Final Progress Report for Research Projects Funded by Health Research Grants

Instructions: Please complete all of the items as instructed. Do not delete instructions. Do not leave any items blank; responses must be provided for all items. If your response to an item is "None", please specify "None" as your response. "Not applicable" is not an acceptable response for any of the items. There is no limit to the length of your response to any question. Responses should be single-spaced, no smaller than 12-point type. The report **must be completed using MS Word**. Submitted reports must be Word documents; they should not be converted to pdf format. Questions? Contact Health Research Program staff at 717-783-2548.

1. **Grantee Institution:** NSABP Foundation, Inc.

2. **Reporting Period (start and end date of grant award period):** 1/1/10 – 12/31/13

3. **Grant Contact Person (First Name, M.I., Last Name, Degrees):** Joshua Cortina

4. **Grant Contact Person's Telephone Number:** 412-339-5316

5. **Grant SAP Number:** 4100050903

6. **Project Number and Title of Research Project:** 1 - Discovery and Validation of MicroRNAs as Biomarkers in Breast and Colon Cancer

7. **Start and End Date of Research Project:** 1/1/10 – 12/31/13

8. **Name of Principal Investigator for the Research Project:** Katherine Pogue-Geile**,** PhD

9. **Research Project Expenses**.

   9(A) Please provide the total amount of health research grant funds spent on this project for the entire duration of the grant, including indirect costs and any interest earned that was spent:

   $ 1,258,856.28

   9(B) Provide the last names (include first initial if multiple individuals with the same last name are listed) of **all** persons who worked on this research project and were supported with health research funds. Include position titles (Principal Investigator, Graduate Assistant, Post-doctoral Fellow, etc.), percent of effort on project and total health research funds expended for the position. For multiple year projects, if percent of effort varied from year to year, report in the % of Effort column the effort by year 1, 2, 3, etc. of the project (x% Yr 1; z% Yr 2-3).

Last Name, First Name	Position Title	% of Effort on Project	Cost
Blackmon	Research Assistant	6% Yr 1; 35% Yr 2; 33% Yr 3	$28,670
Gavin	Molecular Supervisor Tech	3% Yr 1; 16% Yr 2; 38% Yr 3; 11% Yr 4	$30,575
Kim, Chung Yeul	Manger of Molecular Pathology Lab	8% Yr 1; 7% Yr 2; 11% Yr 3	$20,988
Kim, Seong Rim	Research Fellow	10% Yr 2; 9% Yr 3	$10,558
Lipchik	Research Assistant	2% Yr 1; 15% Yr 2; 33% Yr 3	$18,095
Paik	Director of Division of Pathology	6% Yr 1; 14% Yr 2; 16% Yr 3; 9% Yr 4	$60,182
Pogue-Geile	Principal Investigator	10% Yr 1; 21% Yr 2; 33% Yr 3; 21% Yr 4	$53,060
Reilly	Research Tech	2% Yr 2	$558
Remillard	Research Assistant	2% Yr 1	$840
Song	Biostatistician	3% Yr 3; 4% Yr 4	$6,787
Yamaguchi	Biostatistician	21% Yr 3	$15,452

9(C) Provide the names of **all** persons who worked on this research project, but who *were not* supported with health research funds. Include position titles (Research Assistant, Administrative Assistant, etc.) and percent of effort on project. For multiple year projects, if percent of effort varied from year to year, report in the % of Effort column the effort by year 1, 2, 3, etc. of the project (x% Yr 1; z% Yr 2-3).

Last Name, First Name	Position Title	% of Effort on Project
Cortina, Joshua	Grants and Special Projects Administrator	2% Yrs 1-4

9(D) Provide a list of **all** scientific equipment purchased as part of this research grant, a short description of the value (benefit) derived by the institution from this equipment, and the cost of the equipment.

Type of Scientific Equipment	Value Derived	Cost
HT7900 Fast Real Time Real Time PCR System	The purchase of the HT7900 Fast Real Time PCR System from Life Technologies was essential to this project and was an essential piece of equipment that had been missing from our laboratory. It was best instrument available at the time which sensitively and specifically profiled the vast majority of microRNAs. TaqMan PCR on the 7900 is still considered the gold	$44,710

	standard for assessing gene expression. The NSABP Pathology laboratory possesses many platforms for assessing whole genome expression but we did not have instrumentation that allowed us to compare our whole genome expression analysis to the TaqMan gold standard.	

10. Co-funding of Research Project during Health Research Grant Award Period. Did this research project receive funding from any other source during the project period when it was supported by the health research grant?

Yes_____ No___X_____

If yes, please indicate the source and amount of other funds:

11. Leveraging of Additional Funds

11(A) As a result of the health research funds provided for this research project, were you able to apply for and/or obtain funding from other sources to continue or expand the research?

Yes_____ No____X_____

If yes, please list the applications submitted (column A), the funding agency (National Institutes of Health—NIH, or other source in column B), the month and year when the application was submitted (column C), and the amount of funds requested (column D). If you have received a notice that the grant will be funded, please indicate the amount of funds to be awarded (column E). If the grant was not funded, insert "not funded" in column E.

Do not include funding from your own institution or from CURE (tobacco settlement funds). Do not include grants submitted prior to the start date of the grant as shown in Question 2. If you list grants submitted within 1-6 months of the start date of this grant, add a statement below the table indicating how the data/results from this project were used to secure that grant.

A. Title of research project on grant application	B. Funding agency (check those that apply)	C. Month and Year Submitted	D. Amount of funds requested:	E. Amount of funds to be awarded:
None	☐NIH ☐ Other federal (specify:_____ _____)		$	$

	☐ Nonfederal source (specify: _____)			

11(B) Are you planning to apply for additional funding in the future to continue or expand the research?

Yes___X_____ No_____

If yes, please describe your plans:

I am planning to submit a grant to the NCI in 2014 which will use the micro RNA (mir) and expression data to support a proposal that will allow us to understand how the immune system works to influence the prognosis of colon cancer patients.

12. Future of Research Project. What are the future plans for this research project?

We plan to continue data analysis of the data acquired within this research project to develop prognostic or bevacizumab predictive models. Our main interest is to develop a model that will identify patients who received benefit from bevacizumab. Such a model will be tested in our validation cohort and if the model is validated in the C-08 validation cohort and if the results are clinically meaningful, then we will seek additional partnerships for rigorous validation and commercial development.

13. New Investigator Training and Development. Did students participate in project supported internships or graduate or post-graduate training for at least one semester or one summer?

Yes___X_____ No_____

If yes, how many students? Please specify in the tables below:

	Undergraduate	Masters	Pre-doc	Post-doc
Male			2	
Female			1	3
Unknown				
Total			**3***	**3**

	Undergraduate	Masters	Pre-doc	Post-doc
Hispanic				
Non-Hispanic			3	3
Unknown				
Total			**3***	**3**

	Undergraduate	Masters	Pre-doc	Post-doc
White			3*	
Black				
Asian				3
Other				
Unknown				
Total			**3**	**3**

*NSABP is not an academic institution so access to students is limited; however 3 of our B.A. or B.S. level technicians have pursued or are about to pursue graduate training and I believe their experience at NSABP was instrumental in their acceptance to graduate programs.

14. **Recruitment of Out-of–State Researchers**. Did you bring researchers into Pennsylvania to carry out this research project?

Yes___X_____ No_____

If yes, please list the name and degree of each researcher and his/her previous affiliation:

Mathew Remillard, B.S. was a recent graduate of Carnegie Mellon University but was a resident of Massachusetts when he was hired for this project.

Noriko Yamaguchi, Ph.D. Department of Medical Oncology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA, USA

Seong Rim Kim, M.D. Diagnostic Pathology Dept. Medplan Lab, Seoul, South Korea

15. **Impact on Research Capacity and Quality**. Did the health research project enhance the quality and/or capacity of research at your institution?

Yes___X_____ No_____

If yes, describe how improvements in infrastructure, the addition of new investigators, and other resources have led to more and better research.

This research proposal has improved the quality and capacity of research at our institution by the addition of new investigators and by providing other resources which have led to more and better research. The following new investigators have led to more and better research.

The health research funds have supported the following new investigators:

Noriko Yamaguchi, Ph.D. was recruited from Dana-Farber Cancer Institute and Harvard Medical School as our first Laboratory biostatistician. Noriko was our first biostatistician

with a strong biological background and our first laboratory-based biostatistician. This enabled the laboratory to do survival analysis with anonymized molecular and clinical data and greatly facilitated the analyses not only for this project but for all the molecular based projects in the laboratory. In July of 2012, Noriko left NSABP to become Chief of the Biostatistics Section at the National Centre for Global Health and Medicine in Japan.

Nan Song, Ph.D. was recruited from Precision Therapeutics, a life science company which develops tools and services for individualized cancer treatment. This experience along with her graduate degree from Carnegie Mellon University's Computational Biology Program made her an ideal person to act as our laboratory biostatistician.

Patrick Gavin, who was previously a Production Associate responsible for high-throughput SNP genotyping with MassArray at Sequenom Inc. was responsible for compiling the data, doing quality control of the data and assisting the biostatistician with model development. He also wrote programs to monitor instruments and notify personnel of errors.

Seong-Rim Kim, trained as a pathologist in Seoul, South Korea was essential to this project as she identified tumor areas on our slides as a first step in the isolation of RNA.

Matthew Remillard, a recent graduate from Carnegie Mellon University is now a graduate student in the Department of Molecular and Cellular Biology at Princeton University. Although Matt was a graduate in biological science, he was also able to work through a variety of software and hardware problems that we encountered when we set up the HT7900, autoloader, servers and work station for the profiling of the microRNAs.

Nicole Blackmon was recruited from the University of Pittsburgh Drug and Discovery Institute and had been responsible for high-throughput screening of drugs. She was ideally suited for carrying out isolation of RNAs utilizing our robotic instruments. Her previous training made her well suited to preforming RT-PCR of microRNAs utilizing the 384-well TaqMan Array cards.

Corey Lipchik was previously an NSABP data manager but as a biology graduate of Allegheny College, he provided technical support for carrying out mir profiling. In addition his expertise in basic programming provided some much needed expertise to automate data extraction from the Oracle data base which housed the microRNA data.

The purchase of the HT7900 Fast Real Time PCR System from Life Technologies was essential to this project and was an essential piece of equipment that had been missing from our laboratory. It was the best instrument available at the time which sensitively and specifically profiled the vast majority of microRNAs. TaqMan PCR on the 7900 is still considered the gold standard for assessing gene expression. The NSABP Pathology laboratory possesses many platforms for assessing whole genome expression but we did not have instrumentation that allowed us to compare our whole genome expression analysis to the TaqMan gold standard.

16. Collaboration, business and community involvement.

16(A) Did the health research funds lead to collaboration with research partners outside of your institution (e.g., entire university, entire hospital system)?

Yes___X_____ No_____

If yes, please describe the collaborations:

Recently, the NCI has mandated that 9 adult cooperative clinical trial groups merge to form 4 new groups. As a result, a part of NSABP which runs NCI clinical trials has merged with the Radiation Therapy Oncology Group (RTOG) and the Gynecology Oncology Group (GOG) and became NRG. A collaboration was established during a recent meeting of the 3 legacy groups. Because we had a great deal of experience with the profiling of micro RNA isolated from formalin-fixed paraffin-embedded tissue (FFPET), we agreed to profile a small number of cervical cancer samples collected by RTOG/GOG investigators (Ann H. Klopp, MD, PhD, from the University of Texas, MD Anderson Cancer Center, Houston, Texas and Joanne Weidhaas, MD, PhD; Yale Cancer Center, Yale University).

We obtained 34 FFPE tissue samples from Drs. Klopp and Weidhaas. The goal of this study was to identify mirs regulated by chemoradiation in human cervical cancer specimens. The 34 samples (17 pairs of tissues) were obtained prior to chemoradiation and 48-72 hours after initiation of chemoradiation. RNA was isolated from pre- and post-chemoradiation samples, and mirs were profiled utilizing TaqMan microRNA fluidic cards. 11 pairs had available tissues with over 50% tumor cellularity for mirs extraction and TaqMan profiling. Analysis of the data revealed several mirs with significantly different levels of expression after chemoradiation. These include miR-130b, miR-15a, miR-125a-3p, and others. Confirmatory assays are being performed.

16(B) Did the research project result in commercial development of any research products?

Yes_____ No____X_____

If yes, please describe commercial development activities that resulted from the research project:

16(C) Did the research lead to new involvement with the community?

Yes_____ No___X_____

If yes, please describe involvement with community groups that resulted from the research project:

17. Progress in Achieving Research Goals, Objectives and Aims.

List the project goals, objectives and specific aims (as contained in the grant agreement). Summarize the progress made in achieving these goals, objectives and aims for the period that the project was funded (i.e., from project start date through end date). Indicate whether or not each goal/objective/aim was achieved; if something was not achieved, note the reasons why. Describe the methods used. If changes were made to the research goals/objectives/aims, methods, design or timeline since the original grant application was submitted, please describe the changes. Provide detailed results of the project. Include evidence of the data that was generated and analyzed, and provide tables, graphs, and figures of the data. List published abstracts, poster presentations and scientific meeting presentations at the end of the summary of progress; peer-reviewed publications should be listed under item 20.

This response should be a DETAILED report of the methods and findings. It is not sufficient to state that the work was completed. Insufficient information may result in an unfavorable performance review, which may jeopardize future funding. If research findings are pending publication you must still include enough detail for the expert peer reviewers to evaluate the progress during the course of the project.

Health research grants funded under the Tobacco Settlement Act will be evaluated via a performance review by an expert panel of researchers and clinicians who will assess project work using this Final Progress Report, all project Annual Reports and the project's strategic plan. After the final performance review of each project is complete, approximately 12-16 months after the end of the grant, this Final Progress Report, as well as the Final Performance Review Report containing the comments of the expert review panel, and the grantee's written response to the Final Performance Review Report, will be posted on the CURE Web site.

There is no limit to the length of your response. Responses must be single-spaced below, no smaller than 12-point type. If you cut and paste text from a publication, be sure symbols print properly, e.g., the Greek symbol for alpha (α) and beta (ß) should not print as boxes (□) and include the appropriate citation(s). DO NOT DELETE THESE INSTRUCTIONS.

From Strategic Plan

The long-term goal of this research project is to develop biomarkers that will improve the treatment of colon and breast cancers. To achieve this goal we propose to examine the expression of microRNAs (miRNAs) which are small, non-coding RNAs that control translation of many mRNAs.

Our goal is to develop a new high-throughput, cost-effective, and simple methodology for the detection of clinically relevant biomarkers to aid in the prognosis and prediction of colon and breast cancers.

Specific Aims from Strategic plan

I. Discover and validate prognostic markers for breast cancer patients with resectable tumors and histologically positive axillary lymph node(s)

To achieve the overall goal of this proposal, which was to improve the prognostication and the treatment of colon and breast cancers, we have chosen to focus our efforts on colon cancer for both pragmatic and scientific reasons. Expression profiling of all well-annotated micro RNA (mirs) in a large cohort of cases and using the best technology required more budget expenditure than we had originally planned, which then necessitated that we choose only one cancer. To maximize the possibility of developing clinically meaningful signatures, we decided to focus on colon cancer because we thought there was more of a need for developing predictive and prognostic signatures in colon cancer. Prognostic signatures to identify breast cancer patients who will receive benefit from chemotherapy are currently a part of clinical practice.

II. Discover and validate predictive and prognostic markers for colon cancer patients with Stage II and III resectable colon cancer

We chose to use NSABP clinical trial C-08[1] for biomarker development for several reasons. NSABP trial C-08 tested the benefit of bevacizumab when added to mFLOFOX in Stage II and III colon cancer. Given the promising results in the colon cancer metastatic setting, there was much optimism at the time that bevacizumab would show benefit in the treatment of Stage II and III colon cancers. We began these experiments before the clinical assessment of bevacizumab was completed; unfortunately bevacizumab did not show benefit in the entire C-08 cohort. Nonetheless, we thought that our mir profiling work might provide a way to identify a subset of patients who did receive benefit from bevacizumab. Furthermore, the control arm of C-08 provided an opportunity to develop a prognostic signature for patients who would need more than mFOLFOX, which is standard chemotherapy for most Stage III patients.

To achieve our goal of developing prognostic or treatment-predictive markers, we chose to profile micro RNAs or mirs for this project for many reasons. Mirs act as major switches that regulate the expression of many mRNAs. The total number of mirs included in the human genome is relatively small, but they control a large number of mRNAs. It has been suggested that mirs may be able to characterize a tumor better than mRNAs would. Importantly, it has been shown that mirs remain largely intact when isolated from formalin fixed paraffin embedded tissue (FFPET), an ideal choice for this proposal since all mirs will be isolated from FFPET. A comparison of mirs isolated from fresh frozen tissues and from FFPET found that the two samples were correlated by 0.86 - 0.89 [2]. Therefore, a relatively small number of mirs may serve as sufficient targets for the development of clinical tests that can aid in cancer detection, prognosis, and prediction and that may aid monitoring of cancer treatment. Thus, the small number and the stability of miRNAs make them an ideal choice for biomarker discovery for colon cancer.

A. Mir profiling of 761 mirs in NSABP trial C-08

In order to be able to discover and validate a signature, we assigned cases from C-08 into two non-overlapping cohorts: The discovery (N=1004) and validation (n=1002) cohorts were representative of the entire C-08 cohort. All samples were profiled with the Megaplex TaqMan Microfluidic cards (Life Technologies), which profile 761 mirs and represent most of the well-annotated mirs in the human genome. We have now completed the mir profiling and have successfully obtained data for 955 cases within the discovery cohort and for 857 cases in the validation cohort.

Isolation of RNA. Initially we used total RNA preparations that we had prepared for another project. The first 450 RNAs that we interrogated were prepared with a modified version of Ambion kits using a low-throughput robotics, the Fisher mL instrument. When it was time to prepare additional RNAs, we prepared them using Omega Bio-Tek kits and utilized a high-throughput robotics instrument, the KingFisher Flex. We found that these changes to our total RNA procedure improved the quantity and quality of total RNA based on RiboGreen quantification and UV 260/280 results, respectively. Furthermore, use of the KingFisher Flex instrument increased the throughput by more than 9-fold. However, when we compared the expression of mirs in RNAs isolated with the Ambion procedure to the Omega Bio-Tek procedure, we found that mir expression was less robust from RNAs isolated from the Omega Bio-Tek procedure. However, we could not rigorously conclude that the difference detected in expression values was due to the type of preparation since the same samples were not used with both isolations. Therefore, we isolated RNA from 10 samples with both the Omega Bio-Tek kits and a modified version of the Ambion procedure. The percentage of mirs that were expressed, meaning the % of mirs that had a Ct value less than 37, was greater in all of the RNAs isolated with the Ambion procedure. The Ct value for U6 snRNA was also lower, ie, greater expression, in all of the samples isolated with the Ambion procedure. Therefore, RNAs were isolated with a modified version of the Ambion (now part of Life Technologies) kits, which were formatted into our own custom RNA preparation kit.

Global mir expression profiling was done by using TaqMan Microfluidic cards, Life Technologies, and following the protocol recommended by the manufacturer. This procedure involves the reverse transcription, pre-amplification and amplification in 2 mega-plexed reactions. Each case was interrogated with these 2 megaplexed reactions each containing 384 assays and then profiled on the A and B TaqMan Microfluidic cards. These 2 cards provided comprehensive coverage of the Sanger miRBase v14 database content for human mirs. U6 snRNA considered to be a control mir was repeated 4 times on each card.

B. *Analysis of clinical co-variates in C-08 identifies a subset of patients that received benefit from bevacizumab*

As a prelude to the analysis of C-08 mir data, a routine examination of clinical covariates to determine their association with prognosis or with prediction of bevacizumab treatment in the entire C-08 cohort was conducted. One of these covariates was mismatch repair (MMR) status, which refers to the stability of small repeat regions of the genome called microsatellites. Approximately 15% of colon tumors show DNA instability resulting in an amplification or deletion of DNA in these regions of the genome. This instability or microsatellite instability (MSI) is due to deficient mismatch repair (dMMR) due to a loss or mutation in mismatch repair

proteins. This is an important clinical variable because dMMR tumors can be the result of a heritable mutation and dMMR status has been shown by us and others to be associated with prognosis[3]. Citation 3 refers to Gavin et al see below. Surprisingly, we found that patients with tumors with dMMR received benefit from bevacizumab on the basis of their overall survival (Figure 1A). Patients with dMMR tumors derived statistically significant survival benefit from the addition of bevacizumab (hazard ratio = 0.52, p=0.03) in contrast to no benefit in patients shown to have mismatch repair proficient tumors (pMMR) (hazard ratio [HR] = 1.03) (Figure 1B) ($p_{interaction}$ = 0.035)[4]. Although a post-hoc finding, these data suggest that a molecularly-defined subset of colon cancer patients may derive clinical benefit from anti-angiogenesis agents. This observation remains to be validated in another clinical trial; however, to our knowledge, it is the only description of a biomarker defining a *subset* of early-stage colon cancer patients who received benefit from bevacizumab within a clinical trial. This is of particular importance since there was no bevacizumab benefit in the entire cohort of Stage II and III patients in C-08. This observation, if validated, could have a significant impact on clinical practice.

We also showed that there was an association between MMR status and BRAF mutation ($p<0.0001$); therefore, we also tested BRAF V600E mutation for interaction with bevacizumab. No significant interaction was seen. However, we examined as an exploratory analysis whether a combination of the two markers could further refine the subset that benefited from bevacizumab. We found that a subset defined by BRAF mutation and dMMR derived the biggest benefit with HR of 0.27 (95% CI: 0.08-0.94, p=0.028), but the sample size was too small to be conclusive (N=51 with 16 deaths) and the 3-way interaction (BRAF, MMR, bevacizumab) was not significant (p=0.121) (Figure 2).

We have proposed a testable hypothesis to explain the above observations. dMMR tumors are highly immunogenic due to the generation of mutated proteins produced as a result of mistakes made in mismatch repair[5, 6]. This strong immunogenic response must be at least in part responsible for the good prognosis associated with dMMR tumors[7-9]. VEGF-A is speculated to be one of the main tumor-derived soluble factors that act as a chemo-attractant for immature myeloid cells from the marrow to the tumor site and suppresses dendritic cell maturation, creating an immune suppressive microenvironment[10-12]. Furthermore, VEGF-A directly induces regulatory T-cell (Treg) proliferation in tumor-bearing mice through VEGFR-2, and blocking VEGF-A alone was sufficient to inhibit Treg cell accumulation in tumor-bearing mice but not in tumor-naïve mice[13]. In colon cancer patients, adding bevacizumab to chemotherapy resulted in a substantial reduction in the proportion of Treg cells in the peripheral blood of colon cancer patients[13]. Thus, we hypothesize that bevacizumab may be particularly effective in dMMR patients because bevacizumab is able to block the immunosuppressive effect of VEGF-A. Since pMMR tumors do not generally produce a strong immunological response, the blockage of VEGF is irrelevant regarding benefit from bevacizumab.

The observation that dMMR tumors receive benefit from bevacizumab in Stage II and III colon cancer was observed in the entire C-08 cohort. Thus to validate these findings we require an independent cohort of patients. Thus, we are pursuing collaboration with Roche, which conducted the AVANT trial to validate our observation.

C. Logistics for model development and validation using molecular profiling data

Initially, we had planned to develop both a prognostic and a predictive model for bevacizumab benefit in the discovery cohort of C-08 using 3-year follow-up data. In previous progress reports, we have reported on our work to develop prognostic and predictive models using 3-year follow-up data. We had indicated that once the models were developed within the discovery cohort, we would submit the models to the NCI before validation. Now, we have changed our plans and have decided to use 5-year follow-up data, which will increase the number of events and will help us to better evaluate patients' outcomes. However, new NCI regulatory rules mandated that we write a protocol describing the process of model discovery before we were able to access the clinical data for the discovery cohort. Thus, we submitted a protocol to the Protocol Review Committee (PRC) of the Cancer Therapy Evaluation Program (CTEP) on February 15, 2013. It was approved on June 17, 2013. This protocol is a 32-page document that details the types of analyses and the specific procedure that we will use to validate the signature. This is an important part of the validation procedure because the NCI will date-stamp our discovery prognostic and predictive signatures. Only after we have submitted these models will we be able to access the merged file containing molecular and clinical data for the *validation* cohort. This process provides documentation that the validation cohort remains untouched until a fully developed signature has been developed.

D. Assessing the prognostic and predictive value of mirs in the discovery cohort

We assessed the prognostic and predictive value of each mir using expression as continuous or categorical values in Cox models in 50 bootstrap data sets, each containing 2/3, randomly selected samples from the discovery cohort. (Tables 1 and 2). Due to the fact that we may use this information to develop patentable prognostic or predictive signatures, we have removed the names of the mirs in Tables 1 and 2, listing them as sequential numbers. Twelve mirs were found to be significant for prognosis at a mean p value of less than 0.05 by continuous or categorical p value. None of the mirs were significant for bevacizumab prediction when they were analyzed as a continuous variable, but eighteen mirs were significant when they were analyzed as categorical values using a p value below 0.1.

We also used mir data to subtype samples using with a Non-negative Matrix Factorization (NMF) unsupervised clustering method. The optimum number of clusters was determined by cophenetic correlation coefficients, which indicated that our samples could best be divided into 3 to 5 clusters. Unsupervised cluster analysis of samples using mir data identified 3 cluster groups with a trend for differential benefit from bevacizumab with cluster group 1 receiving some benefit, group 2 receiving no benefit, and group 3 receiving harm (Figure 3). Red lines are plots for the control patients and black lines are for patients from the bevacizumab arm. However, these results are unlikely to be clinically meaningful since the results were not significant in any of the groups, and the degree of benefit in group 1 is small (HR is 0.857).

III. <u>Initiate the development of an integrated colon cancer molecular database using miRNA data as well as mRNA and mutation data from the C-07 samples which will provide unique information for hypothesis generation. (For this project, we will</u>

explore the use of an integrated molecular software tool using the miRNA data generated in this project.)

We described above why we decided to focus on C-08 for the purpose of identifying prognostic and predictive markers.

We evaluated the prognostic and predictive p values for mirs and for genes within the discovery cohort (Figure 4). The genes were profiled with a custom code set using the nCounter platform and were selected as either prognostic in C-07 using Illumina DASL arrays or predictive for bevacizumab benefit in C-08 using Agilent arrays. Expression of some mirs was undetectable; other technical problems resulted in missing data. Thus, in Figure 4, we have removed mirs for which data were missing from more than 20% of the samples. In general, the p values for prognosis and prediction are lower in the nCounter or NanoString data than in the mir data as would be expected, since there was a pre-selection for such genes in the NanoString code set.

A. *Prognostic and predictive model development within the discovery cohort using mir and NanoString data*

Mir and or NanoString data was used to build prognostic models using only the control arm of C-08 using the SuperPC method (Figure 5). Stable genes were selected by bootstrapping. As discussed above, we calculated a prognostic p-value for each gene in 50 bootstrap datasets, each of which includes 2/3 randomly selected samples. Genes with mean p-values less than 0.05 were treated as stable genes. This analysis demonstrates that mirs or NanoString genes are able segregate patients with good or poor prognosis. However, perhaps the best model is from the NanoString data because it yielded the largest HR (3.898) with a p value <0.001. Combining the NanoString data with the mir data did not add to the prognostic model; the hazard ratio was essentially the same (HR=3.636).

We have compared the clustering samples using NMF with mir data (Figure 3) and with NanoString data (Figure 6). As with the analysis of the mir data, we found that the optimum number of clusters for NanoString data, determined by cophenetic correlation coefficients, was 3 to 5 clusters. Similar to clustering using mir data, the NanoString data also show 3 clusters with differential benefit from bevacizumab: one for benefit; one for harm; and one with no effect. Furthermore, there is significant overlap between clusters identified with mir or NanoString data (Table 3). Group 1 with mir data was the group that received benefit and it is largely found in cluster groups 1 and 2 in NanoString data, but in the NanoString, only group 1 received the benefit; therefore, NanoString was better than mir at identifying patients who received benefit. With NanoString data, group 3 received harm and overlapped with group 2 and 3 of mir data. In this instance, mir data did a better job of identifying patients who could be harmed. This demonstrates that by integrating the mir and NanoString data we may be able to refine the subset with differential benefit from bevacizumab.

i. *Models built with immune mirs and genes*

Several lines of evidence suggest that the immune reaction plays a critical role in the prognosis of colon cancer and any treatment that affects the immune reaction may alter the therapeutic

effect. We and others have shown that patients with dMMR tumors have a much better prognosis and have been shown to be associated with a high density of cytotoxic (CD8+) T and memory (CD45RO+) T cells, which may be responsible for the good prognosis associated with dMMR tumors.[6, 14-16] Our pathway analysis of whole genome expression analysis of 447 tumors from Stage II and III colorectal tumors collected in NSABP clinical trial C-08 is consistent with these observations. The 5 most significantly differentiated pathways between dMMR and pMMR tumors were 4 T-cell pathways (unpublished data).

Thus, 19 immune-related mirs and 4 immune-related NanoString genes were used to cluster samples within C-08. Clustering by K means was used to arrange subsets of patients into 3, 4, 5 or 6 clusters, and the benefit from bevacizumab was evaluated in each cluster. Clustering into either 4 or 5 different clusters identified a subset with a significant benefit from bevacizumab. The best results occurred when cases were clustered into 4 groups. One cluster in this analysis showed significant benefit from bevacizumab (HR =0.484, p=0.006); two other clusters had a trend toward harm, and another showed no effect (Figure 7).

Summary of Work Completed

We have completed mir expression profiling of 1812 patients enrolled in NSABP clinical trial C-08. While we have not yet completed analyzing the discovery cohort, we do have encouraging results suggesting that a subset of patients may get benefit from bevacizumab. Given new regulatory requirements for the development and validation of clinical signatures, it was not possible to develop and validate a signature within the time frame allowed within this grant. In the best of circumstances, it was a bit naïve to propose to develop and *validate* prognostic and predictive signatures. However, we have made significant progress and will continue to optimize models for prognosis and prediction of bevacizumab benefit.

We have also discovered that patients with dMMR tumors benefitted from adjuvant bevacizumab in NSABP C-08 and reported these results in a presentation entitled *DNA mismatch repair deficiency and benefit from adjuvant bevacizumab in NSABP C-08: Molecular profiling results* at the *Markers in Cancer* meeting in Hollywood FL 2012.

Pogue-Geile KL, Tanaka N, Gavin PG, Yothers G, Colangelo LH, Kim C, Kim S, Fumagalli D, Taniyama Y, Sharif S, O'Connell MJ, Wolmark N, Paik S. DNA mismatch repair deficiency and benefit from adjuvant bevacizumab in NSABP C-08: Molecular profiling results. *J Clin Oncol /Markers in Cancer* ASCO-EORTC-NCI 2012; 30(30 Suppl). Abstr 55
Abstract Link: http://meetinglibrary.asco.org/content/103629-127

Table 1. Prognostic mirs

Mirs	Continuous P mean	Categorical P quar1 mean	Categorical P qurar2 mean	Categorical P quar3 mean
1	0.012	0.443	0.245	0.291
2	0.05	0.52	0.561	0.628
3	0.027	0.069	0.324	0.681
4	0.016	0.439	0.45	0.488
5	0.023	0.563	0.587	0.495
6	0.036	0.325	0.288	0.468
7	0.041	0.595	0.493	0.683
8	0.617	0.037	0.07	0.596
9	0.53	0.024	0.123	0.334
10	0.562	0.049	0.037	0.375
11	0.643	0.011	0.036	0.127
12	0.441	0.182	0.036	0.535

Table 2. Assessment of mirs for prediction of bevacizumab benefit or harm

Mirs	Continuous pmean	Categorical P quar1 mean	Categorical P qurar2 mean	Categorical P quar3 mean
1	0.708	0.078	0.426	0.624
2	0.665	0.037	0.07	0.596
3	0.731	0.054	0.408	0.72
4	0.203	0.069	0.324	0.681
5	0.598	0.07	0.233	0.508
6	0.72	0.024	0.123	0.334
7	0.579	0.049	0.037	0.375
8	0.475	0.011	0.036	0.127
9	0.484	0.089	0.126	0.547
10	0.663	0.066	0.136	0.578
11	0.605	0.269	0.099	0.483
12	0.569	0.279	0.097	0.615
13	0.657	0.17	0.059	0.509
14	0.66	0.467	0.054	0.665
15	0.591	0.182	0.036	0.535
16	0.683	0.542	0.241	0.07
17	0.333	0.125	0.417	0.072
18	0.514	0.505	0.247	0.082

Table 3. Overlap in cluster groups identified by micro RNA (MIR) and NanoString using non-negative matrix factorization (NMF)

		NanoString		
		1Benefit	2 None	3 Harm
MIR	1 Benefit	295	212	44
	2 None	42	54	144
	3 Harm	21	22	119

Figure 1. Kaplan-Meier plots of patients in National Surgical Adjuvant Breast and Bowel Project colon clinical trial number 8 (C-08). A - Patients with deficient mismatch repair (dMMR) tumors. B - Patients with proficient mismatch repair (pMMR) tumors. Patients were treated with a modified regimen consisting of fluorouracil, leucovorin, and oxaliplatin (mFF6) alone or with the addition of bevacizumab (+bev)

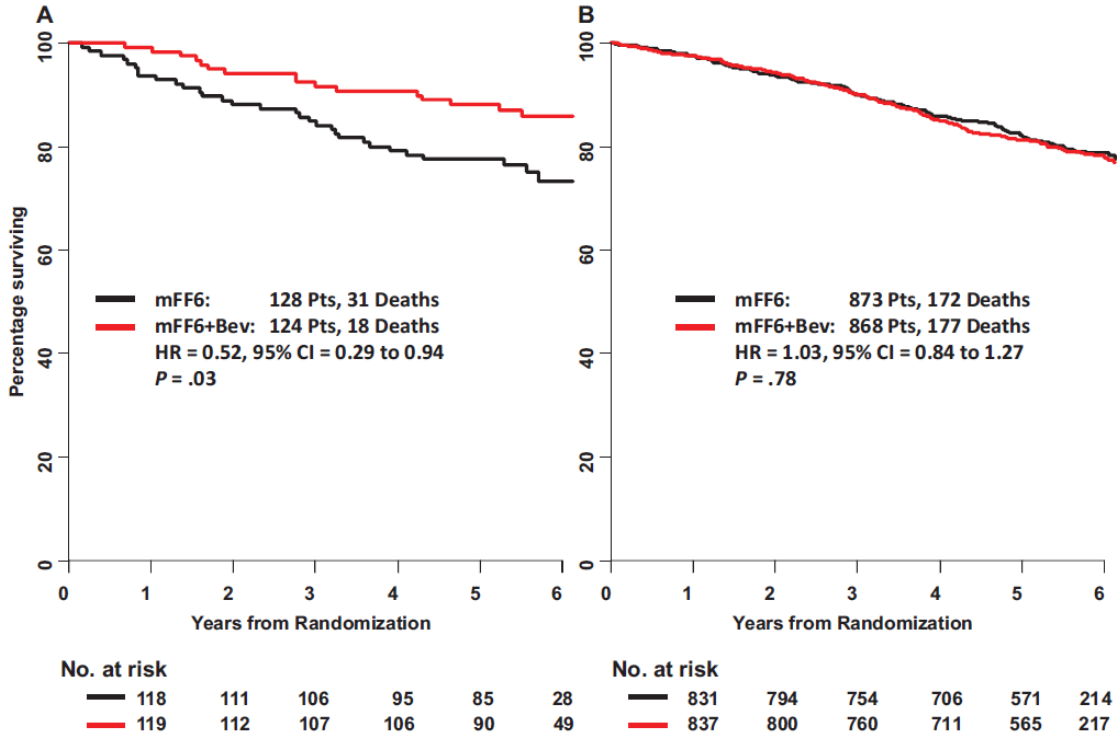


Figure 2. Overall survival (OS) by BRAF and MSI. Kaplan-Meier plots of patients in the National Surgical Adjuvant Breast and Bowel Project colon clinical trial number 8 (C-08) are shown for patients with BRAF mutations (BRAF-Mut) on the left side of the figure and for patients with tumors which were wild type for BRAF (BRAF-Wt) on the right. Patients with tumors which were microsatellite unstable (MS-Instability) are shown in the top 2 plots and patients with tumors that were microsatellite stable (MS-stable) are plotted on the bottom 2 plots. Patients were treated with a modified regimen consisting of fluorouracil, leucovorin, and oxaliplatin (mFF6) alone or with the addition of bevacizumab (mFF6+bev) MMR mismatch repair status Hazard ratios (HR) are indicated.

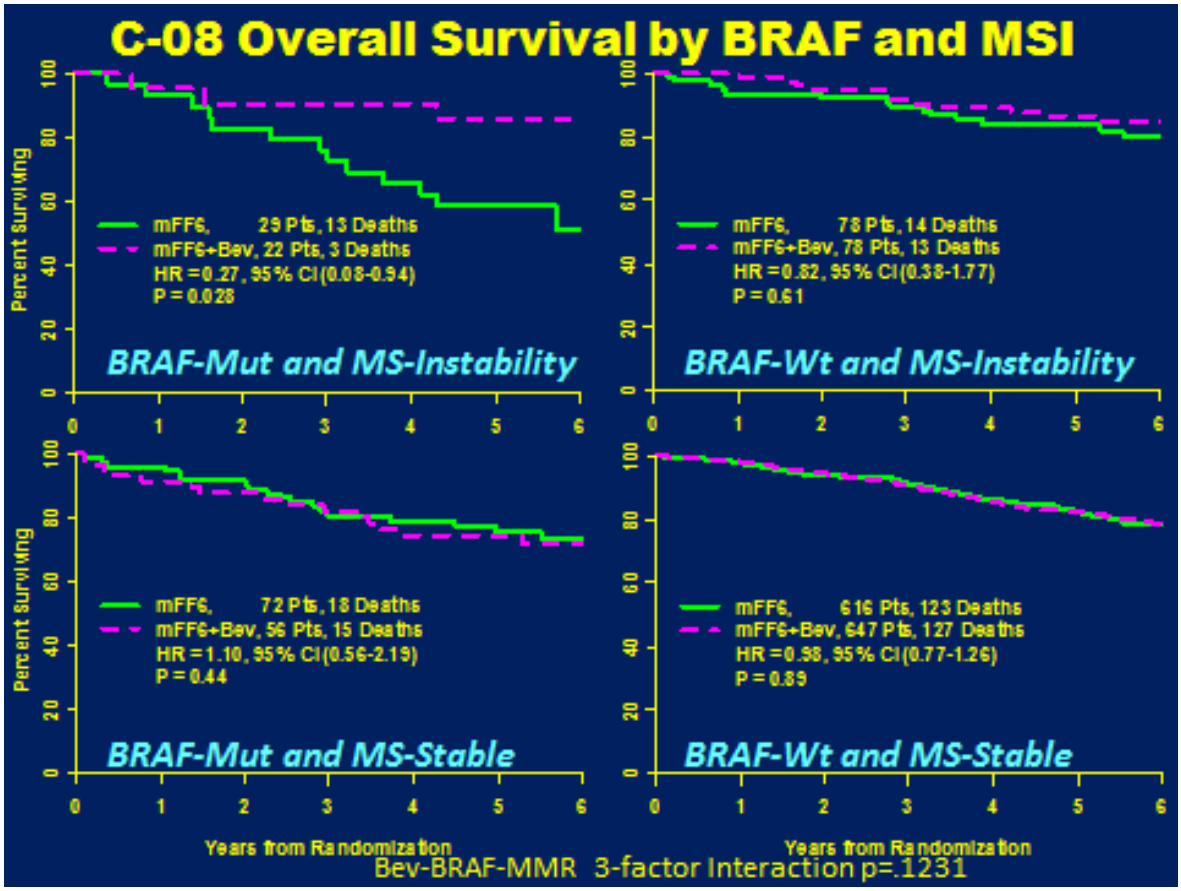


Figure 3. Clustering of micro RNA (MIR) data (mir cluster groups 1, 2 , and 3). Hazard ratios (HR) are indicated.

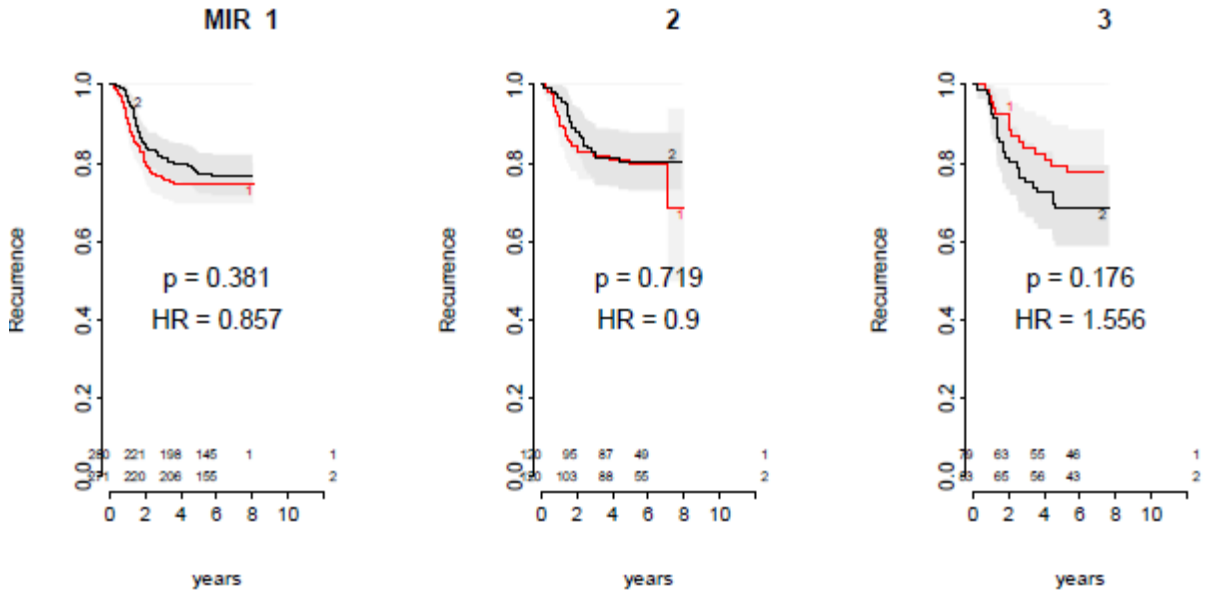


Figure 4. Prognostic and predictive p values for individual micro RNA (mirs) and NanoString genes (nano).

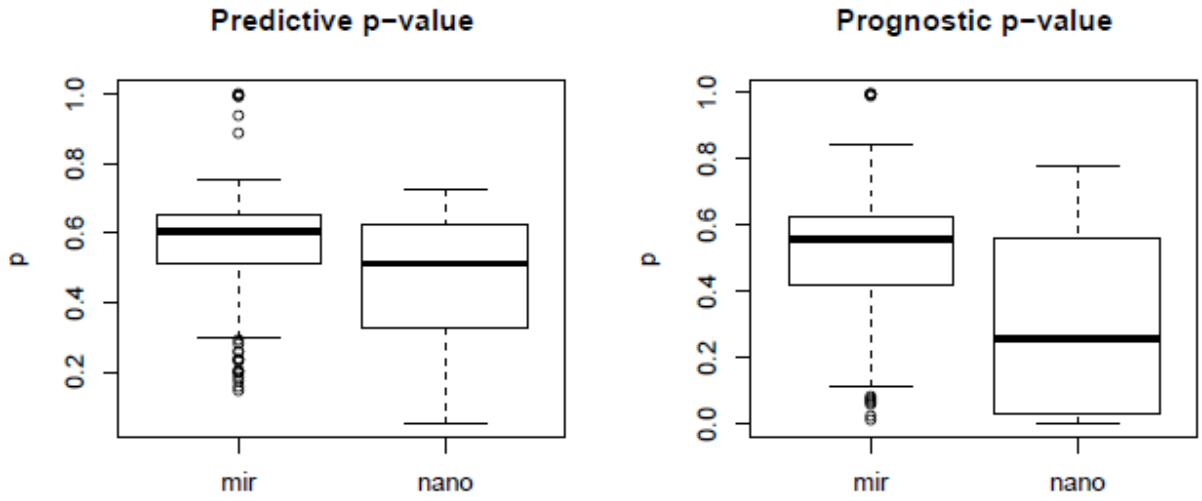


Figure 5. Prognostic models built with SuperPC Clustering by micro RNAs (MIRs) and genes interrogated with Nanostring (Nano genes). Hazard ratios (HR) are indicated.

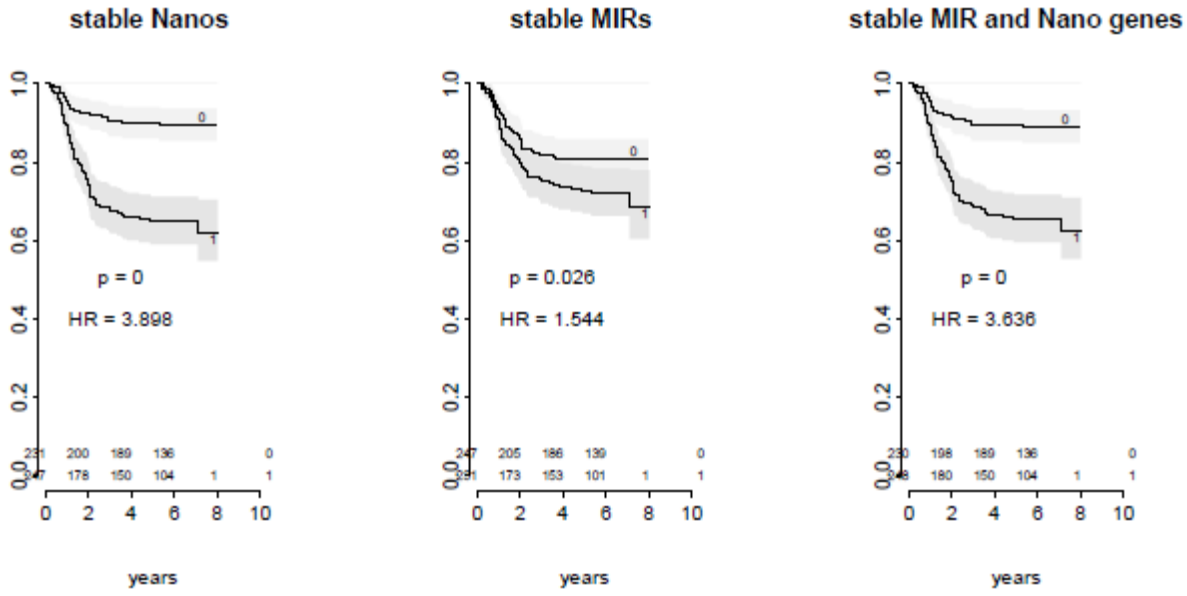


Figure 6. Clustering of National Surgical Adjuvant Breast and Bowel Project colon clinical Trial number 8(C-08) using NanoString data. Hazard ratios (HR) are indicated.

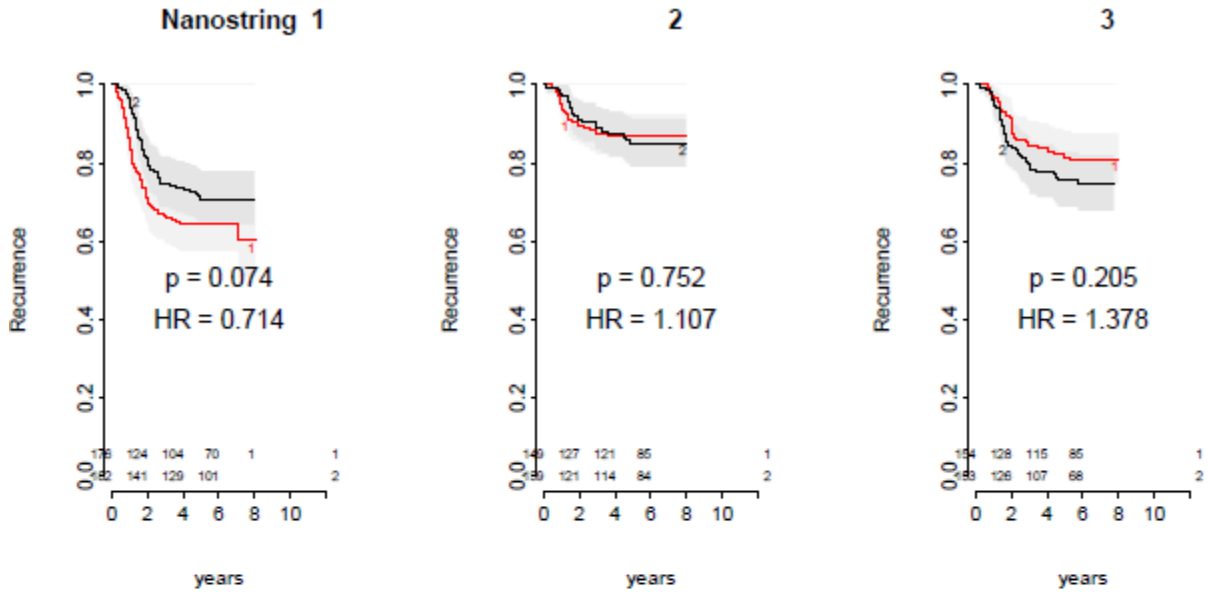
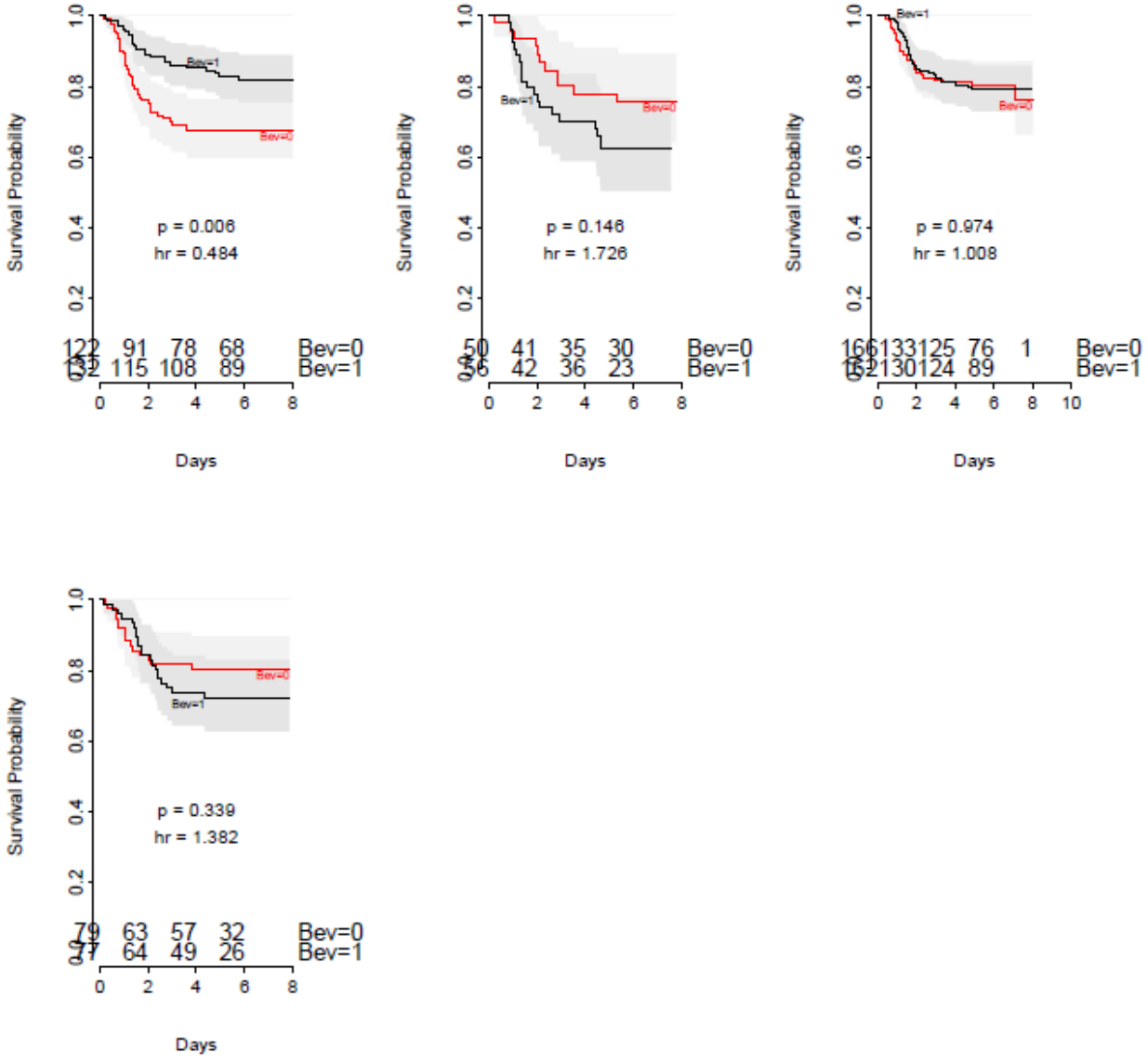


Figure 7. Clustering of samples using immune-related micro RNAs (mirs) and genes. Patients were treated without the addition of bevacizumab (bev=0) or with the addition (bev=1). Hazard ratios (hr) are indicated.



18. Extent of Clinical Activities Initiated and Completed. Items 18(A) and 18(B) should be completed for all research projects. If the project was restricted to secondary analysis of clinical data or data analysis of clinical research, then responses to 18(A) and 18(B) should be "No."

18(A) Did you initiate a study that involved the testing of treatment, prevention or diagnostic procedures on human subjects?

_____Yes
__X___No

18(B) Did you complete a study that involved the testing of treatment, prevention or diagnostic procedures on human subjects?

_____Yes
__X___No

If "Yes" to either 18(A) or 18(B), items 18(C) – (F) must also be completed. (Do NOT complete 18(C-F) if 18(A) and 18(B) are both "No.")

18(C) How many hospital and health care professionals were involved in the research project?

_____Number of hospital and health care professionals involved in the research project

18(D) How many subjects were included in the study compared to targeted goals?

_____Number of subjects originally targeted to be included in the study
_____Number of subjects enrolled in the study

Note: Studies that fall dramatically short on recruitment are encouraged to provide the details of their recruitment efforts in Item 17, Progress in Achieving Research Goals, Objectives and Aims. For example, the number of eligible subjects approached, the number that refused to participate and the reasons for refusal. Without this information it is difficult to discern whether eligibility criteria were too restrictive or the study simply did not appeal to subjects.

18(E) How many subjects were enrolled in the study by gender, ethnicity and race?

Gender:
_____Males
_____Females
_____Unknown

Ethnicity:
_____Latinos or Hispanics

_____Not Latinos or Hispanics
_____Unknown

Race:
_____American Indian or Alaska Native
_____Asian
_____Blacks or African American
_____Native Hawaiian or Other Pacific Islander
_____White
_____Other, specify:_____
_____Unknown

18(F) Where was the research study conducted? (List the county where the research study was conducted. If the treatment, prevention and diagnostic tests were offered in more than one county, list all of the counties where the research study was conducted.)

19. Human Embryonic Stem Cell Research. Item 19(A) should be completed for all research projects. If the research project involved human embryonic stem cells, items 19(B) and 19(C) must also be completed.

19(A) Did this project involve, in any capacity, human embryonic stem cells?
_____Yes
__X___ No

19(B) Were these stem cell lines NIH-approved lines that were derived outside of Pennsylvania?
_____Yes
_____ No

19(C) Please describe how this project involved human embryonic stem cells:

20. Articles Submitted to Peer-Reviewed Publications.

20(A) Identify all publications that resulted from the research performed during the funding period and that have been submitted to peer-reviewed publications. Do not list journal abstracts or presentations at professional meetings; abstract and meeting presentations should be listed at the end of item 17. **Include only those publications that acknowledge the Pennsylvania Department of Health as a funding source** (as required in the grant agreement). List the title of the journal article, the authors, the name of the peer-reviewed publication, the month and year when it was submitted, and the status of publication (submitted for publication, accepted for publication or published.). Submit an electronic copy of each publication or paper submitted for publication, listed in the table, in a PDF version 5.0.5 (or greater) format, 1,200 dpi. Filenames for each publication should include

the number of the research project, the last name of the PI, and an abbreviated title of the publication. For example, if you submit two publications for Smith (PI for Project 01), one publication for Zhang (PI for Project 03), and one publication for Bates (PI for Project 04), the filenames would be:

Project 01 – Smith – Three cases of isolated
Project 01 – Smith – Investigation of NEB1 deletions
Project 03 – Zhang – Molecular profiling of aromatase
Project 04 – Bates – Neonatal intensive care

If the publication is not available electronically, provide 5 paper copies of the publication.

Note: The grant agreement requires that recipients acknowledge the Pennsylvania Department of Health funding in all publications. Please ensure that all publications listed acknowledge the Department of Health funding. If a publication does not acknowledge the funding from the Commonwealth, do not list the publication.

Title of Journal Article:	Authors:	Name of Peer-reviewed Publication:	Month and Year Submitted:	Publication Status (check appropriate box below):
1. Defective mismatch repair and benefit from bevacizumab for colon cancer: findings from NSABP C-08	Pogue-Geile K, Yothers G, Taniyama Y, Tanaka N, Gavin P, Colangelo L, Blackmon N, Lipchik C, Kim SR, Sharif S, Allegra C, Petrelli N, O'Connell MJ, Wolmark N, Paik S.	J Natl Cancer Inst.	December, 2012	☐Submitted ☐Accepted ☒Published
2.				☐Submitted ☐Accepted ☐Published

20(B) Based on this project, are you planning to submit articles to peer-reviewed publications in the future?

Yes___X_____ No_____

If yes, please describe your plans:

We are still developing signatures for prognosis and for prediction of bevacizumab benefit. Our current plan is to publish once we have validated or not validated the prognostic and or predictive signature. Such signatures, if validated, would have dramatic impact on clinical practice.

21. Changes in Outcome, Impact and Effectiveness Attributable to the Research Project. Describe the outcome, impact, and effectiveness of the research project by summarizing its impact on the incidence of disease, death from disease, stage of disease at time of diagnosis, or other relevant measures of outcome, impact or effectiveness of the research project. If there were no changes, insert "None"; do not use "Not applicable." Responses must be single-spaced below, and no smaller than 12-point type. DO NOT DELETE THESE INSTRUCTIONS. There is no limit to the length of your response.

The relevant measures for this research project include the development of methods for expression analysis of mirs in colon cancer and the development and validation of models using mir data to improve colon cancer treatment. The research project was effective in establishing new high through-put protocols for the isolation of RNA which includes robust yields of both mirs and mRNA using routinely processed formalin-fixed paraffin-embedded tissues. Also a moderately high throughput procedure for the expression profiling of these mirs was developed.

22. Major Discoveries, New Drugs, and New Approaches for Prevention Diagnosis and Treatment. Describe major discoveries, new drugs, and new approaches for prevention, diagnosis and treatment that are attributable to the completed research project. If there were no major discoveries, drugs or approaches, insert "None"; do not use "Not applicable." Responses must be single-spaced below, and no smaller than 12-point type. DO NOT DELETE THESE INSTRUCTIONS. There is no limit to the length of your response.

We discovered that patients in NSABP clinical trial C-08 whose tumors were defective for mismatch repair were more likely to receive benefit than patients whose tumors were proficient for MMR. This is an ad hoc observation and remains to be validated in another clinical trial; we are currently pursuing a collaboration to do this. If this observation were to be validated, it may represent an important new and effective treatment option for a subset of colon cancer patients. Importantly, we have also proposed a testable hypothesis to describe our observations and may lead to important new approaches for assessing drug efficacy.

We have developed a promising model for identifying patients who received benefit from bevacizumab in our discovery cohort by selecting mirs and genes that play a role in the immune system. While we are still exploring other models and testing the current one to ensure that the model is the best possible, these early results suggest that it may be possible to identify a group of Stage II and III colon cancer patients who would receive meaningful benefit from bevacizumab.

23. Inventions, Patents and Commercial Development Opportunities.

23(A) Were any inventions, which may be patentable or otherwise protectable under Title 35 of the United States Code, conceived or first actually reduced to practice in the performance of work under this health research grant? Yes__X_____ No_____

If "Yes" to 23(A), complete items a – g below for each invention. (Do NOT complete items a - g if 23(A) is "No.")

a. Title of Invention: Defective Mismatch Repair and benefit from bevacizumab for colon cancer

b. Name of Inventor(s): Katherine Lea Pogue-Geile, and Soonmyung Paik

c. Technical Description of Invention (describe nature, purpose, operation and physical, chemical, biological or electrical characteristics of the invention):

The purpose of the invention is to identify a group of Stage II or III colon cancer patients who derived benefit from treatment with bevacizumab when added to standard chemotherapy, such as FLOX (5-FU plus leucovorin plus oxaliplatin). Bevacizumab did not show benefit in the entire cohort of patients who were enrolled in NSABP C-08, which was the clinical trial conducted to test the usefulness of adding bevacizumab to standard chemotherapy. Through a routine post hoc analysis, we discovered that patients with tumors that were defective for mismatch repair (dMMR) received significant benefit from bevacizumab. The assay that defined dMMR tumors was based on immunohistochemistry of two mismatch repair protiens (MLH1 and MSH2). Any cases that showed negative staining of one of the two proteins in the tumor cells in the presence of positive staining in the surrounding normal cells were classified as dMMR while others were classified as MMR proficient (pMMR). These two IHC markers provide both a sensitive and specific alternative to microsatellite instability in detecting DNA MMR defects (Lindor etal., J Clin Oncol 2001;20(4):1043-8. Details of our work has been published (Pogue-Geile J Natl Cancer Inst. 2013 Jul 3;105(13):989-92).

d. Was a patent filed for the invention conceived or first actually reduced to practice in the performance of work under this health research grant?
Yes__X___ No____

If yes, indicate date patent was filed: 05/24/2013

e. Was a patent issued for the invention conceived or first actually reduced to practice in the performance of work under this health research grant?
Yes_____ No_X_
If yes, indicate number of patent, title and date issued:
Patent number:
Title of patent:
Date issued:

f. Were any licenses granted for the patent obtained as a result of work performed under this health research grant? Yes_____ No__X

If yes, how many licenses were granted?_____

g. Were any commercial development activities taken to develop the invention into a commercial product or service for manufacture or sale? Yes___ No__X

If yes, describe the commercial development activities:

23(B) Based on the results of this project, are you planning to file for any licenses or patents, or undertake any commercial development opportunities in the future?

Yes___X_(it depends)_____ No_____

If yes, please describe your plans:

If we identify and validate a gene signature that is able to identify colon cancer patients who receive benefit from bevacizumab when added to FLOX, then we will pursue a patent and would possibly seek a partner for commercial development. In addition, if we found a clinically useful prognostic signature, we would also pursue a patent.

24. Key Investigator Qualifications. Briefly describe the education, research interests and experience and professional commitments of the Principal Investigator and all other key investigators. In place of narrative you may insert the NIH biosketch form here; however, please limit each biosketch to 1-2 pages. *For Nonformula grants only – include information for only those key investigators whose biosketches were not included in the original grant application.*

Dr. Pogue-Geile's qualifications that make her well suited to this investigation include a sound foundation of molecular analysis gained as a Ph.D. student and graduate of the University of Chicago in the Department of Cell and Molecular Biology. Since leaving Chicago she has gained extensive research experience in utilizing a broad range of molecular techniques to explore the molecular mechanisms of cancer. This is reflected by more than 24 publications, 14 of these as first or senior author, in prestigious journals including *Blood, Molecular and Cellular Biology, Cancer Research, American Journal of Human Genetics, PLOS Medicine,* and *Journal of Clinical Oncology.* Since 2005, she has worked at the NSABP developing methods for mutation and gene expression profiling of nucleic acids isolated from formalin-fixed, paraffin-embedded material. As the Assistant Director of Molecular Profiling at the NSABP Foundation, Inc., she is well suited to use her training and experience in molecular biology to develop assays and explore their clinical meaning as a means to improve outcomes of colon and breast cancer.

1. Allegra CJ, Yothers G, O'Connell MJ, et al. Bevacizumab in Stage II-III Colon Cancer: 5-Year Update of the National Surgical Adjuvant Breast and Bowel Project C-08 Trial. *Journal of Clinical Oncology* 2013; **31**(3): 359-64.
2. Xi Y, Nakajima G, Gavin E, et al. Systematic analysis of microRNA expression of RNA extracted from fresh frozen and formalin-fixed paraffin-embedded samples. *Rna* 2007; **13**(10): 1668-74.
3. Gavin P, Colangelo LH, Fumagalli D, et al. Mutation Profiling and Microsatellite Instability in Stage II and III Colon Cancer: An Assessment of their Prognostic and Oxaliplatin Predictive Value. *Clinical Cancer Research* 2012.
4. Pogue-Geile K, Yothers G, Taniyama Y, et al. Defective Mismatch Repair and Benefit from Bevacizumab for Colon Cancer: Findings from NSABP C-08. *J Natl Cancer Inst* 2013; **105**(13): 989-92.
5. Saeterdal I, Bjorheim J, Lislerud K, et al. Frameshift-mutation-derived peptides as tumor-specific antigens in inherited and spontaneous colorectal cancer. *Proc Natl Acad Sci U S A* 2001; **98**(23): 13255-60.
6. Drescher K, Sharma P, Watson P, Gatalica Z, Thibodeau S, Lynch H. Lymphocyte recruitment into the tumor site is altered in patients with MSI-H colon cancer. *Familial Cancer* 2009; **8**(3): 231-9.
7. Nosho K, Baba Y, Tanaka N, et al. Tumour-infiltrating T-cell subsets, molecular changes in colorectal cancer, and prognosis: cohort study and literature review. *J Pathol* 2010; **222**(4): 350-66.
8. Bindea G, Mlecnik B, Fridman WH, Galon J. The prognostic impact of anti-cancer immune response: a novel classification of cancer patients. *Seminars in immunopathology* 2011; **33**(4): 335-40.
9. Frey DM, Droeser RA, Viehl CT, et al. High frequency of tumor-infiltrating FOXP3+ regulatory T cells predicts improved survival in mismatch repair-proficient colorectal cancer patients. *International Journal of Cancer* 2010; **126**(11): 2635-43.
10. Bellamy WT, Richter L, Sirjani D, et al. Vascular endothelial cell growth factor is an autocrine promoter of abnormal localized immature myeloid precursors and leukemia progenitor formation in myelodysplastic syndromes. *Blood* 2001; **97**(5): 1427-34.
11. Gabrilovich D, Ishida T, Oyama T, et al. Vascular Endothelial Growth Factor Inhibits the Development of Dendritic Cells and Dramatically Affects the Differentiation of Multiple Hematopoietic Lineages In Vivo. *Blood* 1998; **92**(11): 4150-66.
12. Ohm JE, Gabrilovich DI, Sempowski GD, et al. VEGF inhibits T-cell development and may contribute to tumor-induced immune suppression. *Blood* 2003; **101**(12): 4878-86.
13. Gabrilovich DI, Ishida T, Nadaf S, Ohm JE, Carbone DP. Antibodies to vascular endothelial growth factor enhance the efficacy of cancer immunotherapy by improving endogenous dendritic cell function. *Clin Cancer Res* 1999; **5**(10): 2963-70.
14. Mlecnik B, Bindea G, Pages F, Galon J. Tumor immunosurveillance in human cancers. *Cancer Metastasis Rev* 2011; **30**(1): 5-12.
15. Mlecnik B, Tosolini M, Kirilovsky A, et al. Histopathologic-based prognostic factors of colorectal cancers are associated with the state of the local immune reaction. *J Clin Oncol*; **29**(6): 610-8.
16. Nosho K, Baba Y, Tanaka N, et al. Tumour-infiltrating T-cell subsets, molecular changes in colorectal cancer, and prognosis: cohort study and literature review. *The Journal of Pathology*; **222**(4): 350-66.