

## Supplementary Information

# Identification of glycogene-type and validation of ST3GAL6 as a biomarker predicts clinical outcome and cancer cell invasion in urinary bladder cancer

Sumiya Dalangood<sup>1,2\*</sup>, Zhen Zhu<sup>1\*</sup>, Zhihui Ma<sup>2,3</sup>, Jiaxuan Li<sup>1</sup>, Qinghe Zeng<sup>2,3</sup>, Yilin Yan<sup>4</sup>, Bing Shen<sup>4</sup>, Jun Yan<sup>5,6#</sup>, Ruimin Huang<sup>2,3#</sup>

1. MOE Key Laboratory of Model Animals for Disease Study, Model Animal Research Center of Nanjing University, Nanjing 210061, China
2. State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China
3. University of Chinese Academy of Sciences, Beijing 100049, China
4. Department of Urology, Shanghai General Hospital, Shanghai Jiaotong University School of Medicine, Shanghai 200080, China
5. Department of Laboratory Animal Science, Fudan University, Shanghai 200032, China
6. Model Animal Research Center of Nanjing University, Nanjing 210061, China

\*These authors contributed equally to this article.

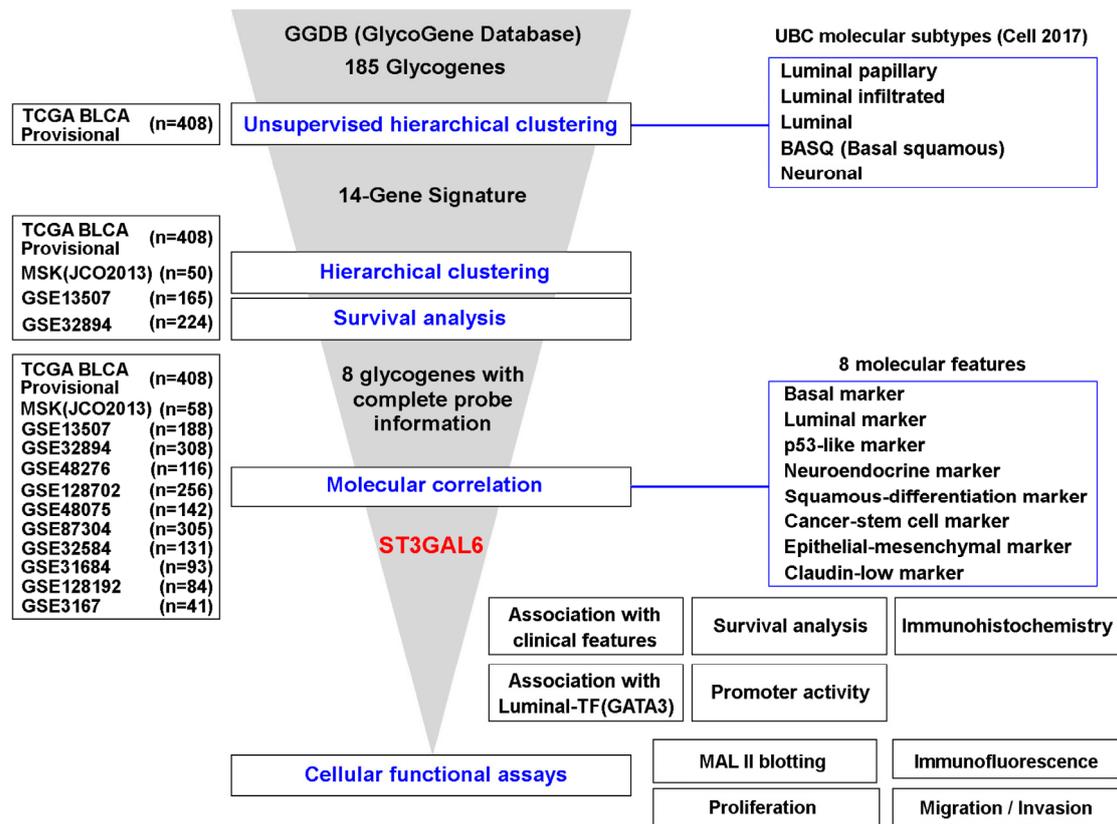
### #Corresponding authors:

Jun Yan, Ph.D., Department of Laboratory Animal Science, Fudan University, Shanghai 200032, China; Phone: 86-21-54237454; Email: [yan\\_jun@fudan.edu.cn](mailto:yan_jun@fudan.edu.cn);

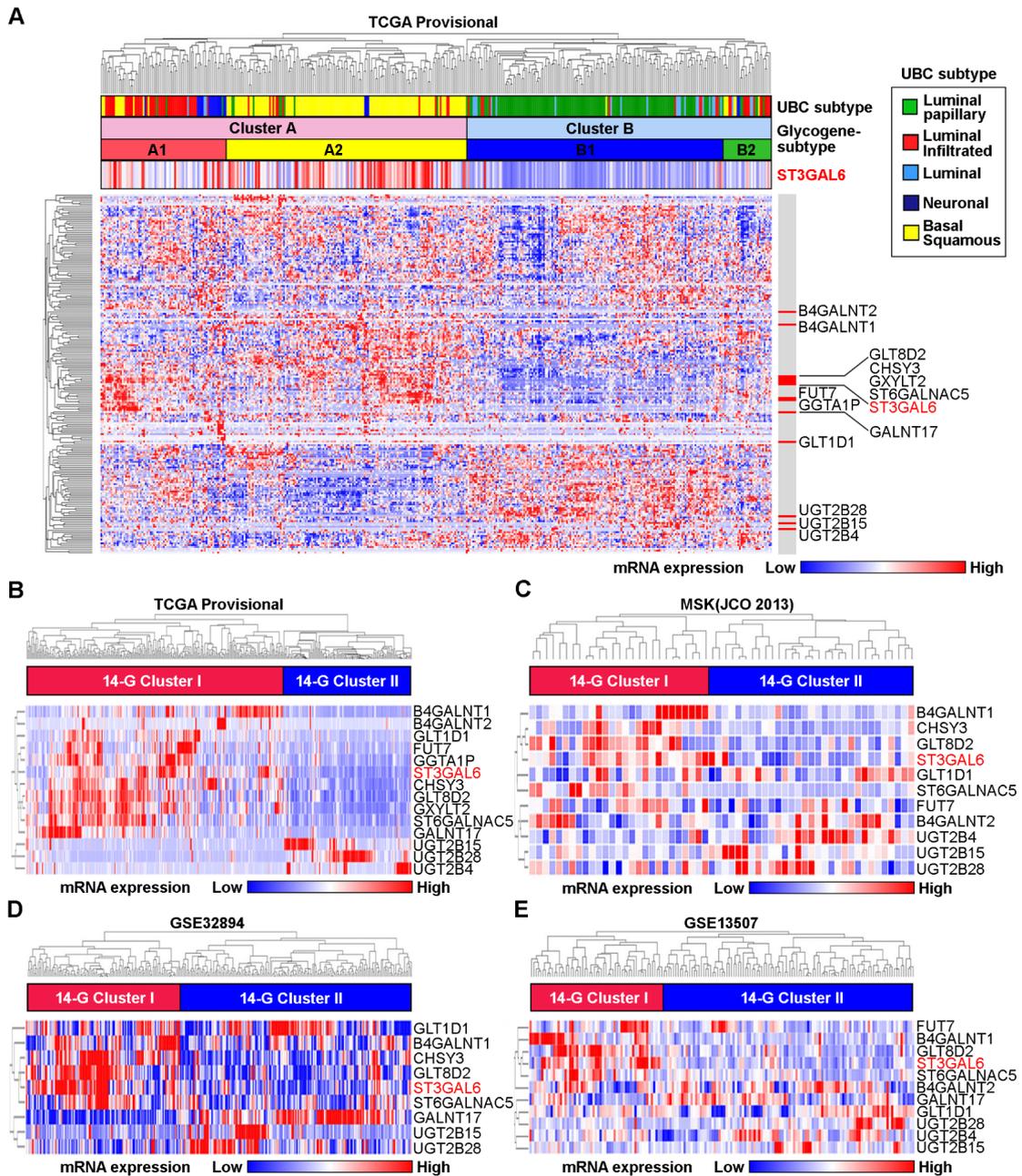
Ruimin Huang, Ph.D., Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China; Phone: 86-21-50817066; Email: [rmhuang@simmm.ac.cn](mailto:rmhuang@simmm.ac.cn)

## Contents

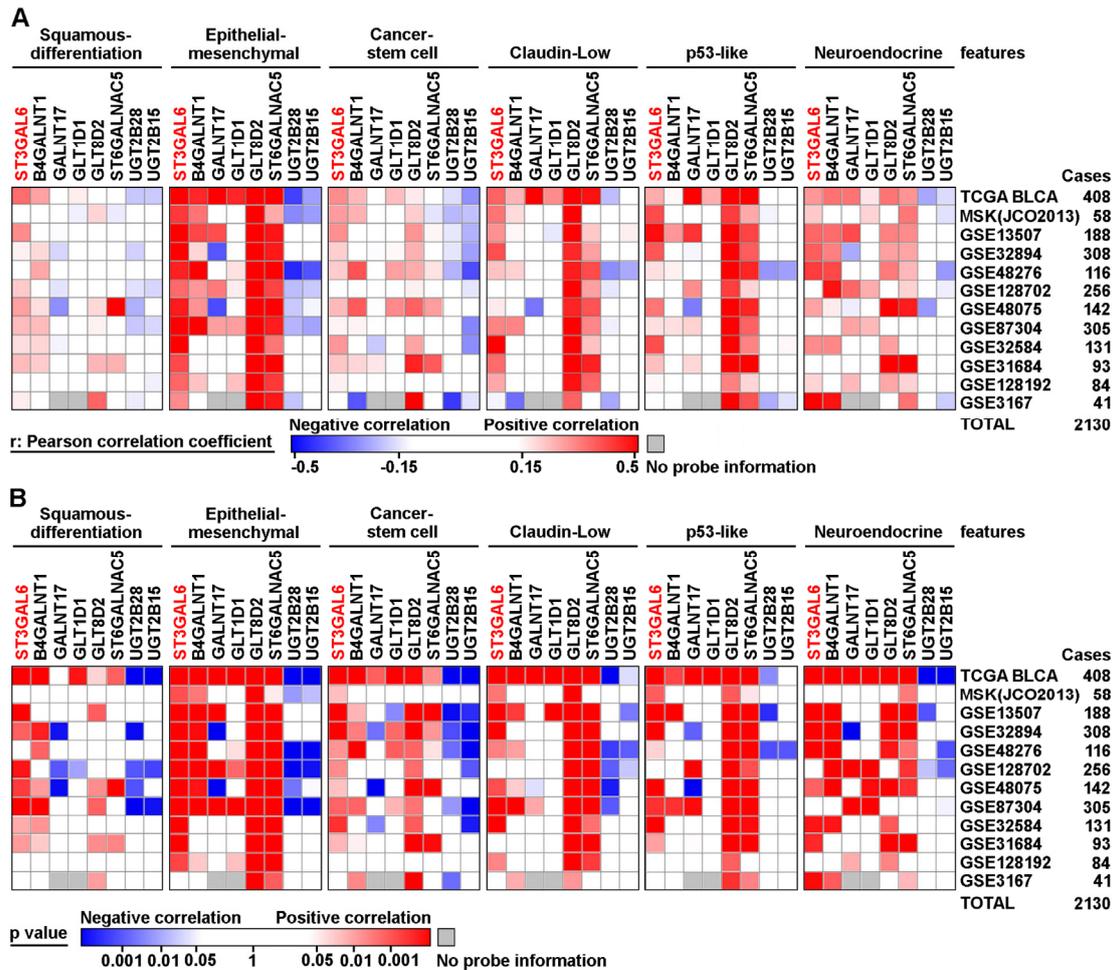
|   |    |
|---|----|
| <b>1. Supplementary Figure S1.</b> Overview for the analysis procedures in this study .....   | 3  |
| <b>2. Supplementary Figure S2.</b> A 14-gene signature for UBC glycogene-type classification .....  | 4  |
| <b>3. Supplementary Figure S3.</b> The correlations between mRNA expressional levels of 14-glycogene signature and 6 molecular features in UBC patients .....                     | 5  |
| <b>4. Supplementary Figure S4.</b> The correlations between mRNA expression level of ST3GAL6 and clinical features in UBC patients .....  | 6  |
| <b>5. Supplementary Figure S5.</b> The Pearson correlation of mRNA expression levels between ST3GAL6 and GATA3 gene in five independent UBC cohorts .....                         | 7  |
| <b>6. Supplementary Figure S6.</b> The conserved GATA3 binding site in the intron 1 of ST3GAL6 is required for GATA3 to downregulate ST3GAL6 expression .....                     | 8  |
| <b>7. Supplementary Figure S7.</b> Biological effects of ST3GAL6 knockdown by shRNAs on UBC cells .....   | 9  |
| <b>8. Supplementary Table S1.</b> List of oligonucleotide sequences and reagents .....  | 10 |
| <b>9. Supplementary Table S2.</b> The clinicopathological associations among 4 clusters based on glycogene-type in TCGA-BLCA provisional dataset .....                            | 12 |
| <b>10. Supplementary Table S3.</b> The 14 glycogenes differentiated the subcluster B1 from other three subclusters in TCGA provisional dataset .....                              | 13 |
| <b>11. Supplementary Table S4.</b> Univariate and multivariate analyses of 14-glycogene signature for patients' outcomes in TCGA provisional dataset .....                        | 14 |
| <b>12. Supplementary Table S5.</b> Univariate and multivariate analyses of 14-glycogene signature for overall survival in GSE32894 dataset .....                                  | 15 |
| <b>13. Supplementary Table S6.</b> Univariate and multivariate analyses of 14-glycogene signature for patients' outcomes in GSE13507 dataset .....                                | 16 |
| <b>14. Supplementary Table S7.</b> The association between the protein expression level of ST3GAL6 and clinicopathological parameters in 52 UBC samples from our IHC cohort ..... | 17 |
| <b>15. Supplementary Table S8.</b> Univariate analysis of ST3GAL6 expression with overall survival in 52 UBC samples from our IHC cohort .....                                    | 18 |



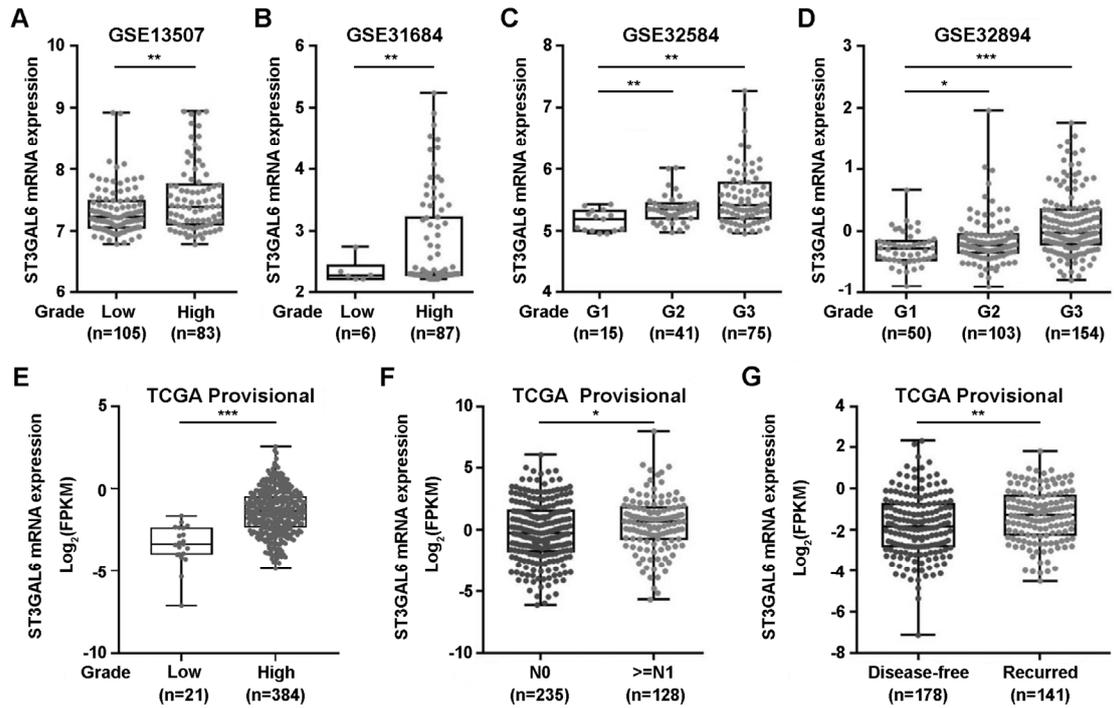
**Supplementary Figure S1. Overview for the analysis procedure in this study.** 1) Unsupervised hierarchical clustering of 185 glycoGenes in TCGA-BLCA provisional dataset, showing the high correlation between glycoGene-types and molecular subtypes of UBC (Cell 2017). 2) Using hierarchical clustering along with survival analysis, a 14-glycoGene signature for glycoGene-type classification was identified and confirmed in TCGA-BLCA provisional, MSK (JCO 2013), GSE13507, and GSE32894 datasets. 3) By Pearson correlation analysis, molecular correlation between 8 glycoGenes from the 14-glycoGene signature with complete probe information and 8 molecular features for UBC patients was analyzed. In 12 datasets, ST3GAL6 gene was shown consistent changes, negative association with Luminal subtype. 4) Correlations between ST3GAL6 gene expression and tumor stage, grade, lymph node invasion, recurrence, survival, and GATA3 gene expression were analyzed. 5) Bioinformatics prediction for transcriptional factor binding site, promoter activity, MAL II blotting, immunofluorescence staining, cell proliferation, cell migration (wound-healing) and cell invasion (transwell) assays were performed to investigate the negative regulation of GATA3, luminal-specific transcription factor, on ST3GAL6 gene expression and further functions.



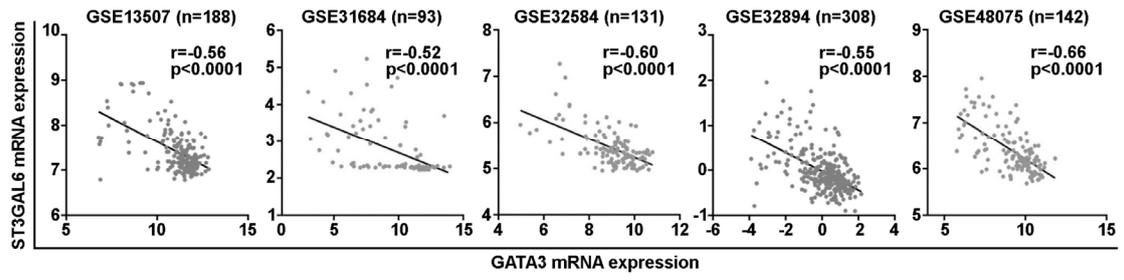
**Supplementary Figure S2. A 14-gene signature for UBC glycogene-type classification.** **A**, 14 glycoenzymes were identified to be significantly dysregulated in subcluster B1 comparing with subclusters A1, A2 and B2 (TCGA provisional dataset). This 14-glycoenzyme signature allowed the glycogene-based classification (14-G Cluster I and II) in UBC patients from TCGA provisional (**B**), MSK (JCO 2013, **C**), GSE32894 (**D**) and GSE13507 (**E**) datasets.



**Supplementary Figure S3. The correlations between mRNA expressional levels of 14-glycogene signature and 6 molecular features in UBC patients.** Pearson correlation analyses ( $r$  value, **A**) with the statistical significance of correlation ( $p$  value, **B**) were performed between the expression levels of 8 glycogenes from the 14-gene signature with the complete probe information and 6 molecular features (Squamous-differentiation, Epithelial-mesenchymal, cancer-stem cell, Claudin-low, p53-like and neuroendocrine) in 12 independent cohorts ( $n=2,130$  in total). Colored-blocks in the heatmaps represented the glycogenes with upregulation (in red) or downregulation (in blue), respectively.

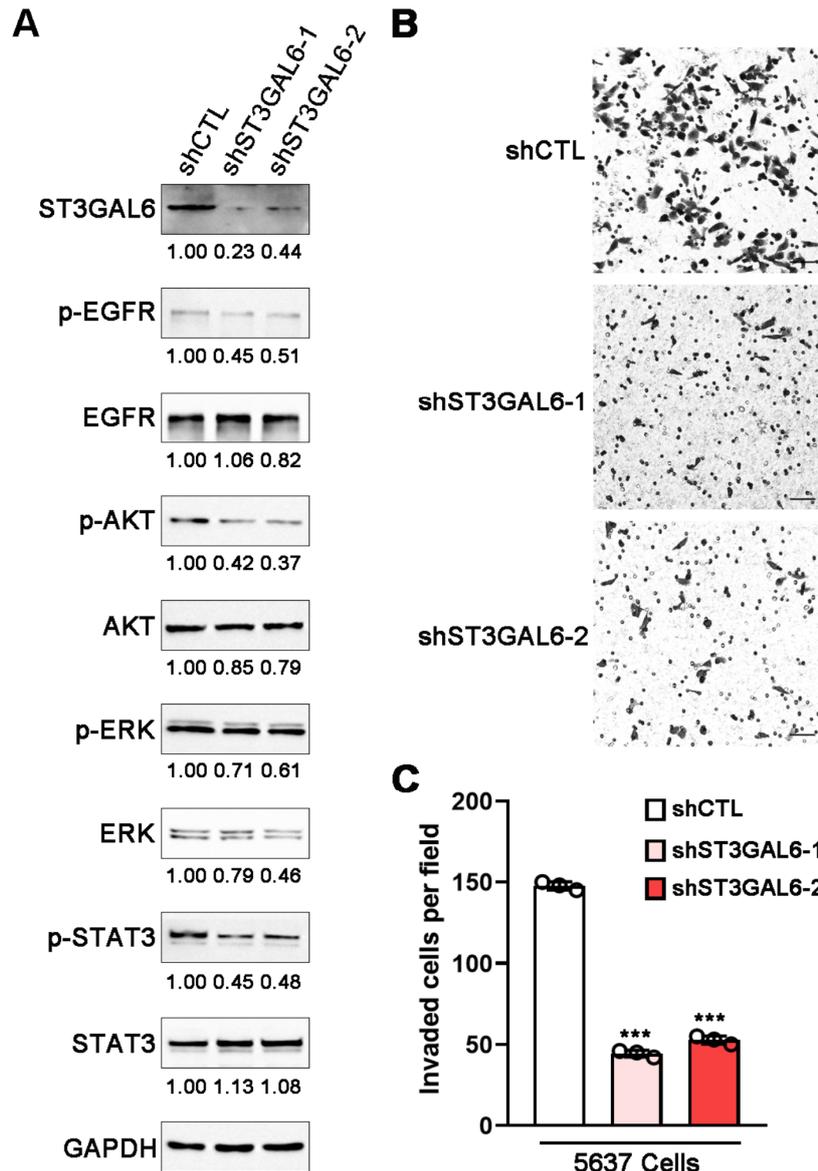


**Supplementary Figure S4. The correlations between mRNA expression level of ST3GAL6 and clinical features in UBC patients. A-E**, The correlation between ST3GAL6 mRNA levels and tumor grade of UBC patients from GES13507 (**A**), GES31684 (**B**), GES32584 (**C**), GES32894 (**D**), and TCGA provisional (**E**) datasets. The association of ST3GAL6 mRNA levels with the status of lymph node metastasis (**F**) and recurrence (**G**) in TCGA provisional dataset. G1, grade 1; G2, grade 2; G3, grade 3. >=N1, with lymph node metastasis; N0, without lymph node metastasis. \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$ .



**Supplementary Figure S5. The Pearson correlation of mRNA expression levels between ST3GAL6 and GATA3 genes in five independent UBC cohorts.**





**Supplementary Figure 7. Biological effects of ST3GAL6 knockdown by shRNAs on UBC cells.** **A**, Effects of ST3GAL6 knockdown by shRNAs on RAS-RAF-MEK-ERK, PI3K-AKT and JAK-STAT pathways in 5637 UBC cells, detected by Western blotting analysis. The normalized protein level (the ratio of test protein to GAPDH protein from same sample) was indicated below each band by ImageJ software. **B**, Effects of ST3GAL6 knockdown by shRNAs on cell invasion in 5637 UBC cells, detected by transwell invasion assay. Scale bars, 50  $\mu$ m. **C**, Quantifications of images from the transwell invasion assay. Data were presented as mean  $\pm$  SD of three independent experiments; \*\*\*,  $p < 0.001$ .

**Supplementary Table S1. List of oligonucleotide sequences and reagents**

|  | Direction              | Sequences (5'-3')  |             |                          |
|--|------------------------|--|-------------|--------------------------|
| <b>qRT-PCR primers</b>                                   |                        |  |             |                          |
| GATA3  | Forward                | GCCCTCATTAAAGCCCAAG  |             |                          |
|  | Reverse                | TTGTGGTGGTCTGACAGTTCCG   |             |                          |
| ST3GAL6  | Forward                | ATTGCCATCACATTGGCGTTT  |             |                          |
|  | Reverse                | GCAAAGGACTCTTGAGGTCAG  |             |                          |
| $\beta$ -actin   | Forward                | CATGTACGTTGCTATCCAGGC  |             |                          |
|  | Reverse                | CTCCTTAATGTCACGCACGA   |             |                          |
| <b>PCR primers for promoter activity assays</b>          |                        |  |             |                          |
| pGL3-ST3GAL6<br>-promoter (WT)                           | Forward                | TTTCTCTATCGATAGGTACCTTCAGAGGTCAGGGATAGAGC                      |             |                          |
|  | Reverse                | CTTAGATCGCAGATCTCGAGCATTCTCTTCTGGGGCAGAG                       |             |                          |
| pGL3-ST3GAL6<br>-promoter (mutant<br>GATA3 binding site) | Overlapping<br>Forward | AACAAATGAGCGGGGCTGCTCCATCTGCTTGGTCGGTTTTTC                     |             |                          |
|  | Overlapping<br>Reverse | GGAGCAGCCCCGCTCATTGTTCATCTATCCATATATCCATGT<br>ATCC             |             |                          |
| pCDH-3XFLAG<br>-GATA3                                    | Forward                | ACAAGCTTGC GGCCGCGATGGAGGTGACGGCGGA                            |             |                          |
|  | Reverse                | TGCCCTCGATATCGAATTCACCCATGGCGGTGACC                            |             |                          |
| <b>siRNA and shRNA sequences (Sense 5'-3')</b>           |                        |  |             |                          |
| siGATA3  | siGATA3-1              | CUCAUUAAGCCCAAGCGAATT  |             |                          |
|  | siGATA3-2              | GGGCUCUACUACAAGCUUCTT  |             |                          |
| siST3GAL6  | siST3GAL6-1            | CAGCCUUAACCUGAUUUATT   |             |                          |
|  | siST3GAL6-2            | CGACUCCUAUGAUGUAAUATT  |             |                          |
| siRNA control  | siNC                   | UUCUCCGAACGUGUCACGUTT  |             |                          |
| shST3GAL6  | shST3GAL6-1            | CCGGTCCTCTATTATGTACTGCATTCTCGAGAATGCAGTAC<br>ATAATAGAGATTTTTG  |             |                          |
|  | shST3GAL6-2            | CCGGGATGAGAACATCAGCGGAATACTCGAGTATTCCGCT<br>GATGTTCTCATCTTTTTG |             |                          |
| shRNA control  | shCTL                  | CCGGGCGCGATAGCGCTAATAATTTCTCGAGAAATTATTAG<br>CGCTATCGCGCTTTTT  |             |                          |
| Antibody   | Company                | Cat #  | RRID        | Application (Dilution)   |
| GATA3  | BD                     | #558686  | AB_2108590  | WB (1: 500)              |
| GAPDH  | Santa Cruz             | sc-47724   | AB_627678   | WB (1: 1,000)            |
| MAL II   | Vector Labs            | B-1265   | AB_2336569  | WB (1:1,000); IF (1:100) |
| ST3GAL6  | Abcam                  | Ab106527   | AB_10865191 | WB (1:500); IHC (1:100)  |
| EGF Receptor   | CST                    | #4267  | AB_2246311  | WB (1:1,000)             |
| phospho-EGF<br>Receptor (Tyr1068)                        | CST                    | #3777  | AB_2096270  | WB (1:1,000)             |
| AKT (pan)  | CST                    | #4691  | AB_915783   | WB (1:1,000)             |
| phospho-AKT<br>(Ser473)                                  | CST                    | #4060  | AB_2315049  | WB (1:1,000)             |
| p44/42 MAPK<br>(ERK1/2)                                  | CST                    | #4695  | AB_390779   | WB (1:1,000)             |

| <b>Cont'd</b>   |                                |                   |            |                   |
|---|--------------------------------|-------------------|------------|-------------------|
| phospho-p44/42<br>MAPK (ERK1/2)<br>(Thr202/Tyr204)          | CST                            | #4370             | AB_2315112 | WB (1:1,000)      |
| STAT3   | CST                            | #9139             | AB_331757  | WB (1:1,000)      |
| phospho-STAT3<br>(Tyr705)                                   | CST                            | #9145             | AB_2491009 | WB (1:1,000)      |
| HRP-conjugated<br>streptavidin                              | Vector Labs                    | SA-5704           | N.A.       | WB (ready to use) |
| FITC-Avidin   | BioLegend                      | #405101           | N.A.       | IF (1:1,000)      |
| Goat $\alpha$ -rabbit HRP<br>2 <sup>nd</sup> antibody       | Jackson<br>Immuno-<br>Research | #846738           | N.A.       | WB (1:5,000)      |
| Goat $\alpha$ -mouse HRP<br>2 <sup>nd</sup> antibody        |                                | #847273           | N.A.       | WB (1:5,000)      |
| <b>Kit and Reagents</b>                                     |                                | <b>Company</b>    |            | <b>Cat #</b>      |
| AceQ Universal SYBR qPCR Master<br>Mix                      |                                | Vazyme            |            | Q511-02           |
| Cell Counting Kit-8 (CCK-8)                                 |                                | Vazyme            |            | A311-01           |
| CloneExpress II One Step Cloning Kit                        |                                | Vazyme            |            | C112              |
| Corning BioCoat 24-Multiwell Tumor<br>Cell Invasion Systems |                                | Corning           |            | #354165           |
| DAB visualization kit                                       |                                | Maixin_Bio        |            | DAB-0031          |
| Dual Luciferase Reporter Gene Kit                           |                                | Beyotime          |            | RG027             |
| EDTA-free Protease Inhibitor Cocktail                       |                                | Roche             |            | #4693132001       |
| Fast Silver Stain Kit                                       |                                | Beyotime          |            | P0017S            |
| Hoechst   |                                | Thermo Scientific |            | #33342            |
| Lipofectamine 3000  |                                | Invitrogen        |            | L3000008          |
| Lipofectamine RNAiMAX reagent                               |                                | Invitrogen        |            | #13778            |
| Mut Express II Fast Mutagenesis Kit<br>V2                   |                                | Vazyme            |            | C214-02           |
| QuantiTect Reverse Transcription Kit                        |                                | QIAGEN            |            | #205311           |
| SuperSignal West Pico PLUS<br>Chemiluminescent Substrate    |                                | Thermo Scientific |            | #34579            |
| Trizol reagent  |                                | Invitrogen        |            | #15596026         |

**Supplementary Table S2. The clinicopathological associations among 4 clusters based on glycogene-type in TCGA-BLCA provisional dataset**

|                    | Total<br>(n=408) | Cluster<br>A1<br>(n=75) | Cluster<br>A2<br>(n=146) | Cluster<br>B1<br>(n=154) | Cluster<br>B2<br>(n=33) | Cluster B1<br>vs Others<br>p |
|--------------------|------------------|-------------------------|--------------------------|--------------------------|-------------------------|------------------------------|
| <b>Age (years)</b> |                  |                         |                          |                          |                         | <b>0.1361</b>                |
| <b>Mean±SD</b>     | 68.07<br>± 10.61 | 68.65<br>± 8.88         | 67.93<br>± 10.67         | 67.08<br>± 11.54         | 72.15<br>± 8.57         |                              |
| <b>Gender</b>      |                  |                         |                          |                          |                         | <b>0.1039</b>                |
| <b>Male</b>        | 301              | 59                      | 98                       | 121                      | 23                      |                              |
| <b>Female</b>      | 107              | 16                      | 48                       | 33                       | 10                      |                              |
| <b>Grade</b>       |                  |                         |                          |                          |                         | <b>&lt; 0.0001</b>           |
| <b>Low</b>         | 21               | 0                       | 1                        | 20                       | 0                       |                              |
| <b>High</b>        | 384              | 74                      | 145                      | 132                      | 33                      |                              |
| <b>NA</b>          | 3                | 1                       | 0                        | 2                        | 0                       |                              |
| <b>Stage</b>       |                  |                         |                          |                          |                         | <b>&lt; 0.0001</b>           |
| <b>I</b>           | 2                | 1                       | 0                        | 1                        | 0                       |                              |
| <b>II</b>          | 130              | 17                      | 32                       | 70                       | 11                      |                              |
| <b>III</b>         | 140              | 27                      | 62                       | 43                       | 8                       |                              |
| <b>IV</b>          | 134              | 30                      | 52                       | 38                       | 14                      |                              |
| <b>NA</b>          | 2                | 0                       | 0                        | 2                        | 0                       |                              |

Numbers in bold indicate p value with statistical difference.

**Supplementary Table S3. The 14 glycogenes differentiated the subcluster B1 from other three subclusters in TCGA provisional dataset**

| Gene Symbol       | GENE ID | Description  | Cluster B1 vs Others p |
|-------------------|---------|--|------------------------|
| <b>B4GALNT1</b>   | 2583    | beta-1,4-N-acetyl-galactosaminyl-transferase 1             | 2.68E-28               |
| <b>B4GALNT2</b>   | 124872  | beta-1,4-N-acetyl-galactosaminyl-transferase 2             | 9.76E-07               |
| <b>CHSY3</b>      | 337876  | chondroitin sulfate synthase 3                             | 1.66E-34               |
| <b>FUT7</b>       | 2529    | fucosyltransferase 7                                       | 1.25E-18               |
| <b>GALNT17</b>    | 64409   | Polypeptide N-acetylgalactosaminyltransferase 17           | 4.96E-19               |
| <b>GGTA1P</b>     | 2681    | glycoprotein, alpha-galactosyltransferase 1 pseudogene     | 4.73E-23               |
| <b>GLT1D1</b>     | 144423  | glycosyltransferase 1 domain containing 1                  | 9.31E-30               |
| <b>GLT8D2</b>     | 83468   | glycosyltransferase 8 domain containing 2                  | 6.30E-33               |
| <b>GXYLT2</b>     | 727936  | glucoside xylosyltransferase 2                             | 1.65E-43               |
| <b>ST3GAL6</b>    | 10402   | ST3 beta-galactoside alpha-2,3 -sialyltransferase 6        | 2.38E-44               |
| <b>ST6GALNAC5</b> | 81849   | ST6 N-acetylgalactosaminide alpha-2,6 -sialyltransferase 5 | 2.15E-17               |
| <b>UGT2B15</b>    | 7366    | UDP glucuronosyltransferase family 2 member B15            | 4.01E-12               |
| <b>UGT2B28</b>    | 54490   | UDP glucuronosyltransferase family 2 member B28            | 1.39E-14               |
| <b>UGT2B4</b>     | 7363    | UDP glucuronosyltransferase family 2 member B4             | 6.45E-03               |

**Supplementary Table S4. Univariate and multivariate analyses of 14-glycogene signature for patients' outcomes in TCGA provisional dataset**

| Variables                            | Univariate analysis |                   | Multivariate analysis |                   |
|--------------------------------------|---------------------|-------------------|-----------------------|-------------------|
|                                      | HR (95% CI)         | p                 | HR (95% CI)           | p                 |
| <b>Overall Survival (OS)</b>         |                     |                   |                       |                   |
| Cluster<br>(Cluster I vs Cluster II) | 1.982(1.380-2.849)  | <b>0.0002</b>     | 1.663(1.145-2.416)    | <b>0.0076</b>     |
| Age (Continuous)                     | 1.032(1.017-1.048)  | <b>&lt;0.0001</b> | 1.031(1.015-1.047)    | <b>&lt;0.0001</b> |
| Gender (Female vs Male)              | 0.882(0.639-1.217)  | 0.445             | 0.913(0.659-1.263)    | 0.581             |
| Grade (Low vs High)                  | 2.958(0.732-11.951) | 0.128             | 1.281(0.307-5.343)    | 0.734             |
| Stage (Stage I-III vs IV)            | 2.282(1.703-3.057)  | <b>&lt;0.0001</b> | 2.057(1.528-2.769)    | <b>&lt;0.0001</b> |
| <b>Disease free survival (DFS)</b>   |                     |                   |                       |                   |
| Cluster<br>(Cluster I vs Cluster II) | 1.718(1.183-2.494)  | <b>0.004</b>      | 1.385(0.944-2.030)    | 0.096             |
| Age (Continuous)                     | 1.014(0.998-1.030)  | 0.084             | 1.010(0.994-1.027)    | 0.224             |
| Gender (Female vs Male)              | 0.905(0.630-1.300)  | 0.589             | 0.868(0.598-1.259)    | 0.455             |
| Grade (Low vs High)                  | 3.076(0.759-12.47)  | 0.116             | 1.802(0.428-7.585)    | 0.422             |
| Stage (Stage I-III vs IV)            | 2.444(1.760-3.395)  | <b>&lt;0.0001</b> | 2.282(1.624-3.205)    | <b>&lt;0.0001</b> |

Abbreviations: HR, Hazard ratio; CI, Confidence interval.

Numbers in bold indicate p value with statistical difference.

**Supplementary Table S5. Univariate and multivariate analyses of 14-glycogene signature for overall survival in GSE32894 dataset**

| Variables                 | Univariate analysis    |                   | Multivariate analysis  |                   |
|---------------------------|------------------------|-------------------|------------------------|-------------------|
|                           | HR (95% CI)            | p                 | HR (95% CI)            | p                 |
| Cluster (Cluster I vs II) | 2.500(1.133-5.513)     | <b>0.023</b>      | 0.776(0.295-2.041)     | 0.608             |
| Age (Continuous)          | 0.975(0.945-1.007)     | 0.125             | 1.003(0.942-1.068)     | 0.921             |
| Gender (Female vs Male)   | 1.538(0.577-4.100)     | 0.389             | 1.387(0.446-4.310)     | 0.572             |
| Grade (G1/G2 vs G3/G4)    | 10.173(3.0333-4.126)   | <b>&lt;0.0001</b> | 0.428(0.076-2.423)     | 0.337             |
| Stage (Ta/T1 vs T2-4)     | 45.087(10.617-191.473) | <b>&lt;0.0001</b> | 92.056(11.082-764.678) | <b>&lt;0.0001</b> |

Abbreviations: HR, Hazard ratio; CI, Confidence interval.

Numbers in bold indicate p value with statistical difference.

**Supplementary Table S6. Univariate and multivariate analyses of 14-glycogene signature for patients' outcomes in GSE13507 dataset**

| Variables                             | Univariate analysis  |                   | Multivariate analysis |                   |
|---------------------------------------|----------------------|-------------------|-----------------------|-------------------|
|                                       | HR (95% CI)          | p                 | HR (95% CI)           | p                 |
| <b>Overall survival (OS)</b>          |                      |                   |                       |                   |
| Cluster (Cluster I vs II)             | 1.585(0.971-2.587)   | 0.065             | 1.525(0.895-2.598)    | 0.12              |
| Age (Continuous)                      | 1.070(1.044-1.096)   | <b>&lt;0.0001</b> | 1.077(1.048-1.106)    | <b>&lt;0.0001</b> |
| Gender (Female vs Male)               | 0.641(0.361-1.138)   | 0.129             | 0.827(0.449-1.520)    | 0.54              |
| Grade (Low vs High)                   | 2.740(1.694-4.433)   | <b>&lt;0.0001</b> | 1.171(0.650-2.107)    | 0.599             |
| Stage (Ta/T1 vs T2-4)                 | 2.730(1.688-4.415)   | <b>&lt;0.0001</b> | 2.491(1.392-4.455)    | <b>0.002</b>      |
| <b>Cancer-specific survival (CSS)</b> |                      |                   |                       |                   |
| Cluster (Cluster I vs II)             | 3.447(1.699-6.993)   | <b>0.001</b>      | 2.547(1.107-5.861)    | <b>0.028</b>      |
| Age (Continuous)                      | 1.051(1.016-1.088)   | <b>0.004</b>      | 1.074(1.027-1.123)    | <b>0.002</b>      |
| Gender (Female vs Male)               | 0.477(0.22-1.032)    | 0.06              | 0.515(0.230-1.152)    | 0.106             |
| Grade (Low vs High)                   | 5.985(2.757-12.992)  | <b>&lt;0.0001</b> | 1.114(0.472-2.629)    | 0.806             |
| Stage (Ta/T1 vs T2-4)                 | 17.897(6.251-51.244) | <b>&lt;0.0001</b> | 14.556(4.57-46.361)   | <b>&lt;0.0001</b> |

Abbreviations: HR, Hazard ratio; CI, Confidence interval.

Numbers in bold indicate p value with statistical difference.

**Supplementary Table S7. The association between the protein expression level of ST3GAL6 and clinicopathological parameters in 52 UBC samples from our IHC cohort**

| Characteristics    | n  | Expression of ST3GAL6 |            | p            |
|--------------------|----|-----------------------|------------|--------------|
|                    |    | high (n, %)           | low (n, %) |              |
| <b>Gender</b>      |    |                       |            | <b>0.761</b> |
| <b>Male</b>        | 40 | 22(55.0%)             | 18(45.0%)  |              |
| <b>Female</b>      | 12 | 6(50.0%)              | 6(50.0%)   |              |
| <b>Age</b>         |    |                       |            | <b>0.634</b> |
| <b>≥ 65</b>        | 30 | 17(56.7%)             | 13(43.3%)  |              |
| <b>&lt; 65</b>     | 22 | 11(50.0%)             | 11(50.0%)  |              |
| <b>Tumor grade</b> |    |                       |            | <b>0.029</b> |
| <b>Low</b>         | 24 | 9(37.5%)              | 15(62.5%)  |              |
| <b>High</b>        | 28 | 19(67.9%)             | 9(32.1%)   |              |
| <b>T stage</b>     |    |                       |            | <b>0.238</b> |
| <b>Ta-T1</b>       | 37 | 18(48.6%)             | 19(51.4%)  |              |
| <b>T2-T4</b>       | 15 | 10(66.7%)             | 5(33.3%)   |              |
| <b>N stage</b>     |    |                       |            | <b>0.856</b> |
| <b>N0</b>          | 47 | 25(53.2%)             | 22(46.8%)  |              |
| <b>≥ N1</b>        | 5  | 3(60.0%)              | 2(40.0%)   |              |

Numbers in bold indicate p value with statistical difference.

**Supplementary Table S8. Univariate analysis of ST3GAL6 expression with overall survival in 52 UBC samples from our IHC cohort**

| Variables                        | HR (95% CI)          | p            |
|----------------------------------|----------------------|--------------|
| Age (<60 vs ≥60)                 | 2.325 (0.852-6.344)  | 0.099        |
| Gender (Female vs Male)          | 1.327 (0.488-3.605)  | 0.579        |
| T stage (Ta-T1 vs T2-T4)         | 2.435 (0.947-6.259)  | 0.065        |
| Grade (Low vs High)              | 4.505 (1.650-12.299) | <b>0.003</b> |
| ST3GAL6 expression (Low vs High) | 2.951 (1.149-7.579)  | <b>0.025</b> |

Abbreviations: HR, Hazard ratio; CI, Confidence interval.

Numbers in bold indicate p value with statistical difference.