Published in final edited form as:

Cancer Res. 2013 March 15; 73(6): 1867–1875. doi:10.1158/0008-5472.CAN-12-0873.

MicroRNA-Related Genetic Variants Associated with Clinical Outcomes in Early Stage Non-Small Cell Lung Cancer Patients

Xia Pu¹, Jack A. Roth², Michelle A.T. Hildebrandt¹, Yuanqing Ye¹, Hua Wei¹, John D. Minna³, Scott M. Lippman⁴, and Xifeng Wu¹

¹Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, TX

²Department of Thoracic and Cardiovascular Surgery, The University of Texas MD Anderson Cancer Center, Houston, TX

³Hamon Center for Therapeutic Oncology Research, The University of Texas Southwestern Medical Center, Dallas, TX

⁴Department of Thoracic/Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX

Abstract

Given the density of single nucleotide polymorphisms (SNPs) in the human genome and the sensitivity of single nucleotide changes in microRNA (miRNA) functionality and processing, we asked whether polymorphisms within miRNA processing pathways and binding sites may influence non-small cell lung cancer (NSCLC) patients' prognosis. We genotyped 240 miRNA-related SNPs in 535 stage I and II NSCLC patients to determine associations with overall recurrence and survival, as well as effect in specific treatment subgroups. After correcting for multiple comparisons, the G allele of *FZD4*:rs713065 displayed a significant association with decreased risk of death in surgery-only patients (HR:0.46, 95%CI:0.32-0.65).

DROSHA:rs6886834 variant A allele (HR:6.38, 95%CI:2.49-16.31) remained significant for increased risk of recurrence in the overall and surgery-only populations, respectively.

FAS:rs2234978 G allele remained significantly associated with survival in all patients (HR:0.59, 95%CI:0.44-0.77), while borderline significant in subgroups (surgery only: HR:0.59, 95%CI:0.42-0.84; surgery plus chemo: HR:0.19, 95%CI:0.07-0.46). Luciferase assays demonstrated that the *FAS* SNP created a miR-651 functional binding site. Survival tree analysis was performed to classify patients into distinct risk subgroups based on their risk genotype combinations. These results indicate that miRNA-related polymorphisms may be associated with NSCLC patients' clinical outcomes through altered miRNA regulation of target genes.

Keywords

NSCLC; recurrence; overall survival; early stage; miRNA; binding site; single nucleotide polymorphism

Correspondence to Xifeng Wu, MD, PhD, Department of Epidemiology, Unit 1340, The University of Texas MD Anderson Cancer Center, 1155 Pressler Blvd., Houston, TX 77030; phone: 713-745-2485; fax: 713-792-4657; xwu@mdanderson.org..

Conflict of interest: The authors have declared that no conflict of interest exists.

INTRODUCTION

Lung cancer is the leading cause of cancer related mortality in the US(1). Most early stage non-small cell lung cancer (NSCLC) patients are treated with curative-intent therapy. However, 50% of surgically resected patients will relapse within 5 years. Thus, there is a strong need to identify reliable prognostic and predictive biomarkers to assist in developing personalized therapy and follow-up care. Germline polymorphisms are characterized by their stability and accessibility. They have been identified as potential prognostic/predictive markers for NSCLC clinical outcomes and treatment response (2).

MicroRNAs (miRNAs) are a class of small non-coding RNAs approximately 22 nucleotides in length. Emerging evidence has shown that miRNAs function as oncogenes or tumor suppressor genes depending on the context (3-5) and have been shown to be potential biomarkers for cancer risk assessment, treatment response, and prognosis (6). MiRNAs undergo a complex processing procedure to produce the mature, functional unit (7), and impaired miRNA processing has been reported to reduce stable miRNA levels and promote tumorigenesis (8). Genetic variations in several miRNA processing genes have been reported to influence risk of several cancers (9-12). In addition, variations in miRNA binding sites within 3' untranslated regions (3'UTR) of target genes have been found to contribute to different outcomes in cancer patients (13-15) which could be a result of altered miRNA-mRNA interactions followed by changes in target gene expression (14). For example, Zhang et al found that the G allele of rs1044129 in the miR-367 binding site of *RYR3* was related to poor survival in 1,125 breast cancer patients (16). Campayo et al showed that several miRNA binding site SNPs in *KRT81* were associated with time to recurrence in 175 surgically resected NSCLC patients (13).

In this study, we performed an analysis of 77 SNPs in eight miRNA processing genes and 163 SNPs in predicted miRNA binding sites for 133 cancer-related genes. We evaluated associations between these variants with overall survival and time to recurrence in early stage (I and II) NSCLC patients treated with curative therapy and also in subgroups of patients who received surgery-only or surgery plus chemotherapy. We also performed luciferase reporter assays to determine the effect of selected binding site SNPs on gene regulation.

MATERIALS AND METHODS

Study population and data collection

All the subjects included in the analysis were histologically confirmed NSCLC patients recruited at MD Anderson Cancer Center from September 1995 to February 2008, which is part of an ongoing lung cancer study initiated in 1991. Among all potential participants approached, 75% consented and were enrolled into the study (17). Blood samples were drawn from each participant. We restricted to early stage patients (stage I and II) who received curative-intent therapy (i.e. surgical resection, chemotherapy, and/or radiation therapy). The last day of follow-up for this study was December 31st, 2009. At the time of last follow-up, 38 patients were lost to follow-up and 284 patients were alive. Staging was based on AJCC staging system (version 6). A structured questionnaire was used to collect epidemiologic data during an in-person interview. Medical records were reviewed to collect clinical and follow-up information. Status of recurrence was ascertained by medical chart review. Vital status was ascertained by linking patient records to MD Anderson Tumor Registry that conducts annual follow up on all cancer patients. Deaths of patients were further confirmed by checking the social security death index. All patients signed an informed consent form and the study was approved by the Institutional Review Board of MD Anderson Cancer Center.

SNP selection and genotyping

SNPs were genotyped on a custom Illumina iSelect Infinium II genotyping platform (Illumina, San Diego, CA, USA) containing a comprehensive panel of approximately 10,000 SNPs from 998 cancer-related genes. The details for the design of this chip, including the SNP and gene selection, were described previously, duplicates were included for 2% of all samples; the concordance rates were greater than 99% (18). Eight miRNA processing genes (*DDX20*, *DGCR8*, *DICER1*, *DROSHA*, *EIF2C1*, *GEMIN4*, *RAN*, and *XPO5*) were among the genes on this chip with 77 tagging (10 kb flanking and within each gene, linkage disequilibrium $r^2 > 0.8$) and potential functional SNPs genotyped. We used the PolymiRTS v1.0 database to identify SNPs in predicted miRNA binding sites (all the 3'-UTR SNPs located within the seed region of responding miRNAs)(19) for the genes included on the chip and identified a total of 163 SNPs from 133 genes. These SNPs were selected to test the hypothesis that miRNA related genetic variations could influence NSCLC patients' clinical outcomes. All of the selected SNPs had a minor allele frequency greater than 0.01 in the Caucasian population. Genomic DNA was extracted from peripheral blood samples using the QIAamp DNA extraction kit (Qiagen, Valencia, CA). Only SNPs with sample call rate > 95% and samples with SNP call rate > 95% were included in the analysis.

Luciferase reporter assay

Luciferase reporter constructs for wildtype and variant binding site regions for *FAS*:rs2234978 and *SP1*:rs17695156 were generated. Constructs were sequenced to ensure the correct insert. NSCLC cell lines (NCI-H460 and NCI-H2444) were purchased from ATCC in 2003 and were validated for identity by short tandem repeat DNA fingerprinting by the Characterized Cell Line Core facility at MD Anderson in April 2012. Cell lines were cultured in RPMI-1640 medium (Mediatech, Manassas, VA) supplemented with 10% FBS (Invitrogen, Carlsbad, CA) in 48-well plates. Cells were transfected with 0.5 mg of each reporter construct, 5 pmol of negative control (scrambled sequence), or predicted targeting miRNAs (Sigma-Aldrich, St. Louis, MO) and 8 ng of pGL4 (Ambion, Austin, TX) Renilla luciferase reporter using Lipofectamine 2000 (Invitrogen). After 36 hours of incubation, cell lysates were harvested and measured for activity using the Dual-Luciferase Reporter Assay System (Promega, Madison, WI) on a FLUOstar Optima microplate reader (BMG Labtech, Cary, NC). Each assay was repeated independently at least twice with four replicates. Firefly luciferase activity was normalized to the Renilla luciferase activity to derive the relative luciferase activity.

Statistical analysis

Cox proportional hazard modeling was used to assess the association of SNPs with overall survival (the time between recruitment and death or last follow-up) and recurrence (time from recruitment to recurrence, progression at the local/distant site/lymph nodes or last follow-up). Patients who came to MD Anderson for treatment due to recurrence were excluded from the recurrence analysis. Hazard ratios (HRs) and 95% confident intervals (CIs) were estimated while adjusting for age, gender, ethnicity, stage, pack year of smoking, and treatment regimens. Kaplan-Meier curves and log-rank tests were used to assess effect of individual SNP on time to recurrence or overall survival. Likelihood ratio test was used to analyze the effect modifications of treatment types on SNP associations. Statistical analyses were performed using STATA software (10.1, Stata Corporation, College Station, TX). Survival tree analysis was performed using the STREE program (20) to build a decision tree using recursive partitioning method. Briefly, the root node contained all the patients, and we defined the measure for goodness of split using log rank P value to select the optimal initial split and subsequent splits of the dataset until no statistically significant split was identified (20). All test were two-sided and associations with P value < 0.05 considered significant. Multiple hypothesis testing was performed using the "q-value" package in R (21) based on a

false discovery rate (FDR) of 10%. Since 240 SNPs were tested for three models in each outcome analysis, therefore the P value was adjusted for multiple comparisons based on 720 tests. Bootstrap re-sampling was done for 500 iterations. In each resampling run, sampling with replacements was used to obtain same number of patients as the original analysis.

RESULTS

Characteristics of patients

This study included 535 early stage NSCLC patients with an overall median survival time (MST) of 90.2 months and median follow-up time of 62.1 months. At the time of analysis, 213 (40%) of the patients had died and 133 (33%) had a recurrence of their disease. The majority of the NSCLC cases were adenocarcinomas (59%). Of the 535 participants, 340 patients received surgery-only, 127 patients were treated with surgery plus neoadjuvant and/or adjuvant chemotherapy, and the remaining was treated with radiation therapy with/without surgery (Table 1).

Associations between individual SNPs and NSCLC clinical outcomes

Eleven processing and 23 binding site SNPs were significantly associated with survival. One SNP, *FAS*:rs2234978 (HR:0.59, 95% CI:0.44-0.77, $P=1.67\times10^{-4}$), remained significant after multiple comparison correction, with the GA+AA genotype resulting in a significant increase in median survival time (MST) from 59 to 118 months (log-rank $P=1.0\times10^{-4}$, Figure 1a).

Five SNPs in processing genes and 23 SNPs in binding sites were significantly associated with time to recurrence. The most significant association was *SP1*:rs17695156 (HR:2.22, 95% CI:1.44-3.41, $P=3.00\times10^{-4}$). Patients with at least one variant allele had a much shorter median recurrence-free time (MRFT) than patients who had common homozygous genotype (45.3 months vs. >270 months, log-rank $P=7.0\times10^{-4}$, Figure 2a). However, this association did not reach significance after correcting for multiple comparisons.

Effects of treatments on association of clinical outcomes

We performed subgroup analysis focusing on two groups of patients: surgery-only and surgery plus chemotherapy.

Effect on overall survival—Eighteen SNPs were significantly associated with overall survival in surgery-only patients (Supplementary Table 1). *FZD4*:rs713065 (HR:0.46, 95% CI:0.32-0.65, $P=2\times10^{-5}$) was the only SNP that remained significant after adjustment for multiple comparisons. Patients with at least one variant allele had significantly decreased risk of death and increased MST compared those patients with the common genotype (MST: 117 vs. 59 months, log-rank $P=1.05\times10^{-5}$; Figure 1b). Similar to the results from the overall analysis, for patients who received surgery plus chemotherapy *FAS*:rs2234978 displayed the most significant association with survival (Supplementary Table 1). Patients with at least one variant allele had a 81% lower risk of death (HR:0.19, 95% CI:0.07-0.46) and significantly longer MST (137 months) than patients who carried the homozygous common genotype (65 months) (log-rank $P=1.05\times10^{-4}$, Figure 1c). The association of this SNP with survival was borderline significant after correction for multiple comparisons in surgery-only patients (HR:0.59, 95% CI:0.42-0.84, $q=0.062$), with increased MST (61 months vs. 102 months, log-rank $P=4.02\times10^{-3}$, Figure 1d) in these surgery-only patients.

By comparing the findings between two subgroups, seventeen SNPs were found to have significant effects on survival in surgery-only patients but not in patients receiving surgery plus chemotherapy. In contrast, 28 SNPs were significantly associated with risk of death

only in patients receiving surgery plus chemotherapy. Intriguingly, within each cluster of SNPs that were significantly associated with outcomes in each treatment subgroup, we identified SNPs with differential effects. A group of 29 SNPs has the same trend in both treatment subgroups (either associated with increased or decreased risk), while 15 SNPs conferred opposite effects (Supplementary Table 1). For example, *FZD4*:rs713065 was associated with significantly decreased risk of death and prolonged MST in surgery-only patients; however, in the surgery plus chemotherapy subgroup, this SNP was associated with increased risk of death and a shortened MST (P for interaction=0.004, $q=0.030$, Figure 1b and 1e).

Effect on recurrence—*SP1*:rs17695156, which was the top SNP associated with recurrence in the overall population, was the most significant SNP in surgery plus chemotherapy group (HR:3.36, 95% CI:1.62-6.69, $P=1.10\times 10^{-3}$) (Supplementary Table 2). In surgery only patients, one miRNA processing SNP, *DROSHA*:rs6886834, remained significant after multiple comparisons corrections; it was significantly associated with more than a 6-fold increased risk for recurrence in surgery-only patients (HR:6.38, 95% CI: 2.49-16.31, $P=1.10\times 10^{-4}$) (Table 2). Patients who carried at least one variant allele of this SNP had significant reduction in MFST compared with patients with common genotype (23 vs. >270 months, log-rank $P=5.0\times 10^{-4}$; Figure 2b).

When comparing results of subgroup analysis, 28 SNPs and 16 SNPs were exclusively associated with altered risk for recurrence in surgery-only patients or surgery plus chemotherapy patients, respectively. Of these, 17 SNPs were found to have opposite effects in both subgroups (Supplementary Table 2). For example, in patients receiving surgery plus chemotherapy, *RRM2B*:rs5005121 was associated with significantly increased risk of recurrence and a shortened MFST, but in surgery-only patients, this SNP was associated with decreased risk for recurrence with an increased MFST (P for interaction=0.015, $q=0.170$, Figure 2c and 2d), however, the interaction between this SNP with treatments was not significant after multiple comparison corrections.

Survival tree analysis

Figure 3 shows the survival-tree structure classifying patients into subgroups with distinct risk of dying based on their risk genotype combinations. SNPs that displayed at least borderline significant association with survival in the main effect analysis after multiple comparisons ($q<0.15$) were included in the analysis, and none of these SNPs were in high LD. The MSTs based on these groupings varied from > 86 months for the low risk group to 41.7 months for the high risk group in surgery-only patients, and from > 118 months to 36.8 months for the low and high risk groups, respectively in patients receiving surgery plus chemotherapy. Moreover, the initial splits in the tree structure for each subgroup, *FZD4*:rs713065 and *FAS*:2234978, were also the two SNPs that remained significant after multiple comparisons in the treatment subgroup analyses. Survival tree analysis was not performed for recurrence due to limited number of SNPs with $q<0.15$.

Bootstrap re-sampling analysis

All the SNPs that were significant after multiple comparisons at an FDR of 5% remained significant in the bootstrap analysis for at least 450 out of 500 re-samplings, providing internal validation to these results. Bootstrap re-sampling analysis was also performed for survival tree analyses, and the results were significant in both the subgroups analysis for entire 500 re-samplings at $P<0.05$ (Supplementary Figure 1).

The effect of selected miRNA binding site variants on miRNA-regulation

Luciferase reporter assays were performed to determine the effect of these predicted binding site variants on miRNA regulation of target genes. *FAS*:rs2234978, which was consistently associated with a favorable prognosis, was predicted to create a new miRNA binding site for miR-561. In two lung cancer cell lines, H460 and H2444, a significant reduction in luciferase activity was observed when miR-561 was transfected with the variant reporter (T) (H460: $P=0.0294$; H2444: $P=0.025$), but not with wildtype allele (C) construct ($P>0.5$ for both cell lines), when compared to the scrambled sequence control transfections. Furthermore, there was a significant difference in luciferase activities between the variant and the wildtype constructs when co-transfected with miR-561 (H460: $P=0.0015$; H2444: $P=0.0040$) (Figure 4)

SPI:rs17695156 was predicted to disrupt a conserved miR-545 binding site; however, in our *in vitro* assays suppression of luciferase activity was observed in both variant and wildtype constructs co-transfected with miR-545. There was no significant difference in reporter activities between the two alleles and the extent of signal decrease varied between cell lines (data not shown).

DISCUSSION

In this study, we identified genetic variants in miRNA processing genes and miRNA binding sites near cancer-related genes that were associated with overall survival and recurrence in early stage NSCLC patients. *FAS*:rs2234978 was identified as a potential prognostic factor in our results and functional data provided evidence that this SNP alters miRNA regulation of *FAS*. We also found evidence that some SNPs exhibited associations that were treatment-specific. These results suggest that genetic variants in miRNA processing genes and miRNA binding sites may serve as potential prognostic markers for survival and predictive markers of response to treatment.

The most significant SNP associated with survival was *FAS*:rs2234978, which was consistent regardless of treatment regimens. *FAS* is a cell-surface receptor of the tumor necrosis family which plays an important role in the regulation of apoptosis. Evidence has shown *FAS* expression and polymorphisms could influence lung cancer patients' prognosis (22, 23). Rs2234978 is a synonymous SNP located in the seventh exon of *FAS*. MiRNA binding sites for *FAS* are located in exon 7 instead of the typical 3'UTR. Alternative splicing of *FAS* results in several transcribed isoforms that are involved in nonsense-mediated mRNA decay (NMD), including a transcript where exon 7 serves as the 3'UTR. NMD plays important roles in limiting the synthesis of truncated or mutant proteins which can negatively regulate apoptosis mediated by the full length protein. This SNP is predicted to create a new miRNA binding site for miR-561, which was supported *in vitro* by our luciferase assay, suggesting in decreased expression of *FAS* alternative transcripts. Since the NMD transcripts may negatively regulate normal *FAS* expression, this would ultimately result in increased level of *FAS* in tissues that express the targeting miRNA. It has also been reported that cisplatin treatment can increase *FAS*-mediated apoptosis (24). It is possible that in patients who carry the variant allele, higher expression of *FAS* could increase tumor cell death resulting in better overall survival independent of treatment regimen. This locus might even be synergistic with chemotherapy agents such as cisplatin, thus conferring a more extreme effect on patients' survival. Further studies are needed to confirm whether this SNP has any influence on *FAS* protein level and apoptotic activity *in vivo*.

FZD4:rs713065 is the only SNP associated with significant decreased risk of death after multiple comparison correction in surgery-only patients. *FZD4* (frizzled homolog 4) is a member of the frizzled gene family of transmembrane receptors, which help to transduce

WNT signals and activate downstream WNT/beta-catenin pathway components in cancer stem cell homeostasis(25). This SNP may down-regulate *FZD4* expression by creating a miRNA binding site, thereby inhibiting transduction of the WNT signal, leading to enhanced survival through decreased WNT signaling. However, due to difficult sequence characteristics of this region, luciferase assays were not possible.

SP1:rs17695156 is the most significant SNP associated with increased risk for death. SP1 is a transcription factor known to regulate expression of many genes, thus having a general regulatory role within the cell. This SNP is predicted to disrupt a conserved miRNA site; however, in our *in vitro* experiments, we did not observe any significant difference between the two alleles in miRNA-induced repression of reporter activity. It is possible that this 3'UTR SNP might affect *SP1* expression independent of its putative role as a miRNA target site (e.g. affecting RNA stability or post-transcriptional regulation) and influence cellular components at physiological levels. And it is also possible that rs17695156 is tagging some other miRNA binding site polymorphism that affects transcription but has not yet been identified by binding-site prediction algorithms, or regulated by miRNAs that have not yet been discovered.

In this study, we identified panels of SNPs from the miRNA processing pathway and miRNA binding site SNPs in major cancer-related pathways exclusively associated with clinical outcomes in either of the two treatment subgroups. Although several GWAS have been published for NSCLC survival (26-31), the SNPs identified in our study were not covered, indicating there is still a need for complementary pathway-based analysis to identify novel loci associated with clinical outcomes. Most microRNA-related variants are not well covered by current GWAS platforms, meaning that it is very likely that our identified potential functional SNPs would not have been identified through GWAS. In addition, the previous studies were often limited to advanced stage NSCLC and relatively small patient populations. In this study, the adoption of a pathway-based approach in a large, well-characterized population allowed for the discovery of novel loci of interest and enhances our understanding of the genetic mediators of clinical outcomes in NSCLC. Our results provide supportive evidence that genetic variations could potentially interact with treatment to influence patients' clinical outcomes, especially survival, therefore highlighting the necessity of personalized treatment decisions based on patients' genetic background. Meanwhile, the tree structure identified could potentially assist the decision by classifying patients into different risk groups in a more intuitive manner. Due to the relatively favorable prognosis and long survival time, around 30-40% of early stage NSCLC patients eventually do not die due to lung cancer (32, 33). In this study, because we focused on overall survival of early stage NSCLC patients, additional investigation will be necessary to further elucidate whether or not these associations are specific to lung cancer clinical outcomes and the mechanism of the identified associations.

The strengths of the current study include the comprehensive query of SNPs from genes involved in cancer-related miRNA regulation and the evidence for biological plausibility provided by *in vitro* functional assays. Although the luciferase assay, by design, does not prove that altered miRNA binding results in changes in host gene expression or protein level, it provides evidence supporting the effect of this SNP on the function of this specific miRNA binding sites, suggesting a downstream effect on gene expression and protein levels. The analysis also took into account the effect of treatment and adopted FDR and Bootstrap resampling methods to exclude potential false positive results.

Overall, the current study provides evidence that genetic variants in the miRNA processing pathway and miRNA binding sites influence clinical outcomes for early stage NSCLC patients. Specifically, we identified the potential prognostic role of a *FAS* SNP in predicting

overall survival in these patients and supported this observation with *in vitro* functional analyses. Following validation in an independent population, our results could provide a basis for future personalized medicine whereby those early stage NSCLC patients with high probability for favorable outcomes can be identified and treated with optimal regimens.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding: This research was supported in part, by National Cancer Institute (R01 CA111646, P50 CA070907, and R01 CA055769).

REFERENCES

1. Jemal A, Siegel R, Xu J, Ward E. Cancer statistics. *CA Cancer J Clin.* 2010; 60:277–300. [PubMed: 20610543]
2. Hildebrandt MA, Gu J, Wu X. Pharmacogenomics of platinum-based chemotherapy in NSCLC. *Expert Opin Drug Metab Toxicol.* 2009; 5:745–55. [PubMed: 19442035]
3. Deng S, Calin GA, Croce CM, Coukos G, Zhang L. Mechanisms of microRNA deregulation in human cancer. *Cell Cycle.* 2008; 7:2643–6. [PubMed: 18719391]
4. Hammond SM. RNAi, microRNAs, and human disease. *Cancer Chemother Pharmacol.* 2006; 58(Suppl 1):s63–8. [PubMed: 17093929]
5. Croce CM. Oncogenes and cancer. *N Engl J Med.* 2008; 358:502–11. [PubMed: 18234754]
6. Yanaihara N, Caplen N, Bowman E, Seike M, Kumamoto K, Yi M, et al. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell.* 2006; 9:189–98. [PubMed: 16530703]
7. Gregory RI, Shiekhattar R. MicroRNA biogenesis and cancer. *Cancer Res.* 2005; 65:3509–12. [PubMed: 15867338]
8. Calin GA, Croce CM. MicroRNA signatures in human cancers. *Nat Rev Cancer.* 2006; 6:857–66. [PubMed: 17060945]
9. Yang H, Dinney CP, Ye Y, Zhu Y, Grossman HB, Wu X. Evaluation of genetic variants in microRNA-related genes and risk of bladder cancer. *Cancer Res.* 2008; 68:2530–7. [PubMed: 18381463]
10. Ye Y, Wang KK, Gu J, Yang H, Lin J, Ajani JA, et al. Genetic variations in microRNA-related genes are novel susceptibility loci for esophageal cancer risk. *Cancer Prev Res (Phila Pa).* 2008; 1:460–9.
11. Horikawa Y, Wood CG, Yang H, Zhao H, Ye Y, Gu J, et al. Single nucleotide polymorphisms of microRNA machinery genes modify the risk of renal cell carcinoma. *Clin Cancer Res.* 2008; 14:7956–62. [PubMed: 19047128]
12. Liang D, Meyer L, Chang DW, Lin J, Pu X, Ye Y, et al. Genetic variants in MicroRNA biosynthesis pathways and binding sites modify ovarian cancer risk, survival, and treatment response. *Cancer Res.* 70:9765–76. [PubMed: 21118967]
13. Campayo M, Navarro A, Vinolas N, Tejero R, Munoz C, Diaz T, et al. A dual role for KRT81: a miR-SNP associated with recurrence in non-small-cell lung cancer and a novel marker of squamous cell lung carcinoma. *PLoS One.* 2011; 6:e22509. [PubMed: 21799879]
14. Chin LJ, Ratner E, Leng S, Zhai R, Nallur S, Babar I, et al. A SNP in a let-7 microRNA complementary site in the KRAS 3' untranslated region increases non-small cell lung cancer risk. *Cancer Res.* 2008; 68:8535–40. [PubMed: 18922928]
15. Teo MT, Landi D, Taylor CF, Elliott F, Vaslin L, Cox DG, et al. The role of microRNA-binding site polymorphisms in DNA repair genes as risk factors for bladder cancer and breast cancer and their impact on radiotherapy outcomes. *Carcinogenesis.* 2012; 33:581–6. [PubMed: 22166496]

16. Zhang L, Liu Y, Song F, Zheng H, Hu L, Lu H, et al. Functional SNP in the microRNA-367 binding site in the 3'UTR of the calcium channel ryanodine receptor gene 3 (RYR3) affects breast cancer risk and calcification. *Proc Natl Acad Sci U S A*. 2011; 108:13653–8. [PubMed: 21810988]
17. Spitz MR, Hong WK, Amos CI, Wu X, Schabath MB, Dong Q, et al. A risk model for prediction of lung cancer. *J Natl Cancer Inst*. 2007; 99:715–26. [PubMed: 17470739]
18. Wu X, Spitz MR, Lee JJ, Lippman SM, Ye Y, Yang H, et al. Novel susceptibility loci for second primary tumors/recurrence in head and neck cancer patients: large-scale evaluation of genetic variants. *Cancer Prev Res (Phila Pa)*. 2009; 2:617–24.
19. Bao L, Zhou M, Wu L, Lu L, Goldowitz D, Williams RW, et al. PolymiRTS Database: linking polymorphisms in microRNA target sites with complex traits. *Nucleic Acids Res*. 2007; 35:D51–4. [PubMed: 17099235]
20. Zhang, H.; Singer, BH. *Recursive Partitioning and Applications*. 2nd ed. 2010.
21. Storey JD, Tibshirani R. Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A*. 2003; 100:9440–5. [PubMed: 12883005]
22. Uramoto H, Osaki T, Inoue M, Taga S, Takenoyama M, Hanagiri T, et al. Fas expression in non-small cell lung cancer: its prognostic effect in completely resected stage III patients. *Eur J Cancer*. 1999; 35:1462–5. [PubMed: 10673978]
23. Park JY, Lee WK, Jung DK, Choi JE, Park TI, Lee EB, et al. Polymorphisms in the FAS and FASL genes and survival of early stage non-small cell lung cancer. *Clin Cancer Res*. 2009; 15:1794–800. [PubMed: 19240174]
24. Huang CR, Jin ZX, Dong L, Tong XP, Yue S, Kawanami T, et al. Cisplatin augments FAS-mediated apoptosis through lipid rafts. *Anticancer Res*. 30:2065–71. [PubMed: 20651352]
25. Katoh M. WNT signaling in stem cell biology and regenerative medicine. *Curr Drug Targets*. 2008; 9:565–70. [PubMed: 18673242]
26. Hu L, Wu C, Zhao X, Heist R, Su L, Zhao Y, et al. Genome-wide association study of prognosis in advanced non-small cell lung cancer patients receiving platinum-based chemotherapy. *Clin Cancer Res*. 2012; 18:5507–14. [PubMed: 22872573]
27. Huang YT, Heist RS, Chirieac LR, Lin X, Skaug V, Zienolddiny S, et al. Genome-wide analysis of survival in early-stage non-small-cell lung cancer. *J Clin Oncol*. 2009; 27:2660–7. [PubMed: 19414679]
28. Niu NMDPD, Schaid DJPD, Abo RPPD, Kalari KPD, Fridley BLPD, Feng QPD, et al. Genetic association with overall survival of taxane-treated lung cancer patients - A genome-wide association study in human lymphoblastoid cell lines followed by a clinical association study. *BMC Cancer*. 2012; 12:422. [PubMed: 23006423]
29. Sato Y, Yamamoto N, Kunitoh H, Ohe Y, Minami H, Laird NM, et al. Genome-Wide Association Study on Overall Survival of Advanced Non-small Cell Lung Cancer Patients Treated with Carboplatin and Paclitaxel. *J Thorac Oncol*.
30. Tan XL, Moyer AM, Fridley BL, Schaid D, Niu N, Batzler A, et al. Genetic variation predicting cisplatin cytotoxicity associated with overall survival in lung cancer patients receiving platinum-based chemotherapy. *Clin Cancer Res*. 2011
31. Wu X, Ye Y, Rosell R, Amos CI, Stewart DJ, Hildebrandt MA, et al. Genome-wide association study of survival in non-small cell lung cancer patients receiving platinum-based chemotherapy. *J Natl Cancer Inst*. 2011; 103:817–25. [PubMed: 21483023]
32. Reed MF, Molloy M, Dalton EL, Howington JA. Survival after resection for lung cancer is the outcome that matters. *The American Journal of Surgery*. 2004; 188:598–602.
33. Butts CA, Ding K, Seymour L, Twumasi-Ankrah P, Graham B, Gandara D, et al. Randomized Phase III Trial of Vinorelbine Plus Cisplatin Compared With Observation in Completely Resected Stage IB and II Non-Small-Cell Lung Cancer: Updated Survival Analysis of JBR-10. *Journal of Clinical Oncology*. 2010; 28:29–34. [PubMed: 19933915]

Figure 1.

Kaplan-Meier estimates of selected SNP on overall survival in early stage NSCLC patients treated with curative intended therapy recruited from 1995 to 2008 at MD Anderson Cancer Center: (a) *FAS*:rs2234978 among total population; (b) *FZD4*:rs713065 among surgery-only patients; (c) *FAS*:rs2234978 among surgery plus chemotherapy patients; (d) *FAS*:rs2234978 among surgery-only patients; (e) *FZD4*:rs713065 among surgery plus chemotherapy patients. MST: median survival time in months. N=A/B, A: number of patients with event, B: total number of patients.

Figure 2.

Kaplan-Meier estimates on time to recurrence in early stage NSCLC patients treated with curative intended therapy recruited from 1995 to 2008 at MD Anderson Cancer Center: (a) *SPI*:rs17695156 among total population; (b) *DROSHA*:rs6886834 among surgery-only patients; (c) *RRM2B*:rs5005121 among surgery-only patients; (d) *RRM2B*:rs5005121 among surgery plus chemotherapy patients. MRFT: median recurrence-free time in months. N=A/B, A: number of patients with event, B: number of patients in subgroup.

Figure 3.

Survival tree analysis based on risk genotype combinations (a) survival tree structure classifying patients into risk groups defined by percentage of patients with event : low risk: <33%; median risk: 33-67%; high risk: >67% (b) Kaplan-Meier curves of survival time for surgery-only patients in three risk groups identified by the survival tree analysis, SNPs analyzed included (rs713065, rs1558496, rs197412, rs2234978, rs3790611, and rs854552) (c) Kaplan-Meier curves of survival time for surgery plus chemotherapy patients among the three risk groups, SNPs analyzed included (rs2234978, rs1047312, rs669702, and rs1133043). MST: median survival time in months. N=A/B, A: number of patients with event, B: total number of patients.

Figure 4.

Effect of the *FAS* variant allele on miR-561 targeting and luciferase reporter expression: (a) Relative luciferase reporter activity of the wildtype and variant *FAS* allele in the presence of control (Ctrl) or miR-561 in lung cancer cell line NCI-H460; (b) Relative luciferase reporter activity of the wildtype and variant *FAS* allele in the presence of control (Ctrl) or miR-561 in lung cancer cell line NCI-H2444.

Table 1

Host characteristics for NSCLC patients recruited from 1995 to 2008 at MD Anderson included in this study

Characteristics	Surgery-only N=340	Surgery & chemotherapy N=127	Radiotherapy, +/- surgery N=74
Age, mean (sd)	65.8 (9.9)	62.9 (10.2)	69.6 (10.2)
Gender			
<i>Male</i>	166 (49)	68 (54)	32 (43)
<i>Female</i>	174 (51)	59 (46)	42 (57)
Ethnicity			
<i>Caucasian</i>	305 (90)	109 (86)	61 (82)
<i>African-American</i>	25 (7)	10 (8)	7 (9)
<i>Others</i>	10 (3)	8 (6)	6 (8)
Pack-year smoking, mean (sd)	45.1 (37.6)	39.4 (35.2)	54 (36)
Histology			
<i>Adenocarcinoma</i>	213 (63)	74 (58)	39 (39)
<i>Squamous cell carcinoma</i>	87 (25)	34 (27)	32 (43)
<i>Unclassified or other</i>	40 (12)	19 (15)	13 (18)
Clinical stage			
<i>Stage IA</i>	181 (53)	23 (18)	41 (55)
<i>Stage IB</i>	113 (33)	55 (43)	22 (30)
<i>Stage IIA</i>	10 (3)	14 (11)	2 (3)
<i>Stage IIB</i>	36 (11)	35 (28)	9 (12)
Surgery-only	340 (100)	0	0
Surgery & adjuvant chemo	0	127 (100)	6 (8)
Surgery & radiotherapy	0	0	21 (28)
Treatment without surgery	0	0	53 (72)
Vital status			
<i>Alive</i>	210 (62)	88 (69)	27 (36)
<i>Dead</i>	130 (38)	39 (31)	47 (64)
Recurrence			
<i>No</i>	233 (69)	85 (67)	46 (62)
<i>Yes</i>	107 (31)	42 (33)	28 (38)

Table 2

Significant SNPs after multiple comparisons correction

Outcomes	Gene SNP	MiRNA	Model	HR (95%CI) *	P-value *	HR (95%CI) #	P-value#	Q-value ^
Survival, overall	<i>FAS</i> rs2234978	miR-561	DOM	0.58 (0.44-0.76)	8×10 ⁻⁵	0.59 (0.45-0.78)	2×10 ⁻⁴	0.062
Survival, surgery only	<i>FZD4</i> rs713065	miR-494, 302a *	DOM	0.48 (0.33-0.66)	2×10 ⁻⁵	0.47 (0.33-0.68)	2×10 ⁻⁵	0.009
Progression, surgery only	<i>DROSHA</i> rs6886834	processing	REC	4.00 (1.72-9.29)	1×10 ⁻³	6.38 (2.49-16.10)	1×10 ⁻⁴	0.021

* Un-adjusted

Adjusted by age, gender, ethnicity, stage, pack year and treatment regimens

^ based on 720 tests