Differential expression of transcripts from the gene encoding the vascular endothelial growth factor D, FIGF, in non-small cell lung cancer and association 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 three transcripts produced by the FIGF gene - a FIGF transcript and two read-through transcripts containing sequences from FIGF and a neighboring gene, PIR (PIR-FIGF) - among those whose expression was most quantitatively different in human NSCLC tumors as compared to the lung. FIGF and PIR-FIGF expression levels were significantly decreased in NSCLC tumors as compared to the lung, and lower expression of FIGF in patient tumors was significantly associated with worse overall survival. Vascular endothelial growth factor D may be important for initiation or progression of non-small cell lung cancer in humans.

Keywords: FIGF, vascular endothelial growth factor D, VEGF-D, 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 transcripts
10 from the gene encoding the vascular endothelial growth factor D, FIGF, in adenocarcinoma
11 tumors from patients with NSCLC, suggesting FIGF may be important for NSCLC tumor
12 initiation or progression.
13
14
15
16
17

18 **Methods**

19 We utilized microarray datasets GSE74706², GSE43458³, GSE33532⁴ and for this
20 differential gene expression analysis of NSCLC tumors in conjunction with GEO2R. GSE74706
21 was generated using Agilent-026652 Whole Human Genome Microarray 4x44K v2 technology;
22 for this analysis, we used $n=18$ control lung tissue and $n=10$ NSCLC tumors, and the analysis
23 was performed using platform GPL13497. GSE43458 was generated using Affymetrix Human
24 Gene 1.0 ST Array technology; for this analysis, we used with $n=30$ control lung tissue and $n=80$
25 NSCLC tumors, and the analysis was performed using platform GPL6244. GSE33532 was
26
27
28

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. 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 FIGF or PIR-FIGF 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 tool⁷ for correlation of FIGF mRNA expression levels with overall survival in $n=1925$ 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.

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

We found significant differential expression of the gene encoding the vascular endothelial

growth factor D, FIGF, 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 FIGF between NSCLC tumors and the normal lung, in this dataset, FIGF ranked 118 out of 34183 total transcripts (Table 1). Differential expression of FIGF in NSCLC tumors was statistically significant (Table 1; $p=2.87E-12$).

We queried a second microarray dataset³ to determine if we could validate differential expression of FIGF in non-small cell lung cancers (Table 2). We found significant differential expression of PIR-FIGF, which may be identical to the long non-coding RNA lncVEGFD1.2, 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 PIR-FIGF between NSCLC tumors and the normal lung, PIR-FIGF ranked 270 of 33252 total transcripts (Table 2). Differential expression of PIR-FIGF in NSCLC tumors was statistically significant (Table 2; $p=1.63E-19$).

Analysis of a third microarray dataset⁴ again revealed significant differential expression of PIR-FIGF in NSCLC tumors of the adenocarcinoma type. When sorting each of the transcripts measured based on significance of difference in expression of PIR-FIGF between NSCLC tumors and the normal lung, PIR-FIGF ranked 44 out of 25906 total transcripts (Table 3). Differential expression of PIR-FIGF in NSCLC tumors was statistically significant (Table 3; $p=2.35E-16$).

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

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

1 directly compare FIGF expression between tumor and control lung tissue and assess for
2 statistical significance. FIGF and a PIR-FIGF read-through transcript were expressed at
3 significantly lower levels in NSCLC tumors as compared to normal lung in each dataset queried,
4 respectively (Figure 1: $p < 0.0001$, Figure 2: $p < 0.0001$). We calculated a mean fold change of
5 0.2832 ± 0.1049 in FIGF expression (Table 1) and 0.5245 ± 0.1240 in PIR-FIGF expression
6 when comparing NSCLC tumors to the lung (Table 3).

7 FIGF expression in NSCLC tumors correlates with overall survival.

8 We performed Kaplan-Meier survival analysis using FIGF mRNA expression in NSCLC
9 tumors coupled with paired overall survival data from each patient, in 1925 NSCLC patients in
10 total, to determine whether FIGF tumor expression was correlated with survival outcomes in
11 NSCLC. We found that patients whose tumors expressed lower levels of FIGF possessed
12 significantly shorter overall survival than patients with high tumor expression of FIGF (Figure
13 3). Median overall survival (OS) of patients in the low expression cohort was 55 months, while
14 median OS in patients in the high FIGF expression cohort was 110.27 months (Table 4); this
15 difference in median OS based on FIGF tumor expression in NSCLC was statistically significant
16 (Figure 3; logrank p -value: $7.4e-14$; hazard ratio: 0.53 (0.45 - 0.63); false discovery rate=0.01).

17 Thus, blind comparative transcriptome analysis of non-small cell lung cancers revealed
18 differential expression of transcripts encoded by the FIGF gene as among the most significant
19 transcriptional features of NSCLC tumors, and FIGF expression was significantly correlated with
20 patient outcomes, as patients with lower tumor expression of FIGF possessed significantly worse
21 overall survival.

Discussion

The vascular endothelial growth factor (VEGF) D, VEGF-D, also known as the c-fos-induced growth factor, can function in support of angiogenesis⁶. VEGF-D can bind both VEGFR-2 and VEGFR-3, functioning as a ligand for these VEGF receptors⁷. Stimulation of human umbilical vein endothelial cells (HUVEC) with VEGF-D results in phosphorylation of tyrosine residues on VEGFR-2 and VEGFR-3⁶. Implantation of chinese hamster ovary (CHO) cells producing VEGF-D into rabbit corneas induces neovascularization, indicating that VEGF-D can support neovascularization *in vivo*⁶. VEGF-D is cleaved at the carboxy- and amino-termini to produce an active form of the VEGF-D polypeptide which is reported to form a homodimer⁸. During murine development, FIGF displays expression in the lung mesenchyme⁹. A male patient with intrapulmonary vasculopathy, hypoxemia, and tachypnea was found to possess mutation in the FIGF gene, with histological examination revealing ectasia in the pulmonary arteries¹⁰. The mutant VEGF-D protein produced in this patient, VEGF-D V118M, was more effective at promoting proliferation of VEGFR-2+ human aortic endothelial cells and of inducing phosphorylation of VEGFR-2 but not VEGFR-3¹⁰. VEGF-D V118M was also more effective at supporting angiogenesis *in vivo* in a chick embryo chorioallantoic membrane assay, evidenced by significant increases in vessel diameter and sprouting¹⁰. This demonstrates that mutation of the gene encoding vascular endothelial growth factor D can result in a human disease affecting the lung.

Expression of FIGF and its gene product, VEGF-D, have been implicated in a number of human cancers. VEGF-D expression can serve as a prognostic indicator in human colorectal and

1 epithelial ovarian cancer^{11,12}, and in breast cancer, VEGF-D expression correlates with
2 metastasis to the lymph nodes as well as both disease-free and overall survival¹³. In a
3 investigation of VEGF-D function in mouse models of cancer, study of the VEGF-D-293 NOD/
4 SCID mouse model revealed that 0/14 mice grafted with 293 cells developed lymph node
5 metastases, 0/16 mice grafted with VEGF-293 cells developed lymph node metastases, but 14
6 out of 23 mice grafted with VEGF-D-293 developed lymph node metastases. Thus, in this solid
7 tumor model, VEGF-D but not VEGF support metastasis to the lymph nodes.
8
9

10
11 It is possible that the PIR-FIGF transcript we identified as differentially expressed in
12 NSCLC tumors in two separate datasets is identical to a long non-coding RNA named
13 lncVEGFD-1.2^{14,15}. We found, by mining multiple independently published microarray
14 datasets, that FIGF was among the genes most differentially expressed in the primary tumors of
15 patients with non-small cell lung cancer; FIGF was expressed at significantly lower levels in
16 NSCLC tumors as compared to the lung, and decreased expression of FIGF was significantly
17 associated with worse overall survival in NSCLC patients. PIR-FIGF, a transcript containing
18 portions of both the pirin and vascular endothelial growth factor D genes, was also significantly
19 differentially expressed in NSCLC tumors. FIGF may be of relevance to the biology of tumor
20 initiation or progression in patients with NSCLC of the adenocarcinoma type, the most common
21 type of the leading cause of cancer death in the United States and worldwide.
22
23
24
25
26
27
28

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., Mairböurl, 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. Marconcini, L., Marchiò, S., Morbidelli, L., Cartocci, E., Albini, A., Ziche, M., Bussolino, F. and Oliviero, S., 1999. c-fos-induced growth factor/vascular endothelial growth factor D induces angiogenesis in vivo and in vitro. *Proceedings of the National Academy of Sciences*, 96(17), pp.9671-9676.
7. Achen, M.G., Jeltsch, M., Kukk, E., Mäkinen, T., Vitali, A., Wilks, A.F., Alitalo, K. and Stacker, S.A., 1998. Vascular endothelial growth factor D (VEGF-D) is a ligand for the tyrosine kinases VEGF receptor 2 (Flk1) and VEGF receptor 3 (Flt4). *Proceedings of the National Academy of Sciences*, 95(2), pp.548-553.
8. Stacker, S.A., Stenvers, K., Caesar, C., Vitali, A., Domagala, T., Nice, E., Roufail, S., Simpson, R.J., Moritz, R., Karpanen, T. and Alitalo, K., 1999. Biosynthesis of vascular endothelial growth factor-D involves proteolytic processing which generates non-covalent homodimers. *Journal of Biological Chemistry*, 274(45), pp.32127-32136.
9. Avantaggiato, V., Orlandini, M., Acampora, D., Oliviero, S. and Simeone, A., 1998. Embryonic expression pattern of the murine figf gene, a growth factor belonging to platelet-derived growth factor/vascular endothelial growth factor family. *Mechanisms of development*, 73(2), pp.221-224.

10. Bailey, E., Cui, Y., Casey, A., Stoler, J.M., Ai, X., Ma, D., Handin, R., Sliz, P., Vargas, S.O. and El-Chemaly, S.Y., 2017. Pulmonary vasculopathy associated with FIGF gene mutation. *The American journal of pathology*, 187(1), pp.25-32.
11. White, J.D., Hewett, P.W., Kosuge, D., McCulloch, T., Enholm, B.C., Carmichael, J. and Murray, J.C., 2002. Vascular endothelial growth factor-D expression is an independent prognostic marker for survival in colorectal carcinoma. *Cancer Research*, 62(6), pp. 1669-1675.
12. Yokoyama, Y., Charnock-Jones, D.S., Licence, D., Yanaihara, A., Hastings, J.M., Holland, C.M., Emoto, M., Umemoto, M., Sakamoto, T., Sato, S. and Mizunuma, H., 2003. Vascular endothelial growth factor-D is an independent prognostic factor in epithelial ovarian carcinoma. *British journal of cancer*, 88(2), pp.237-244.
13. Nakamura, Y., Yasuoka, H., Tsujimoto, M., Yang, Q., Imabun, S., Nakahara, M., Nakao, K., Nakamura, M., Mori, I. and Kakudo, K., 2003. Prognostic significance of vascular endothelial growth factor D in breast carcinoma with long-term follow-up. *Clinical Cancer Research*, 9(2), pp.716-721.
14. Shovlin, C.L. and Vizcaychipi, M.P., 2020. Vascular inflammation and endothelial injury in SARS-CoV-2 infection: the overlooked regulatory cascades implicated by the ACE2 gene cluster. *QJM: An International Journal of Medicine*.
15. Volders, P.J., Anckaert, J., Verheggen, K., Nuytens, J., Martens, L., Mestdagh, P. and Vandesompele, J., 2019. LNCipedia 5: towards a reference set of human long non-coding RNAs. *Nucleic acids research*, 47(D1), pp.D135-D139.

Rank	ID	p-value	t	B	FC	Gene
118	A_23_P45185	2.87E-12	-1.15E+01	18.0634272	0.2832 ± 0.1049	FIGF

Table 1: FIGF 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 FIGF 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
270	8171427	1.63E-19	-11.010746	33.860014	PIR-FIGF///FIGF	PIR-FIGF readthrough///c-fos induced growth factor

Table 2: PIR-FIGF 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	FC	Gene	Gene name
44	206742_at	2.35E-16	-15.736582	27.2972271	0.5245 ± 0.1240	PIR-FIGF///FIGF	PIR-FIGF readthrough ///c-fos induced growth factor

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

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

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

Figure 2: A PIR-FIGF read-through transcript is expressed at significantly lower levels in NSCLC tumors when compared to the lung.

The mRNA expression level of a PIR-FIGF read-through transcript 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 FIGF expression between NSCLC tumors and the lung, a *p*-value, listed above.

Figure 3: FIGF 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 FIGF 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 FIGF 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)
55	110.27

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