

· 论著 ·

烧伤后瘢痕组织内甲基化基因的加权基因共表达网络分析

郭鹏

河北医科大学第四医院整形外科, 石家庄 050011

通信作者: 郭鹏, Email: 277051445@163.com



【摘要】 **目的** 采用加权基因共表达网络分析(WGCNA)探讨烧伤后瘢痕组织内甲基化基因, 发掘瘢痕形成过程中的分子标志物及治疗靶点。 **方法** 采用观察性研究方法。从美国国家生物技术信息中心基因表达综合数据库筛选出同一批次分别进行了 mRNA 测序和甲基化测序的 GSE136906 和 GSE137134 数据集(数据集的样本数均为 6), 并利用数据集 GSE108110(样本数为 9)纳入支持向量机及建模分析。采用“Limma”软件包对烧伤后瘢痕和正常组织之间的差异表达基因和差异甲基化基因进行鉴定。使用 WGCNA 筛选和瘢痕组织临床特征相关性强且基因数目多的模块。对模块内的基因进行功能富集分析并寻找模块内异常甲基化状态的基因, 使用受试者操作特征(ROC)曲线判断异常甲基化基因对瘢痕的诊断效能, 并使用支持向量机模型进行验证。 **结果** 共鉴定出 10 个共表达基因模块, 棕色模块与烧伤后瘢痕组织形成相关性高且基因数目多。棕色模块内的基因主要富集在雄激素受体信号通路的调控、细胞因子-细胞因子受体相互作用、胰岛素分泌的正调节等方面。棕色模块内显示 35 个基因甲基化状态异常, ROC 曲线(曲线下面积>0.9)与支持向量机模型(准确率为 93.3%)表明基因 *CCR2*、*LMO7*、*STEAP4*、*NNAT* 和 *TCF7L2* 具有良好的瘢痕诊断效能。 **结论** *CCR2*、*LMO7*、*STEAP4*、*NNAT* 和 *TCF7L2* 可作为烧伤后瘢痕治疗的潜在靶点。

【关键词】 烧伤; 瘢痕; 甲基化; 加权基因共表达网络; *CCR2*; *LMO7*

Weighted gene co-expression network analysis of methylated genes in burn scar tissue

Guo Peng

Department of Plastic Surgery, the Fourth Hospital of Hebei Medical University, Shijiazhuang 050011, China

Corresponding author: Guo Peng, Email: 277051445@163.com

【Abstract】 **Objective** To investigate the methylated genes in burn scar tissue by weighted gene co-expression network analysis (WGCNA), and to discover molecular markers and therapeutic targets of scar formation. **Methods** An observational research method was used. Datasets were downloaded from the National Center for Biotechnology Information Gene Expression Omnibus Database of America. The GSE136906 ($n=6$) and GSE137134 ($n=6$) datasets in the same batch were screened out for mRNA sequencing and methylation sequencing respectively, and the dataset GSE108110 ($n=9$) was incorporated into support vector machine and modeling analysis. The Limma software package was used to identify the differentially expressed genes and differentially methylated genes between scar tissue after burn and normal tissue. WGCNA was used to select the module with strong correlation with clinical features of scar tissue and large number of genes. Functional enrichment analysis of the genes in the module was performed to find genes with abnormal methylation. The receiver operating characteristic (ROC) curve was used to judge diagnostic efficacy of genes with abnormal methylation for scar, and support vector machine (SVM) was used to verify. **Results** A total of 10 modules were identified, and the brown module with large number of genes was highly correlated to burn scar tissue formation. The genes in the brown module were mainly concentrated in "regulation of androgen receptor signaling pathway", "cytokine-cytokine receptor

DOI: 10.3760/cma.j.cn501120-20200311-00150

本文引用格式: 郭鹏. 烧伤后瘢痕组织内甲基化基因的加权基因共表达网络分析[J]. 中华烧伤杂志, 2021, 37(12): 1185-1190. DOI: 10.3760/cma.j.cn501120-20200311-00150.

Guo P. Weighted gene co-expression network analysis of methylated genes in burn scar tissue[J]. Chin J Burns, 2021, 37(12): 1185-1190. DOI: 10.3760/cma.j.cn501120-20200311-00150.



interaction", "positive regulation of insulin secretion", and so on. The model showed 35 genes with abnormal methylation status. The ROC curve (area under the curve>0.9) and SVM modeling (accuracy=93.3%) indicated that *CCR2*, *LMO7*, *STEAP4*, *NNAT*, and *TCF7L2* genes had good diagnostic performance for scar. **Conclusions** *CCR2*, *LMO7*, *STEAP4*, *NNAT*, and *TCF7L2* can be used as potential targets for burn scar treatment.

【Key words】 Burns; Cicatrix; Methylation; Weighted gene co-expression network; *CCR2*; *LMO7*

烧伤是常见的损伤之一,我国每年有数十万人因为烧伤而致残或致畸。烧伤后病理性瘢痕组织的形成是烧伤患者康复过程中的重要并发症之一,不仅严重影响患者的外观,而且可引起器官及组织的功能障碍,严重影响患者生活质量及身心健康^[1]。对于瘢痕产生的机制,目前在微观方面的研究仅涉及细胞因子、ECM等,而其分子机制尚不十分明确^[2]。

有研究显示,青光眼滤过术后使用表观遗传修饰可减少瘢痕组织的产生^[3],表明瘢痕产生的过程中可能存在表观遗传改变。DNA甲基化是一种重要的表观遗传修饰机制,可在不改变DNA序列的情况下改变基因表达量^[4-5]。甲基化水平的高低会显著影响蛋白的最终翻译。生物信息学分析是一项新兴的研究方法,可结合测序数据进行全基因组分析,有针对性地寻找与疾病相关的基因或基因集^[6-7]。甲基化研究在瘢痕中较为常见,但检索显示以生物信息学为基础,探究甲基化与瘢痕形成关系的研究罕见,因此以生物信息学为基础探讨甲基化与瘢痕形成的关系显得十分必要。

加权基因共表达网络分析(weighted gene co-expression network analysis, WGCNA)可对大量芯片来源的基因数据进行有效分析,是生物信息学研究的重要工具^[8]。本观察性研究使用WGCNA对正常组织和瘢痕组织的mRNA表达谱进行分析并结合甲基化数据集对选定模块内的基因进行甲基化状态判断,对烧伤后瘢痕组织形成的分子机制进行探讨,并使用支持向量机(support vector machine, SVM)模型进行进一步验证,为临床治疗烧伤后瘢痕组织提供新思路。

1 资料及方法

1.1 数据收集

从美国国家生物技术信息中心基因表达综合数据库(<https://www.ncbi.nlm.nih.gov/geo/>)内检索并下载相关数据集,即检索基因表达数据库(gene expression omnibus, GEO)中烧伤后瘢痕组织相关数

据集,结果显示 mRNA 数据集 GSE136906 及甲基化数据集 GSE137134 是同一批次的 12 个样本(均由澳大利亚大学生物医学院烧伤研究组上传,每个数据集均为 6 个样本),分别进行了 mRNA 测序和甲基化测序,因此纳入本研究;数据集 GSE108110 在剩余数据集中样本量相对较大,因此纳入 SVM 及建模分析(表 1)。

表 1 烧伤后瘢痕组织相关数据集的详细信息

数据集	测序平台	Affymetrix eneChip 文件	样本数 (个)	瘢痕数 (个)
GSE137134	GPL13534-11288	1	6	6
GSE136906	GPL16686	2	6	6
GSE108110	GPL570	3	9	9

注:1、2、3 分别为 Illumina HumanMethylation450 BeadChip (HumanMethylation450_15017482)、[HuGene-2_0-st] Affymetrix Human Gene 2.0 ST Array [transcript (gene) version]、[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array

1.2 GEO 数据集的分析

下载 GSE137134 与 GSE136906 数据集并将数据集中的探针代码转换成基因符号。使用 R 语言中的“Limma”软件包寻找 GSE137134 中瘢痕组织与正常组织的差异甲基化基因(differentially methylated genes, DMG)和 GSE136906 中的差异表达基因(differently expressed genes, DEG),即寻找 $P < 0.05$ 相关基因并记录,将这些基因纳入 WGCNA 分析。

1.3 WGCNA 构建及临床相关性分析

使用 R 语言软件中的“FlashClust”软件包对样本进行聚类分析。使用“pickSoftThreshold”函数来调节参数 β 的权重以尽量符合无尺度网络。使用 WGCNA 软件包将有相关性和相邻关系的 DMG 计算成为拓扑重叠矩阵(topological overlap matrix, TOM)并计算其相应的相异度 1-TOM。使用 1-TOM 作为距离度量,进行分层聚类以识别模块。设定模块内最少的基因个数为 30 个,将高度相似的模块通过聚类标记合并。使用“plotDendroAndColors”函数对基因模块进行可视化,并选取模块内基因绘制热图。将临床特征和模块内的基因联合分析,寻找

与瘢痕组织密切相关的模块并探究模块内基因的生物学意义,对生成模块与临床特征之间的相关性热图进行聚类。

1.4 模块内基因功能富集分析

使用 Metascape 数据库和基因探针富集分析 (gene set enrichment analysis, GSEA) 软件对模块内的基因进行基因本体论 (gene ontology, GO) 分析和京都基因与基因组百科全书 (Kyoto Encyclopedia of Genes and Genomes, KEGG) 富集分析和注释。

1.5 模块内基因甲基化状态的探究

基于瘢痕组织 mRNA 数据集 GSE136906 及其甲基化数据集 GSE137134 属于同一批次的样本,本研究利用数据集 GSE136906 探究模块内基因的甲基化状态。使用 Funrich 软件将模块内高甲基化状态的基因与低表达的基因取交集,将模块内低甲基化状态的基因与高表达的基因取交集,在交集内的基因是具有异常甲基化状态的基因。

1.6 统计学处理

依据受试者操作特征 (ROC) 曲线下面积,评估基因的诊断价值。使用数据集 GSE136906 及 GSE108110 进行机器学习,将这 2 个数据集内异常甲基化基因的表达量依次纳入 SVM 模型并进行标准化,构建 SVM 分类器,使用交叉验证的方法选取惩罚参数 c 和核函数参数 g ,随机选取 2/3 的样本作为训练集,其余 1/3 则作为验证集,当 SVM 分类器的效能达到 90% 时则不再纳入新基因。

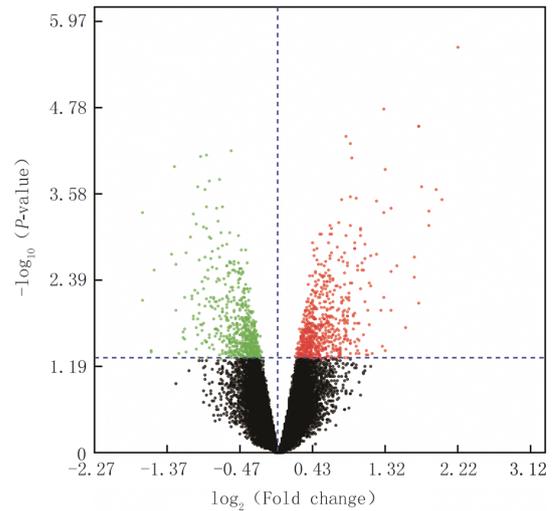
使用 SPSS 22.0、GraphPad Prism 8 统计软件进行数据处理及统计学分析,使用 Matlab 软件进行支持向量机模型构建, $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 WGCNA 网络构建及临床相关性分析

从数据集 GSE136906 中筛选出 1 061 个 DEG,见图 1。从数据集 GSE137134 中筛选出 16 141 个 DMG。聚类结果显示没有明显离群的样本,因此 2 个数据集中 12 个样本均被纳入共表达网络。“pickSoftThreshold”函数的结果显示,当权重参数 $\beta=9$ 时, $\log(k)$ 与 $\log [p(k)]$ 之间相关系数的平方 > 0.9 , k 表示节点的连接度, $p(k)$ 表示 WGCNA 网络具有无标度网络的特征。选取软阈值 $\beta=9$ 来构建 WGCNA 网络 (图 2)。模块的相关性热图与聚类分析展示在图 3 中。筛选后得到 10 个相应的模块,临床相关性结果显示棕色模块与瘢痕组织形成相关

性较高 (图 4), 该模块内共有 2 646 个基因。在所有相关系数 > 0.7 的模块中,棕色模块内基因数目最多,因此选择该模块进行进一步研究。



注: P -value 为差异表达分析后未经校正的 P 值, Fold change 为瘢痕组织与正常组织差异表达倍数;图中绿色为低表达且差异有统计学意义的基因,红色为高表达且差异有统计学意义的基因,黑色为差异无统计学意义的基因

图 1 GSE136906 数据集中烧伤后瘢痕和正常组织的差异表达基因火山图

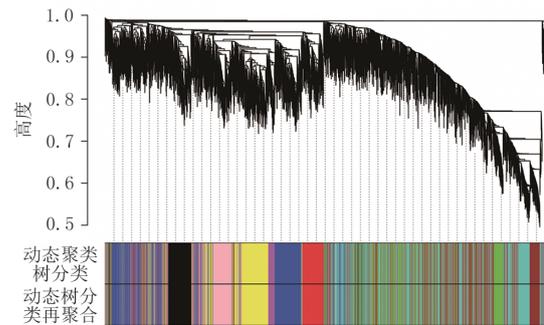
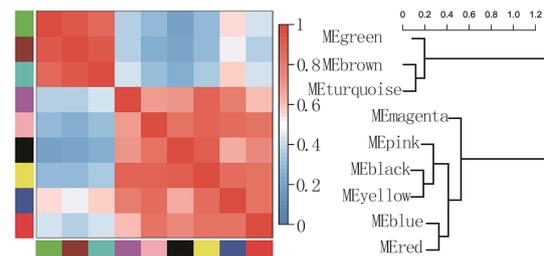
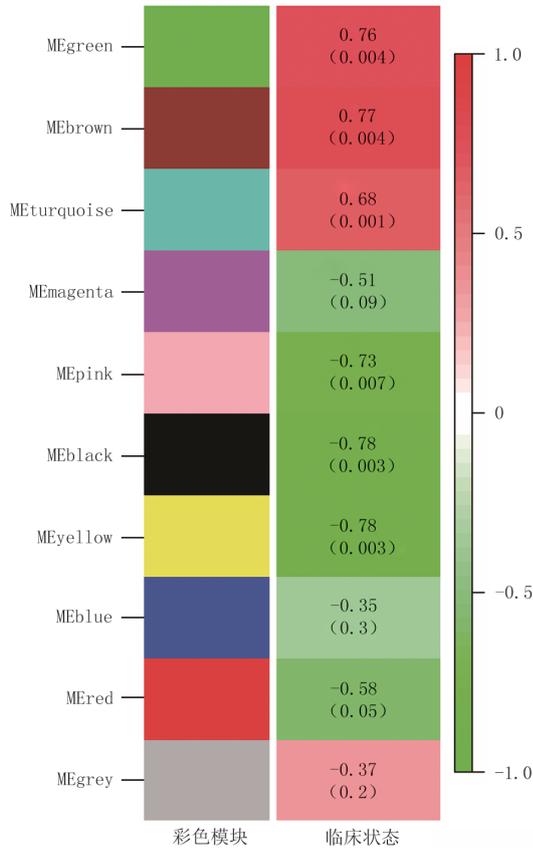


图 2 烧伤后瘢痕和正常组织的加权基因共表达网络分析网络模块



注: MEgreen、MEbrown、METurquoise、MEMagenta、MEpink、MEblack、MEyellow、MEblue、MEdred 分别表示绿、棕、蓝绿、洋红、粉、黑、黄、蓝、红色模块;图中横向刻度值为各模块聚类分析相对聚类距离;图中纵向刻度值为数据模块色值

图 3 加权基因共表达网络分析网络模块之间的相互作用

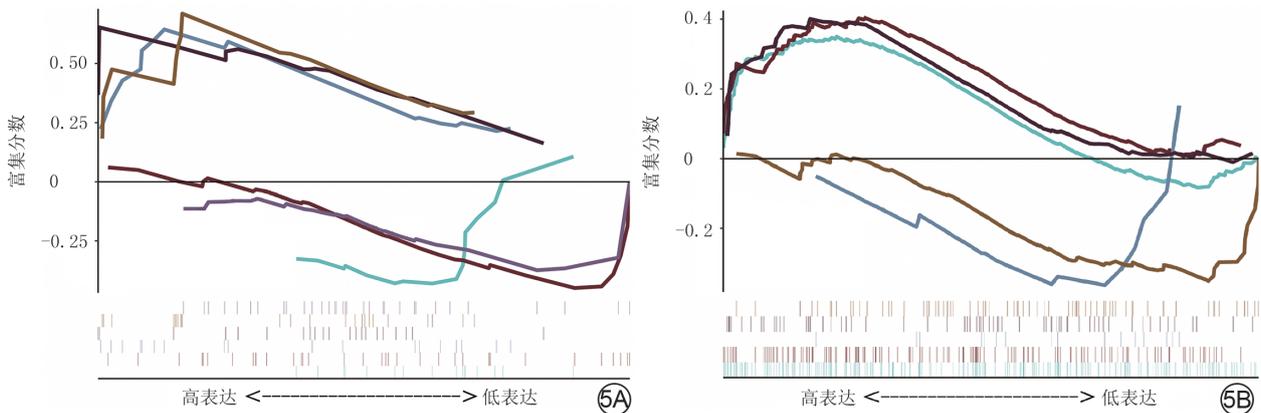


注:图左侧标注为加权基因共表达网络分析结果所示颜色模块名称;模块中数据表示对应加权基因共表达网络分析模块与临床性状的相关性系数以及其对应P值,括号内为P值

图4 烧伤后瘢痕组织临床特征与模块之间的关联

2.2 棕色模块中基因功能富集分析

GO分析结果显示,模块内的基因主要富集在“ α -氨基-3-羟基-5-甲基-4-异恶唑-丙酸酯选择性



注:图5A中的绿、棕、蓝、黑、浅棕、紫色曲线分别表示对核因子 κ B导入细胞核的负向调节、对白细胞介素1产生的正向调节、对 α -氨基-3-羟基-5-甲基-4-异恶唑-丙酸酯选择性谷氨酸受体的活性调控、雄激素受体信号通路调控、脂肪酸氧化的调控、泛酸样蛋白结合酶的调节通路,图5B中的绿、棕、蓝、黑、浅棕色曲线分别表示细胞因子-细胞因子受体的相互作用、Janus激酶/信号转导与转录激活子信号通路、泛酸和辅酶A生物合成、过氧化物酶体增殖物激活受体信号途径、转化生长因子 β 信号途径

图5 对烧伤后瘢痕组织相关性高的模块通过基因探针富集分析进行基因功能和通路分析。5A.基因本体论功能富集;5B.京都基因与基因组百科全书功能富集

谷氨酸受体活性的调控”“雄激素受体信号通路的调控”等,见图5A。KEGG通路富集分析结果显示,模块内的基因主要与“细胞因子-细胞因子受体相互作用”“过氧化物酶体增殖物激活受体信号通路”等相关,见图5B。Metascape分析结果显示,模块内基因主要富集在“急性粒细胞白血病”“半胱氨酸型内肽酶抑制剂的活性”“胰岛素分泌的正调节”等方面,见图6。富集分析结果表明,这些基因可通过细胞因子受体、胰岛素受体等调节瘢痕组织。

2.3 棕色模块内基因甲基化状态

共有35个具有异常甲基化状态的基因,其中的高表达-低甲基化基因21个(图7A、7B),低表达-高甲基化基因14个(图7C、7D)。

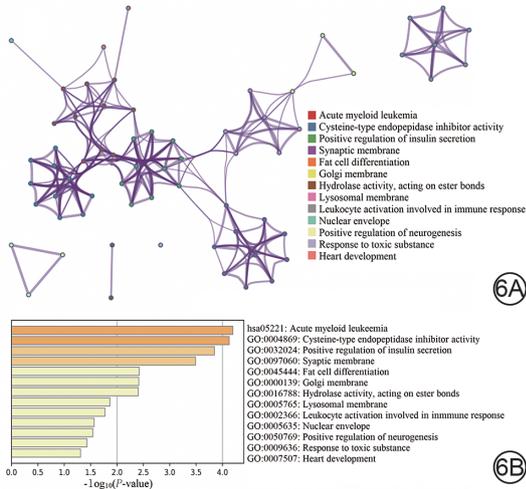
2.4 统计分析结果

ROC曲线分析结果显示,CCR2、LMO7、STEAP4、NNAT、TCF7L2、ZC3H12B等基因具有较好的瘢痕诊断效能,其曲线下面积(AUC)值居于前列。见表2。

将这些基因按照AUC值从大到小依次纳入SVM模型中,当纳入CCR2、LMO7、STEAP4、NNAT和TCF7L2这5个基因时,惩罚参数c和核函数参数g都 <0.01 ,此时构建的分类器模型分类准确率为93.3%,表明上述基因有较好的分类功能(图8)。

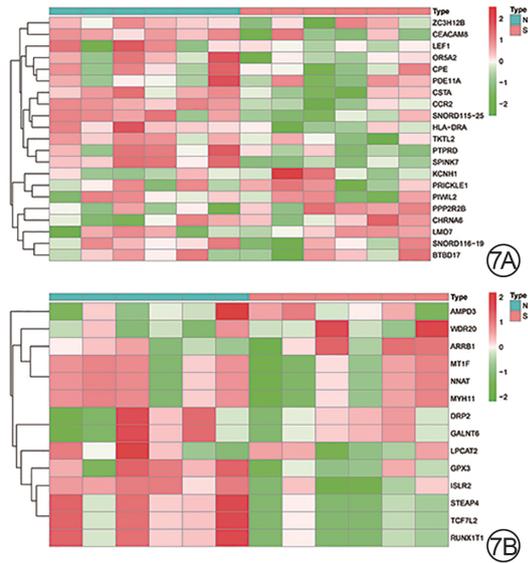
3 讨论

烧伤是临床上常见的疾病,其中相当一部分患者为深II度及以上烧伤。瘢痕是烧伤后康复期患者最常见的并发症之一,该疾病会严重影响患者的



注: P -value 为差异表达分析后未经校正的 P 值; 图 6A 的红、蓝、绿、紫、橙、黄、棕、浅紫、灰、浅绿、浅黄、浅紫、浅橙色圆点分别与图 6B 从上至下的条带一一对应, 依次表示急性粒细胞白血病、半胱氨酸型内肽酶抑制剂的活性、胰岛素分泌的正调节、突触膜、脂肪细胞分化、高尔基体膜、参与免疫反应的水解酶激活、溶酶体膜、参与免疫反应的白细胞激活、核包膜、神经发生正向调节、中毒反应、心脏发育通路

图 6 对烧伤后瘢痕组织相关性高的模块通过 Metascape 数据库进行基因功能和通路分析。6A. 基因本体论功能富集; 6B. 京都基因与基因组百科全书功能富集



注: “type” 表示组织类型, “N” 表示正常组织, “S” 表示瘢痕组织

图 7 烧伤后瘢痕组织差异表达甲基化基因的热图。7A. 21 个高表达-低甲基化基因的热图; 7B. 14 个低表达-高甲基化基因的热图

外貌和生活质量。因此, 探究烧伤后瘢痕组织的形成机制并寻找潜在的治疗靶点对改善烧伤患者的预后具有重要意义。本研究利用 WGCNA 分析寻找与烧伤后瘢痕形成相关的基因模块, 并探究了模块

表 2 烧伤后瘢痕组织与正常组织异常甲基化基因受试者操作特征曲线下面积及 P 值

基因	曲线下面积	P 值
<i>CCR2</i>	0.972	<0.001
<i>LMO7</i>	0.972	0.011
<i>STEAP4</i>	0.972	<0.001
<i>NNAT</i>	0.944	<0.001
<i>TCF7L2</i>	0.944	0.003
<i>ZC3H12B</i>	0.944	0.010
<i>AMPD3</i>	0.917	<0.001
<i>BTBD17</i>	0.917	0.039
<i>PRICKLE1</i>	0.917	<0.001

内的基因生物功能和甲基化状态, 该研究为治疗烧伤后瘢痕提供了有意义的临床参考价值。

根据富集分析的结果, 棕色模块内的基因具有“雄激素受体信号通路的调控”功能, 因此本研究认为雄激素受体在瘢痕组织生成过程中具有一定作用; 且由于 DNA 甲基化状态的改变, 该通路状态发生改变。而 Schierle 等^[9] 研究结果显示, 雄激素受体 DNA 的含量在瘢痕组织与正常组织之间具有显著差异; 同时指出抗雄激素物质可通过竞争结合雄激素受体导致瘢痕组织中 I 型前胶原 mRNA 低表达, 因此本研究认为抗雄激素物质具有治疗瘢痕的潜力。“细胞因子-细胞因子受体相互作用”在 KEGG 分析中被富集到, 研究表明多种细胞因子如血管源性生长因子、IL 家族等对瘢痕生成有重要影响^[10-13]。因此本研究认为模块内的基因可调节细胞因子受体, 与微环境内的细胞因子相互作用并诱导瘢痕组织的形成。Metascape 数据库的分析显示, “胰岛素分泌的正调节”被显著富集, 表明这些异常甲基化的基因可能导致胰岛素相关通路的改变。因此本研究认为胰岛素在瘢痕生成方面具有重要作用, Hallam 等^[14] 进行的一项随机对照实验显示, 皮下注射胰岛素可减少瘢痕产生。因此, 本研究认为胰岛素注射也可作为瘢痕治疗的潜在疗法。

在棕色模块内共显示了 35 个异常甲基化基因, 其中 *CCR2*、*LMO7*、*STEAP4*、*NNAT* 和 *TCF7L2* 具有良好的瘢痕诊断效能并被纳入到 SVM 模型中。*CCR2* 编码单核细胞趋化蛋白-1, Frik 等^[15] 研究显示 *CCR2* 敲除小鼠脑损伤后单核细胞浸润减少且瘢痕形成减少。本研究中观察到瘢痕组织中 *CCR2* 是高表达的, 推测 *CCR2* 可诱导瘢痕产生, 是瘢痕治疗的潜在靶点。本研究结果显示 *LMO7* 在瘢痕组织中高表达, 推测 *LMO7* 可促进瘢痕组织产生。Xie 等^[16] 研

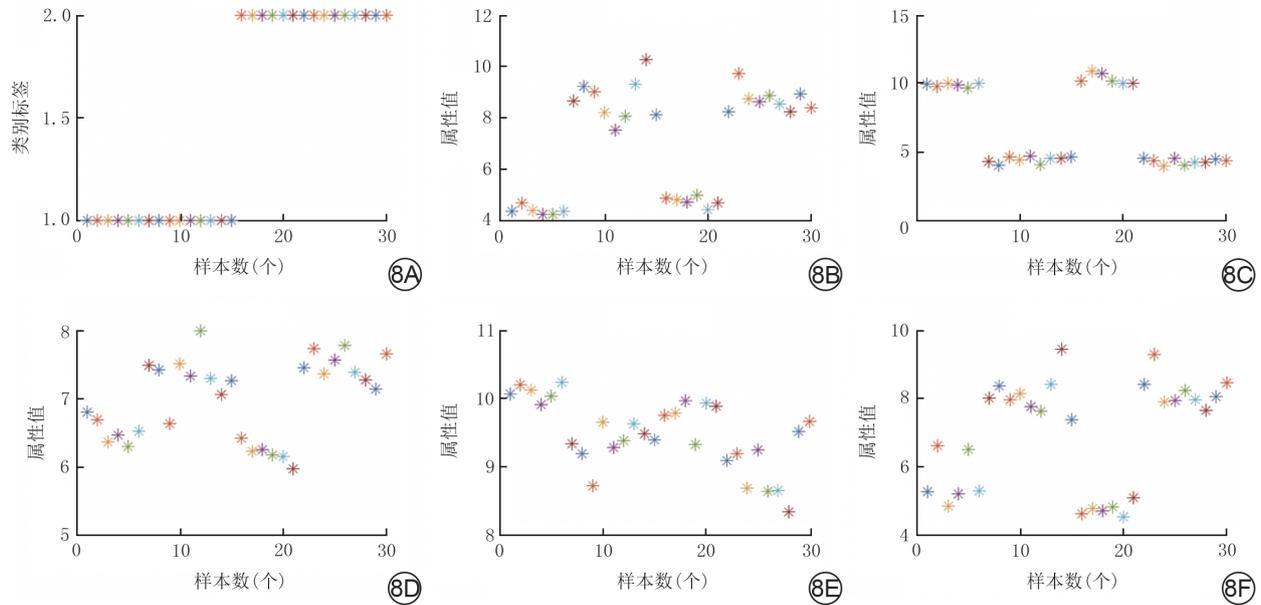


图8 烧伤后瘢痕组织差异表达甲基化基因的数据展示及5个差异甲基化基因可视图。8A、8B、8C、8D、8E、8F.分别为数据的整体展示及 *TCF7L2*、*NNAT*、*LMO7*、*CCR2*和 *STEAP4*基因的分维可视化图

究表明 *LMO7* 由 TGF- β 诱导并通过 TGF- β 途径负反馈调节在伤口愈合和瘢痕生成中发挥重要作用,这与本研究结果一致。*STEAP4*、*NNAT*和 *TCF7L2* 则未见相应报道,其在瘢痕产生中的作用可能需要进一步实验研究进行验证。

总之,本研究通过 WGCNA 初步分析了烧伤后瘢痕产生的分子机制和异常甲基化基因,结果显示基因 *CCR2*、*LMO7*、*STEAP4*、*NNAT*和 *TCF7L2* 在瘢痕形成中具有重要作用,可作为治疗的潜在靶点。

利益冲突 作者声明不存在利益冲突

参考文献

- [1] Liu HF, Zhang F, Lineaweaver WC. History and advancement of burn treatments[J]. *Ann Plast Surg*, 2017, 78(2 Suppl 1):S2-8. DOI: 10.1097/SAP.0000000000000896.
- [2] Li H, Yao Z, Tan J, et al. Epidemiology and outcome analysis of 6325 burn patients: a five-year retrospective study in a major burn center in Southwest China[J]. *Sci Rep*, 2017, 7:46066. DOI: 10.1038/srep46066.
- [3] Sharma A, Anumanthan G, Reyes M, et al. Epigenetic modification prevents excessive wound healing and scar formation after glaucoma filtration surgery[J]. *Invest Ophthalmol Vis Sci*, 2016, 57(7):3381-3389. DOI:10.1167/iops.15-18750.
- [4] Moore LD, Le T, Fan G. DNA methylation and its basic function[J]. *Neuropsychopharmacology*, 2013, 38(1): 23-38. DOI: 10.1038/npp.2012.112.
- [5] Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing[J]. *Nat Rev Genet*, 2018, 19(6): 371-384. DOI:10.1038/s41576-018-0004-3.
- [6] Author unknown. A bioinformatics workshop in a box[J]. *Nature*, 2018, 554(7690):134. DOI:10.1038/d41586-018-01424-4.
- [7] Altman R. Current progress in bioinformatics 2016[J]. *Brief Bioinform*, 2016, 17(1):1. DOI:10.1093/bib/bbv105.
- [8] Pei G, Chen L, Zhang W. WGCNA Application to proteomic and metabolomic data analysis[J]. *Methods Enzymol*, 2017, 585: 135-158. DOI:10.1016/bs.mie.2016.09.016.
- [9] Schierle HP, Scholz D, Lemperle G. Elevated levels of testosterone receptors in keloid tissue: an experimental investigation[J]. *Plast Reconstr Surg*, 1997, 100(2): 390-395; discussion 396. DOI: 10.1097/00006534-199708000-00017.
- [10] Liu W, Wang DR, Cao YL. TGF-beta: a fibrotic factor in wound scarring and a potential target for anti-scarring gene therapy[J]. *Curr Gene Ther*, 2004, 4(1):123-136. DOI:10.2174/1566523044578004.
- [11] Lian N, Li T. Growth factor pathways in hypertrophic scars: molecular pathogenesis and therapeutic implications[J]. *Biomed Pharmacother*, 2016, 84:42-50. DOI:10.1016/j.biopha.2016.09.010.
- [12] Nguyen JK, Austin E, Huang A, et al. The IL-4/IL-13 axis in skin fibrosis and scarring: mechanistic concepts and therapeutic targets[J]. *Arch Dermatol Res*, 2020, 312(2): 81-92. DOI: 10.1007/s00403-019-01972-3.
- [13] Wynn TA. Cellular and molecular mechanisms of fibrosis[J]. *J Pathol*, 2008, 214(2):199-210. DOI:10.1002/path.2277.
- [14] Hallam MJ, Pitt E, Thomas A, et al. Low-dose insulin as an antiscarring therapy in breast surgery: a randomized controlled trial[J]. *Plast Reconstr Surg*, 2018, 141(4): 476e-485e. DOI: 10.1097/PRS.0000000000004199.
- [15] Friik J, Merl-Pham J, Plesnila N, et al. Cross-talk between monocyte invasion and astrocyte proliferation regulates scarring in brain injury[J]. *EMBO Rep*, 2018, 19(5):e45294. DOI:10.15252/embr.201745294.
- [16] Xie Y, Ostriker AC, Jin Y, et al. LMO7 is a negative feedback regulator of transforming growth factor β signaling and fibrosis[J]. *Circulation*, 2019, 139(5): 679-693. DOI: 10.1161/CIRCULATION-AHA.118.034615.

(收稿日期:2020-03-11)