The inter- α -trypsin inhibitor heavy chain family member 5, ITIH5, is differentially expressed in non-small cell lung cancer and associates with patient survival.

Shahan Mamoor, MS¹

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 the gene encoding the inter- α -trypsin inhibitor heavy chain family member 5, ITIH5, was among those whose expression was most different in human NSCLC tumors as compared to the lung. ITIH5 expression levels were significantly decreased in NSCLC tumors as compared to the lung, and lower expression of ITIH5 in patient tumors was significantly associated with worse overall survival. ITIH5 may be important for initiation or progression of non-small cell lung cancer in humans.

Keywords: ITIH5, 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 ITIH5 in
10 adenocarcinoma tumors from patients with NSCLC, suggesting ITIH5 may be important for
11 NSCLC tumor initiation or progression.

12 **Methods**

13 We utilized microarray datasets, GSE43458², GSE33532³ and GSE74706⁴ for this
14 differential gene expression analysis of NSCLC tumors in conjunction with GEO2R. GSE43458
15 was generated using Affymetrix Human Gene 1.0 ST Array technology; for this analysis, we
16 used with $n=30$ control lung tissue and $n=80$ NSCLC tumors, and the analysis was performed
17 using platform GPL6244. GSE33532 was generated using Affymetrix Human Genome U133
18 Plus 2.0 Array technology; for this analysis, we used $n=20$ control lung tissue and $n=10$ NSCLC
19 tumors, and the analysis was performed using platform GPL570. GSE74706 was generated
20 using Agilent-026652 Whole Human Genome Microarray 4x44K v2 technology; for this

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

The Benjamini and Hochberg method of p -value adjustment was used for ranking of differential expression but raw p -values were used to assess statistical significance of global differential expression. Log-transformation of data was auto-detected, and the NCBI generated category of platform annotation was used. A statistical test was performed to evaluate whether ITIH5 expression was significantly between normal lung tissue and NSCLC tumors using a two-tailed, unpaired t-test with Welch's correction. We used PRISM for all statistical analyses of differential gene expression in NSCLC tumors (Version 8.4.0)(455). For Kaplan-Meier survival analysis, we used the Kaplan-Meier plotter online tool⁶ for correlation of ITIH5 mRNA expression levels with overall survival in $n=1144$ non-small cell lung cancer patients.

Results

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

ITIH5 is differentially expressed in non-small cell lung cancers.

We found significant differential expression of the gene encoding the inter- α -trypsin inhibitor heavy chain family member 5, ITIH5, in NSCLC tumors when compared to the lung²

(Table 1). When sorting each of the transcripts measured by microarray based on significance of difference in ITIH5 expression between NSCLC tumors and the normal lung, in this dataset, ITIH5 ranked 304 out of 25906 total transcripts, equating to 98.8% differential expression (Table 1). Differential expression of ITIH5 in NSCLC tumors was statistically significant (Table 1; $p=7.87E-19$).

We queried a second microarray dataset³ to determine if we could validate differential expression of ITIH5 in non-small cell lung cancers (Table 2). We found significant differential expression of ITIH5 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 ITIH5 between NSCLC tumors and the normal lung, ITIH5 ranked 56 of 33252 total transcripts (Table 2), equating to 99.8 % differential expression. Differential expression of ITIH5 in NSCLC tumors was statistically significant (Table 2; $p=6.08E-16$).

Analysis of a third microarray dataset⁴ again revealed significant differential expression of ITIH5 in NSCLC tumors of the adenocarcinoma type. When sorting each of the transcripts measured based on significance of difference in expression of ITIH5 between NSCLC tumors and the normal lung, ITIH5 ranked 29 out of 34183 total transcripts (Table 3), equating to 99.9% differential expression. Differential expression of ITIH5 in NSCLC tumors was statistically significant (Table 3; $p=6.03E-14$).

ITIH5 is expressed at significantly lower levels in NSCLC tumors as compared to the lung.

We obtained exact mRNA levels for ITIH5 from NSCLC tumors and from the lung to

1 directly compare ITIH5 expression between tumor and control lung tissue and assess for
2 statistical significance. ITIH5 was expressed at significantly lower levels in NSCLC tumors as
3 compared to the normal lung in each dataset queried, respectively (Figure 1: $p < 0.0001$ and
4 Figure 2: $p < 0.0001$). We calculated a mean fold change of 0.5980 ± 0.0568 in ITIH5 expression
5 when comparing NSCLC tumors to the lung (Table 2).
6
7

9 ITIH5 expression in NSCLC tumors correlates with overall survival.

10 We performed Kaplan-Meier survival analysis using ITIH5 mRNA expression in NSCLC
11 tumors coupled with paired overall survival data from each patient, in 1144 NSCLC patients in
12 total, to determine whether ITIH5 tumor expression was correlated with survival outcomes in
13 NSCLC. We found that patients whose tumors expressed lower levels of ITIH5 possessed
14 significantly shorter overall survival than patients with high tumor expression of ITIH5 (Figure
15 3). Median overall survival (OS) of patients in the low expression cohort was 46.7 months,
16 while median OS in patients in the high ITIH5 expression cohort was 98.5 months (Table 4); this
17 difference in median OS based on ITIH5 tumor expression in NSCLC was statistically
18 significant (Figure 3; logrank p -value: $1.5e-11$, hazard ratio: 0.57 (0.49 -0.67); false discovery
19 rate=0.01).
20
21
22
23

24 Thus, blind comparative transcriptome analysis of non-small cell lung cancers revealed
25 differential expression of ITIH5 as among the most significant transcriptional features of NSCLC
26 tumors, and ITIH5 expression was significantly correlated with patient outcomes, as patients
27 with lower tumor expression of ITIH5 possessed significantly worse overall survival.
28

Discussion

ITIH5 was cloned as a member of the inter- α -trypsin inhibitor (ITIH) and heavy chain gene families, and contains the vault protein inter-alpha-trypsin (VIT) and von Willebrand type A (vWA) domains conserved in all ITIHs⁷. ITIs contain a light chain termed bikunin and two heavy chains known as the ITI heavy chains. There is relatively limited information available regarding the basic biology of ITIH5. ITIH5 expression has been observed in the placenta, adipose tissue and adipocytes. ITIH5 is expressed at high levels in the subcutaneous adipose tissue of obese humans as compared to that of lean humans⁸. Hyaluronan (HA) is an interaction partner of ITIH5, and functions to stabilize HA. In an *in vitro* skin sensitizing assay, ITIH could decrease the potential of contact sensitizers to induce sensitization⁹.

Perturbed expression of ITIH5 has been reported in multiple human cancers. In 14.4% of acute myeloid leukemia samples (15/104), the ITIH5 promoter was hypermethylated¹⁰. A study using the S2-028 pancreatic ductal adenocarcinoma (PDAC) cell line, which is otherwise non-metastatic, recovered a hairpin targeting ITIH5 from a liver metastasis in a genome-wide shRNA screen in mice, suggesting ITIH5 could function as a suppressor of metastasis in PDAC¹¹. PDAC cell lines of metastatic nature displayed higher expression of ITIH5, and ectopic expression of ITIH5 *in vitro* resulted in decreased invasion and cell motility¹¹. In breast cancer, immunohistochemical analysis of patient tumors revealed that expression of ITIH5 was decreased in 42% of invasive breast cancers (92/217)¹². Breast cancer patients whose tumors had retained ITIH5 expression possessed increased recurrence-free survival and overall survival

1 as opposed to those patients whose tumors had decreased ITIH5 expression¹². The ITIH5
2 promoter was methylated in 41% of primary breast tumors and ITIH5 promoter methylation was
3 associated with decreased ITIH5 mRNA expression¹². Expression of ITIH5 in MDA-MB-231
4 cells, which otherwise do not express ITIH5, resulted in decreased cell proliferation¹². We
5 found one report of ITIH5 expression in NSCLC, describing decreased mRNA expression in
6 NSCLC patient tumor tissues as compared to normal lung, as well as increased methylation of
7 the ITIH5 promoter in 50% of squamous cell lung carcinomas (7/14) and 54% of
8 adenocarcinomas (7/13)¹³. In lung tumor tissue from the Cancer Genome Atlas, methylation of
9 the ITIH promoter was localized proximal to the transcription start site. Higher expression of
10 ITIH5 in squamoid but not bronchoid adenocarcinoma was associated with greater recurrence-
11 free survival¹³.

12 We found, by mining multiple independently published microarray datasets, that the
13 endothelial cell adhesion molecule, ITIH5, was among the genes most differentially expressed in
14 the primary tumors of patients with non-small cell lung cancer; ITIH5 was expressed at
15 significantly lower levels in NSCLC tumors as compared to the lung, and decreased expression
16 of ITIH5 was significantly associated with worse overall survival in NSCLC patients. ITIH5
17 may be of relevance to the biology of tumor initiation or progression in patients with NSCLC of
18 the adenocarcinoma type, the most common type of the leading cause of cancer death in the
19 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. 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.
3. 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.
4. 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.
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. Himmelfarb, M., Klopocki, E., Grube, S., Staub, E., Klamann, I., Hinzmann, B., Kristiansen, G., Rosenthal, A., Dürst, M. and Dahl, E., 2004. ITIH5, a novel member of the inter- α -trypsin inhibitor heavy chain family is downregulated in breast cancer. *Cancer letters*, 204(1), pp. 69-77.
8. Anveden, Å., Sjöholm, K., Jacobson, P., Palsdottir, V., Walley, A.J., Froguel, P., Al-Daghri, N., McTernan, P.G., Mejhert, N., Arner, P. and Sjöström, L., 2012. ITIH-5 Expression in Human Adipose Tissue Is Increased in Obesity. *Obesity*, 20(4), pp.708-714.
9. Huth, S., Huth, L., Marquardt, Y., Fietkau, K., Dahl, E., Esser, P.R., Martin, S.F., Heise, R., Merk, H.F. and Baron, J.M., 2020. Inter- α -Trypsin Inhibitor Heavy Chain 5 (ITIH5) Is a Natural Stabilizer of Hyaluronan That Modulates Biological Processes in the Skin. *Skin Pharmacology and Physiology*, 33(4), pp.198-206.

- 1 10.Oing, C., Jost, E., Dahl, E., Wilop, S., Brümmendorf, T.H. and Galm, O., 2011. Aberrant
2 DNA hypermethylation of the ITIH5 tumor suppressor gene in acute myeloid
3 leukemia. *Clinical epigenetics*, 2(2), pp.419-423.
- 4 11.Sasaki, K., Kurahara, H., Young, E.D., Natsugoe, S., Ijichi, A., Iwakuma, T. and Welch, D.R.,
5 2017. Genome-wide in vivo RNAi screen identifies ITIH5 as a metastasis suppressor in
6 pancreatic cancer. *Clinical & experimental metastasis*, 34(3-4), pp.229-239.
- 7 12.Veeck, J., Chorovicer, M., Naami, A., Breuer, E., Zafrakas, M., Bektas, N., Dürst, M.,
8 Kristiansen, G., Wild, P.J., Hartmann, A. and Knuechel, R., 2008. The extracellular matrix
9 protein ITIH5 is a novel prognostic marker in invasive node-negative breast cancer and its
aberrant expression is caused by promoter hypermethylation. *Oncogene*, 27(6), pp.865-876.
- 10 13.Dötsch, M.M., Kloten, V., Schlensog, M., Heide, T., Braunschweig, T., Veeck, J., Petersen, I.,
11 Knüchel, R. and Dahl, E., 2015. Low expression of ITIH5 in adenocarcinoma of the lung is
12 associated with unfavorable patients' outcome. *Epigenetics*, 10(10), pp.903-912.

Rank	ID	p-value	t	B	Gene	Gene name
304	7931977	7.87E-19	-10.713792	32.302453	ITIH5	inter-alpha-trypsin inhibitor heavy chain family member 5

Table 1: ITIH5 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 ITIH5 expression in NSCLC tumors as compared to the lung, gene and gene name are listed in this chart.

Rank	ID	p-value	t	B	FC	Gene	Gene name
56	1553243_at	6.08E-16	-15.205128	26.3675456	0.5980 ± 0.0568	ITIH5	inter-alpha-trypsin inhibitor heavy chain family member 5

Table 2: ITIH5 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 ITIH5 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
29	A_33_P3420380	6.03E-14	-1.34E+01	21.8264141	ITIH5

Table 3: ITIH5 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, and gene are listed in this chart.

Figure 1: ITIH5 is expressed at significantly lower levels in NSCLC tumors when compared to the lung.

The mRNA expression level of ITIH5 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 ITIH5 expression between NSCLC tumors and the lung, a *p*-value, listed above.

Figure 2: ITIH5 transcript is expressed at significantly lower levels in NSCLC tumors when compared to the lung.

The mRNA expression level of ITIH5 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 ITIH5 expression between NSCLC tumors and the lung, a *p*-value, listed above.

Figure 3: ITIH5 expression in NSCLC tumors significantly correlates with overall survival.

Depicted in this Kaplan-Meier plot is the probability of overall survival for $n=1144$ total patients stratified into two groups, based on low or high expression of ITIH5 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 ITIH5 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)
46.7	98.5

Table 4: ITIH5 expression in NSCLC tumors significantly correlates with patient survival.

The median overall survival of $n=1144$ NSCLC patients based on stratification into low or high expression of ITIH5 in tumors is listed in this chart.