The NCK associated protein 5, NCKAP5, 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 the gene encoding the NCK associated protein 5, NCKAP5, was among those whose expression was most different in human NSCLC tumors as compared to the lung. NCKAP5 expression levels were significantly decreased in NSCLC tumors as compared to the lung, and lower expression of NCKAP5 in patient tumors was significantly associated with worse overall survival. NCKAP5 may be important for initiation or progression of non-small cell lung cancer in humans.

Keywords: NCKAP5, 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 NCKAP5
10 in adenocarcinoma tumors from patients with NSCLC, suggesting NCKAP5 may be important
11 for NSCLC tumor initiation or progression.
12
13
14
15

16 **Methods**

17 We utilized microarray datasets, GSE43458², GSE33532³ and GSE74706⁴ for this
18 differential gene expression analysis of NSCLC tumors in conjunction with GEO2R. GSE43458
19 was generated using Affymetrix Human Gene 1.0 ST Array technology; for this analysis, we
20 used with $n=30$ control lung tissue and $n=80$ NSCLC tumors, and the analysis was performed
21 using platform GPL6244. GSE33532 was generated using Affymetrix Human Genome U133
22 Plus 2.0 Array technology; for this analysis, we used $n=20$ control lung tissue and $n=10$ NSCLC
23 tumors, and the analysis was performed using platform GPL570. GSE74706 was generated
24 using Agilent-026652 Whole Human Genome Microarray 4x44K v2 technology; for this
25
26
27
28

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 NCKAP5 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 NCKAP5 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.

NCKAP5 is differentially expressed in non-small cell lung cancers.

We found significant differential expression of the gene encoding the NCK associated

protein 5, NCKAP5 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 NCKAP5 expression between NSCLC tumors and the normal lung, in this dataset, NCKAP5 ranked 19 out of 25906 total transcripts, equating to 99.9% differential expression (Table 1). Differential expression of NCKAP5 in NSCLC tumors was statistically significant (Table 1; $p=1.69E-29$).

We queried a second microarray dataset³ to determine if we could validate differential expression of NCKAP5 in non-small cell lung cancers (Table 2). We found significant differential expression of NCKAP5 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 NCKAP5 between NSCLC tumors and the normal lung, NCKAP5 ranked 8 of 33252 total transcripts (Table 2). Differential expression of NCKAP5 in NSCLC tumors was statistically significant (Table 2; $p=1.2E-18$).

Analysis of a third microarray dataset⁴ again revealed significant differential expression of NCKAP5 in NSCLC tumors of the adenocarcinoma type. When sorting each of the transcripts measured based on significance of difference in expression of NCKAP5 between NSCLC tumors and the normal lung, NCKAP5 ranked 139 out of 34183 total transcripts (Table 3). Differential expression of NCKAP5 in NSCLC tumors was statistically significant (Table 3; $p=4.84E-12$).

NCKAP5 is expressed at significantly lower levels in NSCLC tumors as compared to the lung.

We obtained exact mRNA levels for NCKAP5 from NSCLC tumors and from the lung to

1 directly compare NCKAP5 expression between tumor and control lung tissue and assess for
2 statistical significance. NCKAP5 was expressed at significantly lower levels in NSCLC tumors
3 as compared to the normal lung in each dataset queried, respectively (Figure 1: $p < 0.0001$, Figure
4 2: $p < 0.0001$, and Figure 3: $p < 0.0001$). We calculated a mean fold change of 0.6918 ± 0.1041
5 and 0.5842 ± 0.0841 in NCKAP5 expression when comparing NSCLC tumors to the lung (Table
6 1 and Table 2, respectively).

7 NCKAP5 expression in NSCLC tumors correlates with overall survival.

8 We performed Kaplan-Meier survival analysis using NCKAP5 mRNA expression in
9 NSCLC tumors coupled with paired overall survival data from each patient, in 1144 NSCLC
10 patients in total, to determine whether NCKAP5 tumor expression was correlated with survival
11 outcomes in NSCLC. We found that patients whose tumors expressed lower levels of NCKAP5
12 possessed significantly shorter overall survival than patients with high tumor expression of
13 NCKAP5 (Figure 4). Median overall survival (OS) of patients in the low expression cohort was
14 42 months, while median OS in patients in the high NCKAP5 expression cohort was 93 months
15 (Table 4); this difference in median OS based on NCKAP5 tumor expression in NSCLC was
16 statistically significant (Figure 4; logrank p -value: $3.3e-9$; hazard ratio: 0.61 (0.51 - 0.72); false
17 discovery rate=0.01).

18 Thus, blind comparative transcriptome analysis of non-small cell lung cancers revealed
19 differential expression of transcripts encoded by the NCKAP5 gene as among the most
20 significant transcriptional features of NSCLC tumors, and NCKAP5 expression was significantly
21

1 correlated with patient outcomes, as patients with lower tumor expression of NCKAP5 possessed
2 significantly worse overall survival.

3 4 **Discussion**

5
6 NCKAP5 is one member of a family of molecules associated with Nck⁷, known as the
7 Nck-associated proteins, or NCKAP. Information of NCKAP5 is relatively limited but
8 NCKAP1, also known as hematopoietic protein-2 (Hem-2), and NCKAP1-like, also known as
9 Hem-1, are components of the Wiskott-Aldrich syndrome protein (WASP)-family verprolin
10 homology complex, or WAVE complex, in conjunction with Sra, Abi, Brk1 and Wave (SCAR)⁸.
11 The WAVE complex functions to induce polymerization of F-actin downstream of Rac in
12 changes of cellular architecture during processes including membrane ruffling^{9,10}. Neutrophils
13 from Hem-1-deficient mice manifest significant defects in F-actin polymerization in response to
14 the chemoattractant MLP as well as migration and phagocytosis¹¹. In addition, lymphocyte
15 development is altered, with a block in T-lymphocyte development between stages double
16 negative 3 (DN3 CD25^{mid} CD44⁻) and DN4 (CD25⁻ CD44⁻), numbers of $\alpha\beta$ as well as $\gamma\delta$ T-cells
17 are significantly reduced in NCKAP5-deficient mice, and production of the IL-17 cytokine is
18 elevated in T-lymphocytes¹¹. Together, the data demonstrate that an NCKAP protein has
19 important functions in immunity *in vivo* through control of the actin cytoskeleton and is critical
20 for lymphocyte development and function.
21
22
23
24
25

26 While information regarding NCKAP5 function in normal biological function and
27 development as well as in human cancers is limited, one study described decreased expression of
28

Arpin, an inhibitor of Arp2/3 that functions to antagonize the WAVE complex in human breast cancer¹². Decreased tumor expression of Arpin was associated with worse metastasis-free survival (MFS), and this study also described correlation of high tumor expression of NCKAP1 with worse MFS in human breast cancer patients¹². Thus, decreased expression of an NCKAP family member, NCKAP1, has been described in a human cancer, cancer of the breast, and decreased NCKAP1 expression significantly correlated with patient survival outcome.

We found, by mining multiple independently published microarray datasets, that NCKAP5 was among the genes most differentially expressed in the primary tumors of patients with non-small cell lung cancer; NCKAP5 was expressed at significantly lower levels in NSCLC tumors as compared to the lung, and decreased expression of NCKAP5 was significantly associated with worse overall survival in NSCLC patients. NCKAP5 may be of relevance to the biology of tumor initiation or progression in patients with NSCLC of the adenocarcinoma type, the most common 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. Meister, M., Belousov, A., Xu, E.C., Schnabel, P., Warth, A. and Hoofmann, H., 2014. Intratumor 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. Lehmann, J.R.M., Riethmüller, G. and Johnson, J.P., 1990. Nck, a melanoma cDNA encoding a cytoplasmic protein consisting of the src homology units SH2 and SH3. *Nucleic Acids Research*, 18(4), p.1048.
8. Gautreau, A., Hsin-yi, H.H., Li, J., Steen, H., Gygi, S.P. and Kirschner, M.W., 2004. Purification and architecture of the ubiquitous Wave complex. *Proceedings of the National Academy of Sciences*, 101(13), pp.4379-4383.
9. Miki, H., Suetsugu, S. and Takenawa, T., 1998. WAVE, a novel WASP-family protein involved in actin reorganization induced by Rac. *The EMBO journal*, 17(23), pp.6932-6941.
10. Derivery, E. and Gautreau, A., 2010. Generation of branched actin networks: assembly and regulation of the N-WASP and WAVE molecular machines. *Bioessays*, 32(2), pp.119-131.

- 1 11. Park, H., Staehling-Hampton, K., Appleby, M.W., Brunkow, M.E., Habib, T., Zhang, Y.,
2 Ramsdell, F., Liggitt, H.D., Freie, B., Tsang, M. and Carlson, G., 2008. A point mutation in
3 the murine Hem1 gene reveals an essential role for Hematopoietic protein 1 in lymphopoiesis
4 and innate immunity. *The Journal of experimental medicine*, 205(12), pp.2899-2913.
- 5 12. Lomakina, M.E.,ALLEMAND, F., VACHER, S., MOLINIE, N., DANG, I., CACHEUX, W., CHIPYSHEVA,
6 T.A., ERMILOVA, V.D., DE KONING, L., DUBOIS, T. and BIÈCHE, I., 2016. Arpin downregulation in
breast cancer is associated with poor prognosis. *British journal of cancer*, 114(5), pp.545-553.

Rank	ID	p-value	t	B	FC	Gene	Gene name
19	8055323	1.69E-29	-15.472456	56.525853	0.6918 ± 0.1041	NCKAP5	NCK associated protein 5

Table 1: NCKAP5 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 NCKAP5 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
8	239650_at	1.2E-18	-18.972885	32.4199481	0.5842 ± 0.0841	NCKAP5	NCK associated protein 5

Table 2: NCKAP5 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.

Rank	ID	p-value	t	B	Gene
139	A_23_P360079	4.84E-12	-1.12E+01	17.5484242	NCKAP5

Table 3: NCKAP5 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 NCKAP5 expression in NSCLC tumors as compared to the lung, gene and gene name are listed in this chart.

Figure 1: NCKAP5 is expressed at significantly lower levels in NSCLC tumors when compared to the lung.

The mRNA expression level of NCKAP5 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 NCKAP5 expression between NSCLC tumors and the lung, a *p*-value, listed above.

Figure 2: NCKAP5 transcript is expressed at significantly lower levels in NSCLC tumors when compared to the lung.

The mRNA expression level of NCKAP5 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 NCKAP5 expression between NSCLC tumors and the lung, a *p*-value, listed above.

Figure 3: NCKAP5 transcript is expressed at significantly lower levels in NSCLC tumors when compared to the lung.

The mRNA expression level of NCKAP5 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 NCKAP5 expression between NSCLC tumors and the lung, a *p*-value, listed above.

Figure 4: NCKAP5 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 NCKAP5 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 NCKAP5 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)
42	93

Table 4: NCKAP5 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 NCKAP5 in tumors is listed in this chart.