

Article

Dominant Gene Expression Profiles Define Adenoid Cystic Carcinoma (ACC) from Different Tissues: Validation of a Gene Signature Classifier for Poor Survival in Salivary Gland ACC

Kathryn J. Brayer ^{1,2}, Huining Kang ^{2,3} , Adel K. El-Naggar ⁴ , Simon Andreassen ^{5,†}, Preben Homøe ⁵, Katalin Kiss ⁶, Lauge Mikkelsen ^{6,7}, Steffen Heegaard ⁷ , Daniel Pelaez ^{8,9,10} , Acadia Moeyersoms ^{8,9,10,11}, David T. Tse ^{8,9,10}, Yan Guo ^{1,2} , David Y. Lee ^{2,12} and Scott A. Ness ^{1,2,*} 

- ¹ Department of Internal Medicine, Division of Molecular Medicine, University of New Mexico School of Medicine, Albuquerque, NM 87131, USA
 - ² University of New Mexico Comprehensive Cancer Center, Albuquerque, NM 87131, USA
 - ³ Department of Internal Medicine, Division of Epidemiology, University of New Mexico School of Medicine, Albuquerque, NM 87131, USA
 - ⁴ Department of Pathology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
 - ⁵ Department of Otorhinolaryngology and Maxillofacial Surgery, Zealand University Hospital, 4600 Køge, Denmark
 - ⁶ Department of Pathology, Rigshospitalet, University of Copenhagen, 1165 Copenhagen, Denmark
 - ⁷ Department of Ophthalmology, Rigshospitalet-Glostrup, University of Copenhagen, 1165 Copenhagen, Denmark
 - ⁸ Dr. Nasser Al-Rashid Orbital Vision Research Center, Department of Ophthalmology, Bascom Palmer Eye Institute, University of Miami Miller School of Medicine, Miami, FL 33136, USA
 - ⁹ Sylvester Comprehensive Cancer Center, University of Miami Miller School of Medicine, Miami, FL 33136, USA
 - ¹⁰ Department of Ophthalmology, Bascom Palmer Eye Institute, University of Miami Miller School of Medicine, Miami, FL 33136, USA
 - ¹¹ The Sheila and David Fuente Graduate Program in Cancer Biology, University of Miami Miller School of Medicine, Miami, FL 33136, USA
 - ¹² Department of Internal Medicine, Division of Hematology/Oncology, Section of Radiation Oncology, University of New Mexico School of Medicine, Albuquerque, NM 87131, USA
- * Correspondence: sness@salud.unm.edu
† Deceased.



Citation: Brayer, K.J.; Kang, H.; El-Naggar, A.K.; Andreassen, S.; Homøe, P.; Kiss, K.; Mikkelsen, L.; Heegaard, S.; Pelaez, D.; Moeyersoms, A.; et al. Dominant Gene Expression Profiles Define Adenoid Cystic Carcinoma (ACC) from Different Tissues: Validation of a Gene Signature Classifier for Poor Survival in Salivary Gland ACC. *Cancers* **2023**, *15*, 1390. <https://doi.org/10.3390/cancers15051390>

Academic Editors: Riccardo Dolcetti and Chamindie K. Punyadeera

Received: 28 November 2022

Revised: 11 February 2023

Accepted: 17 February 2023

Published: 22 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Simple Summary: Adenoid cystic carcinoma (ACC) is a pathologically distinctive tumor that most often occurs in major or minor salivary glands, but can also occur in other tissues. We compared the gene expression profiles of ACC tumor samples from salivary gland, lacrimal gland, breast or skin. Despite their different tissues of origin, the ACC tumors displayed highly related patterns of gene expression. Indeed, gene expression patterns could not distinguish ACC tumors from different tissues, suggesting that genetic and epigenetic regulatory events induce a dominant ACC ‘phenotype’. We also used the new cohort of salivary gland ACC tumors to validate a gene expression biomarker developed with a previously analyzed cohort. The 49-gene classifier correctly identified 98% of the poor survival patients, validating the biomarker and suggesting that a clinical test should be developed so patients at highest risk of poor survival can be identified and provided additional treatment.

Abstract: Adenoid cystic carcinoma (ACC) is an aggressive malignancy that most often arises in salivary or lacrimal glands but can also occur in other tissues. We used optimized RNA-sequencing to analyze the transcriptomes of 113 ACC tumor samples from salivary gland, lacrimal gland, breast or skin. ACC tumors from different organs displayed remarkably similar transcription profiles, and most harbored translocations in the *MYB* or *MYBL1* genes, which encode oncogenic transcription factors that may induce dramatic genetic and epigenetic changes leading to a dominant ‘ACC phenotype’. Further analysis of the 56 salivary gland ACC tumors led to the identification of three distinct groups of patients, based on gene expression profiles, including one group with worse survival. We tested whether this new cohort could be used to validate a biomarker developed previously with a different

set of 68 ACC tumor samples. Indeed, a 49-gene classifier developed with the earlier cohort correctly identified 98% of the poor survival patients from the new set, and a 14-gene classifier was almost as accurate. These validated biomarkers form a platform to identify and stratify high-risk ACC patients into clinical trials of targeted therapies for sustained clinical response.

Keywords: oral cancer; biomarker; MYB oncogene; transcriptome analysis; bioinformatics; survival analysis

1. Introduction

Adenoid cystic carcinoma (ACC) is one of the most common salivary gland malignancies, arising mainly in minor and major salivary glands, but ACC also occurs less frequently in other organs, and the clinical behavior of non-salivary ACC varies widely. This suggests that the ACC tumors arising in different organs may be biologically distinct or that they are affected by different host factors. Molecular analyses have shown that most ACC tumors have recurrent chromosomal translocations that activate the *MYB* oncogene or the related *MYBL1* gene [1–4], resulting in characteristic gene expression changes [5,6]. The translocations frequently relocate a distant, salivary gland-specific enhancer in proximity to the *MYB* or *MYBL1* genes, leading to their overexpression [7]. Many of the translocations occur within the *MYB* or *MYBL1* genes, leading to truncation and overexpression of the genes and their gene products [5,6]. The *MYB* and *MYBL1* genes encode the DNA-binding transcription factor proteins Myb (c-Myb) and A-Myb, which are important for normal development [8,9]. Relatively small changes in these proteins, such as truncations of the N- or C-terminal domains, can lead to profound differences in the genes they regulate [10–13], suggesting that the proteins perform complex regulatory functions. Indeed, the Myb protein can function as a ‘pioneer’ transcription factor capable of initiating the formation of new enhancers that can modify the expression of distant genes [14–16]. Thus, Myb or A-Myb proteins activated by C-terminal truncations may induce a specific ACC tumor phenotype [17], similar to the actions of Myb proteins in other types of cells and malignancies [18–24]. Identifying the ACC-specific regulatory mechanisms that induce an ‘ACC phenotype’ could lead to new types of therapies.

ACC patients often have a slow clinical course with a poor long-term prognosis [4,25]. However, clinical outcomes can vary dramatically; unpredictable aggressive and progressive disease is not uncommon. Post-surgical survival ranges from just a few months to 15 years or longer. The protracted temporal progression of ACC tumors necessitates using archived samples at least 5–10 years old for studies linking genomic changes to outcomes. However, standard genomic methods are largely unsuitable for reliable RNA-sequencing (RNA-seq) analysis of archived samples, because the recovered RNA is often highly fragmented, necessitating the use of specialized approaches [5,6,26]. Despite these complications, several studies have identified subgroups of ACC patients with distinct molecular features linked to differences in prognosis and survival [5,6,27–30], suggesting that applying these approaches to well-structured retrospective cohorts of ACC tumors could produce biomarkers for identifying poor prognosis patients and recommending them for targeted therapy.

In previous studies, we were able to use optimized RNA-seq approaches to successfully analyze the transcriptomes of archived, formalin-fixed, paraffin-embedded ACC tumor samples up to 25 years old, which led to the identification of the first ACC tumors with *MYBL1* translocations [5]. Extending those studies to a larger cohort of 68 samples (the TX cohort) led to the identification of several subgroups of ACC tumors with unique gene expression signatures, including one subgroup with poor survival and a ‘No Myb’ group that expressed neither *MYB* nor *MYBL1* [6]. Although we identified gene expression signatures that correlated with poor survival, it was not possible to validate the results using only one cohort of samples. Here we describe the analysis of a new cohort of ACC tumor samples, from Denmark and Florida (the DK cohort), primarily from salivary gland but

also including some ACC tumors from the lacrimal gland, breast, and skin. The availability of the large cohorts allowed us to designate a training set for defining a gene expression classifier to distinguish poor survival samples, which we validated using the second cohort. These results set the stage for using clinical RNA-seq assays for identifying patients who are likely to be in the poor survival subgroup, so they can be offered clinical trials or additional treatment to improve their outcomes.

2. Materials and Methods

2.1. Human Salivary Gland ACC Samples

De-identified adenoid cystic carcinoma tumor samples were obtained from several institutions: the Department of Otorhinolaryngology and Maxillofacial Surgery, Zealand University Hospital; the Department of Otorhinolaryngology, Head and Neck Surgery and Audiology, Rigshospitalet; the Department of Pathology, Rigshospitalet, University of Copenhagen; and the Department of Ophthalmology, Rigshospitalet-Glostrup, University of Copenhagen, Copenhagen, Denmark. Some lacrimal gland samples were obtained from the Dr. Nasser Al-Rashid Orbital Vision Research Center and the Bascom Palmer Eye Institute, Department of Ophthalmology, University of Miami Miller School of Medicine. All samples were provided Formalin-Fixed and Paraffin-Embedded (FFPE) as 5-micron sections baked onto glass slides. Salivary gland samples with survival information had at least 5-year follow-up. All samples were collected in accordance with the principle of the Declaration of Helsinki and with Institutional Review Board-approved protocols: Danish Regional Ethics Committee (H-6-2014-086) and the Danish Data Protection Agency (Journal no. REG-94-2014).

2.2. RNA Isolation and Sequencing

Total RNA was isolated from one or two 5-micron slide-mounted FFPE sections using the RNeasy FFPE kit (Qiagen, 19300 Germantown Rd Germantown MD 20874, USA). cDNA synthesis and library preparation were performed using the SMARTer Universal Low Input RNA Kit for Sequencing (Takara 1290 Terra Bella Avenue Mountain View, CA 94043, USA) and the Ion Plus Fragment Library Kit (ThermoFisher, 168 Third Avenue, Waltham, MA 02451, USA), as previously described [5,6]. Ion Torrent sequencing using Ion S5/XL systems (ThermoFisher) was performed in the Analytical and Translational Genomics Shared Resource at the University of New Mexico Comprehensive Cancer Center. Data are available for download from the NCBI BioProject database using study accession number PRJNA287156. The TX cohort of samples has been described previously [6].

2.3. Data Analysis

Low quality and non-human RNA-seq reads were filtered and removed using the Kraken2 suite [31–33]. High-quality reads were aligned to the human genome (hg38) using TMAP (v5.10.11), and gene counts were calculated using HT-Seq, as previously described [5,6]. Poor quality samples with fewer than 10% of the median number of reads of all samples were removed. Samples that failed other quality control tests were also removed. The same parameters were used when the data from the new (DK) cohort were combined with the previously described samples in the TX cohort [6]. (Software versions are provided in File S1).

2.4. Unsupervised Hierarchical Clustering

For identifying clusters, analyses were limited to genes that were expressed above a threshold level in a number of samples (e.g., 250 reads in at least 10 samples). These thresholds were reduced (e.g., to 50 reads in at least 10 samples) to generate the final heatmaps, to include as many relevant genes as possible, while retaining the clusters and the sample order. Multi-dimensional scaling was performed using plotMDS from the limma package version 3.48.0. Hierarchical clustering was performed using hclust from the stats package (R/Bioconductor version 4.1.0) [34].

2.5. Statistical Analysis

Overall survival (OS), defined as time from the date of diagnosis to the date of death, was the primary endpoint for outcome. Subjects who were lost to follow-up or alive within the follow-up period were censored at the date of the last contact. OS was estimated using the Kaplan–Meier method. Differences in OS were examined using the log-rank test. All the statistical analyses were performed using statistical software R (version 4.1.0) [34]. Gene Ontology analyses were performed with Bioconductor package GO.db version 3.13.0, as previously described [5,6,10,26].

2.6. Development and Evaluation of Prognostic Classifiers

Personalized logistic regression with the elastic net and LASSO regulations implemented in R package glmnet [35] was used to develop the classifiers that distinguish between the poor prognosis group and the remaining samples. ROC curve and the area under the curve (AUC) were used to evaluate the performance of the classifiers. The prognostic models were built using the training data set, i.e., the TX cohort, through a 10-fold cross-validation (CV) procedure. An unbiased estimate for each of the final models was obtained by performing a nested CV procedure that included the full cycle of the 10-fold inner loop CV followed by a 100 × 6-fold outer loop CV using the training set. The prediction accuracy of each model was further validated using an independent test set, i.e., the DK cohort.

3. Results

3.1. ACC Tumors from Different Organ Sites Share a Common Transcriptional Profile

Although they most commonly occur in major and minor salivary glands, ACC tumors can also arise in lacrimal, bronchial, mammary, or skin adnexal glands. To assess the similarities and differences in gene expression patterns in ACC tumors from different tissues, we performed RNA-sequencing (RNA-seq) on a cohort of 113 ACC tumors, comprised of 17 samples from breast tissue, 24 from cutaneous tissue, 16 from lacrimal glands, and 56 from salivary glands (Table 1). Most of the samples came from Denmark (DK cohort), with the exception of 6 lacrimal gland samples from Florida (FL). We used optimized methods for RNA analysis and Ion Proton sequencing that we developed previously [5,26]. For the DK cohort, 91 samples (81%) clearly expressed *MYB*, while 13 (12%) expressed *MYBL1* and 9 (8%) expressed neither *MYB* nor *MYBL1*.

Table 1. ACC Tumor Cohorts.

Tissue	Number	Form	Gender	Stage	Age at Surgery	Source	Oncogene *
New Cohort (DK)							
Breast	17	NA	NA	NA	NA	DK	<i>MYB</i> : 16 <i>MYBL1</i> : 1
Cutaneous	24	NA	NA	NA	NA	DK	<i>MYB</i> : 17 <i>MYBL1</i> : 3 No_ <i>MYB</i> : 4
Lacrimal	16	NA	NA	NA	NA	DK: 10 FL: 6	<i>MYB</i> : 11 <i>MYBL1</i> : 2 No_ <i>MYB</i> : 3
Salivary	56	Solid: 11 Tubulocribiform: 43 NA: 2	F: 29 M: 27	Stage I-II: 30 Stage III-IV: 26	Age: 32–87 NA: 17	DK	<i>MYB</i> : 47 <i>MYBL1</i> : 7 No_ <i>MYB</i> : 2
Cohort Total	113						<i>MYB</i> : 91 <i>MYBL1</i> : 13 No_ <i>MYB</i> : 9

Table 1. Cont.

Tissue	Number	Form	Gender	Stage	Age at Surgery	Source	Oncogene *
Previous Cohort (TX, Frerich et al. 2018 [6])							
Salivary	68	NA	F: 30 M: 38	NA	NA	TX	MYB: 52 MYBL1: 7 No_MYB: 9
Combined	181						MYB: 143 MYBL1: 20 No_MYB: 18

* No MYB indicates no expression of either MYB or MYBL1.

We first performed Multi-Dimensional Scaling (MDS, i.e., principal component analysis) for the ACC samples. As shown in Figure 1A, the dots representing ACC tumor samples from different organs are shaded with different colors (see legend). For comparison, we included RNA-seq results from several normal salivary gland tissues (shaded black) and several from a histologically different salivary gland tumor, acinic cell carcinoma (shaded gray), that have been described previously [5,28]. All the ACC tumor samples clustered into a large group at the right, suggesting that the transcription profiles of the ACC tumors are more similar to each other than they are to normal tissue or other salivary gland tumors, despite the different tissues of origin.

Next, we used unsupervised hierarchical clustering to group samples that were similar and generated the heatmap comparing the transcriptional profiles shown in Figure 1B. Interestingly, as shown by the dendrogram at the top of the heatmap, the ACC tumor samples formed several large subgroups, but each cluster contained samples from all the tissue types: salivary gland, lacrimal gland, breast and cutaneous (dark blue, orange, pink and green in the color bar at top, respectively). Thus, although the ACC samples in this cohort are heterogeneous and formed distinct subgroups, the groups are not defined by the tissue of origin. Instead, the subgroups could represent biological differences amongst the ACC samples, irrespective of the tissue from which they were derived. These results differ somewhat from a previous report showing that microRNA expression profiles could distinguish ACC tumors from different tissue types [36], suggesting that the biological mechanisms leading to the formation of ACC tumors may have a more profound impact on the overall mRNA transcriptional profile than on microRNAs, which may be more tissue-specific. Several genes that are known to be important in ACC tumors are highlighted with black dots at right, including (from top to bottom) *EN1*, *GABRP*, *MYB*, *MYBL1* and *NFIB*, all of which are expressed more highly in the ACC tumors than in the normal salivary gland or acinic cell carcinoma samples (at left).

3.2. Many ACC Tumors Fail to Express Tissue-Specific Gene Expression Markers

Since our initial hierarchical clustering did not separate ACC tumors by tissue type, we reanalyzed the data to see if we could identify tissue-specific gene expression patterns in the tumors. We specifically selected genes that were differentially expressed in the ACC tumors originating in different tissues. The heatmap in Figure 2 summarizes the results when the most tissue-specific genes are chosen for display (and the normal salivary gland and acinic cell carcinoma samples are left out). This type of analysis led to better clustering of the ACC samples from lacrimal gland (orange, left), salivary gland (blue), cutaneous (green) and breast (pink). There were specific genes that were up-regulated in some ACC samples compared to the others. For example, at the far left of the heatmap (labeled A at bottom, orange color bar at top) is a group of ACC tumors, mostly from the lacrimal gland, that overexpress the *OPRPN* gene, which encodes the Opiorphin Proline-Rich Lacrimal Protein 1 (previously named *PROL1*). The next group (labeled B) are salivary ACC (blue color bar at top) and overexpress several salivary-gland specific genes including *MUC19*, *CA6*, *MUC7*, *SMR3B*, *LPO*, *BPIFA2*, *CST5*, *CST2*, *CST1* and *CST4* (marked by blue dots at

lower right of figure). Group C also contains salivary ACC samples with a few from other tissues, and several overexpress *KRT4* and *KRT13*. Most of the cutaneous samples clustered in a group (D, green color bar) and are identified by overexpression of *FLG*, *KRT10*, *FLG2*, *KRT6A* and *KRT1*. However, the remaining ACC tumors formed a large cluster (labeled E), which included most of the breast ACC tumors as well as samples from salivary, lacrimal and cutaneous adnexal glands. The last group was notable because the samples failed to express the gland-specific marker genes that defined the other groups, suggesting that they had a more de-differentiated or perhaps more stem cell-like phenotype. Thus, while we were able to identify tissue-specific marker genes in some ACC tumors, the specificity was not absolute and there remained significant heterogeneity in the gene expression patterns of different samples. Also, since tumor samples always contain some normal cells, the tissue-specific differences that were detected could be due to the non-tumor cells in the samples. We conclude that an ACC-specific gene expression pattern dominated the tumors, apparently overriding the tissue-specific differences.

3.3. The New DK Cohort of ACC Tumors Also Contains a Poor-Survival Subgroup of Patients

Previous studies of ACC tumor samples identified subgroups of tumors with distinct gene expression and survival characteristics [6]. We carefully evaluated the new cohort of salivary gland ACC tumors for evidence of subgroups with distinct gene expression patterns. Figure 3A shows a multi-dimensional scaling plot of the 56 DK cohort salivary gland samples. Most of the tumors (shaded light blue) form a large cluster but a small group of tumor samples (brown) formed a separate group at the upper left corner of the plot. Figure 3B shows a Kaplan–Meier survival analysis: the samples in the brown group had a median survival of only 8 months, compared to the main group (light blue), which showed a median survival of 80 months (p -value = 0.006). This is reminiscent of our previous results with the TX cohort, where a subgroup of ACC tumors displayed similarly poor survival [6]. Some ACC tumors display a ‘solid form’ morphology, which has been associated with worse prognosis [37,38]. Although ‘solid morphology’ ACC tumors were excluded from the TX cohort, the DK cohort contains 11 such samples: 5 in the poor prognosis group and 6 in other group. This suggests that the poor prognosis group is not defined simply by solid tumor morphology.

The differential gene expression analysis identified 273 genes at least 2-fold up- or down-regulated in the brown subgroup (adjusted p -value < 0.05). The results for 85 of the genes are summarized in the heatmap in Figure 3C (The poor-survival subgroup cluster is at the left side of the heatmap, marked by the brown color bar at the top). In the heatmap, all the genes that were regulated in similar directions both in this new DK cohort and also in the previously described TX cohort (e.g., up-regulated in both poor survival groups) are marked by bars along the right edge of the figure. Genes that were up-regulated in both poor-survival subgroups include (from bottom of the heatmap) *CD37*, *SERPINE2*, *CDK19*, *PRLR*, and *RPL23*. Down-regulated genes include *AQP3*, *SCNN1A*, *LTF*, *ELL2*, and *DKK3*. These results further indicate that a subgroup of ACC tumors from patients with poor survival display a unique gene expression profile that could potentially be used for prognostication [6].

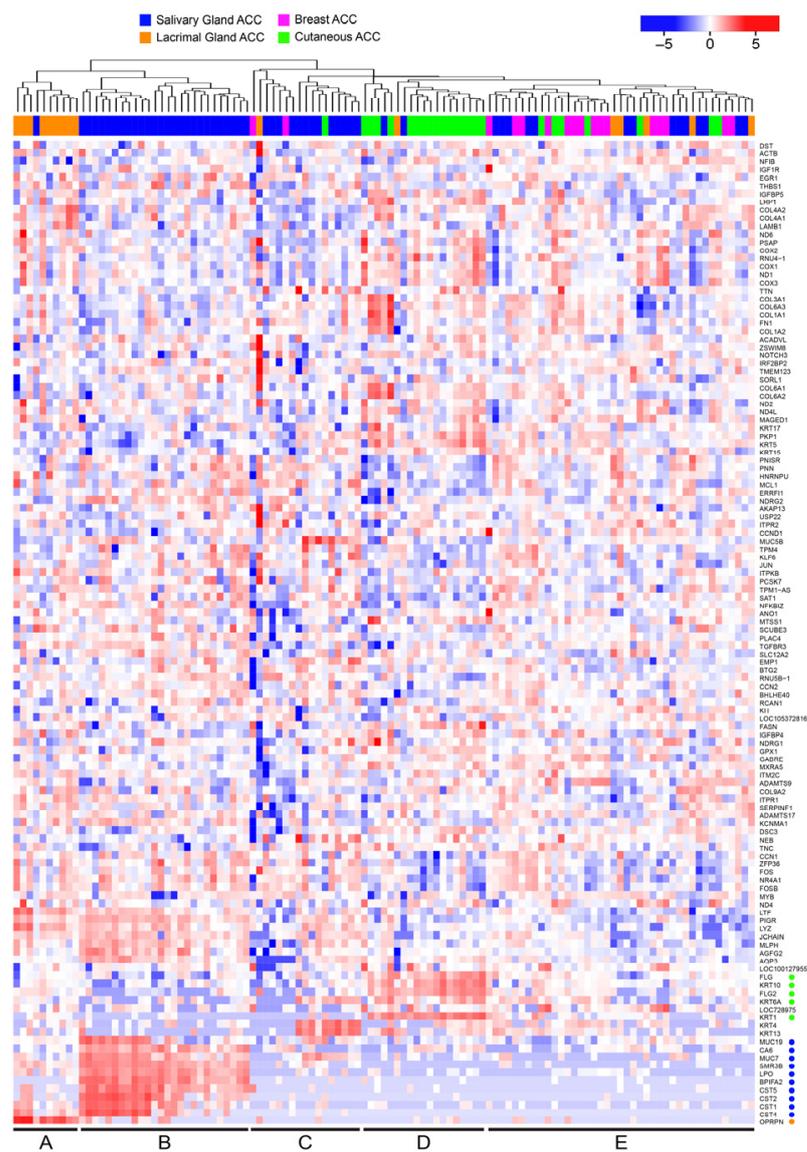


Figure 2. Tissue-specific gene expression differences in ACC tumors. The ACC tumors from the DK cohort were analyzed for tissue-specific gene expression by specifically selecting genes that marked tumors derived from different tissues. A total of 1089 differentially expressed genes were identified by comparing all the tissue groups to each other (at least 2-fold up- or down-regulated with adjusted p -value < 0.05). The heatmap summarizes the gene expression differences for 123 of the most highly expressed genes. The tissues of origin are indicated in the color bar at top: lacrimal gland, salivary gland, cutaneous, and breast are indicated by orange, blue, green and pink, respectively. Notable genes mentioned in the text are marked by dots along the right edge. (Note: a larger version of this heatmap is provided in Figure S2 in the Supplementary Materials).

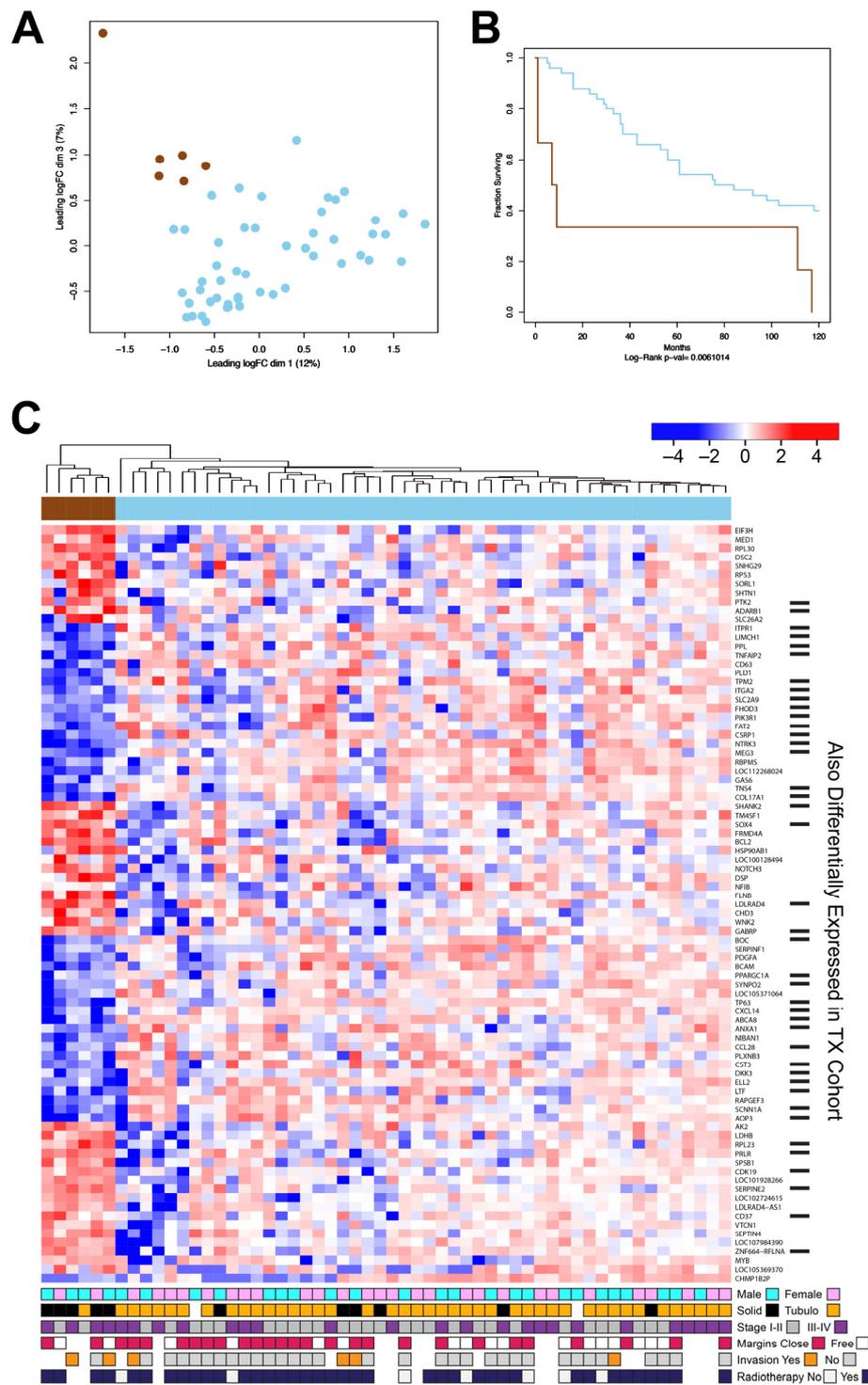


Figure 3. The DK cohort of ACC tumors contains a poor-outcome subgroup. **(A)** Multi-dimensional scaling plot of the gene expression data from the DK cohort of salivary gland ACC tumor samples. **(B)** Kaplan–Meier survival analysis shows that the samples in the brown subgroup had a significantly (p -value = 0.006) worse survival. **(C)** The heatmap summarizes the differential gene expression analysis comparing the poor survival (brown) subgroup to the rest of the samples. Genes marked by bars at the right were also identified previously in a poor-survival subgroup from the TX cohort [6]. The color bars at the bottom summarize the available clinical information for gender, solid or tubulocribiform morphology, tumor stage, margins, vascular invasion, and radiotherapy. A larger version of the heatmap is provided in Supplementary Figure S3.

3.4. Combining the TX and DK Cohorts Provides Additional Details about Subgroups of ACC Tumors

The results described above suggest that the new DK cohort of ACC tumor samples contains subgroups of patients that are very similar to the subgroups we identified previously in the TX cohort [6]. To compare the subgroups we combined the RNA-seq results of the two independent cohorts and performed a unified analysis of 124 salivary gland ACC samples (56 from DK and 68 from TX). Figure 4A shows the multi-dimensional scaling plot of the combined data sets, which form three main groups. The largest group, in the middle of the plot, have been shaded dark blue or cyan to indicate that they overexpress *MYB* or the related *MYBL1* gene, respectively. A group at the upper left is shaded red and contains the poor survival samples from both cohorts, all of which express *MYB*. Finally, a group of samples at the right, shaded orange, express neither *MYB* nor *MYBL1* ('no *MYB*'). This group was described previously, and the 'driver' oncogenes or mutations responsible for that group remain unknown [6]. Although these cohorts of ACC samples were completely independent and the patients came from different countries, both cohorts formed similar major subgroups when analyzed together. A Kaplan–Meier survival analysis of these groups is shown in Figure 4B. The overall survival for patients in the orange 'no Myb' group was similar to the main group of samples expressing either *MYB* or *MYBL1*. As described above, the red group displayed much worse survival compared to the other patients. While the median survival for most patients exceeded 120 months, including the orange 'no *MYB*' group, median survival for the red group was only 16.8 months (p -value $< 1 \times 10^{-6}$). These groups were segregated using only their different gene expression characteristics, suggesting that biomarkers could be developed to identify the patients in the poor survival group at the time of surgery. The MDS plot in Figure 4C shows the large overlap in the DK and TX cohorts, despite being analyzed separately and several years apart.

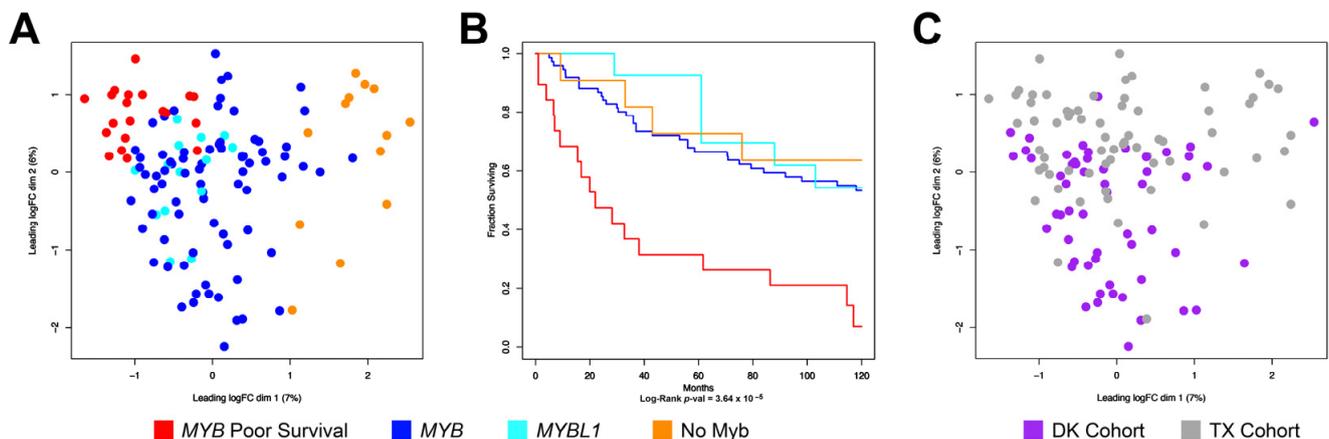


Figure 4. Combined analysis of the DK and TX cohorts of ACC tumor samples. The RNA-seq data from the new DK cohort was combined with previously described (Frerich et al., 2018 [6]) data from the TX cohort for a combined gene expression analysis. (A) Combined multi-dimensional scaling plot. Orange shading indicates 'No Myb' samples that express neither *MYB* nor *MYBL1*, blue and cyan indicate *MYB*- or *MYBL1*-expressing samples in the main group, red indicates the *MYB*-expressing samples with poor survival. (B) Kaplan–Meier survival plot of samples in the four groups of ACC samples. (C) Multi-dimensional scaling plot similar to panel A, but with samples from DK or TX cohorts labeled purple or gray, respectively.

3.5. ACC Tumors That Do Not Express *MYB* or *MYBL1* Have a Unique Transcription Profile

Most ACC tumors have recurrent chromosomal translocations that activate the *MYB* oncogene or the related *MYBL1* gene [1–4], but this raises questions about the underlying biology and driver genes active in the remaining ACC tumors that do not express *MYB* or *MYBL1*. As shown in Figure 4B, the 'no *MYB*' subgroup of tumors (orange line) had

survival similar to the bulk of ACC samples (blue and cyan lines). To further explore the potential driver genes in these samples, we compared them to the rest of the ACC samples in the combined cohort and performed a differential gene expression analysis. In the heatmap shown in Figure 5, the dendrogram at the top shows the hierarchical clustering that was used to arrange the samples from left to right. The ‘no MYB’ samples are at the far right (marked by orange at the top). The heatmap summarizes the gene expression differences for 124 of the 881 genes that were differentially expressed (at least 2-fold up- or down-regulated, adjusted p -value < 0.05) when the ‘no MYB’ samples were compared to all the others. The top 10 up- or down-regulated genes are listed in Table 2, and the full list is provided in Supplementary Data (Table S1).

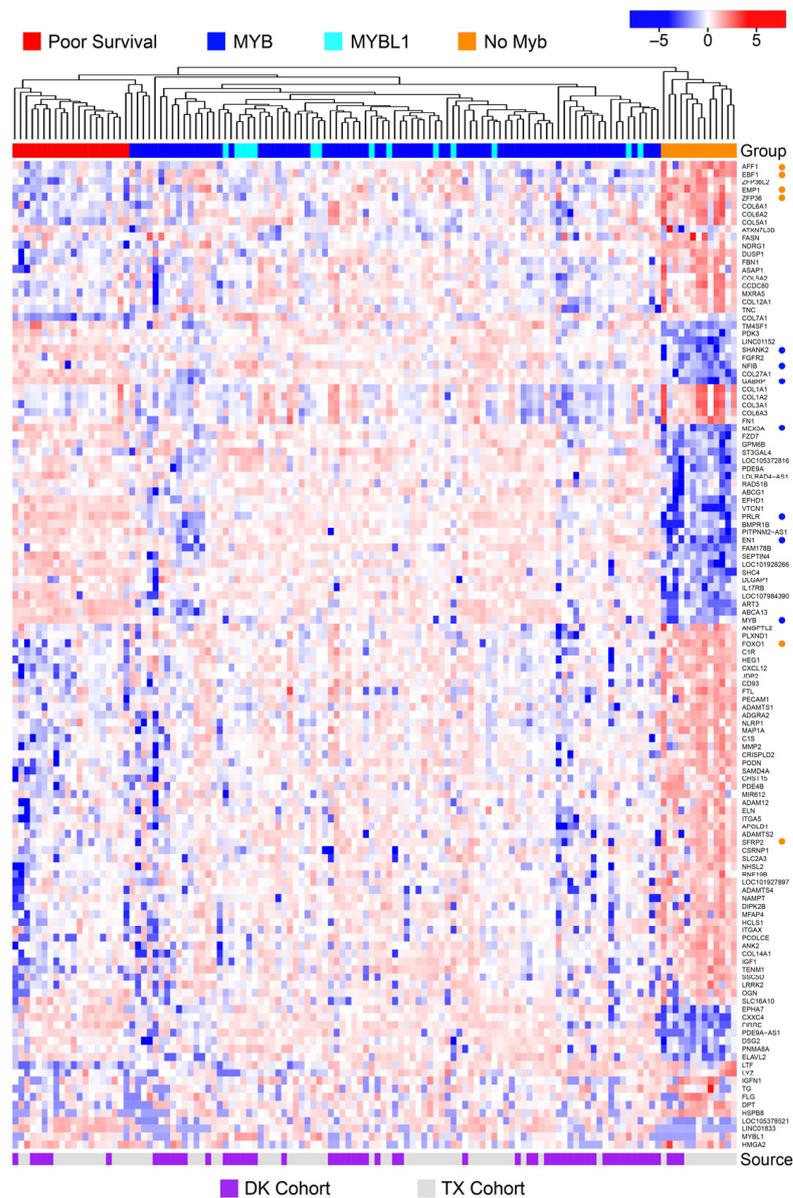


Figure 5. Differential gene expression analysis: ‘No Myb’ samples. The heatmap summarizes the differential gene expression analysis using the combined cohorts of ACC samples from DK and TX, comparing the ‘no Myb’ group (orange color bar at top) to the rest of the samples, all of which express either MYB or MYBL1. Notable genes mentioned in the text are marked by dots at the right. The purple and white color bar at the bottom indicates samples from the DK and TX cohorts, respectively. A larger version of this heatmap is provided in the Supplementary Materials, as Figure S4.

Table 2. Differentially Expressed Genes in No MYB Samples vs. All Others.

DE Genes	Top 10 Up-Regulated	FC	Adj <i>p</i> -Value	Top 10 Down-Regulated	FC	Adj <i>p</i> -Value
881	<i>TG</i>	136.26	1.41×10^{-32}	<i>LINC01833</i>	0.02	1.39×10^{-5}
	<i>HMGA2</i>	42.11	1.66×10^{-14}	<i>MUC7</i>	0.03	1.25×10^{-2}
	<i>IGFN1</i>	14.91	7.41×10^{-8}	<i>LOC643201</i>	0.03	2.04×10^{-3}
	<i>LYZ</i>	12.38	4.19×10^{-6}	<i>MUC19</i>	0.03	1.93×10^{-3}
	<i>FLG2</i>	12.08	5.87×10^{-5}	<i>CTNND2</i>	0.04	6.12×10^{-5}
	<i>COL1A1</i>	9.62	9.72×10^{-20}	<i>FIRRE</i>	0.04	4.73×10^{-10}
	<i>FLG</i>	8.80	5.38×10^{-6}	<i>LOC105378521</i>	0.04	8.12×10^{-6}
	<i>LTF</i>	8.05	1.37×10^{-6}	<i>ART3</i>	0.04	7.41×10^{-8}
	<i>HSPB8</i>	7.97	5.59×10^{-6}	<i>LOC107984390</i>	0.04	1.42×10^{-11}
	<i>TNNT1</i>	7.74	2.29×10^{-3}	<i>SIX3</i>	0.04	5.12×10^{-4}

There are several important conclusions from this analysis. First, as described previously [5,6], the ACC samples that express *MYBL1* do not form their own subgroup, but mix in with the samples expressing *MYB*, suggesting that the two oncogenes have similar effects on gene expression patterns [6]. Second, each of the three main groups (orange, red, blue) contains samples from both the TX and DK cohorts, suggesting that these subgroups are consistent in ACC tumors and are not a characteristic unique to just one cohort or one analysis. Several interesting genes are marked along the right side of the heatmap, including *AFF1*, *EBF1*, *EMP1*, *ZFP36*, *FOXO1*, and *SFRP2*, which are all up-regulated in the ‘no *MYB*’ tumors (marked by orange dots). The *SHANK2*, *NFIB*, *GABRP*, *MEX3A*, *PRLR*, and *MYB* genes were down-regulated in the ‘no *MYB*’ tumors (marked by blue dots). A gene set enrichment analysis identified a number of Gene Ontology Cellular Process categories that were over-represented in the differentially expressed genes. The top six categories are described in Table 3. The finding that the ‘no *MYB*’ samples have such a dramatically different gene expression profile reinforces the conclusion that the ACC phenotype can be achieved through different regulatory pathways.

Table 3. Gene Ontology Categories of DE Genes from No *MYB* Group Analysis.

GO.ID	Term	Annotated	Signif	Expected	Adj <i>p</i> -Value
GO:0043062	extracellular structure organization	218	87	31.6	2.40×10^{-19}
GO:0045229	external encapsulating structure organization	219	87	31.74	3.50×10^{-19}
GO:0045765	regulation of angiogenesis	128	47	18.55	2.40×10^{-10}
GO:0022610	biological adhesion	658	166	95.37	3.10×10^{-10}
GO:0008284	positive regulation of cell population prolifer.	354	93	51.31	7.80×10^{-8}
GO:0009617	response to bacterium	178	54	25.8	2.10×10^{-7}

3.6. Poor Survival ACC Samples Have a Unique Transcription Profile

For the combined cohorts, the poor survival subgroup, marked by red at the top of the heatmap in Figure 5 and in the survival plot in Figure 4B, displayed a median survival of only 22 months, compared to greater than 123 months for the other patients. This suggests that a gene expression panel could be developed to identify patients at highest risk of poor survival.

To characterize the poor survival subgroup in more detail, we performed in depth analysis of the gene expression patterns of tumors from these patients. The heatmap in Figure 6 summarizes the results of a differential gene expression analysis comparing the poor survival subgroup samples to all the other ACC tumor samples in the combined

Table 4. Poor Survival Samples vs. All Others.

DE Genes	Top 10 Up-Regulated	FC	Adj <i>p</i> -Value	Top 10 Down-Regulated	FC	Adj <i>p</i> -Value
729	<i>ANKRD1</i>	10.28	3.99×10^{-5}	<i>CST4</i>	2.2×10^{-04}	3.46×10^{-3}
	<i>LINC02275</i>	8.57	2.77×10^{-5}	<i>CST5</i>	3.0×10^{-03}	3.82×10^{-3}
	<i>LINC02515</i>	5.66	4.90×10^{-5}	<i>CST2</i>	3.0×10^{-03}	8.83×10^{-3}
	<i>NCAN</i>	5.31	5.79×10^{-3}	<i>CST1</i>	3.1×10^{-03}	2.66×10^{-3}
	<i>HPSE2</i>	4.54	4.08×10^{-3}	<i>SPRR3</i>	3.3×10^{-03}	6.00×10^{-3}
	<i>NPY5R</i>	4.51	9.50×10^{-7}	<i>KRT13</i>	4.2×10^{-03}	3.94×10^{-4}
	<i>LINC01833</i>	4.22	2.95×10^{-3}	<i>SPRR1A</i>	5.3×10^{-03}	2.05×10^{-3}
	<i>SOX8</i>	4.04	8.42×10^{-7}	<i>KRT6C</i>	5.5×10^{-03}	6.69×10^{-4}
	<i>ASL</i>	3.88	3.82×10^{-5}	<i>SMR3B</i>	5.9×10^{-03}	2.82×10^{-3}
	<i>PRLR</i>	3.81	1.99×10^{-8}	<i>BPIFA2</i>	7.1×10^{-03}	1.67×10^{-2}

3.7. A Multi-Gene Classifier to Identify Poor Survival Patients

As shown in Figure 3, the DK cohort of ACC samples contained a subgroup of patients with poor survival, similar to one that was originally identified in the TX cohort [5,6]. Having patients from two independent cohorts allowed us to use the original TX cohort as a training set to develop a multi-gene biomarker panel, which could be validated with the DK cohort. Starting with expression data for 3597 genes expressed above a threshold level in the training set (TX cohort), we used an elastic net type of penalized logistic regression model to identify genes that could distinguish the poor prognosis cohort from the rest of the patients. The model selection was performed with a 10-fold cross-validation and yielded a 49-gene classifier developed solely with data from the TX cohort (Table 5). The ROC curve analysis was used to evaluate the classifier's accuracy.

Table 5. Elastic Net 49-Gene Classifier.

Symbol	Value	Symbol	Value	Symbol	Value
<i>A2M</i>	−0.11	<i>HMCN1</i>	−0.065	<i>PLA2R1</i>	−0.0434
<i>ABCA8</i>	−0.0129	<i>IPO9</i>	0.376	<i>PLAT</i>	−0.0448
<i>ACTA2</i>	−0.1264	<i>ITPR1</i>	−0.1121	<i>PLD1</i>	−0.0356
<i>ADAMTS9</i>	−0.0986	<i>KRT14</i>	−0.0359	<i>PPARGC1A</i>	−0.0378
<i>ALDOA</i>	−0.018	<i>LDLRAD4</i>	0.0354	<i>PRUNE2</i>	−0.0026
<i>ANO1</i>	−0.099	<i>LGR4</i>	−0.0391	<i>RASSF6</i>	−0.1461
<i>APOL6</i>	−0.0977	<i>LIMA1</i>	−0.0793	<i>SEMA3C</i>	−0.0582
<i>CARMN</i>	−0.0091	<i>LIMCH1</i>	−0.145	<i>SH3D19</i>	−0.0562
<i>CD9</i>	−0.0886	<i>LOC107987158</i>	0.0363	<i>SLPI</i>	−0.1399
<i>CFH</i>	−0.0226	<i>LTF</i>	−0.0048	<i>SVIL</i>	−0.168
<i>COL17A1</i>	−0.0151	<i>MAMLD1</i>	0.0339	<i>SYNPO2</i>	−0.0159
<i>COL7A1</i>	−0.0095	<i>MIR205HG</i>	−0.0511	<i>TAGLN</i>	−0.0321
<i>COL9A2</i>	−0.0072	<i>MLPH</i>	−0.013	<i>TNFRSF19</i>	−0.0054
<i>DMD</i>	−0.3178	<i>MTUS1</i>	−0.0917	<i>TP63</i>	−0.1442
<i>EFS</i>	0.0074	<i>PARP14</i>	−0.0746	<i>TPM2</i>	−0.0175
<i>EGFR</i>	−0.0288	<i>PCYOX1</i>	−0.1327		
<i>FRMD4B</i>	−0.0279	<i>PIK3R1</i>	−0.0097		

Note: Genes used in the 14-gene classifier are shown in bold.

As shown in Figure 7, the elastic net classifier could distinctly separate the poor prognosis samples from others in the training set (TX cohort, left panel) as AUC = 1. An unbiased estimate for the AUC (AUC = 1) was also achieved through a double loop nested

cross-validation, which showed a perfect classification performance. However, the accuracy achieved was expected because the same gene expression data were used to develop and test the classifier. Importantly, the classifier developed with the TX cohort also gave nearly perfect (AUC = 0.984) separation on the independent DK cohort test set (right panel). We also generated a 14-gene subset of the classifier using a Least Absolute Shrinkage Selection Operator (LASSO) approach. The 14-gene classifier containing genes *A2M*, *ACTA2*, *ANO1*, *APOL6*, *DMD*, *IPO9*, *LIMCH1*, *MAMLD1*, *MIR205HG*, *PLAT*, *RASSF6*, *SEMA3C*, *SLPI*, and *TP63* (shown in bold in Table 4) was only slightly less accurate, with AUC = 0.976. These results suggest that a gene classifier can be used to identify ACC patients with poor prognosis.

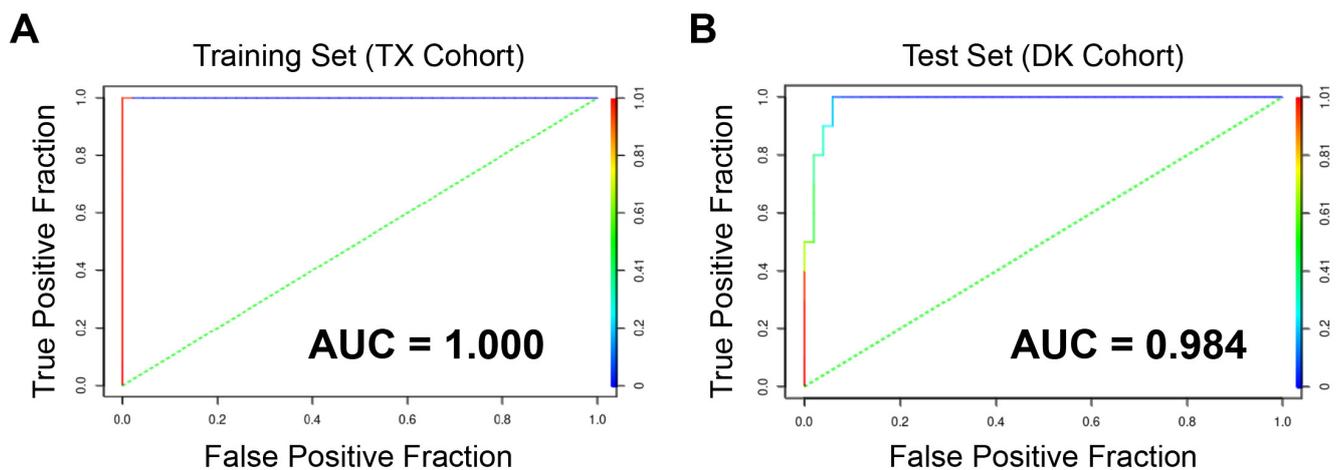


Figure 7. Elastic net ROC curves. The elastic net classifier developed with the TX cohort training set (A, Left) produced an Area Under the Curve (AUC) of 0.984 in the DK cohort test set (B, Right).

To illustrate their usefulness, we tested the genes in the classifiers on the combined DK and TX cohorts of salivary gland ACC samples. We limited the data sets to only the 49-gene or 14-gene lists (except that *MYB*, *MYBL1*, and *NFIB* were added back for comparison), and performed hierarchical clustering, which identified the two major clusters shown in the dendrograms in Figure 8A,B. The heatmaps display the differences in gene expression. Interestingly, most of the classifier genes were down-regulated in the poor prognosis tumors compared to the other samples. As an example, the *TP63* tumor suppressor gene is significantly down-regulated in the poor prognosis group. The poor prognosis tumors appear to lack the expression of specific genes that are expressed by the other ACC samples. Notably, only 4 of the 11 ‘solid form’ morphology samples from the DK cohort were in the poor prognosis subgroup, suggesting that solid morphology is insufficient to classify samples as poor prognosis [37,38].

As shown by the color bar at the top of each heatmap, some of the samples that were in the poor prognosis subgroup described in Figure 4 (marked red) did not cluster with the poor survival samples identified by the gene classifiers, and a few samples that were not included above did cluster in the poor survival group in this analysis. However, as shown in the Kaplan–Meier survival plots in Figure 8C,D, the classifiers did identify a poor-prognosis group with median survival of less than 20 months, compared to a median of 125 months for the rest of the samples. In addition, none of the poor-prognosis patients identified by the classifiers survived 10 years, while more than half of the other patients survived at least 10 years. Thus, the multi-gene classifiers identified using the TX cohort samples were able to identify a subset of ACC patients in the independent DK cohort, which validates the classifiers and suggests that adapting them to the clinic could be useful.

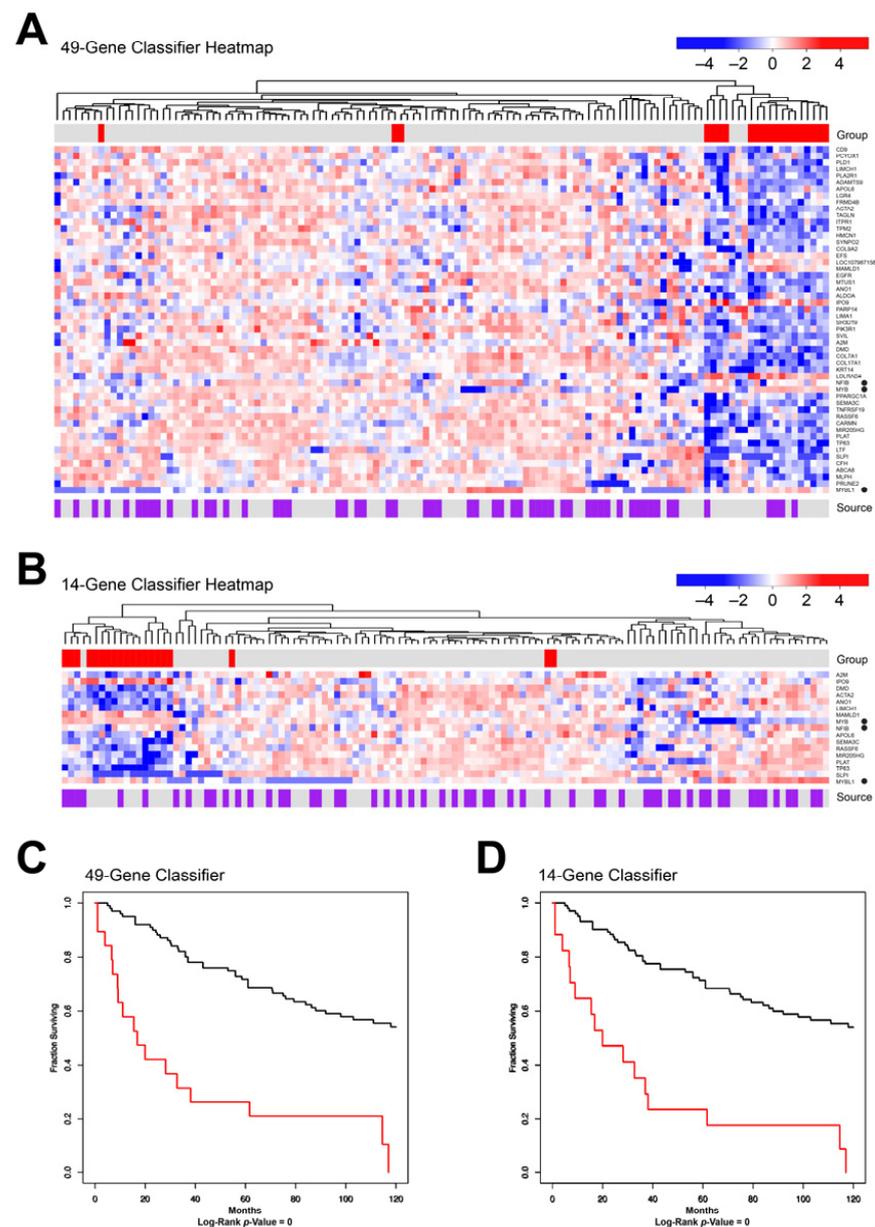


Figure 8. Classifier Groups. The 49-gene (A) or 14-gene (B) classifiers were used to separate ACC samples into groups by hierarchical clustering, as illustrated in the dendrograms at the top of each heatmap, which compare the gene expression profiles of the two groups. (The *MYB*, *MYBL1*, and *NFIB* genes were added to the analysis for comparison. They are marked by black dots at the right). The samples from the original poor survival group (see Figure 4) are marked by red in the color bars. At the bottom, Kaplan–Meier plots compare the survival of patients in the groups defined by the 49-gene (C) or 14-gene (D) classifiers. Larger versions of these heatmaps are in Supplementary Figure S6.

To examine whether the gene classifier provides more information for survival outcomes beyond that contained in the clinical covariates, we performed univariate and multivariate Cox regression analyses, with the gene classifier and clinical covariables deemed to be the risk factors as predictors. (Details of the analysis are in File S1). The available clinical covariables include Margins (free or close), Vascular Invasion (yes or no), Radiotherapy (yes or no), Cribriform (tubular or solid), and Stage (I-II or III-IV). The analyses were restricted to 56 samples; a union of the subsets to which the data of each variable are available. However, the number of samples used by each Cox regression

analysis varied subject to data availability. The univariate and bivariate analyses are in the Supplementary Materials, and the multivariate analysis is reported in Table 6.

Table 6. Multivariate Cox Regression Analysis.

Clinical Covariates & Gene Classifier	Values	Hazard Ratio	95% Confidence Interval	<i>p</i> -Value
Vascular Invasion	No	1		0.097
	Yes	2.703	0.83–8.76	
Stage	I–II	1		0.271
	III–IV	1.601	0.69–3.70	
Gene Classifier	Group 1	1		0.010
	Group 2	26.01	2.19–309.3	

The univariate analysis (see Supplementary Materials) showed that the two variables, Vascular Invasion and Cribriform, were significantly associated with survival outcomes ($p < 0.05$), while the variable Stage was marginally significant ($p = 0.084$). We compared these three variables with the gene classifier through bivariate Cox regression (see Supplementary Materials). The result showed a remarkable association between our gene classifier and survival after adjusting for each clinical covariate's effect. We further performed a multivariate Cox regression (Table 6), and our gene classifier was still significantly correlated with the survival outcomes after adjusting for Vascular Invasion and Stage effects. Note that the Cox regression with three or more variables will not converge if we include Cribriform in the model, which limits our ability to conduct further investigation in this respect. However, the results have given sufficient statistical evidence that our gene classifier provided more information about the survival outcome than the available clinical parameters.

4. Discussion

We compared the transcription profiles of ACC tumor samples that arose in very different tissues: salivary gland, lacrimal gland, breast, and skin. Despite being from different tissues, all ACC tumors had markedly similar gene expression profiles. Indeed, the ACC samples were much more similar to each other than they were to normal salivary gland tissue or another type of salivary gland tumor, acinic cell carcinoma [28]. These results demonstrate that ACC tumors arising in different tissues are highly related and are difficult to distinguish using gene expression patterns alone. Interestingly, different types of ACC tumors were shown previously to have distinct patterns of microRNA expression [36]—a result that we could not reproduce using gene expression results. This suggests that the activated *MYB* or *MYBL1* oncogenes may induce an ACC-specific gene expression pattern that affects protein-coding genes much more than microRNAs. This is a fascinating biological difference that could be important for explaining tumor phenotypes and some aspects of tissue differentiation.

Having RNA-seq data from a new set of ACC samples provided us with the opportunity to perform a validation cohort analysis. Despite the challenges that exist for translating RNA sequencing (RNA-seq) results into widely used clinical assays [39], several types of gene expression signatures have been developed for clinical use [40–42]. In this study, we used RNA-seq data from a previous cohort of 68 salivary gland ACC samples to develop a 49-gene expression classifier for identifying a subgroup of patients with poor survival. We then validated the result using results from the new cohort of 56 salivary ACC samples, finding that the biomarker was able to distinguish 98% of the poor survival patients. A smaller 14-gene classifier achieved similar results with slightly less accuracy. Salivary gland ACC patients display widely variable outcomes, with some patients surviving decades after surgery and others succumbing after only a few months [5,30,43]. It seems clear that the development of new clinical trials should be targeted to the ACC patients that are most

likely to have a recurrence and die from the disease. The validated biomarker we have described should have important utility, if it can be developed into an assay suitable for clinical laboratory use. Although our results do not suggest a new or modified treatment for ACC patients, they do suggest that developing a suitable biomarker assay to identify the worse prognosis patients is worthwhile so a new therapeutic strategy could be developed for them. Clinical RNA-seq is fast becoming a routine assay for cancer patients, so these biomarkers should be adaptable to clinical laboratories.

Some ACC tumors display a ‘solid form’ morphology, which has been associated with worse prognosis [37,38]. Other clinical features, such as advanced tumor stage, lack of clean margins during surgery, or vascular invasion, might also be used to identify higher risk patients. However, in our analysis, none of these other markers were able to identify the poor prognosis group of patients that we identified using gene expression patterns. Therefore, we conclude that the gene classifiers provide a novel and independent means of distinguishing poor prognosis ACC patients that should be pursued and studied further. The next step will be to develop assays that work in clinical laboratories so that these classifiers can be used to identify patients that should be targeted for clinical trials or more aggressive therapy to improve their survival.

In addition to identifying and validating a multi-gene classifier for ACC patients, analyzing a new cohort of 56 salivary gland ACC samples from Denmark (DK) also validated important biological results that we described previously using 68 ACC patient samples from the Salivary Gland Tumor Bank in Texas (TX) [5,6]. The main result is that ACC patients can be divided into at least three distinct groups based on gene expression signatures. These groups are easily discernable in the multi-dimensional scaling plots (e.g., see Figure 4). The samples in the main group, comprising 76% of the total, express either *MYB* or *MYBL1* and have a median survival of more than 10 years after surgery. A second group, about 10% of the samples with survival similar to the main group, express neither *MYB* nor *MYBL1*. These samples have a unique gene expression signature, suggesting a different mechanism driving the malignancy. The samples in the final group, about 14% of the total, are the focus of the multi-gene classifier because they have much worse survival than the rest of the ACC patients.

Although the detailed transcriptome analyses that we performed were able to discern distinct subgroups of ACC tumor types, the bulk RNA-sequencing does not provide information on cell lineage composition within tumors. Thus, it is not clear if the different subgroups result from the unique features of different types of ACC tumor cells or whether the subgroups are due to differences in cellular composition in the tumors. Addressing those questions will require using single-cell genomics assays or spatial genomics approaches that can discern different cell types in the tumors.

5. Conclusions

Our somewhat surprising result is that ACC tumors arising in different tissues or organs have remarkably similar transcriptional profiles. Indeed, we were unable to identify gene signatures that distinguished the ACC tumors from different organs. This may point to an important underlying biology in ACC tumors that makes them so similar. Since the majority of ACC tumors overexpress the *MYB* or *MYBL1* genes, the dominant ACC phenotype may be induced by the activated Myb transcription factors.

A second, but very important finding is that RNA sequencing analysis can be used to identify a subgroup of *MYB*-expressing salivary gland ACC patients with poor prognosis. We were able to use the new DK cohort of ACC samples to validate a biomarker developed with an earlier (TX) cohort. This is especially important for diseases like ACC, in which many patients survive more than 10 years post-surgery. Our results provide a tool for identifying the patients that should be enrolled in clinical trials of targeted therapies to improve their outcomes.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cancers15051390/s1>, Figures S1–S6: Larger versions of the heatmaps in Figures 1–3, Figures 5, 6 and 8, respectively. Table S1: No Myb group differential gene expression analysis results. Table S2: Poor survival differential gene expression analysis results; File S1: Clinical data analysis details with univariate and bivariate Cox regression results and details of software packages used.

Author Contributions: Conceptualization, S.A.N. and S.A.; methodology, K.J.B. and S.A.N.; software, S.A.N.; validation, K.J.B. and S.A.N.; formal analysis, H.K. and Y.G.; resources, S.A., A.K.E.-N., P.H., K.K., L.M., S.H., D.P., A.M. and D.T.T.; data curation, K.J.B. and S.A.N.; writing—original draft preparation, S.A.N.; writing—review and editing, K.J.B., A.K.E.-N., D.Y.L., P.H. and K.K.; visualization, S.A.N. and K.J.B.; supervision, S.A.N.; project administration, K.J.B. and S.A.N.; funding acquisition, S.A.N. All authors have read and agreed to the published version of the manuscript.

Funding: The RNA-seq analysis was partially supported by UNM Comprehensive Cancer Center Support Grant from NIH NCI P30CA118100. K.J.B. and S.A.N. were partially supported by grants from the Adenoid Cystic Carcinoma Research Foundation and from NIH NIDCR R01DE023222 (to S.A.N.). The DK studies were substantially supported by Region Zealand’s Research Fund.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of University of New Mexico (protocol code 20-120 approved 4 June 2020) and the Danish Regional Ethics Committee (H-6-2014-086) and the Danish Data Protection Agency (Journal no. REG-94-2014).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: RNA sequencing data is available for download from the NCBI Bio-Project database using study accession number PRJNA287156.

Acknowledgments: S.A.N. and K.J.B. acknowledge the outstanding technical support from J. Padilla, J. Woods, and M. Cyphery in the Ness laboratory and the Analytical and Translational Genomics Shared Resource.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Persson, M.; Andren, Y.; Mark, J.; Horlings, H.M.; Persson, F.; Stenman, G. Recurrent fusion of MYB and NFIB transcription factor genes in carcinomas of the breast and head and neck. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 18740–18744. [[CrossRef](#)] [[PubMed](#)]
2. Mitani, Y.; Li, J.; Rao, P.H.; Zhao, Y.J.; Bell, D.; Lippman, S.M.; Weber, R.S.; Caulin, C.; El-Naggar, A.K. Comprehensive analysis of the MYB-NFIB gene fusion in salivary adenoid cystic carcinoma: Incidence, variability, and clinicopathologic significance. *Clin. Cancer Res.* **2010**, *16*, 4722–4731. [[CrossRef](#)] [[PubMed](#)]
3. Bell, D.; Roberts, D.; Karpowicz, M.; Hanna, E.Y.; Weber, R.S.; El-Naggar, A.K. Clinical significance of Myb protein and downstream target genes in salivary adenoid cystic carcinoma. *Cancer Biol. Ther.* **2011**, *12*, 569–573. [[CrossRef](#)] [[PubMed](#)]
4. Brill, L.B., II; Kanner, W.A.; Fehr, A.; Andren, Y.; Moskaluk, C.A.; Loning, T.; Stenman, G.; Frierson, H.F., Jr. Analysis of MYB expression and MYB-NFIB gene fusions in adenoid cystic carcinoma and other salivary neoplasms. *Mod. Pathol.* **2011**, *24*, 1169–1176. [[CrossRef](#)]
5. Brayer, K.J.; Frerich, C.A.; Kang, H.; Ness, S.A. Recurrent Fusions in MYB and MYBL1 Define a Common, Transcription Factor-Driven Oncogenic Pathway in Salivary Gland Adenoid Cystic Carcinoma. *Cancer Discov.* **2016**, *6*, 176–187. [[CrossRef](#)]
6. Frerich, C.A.; Brayer, K.J.; Painter, B.M.; Kang, H.; Mitani, Y.; El-Naggar, A.; Ness, S.A. Transcriptomes define distinct subgroups of salivary gland adenoid cystic carcinoma with different driver mutations and outcomes. *Oncotarget* **2018**, *9*, 7341–7358. [[CrossRef](#)]
7. Drier, Y.; Cotton, M.J.; Williamson, K.E.; Gillespie, S.M.; Ryan, R.J.; Kluk, M.J.; Carey, C.D.; Rodig, S.J.; Sholl, L.M.; Afrogheh, A.H.; et al. An oncogenic MYB feedback loop drives alternate cell fates in adenoid cystic carcinoma. *Nat. Genet.* **2016**, *48*, 265–272. [[CrossRef](#)]
8. Ness, S.A. The myb oncoprotein: Regulating a regulator. *Biochim. Biophys. Acta* **1996**, *1288*, F123–F139. [[CrossRef](#)]
9. Zhou, Y.; Ness, S.A. Myb proteins: Angels and demons in normal and transformed cells. *Front. Biosci.* **2011**, *16*, 1109–1131. [[CrossRef](#)]
10. Frerich, C.A.; Sedam, H.N.; Kang, H.; Mitani, Y.; El-Naggar, A.K.; Ness, S.A. N-Terminal Truncated Myb with New Transcriptional Activity Produced Through Use of an Alternative MYB Promoter in Salivary Gland Adenoid Cystic Carcinoma. *Cancers* **2019**, *12*, 45. [[CrossRef](#)]
11. Lei, W.; Rushton, J.J.; Davis, L.M.; Liu, F.; Ness, S.A. Positive and negative determinants of target gene specificity in Myb transcription factors. *J. Biol. Chem.* **2004**, *279*, 29519–29527. [[CrossRef](#)] [[PubMed](#)]

12. Liu, F.; Lei, W.; O'Rourke, J.P.; Ness, S.A. Oncogenic mutations cause dramatic, qualitative changes in the transcriptional activity of c-Myb. *Oncogene* **2006**, *25*, 795–805. [[CrossRef](#)] [[PubMed](#)]
13. Rushton, J.J.; Davis, L.M.; Lei, W.; Mo, X.; Leutz, A.; Ness, S.A. Distinct changes in gene expression induced by A-Myb, B-Myb and c-Myb proteins. *Oncogene* **2003**, *22*, 308–313. [[CrossRef](#)] [[PubMed](#)]
14. Fuglerud, B.M.; Ledsaak, M.; Rogne, M.; Eskeland, R.; Gabrielsen, O.S. The pioneer factor activity of c-Myb involves recruitment of p300 and induction of histone acetylation followed by acetylation-induced chromatin dissociation. *Epigenet. Chromatin* **2018**, *11*, 35. [[CrossRef](#)] [[PubMed](#)]
15. Li, Z.; Abraham, B.J.; Berezovskaya, A.; Farah, N.; Liu, Y.; Leon, T.; Fielding, A.; Tan, S.H.; Sanda, T.; Weintraub, A.S.; et al. APOBEC signature mutation generates an oncogenic enhancer that drives LMO1 expression in T-ALL. *Leukemia* **2017**, *31*, 2057–2064. [[CrossRef](#)]
16. Mansour, M.R.; Abraham, B.J.; Anders, L.; Berezovskaya, A.; Gutierrez, A.; Durbin, A.D.; Echin, J.; Lawton, L.; Sallan, S.E.; Silverman, L.B.; et al. Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. *Science* **2014**, *346*, 1373–1377. [[CrossRef](#)]
17. Frierson, H.F., Jr.; Moskaluk, C.A. Mutation signature of adenoid cystic carcinoma: Evidence for transcriptional and epigenetic reprogramming. *J. Clin. Investig.* **2013**, *123*, 2783–2785. [[CrossRef](#)]
18. Introna, M.; Golay, J.; Frampton, J.; Nakano, T.; Ness, S.; Graf, T. Mutations in v-myb alter the differentiation of myelomonocytic cells transformed by the oncogene. *Cell* **1990**, *63*, 1289–1297. [[CrossRef](#)]
19. Ness, S.A.; Beug, H.; Graf, T. v-myb dominance over v-myc in doubly transformed chick myelomonocytic cells. *Cell* **1987**, *51*, 41–50. [[CrossRef](#)]
20. Ness, S.A.; Marknell, Å.; Graf, T. The v-myb oncogene product binds to and activates the promyelocyte-specific mim-1 gene. *Cell* **1989**, *59*, 1115–1125. [[CrossRef](#)]
21. McNagny, K.M.; Graf, T. E26 leukemia virus converts primitive erythroid cells into cycling multilineage progenitors. *Blood* **2003**, *101*, 1103–1110. [[CrossRef](#)] [[PubMed](#)]
22. Jiang, L.; Duan, M.; Guo, F.; Tang, J.; Oybamiji, O.; Yu, H.; Ness, S.; Zhao, Y.Y.; Mao, P.; Guo, Y. SMDB: Pivotal somatic sequence alterations reprogramming regulatory cascades. *NAR Cancer* **2020**, *2*, zcaa030. [[CrossRef](#)] [[PubMed](#)]
23. Kulesa, H.; Frampton, J.; Graf, T. GATA-1 reprograms avian myelomonocytic cell lines into eosinophils, thromboblats, and erythroblats. *Genes Dev.* **1995**, *9*, 1250–1262. [[CrossRef](#)] [[PubMed](#)]
24. Lemma, R.B.; Ledsaak, M.; Fuglerud, B.M.; Sandve, G.K.; Eskeland, R.; Gabrielsen, O.S. Chromatin occupancy and target genes of the haematopoietic master transcription factor MYB. *Sci. Rep.* **2021**, *11*, 9008. [[CrossRef](#)]
25. Coca-Pelaz, A.; Rodrigo, J.P.; Bradley, P.J.; Vander Poorten, V.; Triantafyllou, A.; Hunt, J.L.; Strojjan, P.; Rinaldo, A.; Haigentz, M., Jr.; Takes, R.P.; et al. Adenoid cystic carcinoma of the head and neck—An update. *Oral Oncol.* **2015**, *51*, 652–661. [[CrossRef](#)]
26. Brown, R.B.; Madrid, N.J.; Suzuki, H.; Ness, S.A. Optimized approach for Ion Proton RNA sequencing reveals details of RNA splicing and editing features of the transcriptome. *PLoS ONE* **2017**, *12*, e0176675. [[CrossRef](#)]
27. Feeney, L.; Hapuarachi, B.; Adderley, H.; Rack, S.; Morgan, D.; Walker, R.; Rauch, R.; Herz, E.; Kaye, J.; Harrington, K.; et al. Clinical disease course and survival outcomes following disease recurrence in adenoid cystic carcinoma with and without NOTCH signaling pathway activation. *Oral Oncol.* **2022**, *133*, 106028. [[CrossRef](#)]
28. Lee, D.Y.; Brayer, K.J.; Mitani, Y.; Burns, E.A.; Rao, A.; Bell, D.; Williams, M.D.; Ferrarotto, R.; Pytynia, K.B.; El-Naggar, A.K.; et al. Oncogenic Orphan Nuclear Receptor NR4A3 Interacts and Cooperates with MYB in Acinic Cell Carcinoma. *Cancers* **2020**, *12*, E2433. [[CrossRef](#)]
29. Akbaba, S.; Bostel, T.; Lang, K.; Bahadir, S.; Lipman, D.; Schmidberger, H.; Matthias, C.; Rotter, N.; Knopf, A.; Freudlsperger, C.; et al. Large German Multicenter Experience on the Treatment Outcome of 207 Patients With Adenoid Cystic Carcinoma of the Major Salivary Glands. *Front. Oncol.* **2020**, *10*, 593379. [[CrossRef](#)]
30. Ferrarotto, R.; Mitani, Y.; Diao, L.; Guijarro, I.; Wang, J.; Zweidler-McKay, P.; Bell, D.; William, W.N., Jr.; Glisson, B.S.; Wick, M.J.; et al. Activating NOTCH1 Mutations Define a Distinct Subgroup of Patients With Adenoid Cystic Carcinoma Who Have Poor Prognosis, Propensity to Bone and Liver Metastasis, and Potential Responsiveness to Notch1 Inhibitors. *J. Clin. Oncol.* **2017**, *35*, 352–360. [[CrossRef](#)]
31. Davis, M.P.; van Dongen, S.; Abreu-Goodger, C.; Bartonicek, N.; Enright, A.J. Kraken: A set of tools for quality control and analysis of high-throughput sequence data. *Methods* **2013**, *63*, 41–49. [[CrossRef](#)] [[PubMed](#)]
32. Wood, D.E.; Salzberg, S.L. Kraken: Ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* **2014**, *15*, R46. [[CrossRef](#)] [[PubMed](#)]
33. Wood, D.E.; Lu, J.; Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* **2019**, *20*, 257. [[CrossRef](#)] [[PubMed](#)]
34. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021.
35. Friedman, J.; Hastie, T.; Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* **2010**, *33*, 1–22. [[CrossRef](#)]
36. Andreasen, S.; Tan, Q.; Agander, T.K.; Steiner, P.; Bjorndal, K.; Hogdall, E.; Larsen, S.R.; Erentaite, D.; Olsen, C.H.; Ulhoi, B.P.; et al. Adenoid cystic carcinomas of the salivary gland, lacrimal gland, and breast are morphologically and genetically similar but have distinct microRNA expression profiles. *Mod. Pathol.* **2018**, *31*, 1211–1225. [[CrossRef](#)]

37. Morita, N.; Murase, T.; Ueda, K.; Nagao, T.; Kusafuka, K.; Nakaguro, M.; Urano, M.; Taguchi, K.I.; Yamamoto, H.; Kano, S.; et al. Pathological evaluation of tumor grade for salivary adenoid cystic carcinoma: A proposal of an objective grading system. *Cancer Sci.* **2021**, *112*, 1184–1195. [[CrossRef](#)]
38. Xuan, L.; Yuan, J.; Zhang, H.; Zhang, Y.; Liu, H. Dominant cell type analysis predicts head and neck adenoid cystic carcinoma outcomes. *Ann. Diagn. Pathol.* **2022**, *56*, 151867. [[CrossRef](#)]
39. Byron, S.A.; Van Keuren-Jensen, K.R.; Engelthaler, D.M.; Carpten, J.D.; Craig, D.W. Translating RNA sequencing into clinical diagnostics: Opportunities and challenges. *Nat. Rev. Genet.* **2016**, *17*, 257–271. [[CrossRef](#)]
40. Cheung, R.; Xu, H.; Jin, X.; Tian, W.; Pinney, K.; Bu, L.; Stone, S.; Woodward, R.N.; Agrawal, N.; Dholakia, S.; et al. Validation of a gene expression signature to measure immune quiescence in kidney transplant recipients in the CLIA setting. *Biomark. Med.* **2022**, *16*, 647–661. [[CrossRef](#)]
41. Morris, J.S.; Luthra, R.; Liu, Y.; Duose, D.Y.; Lee, W.; Reddy, N.G.; Windham, J.; Chen, H.; Tong, Z.; Zhang, B.; et al. Development and Validation of a Gene Signature Classifier for Consensus Molecular Subtyping of Colorectal Carcinoma in a CLIA-Certified Setting. *Clin. Cancer Res.* **2021**, *27*, 120–130. [[CrossRef](#)]
42. Rusch, M.; Nakitandwe, J.; Shurtleff, S.; Newman, S.; Zhang, Z.; Edmonson, M.N.; Parker, M.; Jiao, Y.; Ma, X.; Liu, Y.; et al. Clinical cancer genomic profiling by three-platform sequencing of whole genome, whole exome and transcriptome. *Nat. Commun.* **2018**, *9*, 3962. [[CrossRef](#)] [[PubMed](#)]
43. Rettig, E.M.; Tan, M.; Ling, S.; Yonescu, R.; Bishop, J.A.; Fakhry, C.; Ha, P.K. MYB rearrangement and clinicopathologic characteristics in head and neck adenoid cystic carcinoma. *Laryngoscope* **2015**, *125*, E292–E299. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.