AWARD NUMBER: W81XWH-16-1-0543

TITLE: Fusion Genes Predict Prostate Cancer Recurrence

PRINCIPAL INVESTIGATOR: David F. Jarrard

RECIPIENT: University of Wisconsin System
Madison, WI 53715

REPORT DATE: October 2018

TYPE OF REPORT: Annual

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

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
October 2018	Annual	15 Sep 2017 - 14 Sep 2018

4. TITLE AND SUBTITLE

Fusion Genes Predict Prostate Cancer Recurrence

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-16-1-0543

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

David Jarrard

E-Mail: jarrard@urology.wiscx.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Wisconsin System
21 N Park St., Ste. 6401
Madison, WI 53715-1218

8. PERFORMING ORGANIZATION REPORT NUMBER

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

U.S. Army Medical Research and Materiel 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

This Impact Award between the University of Pittsburgh, Stanford University and University of Wisconsin at Madison seeks to provide definitive clinical validation of low copy number transcript fusions as prognostic markers in clinically localized prostate cancer. The University of Wisconsin site is charged with providing formalin fixed paraffin embedded samples for independent validation of markers discovered at University of Pittsburgh. To that end, we have shipped 437 samples for assay testing and have accumulated over 750 FFPE blocks of radical prostatectomy specimens with detailed clinical follow-up. We are performing an update of the clinical follow-up on these patients to maximize length of clinical follow-up.

The 3 centers have been communicating electronically and via phone conference. An in person meeting is planned in the spring to review data for the final report. The University of Pittsburgh has been addressing issues with assay artifacts. We will increase the number of samples accumulated to get to our goals. We will be sending specimens blinded to the recurrence data to the Pittsburgh site for the remaining sets. We have agreed to allow the Stanford and Wisconsin sites to receive data from Pittsburgh to perform correlation with clinical data and correlation with outcomes at the outside sites.

15. SUBJECT TERMS
Prostate cancer, prognosis, gene fusions, fusion transcripts

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	14	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Page

Major Task 1 : We will conduct analysis of MAN2A1-FER, SLC45A2-AMACR, TRMT11-GRIK2, MTOR-TP53BP1, LRRC59-FLJ60017, CCNH-C5orf30, KDM4-AC011523.2, TMEM135-CCDC67 on 5106 prostate cancer samples collected from University of Pittsburgh, Stanford University and University of Wisconsin Madison. We will first establish prostate cancer recurrence model and short PSADT prediction models either by fusion gene status alone or in combination with nomogram based on the cohort from 600 radical prostatectomy samples from UPMC. This model will be locked in and tested on cohorts from University of Pittsburgh, Stanford University and University of Wisconsin. The prediction accuracy, sensitivity and specificity within each cohort will be evaluated.

Subtask 1: In the first 3 months of the funded period, we plan to establish this test in the CLIA certified laboratory at the University of Pittsburgh Medical Center. Fifty-six FFPE samples that were shown to be positive for at least one fusion transcripts in the matched frozen tissues. These FFPE samples had been tested in non CLIA certified laboratory, and achieved 98.9% sensitivity and 100% specificity. We will repeat the same tests on these samples in CLIA certified laboratories. All PCR products will be analyzed through Sanger's sequencing to confirm the authenticity of the fusion products. In addition, all fusion minigene RNA templates will be serially diluted. TAQMAN QRT-PCR will be performed to evaluate the sensitivity of the test. Detection threshold will be obtained. Random selection of 600 prostate cancer samples with definitive clinical outcomes will be carried out in UPMC campus. TAQMAN QRT-PCR on β-actin will be used as RNA quality control. For sites 2 and 3, all relevant institutional review board exempt protocols will be secured and approved.

Progress: We have procured the CLIA certified lab space in the beginning of the funded period. To accommodate the reality of formalin-fixed and paraffin-embedded tissues, we have designed a set of new primers and Taqman PCR probes for highly fragmented RNA species. These sets of primers and probes were subsequently tested and validated on synthetic mini-fusion genes of MAN2A1-FER, TRMT11-GRIK2, MTOR-TP53BP1, CCNH-C5orf30, KDM4-AC011523.2, SLC45A2-AMACR, TMEM135-CCDC67, and LRRC59-FLJ60017. The probe and primers for β-actin were also revised to accommodate a shorter RNA fragment. The analyses showed that these assays detect as low as 600-1000 molecules of these fusion transcripts. We then analyzed 56 FFPE samples whose frozen counterparts have been previous found to contain at least one fusion gene using these sets of probes and primers. All samples that were positive for these fusion genes were also positive in the new Taqman qRT-PCR assays. The positive match rate is 100%. All participating institutes, including University of Pittsburgh, Stanford University and University of Wisconsin Madison, had obtained the institutional approval for the exempt protocols.

Subtask 2: From month 4-9 of the first funded year, we will perform TAQMAN QRT-PCR and Sanger's sequencing on a randomly selected cohort of 600 samples from phase 1 that have at least 5 years clinical follow-up. These tests will be performed in CLIA certified laboratory of University of Pittsburgh. The prediction models of PCa recurrence and PSADT mentioned will be developed based on this large number of samples. For sites 2 and 3, the first 300 prostate cancer cases from each site will be selected and evaluated for sufficient materials for the assay.

Progress: To create a training set, we performed Taqman qRT-PCR using the primers and probes as mentioned from above on 271 samples from University of Pittsburgh, 155 samples from University of Wisconsin Madison, and 150 samples from Stanford University. The results show surprisingly high positive rate of SLC45A2-AMACR in Stanford and Wisconsin cohort, reaching 96% and 92.6% respectively. Among these fusion genes, the lowest frequent

Table 1 Positive rate of fusion in prostate cancers

Cohort	MAN2A1/ FER	TRMT11/ GRIK2	MTOR/ TP53BP1	CCNH/ C5orf30	KDM4B/ AC011523.2	SLC45A2/ AMACR	TMEM135/ CCDC67	LRRC59/ FLJ60017
UPMC	13% (60)	25.8% (119)	2.8% (13)	33.4% (154)	0.4% (2)	50.1% (234)	1% (5)	3.4% (16)
Stanford	18% (9)	20% (10)	10% (5)	12% (6)	4% (2)	96% (48)	6% (3)	22% (11)
UWisc	19% (31)	12.9% (21)	4.3% (7)	76.7% (125)	9.2% (15)	92.6% (151)	0.6% (1)	26% (43)

Table 2
The cutoffs(and OR) of each fusion gene in each cohort

Cohort	MAN2A1/FER	MAN2A1/FER-actin	TRMT11/GRK2	MTOR/TP53BP1	CCNH/C5Orf30	KDM4/AC011523.2	SL45A2/AMACR	TMEM135/CCDC67
UPMC	32(26.3)	0 (25.9)	43(5.53)	42(inf)	39(0.12)	44(inf)	34(1.57)	47(1.54)
Wisconsin	35(14.81)	3(7.57)	42(inf)	40(23.6)	38(0.49)	40(1.5)	31(1.7)	N/A
Stanford	39(1.71)	0(0.34)	39(4.03)	39(inf)				

one is TMEM135-CCDC67: A total of 8 samples were found positive. In addition, high positive rate of CCNH-C5orf30 was also found in the prostate cohort from University of Wisconsin. In general, the rates of fusion gene positive samples are comparable among the 3 cohorts (table 1). Subsequent analyses showed that MAN2A1-FER (or normalized MANA1-FER), TRMT11-GRIK2, and mTOR-TP53BP1 gene fusions have the highest odd ratios for predicting the recurrence of prostate cancer for UPMC and University of Wisconsin cohorts (Table 2).

Table 3

Model	Fusion genes only					2x2 table		
	Sensitivity	Specificity	Youden	Accuracy	AUC		Recurrent (n=107)	Non-Recurrent (N=164)
RF	0.80	0.81	0.61	0.81	0.862	Positive	TP=86	FP=31
	Top 6, cutoff=0.2					Negative	FN=21	TN=133
SVM	0.71	0.87	0.58	0.81	0.77	Positive	TP=76	FP=21
	Top 4, cutoff=0.2					Negative	FN=31	TN=143
LDA	0.71	0.88	0.59	0.81	0.85	Positive	TP=76	FP=20
	Top 6, cutoff=0.4					Negative	FN=31	TN=144

To establish a prediction model, we combined top 6 fusion genes that have prediction power to construct classification models to predict prostate cancer recurrence. As shown in table 3, all three models (Random Forest, Support vector machine and Linear discriminant analysis) yielded very similar accuracy: 81%, even though the specificity and sensitivity may vary. When combined with gleason's score and TNM pathology staging, the accuracy improves to 84-86%.

When the same models were applied to the dataset from University of Wisconsin, the accuracy rate yielded 75-84%. Interestingly, when combined with Gleason's grade and pathology TNM staging, the accuracy rate improved to 88-90%. However, the same model fared worse in Stanford data set: 67-68% accuracy was found. Combination with fusion genes, Gleason's grade and pathology TNM staging improves the accuracy to 75%.

Table 4

Model	Gleason and TNM stage + fusion genes					2x2 table		
	Sensitivity	Specificity	Youden	Accuracy	AUC		Recurrent (n=208)	Non-Recurrent (N=368)
RF	0.79	0.82	0.61	0.81	0.86	Positive	TP=164	FP=67
	top 5 fusion genes, cutoff=0.2					Negative	FN=44	TN=301
SVM	0.75	0.83	0.59	0.81	0.84	Positive	TP=157	FP=61
	Top 6 fusion genes, cutoff=0.2					Negative	FN=51	TN=307
LDA	0.77	0.85	0.61	0.82	0.87	Positive	TP=160	FP=57
	Top 5 fusion genes, cutoff=0.4					Negative	FN=48	TN=311

When all data were pooled, and 10 fold cross validation was performed. The accuracy of fusion gene prediction rate is 76-77%. When combined fusion genes, Gleason's grading and TNM staging, the accuracy is improved to 81-82% (table 4). In contrast, if prediction is only relied on Gleason's grade and TNM staging, the prediction accuracy is 74-76% across all three data set. As a result, we concluded that fusion contains independent prediction value and can assist in predicting the clinical outcomes of prostate cancer.

Subtask 3: From month 10 of the first year to the end of year 3 of the funded period, we will validate predictive models based on the fusion transcript panel and clinical and pathological parameters on independent datasets from the University of Pittsburgh, University of Wisconsin and Stanford University. The UPMC cohort includes up to 1900 well-annotated and -followed radical prostatectomy samples. University of Wisconsin will provide up to 1000 samples for the analysis. All samples will have at least 5 years of clinical follow-up at the end of year 3. The Stanford cohort will select up to 1500 samples from 3100 cases with a minimum of 5 years

of follow-up. Over 70% of these cases will have 7-15 years clinical follow-up. The model established in phase 2 and our preliminary results will be used to assess the risk of recurrence and short PSADT of each case. From month 10 to 36, TAQMAN QRT-PCR on selected genes will be performed on 100 random selected samples in sites 1, 2 and 3 simultaneously to assess the reproducibility of the assays. Each case will have 3 separate loci of prostate cancer. All loci will be tested for the heterogeneity of the prostate cancer in terms of fusion gene distribution.

Progress: Investigators in University of Pittsburgh, Stanford University and University of Wisconsin Madison, are selecting additional samples for validation of the prediction model. In particular, Dr. Nelson in University of Pittsburgh had already selected additional 250 samples for the purpose of a testing cohort. Dr. Brooks have submitted PCa 150 samples, while Dr. Jarrard submitted 166 samples. Both Drs. Brooks and Jarrard are working to select more samples for the testing set. We plan to finish total 1500 samples by the end of third year of funded period.