

生物信息学分析类风湿关节炎的关键生物标志物和免疫浸润

高扬¹, 田龙龙², 毕春花^{1*}

¹青岛大学附属医院感染科, 山东 青岛

²大连医科大学研究生院, 辽宁 大连

收稿日期: 2022年6月4日; 录用日期: 2022年6月28日; 发布日期: 2022年7月5日

摘要

目的: 挖掘类风湿关节炎的潜在基因标志物, 探究类风湿关节炎的免疫浸润情况, 为进一步阐明类风湿关节炎的发生发展提供方向, 并指导临床治疗。方法: 从基因表达综合数据库(GEO)下载GSE55235基因集, 获得RA和健康样本之间的差异表达基因(DEGs)。用R软件进行DEGs的基因本体论(GO)和京都基因与基因组百科全书(KEGG)富集分析。我们还进行基因集富集分析(GSEA)以进一步了解基因的功能, 并利用加权基因共表达网络分析(WGCNA)构建基因共表达网络, 并确定最重要的模块。并运用Cytoscape软件构建蛋白互作网络, 筛选出位于Top 5的关键基因。建立临床预测模型对得到的关键基因进行验证。最后, 使用R软件的“e1071”, “CIBERSORT”和“parallel”包来分析疾病的免疫细胞浸润模式。结果: 总共筛选了5个在RA的发生发展中最具重要意义差异基因(IGLJ3、IGHM、IGKC、IGLV1-44、GUSBP11)。生物功能分析确定了RA中的关键相关通路、基因模块和共表达网络。免疫浸润分析发现, CD8阳性T淋巴细胞和浆细胞浸润增加, CD4阳性的静息记忆T淋巴细胞和活性肥大细胞浸润减少可能与RA的发生有关。结论: IGLJ3、IGHM、IGKC、IGLV1-44、GUSBP11基因在RA的发生发展中具有重要作用, 这将有助于确定RA的新型诊断标志物和治疗靶点, 也为后续的发病机制研究提供可靠的依据和全新的视角。

关键词

类风湿关节炎, 生物信息分析, 免疫细胞浸润, 临床模型

Identification of Key Biomarkers and Immune Infiltration in Rheumatoid Arthritis by Integrated Bioinformatics Analysis

Yang Gao¹, Longlong Tian², Chunhua Bi^{1*}

*通讯作者 Email: qyfuchb@163.com

¹Department of Infectious Diseases, Affiliated Hospital of Qingdao University, Qingdao Shandong

²Graduate School of Dalian Medical University, Dalian Liaoning

Received: Jun. 4th, 2022; accepted: Jun. 28th, 2022; published: Jul. 5th, 2022

Abstract

Objective: Excavate the potential gene markers of rheumatoid arthritis and explore the immune infiltration of rheumatoid arthritis, so as to provide direction for further clarifying the occurrence and development of rheumatoid arthritis and guide clinical treatment. **Methods:** The GSE55235 gene set was downloaded from the Gene Expression Omnibus data base (GEO), and the differentially expressed genes (DEGs) between RA and healthy samples were obtained. The Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichments of DEGs were performed by R software. We performed Gene Set Enrichment Analysis (GSEA) to further understand the functions of the hub gene. Weighted gene co-expression network analysis (WGCNA) would be utilized to build a gene co-expression network, and the most significant module and hub genes were identified. The protein interaction network was constructed by using Cytoscape software to screen the key genes in the top 5. A clinical prediction model is established to verify the key genes. e1071, CIBERSORT, parallel package of R software are helpful to analyze the immune cell infiltration pattern of disease. **Results:** A total of 5 main differential genes (IGLJ3, IGHM, IGKC, IGLV1-44, GUSBP11) were screened. Biological function analysis identified the key related pathways, gene modules and coexpression networks in RA. Immune infiltration analysis showed that plasma cell and CD8 positive T cells infiltration increased and CD4 positive T cell and Mast cells activated infiltration decreased, which may be related to the occurrence of RA. **Conclusion:** IGLJ3, IGHM, IGKC, IGLV1-44, GUSBP11 genes play an important role in the occurrence and development of RA, which will help to determine the new diagnostic markers and therapeutic targets of RA, and also provide a reliable basis and a new perspective for the follow-up pathogenesis research.

Keywords

Rheumatoid Arthritis, Bioinformatics Analysis, Immune Infiltration, Clinical Model

Copyright © 2022 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. 背景

类风湿关节炎(Rheumatoid arthritis, RA)是一种以慢性炎症、持续性滑膜炎和滑膜关节破坏为特征的系统性自身免疫性疾病[1]。它是最常见的炎性关节炎,具有高的发病率和致残率[2]。RA在女性中的患病率高于男性。RA的病因和发病机制复杂,涉及滑膜细胞增殖和纤维化、血管翳形成,以及软骨和骨侵蚀[3]。除了遗传和家族性危险因素外,还存在多种环境、饮食和生活方式因素与RA发病相关[4]。RA是不同发病机制的结果,并表现出不同的临床表现和细胞基础和分子基础。炎症细胞因子,如白细胞介素(IL)-6、IL-1 β 和肿瘤坏死因子(TNF)- α ,已被证实在RA的发病机制中发挥作用。然而,类风湿关节炎的具体病理生理机制尚未被彻底研究。因此,探索RA发生、发展的分子特征和机制,为RA的有效预防、诊断和治疗提供新的策略具有重要意义[5]。

基于高通量平台的基因芯片技术已被广泛用于基因组水平上探索和识别有希望的疾病诊断和预后生物标志物。大量研究表明, RA 发生的病理生理过程与基因突变和异常表达密切相关, 包括 HLA II 类基因、PTPN22、CTLA4、PADI4 等[6] [7] [8] [9]。在非 HLA-DRB1 基因座中, PTPN22 基因是迄今为止与 RA 易感性最显著相关的基因, 与 HLA-DRB1 等位基因一起, 被认为是 RA 50% 的遗传组成部分[10]。最近的一项研究表明, 病毒和细菌可以通过与 TLR 结合增加对 RA 的易感性并加剧 RA。促进破骨细胞成熟由 toll 样受体(TLR)7/8 信号介导, 该信号由 mir-574-5p 结合激活[11]。一项研究表明, RA 患者的外周血单核细胞和血清外泌体中 lncRNAs Hotair 和 NEAT1 的高表达水平, 有助于活性巨噬细胞的迁移[12]。因此, 为了对 RA 的诊断和治疗做出新突破, 有必要探索 RA 发生和进展中包含的准确分子靶点。

我们分析了美国国立生物技术信息中心(National Center for Biotechnology Information, NCBI)的 GEO (Gene Expression Omnibus)数据库的 mRNA 微阵列数据, 以筛选 RA 和健康样本之间的差异表达基因 (DEGs)。然后, 通过对 DEGs 的 GO 和 KEGG 的富集分析, 进一步探讨了 RA 的分子机制。IGLJ3、IGHM、IGKC、IGLV1-44、GUSBP11 基因是与 RA 发生发展密切相关的基因。通过 R 语言 CIBERSORT 包进行免疫浸润评估。用 WGCNA 和 GSEA 分析 IGLJ3、IGHM、IGKC、IGLV1-44、GUSBP11 可能影响 RA 发病的机制。R 软件(版本 3.6.1; <https://www.r-project.org/>)和 Cytoscape 等用于数据处理。这项研究中发现的类风湿关节炎和健康对照之间的新基因和通路有望帮助揭示潜在的分子。

2. 材料和方法

2.1. 获取基因数据集

以类风湿性关节炎为关键词, 在 GEO 数据库(<https://www.ncbi.nlm.nih.gov/geo>)中搜索符合要求的基因芯片数据集 GSE55235, 数据集包括正常对照组和类风湿性关节炎组。利用平台的注释信息将探针转化为同源基因符号。

2.2. 数据校正和差异表达基因筛选

基于 R 软件(版本 4.0.3)的 limma 和 impute 包, 使用稳健多阵列平均算法(RMA) [13]对从 GEO 数据库下载的原始文件进行预处理和标准化。使用 limma 软件包进行样本间差异基因分析, 获得 p 值后进行多重假设检验和校正。筛选标准为 $\log_2(\text{fold change}) > 1$ 或 < -1 , 校正后的 p 值(Q value) < 0.05 。用 ggplot 软件包制作火山图显示差异基因的分布情况。

2.3. GO 和 KEGG 富集分析

为了对获得的 DEGs 进行功能性注释, 使用 R 软件进行基因本体论(Gene Ontology, GO)和京都基因与基因组百科全书(Kyoto Encyclopedia of Genes and Genomes, KEGG)的注释和可视化。R 软件 clusterProfiler、enrichplot、ggplot2、Circize、RColorBrewer、dplyr、ComplexHeatmap 软件包用于分析 DEGs 的 GO 和 KEGG 富集, 并使用 chord plot 软件包以可视化富集结果。

2.4. GSEA 富集分析

基因集富集分析(Gene Set Enrichment Analysis, GSEA)用于评估基因集中一组预定义基因的分布趋势, 以确定它们对表型的贡献。GSEA 4.1.0 版软件用于分析基因功能。我们下载了 GSEA_4.1.0 和 GO 基因集(c5.all.v7.5.1.symbols.gmt)、KEGG 基因集(c2.cp.KEGG.v7.5.1.symbols.gmt)用于功能富集分析。富集分析采用默认加权富集法。随机组合设定为 1000 次。使用 GSEA 分析对 DEG 进行 GO 和 KEGG 途径富集分析。FDR < 0.25 , NOM p 值 < 0.05 和 $|NES| > 1$ 被认为是显著富集[14]。

2.5. 构建加权基因共表达网络及蛋白互作网络

利用 R 软件中的 WGCNA 软件包构建了基于 DEGs 的加权基因共表达网络(Weighted gene co-expression network analysis, WGCNA) [15]。在本研究中,我们构建了一个邻接矩阵来描述基因之间的相关强度(软阈值参数)。随后,我们将邻接矩阵转化为拓扑重叠矩阵(Topological overlap matrix, TOM)来测量基因的网络连通性。接下来,我们执行分层聚类来识别模块,模块由分支和聚类树的不同颜色表示。最后,我们计算特征基因,对模块进行层次聚类,构建模块关系,计算基因模块与表型之间的相关性,识别与临床特征相关的模块。构建蛋白互作网络(Protein protein interaction network, PPI) (筛选阈值 = 0.5), 然后利用 Cytoscape 软件(<https://www.cytoscape.org/>)的 Cytohubba 插件[16], 综合 12 种算法排序得出 Top5 的基因。

2.6. 临床预测模型建立

通过 R 软件的“limma”软件包获得 5 个关键基因的表达量,使用“RMS”软件包,建立列线图模型来预测类风湿关节炎的发生。然后利用校准曲线评价诺模图模型的预测能力。最后通过决策曲线分析评价模型的临床价值。

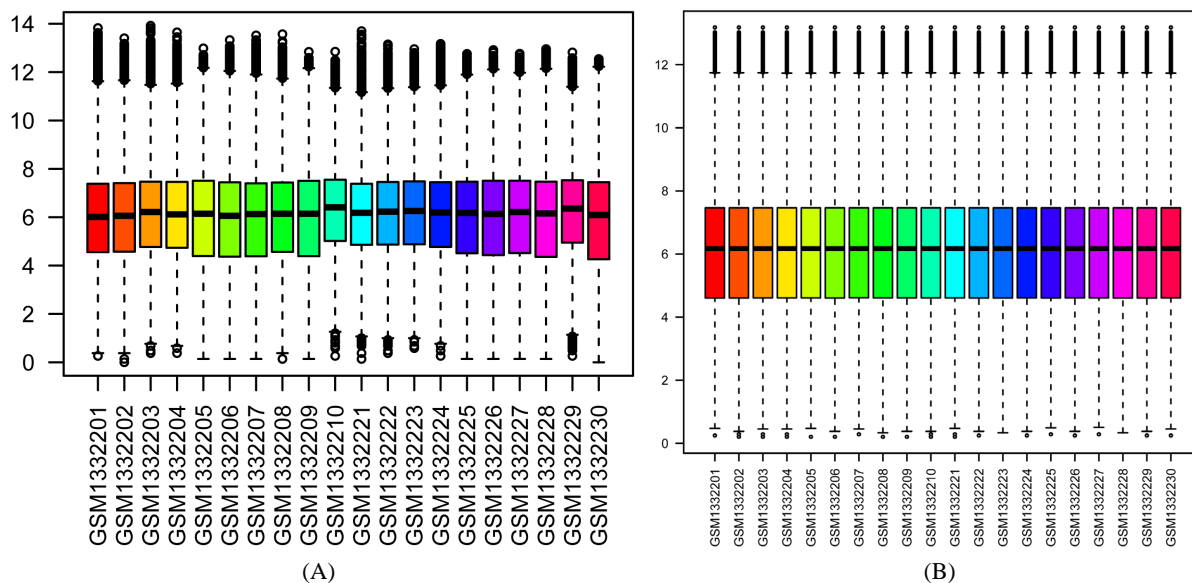
2.7. 免疫细胞浸润分析

我们使用 R 软件“e1071”和“parallel”软件包来评估免疫浸润的丰度,我们从 CIBERSORT 下载了 22 种浸润性免疫细胞的数据(<https://cibersort.stanford.edu/>)。我们使用 R 软件的“corrplot”软件包绘制相关热图,可视化 22 种浸润性免疫细胞的相关性,“ggplot2”软件包对免疫细胞浸润矩阵数据进行 PCA 聚类分析,绘制二维 PCA 聚类图,并绘制小提琴图,以可视化免疫细胞浸润的差异。

3. 结果

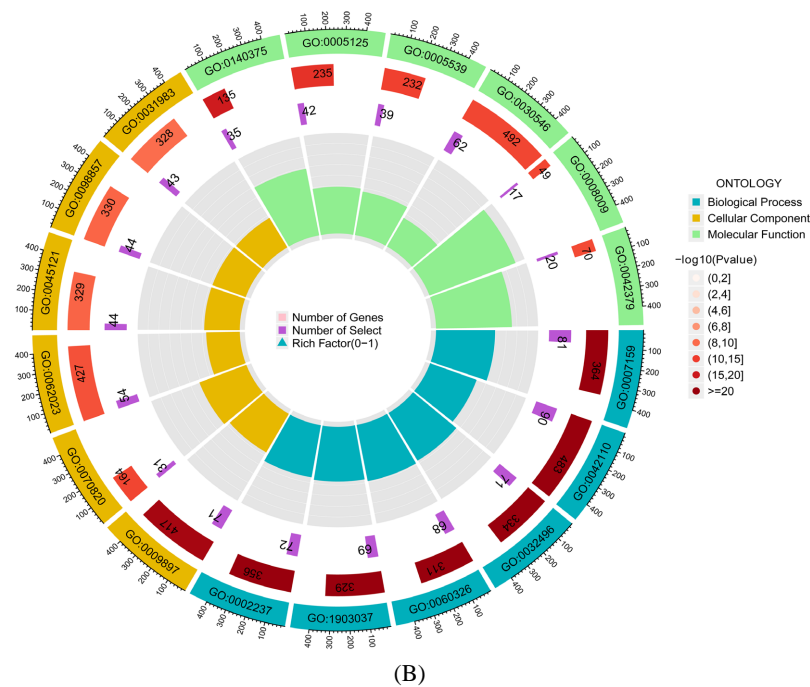
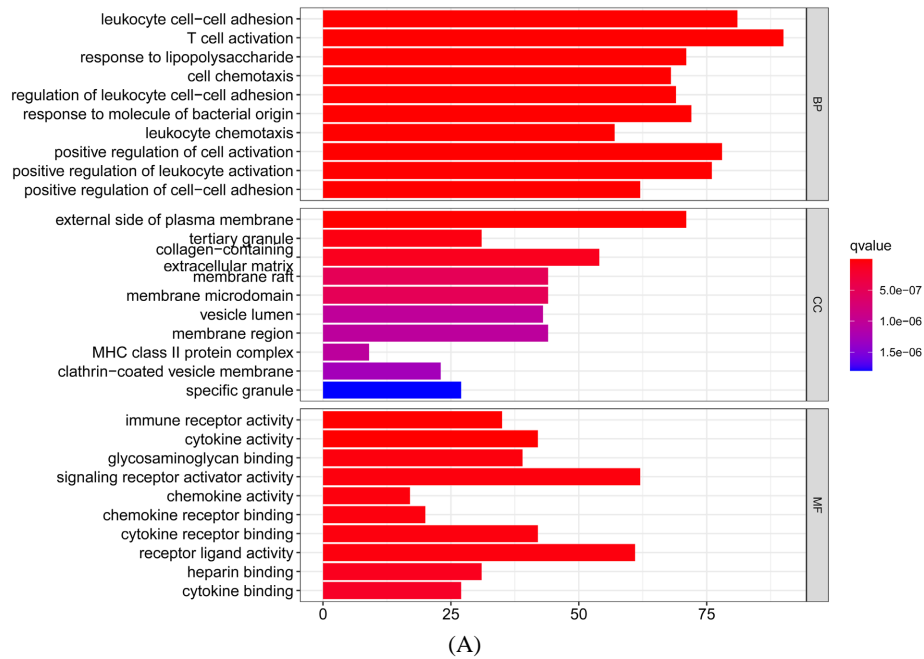
3.1. 数据处理和 DEGs 筛选结果

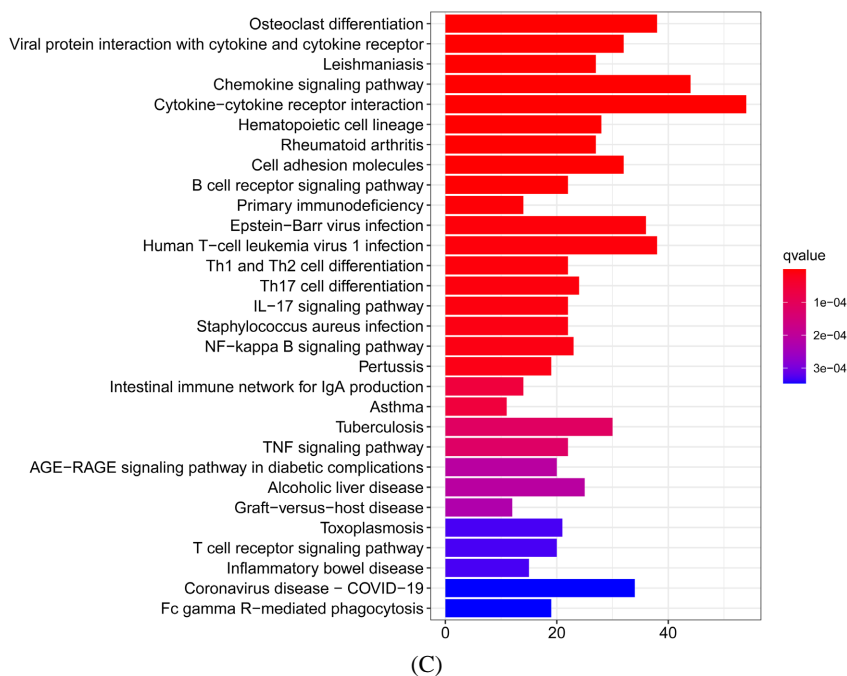
基于高通量分析,对原始数据进行校正和标准化处理,如(图 1(A)、图 1(B))所示。运用“limma”软件包,以 $\log_2(\text{fold change}) > 1$ 或 < -1 ,校正后的 p 值(Q value) < 0.05 为筛选阈值,在 RA 与正常对照组间筛选出 1071 个 DEGs,包括 599 个上调和 472 个下调基因。结果以火山图和热图可视化(图 1(C)、图 1(D))。



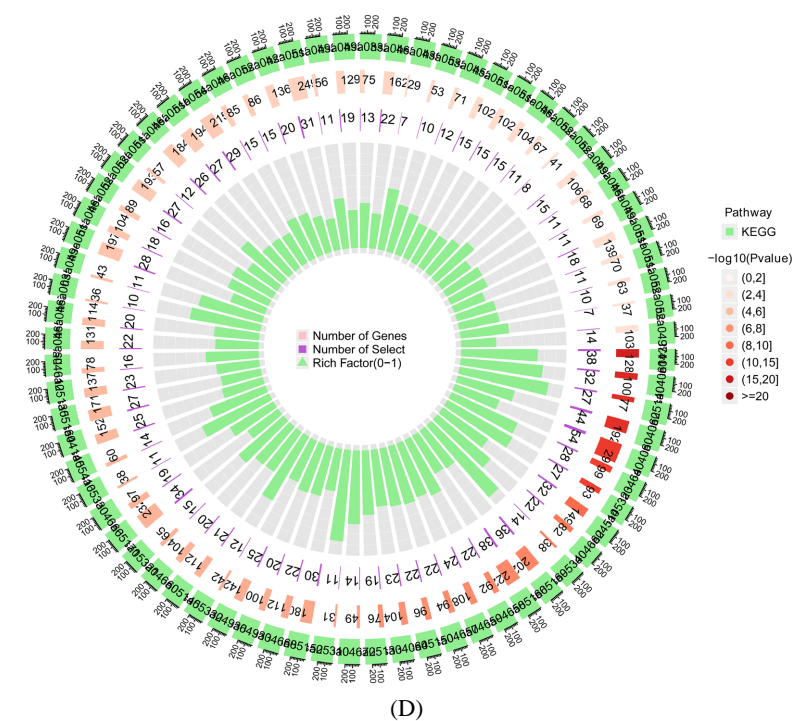
3.2. GO 和 KEGG 富集分析

GO 富集分析用于从分子功能、生物过程和细胞成分评估上述获得的差异基因的潜在机制。GO 富集结果表明，这些基因在生物过程模块上与 T 淋巴细胞活化、细胞活性的正调控、细胞黏附等有关。细胞组成模块富集表明这些基因发挥作用的位置可能位于质膜外侧、细胞外基质等。分子功能模块富集显示基因产物分子的功能主要集中于活化信号受体激活剂、受体 - 配体活化、活化细胞因子及其受体等(图 2(A)、图 2(B))。KEGG 途径分析表明，相关基因参与细胞因子 - 细胞因子受体相互作用、趋化因子信号通路、病毒蛋白与细胞因子及细胞因子受体的相互作用等(图 2(C)、图 2(D))。





(C)



(D)

Figure 2. (A) Bar graph of go enrichment analysis of differential genes; (B) Cycle diagram of go enrichment analysis of differential genes; (C) Bar graph of KEGG enrichment analysis of differential genes; (D) Circle diagram of KEGG enrichment analysis of differential genes

图 2. (A) 差异基因 GO 富集分析条形图; (B) 差异基因 GO 富集分析圈图; (C) 差异基因 KEGG 富集分析条形图; (D) 差异基因 KEGG 富集分析圈图

3.3. GSEA 富集分析

我们使用 GSEA 软件对数据集进行基因富集分析, 结果显示富集主要集中于: 活化细胞参与免疫反

应、B 淋巴细胞活化、B 淋巴细胞受体信号通路、自然杀伤细胞介导的细胞毒性反应、T 淋巴细胞受体信号通路、趋化因子信号通路等(图 3(A)~(F))。综合结果表明, B 淋巴细胞、T 淋巴细胞、自然杀伤细胞水平与 RA 的生物学途径密切相关。

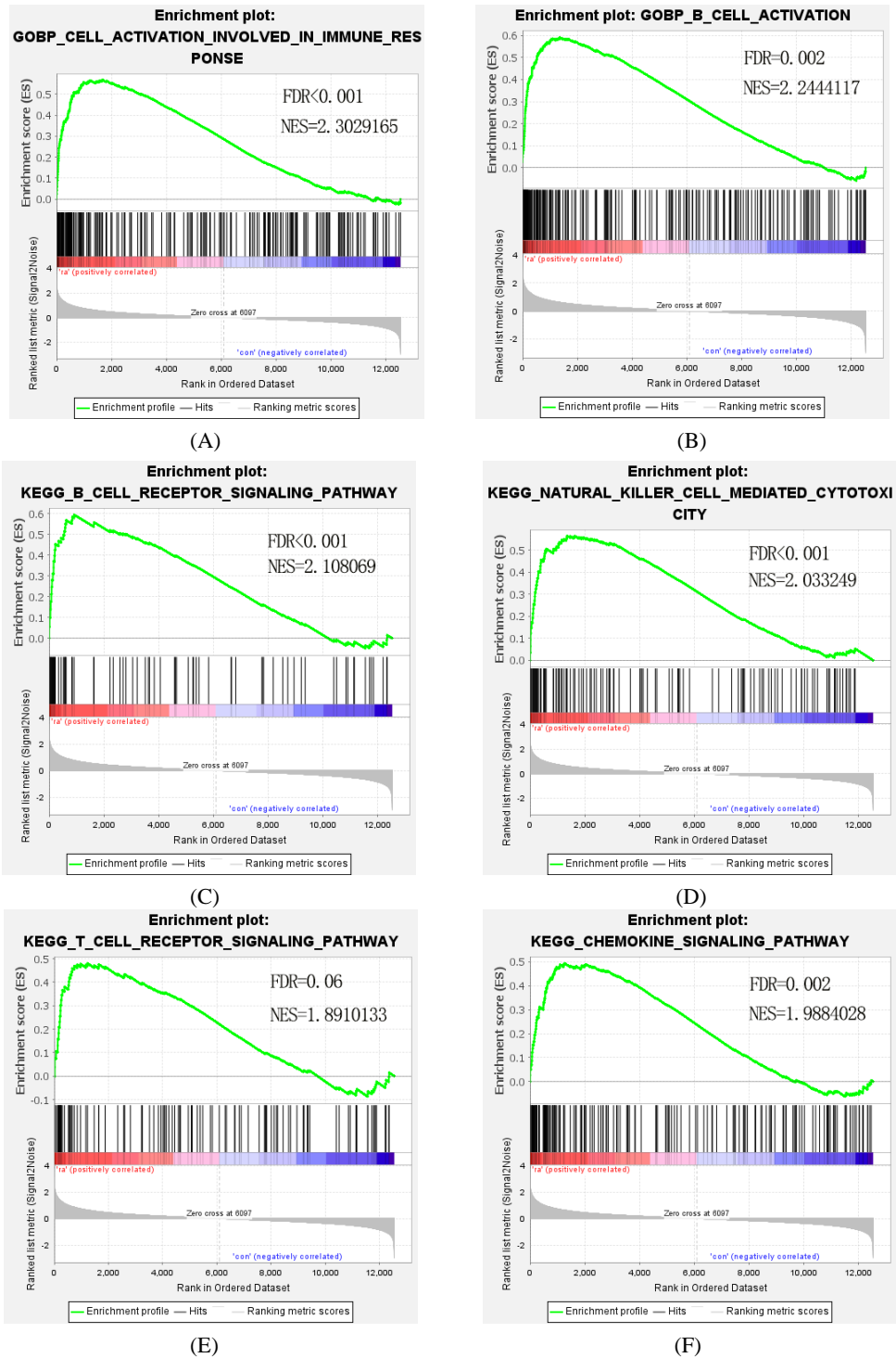
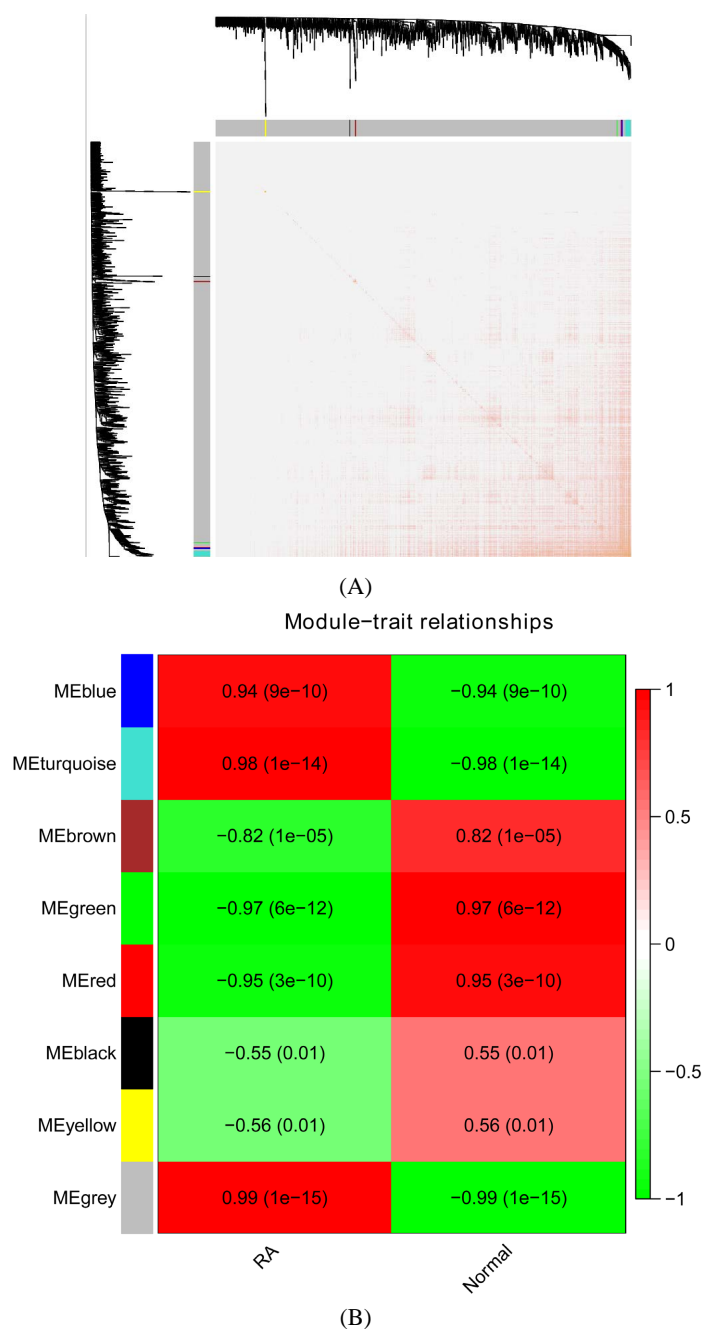


Figure 3. GSEA analyzes the enrichment degree of different signal pathways in the data set
图 3. GSEA 分析不同信号通路在数据集的富集程度

3.4. WGCNA 分析及 PPI 网络构建

我们选择之前筛选的差异基因(少于 5000)进行 WGCNA 分析。拓扑重叠热图显示蓝色、蓝绿色、灰色模块重叠显著,表明该模块的 WGCNA 分析效果好结果较准确(图 4(A))。模块-临床相关性热图显示灰色、蓝色、蓝绿色模块与 RA 的发生最为正相关,绿色、红色模块与 RA 的发生最为负相关(图 4(B))。基因-临床相关分析显示出排名前 10 的差异基因与 RA 发生的相关性,如表 1 所示。使用 WGCNA 分析结果进行基因-基因互作网络预测,将数据导入 Cytoscape 构建蛋白互作网络(图 4(C)),使用 cytoHubba 插件综合 MCC、DMNC、MNC、Degree、EPC 等 12 种算法,获得排名前 5 的关键基因包括 IGLJ3、IGHM、IGKC、IGLV1-44、GUSBP11 (图 4(D))。



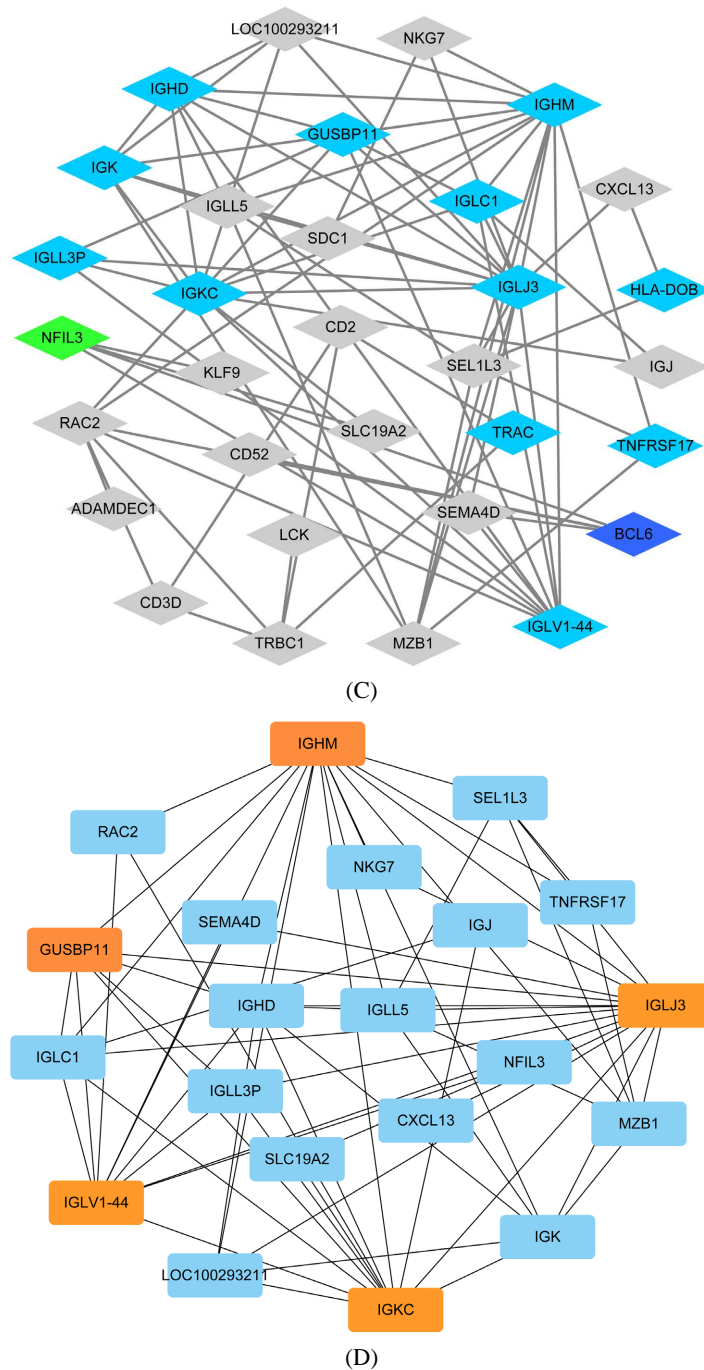


Figure 4. (A) The topology overlapping heat map shows that the blue, blue-green and gray modules overlap significantly, indicating that the analysis results of this module are more accurate; (B) Heat map of gene and module correlation for WGCNA enrichment analysis. Each color represents a common expression module. The value in the module is the correlation coefficient, and the value in brackets is the p value; (C) The gene interaction network was constructed by using the differential genes further screened by WGCNA. The color of genes represents that the genes belong to different modules of WGCNA analysis; (D) The main difference genes obtained by using the cytohubba plug-in, and the top 5 key genes are marked in orange

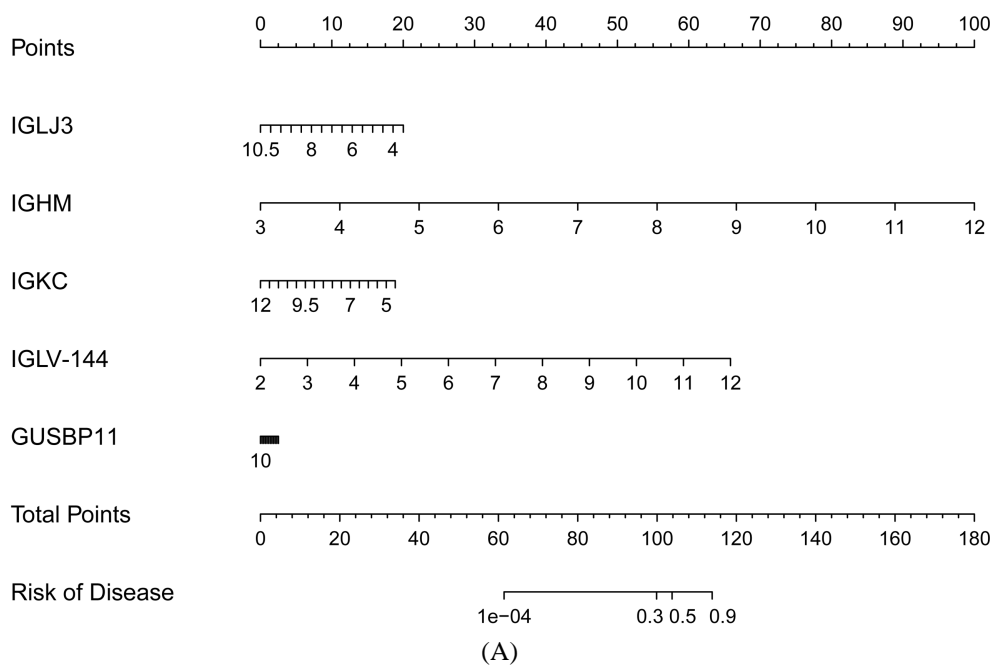
图 4. (A) 拓扑重叠热图显示蓝色、蓝绿色、灰色模块重叠显著，表明该模块的分析结果较准确；(B) WGCNA 富集分析的基因与模块相关性热图。每种颜色代表一个共表达模块，模块内的数值为相关性系数，括号内的为 P 值；(C) 利用 WGCNA 进一步筛选的差异基因构建的基因互作网络，基因的颜色代表基因属于 WGCNA 分析的不同模块；(D) 使用 cytohubba 插件获得的主要差异基因，排名前 5 的关键基因使用橘黄色标注

Table 1. Results of correlation analysis between main differential genes and RA
表 1. 主要差异基因与 RA 相关性分析结果

	RA		Normal		模块
	相关性	P 值	相关性	P 值	
IGHM	0.984963	3.53E-15	-0.98496	3.53E-15	蓝绿色
IGLV1-44	0.982177	1.62E-14	-0.98218	1.62E-14	蓝绿色
IGKC	0.981229	2.57E-14	-0.98123	2.57E-14	蓝绿色
IGLC1	0.976895	1.64E-13	-0.97689	1.64E-13	蓝绿色
GUSBP11	0.971997	9.09E-13	-0.972	9.09E-13	蓝绿色
TNFRSF17	0.971832	9.57E-13	-0.97183	9.57E-13	蓝绿色
IPLL3P	0.970404	1.49E-12	-0.9704	1.49E-12	蓝绿色
IPLJ3	0.970372	1.50E-12	-0.97037	1.50E-12	蓝绿色
LOC100293211	0.96967	1.85E-12	-0.96967	1.85E-12	灰色
IGK	0.968401	2.66E-12	-0.9684	2.66E-12	蓝绿色

3.5. 临床预测模型建立

我们使用 R 软件的“RMS”包建立了 RA 诊断列线图模型(图 5(A))。然后,运用校准曲线评估列线图的预测能力。结果表明,实际 RA 风险与预测风险之间的误差很小,表明列线图模型具有较高预测 RA 的准确性(图 5(B))。决策曲线(Decision Curve Analysis, DCA)显示红色曲线高于灰色曲线,在阈值 0.14 到 0.76 处患者净获益最大,且 C-index 为 0.732 说明模型准确性高(图 5(C))。



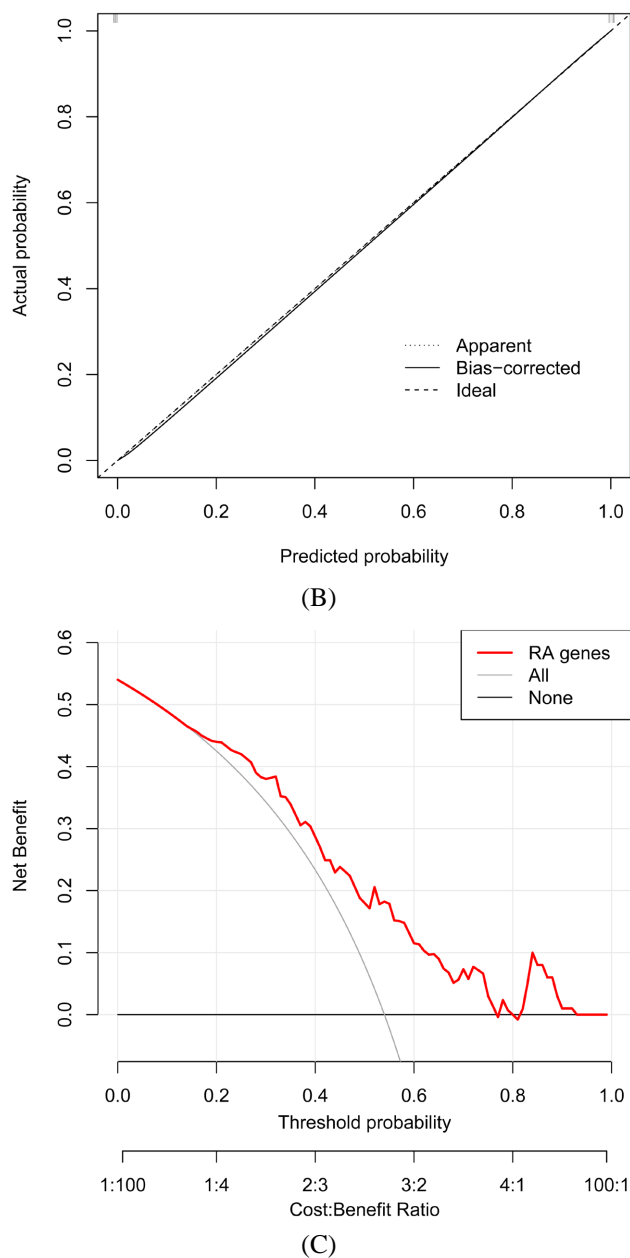


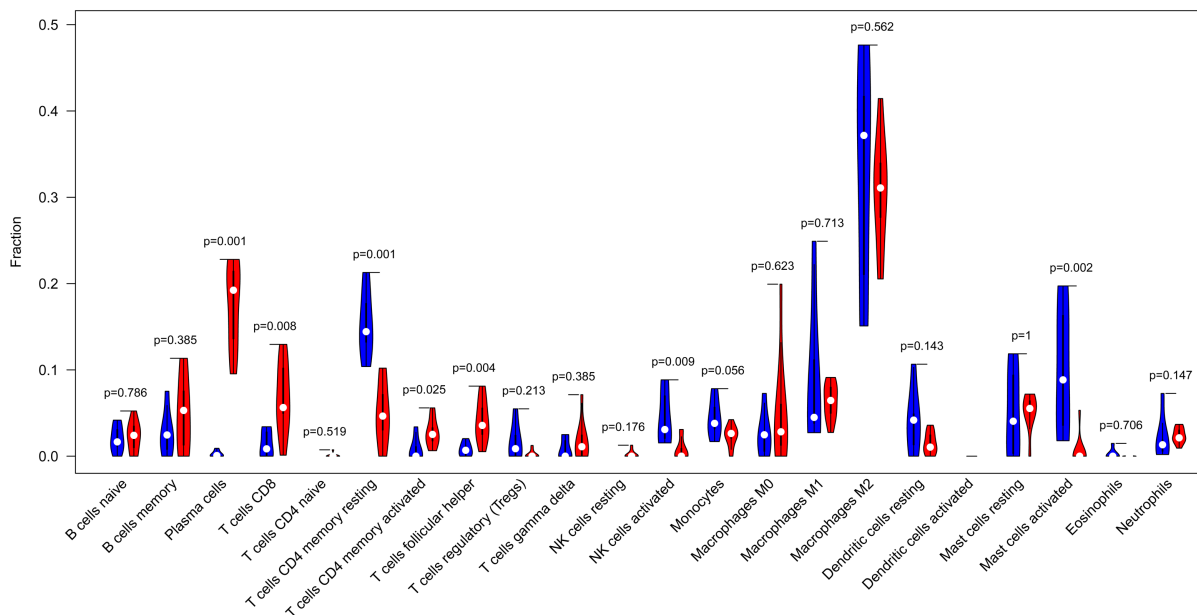
Figure 5. (A) Nomogram for predicting RA occurrence based on differential gene construction; (B) Calibration curve to evaluate the accuracy of nomograph model; (C) The decision curve (DCA) shows that the red curve is higher than the gray curve. When the threshold is 0.14 to 0.76, the net benefit of patients is the largest, and the C-index is 0.732, which proves that the model is very accurate

图 5. (A) 基于差异基因构建的预测 RA 发生的列线图; (B) 校准曲线评估列线图模型的准确性; (C) 决策曲线(DCA)表明, 红色曲线高于灰色曲线, 在阈值为 0.14 至 0.76 时, 患者的净收益最大, 且 C-index 为 0.732 证明该模型非常准确

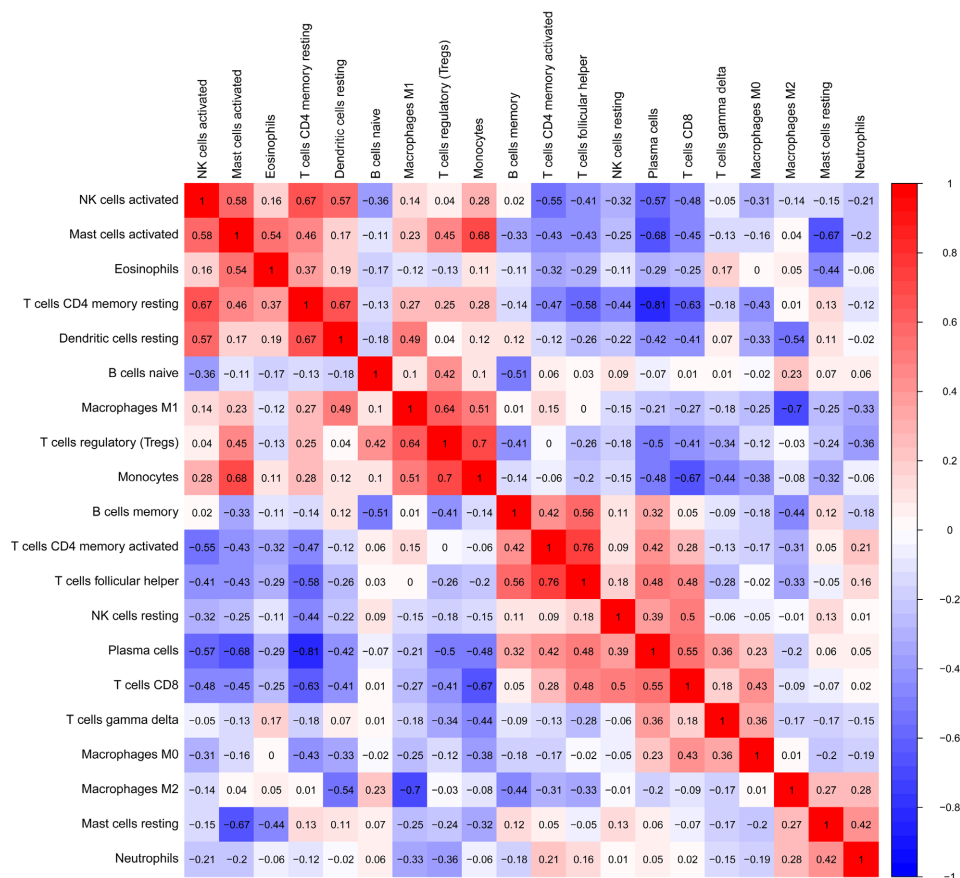
3.6. 免疫细胞浸润分析

免疫细胞浸润的小提琴图显示, 与正常对照样品相比, RA 患者中的浆细胞($p = 0.001$)和 CD8 阳性 T 细胞($p = 0.008$)浸润更多, 活性肥大细胞($p = 0.002$)和 CD4 阳性的静息记忆 T 细胞($p = 0.001$)浸润更少(图 6(A))。22 种免疫细胞的相关热图显示, 单核细胞与调节性 T 细胞呈正相关, 浆细胞与 CD4 阳性

的静息记忆 T 细胞呈负相关(图 6(B))。主成分分析(PCA)显示, RA 患者和正常对照组样本中的免疫细胞比例存在个体差异(图 6(C))。



(A)



(B)

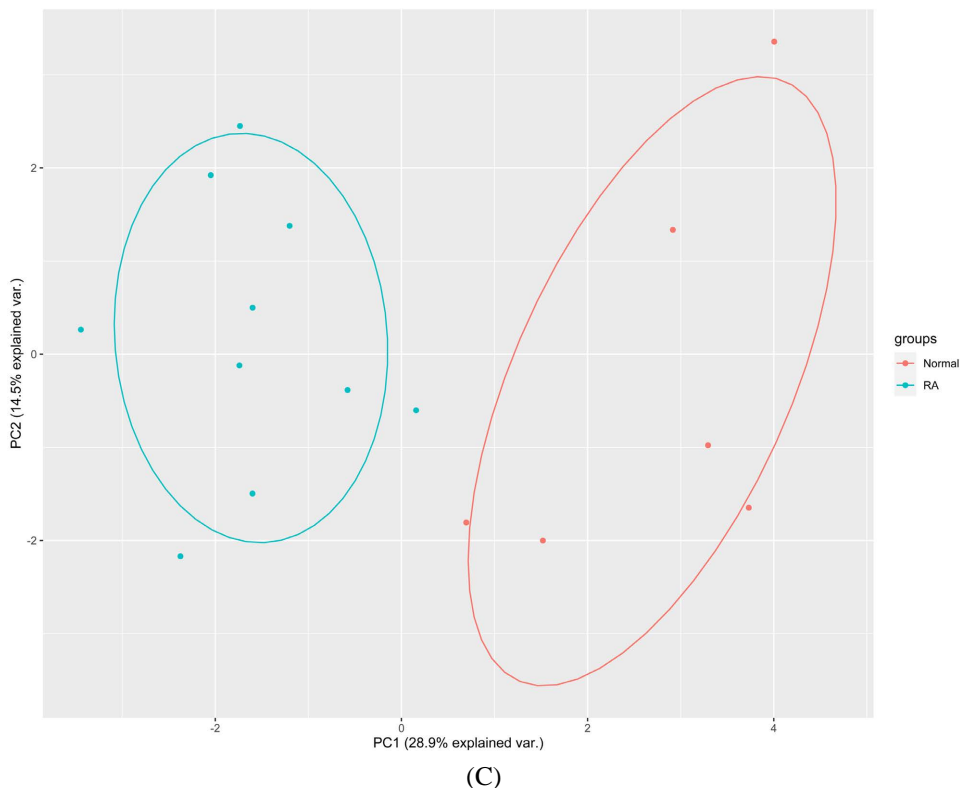


Figure 6. (A) The violin diagram shows the proportion of 22 immune cells in the data set. Plasma cells and CD8 positive T cells were more infiltrated in the data set, while active mast cells and CD4 positive resting memory T cells were less infiltrated; (B) Heat maps of 22 immune cells. The number of colored squares indicates the strength of the correlation. Red represents positive correlation and blue represents negative correlation. The darker the color, the stronger the correlation; (C) Principal component analysis of all samples in the data set showed that there were differences between immunophenotypes

图 6. (A) 小提琴图显示了数据集 22 种免疫细胞的比例。数据集中的浆细胞和 CD8 阳性 T 细胞浸润较多，活性肥大细胞和 CD4 阳性的静息记忆 T 细胞浸润较少；(B) 22 种免疫细胞的相关热图。彩色方块的数量表示相关性的强度。红色代表正相关，蓝色代表负相关。颜色越深，相关性越强；(C) 对数据集的所有样本进行主成分分析，结果表明免疫表型之间存在差异

4. 讨论

RA 是最常见的慢性炎性关节疾病之一，常常由于炎症失控而导致残疾[17]。因此，迫切需要更好地掌握 RA 诊断和治疗新策略的详细机制[18] [19]。在这项研究中，基于从 GEO 数据库数据集获得的基因表达谱，通过一系列生物信息学分析确定了 5 种最具差异的基因(IGLJ3、IGHM、IGKC、IGLV1-44、GUSBP11)。随后，我们对这些差异表达基因的生物学功能进行了进一步的研究，GO 分析显示这些 DEGs 与免疫功能的变化显著相关。GO、KEGG 富集分析和 GSEA 富集分析均表明，B 淋巴细胞通路、T 淋巴细胞通路、趋化因子信号通路是 RA 的关键通路，这与之前的研究一致[20] [21] [22] [23]。这些 DEGs 有助于促进 RA 的诊断和治疗，为了解 RA 的研究指明了新的方向。为了更好地了解 RA 的进展，使用 WGCNA 来识别 RA 的候选生物标志物。最后，通过 WGCNA 分析构建了一些 RA 相关模块。综合各个模块分析及与患者诊断的相关性，我们确定了 IGLJ3、IGHM、IGKC、IGLV1-44、GUSBP11 基因在 RA 疾病发生发展中具有重要作用。

免疫球蛋白在类风湿关节炎的发生发展及诊断治疗中一直有着重要作用，类风湿因子(RF)是一种结合 IgG Fc 段的自身抗体，是自身免疫在 RA 种发挥作用最早的直接证据。与血清 RF 阴性的 RA 患者相比，血清 RF 阳性的 RA 患者有更严重的临床表现和并发症[24]。之前的研究发现，RA 中的 RF 主要由 IgG 重链

可变区 3 (IGVH3)基因和一组轻链可变区(VL)基因编码,而具有 RF 活性的天然抗体则由 IGVH1、IGVH4 和 IGKV3 基因编码[25]。最近的一项研究发现位于 RA 患者上经 4-羟基-2-壬烯醛修饰免疫球蛋白 KappaC (IGKC)轻链构成的 IgG 和 IgM 较正常对照组增多,结果表明高水平的 IGKC 可能会增加患 RA 的风险[26]。我们的研究新发现 IGLJ3、IGHM、IGKC、IGLV1-44 四种免疫球蛋白基因在 RA 的发生发展中起到了重要作用,为进一步探究 RA 的发生机制及诊断治疗提供了新的靶点。

此外,由研究表明,免疫细胞浸润在 RA 的发病过程中起重要作用。因此,寻找特异性诊断标志物,分析 RA 的免疫细胞浸润,对改善 RA 患者的临床症状和预后具有深远意义。我们使用 R 软件的 CIBERSORT 包对 RA 的免疫浸润进行了综合分析。我们发现 CD8 阳性 T 淋巴细胞和浆细胞浸润增加。CD4 阳性的静息记忆 T 淋巴细胞和活性肥大细胞浸润减少可能与 RA 的严重程度显著相关。亚索等人发现记忆性 CXCR4 阳性 CD4 阳性的 T 淋巴细胞与 RA 的严重程度显著相关[27]。单核细胞与调节性 T 细胞呈正相关,提示二者在 RA 的发生发展中具有正向协同作用。浆细胞与 CD4 阳性的静息记忆 T 细胞呈负相关,提示二者在 RA 疾病过程中具有负向协同作用。在 RA 患者中, B 淋巴细胞分化成浆细胞,产生类风湿因子和抗环瓜氨酸肽抗体[28]。临床上利妥昔单抗通过消耗 B 淋巴细胞来减轻使用肿瘤坏死因子抑制剂不能控制病情的 RA 患者。可见,这些免疫细胞浸润机制与 RA 发病相关,为进一步探索 RA 的免疫浸润及临床诊断治疗提供了新的研究思路。

然而本研究尚有不足之处。一是纳入的基因芯片样本量有限,容易造成误差,需进行一定规模的体外实验以验证所得的关键基因;二是 CIBERSORT 算法是基于有限的遗传数据,导致免疫浸润评估时出现计算偏倚,仍需进一步实验验证。今后课题组的相关工作将基于生物样本进行实验研究,以验证本研究结果的准确性,并进一步探讨关键基因和免疫浸润之间的调控关系。

5. 结论

综上所述,根据我们的分析结果及文献证据表明, IGLJ3、IGHM、IGKC、IGLV1-44、GUSBP11 基因在 RA 的发生发展中具有重要作用,应成为进一步研究的重点。这将有助于确定 RA 的新型诊断标志物和治疗靶点,也为后续的发病机制研究提供可靠的依据和全新的视角。

参考文献

- [1] Scott, D.L., Wolfe, F. and Huizinga, T.W. (2010) Rheumatoid Arthritis. *The Lancet*, **376**, 1094-1108. [https://doi.org/10.1016/S0140-6736\(10\)60826-4](https://doi.org/10.1016/S0140-6736(10)60826-4)
- [2] Littlejohn, E.A. and Monrad, S.U. (2018) Early Diagnosis and Treatment of Rheumatoid Arthritis. *Primary Care*, **45**, 237-255. <https://doi.org/10.1016/j.pop.2018.02.010>
- [3] Choy, E. (2012) Understanding the Dynamics: Pathways Involved in the Pathogenesis of Rheumatoid Arthritis. *Rheumatology (Oxford)*, **51**, v3-v11. <https://doi.org/10.1093/rheumatology/kes113>
- [4] Deane, K.D., Demoruelle, M.K., Kelmenson, L.B., et al. (2017) Genetic and Environmental Risk Factors for Rheumatoid Arthritis. *Best Practice & Research: Clinical Rheumatology*, **31**, 3-18. <https://doi.org/10.1016/j.berh.2017.08.003>
- [5] Zhao, X., Zhang, L., Wang, J., et al. (2021) Identification of Key Biomarkers and Immune Infiltration in Systemic Lupus Erythematosus by Integrated Bioinformatics Analysis. *Journal of Translational Medicine*, **19**, Article No. 35. <https://doi.org/10.1186/s12967-020-02698-x>
- [6] The Wellcome Trust Case Control Consortium (2007) Genome-Wide Association Study of 14,000 Cases of Seven Common Diseases and 3,000 Shared Controls. *Nature*, **447**, 661-678. <https://doi.org/10.1038/nature05911>
- [7] Begovich, A.B., Carlton, V.E., Honigberg, L.A., et al. (2004) A Missense Single-Nucleotide Polymorphism in a Gene Encoding a Protein Tyrosine Phosphatase (PTPN22) Is Associated with Rheumatoid Arthritis. *The American Journal of Human Genetics*, **75**, 330-337. <https://doi.org/10.1086/422827>
- [8] Rodríguez, M.R., Núñez-Roldán, A., Aguilar, F., et al. (2002) Association of the CTLA4 3' Untranslated Region Polymorphism with the Susceptibility to Rheumatoid Arthritis. *Human Immunology*, **63**, 76-81. [https://doi.org/10.1016/S0198-8859\(01\)00358-5](https://doi.org/10.1016/S0198-8859(01)00358-5)

- [9] Suzuki, A., Yamada, R., Chang, X., *et al.* (2003) Functional Haplotypes of PADI4, Encoding Citrullinating Enzyme Peptidylarginine Deiminase 4, Are Associated with Rheumatoid Arthritis. *Nature Genetics*, **34**, 395-402. <https://doi.org/10.1038/ng1206>
- [10] Van Der Helm-Van Mil, A.H., Wesoly, J.Z. and Huizinga, T.W. (2005) Understanding the Genetic Contribution to Rheumatoid Arthritis. *Current Opinion in Rheumatology*, **17**, 299-304. <https://doi.org/10.1097/01.bor.0000160780.13012.be>
- [11] Lee, Y.H., Bae, S.C., Kim, J.H., *et al.* (2014) Toll-Like Receptor Polymorphisms and Rheumatoid Arthritis: A Systematic Review. *Rheumatology International*, **34**, 111-116. <https://doi.org/10.1007/s00296-013-2666-7>
- [12] Song, J., Kim, D., Han, J., *et al.* (2015) PBMC and Exosome-Derived Hotair Is a Critical Regulator and Potent Marker for Rheumatoid Arthritis. *Clinical and Experimental Medicine*, **15**, 121-126. <https://doi.org/10.1007/s10238-013-0271-4>
- [13] Cheng, Q., Chen, X., Wu, H., *et al.* (2021) Three Hematologic/Immune System-Specific Expressed Genes Are Considered as the Potential Biomarkers for the Diagnosis of Early Rheumatoid Arthritis through Bioinformatics Analysis. *Journal of Translational Medicine*, **19**, 18. <https://doi.org/10.1186/s12967-020-02689-y>
- [14] Subramanian, A., Tamayo, P., Mootha, V.K., *et al.* (2005) Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 15545-15550. <https://doi.org/10.1073/pnas.0506580102>
- [15] Langfelder, P. and Horvath, S. (2008) WGCNA: An R Package for Weighted Correlation Network Analysis. *BMC Bioinformatics*, **9**, Article No. 559. <https://doi.org/10.1186/1471-2105-9-559>
- [16] Chin, C.H., Chen, S.H., Wu, H.H., *et al.* (2014) cytoHubba: Identifying Hub Objects and Sub-Networks from Complex Interactome. *BMC Systems Biology*, **8**, S11. <https://doi.org/10.1186/1752-0509-8-S4-S11>
- [17] Zhang, Y., Qian, X., Yang, X., *et al.* (2020) ASIC1a Induces Synovial Inflammation via the Ca(2+)/NFATc3/RANTES Pathway. *Theranostics*, **10**, 247-264. <https://doi.org/10.7150/thno.37200>
- [18] Tanaka, H., Arakawa, H., Yamaguchi, T., *et al.* (2000) A Ribonucleotide Reductase Gene Involved in a p53-Dependent Cell-Cycle Checkpoint for DNA Damage. *Nature*, **404**, 42-49. <https://doi.org/10.1038/35003506>
- [19] Jie, L.G., Huang, R.Y., Sun, W.F., *et al.* (2015) Role of Cysteine-Rich Angiogenic Inducer 61 in Fibroblast-Like Synovial Cell Proliferation and Invasion in Rheumatoid Arthritis. *Molecular Medicine Reports*, **11**, 917-923. <https://doi.org/10.3892/mmr.2014.2770>
- [20] Elhage, A., Dargier, C., Mordon, S., *et al.* (1992) Proximal Tubular Recanalization by Endoluminal Laser. II. Functional Results in the Female Rabbit. *Journal de Gynécologie Obstétrique et Biologie de la Reproduction (Paris)*, **21**, 151-154.
- [21] Hirota, K., Yoshitomi, H., Hashimoto, M., *et al.* (2007) Preferential Recruitment of CCR6-Expressing Th17 Cells to Inflamed Joints via CCL20 in Rheumatoid Arthritis and Its Animal Model. *Journal of Experimental Medicine*, **204**, 2803-2812. <https://doi.org/10.1084/jem.20071397>
- [22] Kim, W.J., Kang, Y.J., Koh, E.M., *et al.* (2005) LIGHT Is Involved in the Pathogenesis of Rheumatoid Arthritis by Inducing the Expression of Pro-Inflammatory Cytokines and MMP-9 in Macrophages. *Immunology*, **114**, 272-279. <https://doi.org/10.1111/j.1365-2567.2004.02004.x>
- [23] Tsubaki, T., Takegawa, S., Hanamoto, H., *et al.* (2005) Accumulation of Plasma Cells Expressing CXCR3 in the Synovial Sublining Regions of Early Rheumatoid Arthritis in Association with Production of Mig/CXCL9 by Synovial Fibroblasts. *Clinical & Experimental Immunology*, **141**, 363-371. <https://doi.org/10.1111/j.1365-2249.2005.02850.x>
- [24] Liang, K.P., Kremers, H.M., Crowson, C.S., *et al.* (2009) Autoantibodies and the Risk of Cardiovascular Events. *The Journal of Rheumatology*, **36**, 2462-2469. <https://doi.org/10.3899/jrheum.090188>
- [25] Bouvet, J.P., Wu, Y.X. and Pillot, J. (1987) Restricted Heterogeneity of Polyclonal Rheumatoid Factors. *Arthritis & Rheumatology*, **30**, 998-1005. <https://doi.org/10.1002/art.1780300906>
- [26] Tsai, K.L., Chang, C.C., Chang, Y.S., *et al.* (2021) Isotypes of Autoantibodies against Novel Differential 4-Hydroxy-2-nonenal-Modified Peptide Adducts in Serum Is Associated with Rheumatoid Arthritis in Taiwanese Women. *BMC Medical Informatics and Decision Making*, **21**, Article No. 49. <https://doi.org/10.1186/s12911-020-01380-y>
- [27] Nagafuchi, Y., Shoda, H., Sumitomo, S., *et al.* (2016) Immunophenotyping of Rheumatoid Arthritis Reveals a Linkage between HLA-DRB1 Genotype, CXCR4 Expression on Memory CD4(+) T Cells, and Disease Activity. *Scientific Reports*, **6**, Article No. 29338. <https://doi.org/10.1038/srep29338>
- [28] Cambridge, G., Leandro, M.J., Edwards, J.C., *et al.* (2003) Serologic Changes Following B Lymphocyte Depletion Therapy for Rheumatoid Arthritis. *Arthritis & Rheumatology*, **48**, 2146-2154. <https://doi.org/10.1002/art.11181>