Expression of desmocollin 3 in non-small cell lung cancer distinguishes squamous cell lung carcinoma from adenocarcinoma.

Shahan Mamoor, MS¹

¹shahanmamoor@gmail.com

East Islip, NY USA

Non-small cell lung cancer, classified as adenocarcinomas or squamous cell lung carcinomas, is the major cause of cancer death in the United States and worldwide¹. To understand the most significant transcriptional differences between adenocarcinomas and squamous cell lung carcinomas, we mined published microarray data from two separate studies^{2,3}. We identified desmocollin 3 (DSC3) as a distinguishing transcriptional feature of squamous cell lung carcinomas, suggesting the biology of DSC3 may be relevant to the pathways critical for the development or maintenance of squamous cell lung carcinomas but not of adenocarcinomas. DSC3 expression was significantly correlated with prognosis of patients with NSCLC, as patients with low tumor expression of DSC3 possessed significantly longer median overall survival than those with high tumor expression of DSC3. These analyses will also provide novel tools for diagnostic approaches and for guidance of treatment regimens for a cancer with dismal outlook.

Keywords: desmocollin 3, DSC3, non-small cell lung cancer, adenocarcinoma, squamous cell lung carcinoma, systems biology of NSCLC, targeted therapeutics in NSCLC, targeted diagnostics in NSCLC.

1 Non-small cell lung cancer is the single most common cancer in the United States and
2 the most common cause of cancer death in the United States¹. The two major types of non-
small cell lung cancer are adenocarcinoma and squamous cell lung carcinoma⁴.

3 Treatment approaches and prognosis can differ based on whether a patient is diagnosed
4 with adenocarcinoma or squamous cell lung carcinoma⁴, and current approaches to
distinguishing adenocarcinoma from squamous cell lung carcinoma rely on time-consuming
5 procedures involving immunohistochemical staining with multiple markers including TTF1, p40,
6 p63, CK5 and/or CK7 and subsequent readout by pathologists who will proceed to label each
section, in a binary fashion, as positive or negative for each marker⁵. To understand the
7 molecular nature of ACC and SCLC tumors in an unbiased fashion and at the systems level,
and to facilitate discovery of genes with utility in NSCLC diagnostics in a quantitative real-time
8 workflow providing a quantitative and independent measure rather than qualitative measure
based on pathologist interpretation, we performed comparative transcriptome analysis^{2,3} of the
9 two major types of NSCLC tumors: adenocarcinomas and squamous cell lung carcinomas.

10 This blind, systems-level approach identified the gene encoding desmocollin 3, DSC3,
as among the genes most differentially expressed when comparing adenocarcinomas and
11 squamous cell lung carcinoma tumors, with DSC3 expression distinguishing squamous cell lung
carcinomas from adenocarcinomas, suggesting DSC3 could be a useful tool in NSCLC
12 diagnostics.

13 **Methods**

14
15 We utilized datasets GSE74706² and GSE33532³ for this comparative transcriptome
analysis of adenocarcinomas and squamous cell lung carcinomas. GSE74706 was generated
16 using Agilent-026652 Whole Human Genome Microarray 4x44K v2 technology; for this analysis,
we used $n=10$ adenocarcinomas and $n=4$ squamous cell lung carcinomas, and the analysis was
17 performed using platform GPL13497. GSE33532 was generated using Affymetrix Human
Genome U133 Plus 2.0 Array technology; for this analysis, we used $n=10$ adenocarcinoma
18 tumors and $n=8$ squamous cell lung carcinomas, and the analysis was performed using platform
GPL570.

19
20 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
21 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
22 evaluate whether DSC3 expression was significantly between adenocarcinoma and squamous
cell carcinoma using a two-tailed, unpaired t-test with Welch's correction. We used PRISM for
23 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
24 of DSC3 mRNA expression levels with overall survival in $n=1925$ non-small cell lung cancer
patients.
25
26
27
28

Results

We mined published microarray data from two separate studies^{2,3} to describe in an unbiased fashion and at the systems-level genes whose expression was most specific to either one of the two major types of non-small cell lung cancer: adenocarcinomas and squamous cell lung carcinomas.

Desmocollin 3 expression distinguishes squamous cell lung carcinomas from adenocarcinomas.

By comparing the global gene expression profiles of squamous cell lung carcinomas to adenocarcinomas, we found that desmocollin 3 was among the most differentially expressed gene between the two major sub-types of NSCLC (Table 1). When sorting each of the genes expression by microarray based on significance of change in expression between adenocarcinomas and squamous cell lung carcinomas, desmocollin 3 ranked out of 34183 total transcripts (Table 1). Desmocollin 3 differential expression between adenocarcinomas and squamous cell lung carcinomas was statistically significant (Table 1; $p=1.05E-11-42$).

In a separate dataset, desmocollin 3 was again amongst the most differentially expressed genes when comparing squamous cell lung carcinomas to adenocarcinomas in NSCLC (Table 2). When sorting each of the genes expression by microarray based on significance of change in expression between adenocarcinomas and squamous cell lung carcinomas, desmocollin 3 ranked 4 out of 25906 total transcripts (Table 2). Desmocollin 3 differential expression between adenocarcinomas and squamous cell lung carcinomas was again statistically significant (Table 2; $p=9.9E-42$).

Desmocollin 3 is expressed at significantly higher levels in squamous cell lung carcinomas than in adenocarcinomas.

We obtained exact mRNA values for desmocollin 3 to understand the magnitude and direction of difference in DSC3 mRNA expression between adenocarcinomas and squamous cell lung carcinomas in NSCLC. Desmocollin 3 was expressed at significantly higher levels in squamous cell lung carcinomas as compared to adenocarcinomas, and this difference was statistically significant (Figure 1 and Figure 2: $p<0.0001$ and $p=0.0003$, respectively). We calculated a mean fold change of 2.8274 ± 0.1979 in DSC3 expression when comparing squamous cell lung carcinomas to adenocarcinomas (Table 2).

Expression of desmocollin 3 is associated with patient survival in NSCLC.

We performed Kaplan Meier survival analysis to determine if desmocollin 3 expression was correlated with patient outcomes in NSCLC. We found significant correlation between expression of desmocollin 3 and overall survival in patients with NSCLC (Figure 3 and Table 3). While median overall survival was 85 months for NSCLC patients with low expression of DSC3, median overall survival was 54.3 months for NSCLC patients with high expression of DSC3. Correlation of desmocollin 3 tumor expression with median OS with statistically significant (Figure 3; log rank p -value: $2.2e-07$; hazard ratio: 1.4 (1.23-1.58)).

Thus, by mining published microarray data from independent datasets, we found that desmocollin 3 was among the genes whose expression most significantly distinguished squamous cell lung carcinomas from adenocarcinomas in NSCLC. Moreover, we found a statistically significant correlation between DSC3 tumor expression at the mRNA level and patient survival in NSCLC.

Discussion

To understand the most striking transcriptional features of each of the two major tumor types in NSCLC, we performed comparative transcriptome analysis of ACC and SCLC tumors using published microarray data, discovering desmocollin 3 as amongst the most distinguishing transcriptional features of squamous cell lung carcinomas.

Desmosomal cadherins are proteins that function in cell-cell adhesion of epithelial cells and including desmogleins and desmocollins⁸. Desmocollin 3 resides on 18q21.1 and is comprised of 17 exons. Mice with conditional deletion of desmocollin 3 in the skin using a Keratin 14-Cre display defects in cell-cell adhesion, blistering layers of the epidermis and loss of telogen hair⁹. Mice null for desmocollin 3 are not viable, and desmocollin 3 deletion results in lethality before the formation of desmosomes, suggesting desmocollin 3 is required for development of the embryo prior to implantation¹⁰. In non-small cell lung cancer cell lines, p53 could control the expression of DSC3 depending on the methylation status of the DSC3 gene, as ectopic expression of p53 in a DSC3-non-methylated NSCLC cell line (H2170) could induce expression of p53, but not in a DSC3-methylated NSCLC cell line (H1299)¹¹. Ectopic expression of DSC3 in a NSCLC cell line resulted in significant decreases in anchorage-independent cell growth in soft agar, invasion in transwell migration assays, invasion through a matrigel membrane, as well as cell proliferation in both H2170 and H1299 cell lines¹¹. In a study using immunohistochemical techniques to stain NSCLC tumors found virtually exclusive staining in squamous cell lung cancer tissues as compared to adenocarcinomas, but in 50% of large-cell carcinomas (LCC), and thus investigators assigned squamous differentiation features to anaplastic LCCs¹². In a separate study, investigators demonstrated that expression of DSC3 could be restored in four of seven lung cancer cell lines following 5-aza-2'-deoxycytidine treatment, supporting the notion that in lung cancer, DSC3 expression could be repressed by mechanisms involving DNA methylation¹³. At the protein level, tissue microarrays demonstrated higher expression of DSC3 in squamous cell carcinomas as compared to adenocarcinomas; however, this study described a correlation between low expression of DSC3 and worse clinical outcome, contrary to our findings that median relapse-free survival of NSCLC patients with low tumor expression of DSC3 was more than 2.5 years greater than that of patients with high tumor expression of DSC3¹³.

We found that desmocollin 3 was amongst the most differentially expressed genes when comparing the tumors of patients with the two most common sub-types of NSCLC: adenocarcinomas and squamous cell lung carcinomas. We also found significant correlation between overall survival of NSCLC patients and desmocollin 3 expression. Desmocollin 3 has value as a diagnostic tool, as a prognostic indicator, and the biology of desmocollin 3 may be of value in understanding fundamental differences between the two major types of the most common type of cancer, and the most common 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. Marwitz, S., Depner, S., Dvornikov, D., Merkle, R., Szczygiał, M., Müller-Decker, K., Lucarelli, P., Wäsch, M., Mairböurl, 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. Cetin, K., Ettinger, D.S., Hei, Y.J. and D O'Malley, C., 2011. Survival by histologic subtype in stage IV nonsmall cell lung cancer based on data from the Surveillance, Epidemiology and End Results Program. *Clinical epidemiology*, 3, p.139.
5. Yuan, M., Huang, L.L., Chen, J.H., Wu, J. and Xu, Q., 2019. The emerging treatment landscape of targeted therapy in non-small-cell lung cancer. *Signal Transduction and Targeted Therapy*, 4(1), pp.1-14.
6. Osmani, L., Askin, F., Gabrielson, E. and Li, Q.K., 2018, October. Current WHO guidelines and the critical role of immunohistochemical markers in the subclassification of non-small cell lung carcinoma (NSCLC): Moving from targeted therapy to immunotherapy. In *Seminars in cancer biology* (Vol. 52, pp. 103-109). Academic Press.
7. 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.
8. Whittock, N.V., Hunt, D.M., Rickman, L., Malhi, S., Vogazianou, A.P., Dawson, L.F., Eady, R.A., Buxton, R.S. and McGrath, J.A., 2000. Genomic organization and amplification of the human desmosomal cadherin genes DSC1 and DSC3, encoding desmocollin types 1 and 3. *Biochemical and biophysical research communications*, 276(2), pp.454-460.
9. Chen, J., Den, Z. and Koch, P.J., 2008. Loss of desmocollin 3 in mice leads to epidermal blistering. *Journal of cell science*, 121(17), pp.2844-2849.
10. Den, Z., Cheng, X., Merched-Sauvage, M. and Koch, P.J., 2006. Desmocollin 3 is required for pre-implantation development of the mouse embryo. *Journal of cell science*, 119(3), pp. 482-489.
11. Cui, T., Chen, Y., Yang, L., Knösel, T., Huber, O., Pacyna-Gengelbach, M. and Petersen, I., 2012. The p53 target gene desmocollin 3 acts as a novel tumor suppressor through inhibiting EGFR/ERK pathway in human lung cancer. *Carcinogenesis*, 33(12), pp.2326-2333.

- 1 12.Monica, V., Ceppi, P., Righi, L., Tavaglione, V., Volante, M., Pelosi, G., Scagliotti, G.V. and
2 Papotti, M., 2009. Desmocollin-3: a new marker of squamous differentiation in
3 undifferentiated large-cell carcinoma of the lung. *Modern Pathology*, 22(5), pp.709-717.
- 4 13.Cui, T., Chen, Y., Yang, L., Mireskandari, M., Knösel, T., Zhang, Q., Kohler, L.H., Kunze, A.,
5 Presselt, N. and Petersen, I., 2012. Diagnostic and prognostic impact of desmocollins in
6 human lung cancer. *Journal of clinical pathology*, 65(12), pp.1100-1106.
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20
- 21
- 22
- 23
- 24
- 25
- 26
- 27
- 28

Rank	ID	p-value	t	B	Gene
5	A_24_P344416	1.05E-11	14.484319	15.7445	DSC3

Table 1: Desmocollin 3 is among the most differentially expressed genes when comparing whole transcriptomes of adenocarcinomas and squamous cell lung carcinomas, the two major types of non-small cell lung cancer.

Rank of differential expression, probe ID, p -value with respect to differential expression, t , a moderated t -statistic, B , the log-odds of differential expression between the two groups compared, and gene are listed in this chart.

Rank	ID	p-value	t	B	FC	Gene	Gene name
4	206033_s_at	9.9E-42	37.3207777	82.257942	2.8274 ± 0.1979	DSC3	desmocollin 3

Table 2: Desmocollin 3 is among the most differentially expressed genes when comparing whole transcriptomes of adenocarcinomas and squamous cell lung carcinomas, the two major types of non-small cell lung cancer.

Rank of differential expression, probe ID, *p*-value with respect to differential expression, *t*, a moderated *t*-statistic, *B*, the log-odds of differential expression between the two groups compared, fold change of DSC3 when comparing squamous cell lung carcinomas to adenocarcinomas, gene and gene name are listed in this chart.

Figure 1: Expression of desmocollin 3 distinguishes squamous cell lung carcinoma from adenocarcinoma.

Messenger RNA (mRNA) levels of desmocollin 3 in the tumors of patients with adenocarcinoma (“ACC”; left) and in the tumors of patients with squamous cell lung carcinoma (“SC”; right) are graphically represented here with mean mRNA levels marked and the result of a statistical test evaluating significance of difference in mRNA expression between the tumors of patients with adenocarcinomas and squamous cell lung carcinomas, a *p*-value, listed above.

Figure 2: Expression of desmocollin 3 distinguishes squamous cell lung carcinoma from adenocarcinoma.

Messenger RNA (mRNA) levels of desmocollin 3 in the tumors of patients with adenocarcinoma (“ACC”; left) and in the tumors of patients with squamous cell lung carcinoma (“SC”; right) are graphically represented here with mean mRNA levels marked and the result of a statistical test evaluating significance of difference in mRNA expression between the tumors of patients with adenocarcinomas and squamous cell lung carcinomas, a *p*-value, listed above.

Figure 3: Desmocollin 3 expression in the tumors of patients with NSCLC 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 DSC3 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 DSC3 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)
85	54.3

Table 3: NSCLC patients with low tumor expression of DSC3 possess significantly greater overall survival than patients with high tumor expression of DSC3.

The median overall survival of $n=1925$ NSCLC patients based on stratification into low or high expression of DSC3 in tumors is listed in this chart.