AWARD NUMBER: W81XWH-13-1-0477

TITLE: microRNA in Prostate Cancer Racial Disparities and Aggressiveness

PRINCIPAL INVESTIGATOR: Cathryn Bock

CONTRACTING ORGANIZATION: Wayne State University
Detroit, MI 48202

REPORT DATE: December 2018

TYPE OF REPORT: FINAL

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
Dec 2018	Final	30 Sep 2013 - 29 Sep 2018

4. TITLE AND SUBTITLE

MicroRNA in Prostate Cancer Racial Disparities and Aggressiveness

5a. CONTRACT NUMBER

5b. GRANT NUMBER

W81XWH-13-1-0477

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Cathryn Bock

E-Mail: bockc@wayne.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

WAYNE STATE UNIVERSITY (INC)
CAROLE BACH
5700 CASS AVE STE 4900
DETROIT MI 48202-3692

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

We examined the association of plasma miRNA levels and inherited polymorphisms in genes in the microRNA (miRNA) biogenesis pathway with prostate cancer aggressiveness among African American (AA) and European American (EA) men with prostate cancer (PCa) from the Karmanos Cancer Institute in Detroit, MI. We demonstrated that a previously unreported circulating miRNA signature consisting of two distinct combinations of miRNAs (miR-17/miR-192) and (miR-146a/miR-20a/miR26b) is capable of segregating aggressive and non-aggressive prostate cancer in both AA and EA patients. When we evaluated SNPs in genes involved in the miRNA biogenesis pathway, multiple SNPs in TARBP1, DROSHA and DICER1 were significantly associated with PCa aggressiveness in AA men. Fewer SNPs were significantly associated with PCa aggressiveness in EA men, and were primarily located in TARBP1 and DROSHA. Only 1 SNP was significant for both the AA and EA adjusted analyses: rs4920247, which is an intronic SNP located in TARBP1.

15. SUBJECT TERMS

prostate cancer, microRNA, racial disparities, African American, genetic polymorphisms, biochemical recurrence, epidemiology

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
U					USAMRMC
a. REPORT	b. ABSTRACT	c. THIS PAGE	UU	37	19b. TELEPHONE NUMBER *(include area code)*
U	U	U			

Table of Contents

1. Introduction

In the US, African American (AA) men are at 60% higher risk of developing prostate cancer (PCa) than European American (EA) men, and AA men are 2.4 times more likely to die from PCa than EA men.(1) The objective of this study is to identify novel genetic and epigenetic factors that might contribute significantly to racial/ethnic disparity in PCa risk and progression. We will examine the association of inherited polymorphisms in genes in the microRNA (miRNA) biogenesis pathway as well as the association of plasma miRNA levels with prostate cancer aggressiveness and biochemical recurrence (BCR) among AA and EA men with PCa from the Karmanos Cancer Institute (KCI) in Detroit, MI. Little is known about the role of microRNAs (miRNAs) and their biogenesis in prostate cancer (PCa), and less is understood about the possible race-specific role of miRNAs in PCa aggressiveness and outcomes. We hypothesized that polymorphisms in genes in the miRNA biogenesis pathway and plasma miRNA levels are potential prognostic indicators for PCa aggressiveness and/or outcome and that these associations may be linked to race. The specific aims of this project are to 1) determine the associations between polymorphisms in genes within the miRNA biogenesis pathway and (a) PCa aggressiveness and (b) biochemical recurrence in AA and EA men with PCa, 2) determine the associations between plasma levels of PCa-related miRNAs and PCa aggressiveness, and 3) determine the associations between genetic polymorphisms in miRNA biogenesis pathway genes and plasma levels of miRNAs known to regulate genes in prostate cancer pathways. To increase the potential for translating our results into disease management strategies, we included miRNAs with cell-line evidence of transcriptional regulation by miRNA promoter methylation and evidence of gene-expression regulation within prostate carcinogenic pathways. This project is built on a previously funded study of metabolic syndrome, PCa aggressiveness, and outcomes (CDMRP award W81XWH-09-1-0203, PI: Isaac Powell, MD). We used that study's infrastructure to enroll additional patients recently diagnosed PCa, ~60% of whom are AA and ~35% of whom have aggressive disease. During Years 1-3, 688 study subjects were enrolled, and 668 men completed surveys (*Tasks 1.a and 1.d: completed*), 586 participants provided blood samples for DNA extraction (*Task 1.b*), and 116 men had a plasma sample collected prior to initiation of prostate cancer treatment (*Task 1.c*). Following national trends, the number of prostate cancer patients eligible for the study declined in the clinic over the course of the study. Our enrollment rate among eligible men was ~95%, with no differences by race. Because many miRNAs are regulated by promoter methylation, they are potential targets for treatment with demethylating agents to prevent or slow PCa carcinogenesis;(2,3) target miRNAs may vary by race. Identifying risk profiles of men who may benefit from such treatment, based on race, inherited genotypes and/or plasma miRNA levels will provide momentum for developing the field of personalized medicine.

2. Keywords

prostate cancer, microRNA, racial disparities, African American, genetic polymorphisms, biochemical recurrence, epidemiology

3. Overall Project Summary

Current Objectives

Since final recruitment and data collection activities were completed near the end of year 3, final DNA extractions from blood samples were obtained in year 4 (*Task 1.b: completed*). Further activities during the fourth year of the project were focused on determining circulating miRNA levels in pre-treatment blood, genotyping, analyzing data and preparing manuscripts for publication. We sent the 116 samples that were obtained prior to treatment initiation to Exiqon (with an approved material transfer agreement in place) where miRNA was extracted from plasma (*Task 1.c: completed*). miRNA was quantified by Exiqon using their cancer miRNA panel (*Task 4: completed*). We performed genotyping on DNA samples from 480 men (*Task 3: completed*). The originally proposed genotyping platform (Illumina custom SNP panel) was no longer available, so we genotyped a dense panel of genome-wide SNPs using the Illumina Mega panel. This panel has the advantage of good gene coverage of all of our genes of interest in the miRNA biogenesis pathway, and also includes ancestry informative markers. The cost, however, was substantially higher than for the custom panel, so we selected all EA participants, all high risk AA men, and a random sample of low risk AA men for inclusion in the genotyping sample.

As stated in *Task 2* of our statement of work, we have been actively following all of the enrolled men in the cohort (from the previously funded study and from the current protocol) who did not have extensive disease at diagnosis for PSA outcomes. We continue to abstract the most recent follow-up of PSA test results through medical records and Caisis database, and will perform a final linkage with Metropolitan Detroit SEER registry (MDCSS) for vital status.

Results, Progress and Accomplishments with Discussion.

We evaluated associations between plasma levels of miRNAs and prostate cancer aggressiveness, and a manuscript describing the findings was published in *Carcinogenesis* (4) (*Task 5.c, Appendix 1).* A brief summary of our findings follows.

We investigated the patterns of expression of 68 cancer-related plasma-derived microRNAs (miRNAs) in a cohort of African American (AA) and European American (EA) prostate cancer patients (n=114, Table 1). miRNA qPCR results were evaluated for association with aggressive disease using a novel extension of CART methods. Aggressive disease was defined as Gleason sum $\geq$ 4+3, and non-aggressive disease was defined as Gleason sum $\leq$ 3+4. Our data demonstrate that a previously unreported circulating miRNA signature consisting of two distinct combinations of interacting miRNAs (miR-17/miR-192, Table 2 and Figure 1) and (miR-146a/miR-20a/miR26b, Table 3 and Figure 2) is capable of segregating aggressive and non-aggressive prostate cancer in both AA and EA patients. The interacting miRNAs outperformed independent miRNAs in identifying aggressiveness. Our results suggest that these circulating miRNAs may constitute novel biomarkers of prostate cancer aggressiveness in both races.

Table 1.

Variable	Level	Non-aggressive	Aggressive	χ^2 p-value
Age				0.685
	<=60	37(0.46)	13(0.39)	
	>60	44(0.54)	20(0.61)	
BMI				0.760
	<=25	17(0.21)	8(0.24)	
	(25,30]	33(0.41)	11(0.33)	
	>30	31(0.38)	14(0.42)	
Race				0.449
	African American	68(0.84)	25(0.76)	
	European American	13(0.16)	8(0.24)	

Table 2.

Group	Non-aggressive	Aggressive	% Aggressive
C1	65	14	17.7
C2	13	7	35.0
C3	3	12	80.0

Figure 1.

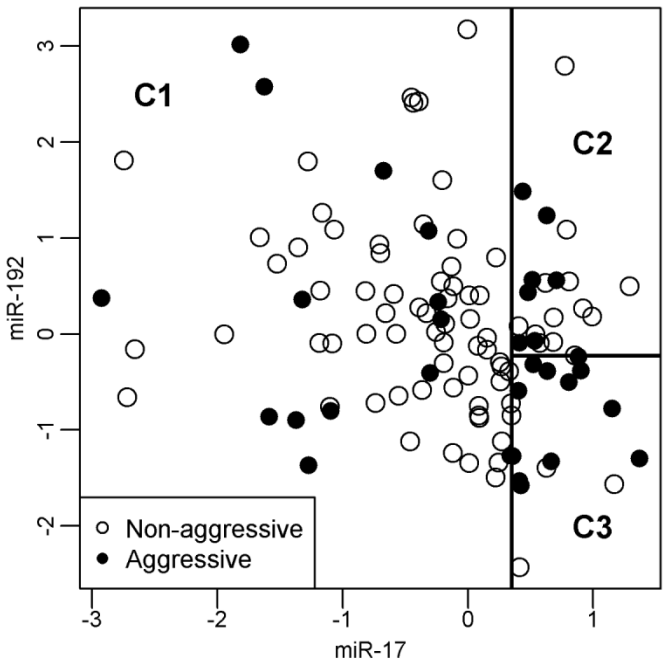
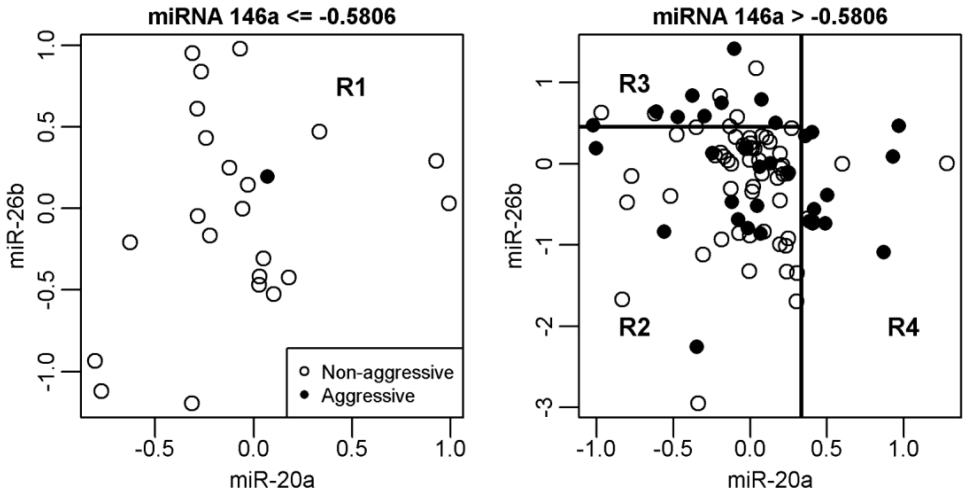


Table 3.

Group	Non-aggressive	Aggressive	% Aggressive
R1	22	1	4.3%
R2	50	13	20.6%
R3	5	9	64.3%
R4	4	10	71.4%

Figure 2.



The methods used for a separate analysis of this data were developed by the biostatistician on this project, and were published using data from this project in the American Journal of Cancer Research; the manuscript is included as Appendix 2 (5). In that paper focused on the African American subset of our cohort, we found that both extrema of circulating miR-17 are associated with aggressive prostate cancer (Figure 3 shows the scatterplot with a smoothed spline fit). This effect was observed in tumor samples from separate datasets measured on different assays representing different populations (Figure 4) of prostate cancer patients. Our result is consistent with the conflicting findings on the impact that miR-17 has in PCa progression, namely that it controls both oncogenic and tumor-suppressive genes.

Figure 3.

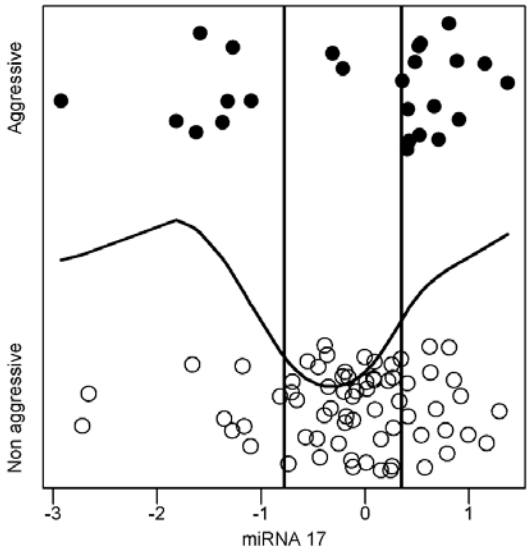
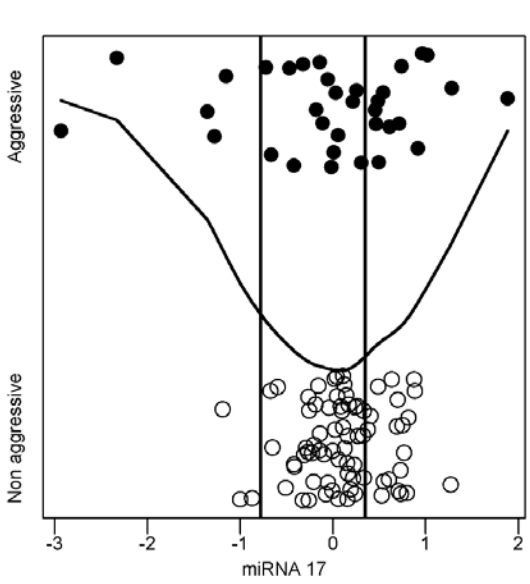
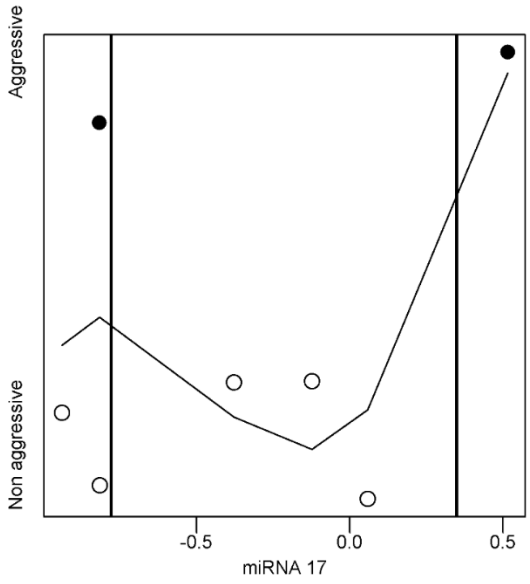


Figure 4.





We have begun our analyses of the SNP data in association with prostate cancer aggressiveness (*Task 5a*), and anticipate completing these and submitting a manuscript for publication soon. Using the methods described in the referenced papers, we estimated global ancestry (https://www.ncbi.nlm.nih.gov/pubmed/18683858) for each sample and local ancestry (https://www.ncbi.nlm.nih.gov/pubmed/23910464) for each gene within each sample to adjust for genetic background in the statistical models. Logistic regression was used to identify germline SNPs from genes in the miRNA biogenesis pathway that are predictive of prostate cancer aggressiveness.

We fit both unadjusted models and models adjusted for age, BMI, global ancestry, and local ancestry (see Appendix 3 for a full listing of the analysis results). Multiple SNPs in *TARBP1*, *DROSHA* and *DICER1* were significant for the AA sample (requiring at least 10 minor allele carriers), adjusted for covariates listed previously. Fewer SNPs were significant for the EA subset, mostly in *TARBP1* and DROSHA. Only 1 SNP was significant for both the African American and European American adjusted analyses: rs4920247, which is an intronic SNP located in *TARBP1*. The interesting finding from the race-specific analyses is that the implicated genes are consistent between the analyses; different locations within those genes are shown to be associated with disease aggressiveness.

We will next evaluate the SNP data in association with prostate cancer recurrence (Task 5b) and examine associations between SNPs and miRNA levels (Task 5d) and prepare manuscripts describing the results.

4. Key research Accomplishments

Nothing to report.

5. Conclusion

Our results suggest that circulating miRNAs may constitute novel biomarkers of prostate cancer aggressiveness in both races. Our data demonstrate that a previously unreported circulating miRNA signature consisting of two distinct combinations of interacting miRNAs (miR-17/miR-192) and (miR-146a/miR-20a/miR26b) is capable of segregating aggressive and non-aggressive prostate cancer in both AA and EA patients.

When we evaluated SNPs in genes involved in the miRNA biogenesis pathway, multiple SNPs in *TARBP1*, *DROSHA* and *DICER1* were significantly associated with PCa aggressiveness in the AA sample. Fewer SNPs were significantly associated with PCa aggressiveness in the EA subset, and were primarily located in *TARBP1* and *DROSHA*. Only 1 SNP was significant for both the AA and EA adjusted analyses: rs4920247, which is an intronic SNP located in *TARBP1*. The interesting finding from the race-specific analyses is that the implicated genes are consistent between the analyses; different locations within those genes were associated with disease aggressiveness.

6. Publications, Abstracts, and Presentations

Presentations:

1. MicroRNA and Prostate Cancer Disparities, Karmanos Cancer Institute Prostate Cancer Research Team seminar, April 22, 2016

2. Exploring Genes and MicroRNAs in Prostate Cancer Disparities, Karmanos Cancer Institute Population Studies and Disparities Research Program Seminar, March 4, 2016

3. MicroRNA in Prostate Cancer Racial Disparities and Aggressiveness, Karmanos Cancer Center Prostate Cancer Research Team SPORE planning retreat, June 25, 2013

Publications:

1. Farran, B., Dyson, G., Craig, D., Dombkowski, A., Beebe-Dimmer, J. L., Powell, I. J., Podgorski, I., Heilbrun, L., Bolton, S., Bock, C. H. (2018). A study of circulating microRNAs identifies a new

potential biomarker panel to distinguish aggressive prostate cancer. Carcinogenesis, 39(4), 556-561. doi:10.1093/carcin/bgy025 (Appendix 1)

2. Dyson, G., Farran, B., Bolton, S., Craig, D. B., Dombkowski, A., Beebe-Dimmer, J. L., Powell, I. J., Podgorski I., Heilbrun, L. K., Bock, C. H. (2018). The extrema of circulating miR-17 are identified as biomarkers for aggressive prostate cancer. Am J Cancer Res, 8(10), 2088-2095. (Appendix 2)

1. Inventions, Patents and Licenses

Nothing to report.

2. Reportable Outcomes

Nothing to report.

3. Other Achievements

The miRNA expression data was used in a manuscript describing the novel extension of CART methods used in this study; this manuscript was published in the *American Journal of Cancer Research* and is included in Appendix 2.

A review paper describing miRNA as a theranostic in prostate cancer has been written, and it is currently under review at *Carcinogenesis*.

4. References

1. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA: a cancer journal for clinicians **2014**;64(1):9-29 doi 10.3322/caac.21208.
2. Kong D, Banerjee S, Ahmad A, Li Y, Wang Z, Sethi S*, et al.* Epithelial to mesenchymal transition is mechanistically linked with stem cell signatures in prostate cancer cells. PloS one **2010**;5(8):e12445 doi 10.1371/journal.pone.0012445.
3. Kong D, Heath E, Chen W, Cher M, Powell I, Heilbrun L*, et al.* Epigenetic silencing of miR-34a in human prostate cancer cells and tumor tissue specimens can be reversed by BR-DIM treatment. American journal of translational research **2012**;4(1):14-23.
4. Farran B, Dyson G, Craig D, Dombkowski A, Beebe-Dimmer JL, Powell IJ*, et al.* A study of circulating microRNAs identifies a new potential biomarker panel to distinguish aggressive prostate cancer. Carcinogenesis **2018**;39(4):556-61 doi 10.1093/carcin/bgy025.
5. Dyson G, Farran B, Bolton S, Craig DB, Dombkowski A, Beebe-Dimmer JL*, et al.* The extrema of circulating miR-17 are identified as biomarkers for aggressive prostate cancer. Am J Cancer Res **2018**;8(10):2088-95.

5. Appendices

OXFORD

Carcinogenesis, 2018, Vol. 39, No. 4, 556–561

doi:10.1093/carcin/bgy025
Advance Access publication February 17, 2018
Original Article

ORIGINAL ARTICLE

A study of circulating microRNAs identifies a new potential biomarker panel to distinguish aggressive prostate cancer

Batoul Farran[1], Gregory Dyson[1], Douglas Craig[1], Alan Dombkowski[2], Jennifer L.Beebe-Dimmer[1], Isaac J.Powell[3], Izabela Podgorski[4], Lance Heilbrun[1], Susan Bolton[1] and Cathryn H.Bock[1,*]

[1]Department of Oncology, [2]Department of Pediatrics, [3]Department of Urology and [4]Department of Pharmacology, Karmanos Cancer Institute, Wayne State University School of Medicine, Detroit, MI, USA

*To whom correspondence should be addressed. Tel: +1 313 578 4203; Fax +1 313 578 4306; Email: bockc@wayne.edu

Abstract

Prostate cancer is one of the most common cancers in men worldwide. Currently available diagnostic and prognostic tools for this disease, such as prostate specific antigen, suffer from lack of specificity and sensitivity, resulting in over- and misdiagnosis. Hence, there is an urgent need for clinically relevant biomarkers capable of distinguishing between aggressive and nonaggressive forms of prostate cancer to aid in stratification, management and therapeutic decisions. To address this unmet need, we investigated the patterns of expression of a panel of 68 plasma-derived microRNAs (miRNAs) in a cohort of African American (AA) and European American (EA) prostate cancer patients (n = 114). miRNA qPCR results were analyzed using in-depth statistical methods, and a bioinformatics analysis was conducted to identify potential targets of the differentially expressed miRNAs. Our data demonstrate that a new previously unreported circulating miRNA signature consisting of a combination of interacting miRNAs (miR-17/miR-192) and an independent miRNA (miR-181a) are capable of segregating aggressive and nonaggressive prostate cancer in both AA and EA patients. The interacting miRNAs outperformed independent miRNAs in identifying aggressiveness. Our results suggest that these circulating miRNAs may constitute novel biomarkers of prostate cancer aggressiveness in both races and warrant further investigation.

Introduction

Prostate cancer is the second most frequently diagnosed cancer and the fifth leading cause of mortality among men worldwide (1). In 2016, an estimated 180 890 new prostate cancer cases were diagnosed in the USA alone, representing up to 10.7% of reported cancer cases, with over 3 million men living with the disease in the USA (2). The highest rates of prostate cancer incidence across five continents occur consistently among African-American (AA) populations (2). Incidence rates are also relatively high in black populations in Africa and the Caribbean, despite the lack of prostate specific antigen (PSA) screening, indicating that genetic predisposition could influence the risk of developing this disease (1,3). While the exact causes of prostate cancer remain largely unknown, established risk factors encompass advancing age, black race, family history of the disease, excess body weight and economic development (1).

MicroRNAs (miRNAs) are small noncoding RNA molecules (~22 nucleotides) that repress protein expression by cleaving mRNA or inhibiting its translation, thus modulating the expression of key oncogenes and tumor suppressor genes implicated in cell-cycle progression, cell growth and apoptosis. Recent studies have revealed the critical impact of alterations in the expression of cancer-related miRNAs on tumor progression. Pioneering research suggests that miRNA expression might be modified during the initial stages of carcinogenesis (4), rendering them early drivers of genomic instability and highlighting

Received: June 22, 2017; **Revised:** January 16, 2018; **Accepted:** February 12, 2018

Abbreviations

AA	African American
CART	classification and regression trees
EA	European American
miRNAs	microRNAs.

their potential as promising biomarkers for several solid cancers, including prostate cancer (5). Building on these findings, Calin and Croce (6) suggested that variations in the expression of miRNAs could be inherited in the germline, thus contributing to cancer predisposition by deregulating the intricate balance between tumor suppressor genes and oncogenes. Applying these ground-breaking concepts to the study of prostate cancer disparities between AA and European-American (EA) males, Calin and Croce (6) were the first to identify a set of miRNAs that are differentially expressed in AA versus EA patients (miR-301, miR-219, miR-26a, miR-1b-1 and miR-30c-1).

In 2008, Mitchell *et al.* published an innovative study demonstrating that miRNAs originating in human prostate cancer xenografts are secreted into the circulation and can be detected in patient plasma (7,8), thus providing the original proof of the presence of miRNA in the plasma of prostate cancer patients. This ground-breaking discovery has ushered in a new age of research aimed at uncovering sensitive noninvasive biomarkers of diagnosis and prognosis that could lessen the shortcomings of PSA-based screening and allow accurate prediction of disease aggressiveness, recurrence and chemoresistance. Improved diagnosis could reduce the mortality of prostate cancer by informing effective targeted therapy choices, but also curtail over-diagnosis and over-treatment, which pose other health threats. Circulating miRNAs constitute attractive diagnosis markers for various reasons. First, miRNAs appear to be frequently deregulated in cancer, where they display unique expression patterns in comparison with healthy controls and contribute to aberrant cell differentiation, proliferation and progression (7,8). They are stable, possess greater longevity than mRNA, are impermeable to degradation by nucleases due to their short sequence length and packaging in sheltering lipid microvesicles and are amenable to reliable extraction and measurement (5). Furthermore, circulating miRNAs overcome the limitations of clonal heterogeneity that reduce the diagnostic precision of tissue markers. They reflect the entire tumor mosaic by encapsulating the sum of heterotypic cellular and clonal interactions with the tumor microenvironment (9). Finally, miRNAs can be utilized as a low-cost noninvasive detection procedure and quantified by a variety of standard techniques, such as qRT-PCR, microarray or small RNA sequencing (5,7).

Recent profiling studies have endeavored to find a discriminatory circulating miRNA signature that can identify patients at risk of developing aggressive prostate disease that requires effective treatment, castration-resistant or chemotherapy refractory disease or simply to guide therapeutic strategies. These efforts have led to the detection of a few recurrent miRNA alterations that associate with prostatic carcinogenesis. In fact, three diagnostic and prognostic biomarkers emerge repeatedly in at least 17 studies (miR-141, miR-375 and miR-21), where one or more miRNAs were found to be concomitantly upregulated in the serum, plasma or urine of prostate cancer patients (7,8). Although discordant results have been published for a number of miRNAs (8), the quest for clinically relevant markers has not subsided as the need for novel diagnostic tools is still unmet. The goal of this study was to examine the association between circulating miRNA levels and prostate cancer aggressiveness to identify a panel of miRNAs that better distinguishes aggressive from nonaggressive disease. The cohort included both AA and EA prostate cancer patients to evaluate potential differences by race in miRNA associations with prostate cancer aggressiveness.

Materials and methods

Sample collection

The current study population was part of a larger prostate cancer cohort study. Cases diagnosed with prostate cancer between January 2004 and December 2016 were recruited through Karmanos Cancer Institute (KCI) clinics in metropolitan Detroit, MI. All participants self-identified as AA or non-Hispanic white were aged 75 years or younger at date of diagnosis and did not have a history of any invasive cancer prior to prostate cancer diagnosis. KCI cases were identified primarily through the hospital's Caisis database as eligible and approached for enrollment into the study at the time of their clinic appointment with consent of their treating physician. All participants provided written informed consent, and the study protocol was approved by the Wayne State University Institutional Review Board.

Participants completed a written, self-administered questionnaire at the time of their consent to participate in the study. This captured demographics, family history of prostate cancer, medical history and prostate cancer screening behaviors. At the time of study entry, each patient had their height, weight, waist and hip circumference measured according to a standardized protocol by trained study personnel. Blood was collected in CPT tubes for plasma isolation. Samples were stored at –80°C.

All medical records related to each patient's prostate cancer diagnosis, treatment and follow-up were reviewed. Age at diagnosis, prediagnostic PSA, clinical stage, biopsy Gleason grade, primary treatment, tumor pathologic stage, margin status and pathologic Gleason grade (for radical prostatectomy patients), nodal status and evidence for metastases were obtained from medical records. Aggressive disease was defined by Gleason score (4 + 3 and higher) from radical prostatectomy (if performed), otherwise from diagnostic biopsy.

RNA sample preparation

Plasma samples were thawed on ice and centrifuged at 3000*g* for 5 min in a 4°C microcentrifuge. An aliquot of 200 µl per sample was transferred to a FluidX tube. Sixty microliters of lysis solution BF, containing 1 µg carrier-RNA per 60 µl lysis solution BF, and RNA spike-in template mixture was added to the sample, mixed for 1 min and incubated for 7 min at room temperature, followed by the addition of 20 µl Protein Precipitation Solution BF. Total RNA was extracted from the samples using miRCURY RNA isolation Kit–Biofluids (Exiqon, Vedbaek, Denmark) in an automated 96-well format. The purified total RNA was eluted in a final volume of 50 µl.

miRNA real-time qPCR

RNA, 4 µl, was reverse transcribed in 20 µl reactions using the miRCURY LNA™ Universal RT miRNA PCR, Polyadenylation and cDNA synthesis kit (Exiqon). cDNA was diluted 50× and assayed in 10 µl PCR reactions according to the protocol for miRCURY LNA™ Universal RT miRNA PCR; each miRNA was assayed once by qPCR on the miRNA Ready-to-Use PCR, CANCer Focus panel using EXILENT SYBR® Green Master mix. Including 5 RNA spike-ins controls and 2 small noncoding RNAs (endogenous controls), 92 miRNAs were measured on the miRNA panel. Negative controls excluding templates from the reverse transcription reaction were performed and profiled like the samples. The amplification was performed in a LightCycler® 480 Real-Time PCR System (Roche) in 384-well plates. The amplification curves were analyzed using the Roche LC software, both for determination of Cq (by the second derivative method) and for melting curve analysis. Internal RNA isolation and cDNA synthesis controls indicated that the extraction efficiency was similar for all samples and that none harbored inhibitors. miR-103, miR-23a and miR-30c, which are known to be expressed at a steady level in many sample types, were employed as controls to assess the miRNA content of the analyzed samples. To verify that the miRNAs detected in our samples did not originate from blood contamination through hemolysis but were shed by the prostatic tumor burden itself, thus accurately mirroring neoplasm stage, two miRNAs were used: miR-451, which is highly expressed in red blood cells, and miR-23-a, which is relatively stable in serum and plasma and not

affected by hemolysis. The dCp value was calculated as the ratio of miR-451 relative miR-23a for all plasma samples.

Statistical methods

Aggressive and nonaggressive cases were evaluated for differences in age, BMI, PSA at diagnosis and race using the chi-square test. Any characteristics that were statistically significant at 0.05 level were included in further models.

The median polish method (10) was used to estimate miRNA and person-specific effects that will allow for normalization of the miRNA array data. These quantities, along with the overall median C_q value across all samples, were subtracted from the observed C_t values, resulting in a score describing the deviation of the observed C_t from expectation. To ease the interpretation (so that higher score values indicate higher miRNA expression), the scores were multiplied by −1. As a result of the median polish, the miRNA values are centered within each miRNA.

We used classification and regression trees (11) (CART) to identify miRNAs along with patient characteristics, which were associated with high or low prostate cancer aggressiveness. CART was implemented using the rpart function in the rpart package (12) in the R programming environment (13). We used the default parameters except that we forced the minimum size of a terminal node to consist of at least 10 observations. Categorized age, race and BMI were included as potential splitting factors in the CART analyses. Standard logistic regression was also used to associate each individual miRNA with the aggressiveness phenotype. Multiple approaches were used to allow for the possibility that heterogeneous mechanisms may lead to aggressive prostate cancer.

In the absence of publically available circulating miRNA datasets on prostate cancer patients, bootstrapping was used to estimate prediction accuracy. Briefly, for each bootstrap sample, a model was developed using the same variables and input parameters as when constructing the model for the original database. Predictions using the model built from the bootstrap sample were made to the observations not included in the bootstrap sample to compute model accuracy. The prediction accuracy was defined as the median accuracy across all bootstrap samples. Given the relatively small sample size, we anticipated that both approaches would have a tendency to overfit the data. We utilized bioinformatic tools (described in the next section) to corroborate any miRNA relationship that we observe.

Targets and pathway analysis

Distinct sets of miRNAs identified as being associated with disease aggressiveness were used to identify associated gene targets in the KEGG (14) prostate cancer pathway (hsa05215) using the DIANA-miRPath v3.0 (15) KEGG reverse search feature. For each set composed of an individual miRNA, a list of gene targets for that miRNA was compiled. For the two signature sets containing multiple miRNAs, we compiled a list of gene targets common to all miRNAs in the set.

Results

Participant characteristics

Plasma samples collected prior to prostate cancer treatment initiation were available for 116 cohort members. One sample was excluded due to suggested hemolysis (dCp > 8) and one sample was excluded because >25% of miRNAs were not detected. After excluding 17 miRNAs that were missing for at least 25% of the samples, 68 plasma-derived miRNAs in 114 patients, n = 93 AA and n = 21 EA, were eligible for inclusion in our analyses. On average, 66 miRNAs were detected per sample, and 44 miRNAs were detected in all 114 samples. Table 1 describes age, BMI, PSA and race characteristics in our sample. Only PSA at the time of diagnosis was significantly associated with aggressiveness.

Table 1. Summary of the patient cohort under investigation in this study

Variable	Level	Nonaggressive	Aggressive	χ^2 P-value
Age				0.685
	≤60	37(0.46)	13(0.39)	
	>60	44(0.54)	20(0.61)	
BMI				0.760
	≤25	17(0.21)	8(0.24)	
	(25, 30)	33(0.41)	11(0.33)	
	>30	31(0.38)	14(0.42)	
Race				0.449
	AA	68(0.84)	25(0.76)	
	EA	13(0.16)	8(0.24)	
PSA				<0.001
	<10	75 (0.93)	14 (0.42)	
	>10	6 (0.07)	19 (0.58)	

A circulating miR-17a/miR-192 signature is associated with disease aggressiveness

CART analysis segregated the prostate cancer cases into three distinct groups based on the combined effects of two miRNAs: miR-17 and miR-192. Age, BMI, PSA and race were not utilized by CART to discriminate aggressive and nonaggressive prostate cancer. We found that the upregulation of miR-192 exerted a moderating effect on prostate cancer aggressiveness for patients with upregulated miR-17. As shown in Figure 1 and Table 2, patients with reduced miR-17 levels (group C1) were classified as having low likelihood of being aggressive, regardless of miR-192 expression level (17.7% aggressive). Patients with upregulated miR-17 and miR-192 levels, shown in group C2, were more likely to have aggressive disease (35% aggressive). Patients who had high levels of miR-17 but reduced levels of miR-192, shown in group C3, were most likely to have aggressive disease (80% aggressive).

Individual miRNAs with marginal association with disease aggressiveness

Standard logistic regression was subsequently used to associate each miRNA with the aggressiveness phenotype in the prostate cancer samples while adjusting for race. As shown in Figure 2, we observed an association between the aberrant expression of three circulating miRNAs (miR-150a, miR-181a and miR-22) and prostate cancer aggressiveness. Plasma miR-181a level [odds ratio = 0.28 (relative to a 1-unit increase in the miR), P = 0.0004] and plasma miR-150a (odds ratio = 0.67, P = 0.032) level were negatively associated with disease aggressiveness. Plasma miR-22 (odds ratio = 1.94, P = 0.017) level was positively associated with disease aggressiveness. When adjusting for age, BMI, PSA at diagnosis and race, we observed only two significant miRNA (miR-181a & miR-93). Plasma miR-181a level (odds ratio = 0.24, P = 0.001) remained negatively associated with disease aggressiveness, while plasma miR-93 (odds ratio = 3.76, P = 0.042) level was positively associated with disease aggressiveness.

The dysregulated miRNAs affect key signaling pathways in prostate cancer development

To identify potential target genes and/or pathways of our miRNAs, we established their putative biological targets using the DianaTools-mirPath v.3 KEGG Reverse Search platform. Focusing on the prostate cancer pathway, we observed that the miR-17/miR-192 and miR-146a/miR-20a/miR-26b signatures affected

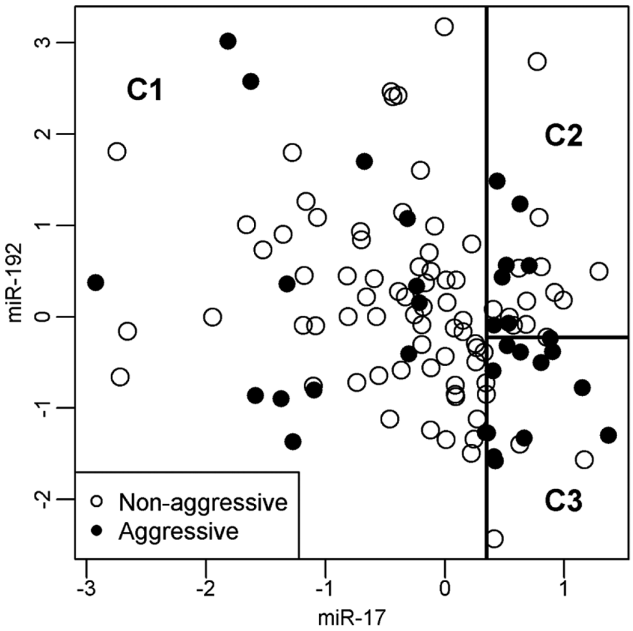


Figure 1. A miR-17/miR-192 signature is associated with prostate cancer aggressiveness. Scatter plot illustrating the graphical relationship between levels of expression of circulating miR-17 and miR-192 and prostate cancer aggressiveness in three patient subgroups constructed by the CART analysis. The error rate from bootstrapping was 36% for this model. Higher values on the x- and y-axis scores indicate higher miRNA expression.

Table 2. Percent of aggressive versus nonaggressive patients in CART subgroups

Group	Nonaggressive	Aggressive	% Aggressive
C1	65	14	17.7
C2	13	7	35.0
C3	3	12	80.0

different tumorigenesis pathways. miR-17 had 31 target genes, encompassing most of the genes in the prostate cancer pathway, while miR-192 possessed 9 targets. These two miRNAs shared three common gene targets, EP300, IGF1R and MDM2. On the other hand, miR-146a possessed 4 potential gene targets, while miR-20a and miR-26b had 23 and 21 targets, respectively. The three miRNAs shared two common gene targets, CDKN1A and MDM2, while miR-146a and miR-26b shared another target, NFKB1. We also analyzed the putative gene targets of our three independent markers and identified three interesting targets of miR-150, CDKN1B, EP300 and FGFR1. miR-181a had 25 gene targets, including the oncogenes NRAS, KRAS, HRAS, HSP90, PI3K3R and BCL2. Finally, miR-22 was predicted to interact with 18 gene targets, including EP300, CDKN1A, PTEN and TP53. The complete lists of target genes identified for each miRNA are shown in Supplementary Material, available at *Carcinogenesis* Online.

Discussion

In this study, we have shown that a previously unreported combination of interacting miRNAs (miR-17/miR-192) and three independent miRNA markers can distinguish between patients with aggressive versus nonaggressive prostate cancer in a cohort of AA and EA men. Bioinformatics analysis unmasked key target genes and cancer pathways potentially impacted by these differentially altered miRNAs. The novelty of our approach lies in the analytical and statistical tools we employed to assess the association between groups of miRNAs and disease aggressiveness, as opposed to identifying individual miRNA targets.

Based on CART analysis, the miR-17/miR-192 pair identified disease aggressiveness according to three distinct expression cut points (downregulated/upregulated; upregulated/upregulated; upregulated/downregulated), indicating that the progressive induction of miR-17a and loss of miR-192a translated into increased incidence of aggressive cancer. Although we established a diagnostic connection between the impaired expression of specific interacting miRNA signatures and prostate cancer stage in both races, we did not identify discrepancies in miRNA expression in AA versus EA men, as suggested by previous studies (9). This limitation indicates that the relatively smaller size of our cohort could have diluted the influence of miRNA variances on racial disparities in prostate cancer incidence, signifying that a larger population might be required to better assess these differences. Furthermore, our miRNA panel only measured a fraction of the total known miRNAs, implying that we might have missed an important miRNA interaction. The lack of circulating miRNA measurements in publically available prostate cancer datasets constituted another limitation. We employed bioinformatics tools to confirm the mechanisms and suggest potential targets. However, we will need to corroborate our findings in another dataset of prostate cancer. Despite these limitations, our findings indicate that the miR-17/miR-192 could provide novel clinically relevant information that might improve the diagnosis and management of prostate cancer in both AA and EA patients. These findings warrant validation in a larger cohort to ascertain their clinical utility.

The upregulation of miR-17 has been consistently reported in various hematopoietic and solid malignancies including breast, lung, pancreas, prostate, stomach cancers and B-cell lymphomas (16). This miRNA belongs to a cluster of six oncogenic miRNAs, miR-17–92 (17), known as oncomir-1 or first oncomir (18). miR-17 has been repeatedly found to be enriched in prostate cancer cells, exosomes (19,20), tissue (21) and the serum of high-risk patients (22,23), indicating that our results agree with previous findings (24). miR-17 exerts its oncogenic effect by inhibiting an intricate network of tumor suppressors, antiproliferative, antiapoptotic and cell cycle genes, including PTEN (25), the loss of which constitutes one of the hallmarks of prostate cancer (26), HIF1A (27) and CDKN1A (28).

miR-192 downregulation has been documented in several cancer types, including multiple myeloma, colorectal, ovarian and renal cell carcinomas (29) but has only been observed in prostate cancer cell lines (30). Our study is the first to report reduced miR-192 levels in the plasma of aggressive prostate cancer patients. miR-192 is directly induced by p53 (31) and functions as a tumor suppressor implicated in the regulation of the angiogenic switch (29) and the cell cycle (32). Its targets comprise genes overexpressed in aggressive prostate cancer, including the oncogene BCL2 (32) and MYLK, described as one of the most informative cancers for prostate tumorigenesis (33).

Our bioinformatics analysis unraveled two common putative target pathways of miR-17 and miR-192, comprising genes such as EP300, IGF1R and MDM2. This suggests that the dysregulation of this miRNA pair could contribute to aberrant G_1 and G_2 arrest, reduced apoptosis and genomic instability through impaired p53 signaling, as well as increased PI3K-AKT signaling. Previous work has shown that the acetyltransferase EP300 is a major promoter of prostate cancer development. It is thought to exercise its oncogenic activity by promoting epithelial–mesenchymal transition (34) and has been linked to worse disease

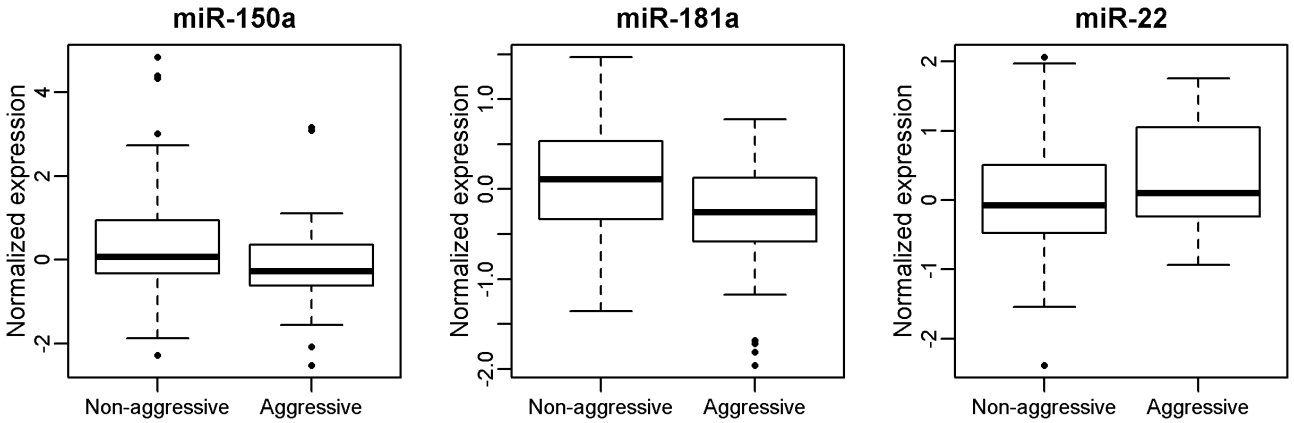


Figure 2. Circulating levels of miR-150a, miR-181a and miR22 are independently associated with prostate cancer aggressiveness by logistic regression. Here, we present normalized expression levels of these three miRNAs by prostate cancer aggressiveness.

prognosis (35). Although the miR-17–92 cluster has been shown to regulate EP300 expression in the myocardium (36), the interaction between miR-192 and EP300 has not been investigated. We hypothesize that miR-17 and miR-192 might be exerting opposing effects by inhibiting key tumor suppressors (miR-17) and upregulating oncogenes (miR-192) in this pathway, resulting in a synergistic effect that promotes tumorigenesis. The paradoxical relation between altered miR-17/miR-192 and EP300 in prostate cancer thus warrants further examination.

To complement our diagnostic model, we identified three independent markers of disease aggressiveness, miR-150a, miR-181a and miR-22. Plasma miR-150 levels were decreased in aggressive prostate cancer cases, in agreement with tissue (37) and urine-based studies (8). Pathway analysis identified three important targets of miR-150, including EP300, FGFR1 and the tumor suppressor CDKN1B (p27), found to be significantly upregulated in clinical samples of prostate cancer (38). Furthermore, reduced levels of circulating miR-181a were associated with disease aggressiveness in our cohort, in agreement with previous findings (39). Putative targets of miR-181a included NRAS, KRAS, HRAS, HSP90, PI3K3R and BCL2, highlighting its involvement in a wide web of signaling pathways, including the PI3K-AKT and MAPK cascades. The final marker, miR-22, has been described as a proto-oncogene in prostate cancer cells (40). However, our study is the first to report its upregulation in the plasma of prostate cancer patients. After adjusting for age, BMI, PSA at diagnosis and race, the statistically significant negative association between miR-181a and aggressiveness was not appreciably changed; however, miR-150a and miR-22 were no longer associated with aggressiveness, and miR-93 was marginally associated with disease aggressiveness.

In this pilot study, we have identified a novel panel of circulating miRNAs, consisting of a combination of interacting miRNAs and three independent markers. Our study has several limitations. First, our cohort of 116 patients is relatively small, which diluted discrepancies in miRNA expression in AA versus EA men. Second, our miRNA panel only measured a fraction of the total known miRNAs. Third, the miRNAs detected in this study were not validated in independent samples. Despite these limitations, our panel was able to successfully distinguish between aggressive versus nonaggressive prostate cancer. Bioinformatics analysis revealed that the dysregulated miRNAs regulate key cellular functions and could potentially drive an aggressive neoplastic phenotype. Hence, this novel miRNA signature warrants further investigation in a larger cohort to assess its potential as a diagnostic tool for prostate cancer.

Supplementary material

Supplementary data can be found at *Carcinogenesis* online.

Funding

Congressionally Directed Medical Research Programs (W81XWH-10-2-0068 to I.P. and W81XWH-13-1-0477 to C.B.); the Michigan Fund for Cancer Research (to C.B).

Acknowledgments

The authors thank Karen Slaven and Nicole Kubinec for their tireless subject recruiting efforts.
Conflict of Interest Statement: None declared.

References

1. Zhou, C.K. *et al.* (2016) Prostate cancer incidence in 43 populations worldwide: an analysis of time trends overall and by age group. Int. J. Cancer, 138, 1388–1400.
2. Howlader, N. *et al.* (1975–2014) SEER Cancer Statistics Review. National Cancer Institute, Bethesda, MD.
3. Parkin, D.M. *et al.* (2014) Cancer in Africa 2012. Cancer Epidemiol. Biomarkers Prev., 23, 953–966.
4. Calin, G.A. *et al.* (2004) Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc. Natl. Acad. Sci. USA, 101, 2999–3004.
5. Fabris, L. *et al.* (2016) The potential of microRNAs as prostate cancer biomarkers. Eur. Urol., 70, 312–322.
6. Calin, G.A. *et al.* (2006) MicroRNA signatures in human cancers. Nat. Rev. Cancer, 6, 857–866.
7. Mitchell, P.S. *et al.* (2008) Circulating microRNAs as stable blood-based markers for cancer detection. Proc. Natl. Acad. Sci. USA., 105, 10513–10518.
8. Filella, X. *et al.* (2016) miRNAs as novel biomarkers in the management of prostate cancer. Clin. Chem. Lab Med. 55, 715–736.
9. Yates, C. *et al.* (2017) miRNAs as drivers of TMPRSS2-ERG negative prostate tumors in African American men. Front. Biosci. (Landmark Ed.), 22, 212–229.
10. Tukey, J.W. (1977) Exploratory Data Analysis. Addison-Wesley, Reading, MA.
11. Breiman, L.F.J. *et al.* (1984) Classification and Regression Trees. Wadsworth International Group, Belmont, CA.
12. Therneau, T.M. *et al.* (2014) rpart: Recursive Partitioning and Regression Trees. R package. https://CRAN.R-project.org/package=rpart
13. R Development Core Team (2011) R: A Language and Environment for Statistical Computing. The R Foundation for Statistical Computing, Vienna, Austria

14. Kanehisa, M. *et al.* (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res., 45, D353–D361.
15. Vlachos, I.S. *et al.* (2015) DIANA-miRPath v3.0: deciphering microRNA function with experimental support. Nucleic Acids Res., 43, W460–W466.
16. van Haaften, G. *et al.* (2010) Tumorigenicity of the miR-17-92 cluster distilled. Genes Dev., 24, 1–4.
17. Ota, A. *et al.* (2004) Identification and characterization of a novel gene, C13orf25, as a target for 13q31-q32 amplification in malignant lymphoma. Cancer Res., 64, 3087–3095.
18. Mogilyansky, E. *et al.* (2013) The miR-17/92 cluster: a comprehensive update on its genomics, genetics, functions and increasingly important and numerous roles in health and disease. Cell Death Differ., 20, 1603–1614.
19. Dhar, S. *et al.* (2015) Resveratrol and pterostilbene epigenetically restore PTEN expression by targeting oncomiRs of the miR-17 family in prostate cancer. Oncotarget, 6, 27214–27226.
20. Ahadi, A. *et al.* (2016) Long non-coding RNAs harboring miRNA seed regions are enriched in prostate cancer exosomes. Sci. Rep., 6, 24922.
21. Volinia, S. *et al.* (2006) A microRNA expression signature of human solid tumors defines cancer gene targets. Proc. Natl. Acad. Sci. USA, 103, 2257–2261.
22. Sapre, N. *et al.* (2014) Curated microRNAs in urine and blood fail to validate as predictive biomarkers for high-risk prostate cancer. PLoS One, 9, e91729.
23. Feng, S. *et al.* (2017) Combinations of elevated tissue miRNA-17-92 cluster expression and serum prostate-specific antigen as potential diagnostic biomarkers for prostate cancer. Oncol. Lett., 14, 6943–6949.
24. Shukla, K.K. *et al.* (2017) Recent scenario of microRNA as diagnostic and prognostic biomarkers of prostate cancer. Urol. Oncol., 35, 92–101.
25. Luan, Y. *et al.* (2017) MiR-17 targets PTEN and facilitates glial scar formation after spinal cord injuries via the PI3K/Akt/mTOR pathway. Brain Res. Bull., 128, 68–75.
26. Wise, H.M. *et al.* (2017) Prostate cancer, PI3K, PTEN and prognosis. Clin. Sci. (Lond.), 131, 197–210.
27. Taguchi, A. *et al.* (2008) Identification of hypoxia-inducible factor-1 alpha as a novel target for miR-17-92 microRNA cluster. Cancer Res., 68, 5540–5545.
28. Li, S. *et al.* (2017) miR-3619-5p inhibits prostate cancer cell growth by activating CDKN1A expression. Oncol. Rep., 37, 241–248.
29. Wu, S.Y. *et al.* (2016) A miR-192-EGR1-HOXB9 regulatory network controls the angiogenic switch in cancer. Nat. Commun., 7, 11169.
30. Sun, J. *et al.* (2016) MiR-192 suppresses the tumorigenicity of prostate cancer cells by targeting and inhibiting nin one binding protein. Int. J. Mol. Med., 37, 485–492.
31. Hünten, S. *et al.* (2015) p53-regulated networks of protein, mRNA, miRNA, and lncRNA expression revealed by integrated pulsed stable isotope labeling with amino acids in cell culture (pSILAC) and next generation sequencing (NGS) analyses. Mol. Cell. Proteomics, 14, 2609–2629.
32. Georges, S.A. *et al.* (2008) Coordinated regulation of cell cycle transcripts by p53-Inducible microRNAs, miR-192 and miR-215. Cancer Res., 68, 10105–10112.
33. Fujita, A. *et al.* (2008) Multivariate gene expression analysis reveals functional connectivity changes between normal/tumoral prostates. BMC Syst. Biol., 2, 106.
34. Isharwal, S. *et al.* (2008) p300 (histone acetyltransferase) biomarker predicts prostate cancer biochemical recurrence and correlates with changes in epithelia nuclear size and shape. Prostate, 68, 1097–1104.
35. Ghosh, A.K. *et al.* (2007) The transcriptional coactivator and acetyltransferase p300 in fibroblast biology and fibrosis. J. Cell. Physiol., 213, 663–671.
36. Shehadeh, L.A. *et al.* (2013) MicroRNA-20a constrains p300-driven myocardial angiogenic transcription by direct targeting of p300. PLoS One, 8, e79133.
37. Kumar, B. *et al.* (2016) MicroRNA expression and function in prostate cancer: a review of current knowledge and opportunities for discovery. Asian J. Androl., 18, 559–567.
38. Lynch, S.M. *et al.* (2016) miR-24 regulates CDKN1B/p27 expression in prostate cancer. Prostate, 76, 637–648.
39. Bryant, R.J. *et al.* (2012) Changes in circulating microRNA levels associated with prostate cancer. Br. J. Cancer, 106, 768–774.
40. Budd, W.T. *et al.* (2015) Dual action of miR-125b as a tumor suppressor and oncomiR-22 promotes prostate cancer tumorigenesis. PLoS One, 10, e0142373.

Am J Cancer Res 2018;8(10):2088-2095
www.ajcr.us /ISSN:2156-6976/ajcr0079878

Original Article

The extrema of circulating miR-17 are identified as biomarkers for aggressive prostate cancer

Greg Dyson[1], Batoul Farran[1], Susan Bolton[1], Douglas B Craig[1], Alan Dombkowski[2], Jennifer L Beebe-Dimmer[1], Isaac J Powell[3], Izabela Podgorski[4], Lance K Heilbrun[1], Cathryn H Bock[1]

[1]Karmanos Cancer Institute and Department of Oncology, Wayne State University, Detroit MI, USA; [2]Karmanos Cancer Institute and Department of Pediatrics, Wayne State University, Detroit MI, USA; [3]Karmanos Cancer Institute and Department of Urology, Wayne State University, Detroit MI, USA; [4]Karmanos Cancer Institute and Department of Pharmacology, Wayne State University, Detroit MI, USA

Received May 18, 2018; Accepted June 3, 2018; Epub October 1, 2018; Published October 15, 2018

Abstract: MicroRNAs (miRNAs) constitute short non-coding RNAs that can post-transcriptionally modulate the expression of many oncogenes and tumor suppressor genes engaged in key cellular processes. Deregulated serum miRNA signatures have been detected in various solid cancers including prostate cancer, suggesting that circulating miRNAs could function as non-invasive biomarkers of tumor emergence and progression. To determine whether serum miRNA expression levels are different between patients with aggressive and non-aggressive prostate cancer, we analyzed a panel of miRNAs from the blood of African American (AA) prostate cancer patients using a new recursive partitioning method that allows hypothesis testing of each split. We observed that both extrema of circulating miR-17, i.e. upregulation and downregulation, are associated with aggressive prostate cancer. A similar effect was observed in tumor samples from a separate dataset representing a different population of prostate cancer patients and in AA prostate cancer samples from the TCGA. The dual effect is consistent with the contradictory findings on the role of miR-17 in prostate cancer progression, whereby it controls important oncogenic and tumor-suppressive genes.

Keywords: Prostate cancer, microRNA, miR-17, classification, hypothesis testing, regression trees

Introduction

Prostate cancer (PCa) constitutes the second most frequently diagnosed cancer and fifth leading cause of mortality among men worldwide [1]. Its incidence is higher in the industrialized world, suggesting that diet and lifestyle represent potential risk factors [2]. In fact, nutrition and exercise modulate serum factors that limit growth and elicit apoptosis in androgen-dependent PCa cells, whereas increased body mass index, blood pressure and metabolic factors provoke an increased risk of PCa death [3]. The highest rates of PCa are observed in African American (AA) and Caribbean populations, suggesting that genetic predisposition might contribute to disease development and aggressiveness [1]. Current screening methods for PCa include digital rectal exam, serum level of prostate-specific antigen and transrectal ultrasound guided biopsy [3]. These methods do not have high sensitivity and specificity, often leading to misdiagnosis or over-diagnosis, highlighting the need for improved methods of early detection and management.

Circulating microRNAs (miRNAs) have emerged as promising diagnostic and prognostic biomarkers in various solid cancers including PCa due to their increased stability, bio-availability, and frequent deregulation during tumorigenesis. miRNAs are small non-coding RNA molecules that repress protein expression by cleaving mRNA or inhibiting its translation, which enables them to fine-tune the expression of an intricate network of oncogenes and tumor suppressor genes implicated in key cellular functions, such as cell cycle progression, cell growth and apoptosis. Aberrant expression of miRNAs contributes to tumorigenesis by de-repressing or silencing key regulatory proteins. Extensive research efforts are currently underway to

uncover clinically relevant tissue and serum miRNA signatures of diagnostic and prognostic value for the stratification and monitoring of PCa [4-6].

Recursive partitioning (RP) is a statistical tool used to create subgroups of objects (individuals) with differential levels of the phenotype; utilizing binary splits derived from values of the input variables to create the subgroups. A thorough review [7] of RP reveals a rich history of statistical developments since the algorithm was first published [8]. Highlights of this review include a brief account of the incremental advancements in RP theory and a description of modifications to apply to various classes of outcome data. A central part of the research on the RP algorithm is defining a rule for when to stop growing a tree (or how much to prune) to yield an optimal prediction model. Classification and regression trees (CART) [9] were developed to classify objects using binary splits with either a continuous or categorical response based in part on the idea formed from research on RP. The key additions of the CART algorithm were pruning large trees to a parsimonious size using a complexity parameter and utilizing crossvalidation to estimate error in prediction. A crossvalidation heuristic is typically used to select the optimal size of the tree via the complexity parameter; e.g., select the tree that results in the smallest crossvalidation error or select the smallest tree where the crossvalidation error is within one standard deviation of the minimum crossvalidation error [10]. Further enhancements which made an effort to incorporate statistical testing into the CART fitting process focused on methods to optimize the pruning process [11, 12]. Subsequent research led to the development of conditional trees, which embeds RP in a conditional inference framework and includes stopping rules based on multiple testing [13]. Ensemble methods like random forests [14] have also been utilized to increase prediction accuracy while consequently sacrificing the interpretability of the resultant model [15].

The objective of the RP with hypothesis testing (RP-HT) method introduced in this paper is to eliminate the parameters, heuristics, and pruning algorithms utilized by RP methods by utilizing hypothesis testing at each potential split in the tree. In doing so, new splitting objective functions for continuous and categorical responses are developed. The proposed RP-HT method builds upon previous work [16] incorporating hypothesis testing into the Patient Rule-Induction Method [17]. A dataset measuring a panel of circulating microRNAs from African American prostate cancer patients is analyzed using this new technique. The models produced by RP-HT will be compared with those produced by CART and a simulation study is utilized to estimate the overall Type I error.

Statistical methods

The new objective function developed for RP-HT will evaluate the splits that maximize the proportion of each of the observed classes for a categorical response variable, or maximizes the mean rank for a continuous response variable. The rank rather than the observed response variable is used for continuous responses to provide a robust estimate that is not sensitive to outlying responses, and to allow for normal distribution theory of order statistics to be applied to data that may not be normally distributed. RP-HT will continue splitting the data and building the tree until there is no statistically significant split available at each of the existing nodes. Like many other RP-based algorithms, RP-HT is greedy in the sense that it selects the optimal split at a given node without looking forward or going back on earlier decisions [18]. RP-HT is especially greedy as it does not identify the split that results in the smallest *p*-value, but rather the largest result proportion or mean rank.

A support parameter and its selection process are also introduced to regularize the RP-HT fitting process. This parameter is the minimum proportion of the observations under consideration (from the mother node) that are required to be in both of the daughter nodes. A range of support parameters between 0.05 and 0.50 incremented by 0.005 will be considered. The lower bound is set at 5% to ensure that terminal nodes have adequate size. The upper bound is the maximum value that the support parameter could be, given that both daughter nodes have to meet the size requirement. The greediness of the RP-HT algorithm is controlled by this support parameter (which is the same at every node in the tree): the smaller the support parameter, the smaller the created sub-nodes

will be. A smaller support parameter will in general lead to a tree with more terminal nodes. For a continuous response, the support parameter that maximizes the Kendall's tau correlation [19] between the observed response and the predicted response is selected (among trees with at least one significant split). For categorical responses, the support parameter that minimizes the sum of one minus the predicted probability for the true class squared is selected (among trees with at least one significant split).

Using a support parameter and the objective function, RP-HT determines the optimal split at a given node. For each potential split, the RP-HT performs exact hypothesis testing for categorical responses and exact or asymptotically exact hypothesis testing for continuous responses to determine if the split was statistically significant, conditional on the number of possible splits from all potential explanatory variables. The derivation of the null distribution for both categorical and continuous responses to construct a *p*-value is detailed in the Supplemental Methods. At each node in the tree-building process, the RP-HT algorithm will determine the most statistically significant split, given the support parameter. If the *p*-value from that split is smaller than the input nominal significance level, then that spilt is incorporated into the definition of the tree. This process will then be conducted on the daughter nodes created from that split. If the *p*-value from a split is not smaller than the nominal significance level, then the algorithm denotes it to be a terminal node and all observations in that node are assigned to that output class. The algorithm continues until there is no split available with a *p*-value less than the nominal significance level at each available node and therefore every observation has been assigned to a terminal node class. For categorical responses, the predicted value for an observation in a particular terminal node is the response category with a plurality of members of that terminal node. For continuous responses, the predicted value for an observation in a particular terminal node is the median of the raw responses of all of the observations that make up the terminal node. The median is used instead of the mean to lessen the impact of outlier observations.

As the implementation of the original CART algorithm is proprietary, we use the version in the rpart package [20] in R programming environment [21] to compare with our proposed algorithm. The rpart version uses the Gini index to select the optimal splits for categorical response variables [22]. Other parameters included in this implementation include "minbucket" (the minimum number of observations required in all daughter nodes), complexity parameter (the minimum decrease in lack of fit required to consider a split), and "minsplit" (the minimum number of observations in a mother node needed to attempt a split) that are used to control the tree-fitting process [22]. For our comparison we used to the default value for theses parameters, with the exception of "minbucket" which we set to 10. We chose not to compare RP-HT with ensemble method like random forests, bagging and boosting to maintain the interpretability of resultant model and ensure a fair comparison.

Clinical methods

miRNA discovery study

A panel of 92 plasma microRNA levels were assayed in a cohort of blood samples from 116 PCa patients from the Karmanos Cancer Institute in Detroit MI, USA using the Exiqon microRNA PCR Cancer Focus panel. Expression levels are normalized using the median polish method [23]. Full details on the cohort, array and pre-processing steps are found in a separate manuscript [24]. As a consequence of the median polish, the median expression within each miRNA will be 0. For this paper we utilize the subset of 93 men with self-identified AA race and the subset of 44 miRNAs with detectable levels on all of them. Of the 93 men, 68 have non-aggressive PCa (defined as a Gleason score of 7 (3 + 4) or lower); while 25 have aggressive PCa (defined as a Gleason score of 7 (4 + 3) or higher). The objective is to identify miRNAs that are associated with disease aggressiveness.

miRNA replication studies

In the absence of publically available circulating miRNA datasets on prostate cancer patients, we chose to use two other studies that examined miRNAs expressed in tumor tissue.

One study (denoted Taylor study) [25] examined miRNAs (on the Agilent-019118 human miRNA microarray) expressed in metastatic and pri-

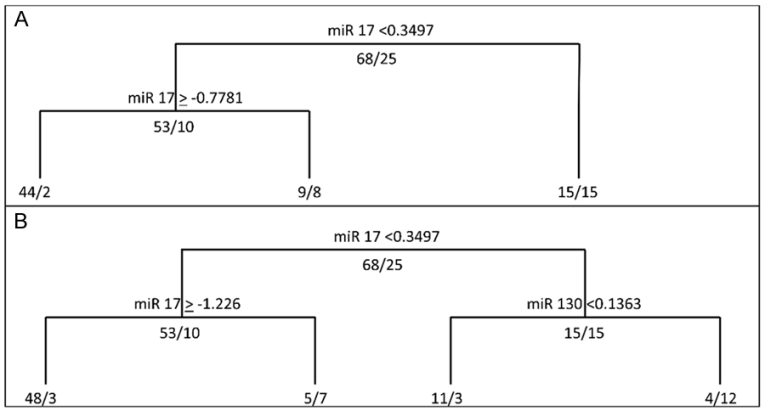


Figure 1. Tree diagrams of the RP-HT (A) and CART (B) analysis of the miRNA dataset. For both trees, the text above a split indicates the Boolean statement that defines the split; taking the left-hand path if the text statement is true and the right-hand path if the statement is false. The two numbers below each split indicate the <number of controls>/<number of cases> at each step in the process.

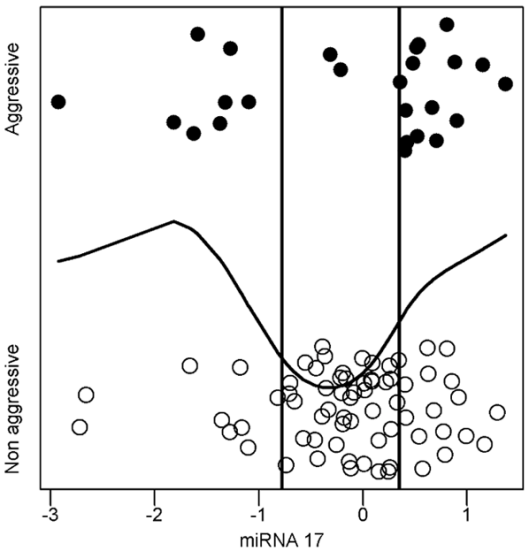


Figure 2. Scatterplot of miR-17 versus jittered aggressiveness status. The vertical lines indicate the cutpoints (-0.778, 0.350) defined by RP-HT. The group of samples at the negative and positive extrema have an aggressiveness incidence of 0.471 (8/17) and 0.500 (15/30), respectively. The middle group of samples has an incidence of 0.043 (2/46). A smoothed spline fit to the non-jittered data emphasizes the U-shaped distribution of aggressive disease incidence.

mary PCa tumors. Combining the data from the supplement to the source paper [25] and the online repository in GEO (GSE21032), we are able to define Gleason's score as we did in our study. As in our study, we normalize all the measured miRNA expression levels using median polish, although we only examine miR-17 as part of this validation exercise. Note that as only 16 of the 110 samples in the Taylor study are AA, results on the whole cohort are presented here. Therefore, although we cannot consider the Taylor study as a true 'replication' dataset, since it measures miRNAs from a different biospecimen type (tumor tissue) in a (mostly) different population using a different assay, we will determine if any significant findings from our miRNA analysis is also detected in this dataset. There are 76 Taylor samples with a Gleason score of 7 (3 + 4) or less and 34 samples with a Gleason score of 7 (4 + 3) or more.

Prostate cancer samples from the Cancer Genome Atlas (TCGA) had miRNA expression quantification measured via miRNA-Seq [26] as part of the data collection process. Linking the phenotype data downloaded from cBioPortal [27], we are able to define Gleason score as we did in our study. Unfortunately (for our validation purposes), only 7 (5 with Gleason score of 7 (3 + 4) or lower and 2 with Gleason score of 7 (4 + 3) or higher) of these samples were AA, while 146 (78 with Gleason score of 7 (3 + 4) or lower and 68 with Gleason score of 7 (4 + 3) or higher) were EA. Like the Taylor study, the TCGA study also cannot be considered a true 'replication' as it is measured in tumor tissue from a diverse population using sequencing (all different than our study). As in our study, we normalize all the measured miRNA expression levels using median polish, although we only examine miR-17 as part of this validation exercise.

Results

The optimal support parameter was identified as 0.26 for this dataset. Three subgroups were created, defined by a single circulating miRNA: miR-17. A CART analysis identified 4 subgroups, defined by 2 miRNAs. A comparison of the RP-HT and CART models is shown in **Figure 1**. **Figure 2** displays a scatterplot of miR-17 versus jittered disease aggressiveness and the identi-

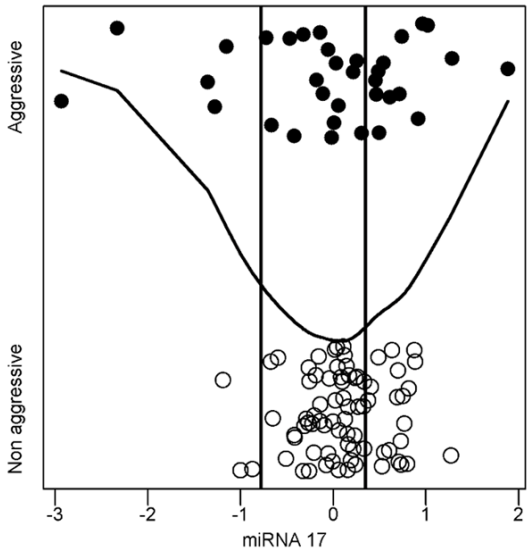


Figure 3. Scatterplot of miR-17 versus jittered aggressiveness status for the Taylor study. A similar U-shaped pattern of aggressiveness incidence is observed in the Taylor dataset. The curve is a smooth spline fit to the non-jittered data, while the vertical lines indicate the thresholds derived from our dataset (-0.778, 0.350). The group of samples at the negative and positive extrema have an aggressiveness incidence of 0.625 (5/8) and 0.406 (13/32), respectively. The middle group of samples has an incidence of 0.229 (16/70).

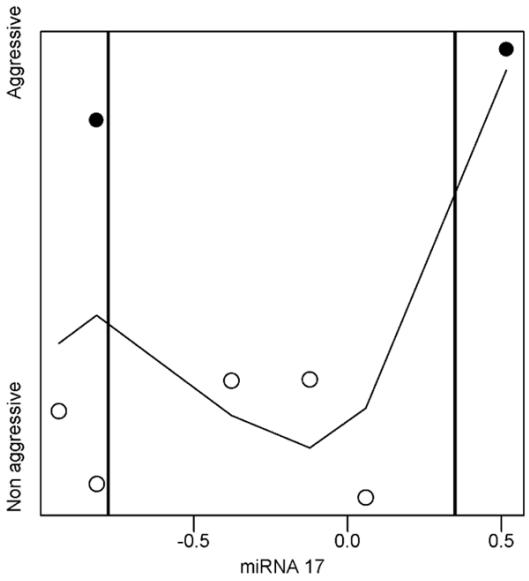


Figure 4. Scatterplot of miR-17 versus jittered aggressiveness status for the AA TCGA study. A similar U-shaped pattern of aggressiveness incidence is observed in the AA TCGA dataset. The curve is a smooth spline fit to the non-jittered data, while the vertical lines indicate the thresholds derived from our dataset (-0.778, 0.350). The group of samples at the negative and positive extrema have an aggressiveness incidence of 0.333 (1/3) and 1.000 (1/1), respectively. The middle group of samples has an incidence of 0.000 (0/3).

fied RP-HT thresholds. The left-hand subgroup consists of 17 individuals, 8 with aggressive disease and 9 without (incidence = 0.47). The right-hand subgroup consists of 30 individuals, half with and half without aggressive PCa (incidence = 0.50). The sample of 46 individuals in the middle of the distribution contains only 2 aggressive cancers (incidence = 0.043). A smoothed spline fit to the non-jittered data (using 0 and 1 as the response variable) emphasizes the U-shaped distribution of aggressive disease incidence.

Taylor dataset

Figure 3 shows a plot of miR-17 versus jittered disease aggressiveness for the Taylor dataset. Utilizing the same thresholds from the analysis of our dataset, we find that the group of samples at the negative and positive extrema have an aggressiveness incidence of 0.625 (5/8) and 0.406 (13/32), respectively. The middle group of samples has an incidence of 0.229 (16/70). The fitted smooth spline (using the non-jittered response) shows a similar, albeit weaker, U-shaped pattern of incidence (even after excluding the right-most point); although the race categorization and biospecimen type were different between the 2 studies. Appling the CART-defined tree to this dataset yielded poor prediction, as miR-130 was not discriminatory (data not shown).

Figure 4 shows a plot of miR-17 versus jittered disease aggressiveness for the TCGA AA subset. Utilizing the same thresholds from the analysis of our dataset, we find that the group of samples at the negative and positive extrema have an aggressiveness incidence of 0.333 (1/3) and 1.000 (1/1), respectively. The middle group of samples has an incidence of 0.000 (0/3). We obviously cannot read too much into this result as the sample size is much too small, but the result is at least promising. The U-shaped pattern was not observed in the EA subset (data not shown).

A simulation study was undertaken to estimate an experiment-wide error rate. To that end, we computed the 5th percentile (out of 1000 iterations) of the logistic regression *p*-value for null

data sets (the predictor variables are not associated with the binary response) under various conditions. The detailed results from the simulation are in Supplementary Table 1. For our miRNA database, utilizing 1,000 iterations with 68 "0" s, 25 "1" s and 44 null predictor variables, we found that 49 (5%) resulted in a smaller or equal *p*-value than observed in the data analysis of the actual data set. A total of 909 (91%) resulted in an outputted tree; while CART produced a tree for each iteration. So even when there is no association between predictors and response, 90% of the null RP-HT analyses and 100% of the null CART analyses returned a tree, highlighting the need for computation of an experiment-wide error rate, regardless of whether an external validation dataset is available.

Discussion

RP-HT overcomes obstacles inherent in an RP analysis including the lack of hypothesis testing of each potential split, the lack of a consistent mechanism to stop growing the tree and the need for subjective input parameters, and instead a nominal *p*-value is the only requirement. Overfitting will still be an issue (typical of greedy classification models); hence the need for *p*-value calculation, external replication, and validation of any significant finding. The RP-HT technique identified a U-shaped pattern in aggressive disease as a function of miR-17, whereby samples at the positive (upregulation) and negative (downregulation) extrema were associated with aggressive PCa. This duality is consistent with the contradictory findings on miR-17's role in PCa progression, as illustrated by studies reporting both its upregulation and downregulation in cells and tissue [28, 29]. miR-17 belongs to the miR-17-92 cluster, which is implicated in the regulation of many oncogenes and tumor suppressor genes involved in distinct and opposing pathways [30]. This duality encapsulates the complexities of cancer progression and enables miR-17-92 to fine-tune the levels of pro-apoptotic and anti-apoptotic genes respectively.

Using a systems biology approach, Cloonan, et al. [31] demonstrated this dichotomy by unmasking an extensive yet contradictory network of genes regulated by miR-17-5p, including inhibitors of cellular proliferation such as TSG101, RBL1 and MAPK9 and promoters of cellular proliferation such as MYCN, NCOA3 and NR4A3. Other studies have similarly captured this duality. For instance, miR-17 upregulation was reported to enhance tumor growth by directly repressing PTEN [32] and TIMP3 [33] and to convey chemoresistance to cisplatin in human PCa cell lines [28]. Conversely, others observed a tumor suppressive role for miR-17-5p in PCa cells, revealing that miR17 loss induced progression to anti-androgen resistance by upregulating FGD4, LIMK1, cyclin D1 and SSH1 cell cycle regulatory proteins [29]. Other targets of miR-17, such as E2 F1, have also been reported to function as oncogenes or tumor suppressor genes depending on the nature of the other driving oncogenic mutations present [5]. These contrasting findings highlight the contextual character of miR-17 function and stress the importance of considering cancerous phenotypes as a product of clonally heterogeneous and complex genetic networks subject to multiple layers of dysregulation [31].

Several limitations exist for our study. Firstly, the RP-HT takes longer to run than a standard CART analysis, as it analyzes 91 potential support parameters. The CART analysis of our miRNA dataset was instantaneous using rpart, while HT-RP took approximately 40 seconds to complete. Datasets with a larger number of variables will intensify this issue, while utilizing parallel processing on multiple computing nodes will mitigate it. Secondly, none of the publically available miRNA PCa datasets had a majority of AA samples and had measurements of circulating miRNA. Thus we were compelled to verify our miR-17 result using a dataset of PCa men with mostly European ancestry, measured from tumor samples and a small number of AA TCGA samples. Checking the incidence of aggressive PCa relative to miR-17 in another circulating miRNA dataset of AA men will be necessary before miR-17 is validated as a biomarker for aggressive PCa. Lastly, the sample sizes of both our discovery and replication datasets were small, increasing the chance of a spurious result. Although the procedure we utilized in this paper, including the estimation of an overall experiment-wise error rate and external replication, should lessen that chance.

In conclusion, utilizing a new statistical tool, we have observed that both extrema of circulating miR-17 are associated with aggressive prostate cancer. This effect was observed in tumor

samples from separate datasets measured on different assays representing different populations of prostate cancer patients. Our result is consistent with the conflicting findings on the impact that miR-17 has in PCa progression, namely that it controls both oncogenic and tumor-suppressive genes.

Acknowledgements

This work was supported by the National Institutes of Health Cancer Center Support Grant 5P30CA022453, Department of Defense grants W81XWH-09-1-0203 and W81XWH-13-1-0477 & Michigan Fund for Cancer Research.

Disclosure of conflict of interest

None.

Address correspondence to: Greg Dyson, Karmanos Cancer Institute and Department of Oncology, Wayne State University, Detroit MI, USA. E-mail: dysong@karmanos.org

References

[1] Torre LA, Siegel RL, Ward EM, Jemal A. Global cancer incidence and mortality rates and trends--an update. Cancer Epidemiol Biomarkers Prev 2016; 25: 16-27.

[2] Zhou CK, Check DP, Lortet-Tieulent J, Laversanne M, Jemal A, Ferlay J, Bray F, Cook MB, Devesa SS. Prostate cancer incidence in 43 populations worldwide: an analysis of time trends overall and by age group. Int J Cancer 2016; 138: 1388-400.

[3] Vanacore D, Boccellino M, Rossetti S, Cavaliere C, D'Aniello C, Di Franco R, Romano FJ, Montanari M, La Mantia E, Piscitelli R, Nocerino F, Cappuccio F, Grimaldi G, Izzo A, Castaldo L, Pepe MF, Malzone MG, Iovane G, Ametrano G, Stiuso P, Quagliuolo L, Barberio D, Perdona S, Muto P, Montella M, Maiolino P, Veneziani BM, Botti G, Caraglia M, Facchini G. Micrornas in prostate cancer: an overview. Oncotarget 2017; 8: 50240-50251.

[4] Wang BD, Ceniccola K, Yang Q, Andrawis R, Patel V, Ji Y, Rhim J, Olender J, Popratiloff A, Latham P, Lai Y, Patierno SR, Lee NH. Identification and functional validation of reciprocal microRNA-mRNA pairings in african american prostate cancer disparities. Clin Cancer Res 2015; 21: 4970-84.

[5] Ambs S, Prueitt RL, Yi M, Hudson RS, Howe TM, Petrocca F, Wallace TA, Liu CG, Volinia S, Calin GA, Yfantis HG, Stephens RM, Croce CM. Genomic profiling of microRNA and messenger RNA reveals deregulated microRNA expression in prostate cancer. Cancer Res 2008; 68: 6162-70.

[6] Yates C, Long MD, Campbell MJ, Suchestom-Campbell L. miRNAs as drivers of TMPRSS2-ERG negative prostate tumors in African American men. Front Biosci (Landmark Ed) 2017; 22: 212-29.

[7] Loh WY. Fifty years of classification and regression trees. Int Stat Rev 2014; 82: 329-48.

[8] Morgan JN, Sonquist JA. Problems in analysis of survey data, and a proposal. Journal of the American Statistical Association 1963; 58: 415.

[9] Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and regression trees. Belmont, Calif: Wadsworth International Group; 1984.

[10] Venables WN, Ripley BD, Venables WN. Modern applied statistics with S. New York: Springer, 2002.

[11] Cappelli C, Mola F, Siciliano R. A statistical approach to growing a reliable honest tree. Computational Statistics & Data Analysis 2002; 38: 285-99.

[12] Zhang H, Singer B. Recursive partitioning in the health sciences. New York: Springer, 1999.

[13] Hothorn T, Hornik K, Zeileis A. Unbiased recursive partitioning: a conditional inference framework. Journal of Computational and Graphical Statistics 2006; 15: 651-74.

[14] Breiman L. Random forests. Machine Learning 2001; 45: 5-32.

[15] Loh WY. Improving the precision of classification trees. Annals of Applied Statistics 2009; 3: 1710-37.

[16] Dyson G, Sing CF. Efficient identification of context dependent subgroups of risk from genome-wide association studies. Stat Appl Genet Mol Biol 2014; 13: 217-26.

[17] Friedman JH, Fisher NI. Bump hunting in high-dimensional data. Statistics and Computing 1999; 9: 123-43.

[18] Curtis SA. The classification of greedy algorithms. Science of Computer Programming 2003; 49: 125-57.

[19] Kendall MG. A new measure of rank correlation. Biometrika 1938; 30: 81-93.

[20] Therneau TM, Atkinson EJ and Ripley B. Rpart: recursive partitioning. R package version 4.1-10. https://CRAN.R-project.org/package=rpart 2015.

[21] R Core Team. R: a language and environment for statistical computing. 2017.

[22] Therneau TM, Atkinson EJ. An introduction to recursive partitioning using the RPART routine. Technical Report. Section of Statistics, Mayo Clinic 1997.

[23] Tukey JW. Exploratory data analysis. Addison-Wesley: Reading, Massachussetts, 1977.

[24] Farran B, Dyson G, Craig D, Dombkowski A, Beebe-Dimmer JL, Powell IJ, Podgorski I, Heilbrun L, Bolton S, Bock CH. A study of circulating microRNAs identifies a new potential biomarker panel to distinguish aggressive prostate cancer. Carcinogenesis 2018; 39: 556-61.
[25] Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao YH, Carver BS, Arora VK, Kaushik P, Cerami E, Reva B, Antipin Y, Mitsiades N, Landers T, Dolgalev I, Major JE, Wilson M, Socci ND, Lash AE, Heguy A, Eastham JA, Scher HI, Reuter VE, Scardino PT, Sander C, Sawyers CL, Gerald WL. Integrative genomic profiling of human prostate cancer. Cancer Cell 2010; 18: 11-22.
[26] Cancer Genome Atlas Research Network. The molecular taxonomy of primary prostate cancer. Cell 2015; 163: 1011-25.
[27] Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012; 2: 401-4.
[28] Zhou P, Ma L, Zhou J, Jiang M, Rao E, Zhao Y, Guo F. miR-17-92 plays an oncogenic role and conveys chemo-resistance to cisplatin in human prostate cancer cells. Int J Oncol 2016; 48: 1737-48.
[29] Ottman R, Levy J, Grizzle WE, Chakrabarti R. The other face of miR-17-92a cluster, exhibiting tumor suppressor effects in prostate cancer. Oncotarget 2016; 7: 73739-53.
[30] Xiang J, Wu J. Feud or friend? The role of the miR-17-92 cluster in tumorigenesis. Curr Genomics 2010; 11: 129-35.
[31] Cloonan N, Brown MK, Steptoe AL, Wani S, Chan WL, Forrest AR, Kolle G, Gabrielli B, Grimmond SM. The miR-17-5p microRNA is a key regulator of the G1/S phase cell cycle transition. Genome Biol 2008; 9: R127.
[32] Dhar S, Kumar A, Rimando AM, Zhang X, Levenson AS. Resveratrol and pterostilbene epigenetically restore PTEN expression by targeting oncomiRs of the miR-17 family in prostate cancer. Oncotarget 2015; 6: 27214-26.
[33] Yang X, Du WW, Li H, Liu F, Khorshidi A, Rutnam ZJ, Yang BB. Both mature miR-17-5p and passenger strand miR-17-3p target TIMP3 and induce prostate tumor growth and invasion. Nucleic Acids Res 2013; 41: 9688-704.

Supplemental methods

Null hypothesis derivation

Categorical responses: The derivation of the null distribution for RP-HT for a categorical response is as follows. Let V indicate the set of valid potential splits (defined by a continuous or categorical explanatoryvariable): subsets that are sized between [$n \times \beta$] and n-[$n \times \beta$] that have a relative frequency ofa response category greater than the input relative frequency of that same response category. Let m_1 denote the frequency of the response category of interest in a sample of size m taken without replacement from the population of size n which has n_1 of the response category of interest. Then the density function of the null distribution [$f(\theta|V)$] as a function of the observed relative frequency values (θ) is defined as:

$$\begin{aligned} f(\theta \mid V) &= \sum_{j=[\beta\times n]}^{n-[\beta\times n]} \{f(\theta, m = j \mid V)\} \\ &= \sum_{j=[\beta\times n]}^{n-[\beta\times n]} \{f(\theta \mid m = j, V) \times p(m = j \mid V)\} \\ &= \sum_{j=[\beta\times n]}^{n-[\beta\times n]} \{f\left(\frac{m_1}{m} \mid m = j, V\right) \times p(m = j \mid V)\} \\ &= \sum_{j=[\beta\times n]}^{n-[\beta\times n]} \{f(m_1 \mid m = j, V) \times p(m = j \mid V)\}, \end{aligned}$$

where m is assumed the be marginally discrete uniform and m_1 follows a hypergeometric distribution. Additionally, define N as the number of potential valid subsets for a response category and $F(\theta|V)$ as the cumulative distribution function of θ values that satisfy V for a response category. The algorithm then proceeds to enumerate all possible combinations of m_1 and m that satisfy V to derive the null distribution. As the number of observable combinations could be very large, we used 10,000 values evenly spaced between the input relative frequency and maximal observable relative frequency to estimate the density functions for a response category. The computed density is rounded up to the nearest of these 10,000 categories for storage. Then the null distribution of the maximum relative frequency is given by $N \times f(\theta|V) \times F(\theta|V)^{N-1}$ for a response category.

From the null distribution for each of the response categories, a p-value is computed given the observed maximal relative frequency obtained from each response category. If multiple splits within a response category result in the same maximal response, the split with the larger number of observations is selected. If there are multiple splits with the same maximal response and number of observations, one of the splits is chosen at random to be selected. The split/response category combination that results in the smallest p-value is then outputted. If multiple splits/response category combinations result in the same minimum p-value, a similar algorithm examining the relative frequency and number of observations is used to select the outputted split/response category, randomly outputting a split if all elements are tied. As the response categories will likely have different input relative frequencies, a single null distribution is not possible. However, to allay concerns about multiple testing, the outputted p-value is multiplied by the number of response classes as a Bonferroni correction. The simulation study in section 5 explores this issue of multiple testing in further detail.

Continuous responses: We rank the response breaking ties randomly, since tied ranks are not consistent with the distribution theory requirements used in the hypothesis testing component of the RP-HT algorithm. As a consequence, continuous response variables with tied values may result in different resultant RP-HT objects, even with the same support parameter. The responses are re-ranked at each node in the tree. Therefore, the rank for the same observation will change at each node.

Mathematically, the objective function at each node is to select from the input set of integers from 1 to n, a subset of size m with a statistically significantly higher mean than in the input population. Potential valid subsets (denoted as V) must be sized between $[n * \beta]$ and n-$[n * \beta]$ and have a mean response greater than the input mean $(n + 1)/2$. The null distribution for this algorithm will be either a weighted enumeration of all valid subsets or a normal approximation if n gets too large. The enumeration of valid subsets is weighted to ensure that the marginal distribution of m is discrete uniform. Other formulations of the marginal distribution of m can be entertained, including an unweighted enumeration which is equivalent to treating all possible combinations of subsets across all possible m values as equally likely. Once the approximation gets close enough (in terms of Kullbeck-Leibler [KL] divergence) to the exact distribution, then RP-HT will switch over to the approximation to save computational time and resources. The exact calculation involves enumerating the mean responses of all possible valid subsets given n and β to construct the null distribution. As with the categorical responses, the enumerated mean values are rounded up to the nearest of the 10,000 values for storage. Finally, the distributions for all possible m values are combined to produce a single density function.

We use normal distributional properties of order statistics to construct the approximation. A randomly selected subset of size m taken from a set of integers from 1 to n without replacement has a normal distribution with mean $(n + 1)/2$ and variance $(n\text{-}m) \times (n + 1)/(12 \times m)$ in the limit. The observable mean responses (θ) given n and m that satisfy V will be the midpoints of the nonoverlapping intervals of width $1/m$ used to estimate the density. The probability that the above normal distribution falls within each of these intervals corresponding to each observable θ is then computed. This value is then scaled by the relative frequency of the marginal distribution of m for valid subsets to ensure that m is marginally discrete uniform. As with the categorical response, 10,000 values evenly spaced between the mean of the input response and maximal observable response across the range of m values are used to estimate the density function. Computed density values from the proper normal distribution are rounded up to the nearest of the 10,000 values for storage. Finally, the densities for all possible m values are combined to produce a single density function. When m is less than 4 or greater than n-4, the exact calculation is used since the approximation poorly matches the exact distribution near the extrema of m. When n is greater than or equal to 22, the RP-HT will use the approximation rather than the exact computation as the KL difference between the two distributions is less than 0.001 for the range of possible support parameters (data not shown). Since the objective function selects the subset with the largest θ from all valid subsets, the null probability distribution function of the largest θ is given by the probability density function of the maximum order statistic of the null distribution, namely $N \times f(\theta|V) \times F(\theta|V)^{N-1}$, where N is the number of potential valid subsets for one of the response categories and $f(\theta|V)$ and $F(\theta|V)$ are the distribution and cumulative distribution functions of observable θ values that satisfy V.

Supplementary Table 1. Estimated experiment-wide error rate utilizing a variety of number of observations, percentage of cases, covariates, alpha level and splits. The first row in the table indicates the result for the dataset analyzed in the manuscript

Number of observations	Percentage of 'cases'	Number of covariates	Alpha	Number of splits	5% experiment wide error threshold
93	0.27	44	0.05	2	1.745E-06
50	0.25	20	0.01	1	1.321E-04
100	0.25	20	0.01	1	1.272E-04
200	0.25	20	0.01	1	8.903E-05
50	0.5	20	0.01	1	1.425E-04
100	0.5	20	0.01	1	7.889E-05
200	0.5	20	0.01	1	6.714E-05
50	0.25	50	0.01	1	5.676E-05
100	0.25	50	0.01	1	3.944E-05
200	0.25	50	0.01	1	3.354E-05
50	0.5	50	0.01	1	5.157E-05

100	0.5	50	0.01	1	3.653E-05
200	0.5	50	0.01	1	3.025E-05
50	0.25	100	0.01	1	5.064E-05
100	0.25	100	0.01	1	1.973E-05
200	0.25	100	0.01	1	1.333E-05
50	0.5	100	0.01	1	1.797E-05
100	0.5	100	0.01	1	1.584E-05
200	0.5	100	0.01	1	1.095E-05
50	0.25	20	0.05	1	1.838E-04
100	0.25	20	0.05	1	1.272E-04
200	0.25	20	0.05	1	8.060E-05
50	0.5	20	0.05	1	1.425E-04
100	0.5	20	0.05	1	1.089E-04
200	0.5	20	0.05	1	6.714E-05
50	0.25	50	0.05	1	5.676E-05
100	0.25	50	0.05	1	6.007E-05
200	0.25	50	0.05	1	2.915E-05
50	0.5	50	0.05	1	5.157E-05
100	0.5	50	0.05	1	4.585E-05
200	0.5	50	0.05	1	2.476E-05
50	0.25	100	0.05	1	3.119E-05
100	0.25	100	0.05	1	2.403E-05
200	0.25	100	0.05	1	2.735E-05
50	0.5	100	0.05	1	2.710E-05
100	0.5	100	0.05	1	1.910E-05
200	0.5	100	0.05	1	1.581E-05
50	0.25	20	0.01	2	7.108E-05
100	0.25	20	0.01	2	3.981E-05
200	0.25	20	0.01	2	9.983E-06
50	0.5	20	0.01	2	1.259E-04
100	0.5	20	0.01	2	1.357E-05
200	0.5	20	0.01	2	1.282E-05
50	0.25	50	0.01	2	5.676E-05
100	0.25	50	0.01	2	5.723E-06
200	0.25	50	0.01	2	1.284E-06
50	0.5	50	0.01	2	5.157E-05
100	0.5	50	0.01	2	3.233E-06
200	0.5	50	0.01	2	1.383E-06
50	0.25	100	0.01	2	3.119E-05
100	0.25	100	0.01	2	1.676E-06
200	0.25	100	0.01	2	4.753E-07
50	0.5	100	0.01	2	1.797E-05
100	0.5	100	0.01	2	4.851E-07
200	0.5	100	0.01	2	3.181E-07
50	0.25	20	0.05	2	1.374E-05
100	0.25	20	0.05	2	4.405E-06
200	0.25	20	0.05	2	3.422E-06
50	0.5	20	0.05	2	8.345E-06
100	0.5	20	0.05	2	3.139E-06

200	0.5	20	0.05	2	2.262E-06
50	0.25	50	0.05	2	3.019E-06
100	0.25	50	0.05	2	1.104E-06
200	0.25	50	0.05	2	7.479E-07
50	0.5	50	0.05	2	1.337E-06
100	0.5	50	0.05	2	3.871E-07
200	0.5	50	0.05	2	2.384E-07
50	0.25	100	0.05	2	7.422E-07
100	0.25	100	0.05	2	1.796E-07
200	0.25	100	0.05	2	1.467E-07
50	0.5	100	0.05	2	4.355E-07
100	0.5	100	0.05	2	1.034E-07
200	0.5	100	0.05	2	1.270E-07
50	0.25	20	0.01	3	1.229E-04
100	0.25	20	0.01	3	1.099E-05
200	0.25	20	0.01	3	3.298E-06
50	0.5	20	0.01	3	1.259E-04
100	0.5	20	0.01	3	9.955E-06
200	0.5	20	0.01	3	3.070E-06
50	0.25	50	0.01	3	5.676E-05
100	0.25	50	0.01	3	3.338E-06
200	0.25	50	0.01	3	6.830E-07
50	0.5	50	0.01	3	4.828E-05
100	0.5	50	0.01	3	1.939E-06
200	0.5	50	0.01	3	4.971E-07
50	0.25	100	0.01	3	1.506E-05
100	0.25	100	0.01	3	7.059E-07
200	0.25	100	0.01	3	1.307E-07
50	0.5	100	0.01	3	1.504E-05
100	0.5	100	0.01	3	1.489E-07
200	0.5	100	0.01	3	9.257E-08
50	0.25	20	0.05	3	3.805E-06
100	0.25	20	0.05	3	1.698E-06
200	0.25	20	0.05	3	5.017E-07
50	0.5	20	0.05	3	2.186E-06
100	0.5	20	0.05	3	2.087E-07
200	0.5	20	0.05	3	1.404E-07
50	0.25	50	0.05	3	6.965E-07
100	0.25	50	0.05	3	9.309E-08
200	0.25	50	0.05	3	2.502E-08
50	0.5	50	0.05	3	2.210E-07
100	0.5	50	0.05	3	3.725E-08
200	0.5	50	0.05	3	1.457E-08
50	0.25	100	0.05	3	8.472E-08
100	0.25	100	0.05	3	2.410E-08
200	0.25	100	0.05	3	3.982E-09
50	0.5	100	0.05	3	9.685E-08
100	0.5	100	0.05	3	3.474E-09
200	0.5	100	0.05	3	2.419E-09

Appendix 3

SNP	Gene	Pooled adjusted effect	Pooled adjusted p-value	AA adjusted effect	AA adjusted p-value	EA adjusted effect	EA adjusted p-value	AA minor allele count	EA minor allele count
rs75271503	AGO1	-0.0705	0.877	0.5048	0.604	-0.5117	0.330	5	18
rs76379471	AGO1	0.5413	0.052	0.5234	0.066	15.0959	0.162	85	1
rs143228790	AGO1	1.5576	0.216	1.8606	0.240	14.1559	0.321	2	1
rs12727197	AGO1	-0.0967	0.843	0.7085	0.480	-0.3081	0.587	5	15
rs636832	AGO1	0.0205	0.934	-0.2229	0.470	0.5345	0.235	209	29
rs34002152	AGO1	0.1471	0.823	-0.2866	0.818	0.3381	0.671	3	7
rs595961	AGO1	-0.0904	0.805	-0.1150	0.765	0.2118	0.880	104	149
rs528594571	AGO1	1.2128	0.127	1.2182	0.131	NA	NA	7	0
rs12751941	AGO1	0.0560	0.845	0.4382	0.363	-0.1579	0.658	26	51
rs116344388	AGO1	-0.0226	0.966	-0.0265	0.960	NA	NA	18	0
rs3738356	AGO1	-0.2918	0.320	-0.3652	0.239	0.4257	0.738	69	6
rs61751004	AGO1	-2.2456	0.008	0.4837	0.753	-17.9044	0.000	2	8
rs41308339	AGO1	0.3563	0.584	-0.8502	0.416	2.0668	0.053	6	5
rs7527165	TARBP1	0.3146	0.175	0.4846	0.085	0.0128	0.976	193	123
rs6669190	TARBP1	0.3460	0.104	0.4608	0.077	0.1512	0.692	136	112
rs1802871	TARBP1	-0.4468	0.068	-0.2246	0.452	-0.8352	0.074	208	27
rs2175593	TARBP1	0.3085	0.177	0.5854	0.044	-0.1461	0.710	209	118
rs143561161	TARBP1	0.5384	0.359	0.5658	0.337	NA	NA	13	0
rs12031728	TARBP1	-0.0989	0.699	-0.6252	0.116	0.3173	0.384	40	47
rs116094080	TARBP1	0.4185	0.375	0.3937	0.406	NA	NA	21	0
rs73101929	TARBP1	0.0861	0.833	0.0286	0.945	NA	NA	31	0
rs3820605	TARBP1	-0.0035	0.989	-0.3200	0.379	0.2645	0.463	46	47
rs12137933	TARBP1	0.6738	0.065	0.2677	0.636	1.1276	0.030	16	20
rs12073170	TARBP1	0.1467	0.552	0.0262	0.928	0.5367	0.270	74	22
rs17517734	TARBP1	-0.5229	0.067	-0.5616	0.248	-0.4789	0.188	28	48
rs74147465	TARBP1	-0.1660	0.580	-0.4710	0.267	0.2072	0.653	33	26
rs270522	TARBP1	0.5279	0.013	0.7612	0.007	0.2111	0.542	194	100
rs61825881	TARBP1	-0.3170	0.406	-1.1426	0.250	-0.2364	0.582	7	29
rs750813	TARBP1	-0.4996	0.025	-0.4533	0.135	-0.5511	0.104	74	70
rs74869729	TARBP1	0.1096	0.855	0.0670	0.911	NA	NA	13	0
rs2273871	TARBP1	0.1997	0.400	0.2329	0.414	0.1693	0.701	79	28
rs78047188	TARBP1	-0.1965	0.701	-0.0237	0.964	-14.9275	0.186	20	1
rs111283864	TARBP1	-0.8156	0.151	0.4184	0.659	-1.4565	0.047	6	11
rs915184	TARBP1	0.1202	0.597	0.1395	0.609	0.0745	0.862	92	31
rs139597411	TARBP1	-12.9735	0.350	NA	NA	-14.2873	0.292	0	1
rs150153532	TARBP1	-0.7578	0.221	-0.7657	0.221	NA	NA	16	0
rs6682410	TARBP1	0.4449	0.051	0.4573	0.139	0.3525	0.319	60	56
rs72763886	TARBP1	0.3403	0.404	0.6196	0.396	0.0908	0.857	9	20
rs2273872	TARBP1	-13.1972	0.307	NA	NA	-14.7677	0.211	0	1
rs116789741	TARBP1	-0.1195	0.892	0.0591	0.950	-14.5304	0.322	5	1
rs111585333	TARBP1	0.0123	0.983	-0.0449	0.958	0.0415	0.959	7	7
rs116387638	TARBP1	0.0752	0.709	0.0699	0.785	0.0227	0.947	154	77

rs3768286	TARBP1	0.0806	0.688	0.0651	0.800	0.0521	0.877	155	77
rs4920246	TARBP1	0.1113	0.581	0.1284	0.618	0.0188	0.956	161	78
rs139816698	TARBP1	-0.8156	0.151	0.4184	0.659	-1.4565	0.047	6	11
rs10489896	TARBP1	-0.0138	0.945	-0.0737	0.772	0.0083	0.980	143	78
rs270524	TARBP1	0.0567	0.798	0.0575	0.831	0.0871	0.831	180	37
rs1619856	TARBP1	0.2137	0.332	0.0248	0.925	0.6737	0.105	138	120
rs4920247	TARBP1	0.7152	0.003	0.7559	0.028	0.6987	0.040	54	62
rs17378272	TARBP1	1.0810	0.045	2.1534	0.033	0.5857	0.366	5	11
rs270519	TARBP1	-0.1286	0.573	-0.0526	0.839	-0.3600	0.481	134	21
rs12068114	TARBP1	-1.0954	0.256	-1.2453	0.193	NA	NA	8	0
rs35562024	TARBP1	-0.1897	0.600	-0.1603	0.661	NA	NA	43	0
rs2273876	TARBP1	0.3137	0.287	0.3422	0.252	-14.4225	0.339	74	1
rs3754308	TARBP1	0.3099	0.278	0.4206	0.192	-0.1629	0.800	57	11
rs116085005	TARBP1	-1.6338	0.059	-1.6846	0.051	NA	NA	11	0
rs76838713	TARBP1	0.4333	0.387	0.4667	0.355	NA	NA	19	0
rs270505	TARBP1	0.0549	0.803	0.2457	0.349	-0.5564	0.200	144	125
rs10910438	TARBP1	-0.2100	0.367	-0.0445	0.867	-0.6325	0.223	127	21
rs10910439	TARBP1	0.2751	0.570	0.3771	0.446	-14.9087	0.189	20	1
rs3736838	TARBP1	-0.0282	0.925	-0.0285	0.932	0.2486	0.753	52	11
rs17378453	TARBP1	-0.9844	0.198	-14.9592	0.067	-0.3727	0.688	4	5
rs57114612	TARBP1	-0.2700	0.394	-0.2709	0.448	0.0326	0.969	46	10
rs60683618	TARBP1	-0.2270	0.318	-0.1633	0.540	-0.2017	0.665	169	27
rs6694349	TARBP1	-0.2091	0.384	-0.0498	0.854	-0.6367	0.254	117	18
rs3754307	TARBP1	-0.4633	0.115	-0.5168	0.124	-0.1158	0.870	57	13
rs6697342	TARBP1	0.2564	0.559	0.3747	0.404	-14.4013	0.342	25	1
rs271776	TARBP1	0.1265	0.558	0.2603	0.319	-0.2027	0.612	134	118
rs75859973	TARBP1	-0.2957	0.316	-0.1061	0.748	-0.9548	0.166	56	11
rs12083221	TARBP1	0.3278	0.499	0.2766	0.570	NA	NA	21	0
rs528774	TARBP1	0.5006	0.018	0.2731	0.309	0.8416	0.017	186	93
rs34912197	TARBP1	0.2894	0.268	0.5300	0.097	-0.2804	0.540	55	24
rs544322	TARBP1	0.0237	0.915	0.2994	0.262	-0.5297	0.212	165	125
rs12082553	TARBP1	0.0392	0.883	-0.0105	0.969	16.3116	0.044	123	3
rs4920255	TARBP1	-0.0792	0.733	0.0599	0.821	-0.3633	0.496	143	21
rs4920256	TARBP1	-0.0897	0.750	0.1433	0.673	-0.5240	0.342	50	20
rs111299145	TARBP1	0.1905	0.577	0.1056	0.761	34.4027	0.012	47	1
rs512643	TARBP1	0.2703	0.260	0.7080	0.020	-0.4918	0.224	93	118
rs56394394	TARBP1	0.2368	0.537	-0.0220	0.973	0.4702	0.343	13	20
rs271748	TARBP1	-0.2054	0.345	-0.0427	0.878	-0.4388	0.216	100	100
rs12080320	TARBP1	-0.2203	0.490	-0.1446	0.656	-16.6609	0.092	58	2
rs78315114	TARBP1	-0.3249	0.453	-0.2921	0.506	-14.4225	0.339	29	1
rs45927	TARBP1	-0.2094	0.320	-0.0862	0.738	-0.4287	0.257	154	108
rs12085070	TARBP1	-0.1226	0.630	-0.0863	0.793	-0.2242	0.585	54	32
rs3951276	TARBP1	0.2810	0.168	0.3629	0.166	0.1820	0.588	171	85
rs79300169	TARBP1	-0.2976	0.477	-0.3255	0.437	NA	NA	31	0
rs140423225	TARBP1	0.0292	0.959	0.0264	0.963	NA	NA	15	0
rs11587891	TARBP1	0.1355	0.667	0.3523	0.342	-0.3516	0.561	38	13
rs2175591	TARBP1	-14.0357	0.098	NA	NA	-15.2828	0.070	0	3

rs6667048	TARBP1	-0.2483	0.257	-0.1372	0.613	-0.4765	0.222	152	116
rs519250	TARBP1	0.1807	0.441	0.3367	0.284	-0.0064	0.986	61	47
rs2793861	TARBP1	0.3736	0.065	0.4628	0.070	0.2345	0.501	148	61
rs141995027	PRKRA	0.1610	0.713	0.3827	0.468	-0.1672	0.832	17	7
rs72953357	PRKRA	0.5466	0.132	0.3886	0.460	0.6705	0.196	18	18
rs334581	DROSHA	0.2876	0.177	0.5392	0.036	-0.2623	0.519	151	32
rs2330696	DROSHA	-0.0611	0.798	0.2544	0.481	-0.3860	0.253	245	68
rs116682738	DROSHA	0.7828	0.198	0.9871	0.121	-16.6651	0.187	11	1
rs75685666	DROSHA	-0.3611	0.241	-0.3343	0.286	-34.3278	0.023	67	2
rs146263195	DROSHA	-0.1430	0.818	-0.2051	0.742	NA	NA	13	0
rs116166271	DROSHA	0.4973	0.416	-13.7496	0.235	0.9021	0.206	2	10
rs458586	DROSHA	-0.1868	0.432	-0.0067	0.982	-0.5284	0.187	69	36
rs115882176	DROSHA	-0.0123	0.985	-0.1361	0.913	-0.2792	0.727	3	7
rs28488546	DROSHA	0.2666	0.361	0.2672	0.372	-0.6615	0.696	73	3
rs17403765	DROSHA	-0.0985	0.627	-0.1010	0.693	-0.1339	0.696	148	92
rs13186629	DROSHA	0.0263	0.912	0.0079	0.981	0.0630	0.858	74	100
rs76562197	DROSHA	0.4262	0.499	15.6722	0.023	-0.1586	0.823	2	10
rs77219111	DROSHA	-0.2993	0.452	-0.2634	0.678	-0.2821	0.587	14	19
rs5015011	DROSHA	-0.0571	0.812	0.0079	0.981	-0.1195	0.738	74	96
rs146917657	DROSHA	0.7573	0.303	0.4705	0.747	0.7628	0.383	2	6
rs115847896	DROSHA	1.2456	0.091	1.2737	0.085	NA	NA	8	0
rs16901060	DROSHA	0.4769	0.158	0.5041	0.139	NA	NA	48	0
rs545046	DROSHA	0.1512	0.487	0.3653	0.178	-0.1987	0.599	92	41
rs12188316	DROSHA	-0.1192	0.571	0.0607	0.817	-0.4476	0.228	179	108
rs60082636	DROSHA	0.5314	0.040	0.6429	0.042	0.1786	0.695	55	24
rs114470207	DROSHA	-0.6135	0.187	-0.6260	0.181	NA	NA	28	0
rs77504242	DROSHA	1.1166	0.037	1.1532	0.034	NA	NA	16	0
rs71627305	DROSHA	0.0409	0.903	0.1922	0.761	-0.0676	0.866	14	35
rs10719	DROSHA	-0.2778	0.279	-0.0574	0.885	-0.4126	0.240	242	60
rs3805525	DROSHA	0.0039	0.987	-0.0610	0.825	0.1619	0.730	123	127
rs2241337	DROSHA	-0.5255	0.011	-0.4569	0.082	-0.6318	0.072	131	56
rs73745744	DROSHA	-0.0373	0.914	0.0251	0.943	-16.5755	0.197	50	1
rs3805521	DROSHA	-0.0922	0.699	-0.1891	0.582	-0.0647	0.849	50	61
rs12186785	DROSHA	0.4419	0.229	0.0685	0.928	0.5541	0.199	9	28
rs624057	DROSHA	0.2649	0.330	0.1984	0.475	15.4315	0.266	131	4
rs3792830	DROSHA	-0.2143	0.385	-0.0319	0.908	-0.7758	0.174	94	17
rs10520980	DROSHA	-12.7300	0.399	-13.9978	0.347	NA	NA	1	0
rs2287584	DROSHA	-0.2675	0.244	-0.2476	0.443	-0.2746	0.417	221	70
rs2279797	DROSHA	-0.2447	0.269	-0.1684	0.555	-0.4071	0.256	82	49
rs111678029	DROSHA	0.2404	0.532	0.2124	0.585	NA	NA	37	0
rs56280959	DROSHA	-0.1404	0.517	-0.3084	0.241	0.1861	0.648	132	118
rs639174	DROSHA	-0.1217	0.599	-0.2065	0.529	-0.0449	0.894	225	74
rs4867329	DROSHA	-0.1601	0.461	-0.3656	0.163	0.2509	0.543	134	118
rs17485810	DROSHA	-0.4004	0.743	-13.5827	0.430	0.1446	0.921	1	2
rs78078120	DROSHA	-0.1435	0.761	-0.1738	0.715	NA	NA	24	0
rs16901168	DROSHA	0.0688	0.861	0.2558	0.544	-0.7182	0.553	28	4
rs7731057	DROSHA	-0.2905	0.189	-0.4082	0.175	-0.1231	0.713	216	78

rs573010	DROSHA	-0.1306	0.533	-0.1073	0.692	-0.1737	0.612	97	65
rs114420423	DROSHA	0.2105	0.742	-14.6263	0.165	0.6836	0.365	3	8
rs10067066	DROSHA	-0.0007	0.997	0.1498	0.564	-0.3898	0.389	120	26
rs140988484	DROSHA	14.8568	0.025	14.7888	0.206	15.4449	0.117	1	2
rs492176	DROSHA	0.0772	0.702	0.3206	0.211	-0.3543	0.294	135	86
rs144787103	DROSHA	0.6225	0.504	-13.9403	0.361	1.3747	0.223	1	4
rs6896679	DROSHA	0.1988	0.638	0.1585	0.740	0.3966	0.669	21	5
rs3805500	DROSHA	0.2216	0.320	0.6726	0.031	-0.3009	0.370	219	90
rs524138	DROSHA	0.1439	0.573	0.2284	0.392	-0.7062	0.477	121	6
rs10472774	DROSHA	-0.1394	0.621	-0.0658	0.829	-0.9768	0.267	203	9
rs10068052	DROSHA	0.0620	0.795	-0.1148	0.670	0.7204	0.192	133	135
rs75466289	DROSHA	0.4040	0.476	-1.0158	0.342	1.1996	0.119	5	10
rs58178446	DROSHA	0.6748	0.216	0.6902	0.211	NA	NA	16	0
rs112750085	DROSHA	-0.0668	0.914	-0.0212	0.973	NA	NA	13	0
rs12515221	DROSHA	-0.0351	0.898	0.4072	0.296	-0.4914	0.230	246	37
rs4867337	DROSHA	0.0769	0.839	0.0762	0.851	0.2620	0.810	31	4
rs7712436	DROSHA	-0.0093	0.972	0.0179	0.949	-0.0918	0.933	95	5
rs78047627	DROSHA	-0.0735	0.862	0.0430	0.946	-0.3423	0.553	13	14
rs116270675	DROSHA	-0.8303	0.146	-16.1597	0.016	-0.4322	0.529	6	11
rs77904217	DROSHA	0.1374	0.748	0.1571	0.716	NA	NA	27	0
rs6876706	DROSHA	-0.2735	0.419	-0.2882	0.398	NA	NA	53	0
rs116425093	DROSHA	-0.3769	0.654	-0.3735	0.660	NA	NA	7	0
rs7737174	DROSHA	-0.0958	0.636	0.2474	0.339	-0.6749	0.046	145	87
rs4323243	DROSHA	0.0737	0.835	-0.0118	0.975	0.6270	0.522	36	5
rs74400537	DROSHA	-0.5872	0.148	-0.7186	0.369	-0.6138	0.202	10	23
rs4374769	DROSHA	0.1125	0.746	-0.0360	0.928	0.8319	0.292	33	8
rs7703802	DROSHA	-0.1450	0.491	0.0784	0.762	-0.6975	0.070	168	112
rs17409275	DROSHA	0.3644	0.129	0.1257	0.693	0.7183	0.062	74	110
rs16901243	DROSHA	-14.0275	0.172	NA	NA	-16.4248	0.131	0	1
rs9292432	DROSHA	0.0064	0.987	-0.0528	0.903	0.2499	0.776	28	6
rs55656741	DROSHA	0.2066	0.465	-0.4759	0.271	0.7886	0.045	50	113
rs1382883	DROSHA	0.1822	0.396	-0.1097	0.674	0.7886	0.045	180	116
rs77034974	DROSHA	0.0672	0.826	-0.0193	0.952	15.4297	0.266	68	4
rs17409624	DROSHA	-0.1145	0.592	0.0196	0.940	-0.5244	0.189	164	118
rs56154830	DROSHA	0.3596	0.308	0.6752	0.275	0.0965	0.824	13	28
rs17409803	DROSHA	0.1382	0.616	-0.0540	0.915	0.3038	0.379	25	75
rs2257082	XPO5	-0.2515	0.240	-0.1250	0.658	-0.4236	0.211	90	72
rs115410573	XPO5	-0.4524	0.695	-0.5089	0.660	NA	NA	4	0
rs61739889	XPO5	0.2610	0.650	0.3356	0.562	NA	NA	14	0
rs145380094	XPO5	-0.8784	0.445	-13.0641	0.528	-0.9370	0.440	1	3
rs1096699	XPO5	0.2917	0.280	0.1055	0.821	0.3386	0.322	36	77
rs34324334	XPO5	-0.4794	0.297	-15.4796	0.064	-0.3395	0.508	6	20
rs4714687	XPO5	0.1931	0.794	0.3512	0.648	-14.9243	0.362	10	1
rs61762966	XPO5	-0.2148	0.673	-0.2124	0.676	NA	NA	20	0
rs6578113	AGO2	-0.5302	0.052	-0.5104	0.110	-0.6679	0.217	66	17
rs9969410	AGO2	0.1258	0.732	0.0603	0.872	32.1850	0.022	38	1
rs78254118	AGO2	-0.0143	0.950	-0.0696	0.805	0.2754	0.508	89	35

rs2944776	AGO2	-0.1423	0.579	-0.2157	0.455	0.5603	0.395	84	14
rs2977462	AGO2	-0.0458	0.835	0.0832	0.759	-0.3839	0.329	191	117
rs6994531	AGO2	0.1513	0.670	0.2066	0.604	-0.1962	0.807	33	7
rs3889488	AGO2	0.2579	0.235	0.5196	0.075	-0.1822	0.586	72	79
rs141277504	AGO2	-13.9310	0.184	NA	NA	-14.9569	0.182	0	1
rs12545019	AGO2	-0.5212	0.015	-0.4769	0.087	-0.4780	0.171	103	57
rs2293939	AGO2	0.2673	0.275	0.6944	0.055	-0.2374	0.488	43	72
rs73360466	AGO2	-0.8559	0.113	-1.0883	0.067	0.9912	0.624	19	2
rs62527843	AGO2	-0.2927	0.252	-0.2644	0.372	-0.6115	0.250	75	18
rs115526370	AGO2	0.0130	0.978	-0.0004	0.999	0.4684	0.774	20	2
rs148079280	AGO2	14.0826	0.162	NA	NA	15.0793	0.164	0	1
rs143987786	AGO2	1.4302	0.019	1.4129	0.197	1.2853	0.090	4	10
rs143571422	AGO2	-0.7856	0.477	-14.5625	0.180	0.1487	0.920	3	2
rs2977482	AGO2	-0.1125	0.648	0.1428	0.674	-0.4367	0.241	69	103
rs10112329	AGO2	-0.0341	0.938	-0.1709	0.705	32.1362	0.023	26	1
rs2271738	AGO2	0.4975	0.017	0.4611	0.076	0.6057	0.089	162	99
rs2977485	AGO2	-0.1524	0.489	-0.1979	0.461	-0.0866	0.828	174	40
rs117377585	AGO2	-0.7536	0.219	-13.5711	0.436	-0.8862	0.165	1	12
rs73362311	AGO2	-0.3994	0.389	-0.7067	0.209	0.6046	0.537	18	6
rs2292775	AGO2	0.3269	0.179	0.4844	0.073	-0.6685	0.272	161	15
rs73362319	AGO2	0.4165	0.093	0.2545	0.378	0.9413	0.074	76	19
rs2977490	AGO2	-0.1306	0.555	0.1756	0.509	-0.7938	0.062	174	122
rs75885567	AGO2	0.0740	0.813	-0.1822	0.770	0.1273	0.734	18	42
rs2944755	AGO2	-0.4049	0.142	0.0792	0.870	-0.5790	0.086	36	71
rs7846322	AGO2	-0.2666	0.511	-0.2408	0.599	-0.5845	0.517	25	6
rs73362336	AGO2	0.3283	0.295	0.3928	0.238	-0.2071	0.832	49	5
rs74500398	AGO2	0.9043	0.225	0.8506	0.258	NA	NA	8	0
rs115627791	AGO2	0.3671	0.314	0.2717	0.489	0.9035	0.420	33	4
rs9694342	AGO2	-0.2752	0.177	-0.0983	0.706	-0.4818	0.154	175	88
rs7828287	AGO2	0.7590	0.149	0.1005	0.934	0.8292	0.186	4	13
rs76170696	AGO2	-0.0815	0.853	-0.0825	0.852	NA	NA	27	0
rs7825416	AGO2	-0.0264	0.903	-0.2330	0.401	0.2351	0.523	94	47
rs73362349	AGO2	0.9171	0.109	1.4511	0.093	0.5265	0.487	6	8
rs78796470	AGO2	-0.5342	0.245	-0.5381	0.243	NA	NA	27	0
rs139665188	AGO2	13.7379	0.213	NA	NA	14.5166	0.252	0	1
rs116283890	AGO2	0.2224	0.706	0.1914	0.745	NA	NA	14	0
rs62529972	AGO2	0.0665	0.762	-0.0669	0.819	0.3395	0.326	211	87
rs4961271	AGO2	0.0613	0.765	0.0791	0.760	-0.0376	0.914	122	56
rs7841188	AGO2	0.0911	0.665	0.1028	0.700	0.0094	0.979	108	52
rs73362386	AGO2	-0.1956	0.376	-0.2569	0.327	-0.1995	0.656	127	28
rs74377429	AGO2	0.3951	0.332	0.9284	0.167	0.1047	0.839	12	18
rs2176397	AGO2	-0.1225	0.544	-0.3229	0.212	0.2209	0.512	172	76
rs10087629	AGO2	-0.1163	0.584	0.0658	0.802	-0.5063	0.176	175	111
rs369649929	AGO2	-0.3565	0.517	0.0051	0.993	-26.9475	0.052	14	3
rs11987214	AGO2	-0.2453	0.385	-0.1988	0.497	-2.0278	0.162	78	4
rs117665211	AGO2	-0.3370	0.655	-13.5784	0.434	-0.4910	0.550	1	7
rs73364314	AGO2	0.3814	0.499	0.8339	0.238	-0.4541	0.636	9	5

rs6983600	AGO2	-0.2034	0.531	-0.2606	0.431	32.1908	0.022	57	1
rs75396503	AGO2	0.1934	0.460	0.2020	0.459	-0.0025	0.998	99	4
rs1878478	AGO2	0.2901	0.169	0.0856	0.752	0.6745	0.052	182	57
rs140337057	AGO2	13.7488	0.211	NA	NA	14.5216	0.251	0	1
rs7824622	AGO2	0.2063	0.592	0.1079	0.785	14.5526	0.246	33	1
rs116587291	AGO2	0.1001	0.805	0.0149	0.971	32.1850	0.022	30	1
rs4961278	AGO2	-0.2970	0.147	-0.2505	0.344	-0.5253	0.124	185	78
rs118159219	AGO2	-0.2795	0.819	-12.6362	0.462	0.0426	0.976	2	2
rs7837679	AGO2	-1.6590	0.075	-14.5220	0.093	-0.7500	0.540	5	3
rs6578129	AGO2	0.0480	0.829	-0.1400	0.634	0.2672	0.451	93	99
rs784567	TARBP2	-0.3356	0.241	-0.2567	0.549	-0.5145	0.196	56	115
rs61499963	TARBP2	-0.1208	0.778	-0.2455	0.579	34.6609	0.045	29	1
rs7307055	RAN	0.1288	0.540	0.2235	0.400	0.0378	0.916	167	94
rs111390894	RAN	-0.0796	0.798	-0.0038	0.995	-0.0592	0.871	18	47
rs7298948	RAN	-0.0941	0.829	-0.1666	0.718	-0.0401	0.978	24	2
rs3809142	RAN	-0.1000	0.698	-0.1348	0.690	-0.0844	0.839	51	32
rs14035	RAN	0.0607	0.765	0.0710	0.783	0.1437	0.678	165	83
rs112363519	RAN	0.5951	0.462	0.5880	0.471	NA	NA	7	0
rs117016470	RAN	0.4836	0.411	0.7180	0.593	0.6979	0.309	3	10
rs11061209	RAN	0.1000	0.631	0.0447	0.867	0.3351	0.344	118	92
rs56085972	RAN	0.1993	0.443	-0.0281	0.947	0.4922	0.154	38	69
rs73469107	RAN	-0.1989	0.449	-0.3111	0.413	-0.0098	0.979	43	43
rs6486610	RAN	0.2038	0.513	0.0244	0.969	0.2416	0.510	15	45
rs77319118	RAN	-0.0983	0.757	-0.0442	0.891	-0.7449	1.000	61	1
rs74242887	RAN	-0.0973	0.730	-0.1103	0.808	0.0080	0.983	31	44
rs10848243	RAN	-0.0620	0.766	-0.0408	0.876	-0.3586	0.326	122	101
rs112976942	RAN	0.0017	0.996	0.0794	0.803	NA	NA	62	2
rs7978711	RAN	0.1794	0.384	0.1044	0.691	0.3094	0.368	178	63
rs7960030	RAN	-0.0859	0.671	-0.0243	0.924	-0.0012	0.997	162	71
rs7135237	RAN	0.1973	0.678	0.3084	0.529	-15.5085	0.250	21	1
rs73469185	RAN	-0.4889	0.472	-0.4865	0.477	NA	NA	11	0
rs34365167	RAN	-0.0876	0.746	-0.1022	0.716	0.8465	0.492	90	4
rs56123038	DICER1	0.2125	0.375	0.5072	0.088	-0.2920	0.486	74	32
rs111676151	DICER1	0.3209	0.543	0.3423	0.521	NA	NA	17	0
rs8018330	DICER1	-0.3529	0.338	-0.5536	0.286	-0.2180	0.690	21	16
rs17091658	DICER1	0.7285	0.215	0.6855	0.247	NA	NA	13	0
rs12892332	DICER1	-0.0108	0.958	-0.1264	0.637	0.2246	0.501	103	71
rs114712920	DICER1	0.3010	0.655	0.5600	0.426	-13.5575	0.467	9	1
rs143827989	DICER1	-1.0806	0.032	-1.0455	0.041	-14.2573	0.365	28	1
rs112369172	DICER1	-0.1073	0.744	-0.0886	0.789	-14.4217	0.339	56	1
rs117190345	DICER1	-15.0628	0.049	-13.7319	0.230	-15.1460	0.156	2	2
rs28542080	DICER1	0.1313	0.620	0.2473	0.362	-15.7330	0.066	125	3
rs28461943	DICER1	-0.1286	0.658	-0.1518	0.611	0.2610	0.878	83	2
rs1959870	DICER1	0.0515	0.833	-0.1662	0.654	0.2103	0.529	51	73
rs116282317	DICER1	-0.3952	0.392	-0.3933	0.398	NA	NA	25	0
rs10148962	DICER1	-0.0432	0.846	-0.5005	0.106	0.5567	0.100	75	66
rs2257824	DICER1	0.0471	0.815	0.3468	0.174	-0.3755	0.265	135	73

rs2798821	DICER1	0.0089	0.965	0.2602	0.313	-0.3597	0.278	157	77
rs2798820	DICER1	-0.1925	0.353	0.0062	0.982	-0.4286	0.200	186	77
rs12889327	DICER1	0.0049	0.981	-0.2449	0.348	0.4567	0.189	138	95
rs56331503	DICER1	0.0652	0.797	-0.0833	0.833	0.1917	0.578	43	56
rs116698691	DICER1	0.1874	0.733	0.1873	0.734	NA	NA	16	0
rs61976498	DICER1	0.2711	0.500	0.2611	0.555	0.1677	0.879	25	6
rs10144811	DICER1	-0.0839	0.706	-0.2034	0.453	0.1662	0.682	192	120
rs74618975	DICER1	0.7900	0.146	0.7669	0.360	0.7793	0.279	7	9
rs76009236	DICER1	-12.8350	0.378	-13.7480	0.395	NA	NA	1	0
rs147432772	DICER1	-0.4422	0.454	-0.4511	0.448	NA	NA	15	0
rs111339908	DICER1	0.9130	0.449	0.8362	0.493	NA	NA	3	0
rs143538814	DICER1	-0.0070	0.990	-0.0083	0.988	NA	NA	16	0
rs2789384	DICER1	-0.0079	0.971	0.1296	0.632	-0.1411	0.701	95	47
rs4905266	DICER1	-0.0009	0.998	-0.1193	0.819	0.0885	0.820	22	37
rs8012576	DICER1	0.0148	0.961	-0.1172	0.842	-0.0599	0.868	19	50
rs35578596	DICER1	0.0625	0.810	-0.0984	0.745	0.6891	0.238	70	17
rs73342562	DICER1	0.5505	0.263	0.4839	0.327	NA	NA	19	0
rs8015314	DICER1	-0.2674	0.327	-0.7184	0.089	-0.0550	0.878	247	53
rs7154873	DICER1	0.0144	0.966	-0.7915	0.238	0.2911	0.485	17	31
rs113434762	DICER1	0.6365	0.295	0.5678	0.352	NA	NA	12	0
rs7141572	DICER1	-0.2904	0.640	-0.2211	0.726	NA	NA	13	0
rs8011194	DICER1	0.0754	0.781	0.3724	0.277	-0.5878	0.190	46	26
rs74078723	DICER1	0.0122	0.960	-0.0790	0.773	0.3540	0.510	104	19
rs2096289	DICER1	-0.1182	0.640	0.0465	0.880	-0.5878	0.190	65	26
rs10139956	DICER1	-0.1420	0.517	-0.0133	0.960	-0.3552	0.388	112	33
rs10140127	DICER1	-0.2166	0.325	-0.2485	0.347	-0.1455	0.722	147	123
rs144182662	DICER1	0.2892	0.664	0.2777	0.679	NA	NA	11	0
rs7160403	DICER1	-0.0164	0.945	0.1803	0.559	-0.4082	0.290	204	41
rs11852060	DICER1	-0.1330	0.544	0.1970	0.484	-0.6954	0.079	206	38
rs1954422	DICER1	-0.0003	0.999	0.0145	0.960	-0.0324	0.924	101	75
rs10140275	DICER1	0.1503	0.469	0.1488	0.564	0.2573	0.483	133	48
rs7146324	DICER1	-0.1131	0.620	-0.1071	0.718	-0.0927	0.799	220	107
rs117886030	DICER1	-0.4227	0.374	0.5852	0.470	-0.9082	0.131	8	14
rs72704678	DICER1	0.4138	0.261	0.1205	0.877	0.4550	0.288	9	29
rs79308079	DICER1	0.0115	0.983	0.8351	0.399	-0.2694	0.663	5	12
rs7148434	DICER1	0.0194	0.943	0.0787	0.787	-0.5400	0.463	78	9
rs12147295	DICER1	0.0564	0.783	-0.0705	0.784	0.2593	0.453	139	55
rs12433435	DICER1	-0.7038	0.403	-0.2217	0.860	-1.0425	0.362	3	4
rs2353567	DICER1	0.2125	0.326	0.2211	0.435	0.2464	0.477	205	87
rs7161418	DICER1	-0.1163	0.577	-0.0665	0.806	-0.2395	0.477	98	70
rs58788238	DICER1	0.2825	0.221	0.2555	0.415	0.3275	0.352	60	53
rs80152868	DICER1	0.1397	0.692	0.2966	0.512	-0.1844	0.753	25	14
rs8021127	DICER1	-0.3466	0.096	-0.3723	0.167	-0.3421	0.309	105	69
rs1954419	DICER1	0.4983	0.081	0.5683	0.215	0.4099	0.272	24	43
rs12879834	DICER1	-0.1522	0.511	-0.2203	0.511	-0.0125	0.971	69	61
rs12891219	DICER1	0.0110	0.960	0.0944	0.729	-0.0200	0.959	154	116
rs143564729	DICER1	-0.4586	0.528	-0.2130	0.871	-0.5037	0.565	3	6

rs75081808	DICER1	0.0049	0.989	-0.0508	0.923	-0.0938	0.852	19	19
rs11621639	DICER1	0.3087	0.179	0.2702	0.385	0.4312	0.219	67	57
rs1891866	DICER1	-0.3327	0.122	-0.2884	0.305	-0.4548	0.181	90	73
rs8015157	DICER1	-0.2739	0.254	-0.2422	0.496	-0.4360	0.199	239	78
rs144830643	DICER1	0.9405	0.104	0.3847	0.800	1.0140	0.114	2	12
rs181025469	DICER1	-15.2190	0.040	-14.9919	0.104	-14.5582	0.301	3	1
rs112264755	DICER1	-0.1867	0.608	-0.1990	0.588	NA	NA	44	0
rs12185038	DICER1	0.4214	0.050	0.5641	0.040	0.2013	0.565	87	54
rs79864812	DICER1	0.0712	0.853	0.1522	0.756	-0.0990	0.878	21	12
rs1891863	DICER1	0.0405	0.840	0.1025	0.688	-0.1593	0.635	142	74
rs1891862	DICER1	0.0297	0.900	0.3143	0.327	-0.1788	0.620	92	106
rs80216092	DICER1	1.4718	0.180	15.1902	0.151	0.8036	0.529	1	3
rs59746208	DICER1	0.0923	0.649	0.1867	0.469	-0.1307	0.696	121	75
rs10131361	DICER1	-0.2095	0.400	-0.2308	0.416	-0.1724	0.748	88	17
rs11622488	DICER1	-1.2646	0.086	-14.2878	0.227	-1.4136	0.077	3	8
rs143234445	DICER1	-0.0659	0.931	0.2210	0.781	-14.5582	0.301	7	1
rs7151227	DICER1	-0.1193	0.583	-0.1555	0.589	-0.1009	0.766	84	63
rs76567590	DICER1	-0.1453	0.848	-14.3881	0.281	0.2389	0.775	1	7
rs1187650	DICER1	-0.0570	0.782	-0.1974	0.446	0.1376	0.698	147	57
rs1057035	DICER1	0.0639	0.807	-0.2862	0.501	0.2943	0.395	43	92
rs142808371	DICER1	-13.1244	0.320	-14.1991	0.306	NA	NA	1	0
rs10144436	DICER1	-0.8899	0.012	-0.8848	0.013	-14.5304	0.322	54	1
rs17784006	DICER1	0.1321	0.688	0.0182	0.971	0.1659	0.711	22	26
rs116247322	DICER1	0.2719	0.719	-0.0970	0.909	34.4027	0.012	7	1
rs140875148	DICER1	-13.2733	0.293	-14.4567	0.261	NA	NA	1	0
rs140891457	DICER1	-0.3993	0.447	-0.8605	0.162	1.1279	0.425	17	3
rs1778057	DICER1	-0.2139	0.394	-0.7585	0.044	0.1922	0.597	49	49
rs149242330	DICER1	-12.4575	0.452	-13.5869	0.427	NA	NA	1	0
rs61729795	DICER1	-0.5332	0.636	-0.5849	0.602	NA	NA	4	0
rs4513027	DICER1	0.0017	0.994	0.2077	0.451	-0.7498	0.159	170	136
rs17091828	DICER1	-0.0583	0.850	-0.0719	0.817	NA	NA	66	0
rs114947750	DICER1	0.7431	0.263	0.7119	0.288	NA	NA	10	0
rs115818409	DICER1	-1.1839	0.215	-1.1642	0.226	NA	NA	8	0
rs12897280	DICER1	0.0520	0.816	0.2608	0.347	-0.3777	0.338	88	38
rs67822993	DICER1	-0.2871	0.205	-0.0977	0.709	-0.8319	0.083	147	24
rs77034506	DICER1	0.1178	0.793	0.7897	0.267	-0.1983	0.731	10	14
rs113745694	DICER1	0.0613	0.852	-0.1994	0.729	0.2470	0.551	17	31
rs73331532	DICER1	0.2494	0.284	0.6275	0.028	-0.5279	0.211	76	32
rs142815547	DICER1	-13.8970	0.113	-13.7522	0.221	-13.5799	0.463	2	1
rs78987194	DICER1	-0.1002	0.813	-0.7340	0.364	0.0945	0.855	11	18
rs11160231	DICER1	-0.3639	0.112	-0.6256	0.041	-0.0630	0.861	220	59
rs11160232	DICER1	0.0567	0.821	0.1540	0.580	-0.4076	0.506	174	140
rs1558496	DGCR8	0.0417	0.881	0.0723	0.889	-0.0152	0.965	26	66
rs2073778	DGCR8	0.0056	0.986	-1.0759	0.133	0.3682	0.324	19	42
rs9606241	DGCR8	-0.3015	0.135	-0.4817	0.063	-0.0066	0.984	137	76
rs117148342	DGCR8	-0.0444	0.954	-12.8950	0.417	0.2414	0.778	2	6
rs35569747	DGCR8	-0.1446	0.855	16.0409	0.079	-0.8002	0.362	1	6

rs443678	DGCR8	-0.2090	0.365	-0.2307	0.486	-0.2502	0.459	68	82
rs77209867	DGCR8	-13.9565	0.180	-14.1231	0.334	NA	NA	1	0
rs417309	DGCR8	-0.3503	0.393	1.0569	0.069	-1.6562	0.006	18	16