


RESEARCH ARTICLE

Open Access



# Glycolysis-associated lncRNAs identify a subgroup of cancer patients with poor prognoses and a high-infiltration immune microenvironment

Kuo-Hao Ho<sup>1,2</sup>, Tzu-Wen Huang<sup>1,3</sup>, Chwen-Ming Shih<sup>1,2</sup>, Yi-Ting Lee<sup>1,2</sup>, Ann-Jeng Liu<sup>4</sup>, Peng-Hsu Chen<sup>1,2</sup> and Ku-Chung Chen<sup>1,2\*</sup> 

## Abstract

**Background:** Long noncoding (lnc)RNAs and glycolysis are both recognized as key regulators of cancers. Some lncRNAs are also reportedly involved in regulating glycolysis metabolism. However, glycolysis-associated lncRNA signatures and their clinical relevance in cancers remain unclear. We investigated the roles of glycolysis-associated lncRNAs in cancers.

**Methods:** Glycolysis scores and glycolysis-associated lncRNA signatures were established using a single-sample gene set enrichment analysis (GSEA) of The Cancer Genome Atlas pan-cancer data. Consensus clustering assays and genomic classifiers were used to stratify patient subtypes and for validation. Fisher's exact test was performed to investigate genomic mutations and molecular subtypes. A differentially expressed gene analysis, with GSEA, transcription factor (TF) activity scoring, cellular distributions, and immune cell infiltration, was conducted to explore the functions of glycolysis-associated lncRNAs.

**Results:** Glycolysis-associated lncRNA signatures across 33 cancer types were generated and used to stratify patients into distinct clusters. Patients in cluster 3 had high glycolysis scores and poor survival, especially in bladder carcinoma, low-grade gliomas, mesotheliomas, pancreatic adenocarcinomas, and uveal melanomas. The clinical significance of lncRNA-defined groups was validated using external datasets and genomic classifiers. Gene mutations, molecular subtypes associated with poor prognoses, TFs, oncogenic signaling such as the epithelial-to-mesenchymal transition (EMT), and high immune cell infiltration demonstrated significant associations with cluster 3 patients. Furthermore, five lncRNAs, namely MIR4435-2HG, AC078846.1, AL157392.3, AP001273.1, and RAD51-AS1, exhibited significant correlations with glycolysis across the five cancers. Except MIR4435-2HG, the lncRNAs were distributed in nuclei. MIR4435-2HG was connected to glycolysis, EMT, and immune infiltrations in cancers.

(Continued on next page)

\* Correspondence: [kuchung@tmu.edu.tw](mailto:kuchung@tmu.edu.tw)

<sup>1</sup>Graduate Institute of Medical Sciences, College of Medicine, Taipei Medical University, Taipei, Taiwan

<sup>2</sup>Department of Biochemistry and Molecular Cell Biology, School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(Continued from previous page)

**Conclusions:** We identified a subgroup of cancer patients stratified by glycolysis-associated lncRNAs with poor prognoses, high immune infiltration, and EMT activation, thus providing new directions for cancer therapy.

**Keywords:** Long noncoding RNAs (lncRNAs), Glycolysis, MIR4435-2HG, Epithelial-to-mesenchymal transition (EMT), Immune infiltrations

## Background

Cancer is regarded as a type of metabolic disease. Tumor cells can drive certain metabolic pathways to sustain their biological processes for growth and to adapt to complex tumor microenvironments (TMEs) [1]. A well-established metabolic pathway that plays a prominent role in cancer progression is glycolysis, which is critical for supplying energy and producing metabolic end products, thus maintaining tumor cell survival [2]. In addition to its functions in sustaining tumor growth, activation of glycolysis affects other phenotypic changes. For example, lactic acid produced by the glycolysis pathways induces the epithelial-to-mesenchymal transition (EMT) in lung cancer cells [3]. Furthermore, a pan-cancer study reported that activated glycolysis is correlated with increasing tumor immunity [4]. Hence, understanding the underlying relationship between glycolysis and cancer progression is a critical goal in cancer research.

Long noncoding (lnc)RNAs, which are longer than 200 nucleotides, can modulate gene expressions through various mechanisms. Several oncogenic signaling pathways, such as the cell cycle [5], immune regulation [6], and EMT mediation [7], are linked to lncRNA regulation. Different types of lncRNAs have been revealed to promote glycolysis activation. In breast cancer, lncRNA-SNHG7, which is promoted by c-MYC regulation, can increase glycolysis through inhibiting miR-34a-5p expression [8]. Another lncRNA, long intergenic noncoding RNA for IGF2BP2 stability, suppresses the degradation of IGF2BP2 by inhibiting the ubiquitination–autophagy pathway, leading to glycolysis upregulation in colorectal cancer [9]. However, most studies have focused only on particular lncRNA candidates and their functions in promoting glycolysis. Therefore, we systematically investigated the association of lncRNAs with glycolysis in different cancer types and explored their related signaling pathways and clinical relevance.

In this study, we used pan-cancer data from The Cancer Genome Atlas (TCGA) to identify glycolysis-associated lncRNAs across 33 tumor types. We performed a consensus clustering analysis to classify these glycolysis-associated lncRNAs into distinct clusters. We then identified glycolysis-associated lncRNAs that exhibited key clinical effects in five cancer types. Finally, we explored the potential pathways and functions of glycolysis-correlated lncRNAs in

association with oncogenic signaling such as EMT and immune regulation.

## Methods

### TCGA pan-cancer data analysis and identification of cancer-expressing lncRNAs

Genomic profiles including RNA sequencing (RNA-Seq) data, gene-level copy number, DNA methylation, and patient clinical characteristics of 33 TCGA cancer types were downloaded from UCSC Xena (<https://xena.ucsc.edu/>). Raw counts of RNA-Seq data were normalized to counts per million (CPMs). Gene-level copy number values were calculated using GISTIC2.0. Beta values derived from Illumina human methylation 450K arrays were used to analyze TCGA DNA methylation changes. lncRNA annotation was retrieved from GENCODE, which contains 17,910 lncRNAs. In total, 15,121 lncRNAs were detected in TCGA RNA-Seq data. We defined an lncRNA as that expressed in a certain cancer type if the gene count was > 10 in more than 90% of patients.

### Identification of glycolysis-associated lncRNAs and categorization of cancers into subtypes

To evaluate the degree of glycolysis activation in cancer patients, expressions of glycolysis-involved genes, derived from the Hallmark glycolysis pathway [10], were examined to infer glycolytic activity through a single-sample gene set enrichment analysis (ssGSEA) method in the GSVA package. In each cancer type, we retained glycolysis-involved genes that were expressed in more than 90% of cancer patients with gene counts of > 10, and gene candidates that did not fit these criteria were excluded (Additional file 1: Table S1). The expressed lncRNAs were correlated with glycolysis scores in each cancer type by performing Pearson correlation analyses. lncRNAs with a false discovery rate (FDR) of < 0.05 and an absolute  $R$  of > 0.3 were considered to be significantly associated with glycolysis signaling. A consensus clustering method, according to de Jong et al.'s study [11], was performed to classify cancer patients into distinct clinically relevant subgroups based on glycolysis-correlated lncRNAs. Partitioning around medoids was used as the clustering algorithm. One thousand permutations with a 0.95 random fraction of lncRNAs in each iteration were

repeated to perform the clustering analysis. Based on the delta area plot, the optimal cluster was selected according to whether no appreciable increase was present (Additional file 2: Figure S1). Then, survival differences among distinct clusters were evaluated using the log-rank test, and glycolysis scores within different groups were compared by employing the Kruskal–Wallis test with post hoc Dunn's test.

#### **Development of genomic classifiers in bladder carcinoma and low-grade gliomas**

To characterize genetic signatures permitting the distinguishing of glycolysis-associated lncRNA-based clusters in cancer patients, gene candidates that exhibited a distinct expression pattern among three clusters were identified. Specifically, we selected gene candidates that were consistently upregulated ( $FC > 1.5$  and  $FDR < 10^{-4}$ ) or downregulated ( $FC < 0.7$  and  $FDR < 10^{-4}$ ) in cluster 3 versus cluster 2, and cluster 2 versus cluster 1. In total, 174 upregulated and 49 downregulated genes were identified in low-grade gliomas (LGGs; Additional file 3: Table S2). In bladder carcinoma (BLCA), 121 genes were upregulated and 112 downregulated (Additional file 3: Table S3). A lasso penalized multinomial logistic regression with 10-fold cross-validation was performed to shrink the number of these gene candidates. After identifying the lambda value that provided a minimum mean cross-validated error, the value was applied to penalize our gene signature and thus build the genomic classifier. The 233 gene candidates in BLCA and 223 gene candidates in LGG were shrunk to 26 and 46 gene candidates, respectively (Additional file 4: Table S4 and S5). Then, this genomic classifier was used with the independent microarray data of glioma (GSE16011 and GSE107850) and BLCA (GSE48075 and GSE13507). However, due to the presence of different platforms for the TCGA RNA-Seq data and GEO microarray data, we performed quantile normalization on both the datasets by using preprocessCore package before building the genomic classifier.

#### **Comparisons of molecular differences including genomic mutations and molecular subtypes within different glycolytic signature-classified groups of cancers**

Genomic mutation data of BLCA, LGGs, mesotheliomas (MESOs), pancreatic ductal adenocarcinomas (PAADs), and uveal melanomas (UVMs) were retrieved from UCSC Xena (<https://xena.ucsc.edu/>). These mutation data, generated by the Multi-Center Mutation Calling in Multiple Cancers (MC3) project, were derived from exon sequencing of TCGA cancer patient samples, and genes were categorized into binary calls as either nonsilent mutation or wild-type. Fisher's exact test was performed to investigate genomic mutations that were significantly enriched (with a  $p$  value of  $< 0.01$ ) in

glycolysis score-classified cluster 3 or cluster 1 cancer patients. The identified mutation within each cancer type was shown as a heatmap. To compare established molecular subtypes with the glycolysis signature-classified groups, we selected two cancer types (LGG and BLCA) to perform the analyses; the types have been classified as different groups in other studies [12, 13]. Fisher's exact test was conducted to examine whether glycolysis score-stratified cluster 3 or cluster 1 patients were enriched in certain molecular types.

#### **Differentially expressed gene analysis, GSEA, and TF activity scoring**

To explore the transcriptome that exhibits distinct expression patterns within glycolysis score-stratified clusters, we used CPM-normalized counts from RNA-Seq data to perform a differentially expressed gene (DEG) analysis by using the edgeR package. Gene candidates were ranked based on  $\log_2$  multiples of change to conduct a GSEA with  $10^4$  permutations to calculate the normalized enrichment score (NES), and a pathway with an FDR of  $< 0.01$  was considered significant enrichment. The Hallmark pathway database [10] was utilized to perform the analysis. To explore signaling pathways that were activated in cluster 3 cancer patients, activated signaling pathways in the poor prognosis patients (cluster 3) compared with good prognosis patients (cluster 1) were shown as a heatmap. To identify signaling pathways exhibiting different degrees of activation among distinct cluster cancer patients, we presented NESs derived from cluster 3 versus cluster 1 and cluster 2 versus cluster 1 in radar plots. To identify activated transcription factors (TFs) in cluster 3 cancer patients compared with cluster 1 patients, we scored TF activity by following a methodology developed by Garcia-Alonso et al. [14]. In brief, they defined a set of high-confidence human TFs and their target genes based on public resources including TF-binding site predictions, text-mining-derived and manually curated TF-target interactions, and chromatin immunoprecipitation coupled with high-throughput data (ChIP-X). RNA expressions of these TF targets were utilized to infer TF activity per patient by using analytical rank-based enrichment analyses (aREAs). TF activation levels among clusters were compared using analysis of variance (ANOVA), and an effect size of  $> 0.5$  was considered a difference. Upregulated TFs in cluster 3 patients were used to perform a functional annotation test by using DAVID (<https://david.ncifcrf.gov/tools.jsp>).

#### **Investigation of localization of glycolysis-associated lncRNAs**

The lncALTAS database [15] was used to investigate the possible localization of glycolysis-associated lncRNAs in

cancer cells. This database contains RNA-Seq data from the cytoplasmic and nuclear compartments of 15 cell lines. In total, 6768 GENCODE-annotated lncRNAs were detected. Furthermore, a relative concentration index (RCI) was developed to characterize the distribution of lncRNAs in nuclear and cytoplasmic regions. In brief, this index was calculated using the log ratio of the concentration of a given RNA molecule per unit mass of RNA between its cytoplasmic and nuclear compartments. A positive RCI value indicates that an lncRNA has a higher concentration in the cytoplasmic region. By contrast, a negative value means that an lncRNA is more abundant in the nuclear region.

#### Calculation of immune cell infiltration scores and an unsupervised hierarchical clustering analysis

To infer immune cell infiltration based on transcriptome profiles, we followed the method reported by Şenbabaoğlu et al. [16]. Briefly, different gene markers of immune cells were utilized as gene sets and used to perform an ssGSEA to obtain an immune cell infiltration score for each patient. Then, these immune cell infiltration scores were utilized to perform unsupervised hierarchical clustering. The distance of each patient was calculated using the Euclidean method, and clustering was performed using Ward's method. Results are illustrated as a heatmap by using the complex heatmap package. For immune checkpoint comparisons, we queried the immune-suppressive checkpoint gene list from HisgAtlas [17], and we conducted ANOVA to compare their expressions in different clusters stratified by glycolysis scores. Checkpoints with an effect size of  $> 0.5$  and a  $p$  value of  $< 0.01$  were considered to significantly differ among the three clusters.

#### First-order partial correlation and multivariate linear regression analysis

A first-order partial correlation was performed to explore interlinks among lncRNAs, glycolysis scores, and glycolysis-associated genes. The glycolysis score was assumed to be  $x$ , and glycolysis-associated gene expression was  $y$ . The first-order partial correlation between  $x$  and  $y$  conditioned on lncRNAs was

$$r_{xy|lncRNA} = \frac{r_{xy} - r_{xlncRNA}r_{ylncRNA}}{\sqrt{(1 - r_{xlncRNA}^2)(1 - r_{ylncRNA}^2)}}$$

We compared the cumulative distribution of Pearson correlation coefficients between glycolysis scores and gene expressions with or without removing the effect of lncRNA expression by using the Kolmogorov–Smirnov test. To identify gene candidates that were correlated with lncRNA expression and were independent of copy number variations and DNA methylation, we performed

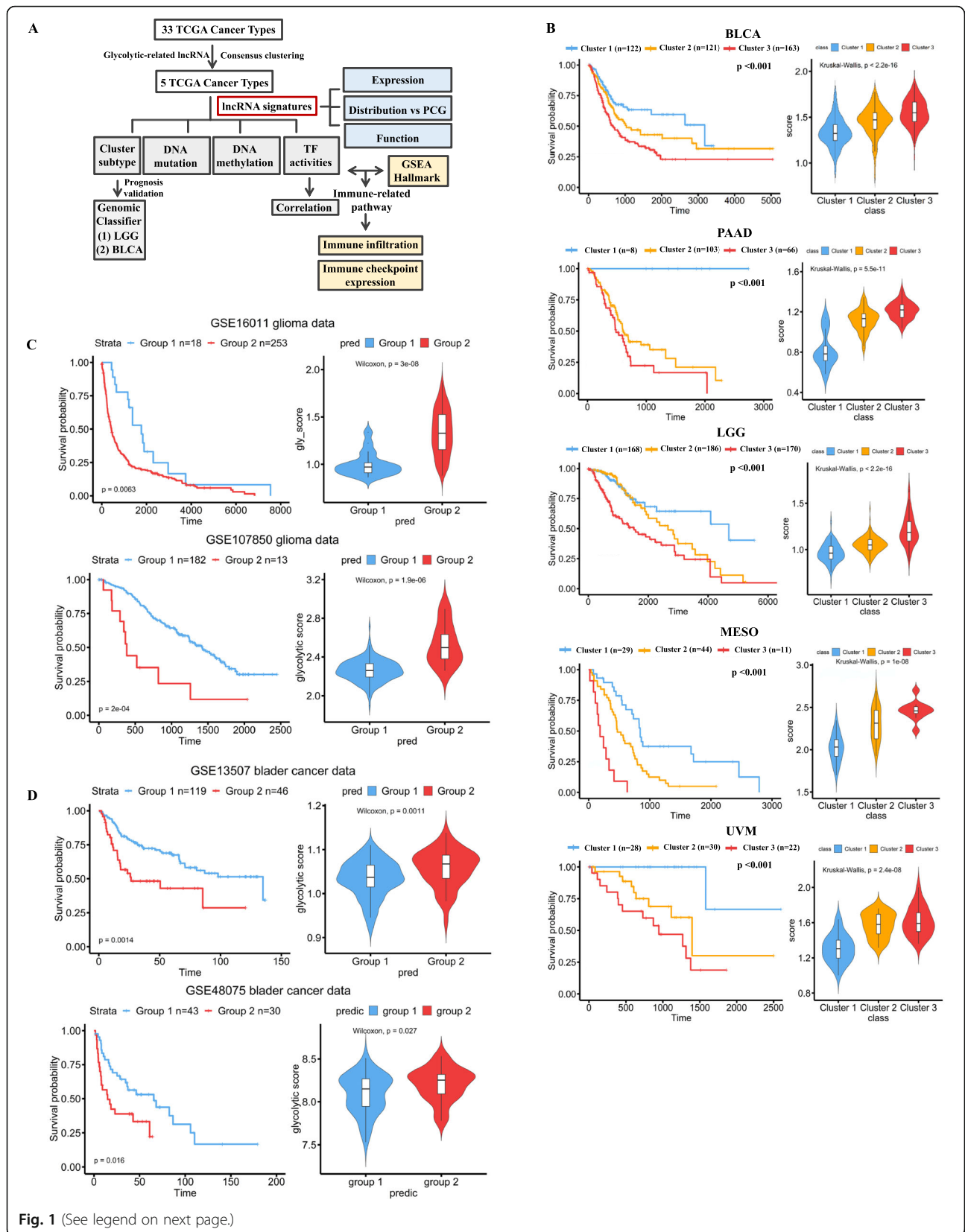
a multivariate linear regression to adjust for these covariates. We considered a gene candidate to be significantly associated with an lncRNA when its absolute correlation coefficient was  $> 0.3$  and its FDR was  $< 10^{-6}$ .

## Results

### Identification of cancer types that can be classified into distinct prognostic groups based on glycolysis-associated lncRNAs

To systematically investigate the roles of glycolysis-associated lncRNAs in different cancer types, we analyzed TCGA pan-cancer RNA-Seq data that contained 33 tumor types from 10,121 tumor samples (Fig. 1a, Additional file 5: Table S6). We used genes derived from Hallmark glycolysis and performed an ssGSEA to infer glycolytic activity (Additional file 6: Figure S2). By performing the Pearson correlation analysis, we identified 1420 lncRNA candidates that were associated with glycolytic activity in at least one cancer type. Most (58.6%) of these lncRNAs were specifically correlated with glycolysis in one cancer type. Because no preexisting clusters were defined based on glycolysis-associated lncRNAs, we employed an unsupervised machine learning algorithm, consensus clustering, to define these subgroups (Additional file 7: Figure S3). Distinct cluster groups were classified for each cancer type. Most of the lncRNAs in cluster 3 were found to be downregulated. The overall survival among these cluster groups in each cancer type was also compared using the log-rank test ( $p < 0.01$ ). Cluster 3 patients in five cancer types—BLCA, LGGs, MESOs, PAADs, and UVMs—demonstrated a significant association with a poor prognosis (Fig. 1b, left panel, Additional file 8: Table S7). Thus, we majorly focused the roles of glycolysis-associated lncRNAs on these five cancer types. The results indicated that the poor prognosis cluster group exhibited the highest glycolysis score across these five cancer types (Fig. 1b, right panel), implying that lncRNAs interact with glycolysis and cancer malignancy in these five cancers.

To validate the clinical significance of these lncRNA-defined groups in cancers, we applied a lasso penalized logistic regression, which was used for the preexisting label classification to train our classifiers. Although we attempted to validate the importance of these glycolysis-associated lncRNA-stratified clusters with other independent data, we discovered that public datasets were established using microarray platforms that lacked the expression information of numerous lncRNA candidates. Therefore, we developed a genomic classifier based on the protein coding gene signature that exhibited a distinct pattern among different clusters for validation with the microarray data. Because sample sizes for PAAD ( $n = 177$ ), MESO ( $n = 84$ ), and UVM ( $n = 80$ ) were small, we mainly focused on the effects of lncRNAs on BLCA



**Fig. 1** (See legend on next page.)

(See figure on previous page.)

**Fig. 1** Identification of glycolysis-associated long noncoding (lnc)RNAs and their roles in cancer prognoses. **a** A flowchart demonstrating our investigation of glycolysis-associated lncRNAs and their putative functions in cancers. **b** Kaplan–Meier plots indicating patient survival rates of distinct clusters stratified by glycolysis-associated lncRNAs in five cancer types. Right panel presents glycolysis scores in different clusters. Cancer patients were clustered into groups based on glycolysis-associated lncRNAs by performing a consensus clustering analysis. Optimal clusters were selected by judging delta area plots where no appreciable increase was evident. Validation of the clinical importance of these glycolysis-associated lncRNAs was derived from genomic classifiers from independent GEO data including GSE16011 and GSE107850 for gliomas (**c**) and GSE13507 and GSE48075 for bladder cancer (**d**). Kaplan–Meier plots demonstrate patient survival rates of different groups. The right panel displays glycolysis scores in different clusters. TCGA, The Cancer Genome Atlas; GSEA, gene set enrichment analysis; PCG, protein coding gene; TF, transcription factor; BLCA, bladder carcinoma; LGG, low-grade glioma; MESO, mesothelioma; PAAD, pancreatic ductal adenocarcinoma; UVM, uveal melanoma

( $n = 406$ ) and LGG ( $n = 524$ ). Furthermore, the microarray data of PAAD, MESO, and UVM cancer patients lacked survival information, limiting our validation of the clinical importance of glycolytic lncRNA-stratified clusters. Thus, we focused mainly on LGG and BLCA to validate the clinical significance of the clusters. We identified 26 gene candidates in BLCA and 45 gene candidates in LGGs by using the developed genomic classifiers, which allowed discrimination within these subgroups based on gene expression levels (Additional file 9: Figure S4, Additional file 4: Table S4 and S5). We then classified LGG and BLCA patients based on distinct clinical features. The results obtained from analyzing two independent databases for LGGs (Fig. 1c) and BLCA (Fig. 1d) indicated that patients with poor prognosis had high glycolysis scores. Moreover, we found that not all the gene candidates that were selected using genomic classifiers belonged to glycolysis pathways, which suggested that glycolysis-associated lncRNAs can be linked to other, nonglycolytic signal pathways to promote cancer malignancy.

#### Molecular characterization of subgroups classified by glycolysis-associated lncRNAs in cancers

Next, we investigated molecular differences in distinct cancers between patient subgroups classified by glycolysis-associated lncRNA signatures. By comparing genomic mutations within these patients, we determined that several gene mutations were associated with different lncRNA-associated groups (Fig. 2a). RB1, KRAS, and BAP1 mutations were enriched in cluster 3 patients with BLCA, PAADs, and UVMS, respectively. As for LGGs, cluster 3 patients generally belonged to the IDH1 wild-type group and had fewer TP53 and CIC mutations.

Previous studies have categorized TCGA BLCA, LGG, PAAD, MESO, and UVM patients into different molecular types based on multiomic profiles [18–22]. Here we compared these defined subgroups and tumor grades with our classified groups by using glycolysis-correlated lncRNA signatures (Fig. 2b). Cluster 3 patients with LGG, BLCA, PAAD, or UVM exhibited higher tumor grades. As for molecular subtypes, cluster 3 patients with BLCA mainly belonged to the basal/squamous subtype,

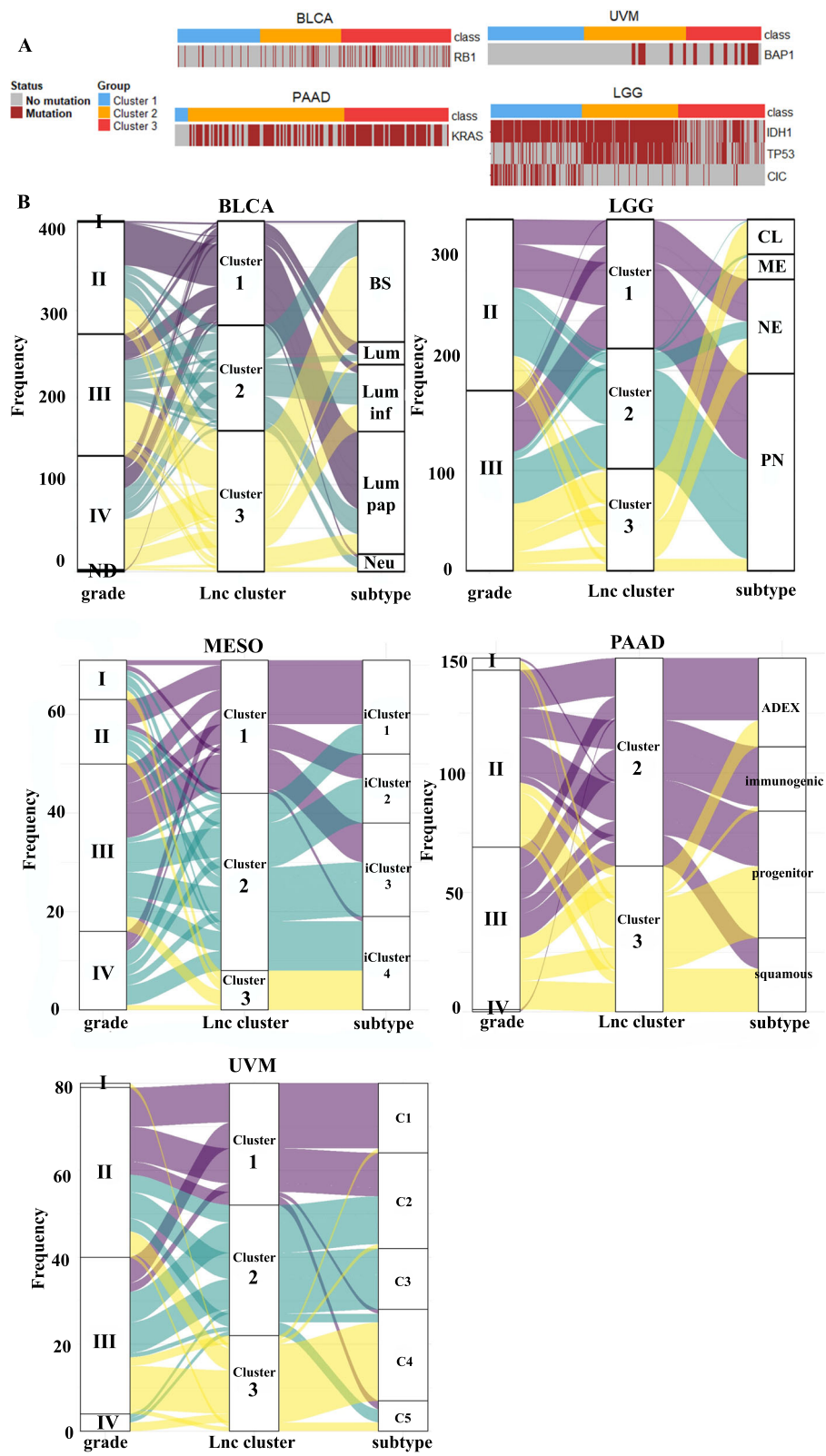
which had the poorest prognosis and exhibited a highly immune-infiltrated characteristic. Meanwhile, cluster 1 BLCA patients were mainly categorized into the luminal papillary type, which had more favorable prognoses. Similarly, in LGG patients, the mesenchymal subtype tended to have poor survival, highly infiltrative immune cells, and EMT activation, and the subtype was mainly enriched in the cluster 3 group. By contrast, patients in cluster 1 were enriched in the proneural and neural subtypes. In PAAD patients, cluster 3 patients were mainly enriched in squamous cell subtypes that possessed activated hypoxia, inflammatory response, TGF- $\beta$  signaling, and MYC pathway activation. For cluster 1 patients, all their histological types belonged to neuroendocrine, which was excluded in the study that defined these molecular subtypes [20]. As for MESO and UVM, no names were used to define their molecular types based on biological features. However, the cluster 4 patients of both UVM and MESO that had the poorest prognosis in previous findings [21, 22] were highly overlapped with cluster 3 patients in the present findings.

#### EMT and inflammation-regulated pathways exhibit distinct patterns in glycolysis-associated lncRNA-classified groups

To delineate underlying signal pathways associated with glycolysis-related lncRNAs, we conducted a DEG analysis within different clusters (Additional file 10: Figure S5). Signal pathways that were upregulated in the poorest prognosis group with the highest glycolysis scores included EMT, inflammatory responses, and interferon gamma signaling (Fig. 3a, b). Therefore, we inferred that lncRNA-based cluster 3 patients exhibited an upregulated status in the EMT and immune regulatory pathways; these lncRNAs might play critical roles in linking glycolytic signaling, EMT, and immune-suppressive microenvironments.

#### Patients with poor prognoses as classified by glycolytic-associated lncRNAs exhibit a highly infiltrative immune microenvironment

To further clarify differences in the immune microenvironments of cluster 1 and cluster 3 patients, we inferred



(See figure on previous page.)

**Fig. 2** Associations of genomic mutations and molecular subtypes in glycolysis-associated long noncoding (lnc)RNA-classified clusters. **a** A heatmap showing gene candidates with nonsilent mutations enriched in certain clusters of bladder carcinoma (BLCA), pancreatic ductal adenocarcinomas (PAADs), uveal melanomas (UVMs), and low-grade gliomas (LGGs). **b** Alluvial plots demonstrating comparisons of lncRNA-stratified subgroups (lnc cluster) with tumor grade (grade) and molecular subtypes (subtype) in BLCA, LGG, mesotheliomas (MESO), PAAD, and UVM. ND, not determined; BS, basal/squamous; LUM, luminal-like; Lum-inf, luminal infiltrated; Lum pap, luminal papillary; Neu, neuroendocrine-like; PN, proneural; NE, neural; CL, classical; ME, mesenchymal

immune cell infiltration scores in the five cancer types by using the expression levels of gene markers in immune cells. We then conducted unsupervised hierarchical clustering to categorize patients into high and low immune infiltration groups (Fig. 4a). We determined that most cluster 3 patients belonged to a high immune infiltration group. By contrast, cluster 1 patients mainly belonged to a low immune infiltration group. Despite the high infiltration of immune cells in cluster 3 patients, they still had the poorest prognosis. Hence, we speculated that cluster 3 cancer patients might have up-regulated immune-suppressive checkpoints that facilitated cancer cells to evade attack by immune cells. Therefore, we compared the expression levels of known immune-suppressive checkpoints in different clusters (Fig. 4b). The results revealed that PDCD1, IDO1, and CTLA4 were significantly upregulated in cluster 3 patients compared with cluster 1 patients in multiple cancer types (Fig. 4c). Taken together, these results suggest that glycolysis-associated lncRNAs can be used to identify a subgroup of cancer patients with the poorest prognosis. They also suggest that lncRNAs activate oncogenic pathways including hypoxia and EMT and promote a highly infiltrative immune tumor environment. Moreover, these cancer subsets may be vulnerable to treatment with immunotherapies due to their enriched immune targets.

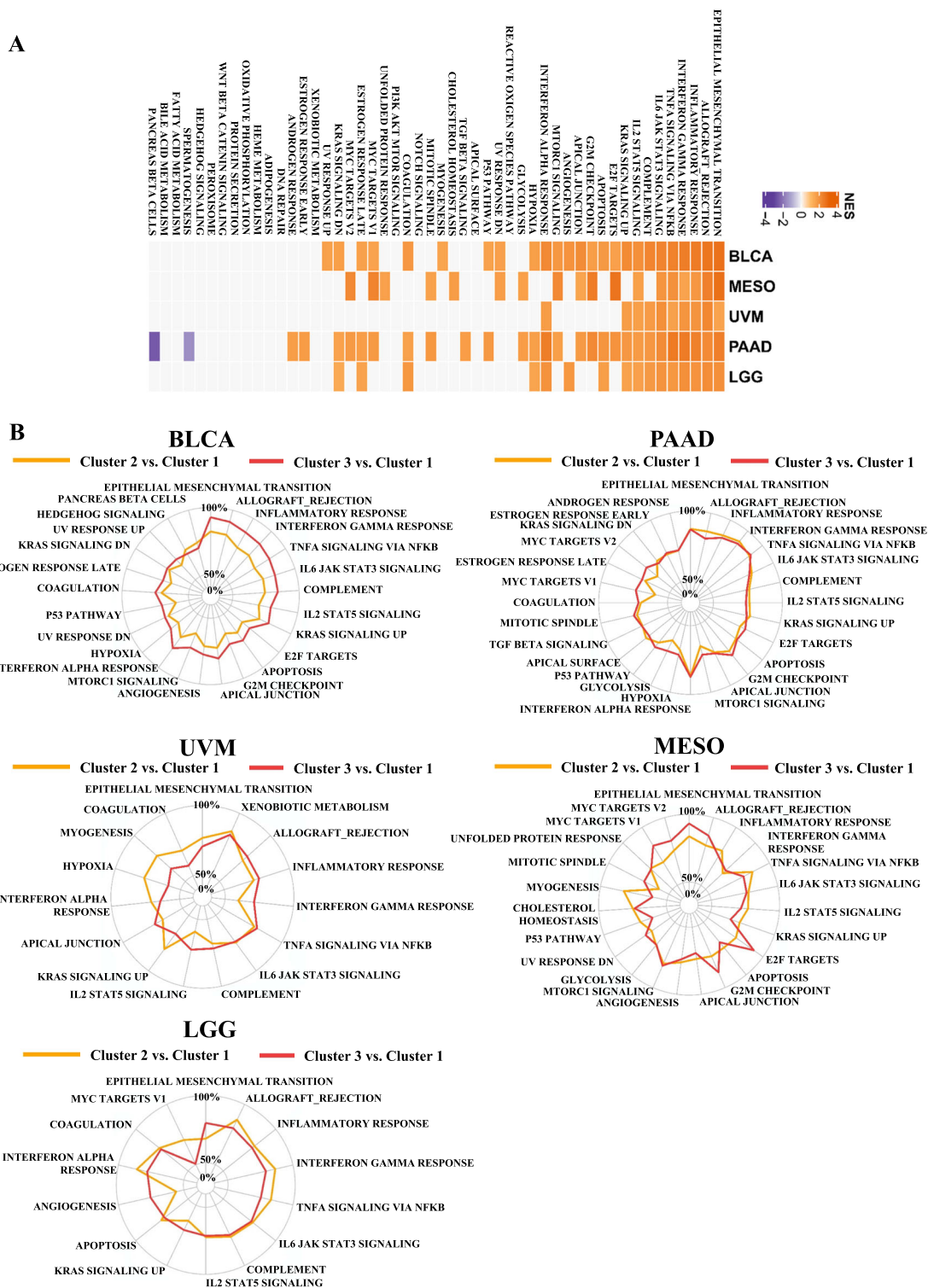
#### Characterization of glycolysis-associated lncRNAs in cancers

Studies have indicated that lncRNAs exhibit tissue-specific expression patterns [23, 24]. Here we found that glycolysis-associated lncRNAs also followed this rule. Most of the lncRNAs were correlated with glycolytic activity in only one cancer type (Fig. 5a). However, common candidates were still associated with glycolytic activity across the five cancers, namely MIR4435-2HG, AC078846.1, AL157392.3, AP001273.1, and RAD51-AS1 (Fig. 5b). The results implied that these five lncRNAs participate in glycolytic signaling in several cancers. Furthermore, most of the glycolysis-associated lncRNAs exhibited a negative correlation with glycolytic activity compared with protein coding genes (Fig. 5c). Because the functionalities of lncRNAs differ based on their location in the cellular compartment, we utilized the LncATLAS database to investigate the localization of these

glycolysis-associated lncRNAs. We determined that 69.22% of the RCIs were negative, which suggests that nearly 70% of lncRNAs were located in nuclei (Fig. 5d). One group of TF activities was highly negatively correlated with glycolysis-associated lncRNAs in nuclei (Fig. 5e). As a previous study indicated [25], nuclear lncRNAs function in chromatin organization, transcriptional and post-transcriptional gene expression regulation, and direct interactions with TFs to facilitate gene expressions.

We then investigated associations between glycolysis-associated lncRNAs and dysregulated TFs among the five cancer types. Despite these cancers having distinct lncRNA patterns correlated with glycolytic activity, they shared common pathways and upregulated TFs in cluster 3 cancer patients. Next, we focused on nuclear lncRNAs that had distinct expressions in cluster 3 patients of the five cancer types. Four lncRNA candidates were selected; MIR4435-2HG was excluded because it was distributed mainly in the cytoplasm (Fig. 5f). Next, we investigated differences in TF activity in glycolysis activity-associated clusters. After conducting an aREA, we inferred TF activity by examining the expressions of TF downstream targets. Then, we identified putative upstream regulators that might be responsible for the poor prognosis of cluster 3 cancer patients (Fig. 5g). The results indicated that 47 TF candidates were dysregulated in the five cancers. Functional annotation analyses of these TFs indicated that most of them were enriched in mothers against decapentaplegic homolog (SMAD) protein complex assembly, cytokine-mediated signal pathways, and the transforming growth factor (TGF)- $\beta$  signal pathway (Fig. 5h). These results further implied the potential roles of glycolysis-associated lncRNAs in cancer immune regulation and EMT activation. Prominent TFs involved in these pathways were HIF1A, FOS, JUN, the SMAD family, and the STAT family (Fig. 5i). In addition, the oncogenic TF, MYC, has been identified as a key regulator of glycolysis. Therefore, we also investigated the potential link between glycolysis-associated lncRNA-stratified clusters and MYC regulation. In pathway analyses (Fig. 3a), MYC targets were significantly activated in PAAD and MESO. In TF analyses, we observed a moderate upregulation of MYC activity in the highly glycolytic groups for BLCA, LGG, and UVM and a significant upregulation in PAAD and MESO (Fig. 5j),





**Fig. 3** Varieties of signaling pathways in distinct clusters stratified by glycolysis-associated long noncoding (lnc)RNAs. **a** A heatmap demonstrating normalized enrichment scores (NESs) of Hallmark pathways calculated by comparing cluster 3 with cluster 1 (with a false discovery rate (FDR) of < 0.01). **b** Radar plots indicating NESs of Hallmark pathways calculated through a gene set enrichment analysis (GSEA) of cluster 3 versus cluster 1 and of cluster 2 versus cluster 1

Additional file 11: Figure S6). The results indicated that these TFs—HIF1A, JUN, LEF1, and FOS—were negatively correlated with four glycolysis-associated lncRNAs across the five cancer types (Fig. 5k), suggesting that the low expression of these lncRNA signatures was accompanied by glycolytic signaling and oncogenic TF activation. Based on the functional annotation analyses of activated TFs, we determined that activated TGF- $\beta$ -mediating SMAD signaling pathways can be used to distinguish these three subtypes. Furthermore, TGF- $\beta$ -mediating SMAD signaling pathways have been identified to be involved in regulating EMT [26] and immune infiltration [27]. Therefore, we identified these three subtypes as cluster 1 (low TGF- $\beta$ /SMAD), cluster 2 (median TGF- $\beta$ /SMAD), and cluster 3 (high TGF- $\beta$ /SMAD).

#### **MIR4435-2HG interlinks glycolysis and its downstream associated genes in cancers**

Finally, we investigated which lncRNAs might potentially be involved in the coexpression of glycolysis-related genes. We performed a first-order partial correlation to adjust correlation coefficients between glycolysis scores and positively associated genes. We observed a significant alteration in correlation coefficients in these glycolysis-associated genes after adjusting for the effect of mutually associated lncRNAs (Fig. 6a). In particular, when the effects of MIR4435-2HG are removed, correlations between glycolysis scores and associated genes significantly decreased in the five cancer types. This suggests that MIR4435-2HG plays a crucial role in linking glycolysis and related signal pathways. Next, we performed a multivariate linear regression to identify MIR4435-2HG-associated genes by adjusting for other covariates, including copy number alterations and DNA methylation (Fig. 6b). MIR4435-2HG-associated genes were linked to EMT and immune signaling across the four cancer types of BLCA, LGG, UVM, and PAAD (Fig. 6c). By contrast, no significantly associated genes were identified in MESO. In addition, we observed that MIR4435-2HG expression was positively associated with MYC activity (which was reported to play prominent roles in glycolysis regulation) across the five cancer types (Additional file 12: Figure S7). Among the top 10 MIR4435-2HG-associated gene candidates in immune signal pathways, cytokines were the most common. In particular, C-X-C motif chemokine 10, a well-established chemotactic cytokine, was strongly correlated with MIR4435-2HG expression in these four cancer types. Consequently, we concluded that lncRNAs, especially MIR4435-2HG, play key roles in linking glycolytic pathways with other oncogenic signal pathways, including EMT and immune regulation, in some cancer types. Besides, all the high-resolution images of results were showed as Additional file 13: Figure S8.

## **Discussion**

Glycolysis, a metabolic pathway, supplies energy, promoting cancer malignancy. Although the evidence of a crosstalk between glycolysis and lncRNA regulation in cancer progression is mounting [28, 29], a clinical link connecting lncRNAs and glycolysis is not fully understood. We conducted a pan-cancer scale analysis to identify glycolysis-associated lncRNAs in 33 cancer types. We determined that the prognoses of five cancer types, BLCA, PAAD, LGG, MESO, and UVM, were significantly correlated with glycolysis-associated lncRNA signatures. Furthermore, we comprehensively analyzed changes in gene mutations, molecular subtypes, TFs, oncogenic signal pathways, and immune cell infiltration in patients stratified by glycolysis-associated lncRNAs. Finally, we identified five lncRNAs, namely MIR4435-2HG, AC078846.1, AL157392.3, AP001273.1, and RAD51-AS1, which exhibited significant correlations with glycolysis across the five cancer types. In particular, MIR4435-2HG was suggested to play a critical role in connecting glycolysis, EMT, and immune cell infiltration in the cancers.

For the glycolysis evaluation of each cancer patient, we chose the ssGSEA-derived glycolysis score instead of the direct correlations of glycolysis-associated genes because of the large number of glycolysis-involved genes in this methodology (200 gene candidates). Directly correlating lncRNAs with these genes would generate hundreds of correlation results and *p* values, making evaluating the degree of association of each lncRNA with glycolysis activation difficult. Therefore, an ssGSEA algorithm was utilized to treat the 200 glycolysis-involved candidates as a gene set to perform a GSEA for individual patients. Through this method, we were able to compress the genes into a single score for each patient and capture the degree of glycolysis activation. We did not directly use the target genes of lncRNAs due to the complexity of the regulatory mechanisms of lncRNAs. The functions of lncRNAs in post-transcriptional regulation include the microRNA sponge, interaction with the chromatin modulator, and direct targeting of downstream RNA. In addition, only focusing on glycolysis-involved genes that belong to lncRNA direct targets would risk our overlooking other promising candidates. Furthermore, no bioinformatics tools that can accurately predict lncRNA target genes are yet available. Therefore, ssGSEA-derived glycolysis scoring may be a suitable tool for identifying glycolysis-associated lncRNAs.

Relationships between glycolysis and the immune resistance of cancers have recently been reported. Tumor cells utilize glucose and metabolically compete with T cells through impairing mammalian target of rapamycin (mTOR) activity and glycolytic activity in T cells, leading to the overriding of the capability of T cell-mediated

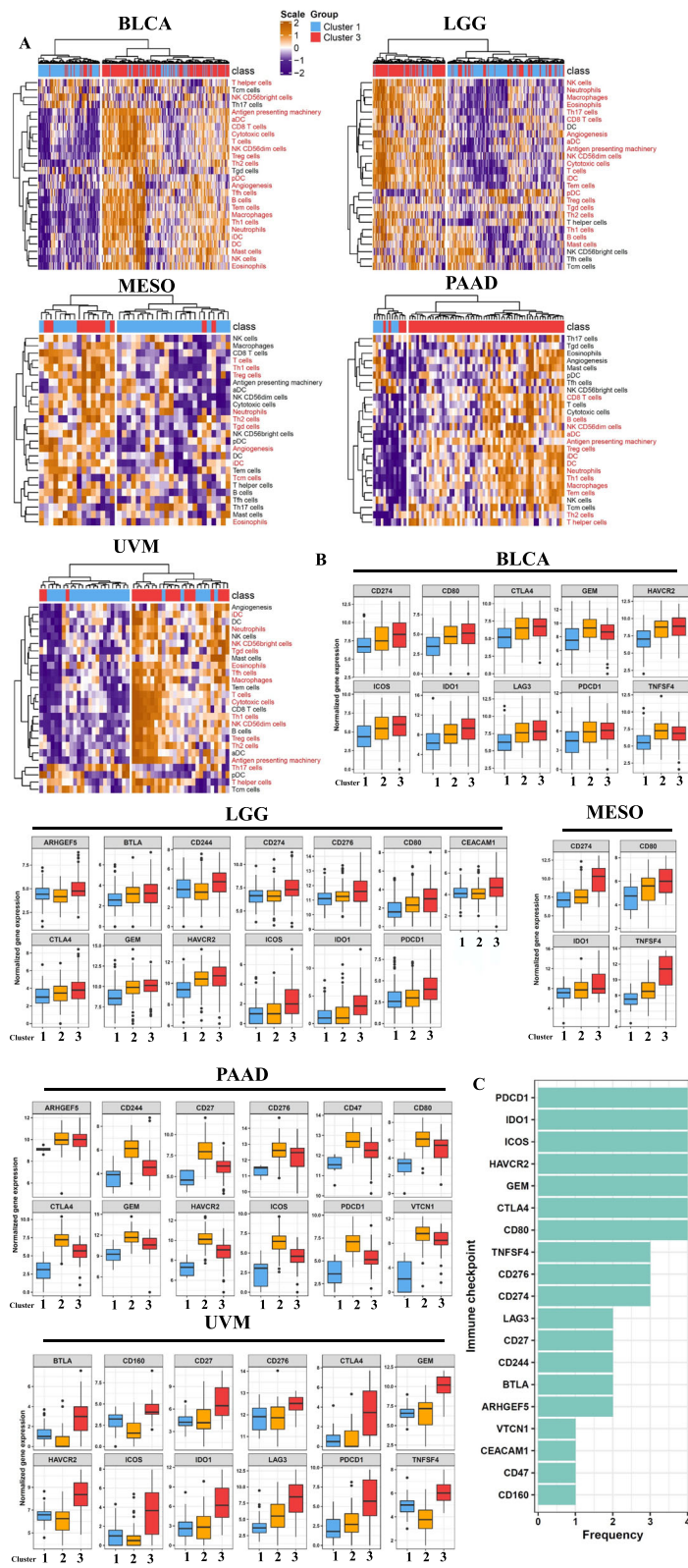
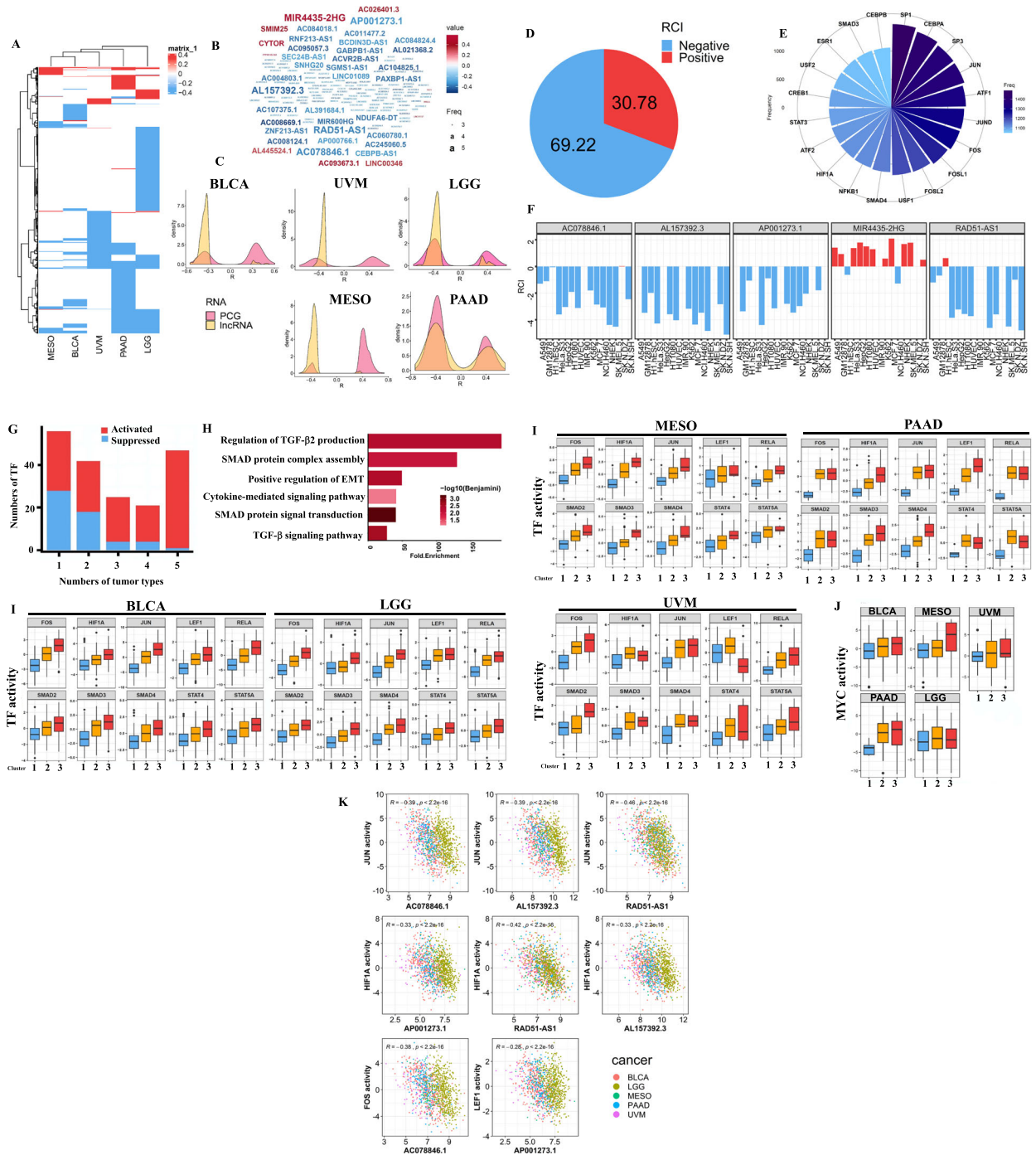


Fig. 4 (See legend on next page.)

(See figure on previous page.)

**Fig. 4** Immune-related signaling is enriched in glycolysis-associated long noncoding (lnc)RNA-stratified cluster 3 cancer patients. **a** Unsupervised clustering of distinct immune cell infiltrations in glycolysis-stratified cluster 3 and cluster 1 cancer patients. **b** A boxplot demonstrating immune checkpoints that were upregulated in cluster 3 (with an effect size of >0.5 and *p* value of <0.001) compared with cluster 1 or cluster 2. **c** Bar plot demonstrating frequencies of immune checkpoints upregulated in cluster 3 cancer patients across the five cancer types. The *y*-axis indicates the names of immune checkpoints, and the *x*-axis represents the number of cancer types



**Fig. 5** (See legend on next page.)

(See figure on previous page.)

**Fig. 5** Characterization of glycolysis-related long noncoding (lnc)RNAs in five cancer types. **a** A heatmap demonstrating glycolysis-associated lncRNAs in five cancer types. Colors in the heatmap indicate the degree of correlation between glycolysis scores and lncRNA expressions. **b** Word cloud of the glycolysis-correlated lncRNAs. Number of cancer types is associated with the text size of the lncRNA names, whereas the color of the lncRNA indicates median correlation coefficients. **c** Density curve showing correlation coefficients of genes (pink) and lncRNAs (yellow) significantly associated with glycolysis scores. Distributions of genes and lncRNAs were compared using the Kolmogorov–Smirnov test. The x-axis represents correlation coefficients with glycolysis scores between genes and lncRNAs, and the y-axis represents numbers of lncRNAs and genes. **d** Pie chart demonstrating percentages of lncRNAs with positive or negative values in the relative concentration index (RCI). **e** A circular bar plot indicating frequencies of transcription factors (TFs) negatively associated with glycolysis-related nuclear lncRNAs. **f** Bar plots demonstrating cellular localizations of five lncRNAs significantly correlated with glycolysis scores across five cancer types. **g** A bar plot showing numbers of activated and suppressed TFs in cluster 3 compared with cluster 1 in five cancer types. The x-axis represents the number of cancer types, and the y-axis represents the number of TFs. **h** Functional annotation analyses of activated TFs involved in more than four cancer types in cluster 3 cancer patients. The y-axis represents significant enriched pathways (with an FDR of < 0.01), and the x-axis represents multiples of enrichment of indicated signal pathways. **i** A boxplot indicating TFs that were upregulated in cluster 3 compared with cluster 1. The y-axis represents transcriptional activity inferred by target gene expression levels by using analytic rank-based enrichment analyses, and the x-axis represents different clusters classified by glycolysis-associated lncRNAs. **j** MYC activity in five cancer types. **k** Dot plots of lncRNA-TF pairs that exhibit negative correlations across five cancer types. The association between TF and lncRNA within each cancer type exhibits a correlation coefficient of < -0.3 and a *p* value of < 0.001. The pooled *p* value and correlation coefficient are indicated in each dot plot

cytotoxicity [30]. In addition, highly activated glucose metabolism in cancer cells promotes lactate accumulation in the TME [31]. This extracellularly accumulating lactate blocks lactate export from T cells, leading to the generation of dysfunctional aerobic glycolysis, a crucial mechanism for maintaining T cell effector function. In addition to immune regulation, metabolic reprogramming (including glycolysis) with highly invasive and drug resistance features was reported to possess the ability to transform cancer cell phenotypes toward EMT [32]. The EMT process has also identified to be linked to immune evasion by cancer cells [33, 34]. However, relationships among glycolysis, EMT, and immune regulation in cancers are still not fully understood. In our analyses, we determined that EMT and inflammatory responses, including immune-suppressive ligand expression, were enriched in cluster 3 cancer patients, as classified by glycolysis-associated lncRNAs. These findings suggest that the TME of cluster 3 cancer patients is surrounded with different types of immune cells, and the upregulated immune-suppressive ligands and EMT activation protect cancer cells from attack by immune cells. Therefore, the “hot” immune microenvironment in cluster 3 cancer patients implies that immune checkpoint blockade therapy might be suitable.

In our genomic mutation analysis, we identified specific gene mutations that were associated with glycolysis-associated lncRNA-stratified clusters in different cancers. Some of these genes have been reported to be involved in glucose metabolism. For example, in mice bearing RB1 null lung cancer [35], upregulation of glucose transporter (GLUT) 1 and two rate-limiting enzymes in glycolysis, hexokinase-2 (HK2) and pyruvate kinase isozymes 2, was observed, suggesting that RB1 regulates glucose metabolism. In the present study, the cluster with high glycolytic activity mainly exhibited RB1 mutation in BLCA. Although loss of RB1 may serve as a poor prognosis predictor in BLCA [36], the relationship

between RB1 and glycolysis in BLCA remains unclear. Additionally, in PAAD and UVM, we identified KRAS and BAP1 mutations, respectively, as enriched in clusters with high glycolytic activity. In PAAD, KRAS mutation activates glycolytic signaling mainly through MEK activation and Myc-dependent transcription, resulting in the upregulation of GLUTs and rate-limiting enzymes of glycolysis, such as HK2, phosphofructokinase-1, and lactate dehydrogenase A (LDHA) [37]. A multi-omics analysis integrating transcriptome, metabolite, and genomic analysis revealed that UVM cells with BAP1 mutation maintain their energy demand through oxidative phosphorylation and glycolytic pathway [38]. Additionally, the distinct metabolic features of mutant BAP1 and wild-type UVM further lead to different responses to metabolic inhibitors. These findings indicate that KRAS and BAP1 mutations respectively drive metabolic alterations in PAAD and UVM. In gliomas, the cluster with high glycolytic activity mainly belongs to IDH1 wild-type cancer patients, as per our findings. A previous study reported that glioma cells with IDH1 mutation produce 2-hydroxyglutarate, which promotes HIF-1 $\alpha$  degradation, accompanied by the downregulation of glycolysis-related genes including *SLC2A1*, *PDK1*, *LDHA*, and *SLC16A3* [39]. By contrast, IDH1 wild-type gliomas exhibited an activated glycolysis pathway. Taken together, our lncRNA-stratified clusters not only exhibit different glycolytic activities but also distinct genomic mutations that have been reported to be involved in glycolysis signaling. However, whether these genomic mutations might be the drivers altering glycolysis-associated lncRNA profiles requires further investigation.

In our TF analysis, we determined that activated prominent TFs in highly glycolytic clusters across BLCA, LGG, PAAD, and UVM were mainly involved in TGF- $\beta$  signaling and SMAD protein complex assembly. SMAD proteins are downstream signal transducers of the TGF-

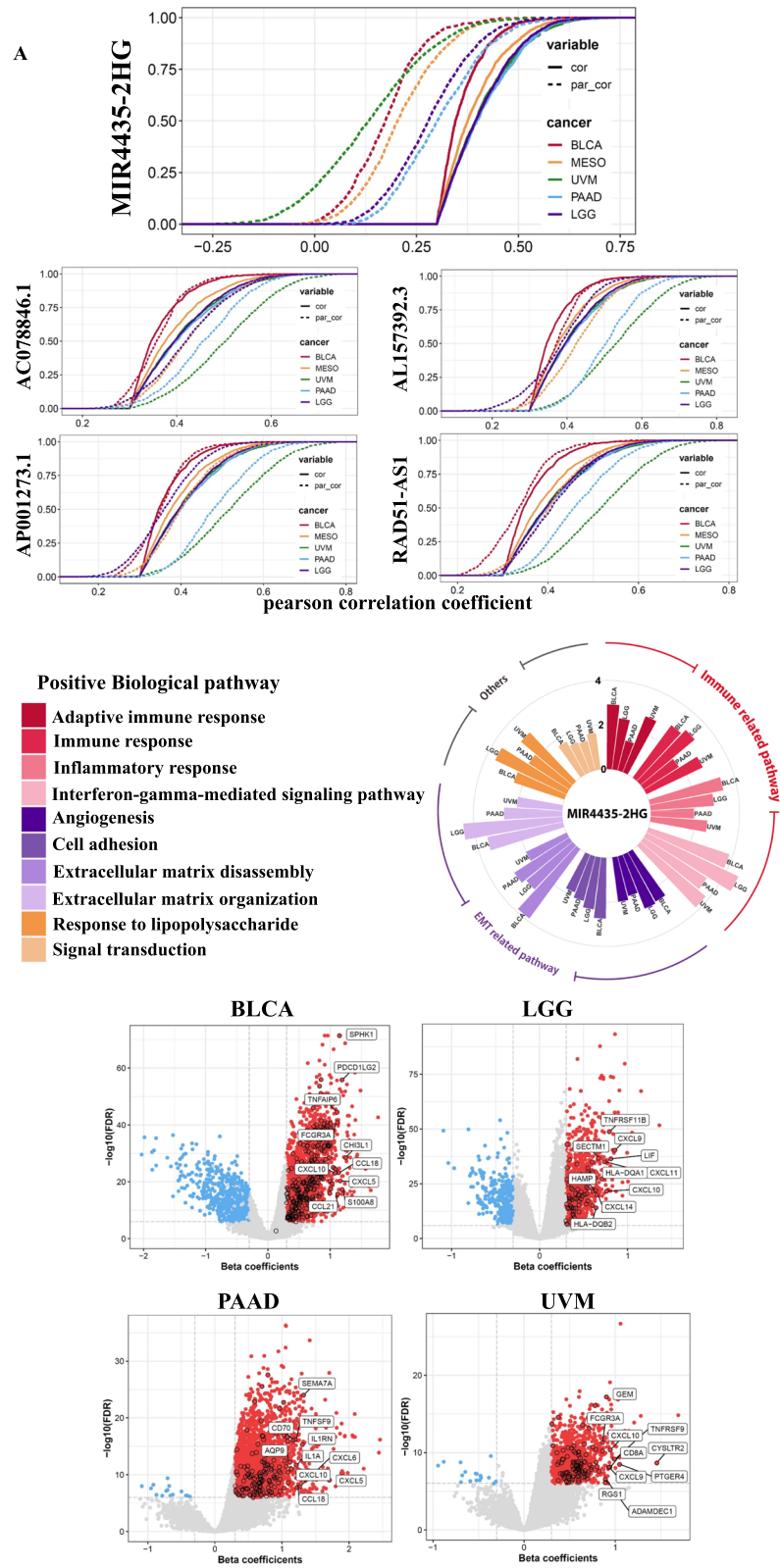


Fig. 6 (See legend on next page.)

(See figure on previous page.)

**Fig. 6** MIR4435-2HG links glycolysis and its related signal pathways including the epithelial-to-mesenchymal transition (EMT) and immune regulation. **a** Cumulative distribution curve (CDF) of glycolysis-gene correlations with or without adjustment for MIR4453-2HG by using a first-order partial correlation. Solid lines indicate CDFs of correlation coefficients between glycolysis scores and gene expressions without adjustment. Dashed lines indicate first-order partial correlation-adjusted relations between glycolysis scores and gene expressions. These two distributions were compared using the Kolmogorov–Smirnov test. The *x*-axis represents Pearson correlation coefficients between glycolysis scores and gene expressions, and the *y*-axis represents cumulative probabilities. **b** A circular bar plot displaying MIR4435-2HG-associated signal pathways. Genes significantly associated with MIR4435-2HG but independent of the methylation status and copy number alterations were identified using a multivariate linear regression. Pathways enriched in MIR4435-2HG-associated genes are shown in the circular bar plot. The height of the bars indicates the degree of enrichment. Only pathways with a false detection rate (FDR) of < 0.01 are included in the plot. **c** A volcano plot indicating MIR4453-2HG-associated genes that were independent of copy number alterations and the DNA methylation status. The *x*-axis represents multivariate linear regression-adjusted coefficients between MIR-4435-2HG and glycolysis scores. Gene candidates involved in immune-related signal pathways are labeled with a black round line, and their names are indicated in the plots

$\beta$  signaling pathway, which functions as an immune-suppressive regulator in cancers. For instance, Smad3-mediated TGF- $\beta$  signaling was reported to suppress the cytotoxic activity of NK cells by blocking the production of CD16-mediated IFN-gamma [40]. Another SMAD protein, SMAD4, has dual roles in regulating NK immunity in a context-specific manner [41]. In the initial phase of tumor formation, SMAD4 has a mainly positive effect in promoting the development and antitumor activity of NK cells. However, at the late stage of cancer development, NK cells are surrounded with TGF- $\beta$  produced by tumor cells. In this condition, SMAD4 cooperates with p-SMAD2 and p-SMAD3 to suppress NK cell-mediated cytotoxicity. Our analyses revealed that the group with high glycolytic activity demonstrated a microenvironment with highly infiltrated NK cells. Furthermore, coordinated upregulation of SMAD2, SMAD3, and SMAD4 was also observed in highly glycolytic patients. Taken together, these results imply that highly glycolytic tumor cells might impede NK-mediated immunity through SMAD signaling, and such tumor cells might be vulnerable to SMAD4 inhibition. However, the relationship between glycolysis-related lncRNAs and SMAD complex regulatory mechanisms in tumor cells needs to be further validated.

Among the glycolysis-associated lncRNAs we identified, some lncRNAs have been reported to be directly involved in glycolysis signaling. For example, plasmacytoma variant translocation 1, which exhibited positive associations with glycolysis scores in six of thirty-three cancer types in our findings, was suggested to function as a microRNA sponge in suppressing miR-497 expression, leading to the promotion of HK2 upregulation and osteosarcoma progression [42]. Nuclear factor (NF)- $\kappa$ B-interacting lncRNA (NKILA), which was positively correlated with glycolytic activity in the five cancer types, activates hypoxia-inducible factor 1 $\alpha$  expression to promote the hypoxia-mediated Warburg effect on gliomas [43]. Some lncRNA candidates in our findings might also be implicated in regulating glycolysis. For instance, cytoskeleton regulator RNA was reported

to interact with Sam68, an RNA-binding protein that possesses an oncogenic function in cancers [44]. Furthermore, Sam68 can upregulate the expression of pyruvate kinase isozymes M2 (PKM2), a key enzyme for glycolysis-dominated energy metabolism [45], through facilitating the transport of PKM2 messenger (m)RNA from the nucleus to cytoplasm [46]. MIR4435-2HG, an lncRNA positively correlated with glycolysis in 20 cancer types, enhances YAP expression for cancer progression [47]. Furthermore, YAP is recognized as a TF and is responsible for upregulating the expression of glucose metabolism enzymes such as HK2 and 6-phosphofructo-2-kinase/fructose-2,6 biphosphatase 3 [48]. However, their roles in the glycolysis process require further investigation.

Among the five lncRNAs that were consistently correlated with glycolysis scores across different cancers in our findings, MIR4435-2HG was identified as being highly associated with glycolysis and its correlated genes. By performing a multivariate linear regression adjustment, we also uncovered that immune and EMT-involved genes are positively correlated with MIR4435-HG expression. Several studies have demonstrated the oncogenic roles of MIR4435-2HG in cancer processes. MIR4435-2HG promotes gastric cancer cell migration and proliferation through Wnt/ $\beta$ -catenin signaling [49]. The upregulation of MIR4435-2HG promotes oral squamous carcinoma cell proliferation through inducing TGF- $\beta$ 1 upregulation [50]. Nevertheless, few studies have indicated links among MIR4435-2HG, glycolysis, and immune regulation in cancers. Our analyses suggest that MIR4435-2HG participates in interrelations among glycolysis, immune resistance, and EMT. The present study has limitations. The lack of public available RNA sequencing data limited us to validate the classification effectiveness of our genome classifier which is predominantly based on the glycolysis-associated lncRNAs or the gene candidates. More clinical data of individuals with other types of cancer or from large cohort studies are required for validation. Several lncRNAs in our findings have rarely been reported, especially those

negatively correlated with glycolysis. These should be further investigated.

## Conclusions

We identified a subgroup of cancer patients stratified by glycolysis-correlated lncRNA signatures with the poorest prognosis, a highly infiltrative immune microenvironment, and EMT activation and thus provide novel aspects for cancer therapy.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12916-021-01925-6>.

**Additional file 1: Table S1.** Glycolysis candidate genes in each cancer types.

**Additional file 2: Figures S1.** Consensus cumulative distribution function (CDF) and delta area for five cancer types.

**Additional file 3: Table S2.** DEGs in LGG. Table S3. DEGs in BLCA.

**Additional file 4: Table S4.** Gene candidates in TCGA BLCA data. Table S5. Ggene candidates in TCGA LGG data.

**Additional file 5: Table S6.** Number of tumor tissues from TCGA pan cancer data.

**Additional file 6: Figures S2.** Flowchart for filtering cancer types and selecting glycolysis-associated long non-coding (lnc) RNAs.

**Additional file 7: Figures S3.** Stratification of patients into different clusters by a consensus clustering analysis.

**Additional file 8: Table S7.** Glycolytic score associated lncRNA in UVM.

**Additional file 9: Figures S4.** Flow chart for establishing genomic classifiers.

**Additional file 10: Figures S5.** Flow chart for Transcription factor activity analyses.

**Additional file 11: Figures S6.** Negative correlation of lncRNA-MYC activity pairs.

**Additional file 12: Figures S7.** Positive correlation of lncRNA-MYC activity pairs.

**Additional file 13: Figures S8.** All the high resolution images in the result section.

## Abbreviations

aREAs: Analytical rank-based enrichment analyses; BLCA: Bladder carcinoma; DEG: Differentially expressed gene; EMT: Epithelial-to-mesenchymal transition; FDR: False discovery rate; GSEA: Gene set enrichment analysis; LGG: Low-grade gliomas; lncRNA: Long noncoding RNA; MESO: Mesotheliomas; NES: Normalized enrichment score; PAAD: Pancreatic ductal adenocarcinoma; RCI: Relative concentration index; RNA-Seq: RNA sequencing; ssGSEA: Single-sample gene set enrichment analysis; TCGA: The Cancer Genome Atlas; TF: Transcription factor; TME: Tumor microenvironment; UVM: Uveal melanomas

## Acknowledgements

We would like to thank other professors in the university for suggestions. This is to evaluate the degree of glycolysis activation.

## Authors' contributions

K-HH and K-CC conceived of the present idea. K-HH, P-HC, and Y-TL carried out the experiments and analyses. K-HH wrote the manuscript. T-WH, C-MS, and A-JL verified the analytical methods. K-CC supervised the findings of this work and contributed to the final version of the manuscript. All authors discussed the results, provided critical feedback, and helped shape the research, analysis and manuscript. The authors read and approved the final manuscript.

## Funding

This study was sponsored by the Ministry of Science and Technology, Taiwan (contract grant no. MOST 109-2320-B-038-014 to Ku-Chung Chen), Taipei City Government (contract grant no. 10801-62-082 to Ann-Jeng Liu), and Taipei City Hospital Ren-Ai Branch (contract grant no. TPCH-109-10 to Ann-Jeng Liu). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Availability of data and materials

The datasets used and/or analyzed are available from the corresponding author on reasonable request.

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that there are no conflicts of interest associated with this work.

## Author details

<sup>1</sup>Graduate Institute of Medical Sciences, College of Medicine, Taipei Medical University, Taipei, Taiwan. <sup>2</sup>Department of Biochemistry and Molecular Cell Biology, School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan. <sup>3</sup>Department of Microbiology and Immunology, School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan. <sup>4</sup>Department of Neurosurgery, Taipei City Hospital Ren-Ai Branch, Taipei, Taiwan.

Received: 24 September 2020 Accepted: 21 January 2021

Published online: 25 February 2021

## References

- Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science*. 2009; 324(5930):1029–33.
- Dang CV. Links between metabolism and cancer. *Genes Dev*. 2012;26(9): 877–90.
- Liu M, Quek LE, Sultani G, Turner N. Epithelial-mesenchymal transition induction is associated with augmented glucose uptake and lactate production in pancreatic ductal adenocarcinoma. *Cancer Metab*. 2016;4:19.
- Jiang Z, Liu Z, Li M, Chen C, Wang X. Increased glycolysis correlates with elevated immune activity in tumor immune microenvironment. *EBioMedicine*. 2019;42:431–42.
- Ali MM, Akhade VS, Koslali ST, Subhash S, Statello L, Meryet-Figuire M, Abrahamsson J, Mondal T, Kanduri C. PAN-cancer analysis of S-phase enriched lncRNAs identifies oncogenic drivers and biomarkers. *Nat Commun*. 2018;9(1):883.
- Li Y, Jiang T, Zhou W, Li J, Li X, Wang Q, Jin X, Yin J, Chen L, Zhang Y, et al. Pan-cancer characterization of immune-related lncRNAs identifies potential oncogenic biomarkers. *Nat Commun*. 2020;11(1):1000.
- Mitra R, Chen X, Greenawalt EJ, Maulik U, Jiang W, Zhao Z, Eischen CM. Decoding critical long non-coding RNA in ovarian cancer epithelial-to-mesenchymal transition. *Nat Commun*. 2017;8(1):1604.
- Zhang L, Fu Y, Guo H. C-Myc-induced long non-coding RNA small nucleolar RNA host gene 7 regulates glycolysis in breast cancer. *J Breast Cancer*. 2019; 22(4):533–47.
- Wang Y, Lu JH, Wu QN, Jin Y, Wang DS, Chen YX, Liu J, Luo XJ, Meng Q, Pu HY, et al. lncRNA LINRIS stabilizes IGF2BP2 and promotes the aerobic glycolysis in colorectal cancer. *Mol Cancer*. 2019;18(1):174.
- Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015;1(6):417–25.
- de Jong JJ, Liu Y, Robertson AG, Seiler R, Groeneveld CS, van der Heijden MS, Wright JL, Douglas J, Dall'Era M, Crabb SJ, et al. Long non-coding RNAs identify a subset of luminal muscle-invasive bladder cancer patients with favorable prognosis. *Genome Med*. 2019;11(1):60.
- Robertson AG, Kim J, Al-Ahmadie H, Bellmunt J, Guo G, Cherniack AD, Hinoue T, Laird PW, Hoadley KA, Akbani R, et al. Comprehensive molecular



- characterization of muscle-invasive bladder cancer. *Cell*. 2017;171(3):540–556.e525.
13. Ceccarelli M, Barthel FP, Malta TM, Sabedot TS, Salama SR, Murray BA, Morozova O, Newton Y, Radenbaugh A, Pagnotta SM, et al. Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse Glioma. *Cell*. 2016;164(3):550–63.
  14. Garcia-Alonso L, Iorio F, Matchan A, Fonseca N, Jaaks P, Peat G, Pignatelli M, Falcone F, Benes CH, Dunham I, et al. Transcription factor activities enhance markers of drug sensitivity in cancer. *Cancer Res*. 2018;78(3):769–80.
  15. Mas-Ponte D, Carlevaro-Fita J, Palumbo E, Hermoso Pulido T, Guigo R, Johnson R. LncAtlas database for subcellular localization of long noncoding RNAs. *Rna*. 2017;23(7):1080–7.
  16. Senbabaoglu Y, Gejman RS, Winer AG, Liu M, Van Allen EM, de Velasco G, Miao D, Ostrovskaya I, Drill E, Luna A, et al. Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. *Genome Biol*. 2016;17(1):231.
  17. Liu Y, He M, Wang D, Diao L, Liu J, Tang L, Guo S, He F, Li D. HisgAtlas 1.0: a human immunosuppression gene database. *Database (Oxford)*. 2017;2017: bax094.
  18. Kamoun A, de Reyniès A, Allory Y, Sjödhall G, Robertson AG, Seiler R, Hoadley KA, Groeneveld CS, Al-Ahmadie H, Choi W, et al. A consensus molecular classification of muscle-invasive bladder cancer. *Eur Urol*. 2020; 77(4):420–33.
  19. Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, Miller CR, Ding L, Golub T, Mesirov JP, et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell*. 2010;17(1): 98–110.
  20. Bailey P, Chang DK, Nones K, Johns AL, Patch AM, Gingras MC, Miller DK, Christ AN, Bruxner TJ, Quinn MC, et al. Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature*. 2016;531(7592):47–52.
  21. Hmeljak J, Sanchez-Vega F, Hoadley KA, Shih J, Stewart C, Heiman D, Tarpey P, Danilova L, Drill E, Gibb EA, et al. Integrative molecular characterization of malignant pleural mesothelioma. *Cancer Discov*. 2018;8(12):1548–65.
  22. Robertson AG, Shih J, Yau C, Gibb EA, Oba J, Mungall KL, Hess JM, Uzunangelov V, Walter V, Danilova L, et al. Integrative analysis identifies four molecular and clinical subsets in uveal melanoma. *Cancer Cell*. 2017;32(2): 204–220.e215.
  23. Ward M, McEwan C, Mills JD, Janitz M. Conservation and tissue-specific transcription patterns of long noncoding RNAs. *J Hum Transcr*. 2015; 1(1):2–9.
  24. Jiang C, Li Y, Zhao Z, Lu J, Chen H, Ding N, Wang G, Xu J, Li X. Identifying and functionally characterizing tissue-specific and ubiquitously expressed human lncRNAs. *Oncotarget*. 2016;7(6):7120–33.
  25. Sun Q, Hao Q, Prasanth KV. Nuclear long noncoding RNAs: key regulators of gene expression. *Trends Genet*. 2018;34(2):142–57.
  26. Valcourt U, Kowanetz M, Niimi H, Heldin CH, Moustakas A. TGF-beta and the Smad signaling pathway support transcriptomic reprogramming during epithelial-mesenchymal cell transition. *Mol Biol Cell*. 2005;16(4):1987–2002.
  27. Battle E, Massagué J. Transforming growth factor-β signaling in immunity and cancer. *Immunity*. 2019;50(4):924–40.
  28. Wei S, Fan Q, Yang L, Zhang X, Ma Y, Zong Z, Hua X, Su D, Sun H, Li H, et al. Promotion of glycolysis by HOTAIR through GLUT1 upregulation via mTOR signaling. *Oncol Rep*. 2017;38(3):1902–8.
  29. Zou ZW, Ma C, Medoro L, Chen L, Wang B, Gupta R, Liu T, Yang XZ, Chen TT, Wang RZ, et al. LncRNA ANRIL is up-regulated in nasopharyngeal carcinoma and promotes the cancer progression via increasing proliferation, reprogramming cell glucose metabolism and inducing side-population stem-like cancer cells. *Oncotarget*. 2016;7(38): 61741–54.
  30. Chang CH, Qiu J, O'Sullivan D, Buck MD, Noguchi T, Curtis JD, Chen Q, Gindin M, Gubin MM, van der Windt GJ, et al. Metabolic competition in the tumor microenvironment is a driver of cancer progression. *Cell*. 2015;162(6): 1229–41.
  31. Fischer K, Hoffmann P, Voelkl S, Meidenbauer N, Ammer J, Edinger M, Gottfried E, Schwarz S, Rothe G, Hoves S, et al. Inhibitory effect of tumor cell-derived lactic acid on human T cells. *Blood*. 2007;109(9):3812–9.
  32. Kang H, Kim H, Lee S, Youn H, Youn B. Role of metabolic reprogramming in epithelial–mesenchymal transition (EMT). *Int J Mol Sci*. 2019;20(8):2042.
  33. Alsuliman A, Colak D, Al-Harazi O, Fitwi H, Tulbah A, Al-Tweigeri T, Al-Alwan M, Ghebesh H. Bidirectional crosstalk between PD-L1 expression and epithelial to mesenchymal transition: significance in claudin-low breast cancer cells. *Mol Cancer*. 2015;14:149.
  34. Datar I, Schalper KA. Epithelial-mesenchymal transition and immune evasion during lung cancer progression: the chicken or the egg? *Clin Cancer Res*. 2016;22(14):3422–4.
  35. Conroy LR, Dougherty S, Krueer T, Metcalf S, Lorkiewicz P, He L, Yin X, Zhang X, Arumugam S, Young LEA, et al. Loss of Rb1 enhances glycolytic metabolism in Kras-driven lung tumors in vivo. *Cancers (Basel)*. 2020;12(1): 237.
  36. Shariat SF, Tokunaga H, Zhou J, Kim J, Ayala GE, Benedict WF, Lerner SP. p53, p21, pRB, and p16 expression predict clinical outcome in cystectomy with bladder cancer. *J Clin Oncol*. 2004;22(6):1014–24.
  37. Ying H, Kimmelman AC, Lyssiotis CA, Hua S, Chu GC, Fletcher-Sananikone E, Locasale JW, Son J, Zhang H, Colloff JL, et al. Oncogenic Kras maintains pancreatic tumors through regulation of anabolic glucose metabolism. *Cell*. 2012;149(3):656–70.
  38. Han A, Purwin TJ, Bechtel N, Liao C, Chua V, Seifert E, Sato T, Schug ZT, Speicher DW, William Harbour J et al: BAP1 mutant uveal melanoma is stratified by metabolic phenotypes with distinct vulnerability to metabolic inhibitors. *Oncogene*. 2021;40:618–32.
  39. Chesnelong C, Chaumeil MM, Blough MD, Al-Najjar M, Stechishin OD, Chan JA, Pieper RO, Ronen SM, Weiss S, Luchman HA, et al. Lactate dehydrogenase A silencing in IDH mutant gliomas. *Neuro-oncology*. 2014; 16(5):686–95.
  40. Trotta R, Dal Col J, Yu J, Ciarlariello D, Thomas B, Zhang X, Allard J 2nd, Wei M, Mao H, Byrd JC, et al. TGF-beta utilizes SMAD3 to inhibit CD16-mediated IFN-gamma production and antibody-dependent cellular cytotoxicity in human NK cells. *J Immunol*. 2008;181(6):3784–92.
  41. Wang Y, Chu J, Yi P, Dong W, Saultz J, Wang Y, Wang H, Scoville S, Zhang J, Wu LC, et al. SMAD4 promotes TGF-β-independent NK cell homeostasis and maturation and antitumor immunity. *J Clin Invest*. 2018;128(11):5123–36.
  42. Song J, Wu X, Liu F, Li M, Sun Y, Wang Y, Wang C, Zhu K, Jia X, Wang B, et al. Long non-coding RNA PVT1 promotes glycolysis and tumor progression by regulating miR-497/HK2 axis in osteosarcoma. *Biochem Biophys Res Commun*. 2017;490(2):217–24.
  43. Chen Z, Li S, Shen L, Wei X, Zhu H, Wang X, Yang M, Zheng X. NF-kappa B interacting long noncoding RNA enhances the Warburg effect and angiogenesis and is associated with decreased survival of patients with gliomas. *Cell Death Dis*. 2020;11(5):323.
  44. Wang X, Yu H, Sun W, Kong J, Zhang L, Tang J, Wang J, Xu E, Lai M, Zhang H. The long non-coding RNA CYTOR drives colorectal cancer progression by interacting with NCL and Sam68. *Mol Cancer*. 2018;17(1):110.
  45. Zhang Z, Deng X, Liu Y, Liu Y, Sun L, Chen F. PKM2, function and expression and regulation. *Cell Biosci*. 2019;9:52.
  46. Zhao J, Li J, Hassan W, Xu D, Wang X, Huang Z. Sam68 promotes aerobic glycolysis in colorectal cancer by regulating PKM2 alternative splicing. *Ann Transl Med*. 2020;8(7):459.
  47. Dong X, Yang Z, Yang H, Li D, Qiu X. Long non-coding RNA MIR4435-2HG promotes colorectal cancer proliferation and metastasis through miR-206/YAP1 Axis. *Front Oncol*. 2020;10:160.
  48. Zhang X, Zhao H, Li Y, Xia D, Yang L, Ma Y, Li H. The role of YAP/TAZ activity in cancer metabolic reprogramming. *Mol Cancer*. 2018;17(1):134.
  49. Wang H, Wu M, Lu Y, He K, Cai X, Yu X, Lu J, Teng L. LncRNA MIR4435-2HG targets desmoplakin and promotes growth and metastasis of gastric cancer by activating Wnt/β-catenin signaling. *Aging (Albany NY)*. 2019;11(7):6657–73.
  50. Shen H, Sun B, Yang Y, Cai X, Bi L, Deng L, Zhang L. MIR4435-2HG regulates cancer cell behaviors in oral squamous cell carcinoma cell growth by upregulating TGF-β1. *Odontology*. 2020;108(4):553–9.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.