast cancer, we propose the following specific aims: 1) determine the best platform for NexGen analysis of RNAs isolated from formalin-fixed, paraffin-embedded (FFPE) tissue material;

2) purchase and install instruments and work flow for NexGen sequencing at the NSABP; and 3) evaluate NexGen gene-expression profiling data from RNAs isolated from formalin-fixed, paraffin-embedded tissues (FFPET) collected in NSABP clinical trial B-20.

To address aim 1: RNAs from cell lines, from both fresh and FFPE preparations, will be isolated. cDNA libraries will be prepared and profiled on both the Illumina® and Life Technologies instruments. Data will be evaluated for reproducibility and accuracy to determine the best platform for use with RNAs isolated from FFPET. In the second phase we will purchase the equipment and perform NexGen RNA sequencing analysis using 200 RNAs isolated from FFPET from NSABP clinical trial B-20. The data will be evaluated for biological accuracy and will provide preliminary data for development of a prognostic gene signature for ER+ and node-negative breast cancer patients.

Principal Investigator

Katherine L. Pogue-Geile, PhD
Assistant Director of Molecular Profiling
NSABP Foundation, Inc.
Four Allegheny Center, 5th Floor
Pittsburgh, PA 15212

Other Participating Researchers

Chung Kim, MD; Patrick Gavin, BS; Noriko Yamaguchi, PhD; – employed by NSABP Foundation, Inc.

Expected Research Outcomes and Benefits

The following are expected outcomes and benefits of this research:

1. To develop a process and protocol for NexGen RNA sequencing (RNA-Seq) analysis which can be used to profile gene expression utilizing RNAs isolated from archival, FFPE clinical trial samples. The development of such a protocol can unleash a wealth of information in a single experiment that is not possible with other technologies.
2. To provide preliminary information for the development of a gene expression signature that will identify ER+, node-negative breast cancer patients who have good prognoses. Such a test may identify breast cancer patients who only need to be treated with tamoxifen and may spare patients the side effects of chemotherapeutic treatment.

Summary of Research Completed

Progress regarding the first specific aim, which is to determine the best platform for NexGen analysis of RNAs isolated from formalin-fixed, paraffin-embedded (FFPE) tissue material.

Preparation of RNA. To make this assessment, as proposed, we have prepared RNA from two cell lines, MCF-7 and SKBR3, and from 3 different preparations of these cells: fresh cells (subsequently frozen), FFPE H&E-stained slides and FFPE unstained slides. One harvest from

each cell line was used to prepare RNA from these 3 different preparation sources to ensure that expression differences were not a result of differences in culture conditions.

Preparing cDNAs ready for RNA-SEQ library construction. For RNAs isolated from fresh preparations the procedures for constructing cDNA libraries is somewhat complex, but detailed protocols for their construction are available. However, the procedures for the preparation of cDNA libraries from degraded RNAs, such as those found in FFPE sources, are not generally available, and the methods are still being optimized by the manufacturers and other researchers. We have obtained pre-release information from Life Technologies, Inc.; Illumina, Inc.; and NuGEN, which supplies kits for the preparation of RNA-Seq libraries for the Illumina platform. We are currently in the process of preparing these RNA-Seq libraries; the details of our progress are outlined below.

Because about 90% of the total RNA is composed of ribosomal RNAs (rRNAs), it is essential that messenger RNA (mRNA) be purified from the total RNA. For pristine RNAs isolated from fresh tissues this process is done by selection of messenger RNA by oligo-dT selection. However, for RNAs isolated from FFPE sources, this selection is not ideal because the oligo-dT selection isolates only the immediate 3-prime end of messages due to degradation of the mRNAs. To circumvent this problem we have experimented with several procedures to enhance the selection of mRNA from our total RNA isolations.

Preparation of RNA-Seq Libraries for Illumina. To enrich our total RNA samples for message, we have tried 2 different procedures. The NuGEN Ovation® RNA-Seq System kit claims to selectively amplify mRNA, combining both rRNA removal and amplification in one kit. The procedure is proprietary so the mechanism for how this works is unclear. The other procedure we tested was the Ribo-Zero™ rRNA Removal Kit, supplied by Epicentre®, now Illumina®. Ribo-Zero employs hybrid selection using oligos homologous to rRNA but in this case the hybrid selected material is removed, and the sample that remains is the ribosomal-depleted RNA (-rRNA RNA).

After removal of rRNA, the RNA must be converted to cDNA and amplified. For this procedure, we tested 2 different amplification kits, the NuGEN Ovation RNA Seq FFPE System amplification kit, supplied by NuGEN and the ScriptSeq™ mRNA-Seq Library Preparation Kit from Epicentre. Five micrograms of total RNA isolated from MCF7 and SKBR3 cell lines from both fresh preparations and FFPE slides were treated with the Ribo-Zero kit following the procedure recommended by Epicentre. To test the Ovation amplification, 200 ng of the same total RNA samples were amplified following the details outlined in the Ovation kit.

Evaluation of NuGEN Ovation using total RNA as input. On the basis of Tables 1 and 2 below, we concluded that Ovation amplification kits were able to amplify and convert the RNA into cDNA: 0.2 ug of input RNA resulted in 1 ug of cDNA for all samples (Table 1). However, selective amplification of mRNA only occurred in the fresh samples. RT-PCR analysis of the Ovation cDNAs showed that the Ct values for 18S increased after RNAs were treated with the Ovation kit (16.05 vs 20.93 for MCF-7 and 15.77 vs 21.47 for SKBR3) and decreased for GUSB (mRNA reference) but only when RNAs were isolated from *fresh* cells (Table 2). This is what we would expect if mRNA was selectively being amplified by Ovation: A higher Ct value

represents lower expression because a greater number of cycles were required to reach the threshold; a lower Ct value represents an increase in expression. However, in the FFPE samples mRNA was NOT selectively amplified; rather rRNA was amplified in both of the cell lines and mRNA (GUSB) was not. On the basis of these findings, we concluded that Ovation amplification was an effective strategy for the amplification of RNA to cDNA but was not selective for mRNA when degraded RNA samples were used. Therefore, other methods were required to remove rRNA in our FFPE samples.

Evaluation of Ribo-Zero and ScriptSeq Amplification. Ribo-Zero was first evaluated for its ability to remove rRNA from fresh and FFPE preparations because it was recommended for preparation of Illumina libraries. Table 3 demonstrates that Ribo-Zero effectively removed nucleic acid from our sample based on the total ng yields before and after the procedure. To quantify the amount of rRNA, RT-PCR assays with 18S and GUSB TaqMan® probes (Life Technologies) were done (Table 2). These assays showed that there was a substantial decrease in the amount of 18S RNA and a substantial increase in the amount of message (GUSB) based on the substantial increase and decrease in the 18S and GUSB Ct values, respectively, for fresh samples (Table 2). Therefore, Ribo-Zero effectively removed rRNAs.

Given that the Ribo-Zero kit appeared to remove rRNA (or at least it has removed the part of 18S rRNA, which is amplified by the TaqMan assay) and that the Ovation kit effectively amplifies RNA to cDNA, we decided to combine the 2 procedures; first removing rRNA and then amplifying with the Ovation procedure. We have completed rRNA removal with Ribo-Zero and amplification and conversion of RNA to cDNA with the Ovation kit. The RT-PCR analysis of these samples is shown in Table 4 and demonstrates that 18S was effectively removed from the samples.

Fragmentation of cDNAs. The final step to make the samples ready for library construction is fragmentation. The necessity of fragmentation for cDNAs from degraded RNAs is dependent on the size of the cDNAs. Therefore, we examined the size of the all the cDNAs by running the preparations on a Bioanalyzer. As expected the RNAs isolated from fresh cell lines had traces that reflected the presence of RNAs greater than the target size of 150 to 200 bps. However, we were surprised to find that the FFPE RNAs also showed a long shoulder indicating the presence of cDNAs greater than 200 bps. Therefore, we decided to fragment all cDNA samples using sonication, which was performed with a Covaris® instrument. Conditions for Covaris treatment are shown in Table 5.

Preparation of RNA-Seq libraries for SOLiD™. Library preparation for samples with the SOLiD platform is distinctly different in that adapters are ligated onto RNAs rather than onto cDNAs as is the case for the Illumina library preparation.

While Ribo-Zero appeared to remove the 18S RNA, we also wanted to test another rRNA removal procedure, which was recommended by Life Technologies for the SOLiD platform. We also acquired pre-release information from Life Technologies for adding additional oligonucleotide probes to the commercially available RiboMinus™ kit. The purpose of these additional probes was to enable a more thorough removal of rRNAs. Because the removal of rRNA is critical to the production of meaningful RNA-Seq data, we thought it wise to also test

the RiboMinus procedure and compare it to Ribo-Zero. Table 3 shows that nucleic acids were removed after treatment with both Ribo-Zero and RiboMinus, although the total amount of nucleic acid removed was greater with Ribo-Zero than with RiboMinus. Quantification of ribosomal RNA with 18S and 28S TaqMan probes revealed that both Ribo-Zero and RiboMinus removed rRNAs. The increase in the Ct value with the 18S probe after removal of the rRNA demonstrated that there was a substantial removal of rRNA.

RNA - rRNA purification and fragmentation. RNA minus rRNA samples were purified with RNA clean and concentration–5 columns (Zymo Research, Irvine, CA) and were fragmented with sonication using the Covaris instrument. Conditions for the Covaris fragmentation are shown in Table 5. The samples can then be used to generate cDNA libraries for RNA-Seq analysis on the SOLiD platform.

Detailed methods for preparation of cell line RNAs. MCF-7 and SKBR3 cell lines were grown in RPMI medium supplemented with 10% fetal bovine serum. When the cells had reached near saturation density, cells were removed from three 150 cm^2 tissue culture flasks with trypsin. Complete media was added to stop trypsinization. For fresh cell preparations, approximately one million cells were removed from the trypsinized flask and were placed in a 1.5 ml Eppendorf tube; preparations were then centrifuged and were subsequently frozen at -80°C for later isolation of RNA. The remainder of the cells, approximately 1 X 10^8 cells, were centrifuged separately, the media was removed, and the cells were fixed in formalin and embedded in paraffin according to standard fixation and embedding procedures.

Summary. During this first report period, we have made significant progress in the preparation of cDNA libraries for NexGen sequence analysis.

Table 1. Results of Ovation Amplification from MCF 7 and SKB R3 Cell Lines

	Total ng		**260/280**	
	Before Ovation	**After Ovation**	**Before Ovation**	**After Ovation**
MCF7 Fresh	200	9544.8	1.97	1.77
MCF7 FFPE	200	10794.9	2.02	1.82
MCF7 H&E	200	11375.7	1.9	1.87
SKBR3 Fresh	200	12547.2	1.88	1.87
SKBR3 FFPE	200	9981.6	1.86	1.85
SKBR3 H&E	200	10459.2	1.86	1.86

Table 2.	Results of Ribo-Zero, Ovation, and ScriptSeq Amplification					
Experimental conditions	18s Ct Values			GUSB Ct Values		
	RT only	RT+ Ovation Amp	RT+ribo-0 + ScriptSeq Amp	RT only	RT+ Ovation Amp	RT + ribo-0 + ScriptSeq Amp
MCF7 FRESH	16.052*	20.93	25.384	29.87*	25.53	34.506
MCF7 FFPE	20.76	18.62	27.781	32.44	32.10	40
SKBR3 FRESH	15.772*	21.47	ND	28.41*	24.86	ND
SKBR3 FFPE	18.364*	17.21	26.274	32.52*	26.37	40

Table 3. Results of Ribo-Zero and Ribo Minus Kits for rNA Removal						
	Total ng		fold decrease (input/ output)	Total ng		fold decrease (input/ output)
	Input Ribo-zero	Output Ribo-zero		Input Ribominus	Output Ribominus	
MCF7 Fresh	5000	943.4	5.3	5000	1482	3.4
MCF7 FFPE	5000	556	9.0	5000	1314	3.8
MCF7 H&E	5000	550.8	9.1	2600	758	3.4
SKBR3 Fresh	5000	2133	2.3	4750	2190	2.2
SKBR3 FFPE	5000	1202.6	4.2	5000	844	5.9
SKBR3 H&E	5000	881.8	5.7	ND	ND	

Table 4. RT-PCR Analysis of rRNA Removal with RiboMinus and Ribo-Zero Kits				
Sample	Detector	No rRNA Removal-	Ribominus Ct	Ribozero Ct
MCF7 Fresh	RN28S1	not tested	>40	35.74
MCF7 Fresh	RN18S1	16.052	23.05	24.18
MCF7 FFPE	RN28S1	not tested	>40	35.79
MCF7 FFPE	RN18S1	19.388	24.08	23.84
MCF7 H&E	RN28S1	not tested	>40	37.56
MCF7 H&E	RN18S1	not tested	27.24	27.22
SKBR3 Fresh	RN28S1	not tested	>40	34.89
SKBR3 Fresh	RN18S1	15.7715	25.59	22.32
SKBR3 FFPE	RN28S1	not tested	>40	39.42
SKBR3 FFPE	RN18S1	18.3635	26.35	28.20
SKBR3 H&E	RN28S1	not tested	not tested	35.96
SKBR3 H&E	RN18S1	not tested	not tested	25.34
In samples with a Ct values of >40, the gene was not detected.				

Table 5. Conditions for Covaris Sonication of cDNA and RNA		
Covaris Conditions	**cDNA for Illumina**	**SOLiD RNA-rRNA**
duty cycle	10%	10%
intesity 5	5	5
cycles/burst	100	200
time	5 minutes	75 seconds
sample volume	120ul	130 ul
temp	4-8 C	4-8 C
power mode	Frequency sweeping	Frequency sweeping
degassing mode	continuous	continuous
buffer	Low TE	Low TE
water level	S2- Level 12, E210-Level6	S2- Level 12, E210-Level6
AFA intesnifier	Yes	Yes