National Surgical Adjuvant Breast and Bowel Project (NSABP) Foundation

Annual Progress Report: 2008 Formula Grant

Reporting Period

July 1, 2009 – June 30, 2010

Formula Grant Overview

The National Surgical Adjuvant Breast and Bowel Project (NSABP) Foundation, Inc. received $1,288,794 in formula funds for the grant award period January 1, 2009 through December 31, 2011. Accomplishments for the reporting period are described below.

Research Project 1: Project Title and Purpose

*Development of Prognostic Index for Colon Cancer Patients Using Gene Expression Profiling* – Currently about a half of patients with a colon cancer diagnosis receive very toxic chemotherapy regimens containing oxaliplatin. However, we know that much milder chemotherapy that does not contain oxaliplatin may be effective in preventing recurrence in a considerable proportion of colon cancer patients. A test that can identify those who do not need oxaliplatin will help many patients avoid unnecessary, toxic therapy. We will utilize a whole genome expression analyses method to screen tumor specimens collected from already finished, large clinical trials conducted by the National Surgical Adjuvant Breast and Bowel Project to develop a prognostic test to achieve this goal. Statistical rigor will be exercised in the project to assure that the test we develop will be highly reliable for actual clinical use.

Anticipated Duration of Project

1/1/2009 - 12/31/2011

Project Overview

The broad research objective of the project is to improve the clinical care of patients with colon cancer diagnoses. The specific aim is to develop a prognostic test using gene expression profiling to identify patients who may not require more than 5-fluorouracil plus leucovorin (FULV) adjuvant chemotherapy. The current standard of care for Stage II or III colon cancer is adjuvant chemotherapy with an oxaliplatin-containing regimen (FULV plus oxaliplatin [FLOX or FOLFOX, based on schedule]). This standard was established by the NSABP trial C-07, which showed the superiority of FLOX over FULV. However, the FOLFOX regimen is very toxic with numerous side effects, including neurotoxicity. Data from NSABP C-07 and preceding trials conducted by our group and others make it clear that many patients may not require oxaliplatin. Yet, there are no reliable prognostic markers in clinical use to identify patients whose prognoses are good enough after treatment with only FULV such that oxaliplatin...
is not required. Developing such a prognostic test will relieve suffering from unnecessary toxic therapy and lead the way to personalized care of colon cancer patients.

Our goal is achievable because 1) we have collected tumor biopsy specimens in paraffin blocks from a majority of patients enrolled in the NSABP C-07 trial that tested oxaliplatin, and 2) we have methods to use these blocks for whole genome expression profiling at relatively low cost to allow screening of a large number of cases with statistical rigor.

We will use the Whole Genome DASL assay (Illumina) to examine expression levels of 25,000 genes in tumor biopsy samples to identify genes that are prognostic for FULV-treated patients in NSABP C-07 (N=1,100). The cohort will be divided into two subcohorts of 550 each; one will be used for gene discovery and the other for refinement of discovered genes in order to minimize false discovery. Linearity and dynamic range of the expression data for each prognostic gene and correlation between the genes will be used to further select candidate genes. We will then develop the nCounter assay (NanoString) to rescreen the selected prognostic genes in the same samples to develop a prognostic algorithm based on the nCounter assay. The algorithm will be prospectively tested in completely independent FULV-treated cohorts from the NSABP C-05 and C-06 trials (N=1000) not used for building of the prognostic algorithm. The statistical rigor built into the project will ensure that the developed prognostic algorithm will be a reliable clinical test for use in routine clinical practice.

Principal Investigator

Soonmyung Paik, MD
Director, Division of Pathology
NSABP Foundation, Inc.
1307 Federal St, Suite 303
Pittsburgh, PA 15212

Other Participating Researchers

Joseph P. Costantino, DrPH – employed by University of Pittsburgh

Expected Research Outcomes and Benefits

Treatment of colon cancer is evolving rapidly and many new agents are being added to what used to be considered a standard chemotherapy (FULV). However, the addition of new agents comes at the expense of increased toxicity to patients and increased health care costs. Since not all patients need such toxic treatments, identifying those who do not need such treatments is a top clinical research priority.

The prognostic index for colon cancer treated with an FULV adjuvant chemotherapy regimen developed from this project is expected to help identify patients with stage 2 or 3 colon cancer diagnoses, who may not need more than FULV treatment and still enjoy good long-term survival. This prognostic test will result in sparing these patients from unnecessary treatment
and toxicity associated with additional therapy such as oxaliplatin, bevacizumab, or cetuximab. It will also reduce the social burden resulting from less discriminate use of the expensive agents.

**Summary of Research Completed**

Whole genome DASL arrays from Illumina were used to profile the gene expression of approximately ½ (N=956) of the total number tumor blocks (N=2222) collected from patients enrolled in NSABP clinical trial C-07. These 956 samples represent the available discovery cohort.

To identify a gene score that could be used for prognosis, the discovery cohort was subdivided into training and test subsets of equal size. Supervised Principal Component Analysis (PCA) (as a part of BRB-Array Tools [Biometrics Research Branch-NCI]) was performed on the training subset (N=478) to develop a prognostic gene score from 68 genes significant at p<0.001. Patients in the test subset (N=478) were classified as high or low risk based on the median value of the prognostic gene score from the training set; risk class was significantly associated with time to recurrence (TTR) (N=478, HR= 2.65, 95% CI: 1.84-3.83, p<0.0001). The Kaplan-Meier plots shown in Figure 1 show how closely the training and test subsets overlap. Subsequent analyses using the full discovery cohort of 956 patients found that 85.1% of patients with low risk were recurrence-free at 5 years versus 58.7% recurrence-free patients in the high risk class.

Although the gene score was developed strictly for prognosis, the gene score provided predictive value for oxaliplatin. When Cox models were used to analyze patient survival based on their gene score and on treatment with oxaliplatin, the high-risk group showed marginal significant benefit from oxaliplatin treatment (HR=0.682, p=0.009) but the low risk gene score group did not (HR=1.17, p=0.49). The Kaplan-Meier plot in Figure 2 demonstrates this difference in treatment response between the 2 risk groups. When the gene score was tested for its interaction with oxaliplatin in the Cox model it was marginally significant (HR=0.586, p=0.046).

Furthermore, multivariate analysis, which included the gene expression score and nodal status, segregated patients into low-, intermediate-, and high-risk categories (Figure 3). The low-risk group (Figure 3, black lines) consisted of tumors with low-risk gene scores and with 3 or fewer positive nodes. This group of patients had an excellent prognosis with a 5-year survival rate of more than 95%, and, because of their excellent prognosis, oxaliplatin would not be recommended. The intermediate group consisted of tumors either with high-risk gene scores and with 3 or fewer positive nodes (Figure 3, brown lines) or with a low-risk gene scores and 4 or more positive nodes (Figure, 3 red lines). This intermediate risk group did not appear to receive benefit from oxaliplatin. If this observation were validated, it would be a critical tool for clinical practice because it could be used to identify patients who should be treated with something other than oxaliplatin. The high-risk group was composed of tumors that had a high- risk gene score and 4 or more positive nodes and appeared to receive benefit from oxaliplatin (Figure 3, green lines).

**Selection of Candidate Prognostic and Oxaliplatin-predictive Genes for Use in Platform Validation**

Whole genome DASL is not an ideal platform for use as a clinical test because it is too expensive, too complex, and too labor intensive. Therefore, we have selected a relatively small
number of genes that will be assayed again on a platform that is more appropriate for a clinical test.

In order to maximize the number of candidate genes that may be useful for prognosis and for prediction of oxaliplatin response, we have analyzed the whole genome DASL data in a variety of ways. In the analysis mentioned above we identified 68 genes with PCA using DASL data in which low-expression genes were removed. In addition, a separate PCA analysis was used to build a prognostic gene signature using unfiltered DASL data. This analysis identified 33 genes which were used for prognosis. There were 22 genes that were included in both the 68- and 33-gene signatures. The total number of genes that were included as prognostic or predictive candidate genes as a result of C-07 analysis was 282. The number and the criteria for selection of prognostic and predictive candidate genes are shown below:

- 79 prognostic genes that made up the prognostic gene score and were identified by PCA;
- 114 prognostic genes that had a hazard ratio (HR) <0.7 or >1.3 and p<0.001;
- 89 interaction genes with a HR <0.7 and >1.3 and p < 0.01

Detailed Methods
We have improved our method of RNA extraction this year by increasing the throughput by switching to instruments that allow 96 samples to be processed in one hour. Approximately 1000 additional C-07 RNAs from the validation cohort have been prepared in this way. Specifically, RNAs were isolated from lysates prepared from 3 slides at a section width of 5µm. Lysates were deparaffinized and digested with protease and then loaded into a 96-well plate with the Tecan EVO robot, and processed with the KingFisher Flex® 96 (deep head) instrument (ThermoFisher; Burlington, ON) and the E.Z.N.A.® FFPE RNA Isolation Kit from Omega BioTek (Norcross, GA). RNAs were quantified with fluorescence, using the Quant-iT ™ RiboGreen® Assay Kit (Invitrogen; Carlsbad, CA) and the Infinite® F200 fluorometer (Tecan; Männedorf, Switzerland).

RNAs (100-200 ng) from 971 samples were profiled for gene expression utilizing the whole genome DASL (cDNA-mediated Annealing, Selection, extension, and Ligation) Arrays from Illumina (San Diego, CA). DASL was specifically designed to interrogate partially degraded RNAs and, therefore, is ideal for samples isolated from FFPET samples. The details of this procedure are provided by the manufacturer. Briefly, DASL involves cDNA synthesis, hybridization of a complex set of oligonucleotides that allow for the synthesis and amplification of more than 24,000 annotated genes derived from RefSeq (Build 36.2). Expression is interrogated by direct hybridization to Illumina Human Ref-8 bead arrays. For our studies, after hybridization BeadChips were scanned using Illumina Bead Arrayer, images were digitized, and data were collected and viewed within Illumina’s Genome Studio software package and exported to third party software for further analysis.

The gene expression profiles for 15 samples were removed from further analysis because the P95 Signal /P05 (signal-to-noise) ratio was below 20, which was a more stringent criteria than Illumina recommends. The P95/P05 ratio provides a way to visualize the overall intensity of the expression signal compared to the background noise.
Conclusion
In conclusion, we have made significant progress in the past year. We have profiled the complete discovery cohort with whole genome DASL from Illumina, analyzed this data to identify prognostic and oxaliplatin-predictive genes, and developed a prognostic gene score that may identify patients that would benefit from oxaliplatin. Furthermore, we have isolated an additional 1000 RNAs from NSABP C-07 samples which will be used for the validation.

Figure 1

Figure 2
Figure 3

C-07 RFS according to gene index, nodal status, and treatment (high risk N+ pts benefit from Oxal)

- Low risk - node negative, and gene score
- Intermediate
- High risk - node positive, low risk gene score
- Low risk - node negative, high risk gene score
- High risk - node positive and gene score

Percent survival vs RFIT