Gene Networks and microRNAs Implicated in Aggressive Prostate Cancer

Liang Wang,[1] Hui Tang,[2] Venugopal Thayanithy,[3] Subbaya Subramanian,[3] Ann L. Oberg,[2] Julie M. Cunningham,[1] James R. Cerhan,[2] Clifford J. Steer,[4] and Stephen N. Thibodeau[1]

[1]Departments of Laboratory Medicine and Pathology and [2]Health Sciences Research, Mayo Clinic, Rochester, Minnesota; and Departments of [3]Laboratory Medicine and Pathology, [4]Medicine, and Genetics, Cell Biology, and Development, University of Minnesota, Minneapolis, Minnesota

Abstract

Prostate cancer, a complex disease, can be relatively harmless or extremely aggressive. To identify candidate genes involved in causal pathways of aggressive prostate cancer, we implemented a systems biology approach by combining differential expression analysis and coexpression network analysis to evaluate transcriptional profiles using lymphoblastoid cell lines from 62 prostate cancer patients with aggressive phenotype (Gleason grade ≥ 8) and 63 prostate cancer patients with nonaggressive phenotype (Gleason grade ≤ 5). From 13,935 mRNA genes and 273 microRNAs (miRNA) tested, we identified significant differences in 1,100 mRNAs and 7 miRNAs with a false discovery rate (FDR) of <0.01. We also identified a coexpression module demonstrating significant association with the aggressive phenotype of prostate cancer ($P = 3.67 \times 10^{-11}$). The module of interest was characterized by overrepresentation of cell cycle–related genes (FDR = 3.50×10^{-50}). From this module, we further defined 20 hub genes that were highly connected to other genes. Interestingly, 5 of the 7 differentially expressed miRNAs have been implicated in cell cycle regulation and 2 (miR-145 and miR-331-3p) are predicted to target 3 of the 20 hub genes. Ectopic expression of these two miRNAs reduced expression of target hub genes and subsequently resulted in cell growth inhibition and apoptosis. These results suggest that cell cycle is likely to be a molecular pathway causing aggressive phenotype of prostate cancer. Further characterization of cell cycle–related genes (particularly, the hub genes) and miRNAs that regulate these hub genes could facilitate identification of candidate genes responsible for the aggressive phenotype and lead to a better understanding of prostate cancer etiology and progression. [Cancer Res 2009;69(24):9490–7]

Introduction

Prostate cancer remains the most commonly diagnosed non–skin cancer in men in the United States. Approximately one in three men over the age of 50 years shows histologic evidence of prostate cancer. However, only ~10% will be diagnosed with clinically significant prostate cancer, implying that most prostate cancers never progress to become life threatening. Thus far, little is known about what makes some prostate cancers biologically aggressive and more likely to progress to metastatic and potentially lethal disease. Prostate cancer is a complex disease, believed to be caused by variations in a large number of genes and their complex interactions. Conventional approaches used to elucidate genetic risk factors and genetic mechanisms include family-based linkage analysis, pathway-based association study, and genome-wide association study. Among these approaches, genome-wide association study has been very successful with over a dozen single nucleotide polymorphisms identified with elevated risk to prostate cancer (1). However, the observed associations have yet to be translated into a full understanding of the genes or genetic elements mediating disease susceptibility. Furthermore, few prostate cancer risk variants identified from genome-wide association study have any association with clinical characteristics. This is not surprising because these risk single nucleotide polymorphisms are identified by comparing prostate cancer cases with controls. Studies using case-case design are clearly needed to identify associations of genetic variants with aggressive prostate cancer.

Traditionally, microarray-based transcriptional profiling analysis produces massive gene lists (usually based on *P* value) without consideration of potential relationships among these genes. The gene-by-gene approach often lacks a coherent picture of disease-related pathologic interactions. To facilitate candidate gene discovery, there is now an increasing interest in using a systems biology approach. This approach allows for a higher order interpretation of gene expression relationships and identifies modules of coexpressed genes that are functionally related, and eventually characterizes causal pathways and genetic variants. Thus far, studies using the approach have successfully identified disease-related transcriptional networks and genetic variants that contribute to the disease phenotypes (2–7). For example, an early study analyzed the gene expression profiles in large population-based adipose tissue cohorts and found a marked correlation between gene expression in adipose tissue and obesity-related traits. The systems biology approach identified a core network module that was causally associated with obesity (2). This study has recently been validated through characterization of transgenic and knockout mouse models of genes predicted to be causal for obesity phenotype (7).

Expression levels of many genes show abundant natural variation in species from yeast to human (8). Studies have shown significant association of genetic polymorphisms with gene expression in a variety of human cell lines and tissues (9). In addition to genetic factors, however, microRNAs (miRNA) are emerging as key players in the regulation of gene expression. miRNAs are small noncoding RNAs that control the expression of protein-coding transcripts. Each miRNA has multiple target genes that are

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Requests for reprints: Liang Wang, Department of Laboratory Medicine and Pathology, Mayo Clinic College of Medicine, 200 First Street Southwest, Rochester, MN 55905. Phone: 507-284-9136; Fax: 507-266-5193; E-mail: wang.liang@mayo.edu.

doi:10.1158/0008-5472.CAN-09-2183

regulated at the posttranscriptional level. They have been implicated in various diseases, and may influence tumorigenesis by acting as oncogenes and tumor suppressors. For example, the miR-17/92 cluster cooperates with *c-MYC* to accelerate tumor development (10, 11). Germline variations in miRNAs and their target genes have been reported to have a profound effect not only on tumor progression but also an individual's risk of developing cancer (12, 13). Hence, miRNAs are related to diverse cellular processes and regarded as important components of the gene regulatory network.

To identify the genes that contribute to the aggressive phenotype of prostate cancer, we implemented a systems biology approach and analyzed whole genome gene expression profiles in 125 lymphoblastoid cell lines derived from 62 aggressive and 63 nonaggressive prostate cancer patients. We identified a set of mRNA genes and miRNAs whose expression levels were associated with not only cell cycle regulation but also aggressive nature of prostate cancer. We then verified the functional role of two miRNAs using prostate cancer cell lines. These results suggested that the cell cycle–related biological process may be genetically dysregulated in prostate cancer patients and that miRNAs may be significantly involved in development of the aggressive phenotype.

Table 1. Clinical characteristics of prostate cancer patients

	Low-grade prostate cancer $n = 63$	High-grade prostate cancer $n = 62$
Patient characteristics:		
Age, median (range)	65 (44–74)	65 (44–74)
Age, quartiles		
40–58	7 (11.1)	7 (11.3)
59–64	22 (34.9)	22 (35.5)
65–69	22 (34.9)	22 (35.5)
70–84	12 (19)	11 (17.7)
PSA		
<4	10 (15.9)	10 (16.1)
4–9.9	34 (54)	32 (51.6)
10–19.9	12 (19)	9 (14.5)
≥20	7 (11.1)	11 (17.7)
Unknown	0	0
Pathologic characteristics:		
Nodal status		
Negative	62 (98.4)	51 (82.3)
Positive	1 (1.6)	11 (17.7)
Unknown	0	0
Stage		
1 or 2	47 (74.6)	16 (25.8)
3 or 4	15 (23.8)	35 (56.5)
Unknown	1 (1.6)	11 (17.7)
Grade		
4	5 (7.9)	0
5	58 (92.1)	0
8	0	30 (48.4)
9	0	30 (48.4)
10	0	2 (3.2)

Abbreviation: PSA, prostate-specific antigen.

Materials and Methods

Study subjects. The patients were selected based on our ongoing clinic-based case-control study (14, 15). The characteristics of these patients were listed in Table 1. All subjects in the study provided written informed consent. The study was approved by the Mayo Clinic Institutional Review Board.

Cell lines and RNA extraction for profiling analysis. Peripheral blood lymphocytes were collected from 125 Caucasian men with median age of 65 y (range, 44–74 y) and transformed with EBV to establish immortalized cell lines. The transformed cell lines were cultured in RPMI 1640 supplemented with 15% fetal bovine serum (FBS), and 1% penicillin/streptomycin at 37°C in humidified incubators in an atmosphere of 5% CO_2. Experimental series were set up by seeding 5-mL cultures in T25 flasks. Each culture was fed with 5 mL of fresh media twice a week until the cell number reached $\sim 10^6$ in a T75 flask. The cells were harvested and suspended in 500 μL of RNA Stabilization reagent (RNAlater) and stored at −80°C for further processing. Total RNA was extracted from each cell culture using miRNeasy Mini kit (QIAGEN) according to the manufacturer's guidelines. This protocol effectively recovered both mRNA and miRNA. The integrity of these total RNAs was assessed using an Agilent 2100 Bioanalyzer.

Messenger RNA and miRNA microarrays. Illumina human-6 V2 gene expression BeadChip and miRNA expression panel (based on miRbase release 9.0) were used for mRNAs and miRNA profiling analyses, respectively (Illumina, Inc.). RNA aliquot of 200 ng from each cell culture was labeled and hybridized to each array using standard Illumina protocols. BeadChips (mRNA) or sample array matrices (miRNA) were scanned on an Illumina BeadArray reader. For mRNA, 30 triplicate samples, 30 duplicate samples, and 65 singleton samples were run for a total of 215 expression profiles. For miRNA, there were 84 duplicate samples and 6 quadruplicate samples for a total of 192 expression profiles. Based on principal component analysis, we removed 26 individual miRNA profiles due to substantial shifts away from a main cluster. However, replicates from each of the 26 individuals were still included in the analysis as they were in the main cluster. These expression profiles have been deposited in National Center for Biotechnology Information's Gene Expression Omnibus.

Data processing. We processed 215 mRNA profiles from a total of 125 independent patients and 166 miRNA profiles from a total of 90 independent patients. For both mRNA and miRNA data, raw data from BeadStudio (Illumina) were first transformed using a variance stabilization transformation algorithm (16) and then normalized using quantile normalization. We averaged samples with replicates and excluded probes with median detection P value of ≥0.01 (the P values were generated in BeadStudio software). This procedure reduced the number of mRNA probes from 48,702 to 13,935 and miRNA probes from 736 to 366. Among the 366 miRNAs, 273 in miRBase database[5] version 9.1 were included in the study. The remaining 93 that were putative miRNAs identified in a RAKE analysis were excluded from further analysis.

Data analysis. The pathologic grades (Gleason Score) of ≤5 and ≥8 were used to dichotomize samples into low-grade (nonaggressive) and high-grade (aggressive) groups. We applied a two sample t test with multiple testing correction to identify genes and miRNAs that were significantly differentially expressed between the two Gleason grade groups. We defined q value of FDR of <0.01 to be statistically significant. Pearson correlation coefficients were also calculated to compare results from the following network analysis.

To explore the phenotype-related genes and their interactions, we applied a systems biology approach using a weighted gene coexpression network analysis (WGCNA; refs. 17–20). Unlike other gene coexpression networks using a binary variable to encode gene coexpression (connected, 1; unconnected, 0), the WGCNA converts coexpression measures into connection weights or topology overlap measures (TOM). Because the program was computationally intensive when running on large numbers of genes, we simplified the computation by selecting a subset of genes for analysis.

[5] http://microrna.sanger.ac.uk

We selected the genes in two steps: first, we selected the genes that showed significant correlation with prostate cancer grade (FDR < 0.01); from the rest of genes, we then selected the top 2,000 most variable genes based on coefficient of variance. We inputted expression profiles of these selected genes to construct weighted gene coexpression modules using the WGCNA R package (18, 19, 21). We defined modules using static method by hierarchically clustering the genes using 1-TOM as the distance measure with a height cutoff of 0.95 and a minimum size (gene number) cutoff of 40 for the resulting dendrogram.

To identify which module is correlated with clinical phenotype, we first calculated module eigengene (i.e., first principal component of the expression values across subjects) using all genes in each module. We then correlated the module eigengenes to prostate cancer grade using the Pearson correlation. We determined intramodular connectivity for each gene by summing the connectivities of that gene with each other gene in that module. We used program VisANT (Integrative Visual Analysis Tool for Biological Networks and Pathways; ref. 22) to construct gene-gene interaction (connections) networks.

Gene ontology analysis. To explore whether genes in each target group share a common biological function, we searched for overrepresentation in gene ontology (GO) categories. We used 13,935 mRNA accession numbers as reference gene list. We inputted each group of genes into The Database for Annotation, Visualization and Integrated Discovery (DAVID) for GO term enrichment analysis. The DAVID is a program that checks for an enrichment of genes with specific GO, KEGG, and SwissProt terms (23).

Nucleofection of miRNA mimics in VCaP and LNCaP cells. We cultured LNCaP cells (24) in RPMI 1640 and VCap (25) cells in DMEM, respectively. Both cell lines were grown in the media containing 10% FBS, 1% penicillin, and streptomycin at 37°C with 5% CO_2. Cells were nucleofected with double stranded synthetic miRNA mimics (syn-hsa-miR-145 miScript miRNA and syn-hsa-miR-331 miScript miRNA) and scrambled controls (Qiagen) using the program T-09 (Lonza). Nucleofection efficiency was monitored by nucleofecting the cells with 2.0 μg of pmaxGFP plasmid DNA in six-well plates. Cells were visualized and tested at 48 h after nucleofection.

Cell viability assay and fluorescence-activated cell sorting. After nucleofection, cells were placed on 24-well plates. Media were changed twice after 10 h of plating and then once every 24 h. Cell viability of treated cells was examined using LIVE/DEAD Viability/Cytotoxicity kit (Invitrogen) after 48 h of treatment and visualized using a fluorescent microscope (×100) after 15 min of staining. Fluorescence-activated cell sorting (FACS) analysis was performed using a FACSCalibur Flow Cytometer (Becton Dikinson) following the method of Riccardi and Nicoletti (26).

Quantitative reverse transcription-PCR. Expression level of target genes were quantified at 48 h after treatment by quantitative reverse transcription-PCR (qRT-PCR) using the Lightcycler 480 SYBR Green I master mix (Roche) in an ABI 7500 real-time PCR system. Primer sequences were listed in Supplementary Table S1. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) expression level was used as normalization control. Relative expression values were calculated following the $2^{-\Delta\Delta}$ Ct method of Schmittgen and Livak (27) using values from three independent experiments.

Results

Correlation between transcripts and pathologic grades. To identify transcripts whose expression traits were associated with aggressive phenotype of prostate cancer, we applied a two sample *t* test using 13,935 detectable gene expression profiles in 62 high-grade and 63 low-grade prostate cancer cases. Among all genes tested, we found significant association in 1,100 genes (FDR < 0.01). For the 125 prostate cancer cases, 90 (45 high-grade and 45 low-grade cases) were also available for miRNA profiling analysis. The two-sample *t* test using 273 detectable miRNA expression profiles identified significant association with prostate cancer grade in 7 miRNAs (FDR < 0.01; Supplementary Table S2). The seven miRNAs included miR-222, miR-221, miR-331-3p, miR-16, miR-145, miR-9*, and miR-551a. Because miR-9 and miR-9* are processed from the same precursor, we also observed an association of miR-9 with the prostate cancer grade (FDR = 0.013). However, we did not find any association of miR-15a with prostate cancer grade ($P = 0.65$), although miR-15a and miR-16 are located in the same miRNA cluster.

To functionally classify these 1,100 significant genes, we used the online biological classification tool DAVID (23) and observed significant enrichment of these genes in multiple GO categories. The most significant enrichment was the GO category of cell cycle biological process with a FDR of 3.40×10^{-23}. The other significant GO categories included DNA replication (FDR = 1.60×10^{-13}) and chromosome (FDR = 2.10×10^{-13}). In fact, all significant GO category clusters were related to cell cycle biological function (Supplementary Table S3).

In an effort to provide additional evidence to support our initial observation, we downloaded gene expression profiles from another study with benign prostate tissues (28). After obtaining the relevant clinical information, we reanalyzed the Affymetrix U95av2-based expression profiles derived from five benign prostate tissues in patients with aggressive phenotype (Gleason Score, ≥8) and four benign prostate tissues in patients with nonaggressive phenotype (Gleason Score, ≤5). Statistical analysis using *t* test revealed significant difference in 1,847 RNA probes ($P < 0.05$). Interestingly, GO analysis of these differential genes showed that cell cycle regulation was the most significantly enriched GO category with $P = 2.97 \times 10^{-5}$ (FDR = 0.056; Supplementary Table S3). We further analyzed these differentially expressed genes and found significant overlap between the benign tissues and the cell lines ($P < 0.01$).

Gene coexpression networks and biological pathways. Because coexpressed genes are biologically related, grouping these highly connected genes by network analysis may shed light on underlying functional processes in a manner complementary to standard differential expression analyses. To ensure that phenotype-related genes were used to construct the network, we included the 1,100 most significant genes with FDR of <0.01 along with the top 2,000 most variable genes (selected from remaining 12,835 genes) determined by their coefficient of variance. The WGCNA analysis identified four modules of genes with high topological overlap (Fig. 1). The modules were defined as a cluster of highly connected genes (nodes). Each major branch in the figure represented a color-coded module containing a group of highly correlated genes. The modules turquoise, brown, blue, and yellow included 265, 106, 229, and 65 genes, respectively.

To examine if these modules were associated with aggressive prostate cancer, we correlated the module eigengene to the Gleason grade and found significant correlation of the prostate cancer grade only with the turquoise module ($P = 3.67 \times 10^{-11}$). The other three modules did not show any correlation (all $P > 0.05$). To biologically characterize those modules, we applied the DAVID tool (23) to classify these genes in each module and observed various level of GO category enrichment in all four modules (Table 2). Specifically, the prostate cancer grade–related turquoise module showed significant enrichment in the biological process of cell cycle (FDR = 3.50×10^{-50}). The blue module showed overrepresentation in protein acetylation (FDR = 8.21×10^{-7}). The brown and yellow modules showed a strong trend but not statistical significance (FDR > 0.01) for GO category enrichment.

Clinical trait–related hub genes. The importance of a gene is often dependent on how well it associates with other genes

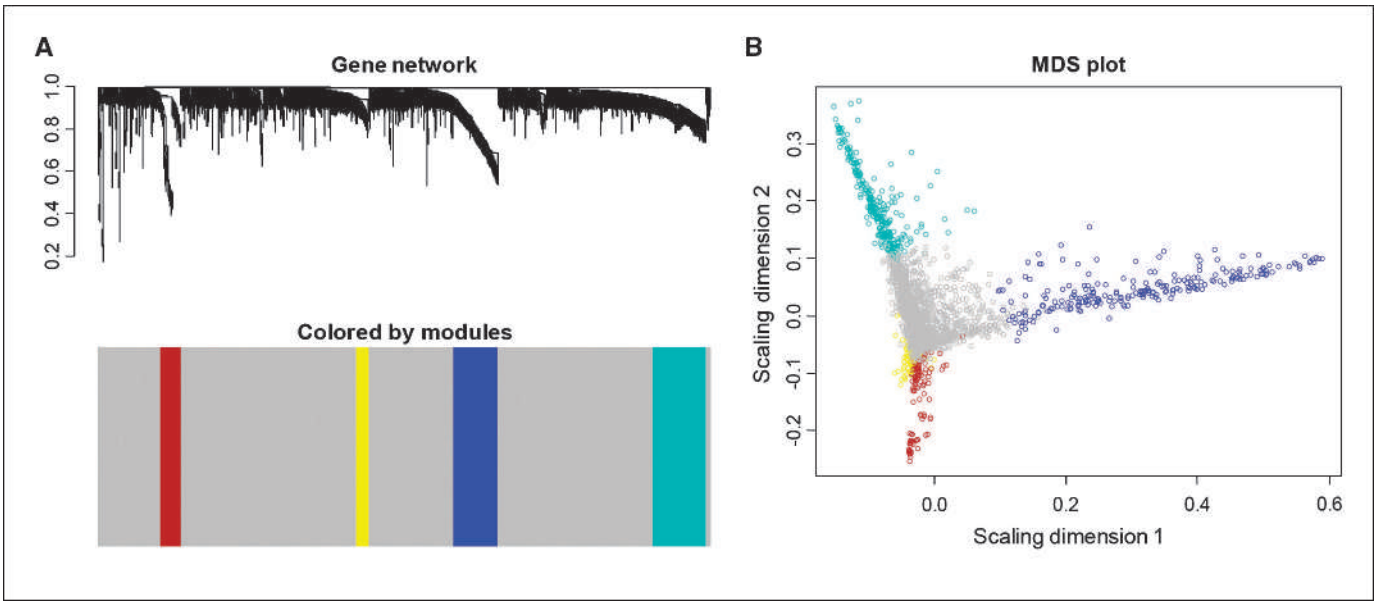

Figure 1. Gene coexpression network analysis. *A,* branches (gene modules) of highly correlated genes by average linkage hierarchical clustering of 3,100 genes. The colored bars directly correspond to the module (color) designation for the clusters of genes. *Gray,* genes that are not part of any module. The remaining colors are used for the four modules. *B,* multidimensional scaling plot of the entire gene expression network. Each dot represents a gene, where the color corresponds to the gene module. The distance between each dot indicates their topological overlap.

in a network. Studies suggest that more centralized genes in the network are more likely to be key drivers to proper cellular function than peripheral genes (nodes; ref. 18). These centralized genes are called hub genes, implying that they are highly connected genes. Intramodular hub genes are defined based on their high correlation with the module eigengene, i.e., as a good representative of a module. We focused our analysis on genes in the turquoise module because of its relevance to clinical trait (Table 2). We used the WGCNA algorithm to calculate intramodular connectivity (connection strength of a given gene with other genes in a particular module). To visualize the relationship between gene significance and intramodular connectiviy, we plotted scaled connectivity on *X*-axis and gene significance (absolute correlation coefficient r value between gene expression and prostate cancer grade) on *Y* axis. We observed significant positive correlation ($r = 0.61$, $P = 7.1 \times 10^{-19}$; Fig. 2*A*). The genes with higher connectivity tended to have stronger correlation with prostate cancer grade, suggesting a potentially important role of highly connected genes (hub genes) in the aggressive phenotype of prostate cancer.

To further visualize gene-gene interactions, we exported the WGCNA-generated connectivity information to the VisANT (22) and observed various degrees of gene-gene connections (interactions). We raised the weighted cutoff value to ≥0.16 to identify hub genes with the strongest connections with other genes. The raised cutoff reduced the total number of connections per gene. Under this criterion, we observed 84 genes, each with at least 1 connection, and 20 genes, each with at least 10 connections (Fig. 2*B*). We defined the 20 highly connected genes as hub genes. The genes *CDC2* and *DTL* were the strongest, each with 55 connections, whereas *CCNA2* had 50. More importantly, all 20 hub genes not only showed significant correlation with pathologic grade but also have been implicated in cell cycle–related functions (Table 3).

Hub genes as miRNA targets. Because each miRNA may regulate multiple mRNA genes, we asked if the expression traits in hub genes were the result of regulatory effects from miRNAs. To explore this, we downloaded all miRNA target genes predicted by TargetScan (29–31). We focused our search on the 20 hub genes and the 7 differential miRNAs. We found that 3 of the 20 hub genes were the predicted targets for 2 differentially expressed miRNAs. The three hub genes *CCNA2, CDCA5*, and *KIF23* were significantly upregulated in aggressive prostate cancer (Table 3). The *miR-145*, significantly downregulated in aggressive prostate cancer, was predicted to bind to 3′ untranslated region of the *CCNA2*. The

Table 2. Module significance in aggressive prostate cancer and GO analysis

Modules	Total gene count	Correlation with pathologic grade		DAVID GO analysis			
		r	*P*	Term	Gene count	*P*	FDR
Turquoise	265	0.548	3.67E-11	Cell cycle	97	1.86E-53	3.50E-50
Brown	106	−0.08	0.377	Nuclear pore complex interacting	3	2.46E-04	0.464
Blue	229	0.058	0.521	Acetylation	30	5.27E-10	8.21E-07
Yellow	65	0.106	0.241	Membrane	27	2.78E-04	0.422

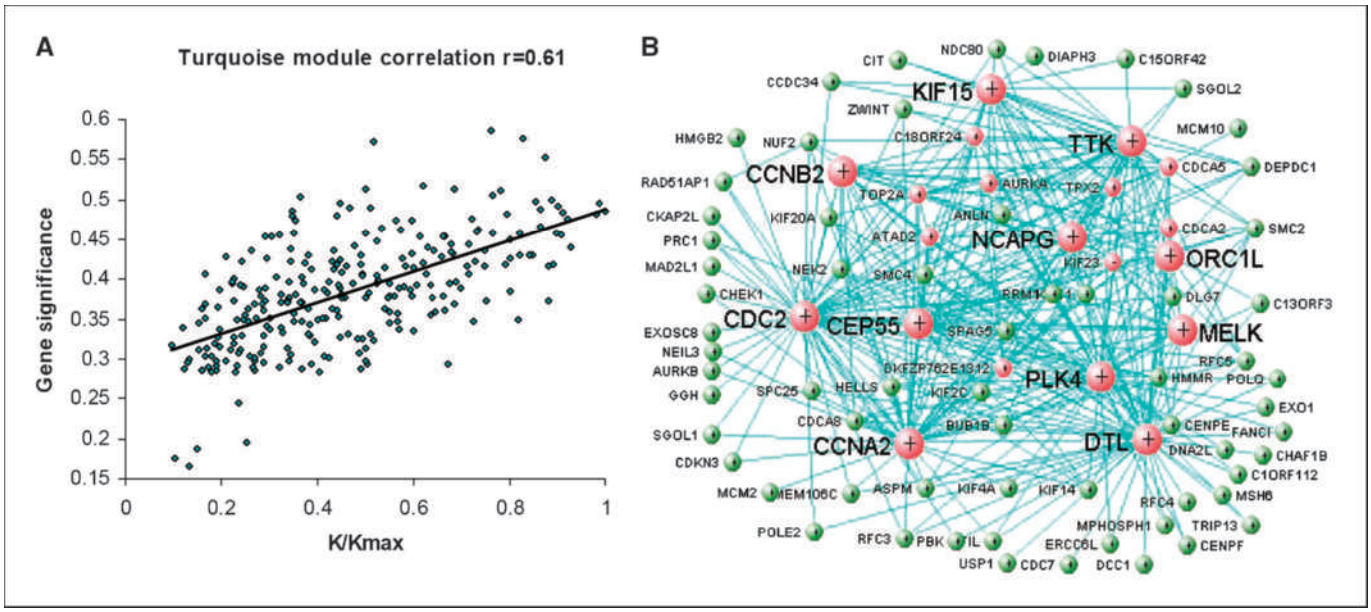

Figure 2. Identification of clinical trait–related hub genes. *A,* scatterplot between gene significance (absolute r; *Y*-axis) and scaled intramodular connectivity (*K/Kmax*). Each point corresponds to a gene in the turquoise module. The intramodular connectivity was significantly correlated with gene significance (r = 0.61, $P = 7.1 \times 10^{-19}$). *B,* visualization of gene-gene interaction within the turquoise module. The connections were drawn using VisANT tool (22). The genes with at least one connection when weighted cutoff value of ≥0.16 are shown. Each node represents a gene. *Red nodes,* hub genes. Bigger nodes indicate more connections.

miR-331-3p, also significantly downregulated in aggressive prostate cancer, was predicted to target the genes *CDCA5* and *KIF23*. More interestingly, we observed significant correlation in expression level for each of these miRNA-gene pairs. The *miR-145:CCNA2* pair showed inverse correlation with $P = 1.48 \times 10^{-4}$. The *miR-331-3p:CDCA5* and *miR-331-3p:KIF23* pairs showed inverse correlation with $P = 2.25 \times 10^{-4}$ and $P = 0.029$, respectively.

Functional evaluation of *miR-145* and *miR-331-3p in vitro*. To evaluate the potential regulatory roles of *miR-145* and *miR-331-3p*, we ectopically expressed these miRNAs in prostate cancer

Table 3. Connectivity and gene significance of 20 selected hub genes

Symbol	Gene name	Accession number	No. of connections	Gene significance*			
				r	*P*	FDR	Rank†
CDC2	Cell division cycle 2	NM_001786	55	0.495	4.49E-09	8.92E-07	43
DTL	Denticleless homologue	NM_016448	55	0.485	1.00E-08	1.68E-06	51
CCNA2	Cyclin A2	NM_001237	50	0.482	1.30E-08	2.03E-06	54
PLK4	Polo-like kinase 4	NM_014264	48	0.441	2.67E-07	1.78E-05	128
TTK	TTK protein kinase	NM_003318	40	0.475	2.26E-08	2.91E-06	66
CEP55	Centrosomal protein 55 kDa	NM_018131	35	0.419	1.16E-06	5.32E-05	186
KIF15	Kinesin family member 15	NM_020242	26	0.484	1.05E-08	1.71E-06	52
CCNB2	Cyclin B2	NM_004701	20	0.416	1.40E-06	6.11E-05	196
ORC1L	Origin recognition complex, subunit 1-like	NM_004153	19	0.491	6.27E-09	1.16E-06	46
MELK	Maternal embryonic leucine zipper kinase	NM_014791	17	0.500	2.82E-09	6.17E-07	39
NCAPG	Non-SMC condensin I complex, subunit G	NM_022346	17	0.373	1.84E-05	4.36E-04	360
HJURP	Holliday junction recognition protein	NM_018410	14	0.487	8.30E-09	1.45E-06	49
Ska1	Spindle and KT associated 1	NM_145060	13	0.370	2.12E-05	4.81E-04	376
TPX2	TPX2, microtubule-associated, homologue	NM_012112	13	0.335	1.36E-04	1.89E-03	613
TOP2A	Topoisomerase (DNA) II α 170 kDa	NM_001067	12	0.497	3.74E-09	7.78E-07	41
CDCA5	Cell division cycle associated 5	NM_080668	12	0.431	5.31E-07	2.95E-05	154
KIF23	Kinesin family member 23	NM_004856	12	0.379	1.29E-05	3.28E-04	334
CDCA2	Cell division cycle associated 2	NM_152562	11	0.392	6.11E-06	1.90E-04	274
ATAD2	ATPase family, AAA domain containing 2	NM_014109	11	0.373	1.86E-05	4.40E-04	361
AURKA	Aurora kinase A	NM_198436	10	0.379	1.34E-05	3.37E-04	340

*Represents statistical significance of Pearson correlations between a specific gene expression and pathologic grade of prostate cancer.
†Rank is based on FDR value among 13,935 genes with most significant gene as 1.

cell lines LNCaP (24) and VCaP (25). We found that ectopic expression of the *miR-145* reduced the *CCNA2* level by 54% in VCaP cells and 45% in LNCaP cells. Ectopic expression of the *miR-331* reduced the *CDCA5* level by 44% in VCaP and 48% in LNCaP cells, and the *KIF23* level by 43% in VCaP and 44% in LNCaP cells (Fig. 3*A*). To investigate the functional consequences of ectopic expression of these miRNAs, we examined cell viability using a flow cytometer. Gene transfer efficiency was monitored in green fluorescent protein transfected control groups and ~80% of transfection was observed in both prostate cancer cell lines. We found significant cell growth arrest and apoptosis by the expression of these miRNAs. Specifically, the *miR-145* and *miR-331* ectopic expression induced 37% and 39% apoptosis in the VCaP cells, and 32% and 33% apoptosis in LNCaP cells, respectively. In contrast, scrambled cells did not show any significant apoptosis (Fig. 3*B* and *C*).

Discussion

Clinical phenotypes of prostate cancer vary from an indolent disease requiring no treatment to one in which tumors metastasize

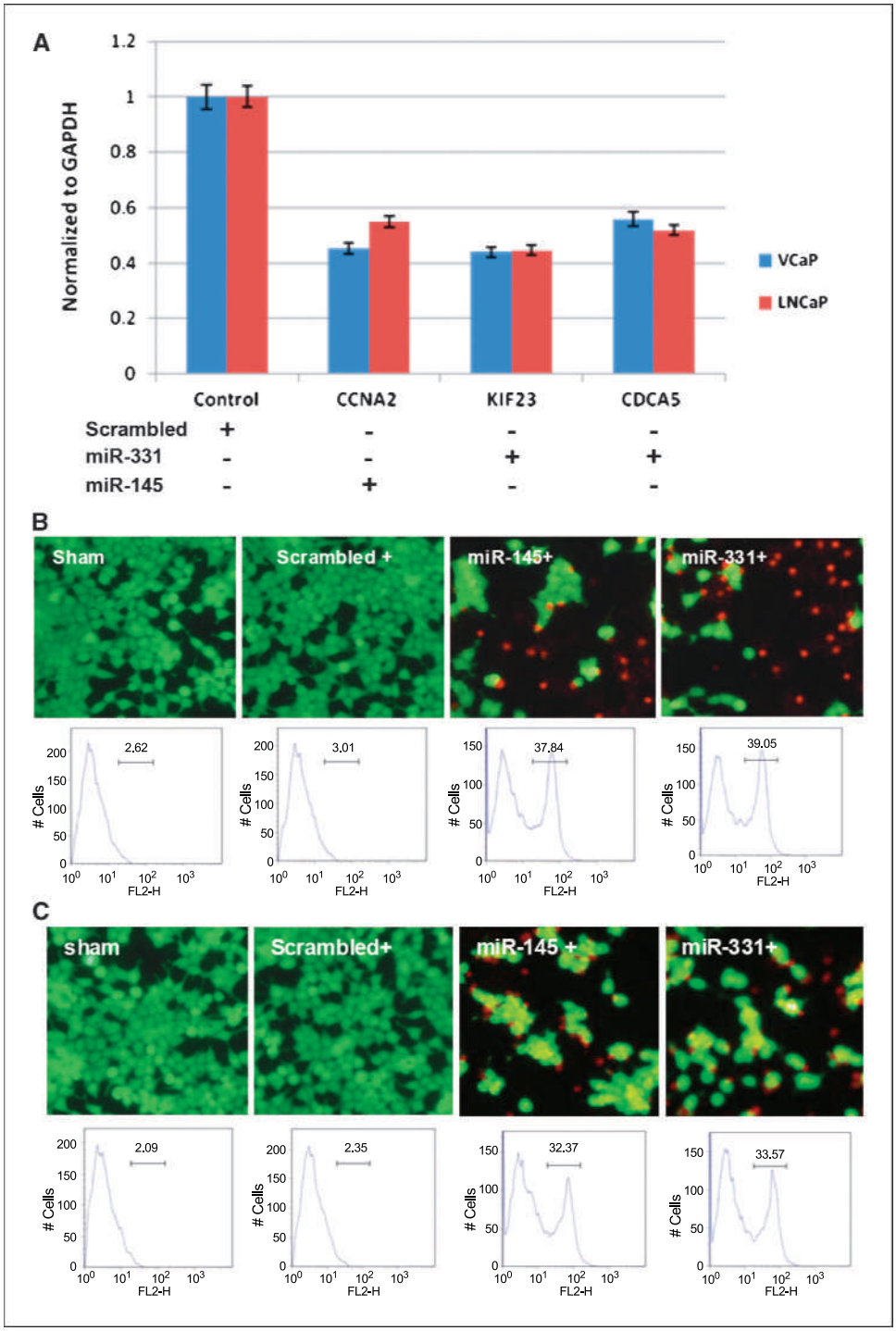

Figure 3. Biological effect of ectopic expression of miR-145 and miR-331 in prostate cancer cell lines. *A*, qRT-PCR was used to measure expression level of target genes using total RNA from nucleofected cells. Expression values were normalized to GAPDH. Expression levels of target genes were significantly reduced by ectopic expression of the two miRNAs. Cell viability was examined in VCaP cells (*B*) and LNCaP cells (*C*). Live and dead cells were stained in green and red, respectively. Percentage of apoptotic cell population measured by FACS was shown below each corresponding cell image.

and escape local therapy even when with early detection. Identification of candidate genes for aggressive prostate cancer has been a difficult task. In this study, we applied a systems biology approach to study the aggressive phenotype of prostate cancer. This approach used gene expression profiles and organized genes into modules based on coexpression. By examining expression profiles in 125 lymphoblastoid cell lines derived from prostate cancer patients, we observed four coexpression modules. Importantly, one of four modules not only enriched genes known to play critical roles in cell cycle regulation but also showed significant correlation with aggressive phenotype of prostate cancer. These results, along with results from benign prostate tissues (Supplementary Table S2), strongly suggested that germline variations of cell cycle–related genes may be a major cause to aggressive prostate cancer.

Hub genes are believed to play major roles in a highly interacted network. In this study, we have defined 20 highly connected hub genes in an aggressive prostate cancer–associated module. Further data mining revealed significant involvement of these hub genes in the cell cycle regulation and the development of various tumors. For example, the gene *CDC2* (connected to 55 other genes) is essential for G_1-S and G_2-M phase transitions of eukaryotic cell cycle. Aberrant activation of the *CDC2* may contribute to tumorigenesis by promoting cell proliferation and survival (32). The gene *DTL* (55 connections) plays important roles in DNA synthesis, cell cycle progression, cytokinesis, proliferation, and differentiation (33). The *DTL* may regulate p53 polyubiquitination (34) and *CDT1* proteolysis in response to DNA damage (35) and may also be essential for early G_2-M checkpoint (36). Suppression of the *DTL* causes accumulation of G_2-M cells, resulting in growth inhibition of cancer cells (37). The gene *CCNA2* (50 connections) belongs to the highly conserved cyclin family. The gene is expressed in all tissues and binds/activates *CDC2* kinases, and thus promotes both cell cycle G_1-S and G_2-M transitions. Overexpression of the gene was associated with high grade (38) and poor prognosis (39) in breast cancer. These data strongly suggest that dysregulation of these cell cycle–related hub genes may be crucial for the development of aggressive phenotype of prostate cancer.

It is worthwhile to mention that none of the 20 hub genes were among the top gene list identified by differential gene expression analysis (Supplementary Table S2). The hub genes with the greatest and least statistical significance are *MELK* (FDR = 6.17×10^{-7}) and *TPX2* (FDR = 1.89×10^{-3}), respectively. The *MELK* is ranked 39th and the *TPX2* is ranked 613th in differential analysis (Table 3). Depending on the purpose of a study, a top gene list approach (based on differential expression *P* value) will be more suitable for biomarker discovery because this type of study is directed at finding disease markers. However, for an understanding of etiology, simply selecting top differential genes identified by two sample *t* test (or similar methods) may miss important genes. Therefore, a systems biology–based network analysis may provide an important alternative and more meaningful tool for candidate gene discovery.

miRNA has been emerged as a crucial regulator of gene expression. In this study, we identified seven differentially expressed miRNAs, five of which have been implicated in regulation of cell cycle. For example, the top two miRNAs (*miR-222/221*) directly targeted cell growth–suppressive cyclin-dependent kinase inhibitors *p27* and *p57* mRNAs, and reduce their protein levels (40, 41). Ectopic expression of the *miR-222/221* also resulted in activation of *CDK2* and facilitation of G_1-S phase transition (42), which agreed with our present study: significant increases of the *miR-222/221* (FDR $\leq 4.73 \times 10^{-6}$) as well as the *CDK2* (FDR = 7.79×10^{-4}) in aggressive prostate cancer. The target gene *p27* (*CDKN1B*), however, only showed slightly decreased expression (mean, 8.78 in high grade and 8.79 in low grade on log2 scale, $P = 0.79$). The lack of significant decrease in the *p27* may be explained by the fact that the miRNAs regulate the target gene at the posttranscriptional level. Another target gene *p57* (*CDKN1C*) was undetectable in our lymphoblastoid cell lines and therefore was not included in the analysis.

Important role of the *miR-222/221* in aggressive prostate cancer was recently confirmed by *in vivo* and *in vitro* studies. For example, *in vivo* overexpression of *miR-221* was able to confer a high growth advantage to LNCaP-derived tumors in severe combined immunodeficient mice, whereas *anti–miR-221/222* treatment in the highly aggressive PC3 cell line reduced tumor growth (43). Furthermore, upregulation of these two miRNAs in prostate cancer–derived primary cell lines showed significant inverse correlation with the *p27* expression. Additionally, both *in vitro* and *in vivo* results implicated that *p21* and *p27* had compensatory roles in advanced prostate cancer cells, and downregulation of both these molecules essentially enhanced the aggressive phenotype (44). These results suggest that the *miR-221/222* may contribute to the oncogenesis and progression of prostate cancer through *p27*(*Kip1*) downregulation.

The other three miRNAs that affect cell cycle regulation include *miR-16, miR-145*, and *miR-331*. The *miR-16* can trigger an accumulation of cells in G_0-G_1 by silencing multiple cell cycle genes simultaneously (45, 46) and negatively regulate two other targets *HMGA1* and *CAPRIN1* involved in cell proliferation (47). In our data set, we observed upregulation of the *miR-16* and downregulation of the target genes *HMGA1* and *CAPRIN1*. Particularly, expression difference of the *HMGA1* was statistically significant (mean, 7.83 in high grade and 7.90 in low grade; FDR = 0.007). The *miR-145* showed inhibition of tumor cell growth by direct silencing *c-Myc* (48). The *MYC* is an oncogenic, nuclear phosphoprotein that plays a key role in cell cycle progression, apoptosis, and cellular transformation. Downregulation of the miR-145 in aggressive prostate cancer was consistent with upregulation of the *MYC* in the same sample set (mean, 11.39 in high grade and 11.30 in low grade; $P = 0.04$; FDR = 0.10). Consequently, we observed significant upregulation of Myc-regulated miRNAs (11) including miR-363 (FDR = 0.016), miR-92a (FDR = 0.022), miR20b (FDR = 0.028), and miR-18b (FDR = 0.030). Additionally, our previous study showed that *miR-331* was significantly associated with cell cycle–related genes (49). By ectopic expression of the *miR-145* and *miR-331-3p*, the current study showed significant reduction of corresponding target genes, inhibition of cell growth and accumulation of apoptotic cells (Fig. 3). These findings suggest that differential expression of these miRNAs at germline level may dysregulate target hub genes, which could lead to an abnormal cell division and proliferation, and eventually developing an aggressive phenotype of prostate cancer.

Overall, this study used a systems biology approach to identify genes that are potentially involved in the aggressive phenotype of prostate cancer. This approach moves beyond single gene investigation to provide a systems level perspective on the potential relationships between members of a network. Our results strongly suggest that dysregulation of cell cycle may significantly contribute to the deadly form of prostate cancer. These findings are important not only because we have discovered a candidate pathway and related hub genes but also because we have identified candidate miRNAs and their predicted target genes. Further studies are needed to determine genetic causes of expression alterations in both differentially expressed miRNAs and mRNA genes. Additional

functional studies will determine whether variations in the selected hub genes and miRNAs are attributable to the aggressive nature of prostate cancer. These studies will facilitate candidate gene discovery and lead to better understanding of the aggressive phenotype of prostate cancer, a more clinically relevant form of the disease.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

Received 6/14/09; revised 10/6/09; accepted 10/9/09; published OnlineFirst 12/8/09.

Grant support: NIH grant CA126786 (L. Wang), Mayo Clinic Specialized Programs of Research Excellence in prostate cancer, NIH CA91956 (L. Wang), and Mayo Clinic Comprehensive Cancer Center Shared Resource grant (P30 CA15083). V. Thayanithy is supported by grant funding from the Academic Health Center, University of Minnesota. Grant support from the Laboratory Medicine and Pathology, University of Minnesota to S. Subramanian is duly acknowledged.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank Drs. Lihua Li and Aaron Sarver for their comments and suggestions. We also thank Mayo Clinic Microarray Shared Resources for microarray service.

References

1. Witte JS. Prostate cancer genomics: towards a new understanding. Nat Rev Genet 2009;10:77–82.
2. Chen Y, Zhu J, Lum PY, et al. Variations in DNA elucidate molecular networks that cause disease. Nature 2008;452:429–35.
3. Fuller TF, Ghazalpour A, Aten JE, Drake TA, Lusis AJ, Horvath S. Weighted gene coexpression network analysis strategies applied to mouse weight. Mamm Genome 2007;18:463–72.
4. Ghazalpour A, Doss S, Zhang B, et al. Integrating genetic and network analysis to characterize genes related to mouse weight. PLoS Genet 2006;2:e130.
5. Miller JA, Oldham MC, Geschwind DH. A systems level analysis of transcriptional changes in Alzheimer's disease and normal aging. J Neurosci 2008;28:1410–20.
6. Ray M, Ruan J, Zhang W. Variations in the transcriptome of Alzheimer's disease reveal molecular networks involved in cardiovascular diseases. Genome Biol 2008; 9:R148.
7. Yang X, Deignan JL, Qi H, et al. Validation of candidate causal genes for obesity that affect shared metabolic pathways and networks. Nat Genet 2009;41:415–23.
8. Emilsson V, Thorleifsson G, Zhang B, et al. Genetics of gene expression and its effect on disease. Nature 2008; 452:423–8.
9. Cookson W, Liang L, Abecasis G, Moffatt M, Lathrop M. Mapping complex disease traits with global gene expression. Nat Rev Genet 2009;10:184–94.
10. He L, Thomson JM, Hemann MT, et al. A microRNA polycistron as a potential human oncogene. Nature 2005;435:828–33.
11. O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT. c-Myc-regulated microRNAs modulate E2F1 expression. Nature 2005;435:839–43.
12. Calin GA, Ferracin M, Cimmino A, et al. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. N Engl J Med 2005;353: 1793–801.
13. Jazdzewski K, Murray EL, Franssila K, Jarzab B, Schoenberg DR, de la Chapelle A. Common SNP in pre-miR-146a decreases mature miR expression and predisposes to papillary thyroid carcinoma. Proc Natl Acad Sci U S A 2008;105:7269–74.
14. Wang L, McDonnell SK, Hebbring SJ, et al. Polymorphisms in mitochondrial genes and prostate cancer risk. Cancer Epidemiol Biomarkers Prev 2008;17:3558–66.
15. Wang L, McDonnell SK, Slusser JP, et al. Two common chromosome 8q24 variants are associated with increased risk for prostate cancer. Cancer Res 2007;67: 2944–50.
16. Lin SM, Du P, Huber W, Kibbe WA. Model-based variance-stabilizing transformation for Illumina microarray data. Nucleic Acids Res 2008;36:e11.
17. Horvath S, Dong J. Geometric interpretation of gene coexpression network analysis. PLoS Comput Biol 2008; 4:e1000117.
18. Horvath S, Zhang B, Carlson M, et al. Analysis of oncogenic signaling networks in glioblastoma identifies ASPM as a molecular target. Proc Natl Acad Sci U S A 2006;103:17402–7.
19. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol 2005;4:Article17.
20. Oldham MC, Konopka G, Iwamoto K, et al. Functional organization of the transcriptome in human brain. Nat Neurosci 2008;11:1271–82.
21. Langfelder P, Zhang B, Horvath S. Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. Bioinformatics 2008;24:719–20.
22. Hu Z, Snitkin ES, DeLisi C. VisANT: an integrative framework for networks in systems biology. Brief Bioinform 2008;9:317–25.
23. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4:44–57.
24. Li H, Lovci MT, Kwon YS, Rosenfeld MG, Fu XD, Yeo GW. Determination of tag density required for digital transcriptome analysis: application to an androgen-sensitive prostate cancer model. Proc Natl Acad Sci U S A 2008;105:20179–84.
25. Korenchuk S, Lehr JE, MClean L, et al. VCaP, a cell-based model system of human prostate cancer. *In vivo* 2001;15:163–8.
26. Riccardi C, Nicoletti I. Analysis of apoptosis by propidium iodide staining and flow cytometry. Nat Protoc 2006;1:1458–61.
27. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc 2008;3:1101–8.
28. Singh D, Febbo PG, Ross K, et al. Gene expression correlates of clinical prostate cancer behavior. Cancer Cell 2002;1:203–9.
29. Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 2009;19:92–105.
30. Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol Cell 2007;27:91–105.
31. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 2005;120:15–20.
32. Liu P, Kao TP, Huang H. CDK1 promotes cell proliferation and survival via phosphorylation and inhibition of FOXO1 transcription factor. Oncogene 2008;27:4733–44.
33. Pan HW, Chou HY, Liu SH, Peng SY, Liu CL, Hsu HC. Role of L2DTL, cell cycle-regulated nuclear and centrosome protein, in aggressive hepatocellular carcinoma. Cell Cycle 2006;5:2676–87.
34. Banks D, Wu M, Higa LA, et al. L2DTL/CDT2 and PCNA interact with p53 and regulate p53 polyubiquitination and protein stability through MDM2 and CUL4A/DDB1 complexes. Cell Cycle 2006;5:1719–29.
35. Higa LA, Banks D, Wu M, Kobayashi R, Sun H, Zhang H. L2DTL/CDT2 interacts with the CUL4/DDB1 complex and PCNA and regulates CDT1 proteolysis in response to DNA damage. Cell Cycle 2006;5: 1675–80.
36. Sansam CL, Shepard JL, Lai K, et al. DTL/CDT2 is essential for both CDT1 regulation and the early G_2-M checkpoint. Genes Dev 2006;20:3117–29.
37. Ueki T, Nishidate T, Park JH, et al. Involvement of elevated expression of multiple cell-cycle regulator, DTL/RAMP (denticleless/RA-regulated nuclear matrix associated protein), in the growth of breast cancer cells. Oncogene 2008;27:5672–83.
38. Aaltonen K, Ahlin C, Amini RM, et al. Reliability of cyclin A assessment on tissue microarrays in breast cancer compared to conventional histological slides. Br J Cancer 2006;94:1697–702.
39. Husdal A, Bukholm G, Bukholm IR. The prognostic value and overexpression of cyclin A is correlated with gene amplification of both cyclin A and cyclin E in breast cancer patient. Cell Oncol 2006;28:107–16.
40. Galardi S, Mercatelli N, Giorda E, et al. miR-221 and miR-222 expression affects the proliferation potential of human prostate carcinoma cell lines by targeting p27Kip1. J Biol Chem 2007;282:23716–24.
41. Medina R, Zaidi SK, Liu CG, et al. MicroRNAs 221 and 222 bypass quiescence and compromise cell survival. Cancer Res 2008;68:2773–80.
42. Kim YK, Yu J, Han TS, et al. Functional links between clustered microRNAs: suppression of cell-cycle inhibitors by microRNA clusters in gastric cancer. Nucleic Acids Res 2009;37:1672–81.
43. Mercatelli N, Coppola V, Bonci D, et al. The inhibition of the highly expressed miR-221 and miR-222 impairs the growth of prostate carcinoma xenografts in mice. PLoS ONE 2008;3:e4029.
44. Roy S, Singh RP, Agarwal C, Siriwardana S, Sclafani R, Agarwal R. Downregulation of both p21/Cip1 and p27/Kip1 produces a more aggressive prostate cancer phenotype. Cell Cycle 2008;7:1828–35.
45. Linsley PS, Schelter J, Burchard J, et al. Transcripts targeted by the microRNA-16 family cooperatively regulate cell cycle progression. Mol Cell Biol 2007;27: 2240–52.
46. Liu Q, Fu H, Sun F, et al. miR-16 family induces cell cycle arrest by regulating multiple cell cycle genes. Nucleic Acids Res 2008;36:5391–404.
47. Kaddar T, Rouault JP, Chien WW, et al. Two new miR-16 targets: caprin-1 and HMGA1, proteins implicated in cell proliferation. Biol Cell 2009;101: 511–24.
48. Sachdeva M, Zhu S, Wu F, et al. p53 represses c-Myc through induction of the tumor suppressor miR-145. Proc Natl Acad Sci U S A 2009;106:3207–12.
49. Wang L, Oberg AL, Asmann YW, et al. Genome-wide transcriptional profiling reveals microRNA-correlated genes and biological processes in human lymphoblastoid cell lines. PLoS One 2009;4:e5878.