ADH1B is differentially expressed in non-small cell lung cancer and associates with patient survival.

Shahan Mamoor¹

¹shahanmamoor@gmail.com

East Islip, NY USA

Non-small cell lung cancer (NSCLC) is the leading cause of cancer death in the United States¹. We mined published microarray data^{2,3,4} to identify differentially expressed genes in NSCLC. We found that ADH1B was among the genes whose expression was most quantitatively different in tumors from patients with NSCLC as compared to the lung. ADH1B expression was significantly decreased in NSCLC tumors as compared to the lung, and lower expression of ADH1B in patient tumors was significantly associated with worse overall survival. ADH1B may be important for initiation or progression of non-small cell lung cancer in humans.

Keywords: ADH1B, NSCLC, non-small cell lung cancer, systems biology of NSCLC, targeted therapeutics in NSCLC.

1 In 2016, lung cancer resulted in the death of 158,000 Americans; 81% of all patients
2 diagnosed with lung cancer will expire within 5 years⁵. Non-small cell lung cancer (NSCLC) is
3 the most common type of lung cancer, diagnosed in 84% of patients with lung cancer, and 76%
4 of all patients with NSCLC will expire within 5 years⁵. The rational development of targeted
5 therapeutics to treat patients with NSCLC can be supported by an enhanced understanding of
6 fundamental transcriptional features of NSCLC tumors. To discover genes associated with
7 NSCLC tumors in an unbiased fashion and at the systems-level, we mined independently
8 published microarray data^{2,3,4} to compare global gene expression profiles of NSCLC tumors to
9 that of the normal lung. We found recurrent and significant differential expression of ADH1B in
10 adenocarcinoma tumors from patients with NSCLC, suggesting ADH1B may be important for
11 NSCLC tumor initiation or progression.

12 **Methods**

13 We utilized microarray datasets GSE33532², GSE43458³ and GSE74706⁴ for this
14 differential gene expression analysis of NSCLC tumors in conjunction with GEO2R. GSE33532
15 was generated using Affymetrix Human Genome U133 Plus 2.0 Array technology; for this
16 analysis, we used $n=20$ control lung tissue and $n=10$ NSCLC tumors, and the analysis was
17 performed using platform GPL570. GSE43458 was generated using Affymetrix Human Gene
18 1.0 ST Array technology; for this analysis, we used with $n=30$ control lung tissue and $n=80$
19 NSCLC tumors, and the analysis was performed using platform GPL6244. GSE74706 was
20 generated using Agilent-026652 Whole Human Genome Microarray 4x44K v2 technology; for

1 this analysis, we used $n=18$ control lung tissue and $n=10$ NSCLC tumors, and the analysis was
2 performed using platform GPL13497. All tumors utilized for differential gene expression
3 analysis here were of the adenocarcinoma type.

4
5 The Benjamini and Hochberg method of p -value adjustment was used for ranking of
6 differential expression but raw p -values were used to assess statistical significance of global
7 differential expression. Log-transformation of data was auto-detected, and the NCBI
8 generated category of platform annotation was used. A statistical test was performed to evaluate
9 whether ADH1B expression was significantly between normal lung tissue and NSCLC tumors
10 using a two-tailed, unpaired t-test with Welch's correction. We used PRISM for all statistical
11 analyses of differential gene expression in NSCLC tumors (Version 8.4.0)(455). For Kaplan-
12 Meier survival analysis, we used the Kaplan-Meier plotter online tool⁶ for correlation of ADH1B
13 mRNA expression levels with overall survival in $n=1925$ non-small cell lung cancer patients.
14
15
16
17
18

19 **Results**

20 We harnessed the power of multiple, independently published microarray datasets^{2,3,4} to
21 discover in an unbiased fashion and at the transcriptome-level the most striking gene expression
22 features of NSCLC tumors.
23
24

25 ADH1B is differentially expressed in non-small cell lung cancers.

26 We found significant differential expression of ADH1B in NSCLC tumors when
27 compared to the lung² (Table 1). When sorting each of the transcripts measured based on
28

significance of difference in expression of ADH1B between NSCLC tumors and the normal lung, ADH1B ranked 59 out of 25906 total transcripts (Table 1). Differential expression of ADH1B in NSCLC tumors was statistically significant (Table 1; $p=7.04E-16$).

We queried a second microarray dataset³ to determine if we could validate differential expression of ADH1B in non-small cell lung cancers (Table 2). We again found significant differential expression of ADH1B in NSCLC tumors of the adenocarcinoma type when compared to the normal lung (Table 2). When sorting each of the transcripts measured based on significance of difference in expression of ADH1B between NSCLC tumors and the normal lung, ADH1B ranked 665 out of 33252 total transcripts (Table 2). Differential expression of ADH1B in NSCLC tumors was statistically significant (Table 2; $p=1.3E-14$).

Analysis of a third microarray dataset⁴ revealed significant differential expression of ADH1B in NSCLC tumors of the adenocarcinoma type. When sorting each of the transcripts measured based on significance of difference in expression of ADH1B between NSCLC tumors and the normal lung, in this dataset, ADH1B ranked 165 out of 34183 total transcripts (Table 3). Differential expression of ADH1B in NSCLC tumors was statistically significant (Table 3; $p=9.55E-12$).

ADH1B is expressed at significantly lower levels in NSCLC tumors as compared to the lung.

We obtained exact mRNA levels for ADH1B from NSCLC tumors and from the lung to directly compare ADH1B expression between tumor and control lung tissue and assess for statistical significance. ADH1B was expressed at significantly lower levels in NSCLC tumors as

1 compared to the normal lung in each dataset queried (Figure 1: $p < 0.0001$, Figure 2: $p < 0.0001$).
2 We calculated a mean fold change of 0.8378 ± 0.0235 in ADH1B expression when comparing
3 NSCLC tumors to the lung, in one dataset (Figure 1; Table 1).
4

6 ADH1B expression in NSCLC tumors correlates with overall survival.

7 We performed Kaplan-Meier survival analysis using ADH1B mRNA expression in
8 NSCLC tumors coupled with paired overall survival data from each patient, in 1925 NSCLC
9 patients in total, to determine whether ADH1B tumor expression was correlated with survival
10 outcomes in NSCLC. We found that patients whose tumors expressed lower levels of ADH1B
11 possessed significantly shorter overall survival than patients with high tumor expression of
12 ADH1B (Figure 4). Median overall survival (OS) of patients in the low expression cohort was
13 52 months, while median OS in patients in the high ADH1B expression cohort was 117.33
14 months (Table 4); this difference in median OS based on ADH1B tumor expression in NSCLC
15 was statistically significant (Figure 4; logrank p -value: $1.2e-16$; hazard ratio: 0.54 (0.46 - 0.63);
16 false discovery rate=0.01).
17
18
19
20

21 Thus, blind comparative transcriptome analysis of non-small cell lung cancers revealed
22 differential expression of ADH1B as among the most significant transcriptional features of
23 NSCLC tumors, and ADH1B expression was significantly correlated with patient outcomes, as
24 patients with lower tumor expression of ADH1B possessed significantly worse overall survival.
25
26
27
28

Discussion

ADH1B belongs to a family of enzymes known as the class I alcohol dehydrogenases⁷. The protein product of the ADH1B gene encodes for the β subunit of alcohol dehydrogenase⁸. ADH1B is expressed most abundantly in the liver, breast, adipose tissues, esophageal mucosa, and stomach⁷. Alcohol dehydrogenase functions in first-pass metabolism of ethanol, and is the major enzyme that functions in ethanol oxidation: this can occur in the liver or in the gastric mucosa⁹. Importantly, the product of ethanol metabolism by ADH1B is acetyldehyde¹⁰.

ADH1B has been implicated in a variety of human cancers, primarily in the context of genetic association but also more mechanistically in supporting tumor function. In a Japanese population, genetic variation in the ADH1B gene, specifically the single nucleotide polymorphism (SNP) rs1229984, can interact with environmental factors like smoking and alcohol consumption to confer increased risk of development of esophageal squamous cell carcinoma¹¹. In a Chinese population the same SNP, rs1229984 also displayed signals of significant association between ADH1B genetic variation and risk of esophageal cancer, with an odds ratio of 1.34 for carriers of the G-allele¹². In a cohort of Han Chinese males, variation in ADH1B was significant associated with overall survival in laryngeal cancer; specifically, the GA genotype of the same SNP, rs1229984, was associated with greater median survival time¹³. In the Netherlands Cohort Study of patients with colorectal cancer, male carriers of SNP rs4147536 possessed significantly increased risk of distal colon cancer risk; however, SNP rs3811802 was associated with decreased colon cancer risk in heavy drinkers (greater than 30g/day) and stable

1 drinkers, those whose use had remained stable as compared to 5 years prior¹⁴. In women, the
2 possession of either of these SNPs (rs4147536 or rs3811802) was correlated with modification
3 of association between alcohol use and risk of development of colon cancer¹⁴. Interestingly,
4 genetic modification of ADH1B was found in 1 case (3%) of ALK-fusion positive non-small cell
5 lung cancer¹⁵. A study analyzing a novel method of identifying differentially expressed genes
6 termed Cross-Value Association Analysis found ADH1B as the single most differentially
7 expressed gene as measured by “dysregulation frequency” with decreased expression in tumor
8 tissue as compared to normal tissue in nearly every cancer type studied¹⁶. A study identifying
9 differentially expressed genes (DEGs) in 131 tumors from patients with non-small cell lung
10 cancer (NSCLC) found, consistent with our analysis, that ADH1B was among the most
11 differentially expressed genes in NSCLC, among the top ten DEGs whose expression was
12 decreased¹⁷. In a mouse model of ovarian cancer using direct injection of A2780 ovarian cancer
13 cells into the ovaries of nude mice, ectopically expressing ADH1B or not, tumor weight of the
14 primary tumor was significantly increased, and in most but not all areas, metastasis was also
15 increased, as measured by number of nodules¹⁸. Conditioned media from A2780 cells was
16 significantly more effective at degradation in a gelatin degradation assay, indicating that ovarian
17 cancer cells might have increased capacity at degradation of extracellular matrix¹⁸. ADH1B-
18 expressing A2780 cells also displayed enhanced invasion capacity *in vitro* as well as clearance
19 capacity, a measure of the ability to clear 80% of mesothelial cells in a layer¹⁸. ADH1B
20 expression was significantly increased in a panel of ovarian cancer cell lines tested, including
21
22
23
24
25
26
27
28

1 HeyA8, IGROV, and SKOV3ip1, and hypoxic conditions supported increased ADH1B
2 expression¹⁸. In patients with ovarian cancer, high expression of ADH1B was significantly
3 correlated with lower overall and progression-free survival¹⁸, in contrast to the data from our
4 analysis of patients with NSCLC. Thus, genetic variation in ADH1B can confer statistically
5 significant changes in risk of development of cancers including esophageal cancer and colon
6 cancer, which can be modified by environmental factors like alcohol consumption; in ovarian
7 cancer, ADH1B is expressed at significantly higher levels in cancer cell lines, ectopic expression
8 of ADH1B supports tumor growth and metastasis, and higher expression of ADH1B is correlated
9 with enhanced patient outcomes. This is in contrast with our findings where ADH1B expression
10 is decreased in tumors from patients with non-small cell lung cancer, and where lower expression
11 of ADH1B is associated with enhanced patient outcomes.

12
13 We found, by mining multiple independently published microarray datasets, that ADH1B
14 was among the genes most differentially expressed in the primary tumors of patients with non-
15 small cell lung cancer; ADH1B was expressed at significantly lower levels in NSCLC tumors
16 as compared to the lung, and decreased expression of ADH1B was significantly associated with
17 worse overall survival in NSCLC patients. ADH1B may be of relevance to the biology of tumor
18 initiation or progression in patients with NSCLC of the adenocarcinoma type, the most common
19 type of the leading cause of cancer death in the United States and worldwide.

References

1. Siegel, R.L., Miller, K.D. and Jemal, A., 2019. Cancer statistics, 2019. *CA: a cancer journal for clinicians*, 69(1), pp.7-34.
2. Kabbout, M., Garcia, M.M., Fujimoto, J., Liu, D.D., Woods, D., Chow, C.W., Mendoza, G., Momin, A.A., James, B.P., Solis, L. and Behrens, C., 2013. Ets2 mediated tumor suppressive function and met oncogene inhibition in human non-small cell lung cancer. *Clinical cancer research*, 19(13), pp.3383-3395.
3. Meister, M., Belousov, A., Xu, E.C., Schnabel, P., Warth, A. and Hoofmann, H., 2014. Intra-tumor heterogeneity of gene expression profiles in early stage non-small cell lung cancer. *J Bioinf Res Stud*, 1, p.1.
4. Marwitz, S., Depner, S., Dvornikov, D., Merkle, R., Szczygiel, M., Müller-Decker, K., Lucarelli, P., Wäsch, M., Mairbäurl, H., Rabe, K.F. and Kugler, C., 2016. Downregulation of the TGF β pseudoreceptor BAMBI in non-small cell lung cancer enhances TGF β signaling and invasion. *Cancer research*, 76(13), pp.3785-3801.
5. Lung Cancer - Non-Small Cell: Statistics. <https://www.cancer.net/cancer-types/lung-cancer-non-small-cell/statistics>.
6. Gyorffy, B., Surowiak, P., Budczies, J. and Lanczky, A., 2013. Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer. *PloS one*, 8(12), pp.e82241-e82241.
7. Edenberg, H.J., 2007. The genetics of alcohol metabolism: role of alcohol dehydrogenase and aldehyde dehydrogenase variants. *Alcohol Research & Health*, 30(1), p.5.
8. Edenberg, H.J. and McClintick, J.N., 2018. Alcohol dehydrogenases, aldehyde dehydrogenases, and alcohol use disorders: a critical review. *Alcoholism: Clinical and Experimental Research*, 42(12), pp.2281-2297.
9. Cederbaum, A.I., 2012. Alcohol metabolism. *Clinics in liver disease*, 16(4), pp.667-685.
10. MATYSIAK-BUDNIK, T.A.M.A.R.A., Jokelainen, K., Kärkkäinen, P., Mäkisalo, H., Ohisalo, J. and Salaspuro, M., 1996. Hepatotoxicity and absorption of extrahepatic acetaldehyde in rats. *The Journal of pathology*, 178(4), pp.469-474.
11. Cui, R., Kamatani, Y., Takahashi, A., Usami, M., Hosono, N., Kawaguchi, T., Tsunoda, T., Kamatani, N., Kubo, M., Nakamura, Y. and Matsuda, K., 2009. Functional variants in ADH1B

and ALDH2 coupled with alcohol and smoking synergistically enhance esophageal cancer risk. *Gastroenterology*, 137(5), pp.1768-1775.

12. Wu, M., Chang, S.C., Kampman, E., Yang, J., Wang, X.S., Gu, X.P., Han, R.Q., Liu, A.M., Wallar, G., Zhou, J.Y. and Kok, F.J., 2013. Single nucleotide polymorphisms of ADH1B, ADH1C and ALDH2 genes and esophageal cancer: A population-based case-control study in China. *International journal of cancer*, 132(8), pp.1868-1877.
13. Li, D., Zhang, R., Jin, T., He, N., Ren, L., Zhang, Z., Zhang, Q., Xu, R., Tao, H., Zeng, G. and Gao, J., 2016. ADH1B and CDH1 polymorphisms predict prognosis in male patients with non-metastatic laryngeal cancer. *Oncotarget*, 7(45), p.73216.
14. Offermans, N.S., Ketcham, S.M., van den Brandt, P.A., Weijenberg, M.P. and Simons, C.C., 2018. Alcohol intake, ADH1B and ADH1C genotypes, and the risk of colorectal cancer by sex and subsite in the Netherlands Cohort Study. *Carcinogenesis*, 39(3), pp.375-388.
15. Liu, S., Huang, T., Liu, M., He, W., Zhao, Y., Yang, L., Long, Y., Zong, D., Zeng, H., Liu, Y. and Liao, W., 2020. The Genomic Characteristics of ALK Fusion Positive Tumors in Chinese NSCLC Patients. *Frontiers in Oncology*, 10.
16. Li, Q.G., He, Y.H., Wu, H., Yang, C.P., Pu, S.Y., Fan, S.Q., Jiang, L.P., Shen, Q.S., Wang, X.X., Chen, X.Q. and Yu, Q., 2017. A normalization-free and nonparametric method sharpens large-scale transcriptome analysis and reveals common gene alteration patterns in cancers. *Theranostics*, 7(11), p.2888.
17. Ahmed, F., 2019. Integrated network analysis reveals FOXM1 and MYBL2 as key regulators of cell proliferation in non-small cell lung cancer. *Frontiers in oncology*, 9, p.1011.
18. Gharpure, K.M., Lara, O.D., Wen, Y., Pradeep, S., LaFargue, C., Ivan, C., Rupaimoole, R., Hu, W., Mangala, L.S., Wu, S.Y. and Nagaraja, A.S., 2018. ADH1B promotes mesothelial clearance and ovarian cancer infiltration. *Oncotarget*, 9(38), p.25115.

Rank	ID	p-value	t	B	FC	Gene	Gene name
59	209614_at	7.04E-16	-15.123601	26.2224983	0.5836 ± 0.1060	ADH1B	alcohol dehydrogenase 1B (class I), beta polypeptide

Table 1: ADH1B is differentially expressed in NSCLC tumors.

The rank of differential expression relative all transcripts measured, probe ID, *p*-value of global differential expression, t, a moderated t-statistic, B, the log-odds of differential expression between the groups compared, fold change of ADH1B expression in NSCLC tumors as compared to the lung, gene and gene name are listed in this chart.

Rank	ID	p-value	t	B	Gene	Gene name
665	8101881	1.3E-14	-8.88285	22.722815	ADH1B	alcohol dehydrogenase 1B (class I), beta polypeptide

Table 2: ADH1B is differentially expressed in NSCLC tumors.

The rank of differential expression relative all transcripts measured, probe ID, *p*-value of global differential expression, *t*, a moderated *t*-statistic, *B*, the log-odds of differential expression between the groups compared, fold change of ADH1B expression in NSCLC tumors as compared to the lung, and gene are listed in this chart.

Rank	ID	p-value	t	B	Gene
165	A_33_P3353737	9.55E-12	-1.09E+01	16.8819094	ADH1B

Table 3: ADH1B is differentially expressed in NSCLC tumors.

The rank of differential expression relative all transcripts measured, probe ID, *p*-value of global differential expression, t, a moderated t-statistic, B, the log-odds of differential expression between the groups compared, gene and gene name are listed in this chart.

Figure 1: ADH1B is expressed at significantly lower levels in NSCLC tumors when compared to the lung.

The mRNA expression level of ADH1B is graphically represented in the lung (left) and in NSCLC tumors of the adenocarcinoma type (right) with mean mRNA expression values marked and the result of a statistical test evaluating significance of difference in ADH1B expression between NSCLC tumors and the lung, a *p*-value, listed above.

Figure 2: ADH1B is expressed at significantly lower levels in NSCLC tumors when compared to the lung.

The mRNA expression level of ADH1B is graphically represented in the lung (left) and in NSCLC tumors of the adenocarcinoma type (right) with mean mRNA expression values marked and the result of a statistical test evaluating significance of difference in ADH1B expression between NSCLC tumors and the lung, a *p*-value, listed above.

Figure 4: ADH1B expression in NSCLC tumors significantly correlates with overall survival.

Depicted in this Kaplan-Meier plot is the probability of overall survival for $n=1925$ total patients stratified into two groups, based on low or high expression of ADH1B in patient tumors. The log rank p -value denoting statistical significance of difference in overall survival when comparing the two groups, as well as hazard ratio for this comparison is listed above. Listed below is the number of patients at risk (number of patients alive) per interval, after stratification based on ADH1B expression; in the first interval, number at risk is number of patients alive; in each subsequent interval, number at risk is the number at risk less those who have expired or are censored.

Low expression cohort (months)	High expression cohort (months)
52	117.33

Table 4: ADH1B expression in NSCLC tumors significantly correlates with patient survival.

The median overall survival of $n=1925$ NSCLC patients based on stratification into low or high expression of ADH1B in tumors is listed in this chart.