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

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

16 **Methods**

17 We utilized microarray datasets GSE74706², GSE43458³, GSE33532⁴ and for this
18 differential gene expression analysis of NSCLC tumors in conjunction with GEO2R. GSE74706
19 was generated using Agilent-026652 Whole Human Genome Microarray 4x44K v2 technology;
20 for this analysis, we used $n=18$ control lung tissue and $n=10$ NSCLC tumors, and the analysis
21 was performed using platform GPL13497. GSE43458 was generated using Affymetrix Human
22 Gene 1.0 ST Array technology; for this analysis, we used with $n=30$ control lung tissue and $n=80$
23 NSCLC tumors, and the analysis was performed using platform GPL6244. GSE33532 was
24 generated using Affymetrix Human Genome U133 Plus 2.0 Array technology; for this analysis,
25
26
27
28

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 MYZAP 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 MYZAP 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.

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

We found significant differential expression of the gene encoding the myocardial zonula adherens protein, MYZAP, in NSCLC tumors when compared to the lung² (Table 1). When

1 sorting each of the transcripts measured based on significance of difference in expression of
2 MYZAP between NSCLC tumors and the normal lung, in this dataset, MYZAP ranked 72 out of
3 34183 total transcripts (Table 1), equating to 99.8% differential expression. Differential
4 expression of MYZAP in NSCLC tumors was statistically significant (Table 1; $p=4.99\text{E-}13$).
5

6 We queried a second microarray dataset³ to determine if we could validate differential
7 expression of MYZAP in non-small cell lung cancers (Table 2). We found significant differential
8 expression of a MYZAP//GCOM transcript in NSCLC tumors of the adenocarcinoma type when
9 compared to the normal lung (Table 2). When sorting each of the transcripts measured based on
10 significance of difference in expression of this MYZAP//GCOM transcript between NSCLC
11 tumors and the normal lung, MYZAP//GCOM ranked 28 of 33252 total transcripts (Table 2),
12 equating to 99.9% differential expression. Differential expression of the MYZAP//GCOM
13 transcript in NSCLC tumors was statistically significant (Table 2; $p=8.6\text{E-}29$).
14
15
16

17 Analysis of a third microarray dataset⁴ again revealed significant differential expression
18 of MYZAP in NSCLC tumors of the adenocarcinoma type. When sorting each of the transcripts
19 measured based on significance of difference in expression of MYZAP between NSCLC tumors
20 and the normal lung, MYZAP ranked 135 out of 25906 total transcripts (Table 3), equating to
21 99.5% differential expression. Differential expression of MYZAP in NSCLC tumors was
22 statistically significant (Table 3; $p=3.95\text{E-}14$).
23
24
25

26 MYZAP is expressed at significantly lower levels in NSCLC tumors as compared to the lung.
27

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

1 directly compare MYZAP expression between tumor and control lung tissue and assess for
2 statistical significance. MYZAP 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: $p < 0.0001$, Figure 3: $p < 0.0001$). We calculated a mean fold change of 0.7719 ± 0.0747 and
5 0.6942 ± 0.0542 in MYZAP expression when comparing NSCLC tumors to the lung (Table 2
6 and Table 3, respectively).

10 MYZAP expression in NSCLC tumors correlates with overall survival.

11 We performed Kaplan-Meier survival analysis using MYZAP mRNA expression in
12 NSCLC tumors coupled with paired overall survival data from each patient, in 1144 NSCLC
13 patients in total, to determine whether MYZAP tumor expression was correlated with survival
14 outcomes in NSCLC. We found that patients whose tumors expressed lower levels of MYZAP
15 possessed significantly shorter overall survival than patients with high tumor expression of
16 MYZAP (Figure 4). Median overall survival (OS) of patients in the low expression cohort was
17 0.7719 ± 0.0747 months, while median OS in patients in the high MYZAP expression cohort was
18 0.6942 ± 0.0542 months (Table 4); this difference in median OS based on MYZAP tumor
19 expression in NSCLC was statistically significant (Figure 4; logrank p -value: $1.3e-10$; hazard
20 ratio: 0.57 ($0.48 - 0.68$); false discovery rate= 0.01).

25 Thus, blind comparative transcriptome analysis of non-small cell lung cancers revealed
26 differential expression of transcripts encoded by the MYZAP gene as among the most significant
27 transcriptional features of NSCLC tumors, and MYZAP expression was significantly correlated
28

1 with patient outcomes, as patients with lower tumor expression of MYZAP possessed
2 significantly worse overall survival.
3
4

5 **Discussion**

6 The myocardial zonula adherens protein (MYZAP), also known as myozap, was initially
7 described as part of the intercalated disc of cardiomyocytes⁷, but has otherwise been described as
8 a component of cell-cell junctions in the endothelium. Myozap has high expression in the lungs,
9 the heart, the skeletal muscle and the placenta⁷. In tissues of the heart, myozap can interact with
10 desmoplakin and zonula occludens-1⁷. Another interaction partner of myozap is a negative
11 regulator of Rho known as MRIP (myosin phosphatase-RhoA interacting protein)⁶. Knockdown
12 of the zebrafish ortholog of myozap results in cardiomyopathy, demonstrating its importance in
13 contractile tissues of the heart⁷. Myozap co-localizes with vascular endothelial cadherin (VE-
14 cadherin). Myozap localization has been described in the lymphatic and blood systems, in
15 cytoplasmic plaques of adherens junctions which connect endothelial cells of vascular tissues
16 including in the virgular cells of *complexus adhaerentes* in lymph node sinus⁹. Myozap can
17 activate the serum response factor (SRF) pathway through its ERM (ezrin/radixin/moesin)
18 domain and dependent on Rho¹⁰.
19
20
21
22
23

24 Over-expression of myozap using a transgene results in cardiomyopathy with
25 hypertrophy and LV dilation, and the hearts of these mice display aggregation of myozap in
26 conjunction with proteins of the intercalated disc including desmoplakin¹¹. Gene expression
27 associated with the serum response factor (SRF) is significantly up-regulated in tissues from
28

1 myozap over-expressing mice, as well as gene expression associated with autophagy and the
2 unfolded protein response¹¹. On the other hand, myozap-deficient mice develop context-specific
3 cardiac hypertrophy following biomechanical stress-induced pressure overload associated with
4 induction of natriuretic peptides A and B, in addition to impaired contractile function, indications
5 of heart failure and increased mortality as compared to wild-type mice¹².

8 A genome-wide association study (GWAS) of patients with atrial fibrillation revealed
9 significant association between a single-nucleotide polymorphism (SNP) resulting in a missense
10 variant of the myozap gene and risk for development of atrial fibrillation, conferring an odds
11 ratio of 1.37¹³. A second GWAS replicated association of a myozap SNP with atrial fibrillation¹⁴,
12 demonstrating that genetic variation in myozap is correlated with defects in electrical
13 conductivity in the human heart. We could not identify literature describing a role for myozap in
14 non-small cell lung cancer (NSCLC), but one study demonstrated that the A549 NSCLC cell line
15 expresses myozap at the protein level⁹.

19 We found, by mining multiple independently published microarray datasets, that
20 MYZAP, also known as myozap was among the genes most differentially expressed in the
21 primary tumors of patients with non-small cell lung cancer; MYZAP was expressed at
22 significantly lower levels in NSCLC tumors as compared to the lung, and decreased expression
23 of MYZAP was significantly associated with worse overall survival in NSCLC patients.
24 MYZAP may be of relevance to the biology of tumor initiation or progression in patients with
25 NSCLC of the adenocarcinoma type, the most common type of the leading cause of cancer death
26 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. Seeger TS, Frank D, Rohr C, Will R, Just S, Grund C, Lyon R, Luedde M, Koegl M, Sheikh F, Rottbauer W, Franke WW, Katus HA, Olson EN, Frey N. Myozap, a novel intercalated disc protein, activates serum response factor-dependent signaling and is required to maintain cardiac function in vivo. *Circ Res*. 2010 Mar 19;106(5):880-90.
8. Rickelt, S., Kuhn, C., Winter-Simanowski, S., Zimbelmann, R., Frey, N. and Franke, W.W., 2011. Protein myozap—a late addition to the molecular ensembles of various kinds of adherens junctions. *Cell and tissue research*, 346(3), pp.347-359.
9. Pieperhoff, S., Rickelt, S., Heid, H., Claycomb, W.C., Zimbelmann, R., Kuhn, C., Winter-Simanowski, S., Kuhn, C., Frey, N. and Franke, W.W., 2012. The plaque protein myozap identified as a novel major component of adhering junctions in endothelia of the blood and the lymph vascular systems. *Journal of cellular and molecular medicine*, 16(8), pp.1709-1719.

- 1 10. Frank, D., Seeger, T., Rohr, C., Will, R., Grund, C., Koegl, M., Franke, W.W., Katus, H.A.,
2 Olson, E.N. and Frey, N., 2009. Myozap, a Novel Cardiac Protein, is a Component of the
3 Intercalated Disc and Activates SRF-dependent Signaling.
- 4 11. Frank, D., Rangrez, A.Y., Poyanmehr, R., Seeger, T.S., Kuhn, C., Eden, M., Stiebeling, K.,
5 Bernt, A., Grund, C., Franke, W.W. and Frey, N., 2014. Mice with cardiac-restricted
6 overexpression of Myozap are sensitized to biomechanical stress and develop a protein-
7 aggregate-associated cardiomyopathy. *Journal of molecular and cellular cardiology*, 72, pp.
196-207.
- 8 12. Rangrez, A.Y., Eden, M., Poyanmehr, R., Kuhn, C., Stiebeling, K., Dierck, F., Bernt, A.,
9 Lüllmann-Rauch, R., Weiler, H., Kirchof, P. and Frank, D., 2016. Myozap deficiency
10 promotes adverse cardiac remodeling via differential regulation of mitogen-activated protein
11 kinase/serum-response factor and β -catenin/GSK-3 β protein signaling. *Journal of Biological
Chemistry*, 291(8), pp.4128-4143.
- 12 13. Thorolfsson, R.B., Sveinbjornsson, G., Sulem, P., Jonsson, S., Halldorsson, G.H., Melsted,
13 P., Ivarsdottir, E.V., Davidsson, O.B., Kristjansson, R.P., Thorleifsson, G. and Helgadottir, A.,
14 2017. Mutations in RPL3L and MYZAP increase risk of atrial fibrillation. *bioRxiv*, p.223578.
- 15 14. Thorolfsson, R.B., Sveinbjornsson, G., Sulem, P., Nielsen, J.B., Jonsson, S., Halldorsson,
16 G.H., Melsted, P., Ivarsdottir, E.V., Davidsson, O.B., Kristjansson, R.P. and Thorleifsson, G.,
17 2018. Coding variants in RPL3L and MYZAP increase risk of atrial
18 fibrillation. *Communications biology*, 1(1), pp.1-9.
- 19
20
21
22
23
24
25
26
27
28

Rank	ID	p-value	t	B	Gene
72	A_33_P3423874	4.99E-13	-1.23E+01	19.7724868	MYZAP

Table 1: MYZAP 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
28	7983890	8.6E-29	-15.143349	54.923339	0.7719 ± 0.0747	MYZAP///GCOM1	myocardial zonula adherens protein///GRINL1A complex locus 1

Table 2: A MYZAP//GCOM1 transcript 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 MYZAP//GCOM1 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
135	228568_at	3.95E-14	-13.026076	22.2523795	0.6942 ± 0.0542	MYZAP	myocardial zonula adherens protein

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

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

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

Figure 2: A MYZAP//GCOM1 transcript is expressed at significantly lower levels in NSCLC tumors when compared to the lung.

The mRNA expression level of a MYZAP//GCOM1 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 MYZAP//GCOM1 expression between NSCLC tumors and the lung, a *p*-value, listed above.

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

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

Figure 4: MYZAP expression in NSCLC tumors significantly correlates with overall patient 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 MYZAP 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 MYZAP 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)
41.37	96.2

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