Novel Prediction of Anticancer Drug Chemosensitivity in Cancer Cell Lines: Evidence of Moderation by microRNA Expressions

Daniel S. Yang

Abstract— **The objectives of this study are (1) to develop a novel "moderation" model of drug chemosensitivity and (2) to investigate if miRNA expression moderates the relationship between gene expression and drug chemosensitivity, specifically for HSP90 inhibitors applied to human cancer cell lines. A moderation model integrating the interaction between miRNA and gene expressions was developed to examine if miRNA expression affects the strength of the relationship between gene expression and chemosensitivity. Comprehensive datasets on miRNA expressions, gene expressions, and drug chemosensitivities were obtained from National Cancer Institute's NCI-60 cell lines including nine different cancer types. A workflow including steps of selecting genes, miRNAs, and compounds, correlating gene expression with chemosensitivity, and performing multivariate analysis was utilized to test the proposed model.**

The proposed moderation model identified 12 significantly-moderating miRNAs: miR-15b*, miR-16-2*, miR-9, miR-126*, miR-129*, miR-138, miR-519e*, miR-624*, miR-26b, miR-30e*, miR-32, and miR-196a, as well as two genes ERCC2 and SF3B1 which affect chemosensitivities of Tanespimycin and Alvespimycin—both HSP90 inhibitors. A bootstrap resampling of 2,500 times validates the significance of all 12 identified miRNAs. The results confirm that certain miRNA and gene expressions interact to produce an effect on drug response. The lack of correlation between miRNA and gene expression themselves suggests that miRNA transmits its effect through translation inhibition/control rather than mRNA degradation. The results suggest that miRNAs could serve not only as prognostic biomarkers for cancer treatment outcome but also as interventional agents to modulate desired chemosensitivity.

I. Introduction

There has been an increased interest in finding customized predictors of a patient's response to anticancer drugs in the drive toward precision medicine [6]. Studying miRNAs' roles in drug chemosensitivity helps (1) to develop "predictive tests of drug efficacy" [20, p. 629] and (2) to understand "mechanisms of drug action" [20, p. 631]. Regarding the first goal, pharmacogenomics research aims to accurately predict a patient's response to drugs in order to deliver individualized treatment. Such personalized medicine is especially important in cancer therapy where drugs often have side effects and may be ineffective in some people [23]. Because miRNAs affect gene functions and expression levels of certain genes can influence drug chemosensitivity [20], miRNAs may play important roles in tumor response to drugs [8] and have potential for predicting response to treatment [22]. Thus, "the pharmacogenetic analysis of miRNAs may represent an innovative field of research for predicting treatment response or chemoresistance." [8, p. 111] Regarding the second goal, testing a novel statistical "moderation" model that is also consistent with the current understanding of how miRNAs function can shed light on how miRNAs transmit their effect upon chemosensitivity. The NCI-60 datasets are useful in this regard since the data is centered on common variants in the entirety of human population and drug response [1]. Based on a search on PubMed, this study is the first analysis of NCI-60 datasets to examine the moderation, as opposed to the direct, effect of miRNA on chemosensitivity. Research on prognostic use of miRNAs is at an "early stage" [8, p. 115], and the incorporation of miRNA to traditional genomics in treatment tailoring is still an incipient field. Because of their role in cancer pathogenesis "miRNAs could represent predictive markers of treatment outcome in oncology" [6].

II. Background

First identified in 1993, miRNAs belong to a category of non-protein encoding RNA molecules whose primary function is to repress the action of specific messenger RNA (mRNA) molecules [20]. As shown in Fig. 1, this function is accomplished either by promoting mRNA degradation or by inhibiting mRNA translation into protein [9, 20]. Specifically, miRNAs operate at the post-transcriptional level in binding to partially complementary sequence motifs at target sites mainly in the 3′-UTRs of corresponding mRNAs, though binding sites also have been reported to exist in the 5′-UTRs as well as the open reading frame [7]. Recent research has also demonstrated a positive regulation ability of miRNA where mRNA translation might be oppositely enhanced. miRNA's positive regulation capabilities may be the result of encouraging mRNA-stabilizing factors [6]. Because gene expression is characterized by mRNA which continues to the ribosome to support protein synthesis, miRNA's regulating effect on mRNA also suggests that miRNA affects gene expression.

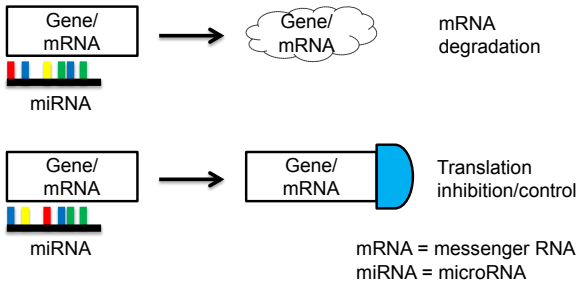

Figure 1. Mechanisms of miRNA regulation (after [20]).

D. S. Yang is with Woodbridge High School, Irvine, CA 92604 USA (e-mail: dspyang@live.com).

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

III. Drug Chemosensitivity Model

Previous studies of predictors of drug chemosensitivity have focused on the link between gene expression and drug chemosensitivity [e.g., 14, 17, 23], and that link is conceptually shown in Fig. 2.

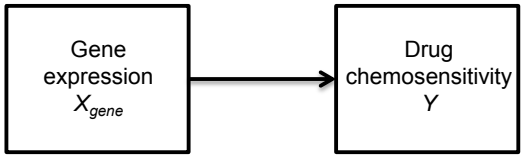

Figure 2. Baseline gene model.

The regression equation characterizing this model is:

$$Y = \beta_0 + \beta_{gene} X_{gene} \quad (1)$$

Y is drug response or chemosensitivity; X_{gene} is gene expression level.

As described above, recent efforts have started looking at the role played by miRNA in drug chemosensitivity, and studies of miRNA pharmacogenomics have examined the link between miRNA expression and drug chemosensitivity [e.g., 17, 22] utilizing the conceptual (baseline miRNA) model shown in Fig. 3.

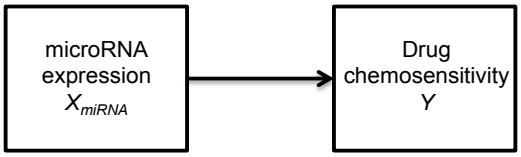

Figure 3. Baseline miRNA model.

The regression equation characterizing the baseline miRNA model is (X_{miRNA} is miRNA expression level):

$$Y = \beta_0 + \beta_{miRNA} X_{miRNA} \quad (2)$$

However, there is not likely a simple relationship between miRNA expression and drug chemosensitivity [22]. Although ultimately it is the genes that code proteins which affect drug chemosensitivity, miRNAs are now understood to play a role in regulating gene functions. Thus, there are fundamentally three variables: miRNA expression, gene expression, and drug chemosensitivity. Past studies have mostly focused on the relationship either between gene expression and chemosensitivity or between miRNA expression and chemosensitivity. As such, it is critical to integrate miRNA expression data with gene expression data to predict drug chemosensitivity, and a *multivariate* analysis in associating miRNA expression and drug resistance [7] is needed. Because "genes *affected* by a miRNA pathway are often involved in cellular response to cancer drugs" [my italics] [2, p. 3129], a "moderation" model [10] specification is proposed by which miRNA expression interacts with gene expression to exert a "moderator" effect on drug chemosensitivity. It is therefore hypothesized that, in the context of prognostic biomarkers, miRNA expression moderates the relationship between gene expression and drug chemosensitivity. To put it another way, the strength of gene expression's effect on chemosensitivity is modulated by miRNA expression. The moderation model is supported conceptually because gene expression levels can modulate drug response, and miRNAs are known to regulate gene functions.

This new moderation model specification (not previous adopted by other studies), shown in Fig. 4, has three advantages: First, it grounds statistical testing of models in the current theory of how miRNAs operate. Second, it clearly shows conceptual links and directions of influence and is not a "black-box" model. Third, it is a parsimonious model that is also multivariate in nature (which responds to the call for multivariate analysis in miRNA studies [7]).

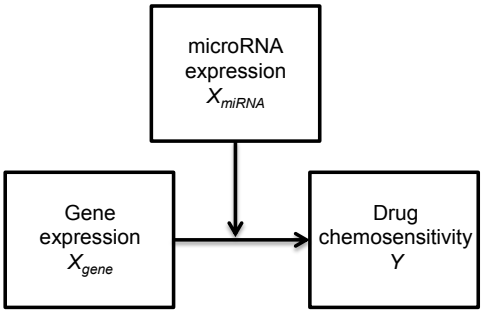

Figure 4. New moderation model.

The moderator effect is expressed by including an interaction term ($\beta_{gene \times miRNA} X_{gene} X_{miRNA}$) in the hypothesized model which can be stated as:

$$Y = \beta_0 + \beta_{miRNA} X_{miRNA} + \beta_{gene} X_{gene} + \beta_{gene \times miRNA} X_{gene} X_{miRNA} \quad (3)$$

Another interpretation of the proposed model is that because miRNA expression affects gene functions, gene expression X_{gene} and miRNA expression X_{miRNA} are not orthogonal (as implied by $Y = \beta_{gene} X_{gene} + \beta_{miRNA} X_{miRNA}$), hence an interaction term is needed. The model suggests a complex, obscure relationship in the link between chemosensitivity and miRNA and gene expressions. The proposed model is novel as a literature search on PubMed revealed no study that explicitly tests for statistical interaction effects between miRNA and gene expressions on drug chemosensitivity.

IV. Methods

Fig. 5 shows the workflow of the steps performed by this study to obtain the results. These steps are described in more detail in the following subsections.

A. Cell Line Data

This study utilizes datasets generated from the NCI-60 cell lines, used for anticancer drug screening, provided by the National Cancer Institute (NCI). NCI-60 consists of 60 cell lines from nine human cancer tissues: breast, central nervous system, colon, kidney, leukemia, lung, melanoma, ovary, and prostate and "remains the most powerful human

cancer cell panel for high throughput screening of anticancer drugs." [16, p. 1947] A study spanning different cancer types is also appropriate in the context of miRNAs because miRNAs "exhibit an ability to shift global gene expression in an appreciable manner" [20, p. 631]. Data on gene expression, miRNA expression, and drug chemosensitivity were obtained in the beginning of October, 2013, from the NCI-60 cell line database maintained by the NCI, and data were obtained in the "standardized" form of z-scores, which included data on 26,065 genes, 360 miRNAs, and 20,503 compounds. Out of these compounds, 110 are Food and Drug Administration (FDA)-approved drugs and 54 are in clinical trial.

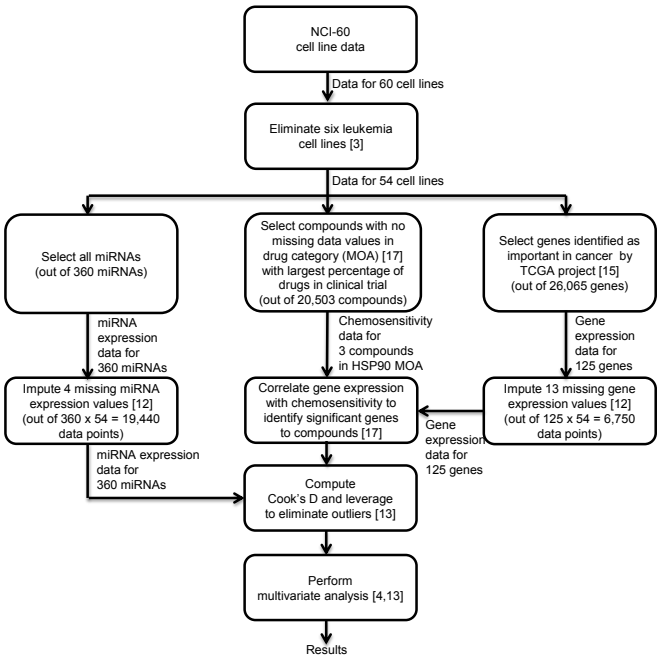

Figure 5. General workflow for assessing if miRNA expression moderates the relationship between gene expression and drug chemosensitivity.

B. Quality Controls: NCI-60

NCI-60 datasets used strictly quality control of gene expression, miRNA expression, and chemosensitivity data [19]. The accuracy and reliability of the data were improved by transforming each probe measurement into a z-score that integrated results produced by the different platforms. Gene expression data were generated using five microarray platforms and integrating all pertinent probes across these five platforms: the Affymetrix Human Genome U95 Set, the Affymetrix Human Geonome U133, the Affymetrix Human Genome U133 Plus 2.0 Arrays, and the Affymetrix GeneChip Human Exon 1.0 ST array, and the Agilent Whole Human Genome Oligo Microarray. Probe intensity values (and their z-scores) must pass additional quality control measures to be included in the datasets [19]. miRNA expression data were produced using the Agilent Technologies 15k feature Human miRNA microarray (V2) and had to pass prescribed quality control metrics; those miRNA microarrays which did not pass were not included in the miRNA data [17].

Drug chemosensitivity data were produced using the sulphorhodamine B assay by the NCI Developmental Therapeutics Program (DTP). The data are z-scores of $-10\log_{10}(GI50)$; GI50 is defined as the concentration required to inhibit (tumor) cell growth by 50% at 48 hours as determined by the DTP [17], so higher values of $-10\log_{10}(GI5)$ means higher drug response/chemosensitivity. Drug response experiments also had to pass similar quality control criteria [19].

C. Data Preprocessing: This Study

To further improve the quality of results, data preprocessing steps were implemented specific to this study and are in addition to NCI's generation of its datasets. First, the six leukemia cell lines in the NCI-60 datasets were expurgated prior to analysis. The hypersensitivity of leukemia to drugs makes data analysis difficult and harder to generalize across different cancer types. Thus, 54 non-leukemia cell lines were utilized for analysis [3], and for this study, there were effectively 54 trials carried out.

Second, this study used an objective set of criteria to eliminate outliers in the data. Specifically, statistical measures of Leverage and Cook's Distance (Cook's D) were used. *Leverage* measures how much a data point is unusual as compared to the independent variables [13], and *Cook's Distance* measures how much a data point substantially influences the regression itself [10]; those data points whose leverage values exceeded $3(p + 1)/n$ were eliminated [13] as outliers where p (= 3) is the number of predictors and n (= 54) is the number of samples. Thus the threshold value for Leverage is $3(3 + 1)/54 = 0.2222$. In addition, those data values whose Cook's D exceeded a threshold value of 1 were eliminated [13] as well. These objective measures were used to identify outliers to minimize subjective bias, and removing outliers minimizes a data point's undue impact on statistical testing.

Lastly, a form of *statistical control* is built into the multivariate structure of the drug chemosensitivity model (Fig. 4), which has standalone terms of gene expression (X_{gene}) and miRNA expression (X_{miRNA}) in addition to the interaction term ($X_{gene}X_{miRNA}$) in (3). This model structure effectively examines the relationship between the interaction term (X_{gene} X_{miRNA}) and drug chemosensitivity (Y) while controlling for (X_{gene}) and (X_{miRNA}) [13].

D. Gene and miRNA Selection

While individualization of treatment is one of oncology's goals, so is the identification of common therapeutic strategies against tumors. Doing so would narrow the search for significant predictors, miRNA included, of outcome such as drug chemosensitivity. Accordingly, out of the 26,065 genes in the NCI-60 datasets, this study emphasized on the 127 genes identified as significant across major cancer types by The Cancer Genome Atlas (TCGA) project based on well-known and emerging cellular processes in cancer [15]. However, two of those 127 genes are not available in the NCI-60 datasets, leaving 125 genes to be analyzed. Of these genes, 13 genes have a null value (gene expression data was missing), each in the central nervous system (CNS)

CNS:SF_539 cell line. Instead of deleting the CNS:SF_539 cell line, these 13 gene expression values were imputed by mean substitution [12], so there are 13 imputed values out of a total of 125 x 54 = 6,750 gene expression values.

For miRNA expression data, all 360 miRNAs for which miRNA expression data are available were utilized. Of these miRNAs, four miRNAs have a null value each (miRNA expression data was missing): miR-29b-1* for the breast (BR) BR:MDA_MB_231 cell line, miR-135a* for the melanoma (ME) ME:LOXIMVI cell line, miR-136 for the CNS:SNB_19 cell line, and miR-455-3p for the ovary (OV) OV:NCI_ADR_RES cell line. These four miRNA expression values were also imputed by mean substitution [12], so there are four imputed values out of a total of 360 x 54 = 19,440 miRNA expression values.

E. Compound Selection

Several criteria were utilized to select the compounds for analysis. First, only those compounds with known mechanism of action (MOA) from the NCI DTP were to be considered [17]. Second, only those compounds that are either FDA-approved or in clinical trial were to be considered. Third, only those drugs that have no null value (missing data) across all 54 NCI-60 cell lines were to be considered. This way, imputation of null values was avoided for the dependent variable, and sole observation of actual response of the dependent variable to changes in the independent variables was preserved. Lastly, special attention was paid to a single class of drug, relatively new, with the largest percentage of drugs in clinical trial in that MOA category (MOA category with two or fewer drugs in clinical trial were not considered). This class of drugs was HSP90 (Heat Shock Protein 90 inhibitor) in the MOA category. Three drugs with complete chemosensitivity data were available for analysis in this category: NSC#255109: Oral HSP90 Inhibitor IPI-493 (in clinical trial), NSC#330507: Tanespimycin (in clinical trial), and NSC#707545: Alvespimycin (in clinical trial). Incidentally, a search of HSP90 and miRNA on the NIH PubMed publication database as of December 2013 turned up no study on the relationship between miRNA and HSP90. This search result suggests that this study is among the first to examine miRNA as prognostic biomarkers to anticancer drugs in the HSP90 MOA category.

Heat Shock Protein 90 (HSP90) serves to regulate several cancer-associated proteins through a chaperoning activity [21]. In fact, the HSP90 molecular chaperone affects over 200 various "client proteins" [11], of which many are necessary components to cell survival and reproduction. Because a considerable number of these client proteins are oncogenic and cancers often have overexpression of HSP90 (2-10 fold higher levels in tumor cells versus normal tissue cells [18]), drugs that inhibit HSP90 function have effective anticancer properties.

F. Statistical Analysis

Similar to the approach taken by [17], genes are first included "…on the basis of high correlation to drug activities in the NCI-60" [17, p. 1087]. In this study, genes were selected for subsequent analysis based on the Bonferroni threshold [13] p-value of 0.05/125=0.0004. Then, a correlation analysis of the gene expression data (for 125 genes) against the drug chemosensitivity data (for three drugs) produced 3 x 125 = 375 Pearson correlations. The following subset of genes (expressions) showed significant correlations with chemosensitivity profiles of HSP90 MOA:

- Between ERCC2 expression and Tanespimycin chemosensitivity.
- Between SF3B1 expression and Tanespimycin chemosensitivity.
- Between ERCC2 expression and Alvespimycin chemosensitivity.

No significant correlation was detected between the drug chemosensitivity of Oral HSP90 Inhibitor IPI-493 and any one of the 125 gene expressions.

Based on these results, multiple regression analysis was performed using the moderation model shown in Fig. 4 and (3). Specifically, the dependent variable is Tanespimycin chemosensitivity, the first independent variable X_{gene} is ERCC2 expression, and the second independent (moderating) variable X_{miRNA} is miRNA expression. The "slope test" [4, p. 6] was used to test if the coefficient (i.e., slope) of the interaction term ($\beta_{gene \times miRNA} X_{gene} X_{miRNA}$) is significant. If the interaction term is significant, then the corresponding miRNA expression is deemed to be significant in moderating the relationship between gene expression and chemosensitivity.

Initially, 41 miRNAs were found to be significantly moderating the relationship between gene expression and chemosensitivity. Cook's D and leverage values were computed. Based on these values, outliers were identified and eliminated. Then multiple regression analysis was performed again without the outliers to minimize Type I error (finding a relationship when there is none) and the slope test applied. After the elimination of outliers, 12 miRNAs were found to be significantly moderating the relationship between gene expression and chemosensitivity.

The process defined in the two paragraphs above was repeated with the drug Tanespimycin and the gene SF3B1, as well as with the drug Alvespimycin and the gene ERCC2.

V. RESULTS

Certain miRNAs moderate the relationship between gene expression and drug chemosensitivity. Table I shows the results on the 12 significantly-moderating miRNAs. The β_{gene} column shows the standardized coefficients of the term X_{gene}—the direct effect from gene expression to chemosensitivity, whereas the $\beta_{gene \times miRNA}$ column shows the standardized coefficients of the term $X_{gene}X_{miRNA}$ —the effect from the interaction between miRNA and gene expressions. The p-values of these standardized coefficients are also shown. In addition, R^2 is the (%) variation in the dependent variable that can be explained by the independent variable [10]. Table I shows the unbiased R^2 [13], or *adjusted* R^2 for each moderation model.

A positive β_{gene} coefficient shows that when the expression of the gene increases, the drug activity increases and cells become more sensitive to the drug, and a negative β_{gene} coefficient shows that when the expression of the gene increases, the drug activity decreases and cells become less sensitive (or more resistant) to the drug. Applying the moderation model (Fig. 4) and including the interaction term $X_{gene}X_{miRNA}$ add another dimension to the understanding of chemosensitivity. For example, when the direct gene coefficient β_{gene} is positive but the interaction coefficient $\beta_{gene\ x\ miRNA}$ is negative, then higher levels of miRNA expression are associated with chemosensitivity being less responsive to gene expression. The interaction between miRNA and gene is examined in more detail in the next section.

TABLE I. MODERATION MODEL: SIGNIFICANCE OF INTERACTION TERMS

Direct	Interaction	β_{gene}	β_{gene} p-value	β_{miRNA}	β_{miRNA} p-value	$\beta_{gene\ x\ miRNA}$	$\beta_{gene\ x\ miRNA}$ p-value	Adj. R^2
ERCC2→Tanespimycin	ERCC2 x miR-15b*	-0.546	0.000	0.167	0.139	0.237	0.043	0.420
ERCC2→Tanespimycin	ERCC2 x miR-16-2*	-0.504	0.000	0.045	0.709	0.267	0.032	0.299
SF3B1→Tanespimycin	SF3B1 x miR-9	0.436	0.000	0.078	0.504	-0.388	0.002	0.402
SF3B1→Tanespimycin	SF3B1 x miR-126*	0.362	0.009	0.094	0.433	-0.313	0.024	0.305
SF3B1→Tanespimycin	SF3B1 x miR-129*	0.258	0.054	0.018	0.881	-0.465	0.001	0.361
SF3B1→Tanespimycin	SF3B1 x miR-138	0.454	0.001	-0.190	0.213	-0.422	0.006	0.397
SF3B1→Tanespimycin	SF3B1 x miR-519e*	0.362	0.007	0.110	0.380	-0.342	0.010	0.289
SF3B1→Tanespimycin	SF3B1 x miR-624*	0.392	0.002	0.099	0.402	-0.353	0.006	0.348
ERCC2→Alvespimycin	ERCC2 x miR-26b	-0.394	0.002	0.275	0.040	0.337	0.012	0.303
ERCC2→Alvespimycin	ERCC2 x miR-30e*	-0.420	0.001	-0.044	0.694	0.353	0.004	0.367
ERCC2→Alvespimycin	ERCC2 x miR-32	-0.434	0.000	0.200	0.061	0.365	0.001	0.494
ERCC2→Alvespimycin	ERCC2 x miR-196a	-0.417	0.003	-0.009	0.931	-0.350	0.014	0.440

$0.000 \le p \le 0.001$
$0.001 < p \le 0.01$
$0.01 < p \le 0.05$

Table I shows the following miRNAs whose expressions significantly moderate the relationship between gene expression and chemosensitivity (i.e., whose coefficient $\beta_{miRNA\ x\ gene}$ is significant):

- miR-15b* and miR-16-2* between ERCC2 expression and Tanespimycin chemosensitivity.
- miR-9, miR-126*, miR-129*, miR-138, miR-519e*, and miR-624* between SF3B1 expression and Tanespimycin chemosensitivity.
- miR-26b, miR-30e*, miR-32, and miR-196a between ERCC2 expression and Alvespimycin chemosensitivity.

The significance of moderating miRNAs is validated by regression bootstrap technique. "Bootstrap" samples can be created by sampling the 54 cell lines with replacement, effectively treating all available 54 cell lines (N=54) as a population. By repeating the above procedure 2,500 times, the procedure yields an ensemble of coefficients whose significance can then be evaluated [13]. Table II confirms that all the interaction term coefficients $\beta_{geneXmiRNA}$ are significant (last column). One advantage of nonparametric tests like bootstrapping is that they are not affected by outliers [13] as an outlier may (or may not) be just one of the data points drawn in one of the many resampling procedures. Based on bootstrapping, all 12 gene-miRNA interaction terms shown above are significant.

TABLE II. MODERATION MODEL: SIGNIFICANCE OF INTERACTION TERMS IN BOOTSTRAPPING

Direct	Interaction	β_{gene} p-value	β_{miRNA} p-value	$\beta_{gene\ x\ miRNA}$ p-value
ERCC2→Tanespimycin	ERCC2 x miR-15b*	0.000	0.162	0.000
ERCC2→Tanespimycin	ERCC2 x miR-16-2*	0.000	0.684	0.001
SF3B1→Tanespimycin	SF3B1 x miR-9	0.004	0.411	0.026
SF3B1→Tanespimycin	SF3B1 x miR-126*	0.001	0.263	0.004
SF3B1→Tanespimycin	SF3B1 x miR-129*	0.001	0.898	0.033
SF3B1→Tanespimycin	SF3B1 x miR-138	0.000	0.086	0.000
SF3B1→Tanespimycin	SF3B1 x miR-519e*	0.001	0.193	0.003
SF3B1→Tanespimycin	SF3B1 x miR-624*	0.000	0.217	0.000
ERCC2→Alvespimycin	ERCC2 x miR-26b	0.001	0.058	0.012
ERCC2→Alvespimycin	ERCC2 x miR-30e*	0.003	0.873	0.006
ERCC2→Alvespimycin	ERCC2 x miR-32	0.036	0.018	0.003
ERCC2→Alvespimycin	ERCC2 x miR-196a	0.012	0.541	0.020

$0.000 \le p \le 0.001$
$0.001 < p \le 0.01$
$0.01 < p \le 0.05$

VI. DISCUSSIONS

In terms of predictor variables of biomarkers, miRNA or gene expression has a column of data containing the profile, across 54 cell lines, of biological characteristics (miRNA or gene expression) that may interact to affect chemosensitivity of a tested drug. In terms of the outcome variable of chemosensitivity, each drug has a unique profile of chemosensitivity, across 54 cell lines, that the moderation model attempts to predict. This design is analogous to a clinical trial with 54 patients (cell lines) [24], each treated with three drugs and profiled on 125 gene expressions and 360 miRNA expressions to identify pertinent prognostic biomarkers of drug chemosensitivity response.

A. *Prognostic Biomarker Discovery: miRNA Interaction with Gene*

For Tanespimycin, two genes: ERCC2 and SF3B1 have the highest correlations with the drug's chemosensitivity. For example, SF3B1 *positively* modulates drug chemosensitivity, so cells with high SF3B1 expression are more sensitive to Tanespimycin. Applying the moderation model reveals that six miRNAs (miR-9, miR-126*, miR-129*, miR-138, miR-519e*, and miR-624*) negatively moderate the relationship between SF3B1 expression and chemosensitivity. This means that at higher expression levels of SF3B1, high miRNA expression is associated with the cancer cells being less chemosensitive while low miRNA expression is associated with the cells being more chemosensitive.

To illustrate how the miRNA expression moderates the relationship between gene expression and drug chemosensitivity, Fig. 6 shows a scatter plot of Tanespimycin chemosensitivity as a function of gene SF3B1 expression—in general positive-sloping. But, the moderation model proposed by this study shows that the positive-sloping effect itself is dependent on the expression level of miRNA miR-138. Because the interaction coefficient is negative ($\beta_{gene\ x\ miRNA}$ = -0.422), a higher miR-138 expression dampens the effect exerted by SF3B1 expression on chemosensitivity. As such, the figure clearly shows that when the miR-138 expression is low, the direct effect of gene SF3B1 expression on drug sensitivity is high (large-sloped regression line). But when the miR-138 expression is high, the direct effect of gene SF3B1 expression on drug sensitivity becomes low

(small-sloped regression line). In this case, because the interaction coefficient is significant (p-value = 0.006 < 0.05), miR-138 expression significantly moderates the relationship between SF3B1 expression and chemosensitivity.

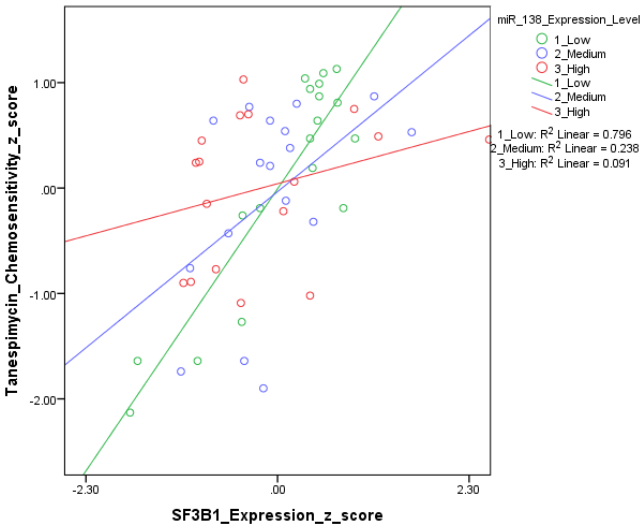

Figure 6. Scatter Plot of Tanespimycin Chemosensitivity and SF3B1 Expression with Regression Lines as functions of miR-138 Expressions.

This behavior, detected by the novel model shown in Fig. 4, suggests that there may be an underlying mechanism of chemoresistance to Tanespimycin, and that mechanism is partly associated with expression levels of the miRNAs.

B. *Evidence of Translation Control of miRNA*

In pharmacogenomics, a basic assumption is that through their levels of expressions, genes that code for drug-metabolizing enzymes, drug transporters, or drug targets can modulate drug response [20], and miRNAs can regulate gene expression by promoting mRNA degradation or inhibiting translation [8, 9, 20]. For HSP90 inhibitors in this study, the evidence points to that miRNA expression is not associated with *mRNA degradation*, at least not in a way that gene expressions are measured by the five platforms used by NCI. This is shown by the correlation values (in Table III) between miRNA expressions and gene expression. None of the correlations (R) in Table III is significant, and this lack of correlation means that gene expressions are not associated with increased or decreased levels of miRNA expressions. Because the correlation between miRNA and gene expressions is not significant but the interaction term of miRNA and gene expressions is, miRNAs most likely transmit their regulatory influence through *translation inhibition/control* (instead of mRNA degradation).

In addition, as shown in Table III, the lack of correlation between the two independent variables of miRNA expression and gene expression means that there is very little multicollinearity. Thus the moderation model results shown in Table I are robust. In multiple regressions, it is desirable to incorporate independent variables that do not correlate much with each other because of the problem of multicollinearity [10].

TABLE III. CORRELATION BETWEEN GENE AND MIRNA EXPRESSIONS

Gene	miRNA	R	R p-value	Number of Samples
ERCC2	miR-15b*	0.021	0.885	51
ERCC2	miR-16-2*	0.062	0.667	51
SF3B1	miR-9	0.079	0.579	52
SF3B1	miR-126*	-0.064	0.652	52
SF3B1	miR-129*	-0.075	0.603	51
SF3B1	miR-138	-0.252	0.071	52
SF3B1	miR-519e*	-0.173	0.226	51
SF3B1	miR-624*	-0.099	0.486	52
ERCC2	miR-26b	-0.180	0.202	52
ERCC2	miR-30e*	0.084	0.555	52
ERCC2	miR-32	-0.248	0.076	52
ERCC2	miR-196a	0.039	0.787	50

$0.000 \le p \le 0.001$
$0.001 < p \le 0.01$
$0.01 < p \le 0.05$

C. *Benefits of this Study*

There are several benefits of this study. First, these results, for the first time, support an important role of miRNAs in structurally moderating the relationship between gene expressions and chemosensitivity. The study explicitly uses an interaction term in the (moderation) model specification to more closely align current understanding of miRNA activities with the model—a novel approach that has not been adopted before in Systems Pharmacology based on a search of the PubMed publication database. Also, the moderation model proposed by this study yields insight into mechanisms by which these drugs' chemosensitivity may be modulated, and the chemosensitivity (and chemoresistance) mechanisms implicated and described above form the basis of a rich set of hypotheses that can be experimentally tested.

Second, this study addresses criticisms of current miRNA research. For example, the inconsistencies in results of prior studies on emerging miRNA signatures may be due to "different specimens (frozen vs paraffin-embedded, micro- vs non-microdissected)" and "experimental platforms used (quantitative PCR vs different miRNA array or in situ hybridization systems)" [8, p. 116]. This study utilizes NCI-60 cell lines that have been grown in consistent laboratory environments and whose expression levels were obtained consistently with high quality controls [19]. In addition, the "lack of multivariate analysis" [7, p. 1665] is addressed explicitly by this study's development of moderation model with two interacting independent variables.

D. *Limitations*

The results of this study are based on the NCI-60 cell lines grown in vitro produced by the NCI. Experiments based on in vitro cells have two disadvantages. First, the cell lines themselves may lack the cancer characteristics which are important to the hypothesis proposed by the model. Second, the conclusion is limited to those cancer cells that can be grown in vitro, and many cancer types cannot [5].

VII. CONCLUSION

This study has proposed a novel moderation model and shown that for two HSP90 inhibitors, 12 miRNA expressions significantly moderate the relationships between genes ERCC2 and SF3B1 and drug chemosensitivities. There are areas of potential applications and future research.

A. Potential Applications in Clinical Setting

Prognostic Biomarkers: A potentially important application of this study is that the identification of prognostic biomarkers (miRNA and gene expressions) can help oncologists decide on the choice of anticancer drugs. One potential utility of the significant miRNA and gene expressions identified in Table I is personalized medicine. For example, if a patient has a combination of miRNA and gene expressions that, after applying the moderation model, shows predicted chemoresistance, then alternate therapeutic regiment could be explored leading to higher chemosensitivity and beneficial response [2].

Intervention and Drug Discovery: miRNA and gene expressions identified by this study's moderation model could serve as targets on which new drugs/compounds can be developed to specifically inhibit or enhance in order to increase anticancer chemosensitivity. Instead of abandoning a specific drug, an miRNA expression could serve as an interventional agent that can be manipulated to effect a desirable level of chemosensitivity. miRNA expression levels can be increased by transfecting their precursors and decreased by transfecting their inhibitors [4]. There is precedence in targeting miRNA for therapeutic purposes. For example, mirvirsen is an inhibitor of miR-122 and is in Phase II clinical trial for treating hepatitis C [6].

B. Future Research

This study can be expanded to examine more miRNAs and more compounds. For the *miRNAs*, the Release 20 (June 2013) of miRBase database shows that there are 1,872 human miRNAs (http://www.mirbase.org/cgi-bin/mirna_summary.pl?org=hsa) in the miRNA catalog—more than five times of the number of miRNAs examined by this study. One avenue of future research is to apply the proposed methodology and moderation model (which yielded new results out of 360 miRNAs in NCI-60 datasets) to a larger population of miRNAs. For the *compounds*, the results of this study are only applicable to HSP90 inhibitors (Tanespimycin and Alvespimycin), but the NCI-60 datasets included 20,503 compounds as of early October 2013 (many of which do not have known MOAs), so the same methodology and model can be applied to more compounds. In addition to identifying pertinent miRNA and gene expressions for specific compounds' chemosensitivities, the model can serve to explore underlying mechanisms by which an miRNA expression influences chemosensitivity (e.g., translation control vs. mRNA degradation).

ACKNOWLEDGMENT

I would like to thank Mr. Antrim for mentoring my research and my parents for their support.

REFERENCES

[1] O. D. Abaan et al., "The Exomes of theNCI-60 Panel: A Genomic Resource for Cancer Biology and Systems Pharmacology," *Cancer Research*, vol. 73, pp. 4372-4380, 2013.

[2] K. E. Allen and G. J. Weiss, "Resistance May Not Be Futile: microRNA Biomarkers for Chemoresistance and Potential Therpeutics," *Molecular Cancer Therapeutics*, vol. 9, pp. 3126-3136, 2010.

[3] P. E. Blower et al., "MicroRNA Expression Profiles for the NCI-60 Cancer Cell Panel," *Molecular Cancer Therapeutics*, vol. 6, pp. 1483-1491, 2007.

[4] P. E. Blower et al., "MicroRNAs Modulate the Chemosensitivity of Tumor Cells," *Molecular Cancer Therapeutics*, vol. 7, pp. 1-9, 2008.

[5] G. Caponigro and W. R. Sellers, "Advances in the Preclinical Testing of Cancer Therapeutic Hypotheses," *Nature Reviews Drug Discovery*, vol. 10, pp. 179-187, 2011.

[6] E. Dreussi et al., "miRNA Pharamcogenomics: The New Frontier for Personalized Medicine in Cancer?" *Pharamcogenomics*, vol. 13, pp. 1635-1650, 2012.

[7] A. E. Frampton et al., "Role of miRNAs in the Response to Anticancer Therapy," *Pharamcogenomics*, vol. 13, pp. 1663-1666, 2012.

[8] E. Giovannetti et al., "Molecular Mechanisms underlying the role of microRNAs (miRNAs) in Anticancer Drug Resistance and Implications for Clinical Practice," *Critical Reviews in Oncology/Hematology*, vol. 81, pp. 103-122, 2012.

[9] W. M. Grady and M. Tewari, "The Next Thing in Prognostic Molecular Markers: miroRNA Signatures of Cancer," *Gut*, vol. 59, pp. 706-708, 2010.

[10] J. F. Hair et al., *Multivariate Data Analysis*, 5th ed. Upper Saddle River, NJ: Prentice Hall, 1998.

[11] D. S. Hong et al., "Targeting the Molecular Chaperone Heat Shock Protein 90 (HSP90): Lessons Learned and Future Directions," *Cancer Treatment Reviews*, vol. 39, pp. 375-387, 2013.

[12] D. C. Howell, "Treatment of Missing Data," 2012; http://www.uvm.edu/~dhowell/StatPages/More_Stuff/Missing_Data/Missing.html

[13] D. C. Howell, *Statistical Methods for Psychology*, 8th ed. Belmont, CA: Wadsworth, 2013.

[14] Y.-C. Hsu et al., "Genome-Wide Analysis of Three-Way Interplay among Gene Expression, Cancer Cell Invasion and Anti-Cancer Compound Sensitivity," *BMC Medicine*, vol. 11, pp. 1-14, 2013.

[15] C. Kandoth et al., "Mutational Landscape and Significance across 12 Major Cancer Types," *Nature*, vol. 502, pp. 333-339, 2013.

[16] D. Kong and T. Yamori, "JFCR39, A Panel of 39 Human Cancer Cell Lines, and its Application in the Discovery and Development of Anticancer Drugs," *Bioorganic and Medicinal Chemistry*, vol. 20, pp. 1947-1951, 2012.

[17] H. Liu et al., "mRNA and microRNA Expression Profiles of the NCI-60 Integrated with Drug Activities," *Molecular Cancer Therapeutics*, vol. 9, pp. 1080-1091, 2010.

[18] L. Neckers, "Hsp90 Inhibitors as Novel Cancer Chemotherapeutic Agents," *Trends in Molecular Medicine*, vol. 8, pp. S55-S61, 2002.

[19] W. C. Reinhold et al., "CellMiner: A Web-based Suite of Genomic and Pharmacologic Tools to Explore Transcript and Drug Patterns in the NCI-60 Cell Line Set," *Cancer Research*, vol. 72, pp. 3499-3511, 2012.

[20] N. Shomron, "MicroRNAs and Pharmacogenomics," *Pharamcogenomics*, vol. 11, pp. 629-632, 2010.

[21] J. R. Smith and P. Workman, "Targeting the Cancer Chaperone HSP90," *Drug Discovery Today: Therapeutic Strategies*, vol. 4, pp. 219-227, 2007.

[22] R. Søkilde et al., "Global microRNA Analysis of the NCI-60 Cancer Cell Panel," *Molecular Cancer Therapeutics*, vol. 10, pp. 375-384, 2011.

[23] J. E. Staunton et al., "Chemosensitivity Prediction by Transcriptional Profiling," *PNAS*, vol. 98, pp. 10787-10792, 2001.

[24] J. N. Weinstein et al., "An Information-Intensive Approach to the Molecular Pharmacology of Cancer," *Science*, vol. 275, pp. 343-349, 1997.