TCF21 is differentially expressed in non-small cell lung cancer.

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^{2,3,4,5} to identify differentially expressed genes in NSCLC. We found that the gene encoding the transcription factor TCF21 was among the genes whose expression was most quantitatively different in tumors from patients with NSCLC as compared to the lung. TCF21 may be important for initiation or progression of non-small cell lung cancer in humans.

Keywords: TCF21, 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 years⁶. 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 years⁶. 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^{2,3,4,5} to compare global gene expression profiles of NSCLC tumors to
9 that of the normal lung. We found recurrent and significant differential expression of the
10 transcription factor TCF21 in adenocarcinoma tumors from patients with NSCLC, suggesting
11 TCF21 may be important for NSCLC tumor initiation or progression.

12 **Methods**

13 We utilized microarray datasets GSE74706², GSE43458³, GSE33532⁴ and GSE62113⁵
14 for this differential gene expression analysis of NSCLC tumors in conjunction with GEO2R.
15 GSE74706 was generated using Agilent-026652 Whole Human Genome Microarray 4x44K v2
16 technology; for this analysis, we used $n=18$ control lung tissue and $n=10$ NSCLC tumors, and the
17 analysis was performed using platform GPL13497. GSE43458 was generated using Affymetrix
18 Human Gene 1.0 ST Array technology; for this analysis, we used with $n=30$ control lung tissue
19 and $n=80$ NSCLC tumors, and the analysis was performed using platform GPL6244. GSE33532
20 was generated using Affymetrix Human Genome U133 Plus 2.0 Array technology; for this

analysis, we used $n=20$ control lung tissue and $n=10$ NSCLC tumors, and the analysis was performed using platform GPL570. GSE62113 was generated using Illumina HumanHT-12 WG-DASL V4.0 R2 expression beadchip technology; for this analysis, we used $n=9$ control lung tissue and $n=7$ NSCLC tumors, and the analysis was performed using platform GPL14951. 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 TCF21 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 (Version 8.4.0)(455).

Results

We harnessed the power of multiple, independently published microarray datasets^{2,3,4,5} to discover in an unbiased fashion and at the transcriptome-level the most striking gene expression features of NSCLC tumors.

TCF21 is differentially expressed in non-small cell lung cancers.

We found significant differential expression of the gene encoding the transcription factor TCF21 in NSCLC tumors when compared to the lung² (Table 1). When sorting each of the

transcripts measured based on significance of difference in expression of TCF21 between NSCLC tumors and the normal lung, TCF21 ranked 247 out of 34183 total transcripts (Table 1). Differential expression of TCF21 in NSCLC tumors was statistically significant (Table 1; $p=4.58E-11$).

Analysis of a second microarray dataset³ again revealed significant differential expression of TCF21 in NSCLC tumors of the adenocarcinoma type (Table 2). When sorting each of the transcripts measured based on significance of difference in expression of TCF21 between NSCLC tumors and the normal lung, TCF21 ranked 215 out of 33252 total transcripts (Table 2). Differential expression of TCF21 in NSCLC tumors was statistically significant (Table 2; $p=1.02E-20$).

We queried a third microarray dataset⁴, again finding significant differential expression of TCF21 in non-small cell lung cancers of the adenocarcinoma type when compared to the normal lung. When sorting each of the transcripts measured based on significance of difference in expression of TCF21 between NSCLC tumors and the normal lung, in this dataset, TCF21 ranked 9 out of 25906 total transcripts (Table 3). Differential expression of TCF21 in NSCLC tumors was statistically significant (Table 3; $p=1.98E-18$).

In a fourth dataset⁵, we again observed significant differential expression of TCF21 in human NSCLC tumors when compared to normal lung. When sorting each gene by significance of difference in expression between NSCLC tumors and the normal lung, TCF21 ranked 121 out of 29376 total transcripts (Table 4). Differential expression of TCF21 in this fourth dataset

was again statistically significant (Table 4; $p=8.36E-06$).

TCF21 is expressed at significantly lower levels in NSCLC tumors as compared to the lung.

We obtained exact mRNA levels for TCF21 from NSCLC tumors and from the lung to directly compare TCF21 expression between tumor and control lung tissue and assess for statistical significance. TCF21 was expressed at significantly lower levels in each of the datasets queried (Figure 1: $p<0.0001$; Figure 2: $p<0.0001$; Figure 3: $p<0.0001$; and Figure 4: $p=0.0002$). We calculated a fold change of 0.7908 ± 0.0905 (Table 2), 0.5922 ± 0.0823 (Table 3) and 0.7789 ± 0.0595 (Table 4) in TCF21 expression when comparing NSCLC tumors to the lung.

Thus, blind comparative transcriptome analysis of non-small cell lung cancers revealed differential expression of TCF21 as among the most significant transcriptional features of NSCLC tumors.

Discussion

TCF21, also known as capsulin and Pod1, was identified based on homology to eHAND and dHAND (Thing-1 and Thing-2) as a basic helix-loop-helix (bHLH) possessing mesodermic expression in epithelium of the developing gastrointestinal tract and respiratory system, and serving as a marker of a structure of the heart known as the spiral septum and in progenitor cells of the pericardium and coronary arteries^{7,8}. In the adult mouse, TCF21 is expressed at highest levels in the lungs, with prominent expression in the kidneys, liver, heart and testis⁷. TCF21, as a heterodimer with the E-protein E12, can bind CANNTG sequences, as is found in the E-box

1 sequence of the muscle creatine kinase enhancer⁷. TCF21-null mice fail to develop a spleen,
2 with extensive apoptotic cell death observable at embryonic day 12.5, and the lungs of TCF21-
3 null mice possess defects in branching morphogenesis and do not contain alveoli^{8,9}. The
4 epicardium of TCF21-null mice are lacking in cardiac fibroblasts with enlargement of the right
5 atrium¹⁰. The epithelial to mesenchymal transition (EMT) is perturbed in TCF21-null mice with
6 up-regulation of epithelial gene expression including Cldn15 and Krt14 as well as down-
7 regulation of mesenchymal gene expression including Slug, Snail, and Zeb1¹⁰. At embryonic
8 day 11.5 (E11.5), progenitors expressing Tcf21 developed into three cell types: fibroblasts,
9 lipofibroblasts and smooth muscle cells. In primary neonatal lung fibroblasts, expression of
10 Tcf21 resulted in increased levels of intracellular lipids, consistent with a role for Tcf21 in
11 lipofibroblast function¹¹. Lineage-labeling in mice revealed a Tcf21+ platelet-derived growth
12 factor (PDGFR) α -expressing alveolar population with the lung parenchyma adjacent to alveolar
13 epithelial type 2 cells in the lungs¹². Thus, Tcf21 is expressed in the lung within discrete cell
14 populations and its deficiency in mice results in severe lung morphological and vascular defects
15 as well as perturbation of the EMT gene expression program.

22 Tcf21 is implicated in a number of human cancers. In bladder cancer, Tcf21 is over-
23 expressed in primary tumors as compared to lymph node metastases and over-expression of
24 Tcf21 in bladder cancer cell lines results in decreases in metastatic and invasive capacity¹³. The
25 Tcf21 locus was found to be hyper-methylated in head and neck cancers, as well as in some non-
26 small cell lung cancers surveyed, and ectopic expression of Tcf21 in A549 NSCLC cell line

1 resulted in decreased cell proliferation and colony formation *in vitro* as well as in decreased
2 tumor weight and tumor volume *in vivo*¹⁴. A separate study found that a majority of NSCLC cell
3 lines tested possessed differential methylation of a CpG-rich sequence in exon 1 of Tcf21,
4 suggesting widespread hyper-methylation of the Tcf21 locus in NSCLC¹⁵. Treatment of the
5 H1299 NSCLC cell line with 5-aza-2'-deoxycytidine in resulted in increased cell death by
6 apoptosis as well as decreases in cell viability and invasive capacity¹⁶. Depletion of Tcf21
7 resulted in up-regulation of gene expression of a number of autophagic markers including Atg9,
8 Beclin, LC-I and LC-II, suggesting that Tcf21 could control expression of essential components
9 of the autophagic pathway in lung cancer cells¹⁶. One previous study has documented
10 differential expression of Tcf21 in NSCLC tumors as compared to the normal lung when
11 comparing tissue samples based on sex, and higher expression of Tcf21 was significantly
12 associated with enhanced patient outcomes including relapse-free survival and overall survival¹⁷.
13 Thus, Tcf21 is differentially methylated in a number of NSCLC cell lines, and differential
14 expression of Tcf21 has been documented in NSCLC patient tumors consistent with our findings
15 here.

22 Tcf21 is an important developmental transcriptional factor with essential roles in the
23 lungs; we found that Tcf21 was among the genes most differentially expressed in tumors from
24 patients with NSCLC of the adenocarcinoma type across four separate datasets. Tcf21 may be of
25 relevance to the initiation or progression of NSCLC, the most common type of the most lethal
26 cancer in the United States.

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. 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 TGFβ pseudoreceptor BAMBI in non-small cell lung cancer enhances TGFβ signaling and invasion. *Cancer research*, 76(13), pp.3785-3801.
3. 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.
4. 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.
5. Li, L., Wei, Y., To, C., Zhu, C.Q., Tong, J., Pham, N.A., Taylor, P., Ignatchenko, V., Ignatchenko, A., Zhang, W. and Wang, D., 2014. Integrated omic analysis of lung cancer reveals metabolism proteome signatures with prognostic impact. *Nature communications*, 5(1), pp.1-12.
6. Lung Cancer - Non-Small Cell: Statistics. <https://www.cancer.net/cancer-types/lung-cancer-non-small-cell/statistics>.
7. Lu, J., Richardson, J.A. and Olson, E.N., 1998. Capsulin: a novel bHLH transcription factor expressed in epicardial progenitors and mesenchyme of visceral organs. *Mechanisms of development*, 73(1), pp.23-32.
8. Quaggin, S.E., Schwartz, L., Cui, S., Igarashi, P., Deimling, J., Post, M. and Rossant, J., 1999. The basic-helix-loop-helix protein pod1 is critically important for kidney and lung organogenesis. *Development*, 126(24), pp.5771-5783.
9. Lu, J., Chang, P., Richardson, J.A., Gan, L., Weiler, H. and Olson, E.N., 2000. The basic helix-loop-helix transcription factor capsulin controls spleen organogenesis. *Proceedings of the National Academy of Sciences*, 97(17), pp.9525-9530.

10. Acharya, A., Baek, S.T., Huang, G., Eskiocak, B., Goetsch, S., Sung, C.Y., Banfi, S., Sauer, M.F., Olsen, G.S., Duffield, J.S. and Olson, E.N., 2012. The bHLH transcription factor Tcf21 is required for lineage-specific EMT of cardiac fibroblast progenitors. *Development*, 139(12), pp.2139-2149.
11. Park, J., Ivey, M.J., Deana, Y., Riggsbee, K.L., Sørensen, E., Schwabl, V., Sjöberg, C., Hjertberg, T., Park, G.Y., Swonger, J.M. and Rosengreen, T., 2019. The Tcf21 lineage constitutes the lung lipofibroblast population. *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 316(5), pp.L872-L885.
12. Rodríguez-Castillo, J.A., Ntokou, A., Atzberger, A., Tallquist, M.D., Seeger, W., Morty, R.E. and Ahlbrecht, K., 2017. Characterization of pulmonary TCF21 positive cells during neo-alveolarisation. *European Respiratory Journal* 50 (suppl 61) PA2091.
13. Mokkalapati, S., Porten, S.P., Narayan, V.M., Lim, A.H., Jayaratna, I.S., Roth, B., Cheng, T., Navai, N., Wszolek, M., Melquist, J. and Manyam, G., 2020. TCF21 promotes luminal-like differentiation and suppresses metastasis in bladder cancer. *Molecular Cancer Research*, 18(6), pp.811-821.
14. Smith, L.T., Lin, M., Brena, R.M., Lang, J.C., Schuller, D.E., Otterson, G.A., Morrison, C.D., Smiraglia, D.J. and Plass, C., 2006. Epigenetic regulation of the tumor suppressor gene TCF21 on 6q23-q24 in lung and head and neck cancer. *Proceedings of the National Academy of Sciences*, 103(4), pp.982-987.
15. Shivapurkar, N., Stastny, V., Xie, Y., Prinsen, C., Frenkel, E., Czerniak, B., Thunnissen, F.B., Minna, J.D. and Gazdar, A.F., 2008. Differential methylation of a short CpG-rich sequence within exon 1 of TCF21 gene: a promising cancer biomarker assay. *Cancer Epidemiology and Prevention Biomarkers*, 17(4), pp.995-1000.
16. Chen, B., Zeng, C., Ye, Y., Wu, D., Mu, Z., Liu, J., Xie, Y. and Wu, H., 2018. Promoter methylation of TCF21 may repress autophagy in the progression of lung cancer. *Journal of cell communication and signaling*, 12(2), pp.423-432.
17. Xiao, J., Liu, A., Lu, X., Chen, X., Li, W., He, S., He, B. and Chen, Q., 2017. Prognostic significance of TCF21 mRNA expression in patients with lung adenocarcinoma. *Scientific reports*, 7(1), pp.1-9.

Rank	ID	p-value	t	B	Gene
247	A_23_P156890	4.58E-11	-1.02E+01	15.3369325	TCF21

Table 1: TCF21 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.

Rank	ID	p-value	t	B	FC	Gene	Gene name
215	8122176	1.02E-20	-11.533628	36.596095	0.7908 ± 0.0905	TCF21	transcription factor 21

Table 2: TCF21 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 TCF21 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
9	204931_at	1.98E-18	-18.64419	31.9382991	0.5922 ± 0.0823	TCF21	transcription factor 21

Table 3: TCF21 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 TCF21 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
121	ILMN_1719351	8.36E-06	-6.5152813	3.890735	0.7789 ± 0.0595	TCF21

Table 4: TCF21 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 TCF21 expression in NSCLC tumors as compared to the lung, gene and gene name are listed in this chart.

Figure 1: TCF21 is expressed at significantly lower levels in NSCLC tumors when compared to the lung.

The mRNA expression level of TCF21 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 TCF21 expression between NSCLC tumors and the lung, a *p*-value, listed above.

Figure 2: TCF21 is expressed at significantly lower levels in NSCLC tumors when compared to the lung.

The mRNA expression level of TCF21 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 TCF21 expression between NSCLC tumors and the lung, a *p*-value, listed above.

Figure 3: TCF21 is expressed at significantly lower levels in NSCLC tumors when compared to the lung.

The mRNA expression level of TCF21 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 TCF21 expression between NSCLC tumors and the lung, a *p*-value, listed above.

Figure 4: TCF21 is expressed at significantly lower levels in NSCLC tumors when compared to the lung.

The mRNA expression level of TCF21 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 TCF21 expression between NSCLC tumors and the lung, a *p*-value, listed above.