

基于血糖相关基因不同亚型乳腺癌的预后模型

吴南平

河北工业大学理学院, 天津

收稿日期: 2025年3月18日; 录用日期: 2025年4月11日; 发布日期: 2025年4月21日

摘要

乳腺癌的发病率在全球范围内呈上升趋势, 也是女性中最常见的恶性肿瘤之一。本研究从TCGA数据库获取1028例乳腺癌患者的转录组数据以及相关的临床数据。在MSigDB数据库使用关键词筛选845个血糖相关的基因。对TCGA数据库中的肿瘤样本和正常样本进行差异表达分析获取7549个差异表达基因, 同时将血糖相关基因集和差异表达基因取交集获得342个交集基因。使用加权基因共表达网络(WGCNA)筛选出与乳腺癌预后相关性高的黄色模块, 并基于黄色模块中的基因使用一致性聚类算法对TCGA-BRCA乳腺癌样本进行最优化分组。通过单因素Cox回归分析和LASSO回归分析筛出10个与预后显著相关的基因(CLIC6、ELOVL2、KLHDC7B、KRT80、MMP12、NEK10、SEC14L2、SHCBP1、SUSD3和TAT), 并基于这10基因构建了乳腺癌的风险评分模型。本研究通过一致性聚类分组后构建的乳腺癌风险评分模型也能够较好地预测患者的预后情况, 为乳腺癌的预后评估及开发可能的治疗靶点提供一定依据。

关键词

乳腺癌, 血糖, 加权基因共表达网络分析, 一致性聚类, 风险评分

Prognostic Model for Breast Cancer Based on Blood Glucose-Related Genes in Different Subtypes

Nanping Wu

College of Science, Hebei University of Technology, Tianjin

Received: Mar. 18th, 2025; accepted: Apr. 11th, 2025; published: Apr. 21st, 2025

Abstract

The incidence of breast cancer is on the rise globally and remains one of the most common malignant tumors among women. In this study, transcriptomic data and relevant clinical data from 1028

breast cancer patients were obtained from the TCGA database. A total of 845 blood glucose-related genes were screened using keywords from the MSigDB database. Differential expression analysis between tumor and normal samples in the TCGA database yielded 7549 differentially expressed genes. The intersection of blood glucose-related genes and differentially expressed genes resulted in 342 overlapping genes. A weighted gene co-expression network was employed to identify the yellow module, which exhibited a high correlation with breast cancer prognosis. Based on the genes within the yellow module, an optimal grouping of TCGA-BRCA breast cancer samples was achieved using a consensus clustering algorithm. Univariate Cox regression analysis and LASSO regression analysis identified 10 genes significantly associated with prognosis (CLIC6, ELOVL2, KLHDC7B, KRT80, MMP12, NEK10, SEC14L2, SHCBP1, SUSD3, and TAT). A risk scoring model for breast cancer was subsequently constructed based on these 10 genes. The risk scoring model developed after consensus clustering grouping demonstrated a robust ability to predict patient prognosis, offering valuable insights for prognostic assessment and the potential development of therapeutic targets in breast cancer.

Keywords

Breast Cancer, Blood Glucose, Weighted Gene Co-Expression Network Analysis, Consensus Clustering, Risk Score

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

乳腺癌严重威胁女性的生命健康，是女性病死率第二高的癌症[1]。同时随着平均体重的增加和生育率的下降，乳腺癌的发病率每年逐渐增加 0.5% [2] [3]。目前已知年龄、性别、肥胖、雌激素水平和家族史是乳腺癌的危险因素[4]。最近的研究表明，高血糖是乳腺癌发展的一个重要危险因素[5]。并且同时糖尿病(主要为 2 型糖尿病)也是全球最常见的慢性病之一，包括儿童、女性、青年和成人都受其影响[6]。越来越多的证据表明，葡萄糖代谢受损可能是各种癌症发展的独立危险因素，并可能对抗癌治疗的结果产生负面影响[7] [8]。流行病学研究报告称，在糖尿病的受试者中，乳腺癌的发病率会增加[9] [10]。

在各种癌症类型中，有多项研究表明了乳腺癌风险与糖尿病之间具有联系。荟萃分析报告显示：(1) 糖尿病女性乳腺癌的风险增加了 20%~28% [9] [11]，报告还指出，50 岁及以上的绝经后女性患糖尿病的乳腺癌风险增加了 20%~27% [11]；(2) 糖尿病与整体生存率差和无病生存率差有关[12]。患有乳腺癌的糖尿病患者死于其他各种原因的风险可能增加高达 50% [13]。有证据表明，16%的乳腺癌患者患有糖尿病，这与糖尿病女性发展乳腺癌的风险增加 10%~20%相关[14]。此外，糖尿病或糖耐量受损会恶化乳腺癌预后[15] [16]，并在一项为期 10 年的随访研究与乳腺癌相关死亡率呈正相关[17]。尽管有多项分析表明糖尿病与乳腺癌发病率和死亡率之间具有关联性，但这一关联背后的分子机制仍不明确。

乳腺癌的异质性和基因表达的广泛变异决定了疾病的进程、对治疗方法的反应、临床治疗结果以及患者的生存情况，并且基于乳腺癌的分子特征及其特征受体的有无情况决定了治疗干预的类型[18]。乳腺癌中的雌激素受体(ER)、孕激素受体(PR)和人类表皮生长因子受体 2 (HER2)中的一种或多种受体的存在或过表达会使这些类型的乳腺癌对激素治疗及靶向治疗敏感[19]。虽有许多队列研究强调了血糖和乳腺癌的联系，但这些研究仅限于某些受体或基因，尚未建立基于多种血糖相关基因的研究。因此，本文基

于血糖相关基因的将乳腺癌样本通过一致性聚类的方式分为两组，并利用这两组之间的关键基因构建预后风险评分模型。首先我们在基因表达综合数据库(GEO)数据库中将筛选的血糖相关基因集与乳腺癌中肿瘤和正常之间的差异表达基因(DEGs)取交集。使用一致性聚类方法将乳腺癌样本根据筛选的血糖相关基因分为两组，并对这两组进行生存差异分析、肿瘤微环境(TME)和通路富集的分析。为了进一步研究血糖相关基因在乳腺癌中的作用，基于这两个聚类分组建立了一个预后风险评分模型。最后使用该模型预测了 TCGA 测试组的生存情况并使用 GEO 数据库中的乳腺癌样本进行外部验证。统计分析结果表明，该基于血糖相关基因的多基因风险评分模型可以一定程度上评估乳腺癌的预后情况。

2. 材料与方法

2.1. 数据来源

本研究从癌症基因组图谱(TCGA)获取 1028 名乳腺癌患者的原始基因型数据，包括转录组数据以及相关的临床数据。临床数据主要包括性别、年龄、TNM 分期、分子分型、生存状态和生存时间等。从 GEO 数据库(<https://www.ncbi.nlm.nih.gov/geo/>)中获取了基因表达谱数据集 GSE20658。该数据集有 327 例乳腺癌患者的转录组数据和对应的临床数据。使用 R 语言加载“msigdb”包，选定参数 species = “Homo sapiens”，同时使用关键词“Insulin、Glycolysis、Gluconeogenesis、Blood Glucose、Glucagon、Glucose”检索到 845 个血糖相关的基因。血糖(Blood glucose)指的是血液中的葡萄糖浓度，葡萄糖(Glucose)是血糖的核心组成部分，胰高血糖素(Glucagon)和胰岛素(Insulin)是维持血糖稳态的核心激素，糖酵解(Glycolysis)是葡萄糖代谢的核心途径，糖异生(Gluconeogenesis)是生物体将非糖物质转化为葡萄糖的过程，特别是在空腹或者饥饿状态下这个过程对于维持血糖稳定非常重要，因此选择这六个关键词。

2.2. 数据分析方法

2.2.1. 血糖相关基因的筛选

为了筛选出能够评估乳腺癌预后的血糖相关基因集，并且检验一致性聚类的分组效果、组间的生存差异和组间免疫浸润分析的差异是否显著，我们采用一系列系统的筛选方式。使用关键词“Insulin、Glycolysis、Gluconeogenesis、Blood Glucose、Glucagon、Glucose”在分子特征数据库(MSigDB)中筛选出血糖相关的基因后，为了祛除冗余基因以及最大化筛选出与乳腺癌患者预后相关的血糖基因，对 MSigDB 中初筛的血糖基因集和 TCGA-BRCA 样本中的差异表达基因取交集。同时为了在取得的基因集中找到与不同分型乳腺癌预后起关键作用的基因，在 R 语言中使用“BiocManager”包进行加权基因共表达网络分析(WGCNA)。将具有相似表达模式的基因划分到同一模块，通过计算基因间的表达相关性，选择重要模块中血糖相关基因作为后续的分析。

2.2.2. 一致性聚类分组及组间的生存分析

使用加权基因共表达网络分析方法(WGCNA)从血糖相关的基因中筛选出 36 个与乳腺癌患者预后密切相关的基因。基于筛选的这 36 个血糖相关基因的表达水平我们将 TCGA-BRCA 样本聚成两组，使用 R 语言中的“ConsensusClusterPlus”包进行了一致性聚类分析。其中关键参数选择了“reps = 1000, pItem = 0.8, clusterAlg = km”，对样本的聚类进行了 1000 次迭代，这样能够保持聚类的一致性，并选择最优的分组方式。使用一致性聚类后我们将 TCGA-BRCA 样本分为了两组。接下来，使用 R 语言中的“survival”包和“survminer”包[20]并使用 Kaplan-Meier(KM)生存分析比较：两组之间的总体生存期(OS)是否存在显著差异。然后利用 R 语言中的“xCell”包对 64 种免疫细胞和基质细胞类型的基因表达数据进行细胞类型富集分析，从而比较两组之间免疫细胞的浸润评分是否存在显著差异。

2.2.3. 差异分析及功能注释

在 R 语言中使用“limma”包进行差异分析并筛选出两组之间的差异表达基因($FDR < 0.05$ 和 $|\log_2 \text{fold chang}| \geq 1$), 为了探索与这两组之间差异表达基因的可能活动并找到与之相关的基因功能和富集途径[21], 我们使用了 R 语言中的“clusterprofiler”包对差异表达基因进行了功能富集分析。

2.2.4. 乳腺癌预后风险评分的构建及验证

为了进一步探索这两组的差异表达基因对乳腺癌预后的价值, 将 TCGA-BRCA 样本按照 1:1 的比例分为训练组和验证组。同时迭代循环分组以找出对训练组和验证组生存差异都显著的分组, 迭代次数设置为 1000 次。使用 GEO 数据库中的 GSE20658 队列作为外部的验证组。然后使用 LASSO 回归降低单因素 Cox 分析中筛选结果的维度, 并且最终基于 LASSO 回归的评价结果, 筛选出 10 个基因作为构建预后风险评分模型的变量。风险评分的构建模型如下:

$$\text{score} = \sum_{n=1}^i (\text{Coef}_i * \text{geneExp})$$

其中 Coef_i 表示的是第 i 个最终利用 LASSO 回归筛选作为变量的相关基因的系数, geneExp 表示第 i 个对应基因的表达水平。并且根据基因表达水平的中位数作为高低风险的分值, 将所有队列样本分为低风险组($\text{score} < \text{中位数}$)和高风险组($\text{score} > \text{中位数}$)。使用 Kaplan-Meier (KM)生存分析比较两组的总体生存期(OS)。使用 R 语言中的“timeROC”包绘制关于时间受试者工作特征(ROC)曲线用于评估基于血糖相关基因的风险评分模型分类效果。生存分析表明基于风险评分的分组的总生存期具有显著差异。并且通过 GEO 数据库的 GSE20658 队列得到了验证。

3. 结果

3.1. 血糖相关基因的筛选

对来自 TCGA-BRCA 队列中的 984 个肿瘤样本和 44 个正常样本进行了差异基因分析($|\log_2 \text{FC}| > 0.5$, $p < 0.05$), 总共筛选出 7891 个差异表达基因(图 1(a))。然后在 MSigDB 中的人类数据库中收集到 845 个血糖相关基因。对差异表达基因和血糖相关基因取交集获得 342 个相交的基因(图 1(b))。随后使用加权基因共表达网络分析识别高度协同表达关系的基因模块, 并将这些模块与乳腺癌的表型特征相关联。对样本进行聚类分析, 用平均距离法计算样本之间的距离, 得到样本聚类树状图(图 1(c))。使用动态树切割算法识别模块, 对模块进行聚类, 找到相似模块并合并相似模块, 然后将模块与外部信息相关联得到模块表型相关图(图 1(d))。最后计算 MM 值(基因与模块之间的相关性)和 GS 值(基因相对性状的重要性), 获得模块间的基因显著性(图 1(e)), 最终筛选出黄色模块中的 36 个血糖相关基因。

3.2. 基于黄色模块基因的一致性聚类分组

为了研究血糖相关基因和乳腺癌之间的关系, 使用一致性聚类算法对 TCGA-BRCA 队列进行聚类分析。为了获得最佳的聚类方式, 取最大聚类数目为 10 并迭代 1000 次进行聚类。比较这 10 组的一致性累计分布函数(图 2(a))和 CDF 曲线面积相对变化(图 2(b)), 并比较这十组的矩阵热图最终将 TCGA-BRCA 队列聚成两组(图 2(c))。这两组乳腺癌样本在总生存期(OS)上存在显著差异($p = 0.0461$; 图 2(e))。同时将这两组在临床性状上的信息连接起来(图 2(d))。为了探究不同亚型乳腺癌与一致性聚类分组之间的关系, 我们使用了 R 语言中的“ggalluvial”包绘制二者之间的冲积图(图 3(a))。可以看出 Luminal A 型乳腺癌患者大部分流向 C1 组, 而 Luminal A 型的分子特征是增值基因表达水平低, 是一种预后较好, 复发风险相对较低的乳腺癌亚型。这也验证了在生存分析中 C1 型预后更好的结果。

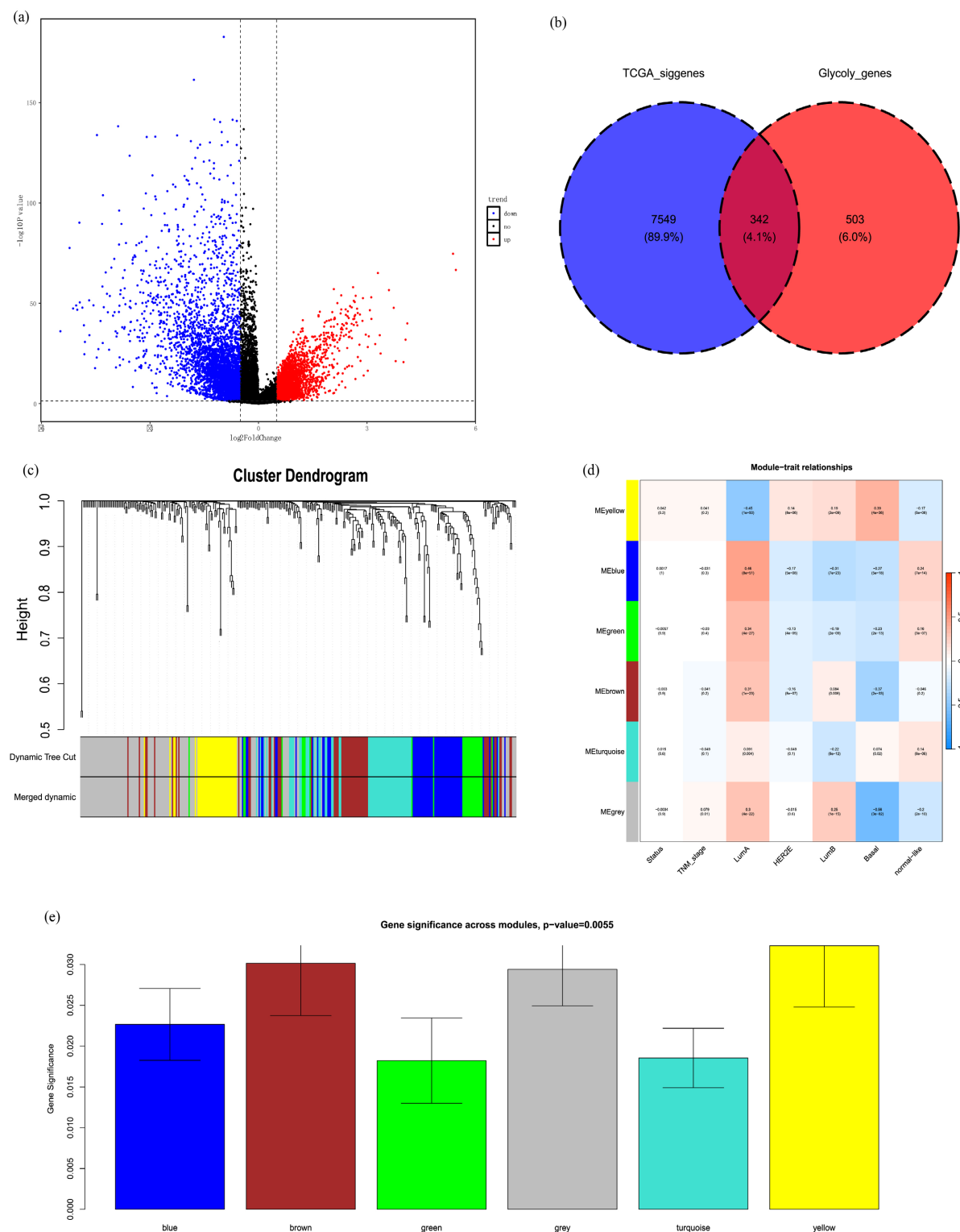


Figure 1. (a): Volcano plot of differentially expressed genes between tumor tissues and normal tissues in TCGA_BRCA; (b): Intersection of breast cancer DEGs and blood glucose-related genes; (c): Gene dendrogram and corresponding module colors; (d): Module-trait relationships; (e): Gene significance across modules

图 1. (a): TCGA_BRCA 肿瘤组织与正常组织差异表达火山图; (b): 乳腺癌 DEGs 和血糖相关基因取交集; (c): 基因树状图及对应模块颜色; (d): 模块 - 特征关系图; (e): 模块间基因显著性

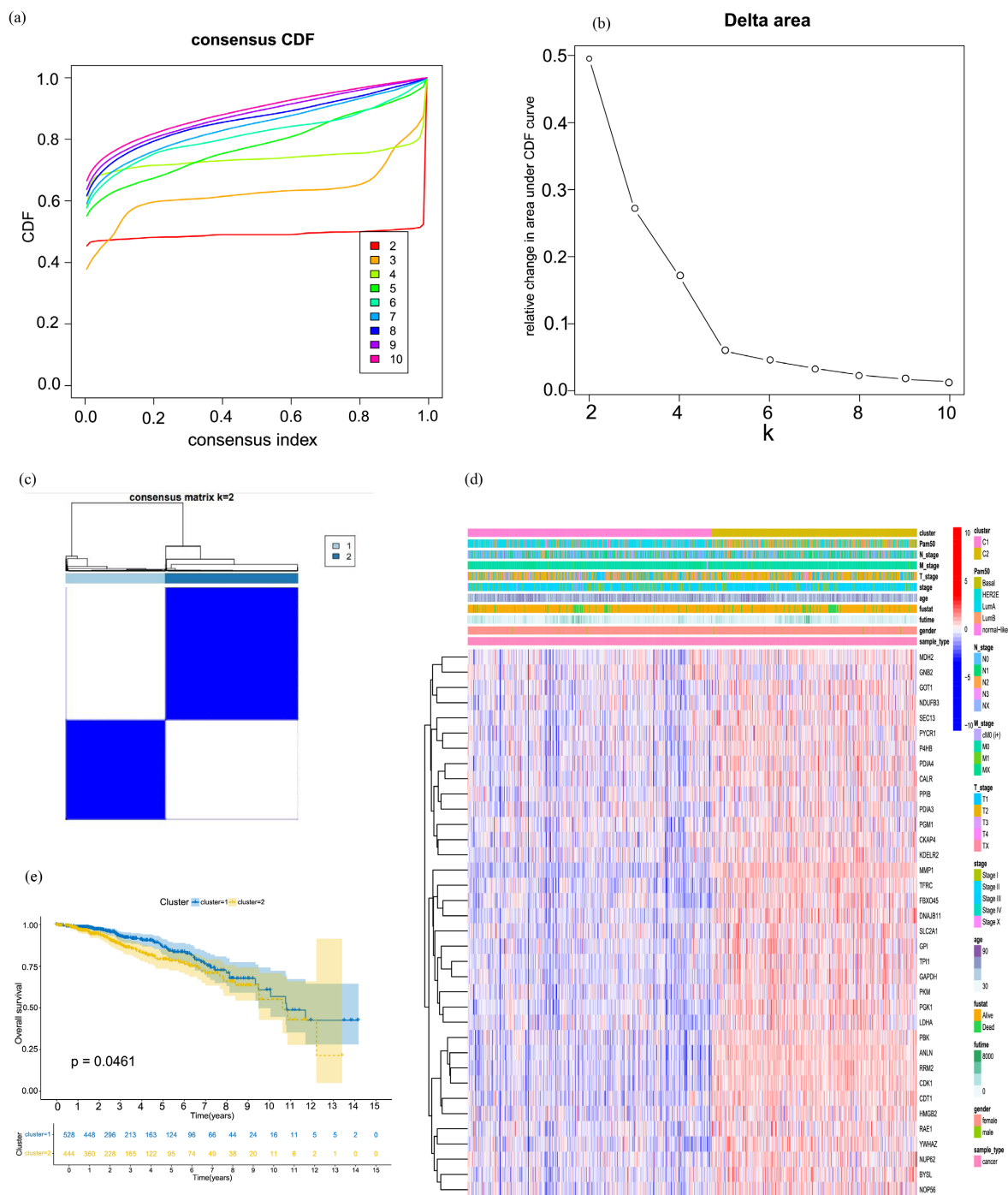


Figure 2. (a): Cumulative distribution function (CDF) plot for consensus clustering at different values of k ; (b): Relative change in the area under the CDF curve; (c): Consensus clustering heatmap at $k = 2$; (d): Heatmap of clinicopathological characteristics and gene expression between two clusters at $k = 2$; (e): Survival curve at $k = 2$

图 2. (a): k 取不同值时的一致性聚类累积分布函数(CDF)图; (b): CDF 曲线下面积相对变化图; (c): $k = 2$ 时一致性聚类热图; (d): $k = 2$ 是两个聚类样本间的临床病理特征和基因表达热图; (e): $k = 2$ 时的生存曲线

3.3. 两组乳腺癌样本的生存分析和免疫浸润分析

检验发现一致性聚类分组的乳腺癌样本的生存情况存在显著差异, 同时为了更好地了解两组之间的

免疫相关基因的差异, 在 R 中使用“xCell”包进行肿瘤微环境(TME)的研究。从而得到两个亚型之间免疫细胞的情况(图 3(c)), 可以清晰看到两组之间的免疫评分(Immune Score)、基质评分(Stroma Score)和微环境评分(Micro environment Score)的差异都是显著的, 在 B-cells、Th2 cells、CD4+ memory T-cells、Endothelial cells 等免疫细胞中的差异显著。在免疫激活相关基因的差异分析中, HSC、NKT 在 C1 组中高表达, aDC、Astrocytes、MEP、DC、iDC、pDC 在 C2 组中具有高表达。基于这个肿瘤微环境(TME)分析, 可以更加具体地了解二者间的差异。

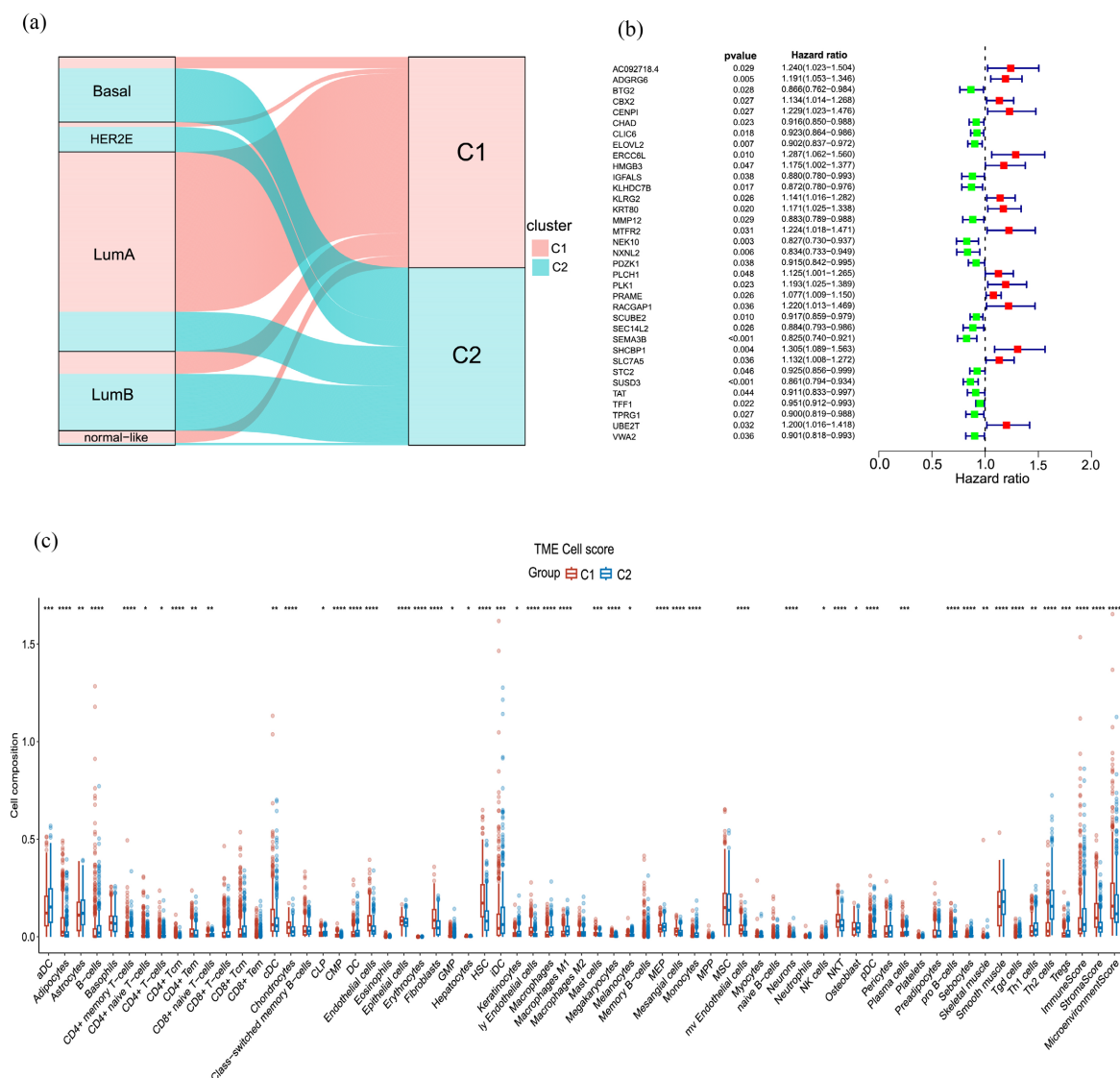


Figure 3. (a): Alluvial diagram illustrating the results of different breast cancer subtypes and consensus clustering grouping; (b): Results of variable selection through univariate Cox regression analysis; (c): Analysis of immune cell infiltration between the two subtypes using the xCell algorithm

图 3. (a): 不同亚型乳腺癌和共识聚类分组结果的冲积图; (b): 单因素 Cox 回归变量筛选结果; (c): 使用 xCell 算法对两个亚型之间免疫细胞的浸润情况分析

3.4. 差异表达基因的鉴定与预后风险评分模型的构建

对这两簇进行差异表达基因分析获得 368 个差异基因以进行下一步分析, 之后这 368 个差异表达基

因进行 GO 富集分析(图 4(a))和 KEGG 富集分析(图 4(b))。同时,使用单因素 Cox 分析筛选出与乳腺癌预后相关的 35 个基因(图 3(b)),再使用 LASSO 回归分析筛选出 10 个基因(图 4(c)、图 4(d))用于预后模型的构建。风险评分 = CLIC6 表达量 * (-0.018) + ELOVL2 表达量 * (-0.027) + KLHDC7B 表达量 * (-0.062) + KRT80 表达量 * 0.072 + MMP12 表达量 * (-0.046) + NEK10 表达量 * (-0.023) + SEC14L2 表达量 * (-0.007) + SHCBP1 表达量 * 0.130 + SUSL3 表达量 * 0.028 + TAT 表达量 * (-0.006), 从之前的分析表明, C1 组的生存情况比 C2 组的更好。

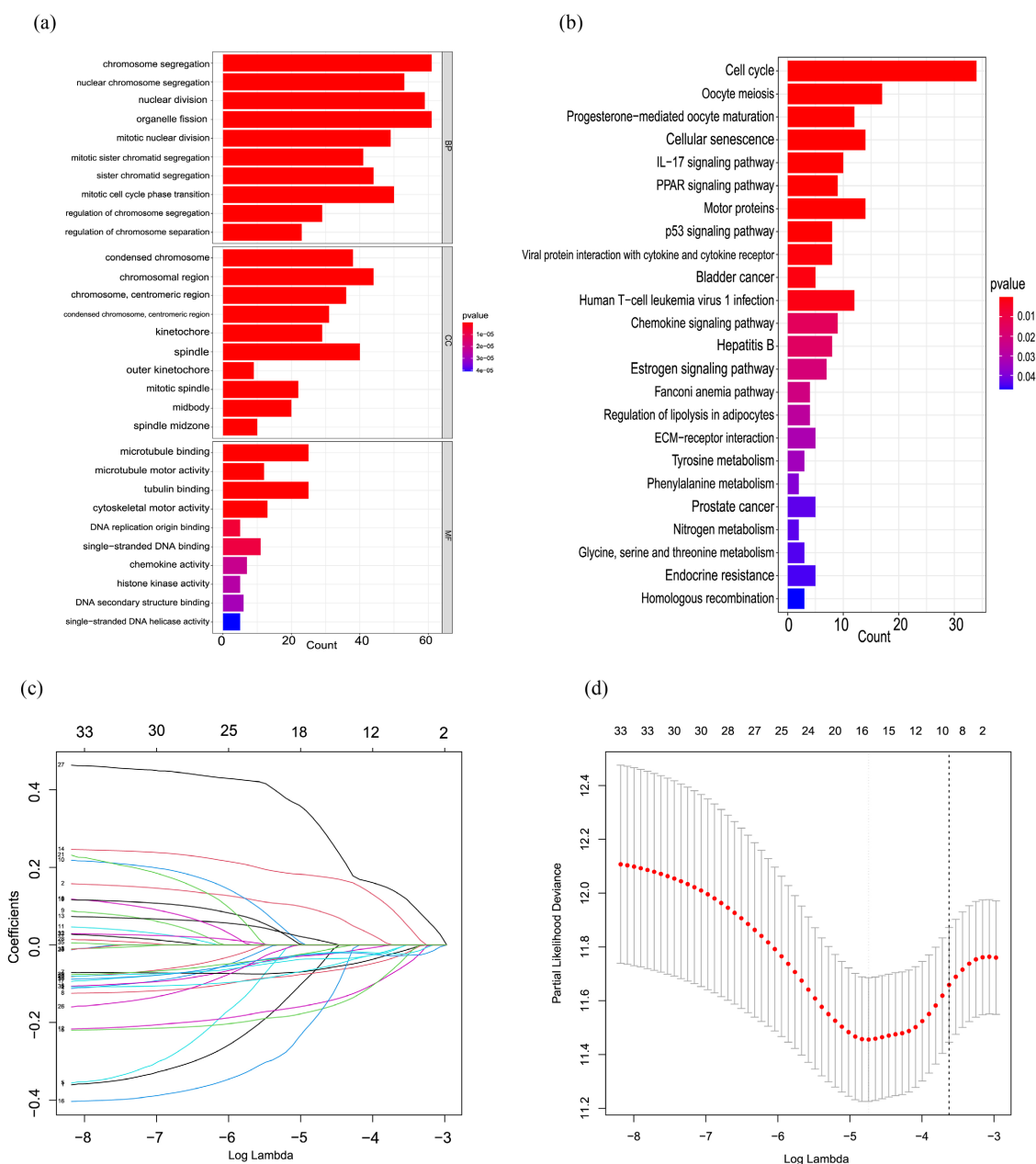


Figure 4. (a): Bar plot of GO gene enrichment analysis for 368 genes; (b): Bar plot of KEGG gene enrichment analysis for 368 genes; (c): Trajectory plot of LASSO regression independent variables; (d): Diagram of 20-fold cross-validation

图 4. (a): 368 个基因 GO 基因富集柱状图; (b): 368 个基因 KEGG 基因富集柱状图; (c): LASSO 回归自变量运动轨迹图; (d): 20 倍交叉验证图

3.5. 风险评分模型的验证

我们将 TCBA-BRCA 样本分为训练组和测试组分别用于评分模型的构建和验证。之后我们使用 KM 生存曲线确定基于该风险评分的分组是否对患者的总生存期具有影响。在高风险组和低风险组的生存曲线(图 5(a))可以看出, 高风险组样本的生存状况要低于低风险组, 并且对应的 p 值在统计学上也是具有显著差异的。同时我们使用 ROC 曲线用于评估该风险评分模型的性能(图 5(b)), 其 AUC 值大于 0.7 表明该分类模型具备了一定的区分能力。

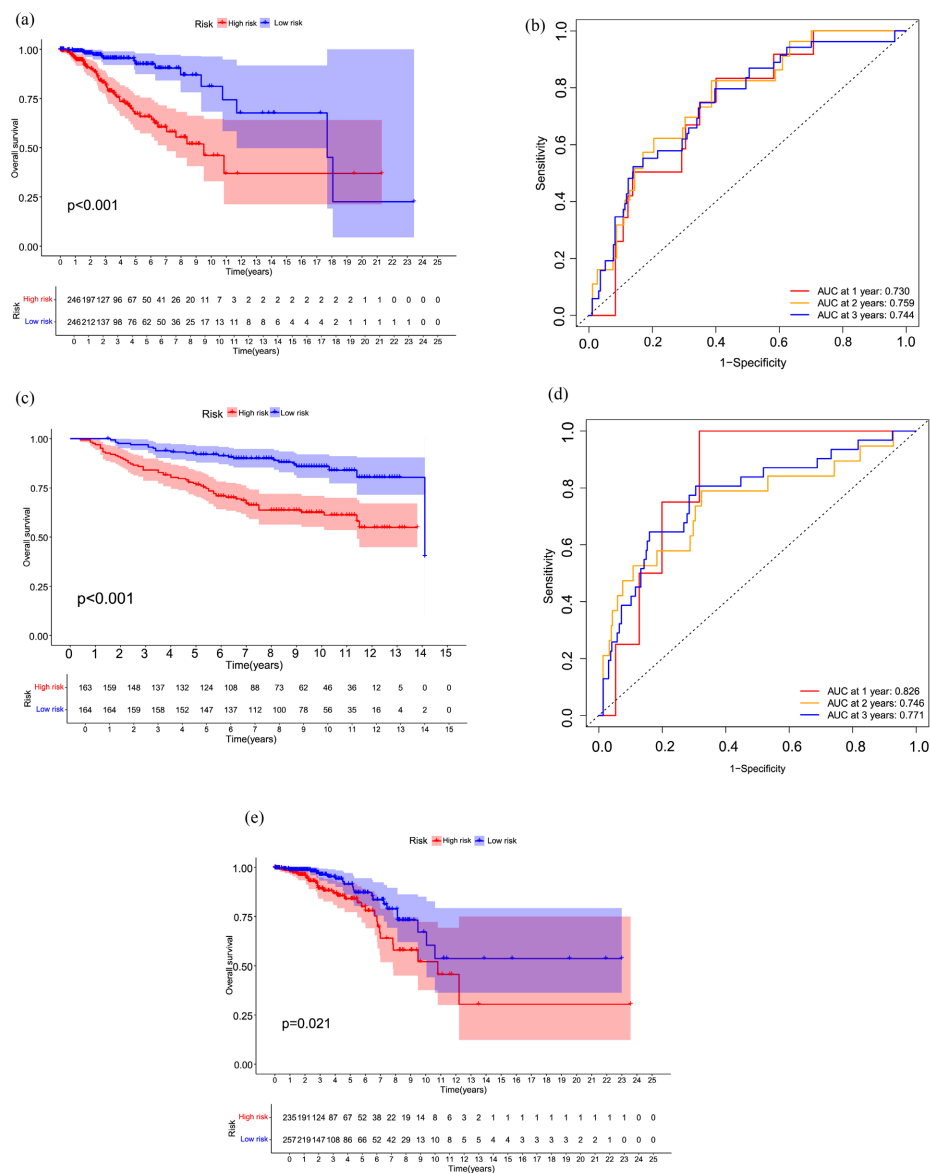


Figure 5. (a): Difference in overall survival time between the two groups in the TCGA-BRCA training set; (b): Time-dependent ROC curve in the TCGA-BRCA training set; (c): Difference in overall survival time between the two groups in the TCGA-BRCA validation set; (d): Time-dependent ROC curve in the TCGA-BRCA validation set; (e): Difference in overall survival time between the two groups in the GEO dataset

图 5. (a): TCGA-BRCA 训练集中两组的总生存期时间差; (b): TCGA 训练集中的时间依赖 ROC 曲线; (c): TCGA-BRCA 验证集中两组的总生存期时间差; (d): TCGA 验证集中的时间依赖 ROC 曲线; (e): GEO 数据集中两组的总生存期时间差

使用 TCGA-BRCA 验证组中的样本对该风险评分模型进行生存分析(图 5(c)), 使用验证中的样本绘制对应的 ROC 曲线(图 5(d)), 其 AUC 值也是大于 0.7 的。最后我们 GSE20658 数据集来验证, 其样本的高低风险组的生存差异显著(图 5(e))。

4. 讨论

血糖在癌症中不仅仅作为能量来源的作用, 更重要的是通过影响肿瘤代谢、炎症和信号通路调控癌症的发生和发展, 比如会促进胃癌、结肠直肠癌、肝细胞癌、胰腺癌和肺癌的发生和发展过程[22]。并且乳腺癌作为女性最常见的肿瘤, 血糖异常也会对患者的预后产生不利的影响[12][23]。尽管不同的乳腺癌亚型表现出不同的临床病理和分子特征, 并且对不同的治疗表现出异质性反应[24][25], 但越来越多的证据表明, 有氧糖酵解是癌细胞的一种重要代谢特征。因此, 针对葡萄糖代谢功能可能是一种有前景的乳腺癌治疗策略。无论氧气量是否充足, 肿瘤细胞主要利用糖酵解进行能量的生产并具有更高的糖酵解速率。这种有氧糖酵解表型被称为 Warburg 效应, 最早由德国生物化学家奥托·华尔堡在 1920 年首次描述[26]。有氧糖酵解也是被科学认可的一个癌细胞代谢的标志, 针对它可能为癌症治疗的策略提供潜在的药物靶点[27]。

乳腺癌中的一些血糖相关的基因也被广泛研究。比如高血糖诱导 PFKFB3 过表达, 并通过 RAS/MAPK 激活并促进乳腺癌的恶性发展[28]。以及高血糖诱导的 miR-467 驱动乳腺癌肿瘤炎症和生长[29]。同时高血糖可通过诱导氧化应激, 导致 DNA 损伤并促进癌细胞突变[30]。高血糖状态也会通过激活促炎信号通路(如 NF- κ B 和 IL-6)促进慢性炎症的发生, 而慢性炎症被认为是癌症的重要治病因素[31]。高血糖状态往往伴随高胰岛素状态, 并且胰岛素和 IGF-1 可通过 PI3K/AKT/mTOR 信号通路促进癌细胞增值和抑制凋亡[32]。尽管单个血糖相关基因已经被广泛的研究以及高血糖状态与乳腺癌预后的关系进行了统计研究, 本研究专注于将多个血糖相关的基因与乳腺癌预后相关联, 以研究乳腺癌患者基于血糖相关基因的亚型分析和预后模型的构建。

研究中尝试过多种聚类及筛选方式, 而使用本研究的一致性聚类分组方式两组在乳腺癌的预后总生存期、肿瘤微环境(TME)以及免疫浸润情况方面都存在显著差异。同时在差异基因的富集分析也富集在与肿瘤相关的免疫、代谢通路上。因此本研究认为筛选出的血糖相关基因与乳腺癌患者预后具有一定的相关性, 肿瘤的 TNM 分期是作为国际的公认标准, 我们也发现不同亚型乳腺癌与聚类分组存在联系。当然本研究还存在一定局限性, 缺少对应的体外实验和体内测试, 所以需要进一步的真实数据进行验证。

5. 结论

基于外部数据集对该乳腺癌风险评分模型的验证, 说明最终 WGCNA 中筛选的黄色模块中的血糖相关基因与乳腺癌的预后具有统计意义上的相关性。同时发现基于该血糖相关基因构建的风险评分模型进行的分组, 其预后情况差异的。虽然通过 ROC 曲线分析表明该模型的 AUC 值还达不到医学诊断等高要求的场景, 但是依旧可以表明该模型具有一定的区分能力。至于该十个基因是否可以作为乳腺癌的治疗靶点以及对伴高血糖的乳腺癌患者是否具有临床上的意义, 又或者基于该十个基因的分组为何会具有生存上的差异还需要更多的临床数据和实验去验证。

参考文献

- [1] Barzaman, K., Karami, J., Zarei, Z., Hosseinzadeh, A., Kazemi, M.H., Moradi-Kalbolandi, S., *et al.* (2020) Breast Cancer: Biology, Biomarkers, and Treatments. *International Immunopharmacology*, **84**, Article 106535. <https://doi.org/10.1016/j.intimp.2020.106535>
- [2] Pfeiffer, R.M., Webb-Vargas, Y., Wheeler, W. and Gail, M.H. (2018) Proportion of U.S. Trends in Breast Cancer

- Incidence Attributable to Long-Term Changes in Risk Factor Distributions. *Cancer Epidemiology, Biomarkers & Prevention*, **27**, 1214-1222. <https://doi.org/10.1158/1055-9965.epi-18-0098>
- [3] Siegel, R.L., Miller, K.D., Fuchs, H.E. and Jemal, A. (2021) Cancer Statistics, 2021. *CA: A Cancer Journal for Clinicians*, **71**, 7-33. <https://doi.org/10.3322/caac.21654>
 - [4] Mohamed, H.T., El-Shinawi, M., Nouh, M.A., Bashtar, A., Elsayed, E.T., Schneider, R.J., *et al.* (2014) Inflammatory Breast Cancer: High Incidence of Detection of Mixed Human Cytomegalovirus Genotypes Associated with Disease Pathogenesis. *Frontiers in Oncology*, **4**, Article 246. <https://doi.org/10.3389/fonc.2014.00246>
 - [5] Samuel, S.M., Varghese, E., Varghese, S. and Büsselberg, D. (2018) Challenges and Perspectives in the Treatment of Diabetes Associated Breast Cancer. *Cancer Treatment Reviews*, **70**, 98-111. <https://doi.org/10.1016/j.ctrv.2018.08.004>
 - [6] Hu, F.B. (2011) Globalization of Diabetes: The Role of Diet, Lifestyle, and Genes. *Diabetes Care*, **34**, 1249-1257. <https://doi.org/10.2337/dc11-0442>
 - [7] Shikata, K., Ninomiya, T. and Kiyohara, Y. (2012) Diabetes Mellitus and Cancer Risk: Review of the Epidemiological Evidence. *Cancer Science*, **104**, 9-14. <https://doi.org/10.1111/cas.12043>
 - [8] Vigneri, P., Frasca, F., Sciacca, L., Pandini, G. and Vigneri, R. (2009) Diabetes and Cancer. *Endocrine-Related Cancer*, **16**, 1103-1123. <https://doi.org/10.1677/erc-09-0087>
 - [9] Larsson, S.C., Mantzoros, C.S. and Wolk, A. (2007) Diabetes Mellitus and Risk of Breast Cancer: A Meta-Analysis. *International Journal of Cancer*, **121**, 856-862. <https://doi.org/10.1002/ijc.22717>
 - [10] Stattin, P., Björ, O., Ferrari, P., Lukanova, A., Lenner, P., Lindahl, B., *et al.* (2007) Prospective Study of Hyperglycemia and Cancer Risk. *Diabetes Care*, **30**, 561-567. <https://doi.org/10.2337/dc06-0922>
 - [11] Boyle, P., Boniol, M., Koechlin, A., Robertson, C., Valentini, F., Coppens, K., *et al.* (2012) Diabetes and Breast Cancer Risk: A Meta-Analysis. *British Journal of Cancer*, **107**, 1608-1617. <https://doi.org/10.1038/bjc.2012.414>
 - [12] Zhao, X. and Ren, G. (2016) Diabetes Mellitus and Prognosis in Women with Breast Cancer. A Systematic Review and Meta-Analysis *Medicine*, **95**, e5602. <https://doi.org/10.1097/md.0000000000005602>
 - [13] Peairs, K.S., Barone, B.B., Snyder, C.F., Yeh, H., Stein, K.B., Derr, R.L., *et al.* (2011) Diabetes Mellitus and Breast Cancer Outcomes: A Systematic Review and Meta-Analysis. *Journal of Clinical Oncology*, **29**, 40-46. <https://doi.org/10.1200/jco.2009.27.3011>
 - [14] Vona-Davis, L. and P. Rose, D. (2012) Type 2 Diabetes and Obesity Metabolic Interactions: Common Factors for Breast Cancer Risk and Novel Approaches to Prevention and Therapy. *Current Diabetes Reviews*, **8**, 116-130. <https://doi.org/10.2174/157339912799424519>
 - [15] Goodwin, P.J., Ennis, M., Pritchard, K.I., Trudeau, M.E., Koo, J., Taylor, S.K., *et al.* (2012) Insulin- and Obesity-Related Variables in Early-Stage Breast Cancer: Correlations and Time Course of Prognostic Associations. *Journal of Clinical Oncology*, **30**, 164-171. <https://doi.org/10.1200/jco.2011.36.2723>
 - [16] Schrauder, M.G., Fasching, P.A., Häberle, L., Lux, M.P., Rauh, C., Hein, A., *et al.* (2011) Diabetes and Prognosis in a Breast Cancer Cohort. *Journal of Cancer Research and Clinical Oncology*, **137**, 975-983. <https://doi.org/10.1007/s00432-010-0960-2>
 - [17] De Bruijn, K.M.J., Arends, L.R., Hansen, B.E., Leeflang, S., Ruiter, R. and van Eijck, C.H.J. (2013) Systematic Review and Meta-Analysis of the Association between Diabetes Mellitus and Incidence and Mortality in Breast and Colorectal Cancer. *British Journal of Surgery*, **100**, 1421-1429. <https://doi.org/10.1002/bjs.9229>
 - [18] Rivenbark, A.G., O'Connor, S.M. and Coleman, W.B. (2013) Molecular and Cellular Heterogeneity in Breast Cancer. Challenges for Personalized Medicine. *The American Journal of Pathology*, **183**, 1113-1124. <https://doi.org/10.1016/j.ajpath.2013.08.002>
 - [19] Mohamed, A., Krajewski, K., Cakar, B. and Ma, C.X. (2013) Targeted Therapy for Breast Cancer. *The American Journal of Pathology*, **183**, 1096-1112. <https://doi.org/10.1016/j.ajpath.2013.07.005>
 - [20] Rich, J.T., Neely, J.G., Paniello, R.C., Voelker, C.C.J., Nussenbaum, B. and Wang, E.W. (2010) A Practical Guide to Understanding Kaplan-Meier Curves. *Otolaryngology—Head and Neck Surgery*, **143**, 331-336. <https://doi.org/10.1016/j.otohns.2010.05.007>
 - [21] Yu, G., Wang, L., Han, Y. and He, Q. (2012) Clusterprofiler: An R Package for Comparing Biological Themes among Gene Clusters. *OMICS: A Journal of Integrative Biology*, **16**, 284-287. <https://doi.org/10.1089/omi.2011.0118>
 - [22] Li, W., Liu, H., Qian, W., Cheng, L., Yan, B., Han, L., *et al.* (2018) Hyperglycemia Aggravates Microenvironment Hypoxia and Promotes the Metastatic Ability of Pancreatic Cancer. *Computational and Structural Biotechnology Journal*, **16**, 479-487. <https://doi.org/10.1016/j.csbj.2018.10.006>
 - [23] García-Esquinas, E., Guinó, E., Castaño-Vinyals, G., Pérez-Gómez, B., Llorca, J., Altzibar, J.M., *et al.* (2016) Association of Diabetes and Diabetes Treatment with Incidence of Breast Cancer. *Acta Diabetologica*, **53**, 99-107. <https://doi.org/10.1007/s00592-015-0756-6>

-
- [24] Tang, Y., Wang, Y., Kiani, M.F. and Wang, B. (2016) Classification, Treatment Strategy, and Associated Drug Resistance in Breast Cancer. *Clinical Breast Cancer*, **16**, 335-343. <https://doi.org/10.1016/j.clbc.2016.05.012>
- [25] Wu, L. and Yang, X. (2018) Targeting the Hippo Pathway for Breast Cancer Therapy. *Cancers*, **10**, Article 422. <https://doi.org/10.3390/cancers10110422>
- [26] Ganapathy-Kanniappan, S. (2017) Taming Tumor Glycolysis and Potential Implications for Immunotherapy. *Frontiers in Oncology*, **7**, Article 36. <https://doi.org/10.3389/fonc.2017.00036>
- [27] Li, L., Liang, Y., Kang, L., Liu, Y., Gao, S., Chen, S., *et al.* (2018) Transcriptional Regulation of the Warburg Effect in Cancer by SIX1. *Cancer Cell*, **33**, 368-385.E7. <https://doi.org/10.1016/j.ccell.2018.01.010>
- [28] Cheng, X., Jia, X., Wang, C., Zhou, S., Chen, J., Chen, L., *et al.* (2023) Hyperglycemia Induces PFKFB3 Overexpression and Promotes Malignant Phenotype of Breast Cancer through RAS/MAPK Activation. *World Journal of Surgical Oncology*, **21**, Article No. 112. <https://doi.org/10.1186/s12957-023-02990-2>
- [29] Gajeton, J., Krukovets, I., Muppala, S., Verbovetskiy, D., Zhang, J. and Stenina-Adognravi, O. (2021) Hyperglycemia-Induced miR-467 Drives Tumor Inflammation and Growth in Breast Cancer. *Cancers*, **13**, Article 1346. <https://doi.org/10.3390/cancers13061346>
- [30] Dizdaroglu, M. (2012) Oxidatively Induced DNA Damage: Mechanisms, Repair and Disease. *Cancer Letters*, **327**, 26-47. <https://doi.org/10.1016/j.canlet.2012.01.016>
- [31] Hirano, T. (2021) IL-6 in Inflammation, Autoimmunity and Cancer. *International Immunology*, **33**, 127-148. <https://doi.org/10.1093/intimm/dxaa078>
- [32] Pollak, M. (2012) The Insulin and Insulin-Like Growth Factor Receptor Family in Neoplasia: An Update. *Nature Reviews Cancer*, **12**, 159-169. <https://doi.org/10.1038/nrc3215>