

***In silico* functional analyses and discovery of survival-associated microRNA signatures in pediatric osteosarcoma**

Patricia C. Sanchez-Diaz^{1,6}, Tzu-Hung Hsiao¹, Yi Zou¹, Aaron J. Sugalski², Josefine Heim-Hall^{3,4}, Yidong Chen^{1,4,5}, Anne-Marie Langevin^{2,4}, and Jaclyn Y. Hung^{1,2,4}

¹ Greehey Children's Cancer Research Institute, University of Texas Health Science Center at San Antonio, San Antonio, Texas, USA

² Division of Hematology and Oncology, Department of Pediatrics, University of Texas Health Science Center at San Antonio, San Antonio, Texas, USA

³ Department of Pathology, University of Texas Health Science Center at San Antonio, San Antonio, Texas, USA

⁴ Cancer Therapy and Research Center, University of Texas Health Science Center at San Antonio, San Antonio, Texas, USA

⁵ Department of Epidemiology and Biostatistics, University of Texas Health Science Center at San Antonio, San Antonio, Texas, USA

⁶ Current address: Rosenberg School of Optometry, University of the Incarnate Word, San Antonio, Texas, USA

Correspondence to: Jaclyn Y. Hung, email: hungj@uthscsa.edu

Yidong Chen, email: chen8@uthscsa.edu

Keywords: osteosarcoma, microRNA expression, prognosis, pathways, pediatric cancers

Received: August 20, 2014

Accepted: September 17, 2014

Published: September 23, 2014

This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Purpose: Osteosarcoma is the most common bone tumor in children, adolescents, and young adults. In contrast to other childhood malignancies, no biomarkers have been consistently identified as predictors of outcome. This study was conducted to assess the microRNAs(miRs) expression signatures in pre-treatment osteosarcoma specimens and correlate with outcome to identify biomarkers for disease relapse.

Results: A 42-miRs signature whose expression levels were associated with overall and relapse-free survival was identified. There were 8 common miRs between the two sets of survival-associated miRs. Bioinformatic analyses of these survival-associated miRs suggested that they might regulate genes involved in ubiquitin proteasome system, TGF β , IGF, PTEN/AKT/mTOR, MAPK, PDGFR/RAF/MEK/ERK, and ErbB/HER pathways.

Methods: The cohort consisted of 27 patients of 70% Mexican-American ethnicity. High-throughput RT-qPCR approach was used to generate quantitative expression of 754 miRs in the human genome. We examined tumor recurrence status, survival time and their association with miR expression levels by Cox proportional hazard regression analysis. TargetScan was used to predict miR/genes interactions, and functional analyses using KEGG, BioCarta, Gene Ontology were applied to these potential targets to predict deregulated pathways.

Conclusions: Our findings suggested that these miRs might be potentially useful as prognostic biomarkers and therapeutic targets in pediatric osteosarcoma.

INTRODUCTION

Osteosarcoma is the most common bone malignancy affecting approximately 400 children, adolescents, and young adults in the United States each year, and for which no appreciable improvement in outcome has been attained

in the past 20 years, indicating that we have reached the limits of classical non-targeted therapies. Despite extensive mutilating surgery followed by aggressive chemotherapy, the survival rate for this malignant bone cancer is just

60-70%. Moreover, 40% of patients with osteosarcoma experience metastatic recurrent disease, in which case the survival rates are even lower. For the survivors, most develop debilitating long-term side effects after intensive curative treatment that impact their quality of life. Osteosarcoma arises primarily in the metaphysis of the long bones with a predilection for the areas of most intense bone growth in skeletally immature patients, i.e. distal femur, proximal tibia and proximal humerus. The peak incidence is seen in the second decade of life and seems to correlate with the growth spurt associated with pubertal development [1]. A monograph on cancer in the adolescent

and young adult (AYA) population published by the NCI in 2006 revealed that Hispanics have the highest incidence of osteosarcoma in the 15 to 29-age group [2]. A retrospective analysis of the characteristics and outcome in a cohort of Hispanics children and AYA patients of predominately Mexican-American ethnicity with localized high-grade osteosarcoma of the extremity finds a striking increased rate of relapse in the preadolescent patients (before the age of 12). The study also finds that percentage of tumor necrosis after neoadjuvant chemotherapy is not directly predictive of outcome for this population of homogenously Mexican-American ancestry [3].

Table 1: Overall survival-associated miR. Cox regression parameters (β , confidence interval (CI) and P value) are provided in the table, as well miRNA's chromosomal location.

miR	β	95% CI	P value	Chromosome
<i>hsa-miR-221</i>	0.53	[0.21 – 0.86]	0.001	Xp11.3
<i>hsa-miR-92a</i>	0.25	[0.08 – 0.42]	0.003	13q31.3 (1) or Xq26.2 (2)
<i>hsa-miR-770-5p</i>	0.5	[0.16 – 0.83]	0.003	14q32.2
<i>hsa-miR-628-3p</i>	0.24	[0.07 – 0.41]	0.01	15q21.3
<i>hsa-miR-505</i>	0.77	[0.23 – 1.31]	0.01	Xq27.1
<i>hsa-miR-541</i>	0.22	[0.05 – 0.39]	0.01	14q32.31
<i>hsa-miR-499-5p</i>	0.39	[0.09 – 0.69]	0.01	20q11.22
<i>hsa-miR-455-3p</i>	0.74	[0.15 – 1.32]	0.01	9q32
<i>hsa-miR-29a</i>	0.77	[0.14 – 1.40]	0.02	7q32.3
<i>hsa-miR-876-5p</i>	0.15	[0.02 – 0.28]	0.02	9p21.1
<i>hsa-miR-890</i>	0.18	[0.03 – 0.34]	0.02	Xq27.3
<i>hsa-let-7c</i>	0.64	[0.10 – 1.18]	0.02	21q21.1
<i>hsa-miR-545</i>	0.47	[0.06 – 0.88]	0.02	Xq13.2
<i>hsa-miR-486-3p</i>	0.19	[0.02 – 0.35]	0.03	8p11.21
<i>hsa-miR-622</i>	0.38	[0.04 – 0.73]	0.03	13q31.3
<i>hsa-miR-1255b</i>	0.33	[0.03 – 0.63]	0.03	4p14 (1) or 1q24.2 (2)
<i>hsa-miR-483-3p</i>	0.29	[0.03 – 0.56]	0.03	11p15.5
<i>hsa-miR-30c</i>	0.19	[0.02 – 0.37]	0.03	1p34.2 (1) or 6q13 (2)
<i>hsa-miR-485-3p</i>	0.24	[0.01 – 0.48]	0.04	14q32.31
<i>hsa-miR-21</i>	0.4	[0.01 – 0.79]	0.04	17q23.1
<i>hsa-miR-374b</i>	0.18	[0.00 – 0.36]	0.04	Xq13.2
<i>hsa-miR-573</i>	0.48	[0.01 – 0.95]	0.05	4p15.2
<i>hsa-miR-1282</i>	0.93	[0.02 – 1.83]	0.05	15q15.3
<i>hsa-miR-656</i>	0.43	[0.01 – 0.86]	0.05	14q32.31
<i>hsa-miR-202</i>	0.34	[0.00 – 0.69]	0.05	10q26.3
<i>hsa-miR-429</i>	-0.29	[-0.51 – -0.07]	0.01	1p36.33
<i>hsa-miR-126</i>	-0.36	[-0.64 – -0.08]	0.01	9q34.3
<i>hsa-miR-888</i>	-0.56	[-1.02 – -0.10]	0.02	Xq27.3
<i>hsa-miR-517b</i>	-0.22	[-0.41 – -0.03]	0.03	19q13.42
<i>hsa-miR-363</i>	-0.43	[-0.80 – -0.05]	0.03	Xq26.2
<i>hsa-miR-216b</i>	-0.26	[-0.50 – -0.03]	0.03	2p16.1
<i>hsa-miR-1285</i>	-0.27	[-0.53 – -0.02]	0.04	7q21.2 (1) or 2p13.3 (2)

In contrast to other childhood malignancies, very few clinically useful indicators have been consistently identified as putative predictors of outcome. To this date, incomplete surgery remains the most important negative prognostic indicator, followed by poor response to preoperative chemotherapy, primary metastases, and axial location [4-8]. While osteosarcoma has wide-ranging histologic appearances, proliferation of malignant mesenchymal cells and production of osteoid and/or bone by the tumor cells are key features and are mandatory for clinical diagnosis. The molecular events underpinning the biology of osteosarcoma remain poorly understood but there is evidence that osteosarcoma is a differentiation disease where normal osteoblast differentiation from the mesenchymal stem cell is disrupted as a result of genetic and epigenetic changes [9-11]. In addition, numerous structural and numerical cytogenetic abnormalities have been described [12]. Yet none have been consistently predictive of survival or response to chemotherapy.

It is well known that microRNA (miR) deregulation is often associated with many cancers [13]. miRs signatures have been suggested to be a better predictor of outcome than mRNA expression profiles. A comparison between miR profiling and mRNA profiling shows that a molecular signature consisting of about 200 miRs is more accurate in classifying highly malignant, poorly differentiated tumors than mRNA profiles containing more than 20,000 genes [14]. *In silico* analyses indicate that a single miR likely regulates over 100 genes [15].

In this study, we assessed the miR expression

profiles in 27 patients diagnosed with osteosarcoma localized to the extremities. The miR signatures of the pre-treatment tumor specimens were used to correlate with clinical outcome. Using Cox proportional hazard regression analyses, 32 miRs were identified to be associated with overall survival (OV-miRs), and 18 miRs were shown to predict relapse (RE-miRs). We then applied bioinformatics methods to predict pathways and functions of these survival-associated miR signatures. Our findings indicated that the survival-associated miR signatures and the putative pathways they regulate might have the potential to improve clinical management and future treatment decision-making for patients with this cancer.

RESULTS

Clinical characteristic of the cohort

The cohort consisted of 27 patients diagnosed with osteosarcoma of the extremities. All except one patient had their tumor specimens collected and banked at the time of diagnosis (Table S1).

The Kaplan–Meier plots of gender and race are included as supplementary data (Figure S1). Although the survival curves of gender were separated, the *P* values of the log-rank test for overall and relapse-free survival did not exhibit statistical significance. Neither the survival curve nor log-rank test showed racial difference.

Table 2: Relapse-free survival-associated miR

miR	β	95% CI	<i>P</i> value	Chromosome
<i>hsa-miR-483-3p</i>	0.36	[0.08 – 0.64]	0.01	11p15.5
<i>hsa-miR-500</i>	0.7	[0.14 – 1.26]	0.01	Xp11.23
<i>hsa-miR-545</i>	0.63	[0.11 – 1.14]	0.02	Xq13.2
<i>hsa-miR-221</i>	0.44	[0.08 – 0.81]	0.02	Xp11.3
<i>hsa-miR-30b</i>	0.45	[0.08 – 0.82]	0.02	8q24.22
<i>hsa-miR-541</i>	0.22	[0.03 – 0.41]	0.03	14q32.31
<i>hsa-miR-30a</i>	0.29	[0.03 – 0.55]	0.03	6q13
<i>hsa-miR-146b-5p</i>	0.83	[0.05 – 1.61]	0.04	10q24.32
<i>hsa-miR-668</i>	0.58	[0.03 – 1.13]	0.04	14q32.31
<i>hsa-miR-770-5p</i>	0.4	[0.01 – 0.79]	0.04	14q32.2
<i>hsa-miR-335</i>	0.42	[0.01 – 0.83]	0.04	7q32.2
<i>hsa-miR-328</i>	0.24	[0.01 – 0.47]	0.04	16q22.1
<i>hsa-miR-449b</i>	0.22	[0.00 – 0.43]	0.05	5q11.2
<i>hsa-miR-499-5p</i>	0.36	[0.00 – 0.71]	0.05	20q11.22
<i>hsa-miR-429</i>	-0.39	[-0.65 – -0.13]	0.003	1p36.33
<i>hsa-miR-151-3p</i>	-0.18	[-0.32 – -0.04]	0.01	8q24.3
<i>hsa-miR-888</i>	-0.83	[-1.48 – -0.17]	0.01	Xq27.3
<i>hsa-miR-511</i>	-0.24	[-0.47 – -0.01]	0.04	10p12.33

Identification of survival-associated miRs

Univariate Cox proportional analysis was performed to identify a miR signature that showed association with patients' survival, we chose P value of Cox regression model <0.05 as criterion. A total of 32 and 18 miRs were associated with overall and relapse-free survival (Tables 1 and 2, respectively). In the OV-miRs, 25 out of 32 had positive hazard ratio ($\beta > 0$), 7 had negative hazard ratio ($\beta < 0$). Among them, *hsa-miR-221* and *hsa-miR-429* had the most significant P value in terms of positive and negative hazard ratio respectively. The Kaplan–Meier plots of *hsa-miR-221* and *hsa-miR-429*, shown as Figure 1A (left and right panels respectively), illustrate the survival difference between the high and low expressed miRs. In our analysis, we found that patients with a high expression of *hsa-miR-429* had better survival, whereas a poorer survival was observed in patients with a high expression of *hsa-miR-221*. For the 18 RE-miRs, 14 had positive hazard ratio and 4 negative hazards. The Kaplan–Meier plots of two of the RE-miRs are shown in Figure 1B. There were 8 overlapping miRs between the two sets of miRs (Figure S2). The possibility of selecting the 8 overlapping miR by chance was $P < 1 \times 10^{-7}$ (Fisher's exact test).

In order to enhance the prediction power, multivariate Cox hazard regression analysis was done to assess the correlation between the survival-associated miRs pairs and patients' survival. Using a selection criteria of P values < 0.01 in the miRs pair (listed in supplementary Table S2 and S3), 30 and 13 pairs of miRs were found to be associated with overall and relapse-free survival, respectively. For example, multivariate analysis improved the P value of the OV-miR pair *hsa-miR-221* and *hsa-miR-1285*; *hsa-miR-221* from 0.001 (Table 1) to 0.0003 (Table S2) and *hsa-miR-1285* from 0.04 to 0.003. The Kaplan–Meier plots of this OV-miR pair is shown on Figure 2A. In Figure 2B, RE-miR pair of *hsa-miR-483-3p* ($P < 0.01$) and *hsa-miR-151-3p* ($P < 0.04$) showed improved predictive power ($P < 0.0001$).

Survival-associated miR predicted target genes

The target genes of the survival-associated miRs were predicted using TargetScan (<http://www.targetscan.org>). The *in silico* analysis indicated 870 and 1,059 target genes regulated by the OV-miRs and RE-miRs, respectively (supplementary Table S4A and S4B). Our analysis revealed that each OV-miR and RE-miR

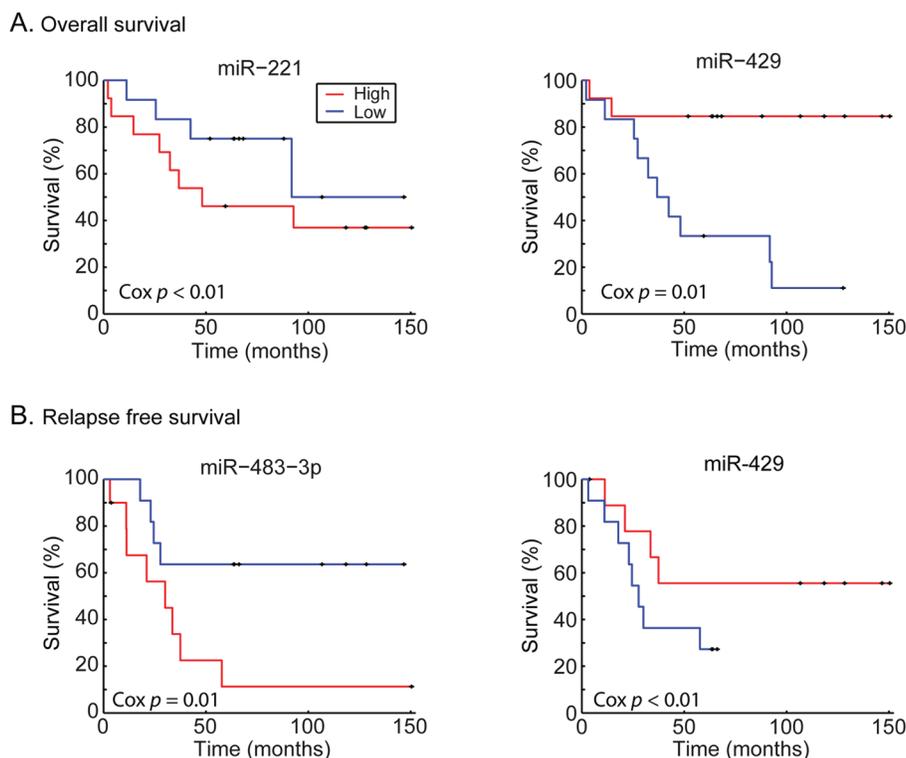


Figure 1: Kaplan–Meier plots of the survival-associated miRs. The Kaplan–Meier plots of 4 miRs associated with patients' survival separated by lower or higher expression abundance. The P values of the 4 miRs in Cox regression model were all statistically significant. In each miRs, sample was assigned to two groups, higher expression (red line) and lower expression (blue line), according to the ranking of the expression level by using the 50th percentile as criterion. (A) *hsa-miR-221* and *hsa-miR-429* were associated with overall survival with positive and negative regression coefficients. The patients with high expression of *hsa-miR-221* had worse survival rate but better survival in *hsa-miR-429*. (B) *hsa-miR-483-3p* had significant P value in Cox model of relapse-free survival. The patients with high expression had worse survival. *hsa-miR-429* was the most significance miRs with negative coefficients

regulated on average 80.3 and 88.3 of the predicted genes, respectively (Figure 3). The OV-miR, *hsa-miR-1284* regulated the largest numbers of predicted genes (462). *hsa-miR-541* was the RE-miR with the most number of predicted target genes, 264.

Enrichment analyses and pathway analysis

Discovery of the function of the target genes of the survival-associated miRs and pathway enrichment analyses were performed using GO BP, KEGG and BioCarta pathways bioinformatic tools. As shown in Tables S5 and S6, the number of KEGG pathways of the OV-miRs and RE-miRs data sets was 36 and 13, respectively. In the KEGG_GLIOMA pathway, oncogenes and tumor suppressor genes, such as *KRAS*, *MDM2*, *E2F1*, *PTEN*, *SOS1* and *MAPK1* were regulated by the OV-miRs (Figure 4A). In the KEGG_ENDOCYTOSIS, shown in Figure 4B, TGFβ signaling and the ubiquitin proteasome pathway containing genes such as ubiquitin E3 ligases and *UBPY* and *AMSH* (deubiquitinating enzymes) were regulated by the RE-miRs. We identified 20 and 8 items in the BioCarta

signaling pathways for OV-miRs and RE-miRs data sets, respectively (Table S7 and S8). The BIOCARTA EIF4 PATHWAY, which was the pathway regulated by eIF4e and p70 S6 Kinase and BIOCARTA CREB PATHWAY were regulated by the OV-miRs and RE-miRs, respectively (Figure S3). Several known oncogenes and tumor suppressor genes, such as p38, *ERK1*, *ERK2*, *PTEN*, and *PI3K*, were regulated by the OV-miRs in this pathway. The GO BP analysis identified 66 and 52 items with statistical significance in the OV-miRs and RE-miRs data sets, respectively (Tables S9 and S10). Our analysis showed some overlapping GO terms between the two data sets including the regulation of transcription, DNA-dependent, synaptic transmission, *in utero* embryonic development, or post-embryonic development ($P < 10^{-4}$).

DISCUSSION

Our population of children and AYA diagnosed with osteosarcoma localized to the extremities was 70% Hispanic, predominantly of Mexican American ancestry, providing us with a unique cohort. Five of these

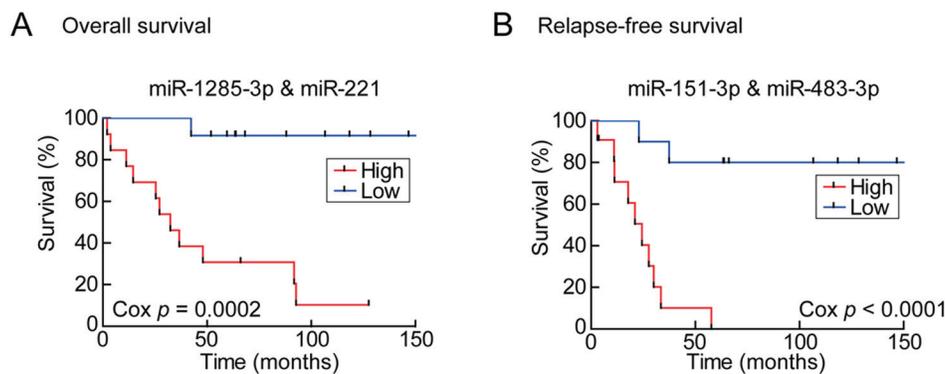


Figure 2: Kaplan–Meier plots of the survival-associated miR pairs. The miRs pairs were identified by multi-variants Cox regression model. Each sample was assigned to two groups, higher (red line) and lower (blue line) hazard, based on the hazard values. (A) The miR pairs of *hsa-miR-1285-3p* and *hsa-miR-221* and (B) *hsa-miR-151-3p* and *hsa-miR-483-3p* (B) had statistical significance in overall and relapse-free survival.

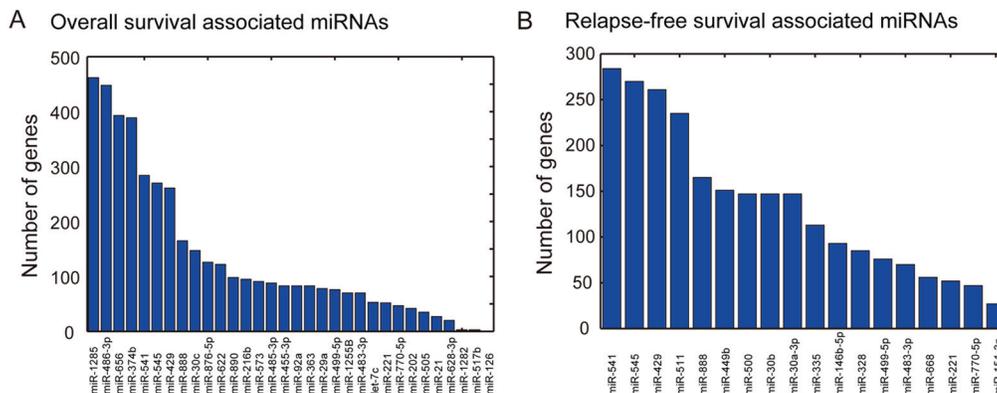


Figure 3: The number of genes targeted by the survival-associated miRs. (A) Bar charts of the number of genes regulated by OV-miRs and (B) RE-miRs. *hsa-miR-1285* and *hsa-miR-541* had the largest number of genes (462 and 264 genes, respectively) targeted by the survival-associated miRs.

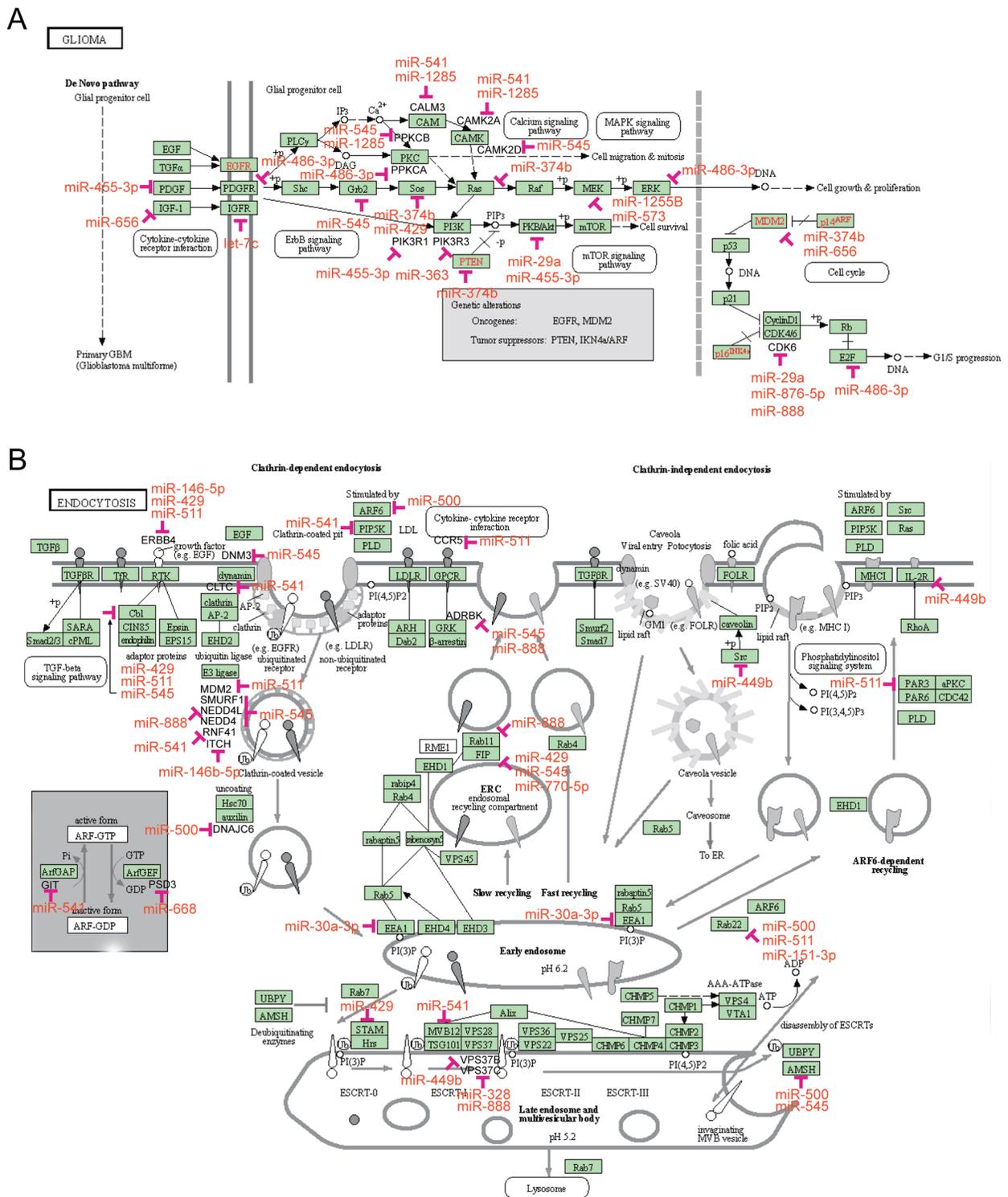


Figure 4: KEGG pathways enrichment analysis of the survival-associated miRNAs target genes. (A) The GLIOMA signaling pathway of KEGG includes 26 genes that were highly regulated by the OV-miRs. Fifteen of the 42 survival-associated miRNAs were involved in this pathway. Eleven of the miRNAs (*hsa-miR-876-5p*, *hsa-miR-1255b*, *hsa-miR573*, *hsa-miR-29a*, *hsa-miR-455-3p*, *hsa-miR374b*, *hsa-miR-545*, *hsa-miR-486-3p*, *hsa-miR-541*, *hsa-miR-656*, and *hsa-miR-lei-7c*) were of poor prognosis. miR/gene interactions included growth factors, oncogenes and tumor suppressor genes, such as *IGF-1*, *PDGF*, *CDK6*, *E2F1*, *MDM2*, *RAS*, *SRC*, *AKT*, *ERK* and *PTEN*. (B) The endocytosis pathway of KEGG was regulated by the RE-miRs. E3 ligases and deubiquitinases such as *MDM2*, *SMUR1*, *NEDD4L*, *NEDD4*, *RNF41*, *ITCH*, *CBL*, *UBPY*, and *AMSH* were targeted by 9 (*hsa-miR-545*, *hsa-miR-541*, *hsa-miR-511*, *hsa-miR-888*, *hsa-miR-146-5p*, *hsa-miR-429*, *hsa-miR-374b*, *hsa-miR-500*, and *hsa-miR-656*) of the predicted poor prognosis miRNAs.

patients were preadolescent (below 12 years of age at diagnosis) of which 90% was Hispanic. Twelve patients in our cohort deceased, including the 5 preadolescents (Table S1). In our cohort, we identified 32 OV-miRs and 18 RE-miRs to be predictive of outcome. There were 8 common miRs (*hsa-miR-429*, *hsa-miR-221*, *hsa-miR-499-5p*, *hsa-miR-888*, *hsa-miR-770-5p*, *hsa-miR-545*, *hsa-miR-541*, *hsa-miR-483-3p*) between the two sets of survival-associated miRs (Figure S2). To the best of our knowledge, there were 29 unique miRs in our 42 survival-associated miR signature that had not been reported to be associated with prognosis in osteosarcoma. This could in part be explained by the nature of our cohort, since 70% of the population treated at UTHSCSA is Hispanic of Mexican American ancestry, suggesting that the underlying tumor biology in this patients' population may be different from previous reports. However, this possibility needs to be confirmed in a larger multi-institutional study with patients of similar demographics. Interestingly, 5 of the 42 survival-associated miRs (*hsa-miR-770-5p*, *hsa-miR-541*, *hsa-miR-485-3p*, *hsa-miR-656*, *hsa-miR-668*) were encoded in the chromosome 14q32 region. These 5 miRs were associated with poor prognosis in our cohort, which is keeping with previous reports that indicate the miRs cluster at 14q32 region drives aggressiveness in several malignancies including osteosarcoma [17-23].

We also used systems biology approaches to predict functions and potential pathways regulated by the two sets of miRs. TargetScan was employed to predict the putative genes targeted by the survival-associated miRs. We applied the KEGG and BioCarta pathways and GO BP to determine *in silico* gene function and performed pathway enrichment analyses of the OV-miRs and RE-miRs data sets. As well, the Fisher's exact test was used to confirm that the potential pathways identified were not due to random chance. The computational analyses of these 42 survival-associated miRs suggested that these OV- and RE- miRs might regulate genes involved in endocytosis, the ubiquitin proteasome system (UPS), TGF beta, IGF, PTEN/AKT/mTOR, MAPK, PDGFR/Raf/MEK/ERK, and ErbB/HER pathways. Interestingly, the bioinformatics analyses of the miR signature that was predictive of poor prognosis revealed miR/gene interactions involved in biological processes such as apoptosis, migration, proliferation and cell cycle. Intriguingly, several of the predictor of poor prognosis miRs targeted genes that belong to the UPS that regulates protein homeostasis. It is well known that the UPS system includes a myriad of regulatory proteins that play essential roles in cell cycle progression, tumorigenesis, DNA transcription and repair, and differentiation [24-26]. In addition, evidence suggests that components of this degradation machinery may act as tumor suppressors

or as oncogenes in a context dependent manner [27-29]. Moreover, targeting specific components of the UPS for the modulation of protein degradation is an emerging and promising therapeutic strategy [30-32]. Strikingly, 9 of the poor prognosis associated miRs: *hsa-miR-545*, *hsa-miR-541*, *hsa-miR-511*, *hsa-miR-888*, *hsa-miR-146-5p*, *hsa-miR-429*, *hsa-miR-374b*, *hsa-miR-500*, and *hsa-miR-656* have functions in E3 ligases and deubiquitinases (Figure 4B). Notably, 3 of these miRs (*hsa-miR-545*, *hsa-miR-541*, and *hsa-miR-511*) targeted more than one component of the UPS system. These observations indicated that further investigations are needed to characterize these miRs to improve our understanding of the biology of osteosarcoma and discover novel targets for therapeutic intervention.

In summary, we surveyed 27 children and AYA diagnosed with osteosarcoma localized to the extremities treated at a single institution for survival-associated miRs. Our cohort was homogeneously 70% Hispanic of Mexican American ancestry. We identify a novel 42-miRs signature that was predictive of outcome. Using computational approaches we also predicted miR/genes interactions and pathways regulated by these miRs. Our findings suggested additional makers for prognostics and therapeutic intervention in these patients. Further studies are warranted to evaluate the use of these miR signatures in the clinical settings.

MATERIALS AND METHODS

Patient cohort

The cohort consisted of 27 children and AYA diagnosed with osteosarcoma localized to the extremities between January 1, 2000 and December 31, 2009 at the University of Texas Health Science Center at San Antonio (UTHSCSA). All patients were treated in a standardized fashion with cisplatin, doxorubicin and methotrexate with or without the addition of ifosfamide and etoposide. The study was conducted after receiving approval from UTHSCSA Institutional Review Board (IRB). Tumor tissue was obtained at the time of the initial biopsy prior to treatment. Specimens were immediately frozen and banked according to national standards in the IRB-approved UTHSCSA Oncological Musculoskeletal tumor bank. For correlation with miRs results, the data elements for demographic and socioeconomic variables such as: patient age, gender, race (NCI-standard), and ethnicity (NCI-standard), and clinical variables such as: histology subtype, location and staging of primary tumor, location and date of relapse date, and date of death were collected.

RNA extraction and high-throughput miR expression analysis

In a -20°C cryostat, 10 slices, each 10 µm thick frozen tissue were cut (total thickness 100 µm) and placed in 50 mL centrifuge tube containing pre-chill RNAlater-ICE solution (Ambion, Applied Biosystems). Total RNA was extracted using the mirVana miR Isolation Kit (Ambion) according to manufacturer's instructions. Sample quantitation and integrity checks (measured as RNA Integrity Number; RIN) were performed using the Agilent 2100 Bioanalyzer (Agilent Technologies). The global profiling for miR expression was performed using the TaqMan Array Human MicroRNA Cards (Applied Biosystems), which includes Cards A (v2) and B (v3) in a 384-well format, containing a total of 754 specific assays specific to miRs in the human genome, annotated in the Sanger miRBase v14 Registry (<http://microrna.sanger.ac.uk>). Each card also contained 3 endogenous and 1 negative control assays. The stem-loop quantitative polymerase chain reaction (qRT-PCR) assay was performed according to the manufacturer's protocol and described in [16]. The data collected were then processed using the Plate Utility and Automation Controller software (Applied Biosystems). For each miR, the expression level was determined by the threshold cycle method ($2^{-\Delta Ct}$) of relative quantification ($\Delta Ct = Ct$ of sample – Ct of endogenous controls of sample). Quantile normalization was applied to compensate for the technical differences present among all the samples so as to determine the real biological differences among sample.

Data analysis

Cox hazard regression model was applied to calculate the hazard risk of the expression value of each miR. Univariate and multivariate analyses were conducted for each miR measured in the 27 patients to predict patient overall and relapse-free survival. The miRs that had P values < 0.05 were identified as either overall or relapse-free survival-associated miRs (OV-miRs or RE-miRs). For univariate Cox analysis, the good and poor prognosis patients were divided through 50% percentile of miR expression. Kaplan–Meier analyses were performed to visualize the predictive value of either of the individual or paired survival-associated miRs. To assess the correlation between two of the survival-associated miRs and patient survival, multivariate Cox hazard regression analysis was done. The patients were separated into two groups based on the hazard values. For patient i the hazard can be calculated as Equation (1)

$$HR_i = e^{(\beta_1 x_{1,i} + \beta_2 x_{2,i})}$$

where β_1 and β_2 were the coefficients of the two miRs and $x_{1,i}$ and $x_{2,i}$ were the miR expression abundances of patient i .

Prediction of miR targets genes and gene sets of KEGG, Biocarta pathway and Gene Ontology

miR target genes were identified based on the numbers of miR binding sites which were downloaded from TargetScan database (<http://www.targetscan.org/>). The gene containing more than 3 binding site of specific miR on its transcript's sequence were considered as the miR target genes. For the gene that have multiple transcripts, the variant 1 recorded in Refseq database was selected as the presented transcript. The analyses of the gene sets of KEGG (<http://www.genome.jp/kegg/pathway.html>) and Biocarta pathways (<http://www.biocarta.com/>) and Biological Processes in Gene Ontology (GO BP) (<http://www.geneontology.org/>) performed as described in [16].

miR regulated genes and enrichment analysis of miR regulated gene set

Pathway enrichment analysis of the survival-associated miR regulated gene set was obtained through Fisher's exact test. The P value of Fisher's exact test between the miR regulated genes and the genes in pathways can be calculated by

$$P(x > z_l) = \sum_{h=z_l}^{\infty} \frac{\binom{r}{h} \binom{M-r}{y_l-h}}{\binom{M}{y_l}}$$

where M is the assume number of genes in the genome, r is the number of miR regulated genes, y_l is the number of genes in gene set l , and z_l is the number of miR regulated genes in gene set l . P value < 0.05 was considered significant.

ACKNOWLEDGMENTS

Funding for this work was provided in part by a National Institutes of Health NIH-NCATS UL1TR000149 CTSA grant through the Institute of Medicine and Science (IIMS) at the University of Texas Health Science Center at San Antonio, the NCI Cancer Center Support Grant 2 P30 CA054174-17 to Cancer Therapy and Research Center (CTRC), the Greehey Cancer Research Institute Permanent Health Fund Permanent Endowments – Higher Education (Children's Cancer Research) at the University of Texas Health Science Center at San Antonio.

REFERENCES

1. Ottaviani G, Jaffe N. The epidemiology of osteosarcoma. *Cancer Treat Res.* 2009; 152: 3-13.
2. Mascarenhas L, Siegel S, Spector L, Arndt C, Femino D, Malogolowkin M. Malignant Bone Tumors. In: Bleyer A, O'Leary M, Barr R, Ries L, editors. *Cancer Epidemiology*

- in Older Adolescents and Young Adults 15 to 29 Years of Age, Including SEER Incidence and Survival: 1975-2000. Bethesda, MD National Cancer Institute. 2006; 97-110.
3. Sugalski AJ, Jiwani A, Ketchum NS, Cornell J, Williams R, Heim-Hall J, Hung JY, Langevin AM. Characterization of Localized Osteosarcoma of the Extremity in Children, Adolescents, and Young Adults From a Single Institution in South Texas. *J Pediatr Hematol Oncol.* 2014; 36: e353-e358.
 4. Bielack SS, Kempf-Bielack B, Delling G, Exner GU, Flege S, Helmke K, Kotz R, Salzer-Kuntschik M, Werner M, Winkelmann W, Zoubek A, Jurgens H, Winkler K. Prognostic factors in high-grade osteosarcoma of the extremities or trunk: an analysis of 1,702 patients treated on neoadjuvant cooperative osteosarcoma study group protocols. *J Clin Oncol.* 2002; 20: 776-790.
 5. Bieling P, Rehan N, Winkler P, Helmke K, Maas R, Fuchs N, Bielack S, Heise U, Jurgens H, Treuner J, Romanowski R, Exner U, Kotz R, Winkler K. Tumor size and prognosis in aggressively treated osteosarcoma. *J Clin Oncol.* 1996; 14: 848-858.
 6. Arndt CA, Rose PS, Folpe AL, Laack NN. Common musculoskeletal tumors of childhood and adolescence. *Mayo Clin Proc.* 2012; 87: 475-487.
 7. Janeway KA, Barkauskas DA, Krailo MD, Meyers PA, Schwartz CL, et al. Outcome for adolescent and young adult patients with osteosarcoma: a report from the Children's Oncology Group. *Cancer.* 2012; 118: 4597-4605.
 8. Ferrari S, Mercuri M, Bacci G. Comment on "Prognostic factors in high-grade osteosarcoma of the extremities or trunk: an analysis of 1,702 patients treated on neoadjuvant Cooperative Osteosarcoma Study Group protocols". *J Clin Oncol.* 2002; 20: 2910; author reply 2910-2911.
 9. Mohseny AB, Szuhai K, Romeo S, Buddingh EP, Briaire-de Bruijn I, de Jong D, van Pel M, Cleton-Jansen AM, Hogendoorn PC. Osteosarcoma originates from mesenchymal stem cells in consequence of aneuploidization and genomic loss of Cdkn2. *J Pathol.* 2009; 219: 294-305.
 10. Gorlick R. Current concepts on the molecular biology of osteosarcoma. *Cancer Treat Res.* 2009; 152: 467-478.
 11. Tang N, Song WX, Luo J, Haydon RC, He TC. Osteosarcoma development and stem cell differentiation. *Clin Orthop Relat Res.* 2008; 466: 2114-2130.
 12. Guillou L, Aurias A. Soft tissue sarcomas with complex genomic profiles. *Virchows Arch.* 2010; 456: 201-217.
 13. Metias SM, Lianidou E, Yousef GM. MicroRNAs in clinical oncology: at the crossroads between promises and problems. *J Clin Pathol.* 2009; 62: 771-776.
 14. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA, Downing JR, Jacks T, Horvitz HR, Golub TR. MicroRNA expression profiles classify human cancers. *Nature.* 2005; 435: 834-838.
 15. Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* 2009; 19: 92-105.
 16. Sanchez-Diaz PC, Hsiao TH, Chang JC, Yue D, Tan MC, Chen HI, Tomlinson GE, Huang Y, Chen Y, Hung JY. Deregulated microRNAs in pediatric cancer stem cells target pathways involved in cell proliferation, cell cycle and development. *PLoS One.* 2013; 8: e61622.
 17. Nadal E, Zhong J, Lin J, Reddy RM, Ramnath N, Orringer MB, Chang AC, Beer DG, Chen G. A MicroRNA Cluster at 14q32 Drives Aggressive Lung Adenocarcinoma. *Clin Cancer Res.* 2014; 20: 3107-3117.
 18. Sarver AL, Thayanithy V, Scott MC, Cleton-Jansen AM, Hogendoorn PC, Modiano JF, Subramanian S. MicroRNAs at the human 14q32 locus have prognostic significance in osteosarcoma. *Orphanet J Rare Dis.* 2013; 8: 7.
 19. Zehavi L, Avraham R, Barzilai A, Bar-Ilan D, Navon R, Sidi Y, Avni D, Leibowitz-Amit R. Silencing of a large microRNA cluster on human chromosome 14q32 in melanoma: biological effects of mir-376a and mir-376c on insulin growth factor 1 receptor. *Mol Cancer.* 2012; 11: 44.
 20. Kelly AD, Haibe-Kains B, Janeway KA, Hill KE, Howe E, Goldsmith J, Kurek K, Perez-Atayde AR, Francoeur N, Fan JB, April C, Schneider H, Gebhardt MC, Culhane A, Quackenbush J, Spentzos D. MicroRNA paraffin-based studies in osteosarcoma reveal reproducible independent prognostic profiles at 14q32. *Genome Med.* 2013; 5: 2.
 21. Luk JM, Burchard J, Zhang C, Liu AM, Wong KF, Shek FH, Lee NP, Fan ST, Poon RT, Ivanovska I, Philippart U, Cleary MA, Buser CA, Shaw PM, Lee CN, Tenen DG, Dai H, Mao M. DLK1-DIO3 genomic imprinted microRNA cluster at 14q32.2 defines a stemlike subtype of hepatocellular carcinoma associated with poor survival. *J Biol Chem.* 2011; 286: 30706-30713.
 22. Haller F, von Heydebreck A, Zhang JD, Gunawan B, Langer C, Ramadori G, Wiemann S, Sahin O. Localization- and mutation-dependent microRNA (miRNA) expression signatures in gastrointestinal stromal tumours (GISTs), with a cluster of co-expressed miRNAs located at 14q32.31. *J Pathol.* 2010; 220: 71-86.
 23. Flores RJ, Li Y, Yu A, Shen J, Rao PH, Lau SS, Vannucci M, Lau CC, Man TK. A systems biology approach reveals common metastatic pathways in osteosarcoma. *BMC Syst Biol.* 2012; 6: 50.
 24. Severe N, Dieudonne FX, Marie PJ. E3 ubiquitin ligase-mediated regulation of bone formation and tumorigenesis. *Cell Death Dis.* 2013; 4: e463.
 25. Geng F, Wenzel S, Tansey WP. Ubiquitin and proteasomes in transcription. *Annu Rev Biochem.* 2012; 81: 177-201.
 26. Weissman AM, Shabek N, Ciechanover A. The predator becomes the prey: regulating the ubiquitin system by ubiquitylation and degradation. *Nat Rev Mol Cell Biol.* 2011; 12: 605-620.

27. Ding F, Xiao H, Wang M, Xie X, Hu F. The role of the ubiquitin-proteasome pathway in cancer development and treatment. *Front Biosci (Landmark Ed)*. 2014; 19: 886-895.
28. Mani A, Gelmann EP. The ubiquitin-proteasome pathway and its role in cancer. *J Clin Oncol*. 2005; 23: 4776-4789.
29. Fraile JM, Quesada V, Rodriguez D, Freije JM, Lopez-Otin C. Deubiquitinases in cancer: new functions and therapeutic options. *Oncogene*. 2012; 31: 2373-2388.
30. Weathington NM, Mallampalli RK. Emerging therapies targeting the ubiquitin proteasome system in cancer. *J Clin Invest*. 2014; 124: 6-12.
31. D'Arcy P, Brnjic S, Olofsson MH, Fryknas M, Lindsten K, De Cesare M, Perego P, Sadeghi B, Hassan M, Larsson R, Linder S. Inhibition of proteasome deubiquitinating activity as a new cancer therapy. *Nat Med*. 2011; 17: 1636-1640.
32. Nalepa G, Rolfe M, Harper JW. Drug discovery in the ubiquitin-proteasome system. *Nat Rev Drug Discov*. 2006; 5: 596-613.