Tox2 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 thymocyte selection-associated high mobility box protein family member 2 transcription factor Tox2 was among the genes whose expression was most quantitatively different in tumors from patients with NSCLC as compared to the lung. Tox2 expression was significantly decreased in NSCLC tumors as compared to the lung, and lower expression of Tox2 in patient tumors was significantly associated with worse overall survival. Tox2 may be important for initiation or progression of non-small cell lung cancer in humans.

Keywords: Tox2, 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
10 transcription factor Tox2 in adenocarcinoma tumors from patients with NSCLC, suggesting Tox2
11 may be important for NSCLC tumor initiation or progression.
12
13
14
15

16 **Methods**

17 We utilized microarray datasets GSE74706², GSE33532³ and GSE43458⁴ 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. GSE33532 was generated using Affymetrix Human
22 Genome U133 Plus 2.0 Array technology; for this analysis, we used $n=20$ control lung tissue and
23 $n=10$ NSCLC tumors, and the analysis was performed using platform GPL570. GSE43458 was
24 generated using Affymetrix Human Gene 1.0 ST Array technology; for this analysis, we used
25
26
27
28

with $n=30$ control lung tissue and $n=80$ NSCLC tumors, and the analysis was performed using platform GPL6244. 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 Tox2 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 Tox2 mRNA expression levels with overall survival in non-small cell lung cancer in $n=1144$ 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.

Tox2 is differentially expressed in non-small cell lung cancers.

We found significant differential expression of the gene encoding the thymocyte selection-associated high mobility box protein family member 2, Tox2, in NSCLC tumors when

1 compared to the lung² (Table 1). When sorting each of the transcripts measured based on
2 significance of difference in expression of Tox2 between NSCLC tumors and the normal lung,
3 Tox2 ranked 148 out of 34183 total transcripts (Table 1). Differential expression of Tox2 in
4 NSCLC tumors was statistically significant (Table 1; $p=6.08E-12$).

6 We queried a second microarray dataset³, to determine if we could validate differential
7 expression of Tox2 in non-small cell lung cancers. We again found significant differential
8 expression of Tox2 in NSCLC tumors of the adenocarcinoma type when compared to the normal
9 lung (Table 2). When sorting each of the transcripts measured based on significance of
10 difference in expression of Tox2 between NSCLC tumors and the normal lung, in this dataset,
11 Tox2 ranked 23 out of 25906 total transcripts (Table 2). Differential expression of Tox2 in
12 NSCLC tumors was statistically significant (Table 2; $p=3.55E-17$).

16 Analysis of a third microarray dataset⁴ again revealed significant differential expression
17 of Tox2 in NSCLC tumors of the adenocarcinoma type (Table 3). When sorting each of the
18 transcripts measured based on significance of difference in expression of Tox2 between NSCLC
19 tumors and the normal lung, Tox2 ranked 1772 out of 33252 total transcripts (Table 3).
20 Differential expression of Tox2 in NSCLC tumors was statistically significant (Table 3;
21 $p=5.32E-10$).

25 Tox2 is expressed at significantly lower levels in NSCLC tumors as compared to the lung.

26 We obtained exact mRNA levels for Tox2 from NSCLC tumors and from the lung to
27 directly compare Tox2 expression between tumor and control lung tissue and assess for statistical
28

significance. Tox2 was expressed at significantly lower levels in NSCLC tumors as compared to the normal lung in both datasets queried (Figure 1: $p < 0.0001$ and Figure 2: $p < 0.0001$). We calculated a mean fold change of 0.6941 ± 0.0676 (Table 2) in Tox2 expression when comparing NSCLC tumors to the lung.

Tox2 expression in NSCLC tumors correlates with overall survival.

We performed Kaplan-Meier survival analysis using Tox2 mRNA expression in NSCLC tumors coupled with paired overall survival data from each patient, in 1144 NSCLC patients in total, to determine whether Tox2 tumor expression was correlated with survival outcomes in NSCLC. We found that patients whose tumors expressed lower levels of Tox2 possessed significantly shorter overall survival than patients with high tumor expression of Tox2 (Figure 3). Median overall survival (OS) of patients in the low expression cohort was 41.6 months, while median OS in patients in the high Tox2 expression cohort was 102 months (Table 4); this difference in median OS based on Tox2 tumor expression in NSCLC was statistically significant (Figure 3; logrank p -value: $6.3e-12$; hazard ratio: $0.57 (0.48-0.67)$).

Thus, blind comparative transcriptome analysis of non-small cell lung cancers revealed differential expression of Tox2 as among the most significant transcriptional features of NSCLC tumors, and Tox2 expression was significantly correlated with patient outcomes, as patients with lower tumor expression of Tox2 possessed significantly worse overall survival.

Discussion

Tox2 is one of four thymocyte-selection high mobility box group transcription factors that contain a high-mobility group (HMG) box domain^{6,7}. Tox2 is less well-described than Tox. The rat homolog of Tox2, GCX-1, displays expression in the granulosa cells of the ovary⁸. The HMG domain of Tox2 displays 92% homology to the HMG domain of other Tox transcription factors⁶. Tox2 is reported to display most abundant expression in natural killer (NK) cells in humans, and its expression is induced during *in vitro* differentiation of CD34+ human umbilical cord stem cells to NK cells⁹. Tox2 could induce expression of the Th1 lineage master transcriptional regulator T-bet, and over-expression of T-bet rescued phenotypic defects resulting from depletion of Tox2, namely in NK cell maturation from early (CD117+ CD94- CD56-) to late (CD117+ CD94- CD56+) stage III NK cells⁶. In chimeric antigen receptor (CAR) tumor-infiltrating CD8+ CAR+ PD-1^{high} TIM3^{high} lymphocytes (CAR TILs), the expression of Tox2 is significantly induced⁹. CAR TILs lacking expression of both Tox and Tox2 (DKO) are significantly more effective in suppression of tumor growth and enhancement of survival in the B16-hCD19 melanoma immunocompetent solid tumor model in C57BL/6 mice⁹. Since Tox DKO CAR TIL display decreased expression of inhibitory receptors like PD-1, TIM3 and LAG3, it was suggested that Tox and Tox2 control gene expression of exhausted CD8+ tumor-infiltration lymphocytes⁹. Tox2 expression in CAR TILs was found to be a consequence of transactivation by the nuclear factor of activated T-cells, or NFAT⁹. Tox2 also displays appreciable expression in T follicular helper cells (Tfh) induced through Bcl6 and STAT3¹⁰.

1 Ectopic expression of Tox2 could drive the development of Tfh lineage cells, and Tox2-deficient
2 mice were deficient in Tfh differentiation¹⁰. Though we could not identify literature describing a
3 role for Tox2 in NSCLC, Tox3 has been reported as a favorable prognostic indicator in NSCLC
4 adenocarcinoma; mRNA expression of Tox3 was increased in lung adenocarcinomas as
5 compared to the lung and increased Tox3 tumor expression was associated with improved
6 progression-free and overall survival in NSCLC patients¹¹.

9 We found significant differential and decreased expression of the transcription factor
10 Tox2 in NSCLC tumors of the adenocarcinoma type, and lower expression of Tox2 was
11 significantly associated with worse overall survival in NSCLC patients. Tox2 may be relevant to
12 the biology of the most common type of the leading cause of cancer death in the United States
13 and worldwide.
14
15
16
17
18
19
20
21
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. 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. Lung Cancer - Non-Small Cell: Statistics. <https://www.cancer.net/cancer-types/lung-cancer-non-small-cell/statistics>.
6. Vong, Q.P., Leung, W.H., Houston, J., Li, Y., Rooney, B., Holladay, M., Oostendorp, R.A. and Leung, W., 2014. TOX2 regulates human natural killer cell development by controlling T-BET expression. *Blood, The Journal of the American Society of Hematology*, 124(26), pp. 3905-3913.
7. Tessema, M., Yingling, C.M., Grimes, M.J., Thomas, C.L., Liu, Y., Leng, S., Joste, N. and Belinsky, S.A., 2012. Differential epigenetic regulation of TOX subfamily high mobility group box genes in lung and breast cancers. *PloS one*, 7(4), p.e34850.
8. Kajitani, T., Mizutani, T., Yamada, K., Yazawa, T., Sekiguchi, T., Yoshino, M., Kawata, H. and Miyamoto, K., 2004. Cloning and characterization of granulosa cell high-mobility group (HMG)-box protein-1, a novel HMG-box transcriptional regulator strongly expressed in rat ovarian granulosa cells. *Endocrinology*, 145(5), pp.2307-2318.
9. Seo, H., Chen, J., González-Avalos, E., Samaniego-Castruita, D., Das, A., Wang, Y.H., López-Moyado, I.F., Georges, R.O., Zhang, W., Onodera, A. and Wu, C.J., 2019. TOX and TOX2 transcription factors cooperate with NR4A transcription factors to impose CD8+ T cell exhaustion. *Proceedings of the National Academy of Sciences*, 116(25), pp.12410-12415.

- 1 10. Xu, W., Zhao, X., Wang, X., Feng, H., Gou, M., Jin, W., Wang, X., Liu, X. and Dong, C.,
2 2019. The transcription factor Tox2 drives T follicular helper cell development via regulating
3 chromatin accessibility. *Immunity*, 51(5), pp.826-839.
- 4 11. Zeng, D., Lin, H., Cui, J. and Liang, W., 2019. TOX3 is a favorable prognostic indicator and
5 potential immunomodulatory factor in lung adenocarcinoma. *Oncology Letters*, 18(4), pp.
6 4144-4152.

Rank	ID	p-value	t	B	Gene
148	A_23_P154566	6.08E-12	-1.11E+01	17.3251805	TOX2

Table 1: Tox2 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
23	228737_at	3.55E-17	-16.840245	29.1438266	0.6941 ± 0.0676	TOX2	TOX high mobility group box family member 2

Table 2: Tox2 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 Tox2 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
1772	8062782	5.32E-10	-6.805915	12.283857	TOX2	TOX high mobility group box family member 2

Table 3: Tox2 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.

Figure 1: Tox2 is expressed at significantly lower levels in NSCLC tumors when compared to the lung.

The mRNA expression level of Tox2 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 Tox2 expression between NSCLC tumors and the lung, a *p*-value, listed above.

Figure 2: Tox2 is expressed at significantly lower levels in NSCLC tumors when compared to the lung.

The mRNA expression level of Tox2 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 Tox2 expression between NSCLC tumors and the lung, a *p*-value, listed above.

Figure 3: Tox2 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 Tox2 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 Tox2 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.

Table 4: Tox2 expression in NSCLC tumors significantly correlates with patient survival.

Low expression cohort (months)	High expression cohort (months)
41.6	102

The median overall survival of $n=1144$ NSCLC patients based on stratification into low or high expression of Tox2 in tumors is listed in this chart.