

类风湿性关节炎是一种常见的慢性炎症性关节疾病，以滑膜炎、关节损伤和全身免疫反应为主要特征。早期诊断，及早治疗是降低疾病活性，改善 RA 患者病情的关键[1]。利妥昔单抗是一种抗 CD20 单克隆抗体，起初用来治疗淋巴瘤、白血病，后应用于治疗 RA。对于处于严重疾病活动期的患者以及对于其它抗风湿药物耐受的 RA 患者有着显著的疗效[2, 3]，并可改善 RA 合并肺间质性疾病患者的病情[4]。但并非所有的 RA 患者对 RTX 都能产生良好的疾病反应，有相当一部分 RA 患者应用 RTX 治疗无效，甚至产生严重的不良反应如肾脏淀粉样变[5]，或者发生严重感染事件[6]。因此，及早识别出不适宜应用 RTX 治疗的 RA 患者，避免病情的延误和不良反应的发生，对患者具有重要的意义。

本研究利用生物信息学方法分析与 RTX 治疗 RA 疗效相关的基因表达数据，对相关基因的分子机制进行综合分析，筛选出 RTX 疗效相关的核心基因，为 RTX 对于 RA 患者的治疗应用提供帮助。

1 材料与方法

1.1 微阵列数据的获取与处理

从基因表达数据库(<http://www.ncbi.nlm.nih.gov/geo/>)中获得与 RTX 治疗 RA 相关的基因表达谱 GSE24742，其中包括 8 例应用 RTX 治疗反应良好的 RA 患者的滑膜样本，以及 3 例应用 RTX 治疗反应较差的 RA 患者的滑膜样本。该芯片的平台文件为 GPL570。

应用在线分析工具 GEO2R (<http://www.ncbi.nlm.nih.gov/geo/geo2r/>)对微阵列数据进行分析[7]，比较 RTX 疗效较好与疗效较差的两组 RA 患者的差异表达基因 (DEGs)。以 $P < 0.05$ ，差异倍数 $|\log_2 FC| \geq 2$ 作为差异基因的筛选条件。

1.2 差异基因的 GO 功能富集分析及 KEGG 通路富集分析

将筛选得到的 DEGs 应用 DAVID 数据库 (<https://david.ncifcrf.gov/>) 进行基因本体 (gene ontology, GO) 功能富集分析和《京都基因与基因组百科全书》(kyoto encyclopedia of genes and genomes, KEGG) 基因通路富集分析[8]， $P < 0.05$ 的富集分析结果有意义。

1.3 PPI 网络的构建及网络中心性分析

应用 STRING 数据库 (<https://string-db.org/>) 对筛选出的 DEGs 进行蛋白-蛋白互作 (PPI) 网络分析[9]，最小有效结合分数设定为 0.7，将分析结果导入 Cytoscape 软件 (<http://www.cytoscape.org/>)，利用 cytoHubba 插件对或网络中每个节点的度 (degree) 进行计算[10]，按度计算结果由高到低进行排序，选取度值排名前 10 的基因为网络核心基因。

2 结果

2.1 微阵列数据差异基因筛选结果

芯片数据 GSE24742 通过 GEO2R 软件的分析，比较 RTX 治疗 RA 疗效较好组与疗效较差组之间的 DEGs，共得出 727 个 DEGs，其中包括上调基因 522 个，下调基因 205 个 ($|\log_2 FC| \geq 2, P < 0.05$)，如图 1 所示。

2.2 GO 和 KEGG 通路富集分析

利用 DAVID 数据库对 727 个 DEGs 进行功能富集分析及通路富集分析。GO 富集分析主要包含生物学过程、细胞组分和分子功能三大模块。GO 富集分析结果如图 2A 所示，DEGs 在生物学过程模块主要富集在基因表达的正调控，转录的正调控以及细胞内信号转导 (图 2B)，DEGs 在细胞组分模块主要富集在细胞

外区域及核内体(图 2C),DEGs 在分子功能模块主要富集在与 NFAT 蛋白结合,与运动活动相关等方面(图 2D)。

KEGG 通路分析结果表明,DEGs 主要富集在 PI3K-Akt 信号通路及 Rap1 信号通路等,对其结果进行可视化,如图 3。

2.3 PPI 网络的构建及网络中心性分析

利用 STRING 数据库对筛选出的 727 个 DEGs 构建 PPI 网络如图 4A 所示。将 DEGs 互作关系结果导入到 Cytoscape,应用 cytoHubba 插件计算每个节点的度值,将度值位于前 10 的基因定义为网络核心基因,分别为 *EGFR*,*HIST1H2BB*,*MMP9*,*CXCR3*,*MEIS1*,*CXCL5*,*IL17A*,*WASL*,*KMT2C*,*PGF*(图 4B)。

3 讨论

本研究选取 RTX 治疗 RA 患者相关的芯片数据 GSE24742,利用生物信息学技术进行处理,在 RTX 疗效较好与较差两组间共筛选出 727 个 DEGs,其中包括上调基因 522 个,下调基因 205 个,对 DEGs 进行 GO 功能富集分析及 KEGG 通路富集分析并构建 PPI 互作网络,识别 10 个核心基因 *EGFR*,*HIST1H2BB*,*MMP9*,*CXCR3*,*MEIS1*,*CXCL5*,*IL17A*,*WASL*,*KMT2C*,*PGF*。

RTX 对于 RA 患者的疗效涉及多种基因及细胞通路,发现 RTX 疗效相关网络的核心基因,了解其在 RTX 治疗 RA 患者相关的分子作用机制中有助于分辨适于应用 RTX 治疗的 RA 患者,在基因学水平上为 RTX 疗效的预测提供帮助。

GO 富集分析结果显示,DEGs 主要参与基因表达的正调控,转录的正调控以及细胞内信号转导。KEGG 通路富集分析发现,DEGs 主要富集在 PI3K-Akt 信号通路及 Rap1 信号通路等。上述结果表明,与 RTX 疗效相关的差异基因主要参与 RA 炎症相关细胞的增殖与凋亡。

对筛选出的 DEGs 构建 PPI 网络,识别处于网络核心位置的基因,其中以 *MMP9* 表达差异最为显著。*MMP* 是由巨噬细胞和滑膜成纤维细胞产生的基质金属蛋白酶,可以降解细胞外基质成分,从而造成关节损伤[11]。既往研究表明,*MMP* 在 RA 患者的损伤软骨部位有着较高的表达量,累积的 *MMP* 水平与关节的炎症有关,降低 *MMP* 水平可以改善关节肿胀,减轻局部炎症,避免加重关节损伤[12]。在本研究中,RTX 疗效差的组相较于疗效好的组,*MMP* 水平表达显著升高。

PPI 网络中的核心基因 *EGFR*,作为表皮因子生长受体,参与 DNA 合成和细胞增殖,在 RA 患者血清中表达量高于正常人群,并与 RA 疾病发展有着紧密的联系[13-15]。趋化因子受体 *CXCR3* 是一种 G 蛋白偶联受体,在 RA 患者的外周血和滑膜中均有明显表达。*CXCR3* 可介导免疫细胞调控,参与 B 细胞向滑膜的迁移,以及滑膜内 B 细胞的局部增殖[16],进而通过抗原呈递、产生自身抗体等多种途径引发免疫和炎症反应,参与 RA 的发病机制[17]。*IL-17* 是由 CD4⁺T 细胞产生的细胞因子,在自身免疫性疾病中表达增加,*IL-17* 作为炎性因子可导致 RA 患者骨及软骨的损伤[18],在 RA 的发病机制中发挥核心作用。*IL-17* 还可以激活 RA 患者滑膜组织成纤维细胞产生大量 CXC 趋化因子,并诱导中性粒细胞内陷和颗粒形成导致炎症反应[19]。尤其是 *CXCL5* 与 *IL-17* 诱导的炎症反应密切相关[20]。瓜氨酸化 *CXCL5* 可招募单核细胞使关节组织发炎肿胀,参与 RA 的发生发展[21]。RTX 作为治疗 RA 的生物制剂,可以抑制由 B 细胞驱动的 T 细胞,进而抑制促炎细胞因子的激活,对 RA 患者的疾病起到缓解作用[22]。但是,不同 RA 患者核心基因表达量不同,炎症反应程度不等,应用 RTX 治疗的疗效

不同。对网络核心基因进行进一步的研究或许可以为判别 RA 患者是否适宜应用 RTX 治疗提供基因学水平的依据。

综上所述，利用生物信息学方法对与 RTX 治疗 RA 患者疗效相关的芯片数据进行分析，可以帮助我们提高对 RTX 疗效相关的潜在分子事件的认识，通过网络构建筛选出的核心基因可能对预先判断 RTX 的疗效具有重要的意义。但是，这些核心基因的具体功能仍有待进一步研究。