



整合生物信息学和机器学习算法筛选心房颤动 炎症标志物与验证*

卜琪^①, 张伟, 黄英, 李铮, 王杨, 赵璇, 贾云莉,
樊晓燕, 杨媛, 王志敏^{①△}

昆明医科大学附属延安医院(昆明 650051)

【摘要】目的 通过整合生物信息学和机器学习算法及实验研究, 筛选并鉴定心房颤动(AF)的潜在生物标志物。**方法** 从GEO和MSigDB数据库中获取AF转录组学并筛选差异表达基因(DEGs), 从炎症相关基因集(IRGs)。取交集形成DE-IRGs基因集。采用机器学习算法过滤AF相关DE-IRGs基因后获得特征基因。取12只SPF级雄性SD大鼠, 分为对照组和AF组。其中AF组大鼠连续5周尾静脉注射乙酰胆碱(66 μ g/mL)和氯化钙(10 mg/kg)构建AF模型; HE染色观察心房肌细胞形态特征、RT-qPCR和免疫组织化学染色测定特征基因表达mRNA和蛋白质变化。**结果** 训练集发现119个DEGs, IRGs与两个共表达模块相关性最高($r=0.6$ 和 0.56 , $P<0.0001$)。DEGs与IRGs取交集获得9个DE-IRGs; SVM-RFE和RF算法筛选出的4个特征基因5羟色胺受体2B基因(5-hydroxytryptamine receptor 2B gene, *HTR2B*)、潜在转化生长因子 β 结合蛋白2基因(latent-transforming growth factor beta-binding protein 2 gene, *LTBP2*)、基质重塑关联蛋白5基因(matrix remodeling associated protein 5 gene, *MXRA5*)和转化生长因子 β 诱导蛋白基因(transforming growth factor β -induced protein gene, *TGFBI*)在训练集和验证集AF组均呈高表达且诊断效能较高($AUC>0.77$)。AF大鼠心电图肢体导联出现大量AF特征性P波。HE染色显示AF组心房肌细胞排列紊乱、炎性细胞浸润。TUNEL荧光检测AF组凋亡率为(55.34 $\pm$ 4.29)%, 而对照组为(8.69 $\pm$ 3.12)%($P=0.0001$)。*LTBP2*和*TGFBI*基因相对表达量在AF组(4.97 $\pm$ 4.20, 2.62 $\pm$ 1.85)和对照组(1.12 $\pm$ 0.21, 1.18 $\pm$ 0.77)间差异存在统计学意义($P=0.0137$, $P=0.0444$)。免疫组织化学染色检测AF组*LTBP2*和*TGFBI*蛋白阳性率分别为(36.50 $\pm$ 1.31)%和(27.39 $\pm$ 4.57)%, 而对照组分别是(22.95 $\pm$ 2.62)%和(18.26 $\pm$ 3.70)%($P=0.0008$, $P=0.0485$)。**结论** 特征基因*LTBP2*和*TGFBI*在AF中高表达, 且较有诊断价值的生物标志物。

【关键词】 心房颤动 机器学习算法 生物标志物 炎症反应

Integrating Bioinformatics and Machine Learning Algorithms to Screen Inflammatory Biomarkers for Atrial Fibrillation and Experimental Validation

BU Qi^①, ZHANG Wei, HUANG Ying, LI Zheng, WANG Yang, ZHAO Xuan, JIA Yunli, FAN Xiaoyan,
YANG Yuan, WANG Zhimin^{①△}. Yan'an Hospital Affiliated to Kunming Medical University, Kunming 650051,
China

△ Corresponding author, E-mail: 2803946798@qq.com

This study was supported by the Health Research Project of Kunming Municipal Health Commission (No. 2023-03-01-007).

【Abstract】 Objective To identify potential biomarkers of atrial fibrillation (AF) using bioinformatics, machine learning (ML) and experimental methods. **Methods** AF transcriptomic data were obtained from the GEO and MSigDB databases to screen for differentially expressed genes (DEGs) and inflammation-related gene sets (IRGs). The overlap between DEGs and IRGs was used to identify the DE-IRGs set. Two machine learning algorithms were used to filter AF-related DE-IRGs. Twelve male SD rats were randomly divided into a control group and an AF group. Rats received acetylcholine (66 μ g/mL) and calcium chloride (10 mg/kg) via tail vein injection for five consecutive weeks to establish the AF model. Morphological characteristics of atrial myocytes were assessed with HE staining, while RT-qPCR and immunohistochemistry (IHC) were used to measure changes in mRNA and protein levels. **Results** In the training set, 119 DEGs were identified, with IRGs showing the highest correlation with two co-expression modules ($r = 0.60$ and 0.56 , $P < 0.0001$). The intersection of DEGs and IRGs yielded nine DE-IRGs. The SVM-RFE and RF algorithms identified 5-hydroxytryptamine receptor 2B gene (*HTR2B*), latent-transforming growth factor beta-binding protein 2 gene (*LTBP2*), matrix remodeling associated protein 5 gene (*MXRA5*), and transforming growth factor β -induced protein gene (*TGFBI*) as highly expressed in both the training and test sets in AF groups with high diagnostic efficacy ($AUC > 0.77$). The

* 昆明市卫生健康委员会卫生科研课题项目 (No. 2023-03-01-007)资助

△ 通信作者, E-mail: 2803946798@qq.com

出版日期: 2026-05-20

electrocardiogram limb leads of AF rats showed numerous f-waves. HE staining revealed disorganized atrial myocyte arrangement and inflammatory cell infiltration in the AF group. The TUNEL fluorescence assay showed an apoptosis rate of $(55.34 \pm 4.29)\%$ in the AF group, compared to $(8.69 \pm 3.12)\%$ in the control group ($P = 0.0001$). There were statistically significant differences in the relative expression levels of *LTBP2* and *TGFBI* between the AF group (4.97 ± 4.20 , 2.62 ± 1.85) and the control group (1.12 ± 0.21 , 1.18 ± 0.77) ($P = 0.0137$, $P = 0.0444$). IHC revealed positive rates of *LTBP2* and *TGFBI* proteins in the AF group of $(36.50 \pm 1.31)\%$ and $(27.39 \pm 4.57)\%$, respectively, compared to $(22.95 \pm 2.62)\%$ and $(18.26 \pm 3.70)\%$ in the control group ($P = 0.0008$, $P = 0.0485$). **Conclusion** The characteristic genes *LTBP2* and *TGFBI* are highly expressed in AF and serve as valuable diagnostic biomarkers.

[Key words] Atrial fibrillation Machine learning algorithms Biomarkers Inflammatory response

心房颤动(atrial fibrillation, AF)是一种常见心律失常类型,严重影响患者生活质量,增加心肌缺血、心力衰竭、心源性猝死等心血管不良事件风险^[1]。约93%AF发作无症状,常规检测手段敏感性和特异性有限^[2-4]。AF还会被其他心脏疾病掩盖,如心力衰竭患者的N末端B型利钠肽原(NT-pro BNP)水平的升高可能会导致AF被忽视^[5]。消融术后患者心房瘢痕改变了预期的P波形态,会增加心电图识别难度^[6]。因此,有必要深入探究AF发病机制,筛选有效便捷的新生物标志物提高诊断率和治疗率。

炎症反应在AF的发生和维持中发挥关键作用,参与心房结构重塑与心脏电生理改变^[7]。C反应蛋白、纤维蛋白原、可溶性细胞间黏附分子-1以及心房肌细胞炎性小体在AF的发病中具有重要作用^[7]。血管紧张素Ⅱ能上调IL-6、TNF- α 等表达,其与JAK/STAT3和NF- κ B等信号通路的激活有关^[8-9]。CD4⁺ T细胞和巨噬细胞浸润与AF易感性增加密切相关^[10]。炎症因子与心房肌细胞之间的相互作用能导致电学特性变化,引起心律失常发生^[11]。炎症细胞因子影响心房肌细胞内离子通道功能,导致电传导紊乱和不期延长,损害心肌细胞收缩的同步性,形成促AF发展微环境^[12]。深入了解炎症机制有助于发现新的AF诊断标志物。

本研究基于GEO数据库AF患者和窦性心律志愿者的样本,联合生物信息学技术和机器学习算法筛选AF相关炎性标志物,并通过大鼠AF模型加以验证,旨在为AF的临床诊断和治疗提供新的理论依据和治疗靶点。

1 材料与方法

1.1 实验材料

1.1.1 实验动物

SPF级雄性SD大鼠,共12只,周龄8周,体质量250 g/只,购自SBF(北京)公司(许可证号SCXK-京2019-0010),实验在云南贝斯泰生物科技有限公司动物实验平台进行,并通过该公司实验动物伦理委员会批准(BST-PZ-RAT-20240815-01),遵循NIH实验动物使用指南。

1.1.2 主要试剂

苏木精-伊红染液和TUNEL细胞凋亡检测试剂盒购于Servicebio; SweScript RT II 第一链cDNA合成试剂盒(含gDNA去除剂)购于Servicebio; 抗TGFBI抗体和抗LTBP2抗体购自于Abcam; Alexa Fluor 488标记的山羊抗兔IgG二抗购于Abcam; 水合氯醛溶液购于生工; 乙酰胆碱(Ach, 66 μ g/mL)和氯化钙(CaCl_2 , 10 mg/kg); 2 $\times$ T5 Fast qPCR Mix(SYBR Green I)试剂购于擎科生物。

1.1.3 主要设备

生理数据采集设备Powerlab 15T购自ADInstruments; 组织研磨仪三维冷冻型SWE-3D购于Servicebio; NanoPhotometer N50分光光度计购自于Implen公司; HistoCore BIOCUT石蜡切片机购于莱卡公司; ECLIPSE Ts2R荧光显微镜购于尼康; CFX96TM实时荧光定量PCR仪购于Bio-rad公司。

1.2 实验方法

1.2.1 数据获取

从公共数据库GEO下载AF相关数据集GSE115574和GSE41177。GSE115574含28例AF患者和31例窦性心律(sinoatrial rhythm, SR)对照心房组织样本; GSE41177数据集共38例样本: AF患者32例和SR对照者6例。此外,从分子特征数据库MSigDB获取炎症相关基因集(Inflammation Related Genes, IRGs),共200个基因。

1.2.2 生物信息学分析

使用R语言limma软件包对GSE115574数据集的AF组与对照组进行差异表达基因(differentially expressed genes, DEGs)分析,筛选标准设定为 $|\log_2\text{FC}| > 0.5$ 、 $P < 0.05$ 且 $Q < 0.05$ 。IRGs构建使用加权基因共表达网络分析(WGCNA, Weighted correlation network analysis)^[13]。基于200个IRGs采用GSVA包计算GSE115574集的单样本基因集富集分析(single-sample gene set enrichment analysis, ssGSEA)评分,并分析其与基因标志物的相关性,筛选标准为相关系数绝对值大于0.3且P值小于0.05,再调用stats包“Hclust”函数进行分层聚类分析以排除异常值样本^[14]。此

后, 据无标度网络的标准(scale-free networks)signed R^2 大于 0.85, 用WGCNA包的“pickSoftThreshold”函数选择合适的软阈值 β , 再采用Pearson分析计算出的软阈值 β 对应基因相关邻接矩阵后, 转换为拓扑重叠矩阵, 而后构建分层聚类树状图, 将相似的基因表达分为不同的共表达模块^[14]。最后, 通过Pearson分析计算各个模块与ssGSEA评分之间的相关性。依据相关系数 $|\text{cor}| > 0.5$ 且 $P < 0.05$ 筛选关键模块, 并提取关键模块中的全部基因作为IRGs进行后续分析。

1.2.3 机器学习算法筛选特征基因

采用支持向量机递归特征消除(support vector machine-recursive feature elimination, SVM-RFE)和随机森林(random forest, RF)算法对DEGs与IRGs取交集获得DE-IRGs行最佳特征基因提取并取交集, 作为“AF相关特征基因”。GSE115574为训练集构建机器学习算法模型, GSE41177作为验证集检测模型效能并绘制受试者工作特征(receiver operating characteristic, ROC)曲线, 提取特征基因。在GSE115577和GSE41170集中, 同时满足以下3个条件的特征基因定义为AF相关特征基因: (1)基因表达均存在明显差异; (2)基因表达趋势在不同数据集中一致; (3)ROC曲线下面积(area under the curve, AUC)值大于0.7。

1.2.4 基因富集分析和KEGG通路分析

将GSE115574数据集依照特征基因分为高表达和低表达组, limma包分析组间的差异表达基因并按 $\log_2$ FC值降序排列生成基因列表, 利用clusterProfiler软件包, 以MSigDB数据库中KEGG基因集(C2类别)为参考进行GSEA, 且对 $Q < 0.05$ 的显著富集通路并进行可视化^[15]。最后通过Spearman相关性分析评估特征基因与KEGG通路关键基因的相关性。

1.2.5 大鼠AF模型构建

将SD大鼠分为两组: 对照组和AF组(每组 $n=6$)。对照组SD大鼠均为窦性心律、无心电图异常。AF组SD大鼠连续5周行尾静脉注射ACh(66 $\mu\text{g/mL}$)和 CaCl_2 (10 mg/kg)混合溶液(每100 g体重注射0.1 mL), 然后麻醉后仰卧位固定, 行心电图检测^[16]。AF模型构建成功判断标准需同时满足以下条件: (1)形态、振幅和频率不一致的P波出现, (2)P波消失, (3)R-R间期绝对不齐, (4)QRS波变异。建模成功后采用戊巴比妥钠联合氯胺酮麻醉处死两组大鼠并收集心脏组织用于后续实验。

1.2.6 组织学染色

体积分数为4%多聚甲醛固定心脏组织24 h后用PBS冲洗, 再采用石蜡包埋后制备成厚度为4 μm 切片, 经二甲苯脱蜡、梯度乙醇水化, 行HE染色, 脱水透化后中性树胶封片, 显微镜观察并采集5个非重叠视野图像。免疫

组织化学染色切片制备同HE染色, 后将切片依次置于二甲苯和梯度乙醇中脱蜡水化, 再移入10 mmol/L EDTA($\text{pH}=8.0$)溶液中, 在微波中火下修复8 min, 停火7 min, 重复2次后置室温冷却30 min; 再在3% H_2O_2 溶液下室温处理10 min, PBS漂洗后10%FBS封片处理10 min; 漂洗后分别滴加TGFBI抗体(1:100)及LTBP2抗体(1:100)4 $^\circ\text{C}$ 孵育24 h; 加入HRP标记二抗和显色剂; 最后用苏木精复染细胞核, 经脱水处理后封片完成染色。显微镜下观察记录5个非重叠视野, 计算阳性细胞百分比。

1.2.7 TUNEL染色

操作参照TMR(红色)标记的TUNEL细胞凋亡检测试剂盒说明书进行。Alexa Fluor 488标记山羊抗兔二抗(1:2000)室温避光孵育1 h, PBS漂洗3次, 每次15 min, 荧光显微镜采集图像并分析。

1.2.8 RT-qPCR实验

收集AF模型组和对照组心房组织4 $^\circ\text{C}$ 下研磨, TRIzol试剂提取组织总RNA, 测定RNA浓度。以10 ng RNA为模板, 在55 $^\circ\text{C}$ 条件下将RNA逆转录成cDNA。随后使用2 $\times$ T5 Fast qPCR Mix (SYBR Green I)试剂检测mRNA表达, 反应条件如下: 95 $^\circ\text{C}$ 预变性60 s, 然后以95 $^\circ\text{C}$ 持续10 s, 不同退火温度持续5 s, 72 $^\circ\text{C}$ 持续15 s为一个循环, 重复40个循环。LTBP2、TGFBI、HTR2B、MXRA5和GAPDH退火温度分别为60 $^\circ\text{C}$ 、65 $^\circ\text{C}$ 、60 $^\circ\text{C}$ 、55 $^\circ\text{C}$ 和60 $^\circ\text{C}$ 。以GAPDH作为内参基因, 采用 $2^{-\Delta\Delta\text{Ct}}$ 法对比AF组与对照组间基因表达差异。引物序列见表1。

表 1 PCR检测基因相对表达量引物
Table 1 PCR primers for genes expression

Gene name	Primer	Primer sequences from 5' to 3'
LTBP2	Forward	AGCACCAACCACTGTATCAAAC
	Reverse	CTCATCGGGAATGACCTCCTC
TGFBI	Forward	TACAGCACGGTATGCAAGCC
	Reverse	GCAACCGATCTAGCTCACAGAG
HTR2B	Forward	TGATTTGCTGGTTGGATTGTTT
	Reverse	ATGGATGCGGTTGAAAAGAGAA
MXRA5	Forward	CATCCCCGATGGAGCTTTAAG
	Reverse	TCAATGTGCAGCCTCATTAAAGTT
GAPDH	Forward	GGAGCGAGATCCCTCCAAAAT
	Reverse	GGCTGTTGTCATACTTCTCATGG

LTBP2: latent-transforming growth factor beta-binding protein 2 gene; TGFBI: transforming growth factor β -induced protein gene; HTR2B: 5-hydroxytryptamine receptor 2B gene; MXRA5: matrix remodeling associated protein 5 gene.

1.2.9 统计学方法

使用R(4.1.3)和GraphPadPrism 9.0进行统计学分析。计量资料以均值 $\pm$ 标准误进行统计描述; 符合正态分

布且方差齐的, 两组间比较采用Student's T检验; 若不符合正态分布或方差不齐, 则采用Mann-Whitney U检验。单次假设检验 $P < 0.05$ 为差异具有统计学意义, 多重检验假设 P 校正采用Benjaminiand-Hochberg方法进行, $Q < 0.05$ 为差异有统计学意义。

2 结果

2.1 获取AF的DE-IRGs

GSE115574数据集共鉴定出119个DEGs, 对比SR组, AF组基因表达上调52个基因, 下调67个(附图S1A、1B)。ssGSEA评分前后的分层聚类分析显示, AF组与对照组样本具有显著区分度, 故未剔除任何样本。WGCNA软阈值为7(signed $R^2=0.85$)时, 网络呈现无尺度拓扑特征且邻接函数均值趋近于0, 见附图S1C, 最终识别出7个共表达模块(不含灰色模块)(附图S1D)。相关性分析筛选出与ssGSEA评分显著相关的黑色模块(含153个基因)和红色模块(含703个基因), 其相关系数和 P 值分别为 $R=0.6$, $P=6.0 \times 10^{-7}$ 和 $R=0.56$, $P=5.0 \times 10^{-6}$ (附图S1E)。856个模块基因与DEGs取交集, 最终获得9个DE-IRGs, 分别是*HTR2B*、*LTBP2*、*MXRA5*、*TGFBI*、*LOXL1*、*FMOD*、*COL1A2*、*COL3A1*和*CFD*(附图S1F)。附图见网络资源附件。

2.2 机器学习算法筛选AF特征基因

SVM-RFE算法从9个DE-IRGs筛选出5个特征基因(*HTR2B*、*TGFBI*、*MXRA5*、*LTBP2*和*LOXL1*)(图1A); RF模型依据基尼系数从9个DE-IRGs选出5个特征基因(*HTR2B*、*TGFBI*、*CFD*、*LTBP2*和*MXRA5*)(图1B)。取交集后确定4个与AF相关的特征基因(*HTR2B*、*LTBP2*、*MXRA5*和*TGFBI*)(图1C)。4个基因在区分训练集AF与SR组时判别效能较好, ROC的AUC均大于0.7, 且在AF组的表达水平高于SR组($P < 0.01$), 见图2。此外, 验证集

4个基因诊断效能较高, *HTR2B*、*LTBP2*和*TGFBI*的AUC均高于0.9, *MXRA5*则为0.771(图2)。而且该数据集中, 对比SR组, AF组的4个特征基因表达水平较高($P < 0.001$ 和 $P < 0.05$), 见图2。

2.3 特征基因的生物学功能

4个特征基因GSEA结果显示, *HTR2B*、*LTBP2*和*MXRA5*在细胞外基质受体相互作用(Extracellular Matrix Receptor Interaction)通路中富集, 而*TGFBI*在同种异体排斥(Allografts Rejection)通路中呈现显著的富集(图3A)。*HTR2B*、*LTBP2*和*MXRA5*与细胞外基质受体相互作用通路包含的基因呈正相关, 相关系数分别是0.52、0.63和0.65($P=2.72 \times 10^{-5}$ 、 8.58×10^{-8} 和 2.15×10^{-8}), *TGFBI*与同种异体排斥通路包含的基因也呈正相关($r=0.61$, $P=2.21 \times 10^{-7}$), 见图3B。

2.4 大鼠AF模型特征和组织病理学改变

与对照组SD大鼠心电图对比, AF组I、II、III导联显示P波消失, f波大量出现, RR间期绝对不齐, QRS波的形态和振幅不一致, 见图4A。HE染色显示, 对照组心房肌细胞形态规则、排列紧密规整、胞质均匀, 炎性细胞浸润少、无组织水肿; AF组心房肌细胞排列紊乱、细胞萎缩变形、核深染或空泡化、胞浆疏松, 炎性细胞浸润明显增多(图4B)。TUNEL荧光检测提示, 对照组凋亡率为(8.69±3.12)%, AF组为(55.34±4.29)%, 组间差异有统计学意义($P=0.0001$), 见图4C和4D。

2.5 生物标志物的表达验证

AF组*LTBP2*基因相对表达量为 4.97 ± 4.20 , 对照组较低为 1.12 ± 0.21 ($P=0.0137$)。AF组和对照组*TGFBI*基因相对表达量分别为 2.62 ± 1.85 和 1.18 ± 0.77 , 且差异具有统计学意义($P=0.0444$)。然而, *HTR2B*和*MXRA5*基因相对表达量, AF组低于对照组, P 值分别为0.0154和0.0034

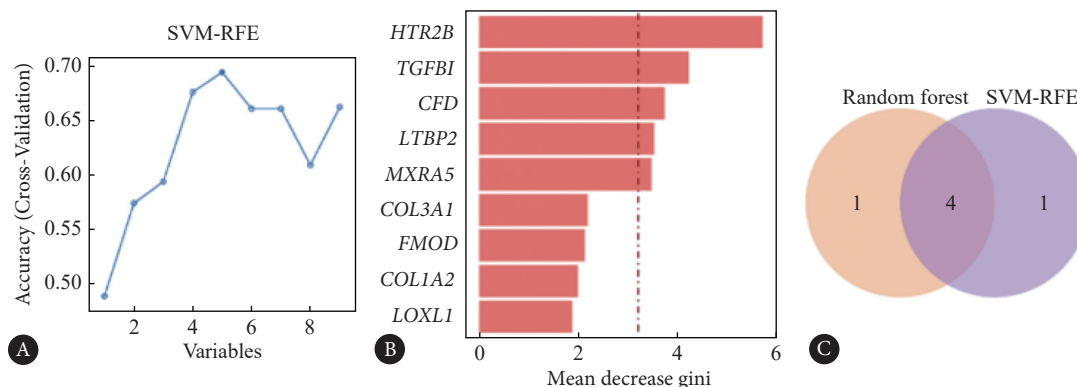


图1 机器学习算法筛选特征基因

Fig 1 Feature genes screen by machine learning

A, Feature gene selection using the SVM-RFE algorithm. B, Feature gene selection using the RF algorithm. C, Venn diagram showing the intersection of feature genes selected by SVM-RFE and RF algorithms.

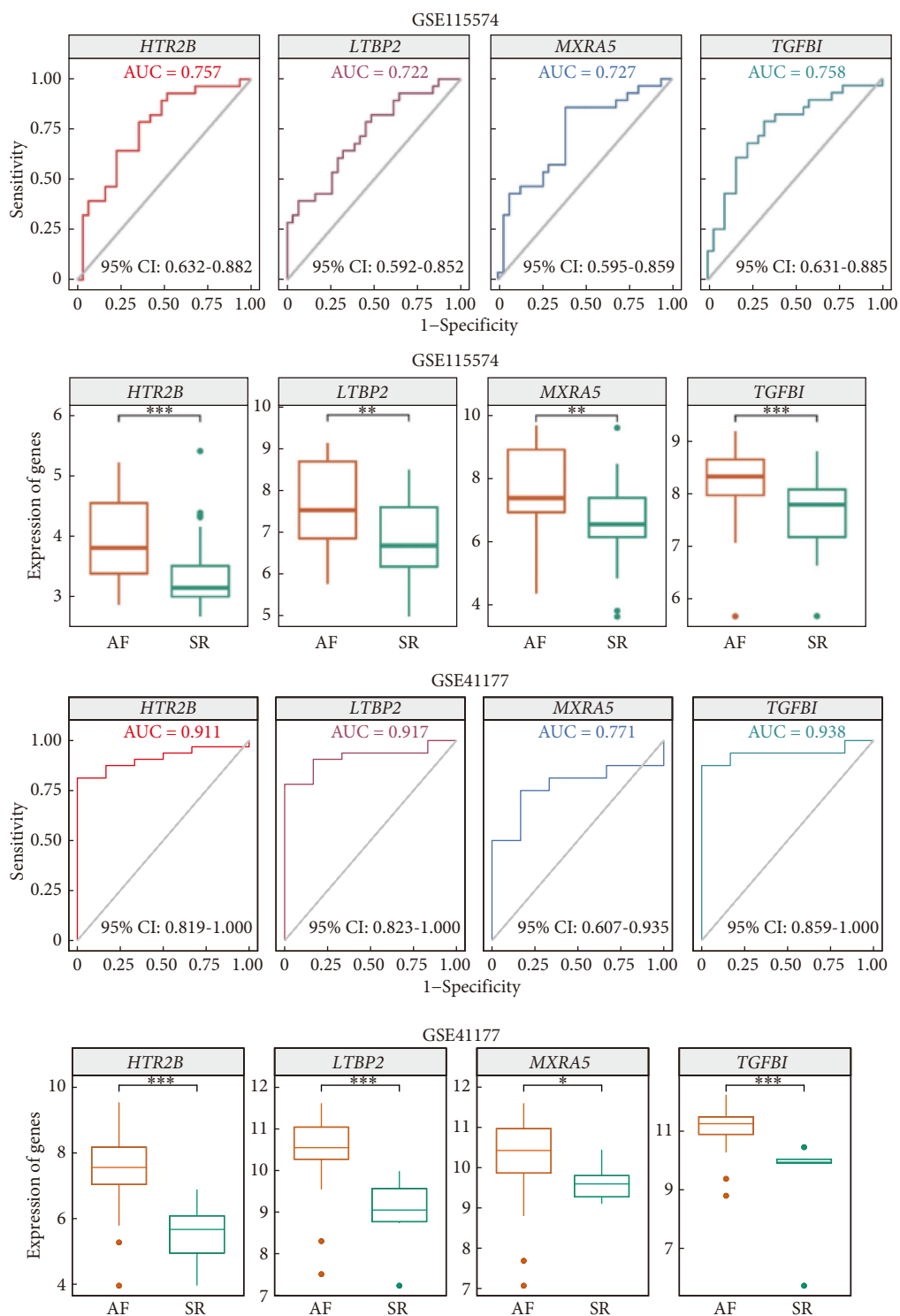


图 2 4个特征基因在训练集和验证集中的诊断效能

Fig 2 Diagnostic performance of the four feature genes in the training and validation sets

AF: atrial fibrillation; SR: sinoatrial rhythm. GSE115574: AF group $n = 28$, SR group $n = 31$; GSE41177: AF group $n = 32$, SR group $n = 6$. *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$.

(图5A)。免疫组织化学染色与RT-qPCR检测结果一致,与对照组[(22.95±2.62)%]相比,AF组中LTBP2蛋白阳性细胞率达到(36.50±1.31)%($P=0.0008$);AF组和对照组TGFBI蛋白阳性细胞率分别为(27.39±4.57)%和

(18.26±3.70)%($P=0.0485$)(图5B和5C)。

3 讨论

AF是一种复杂的心律失常,其特征是心房内出现混

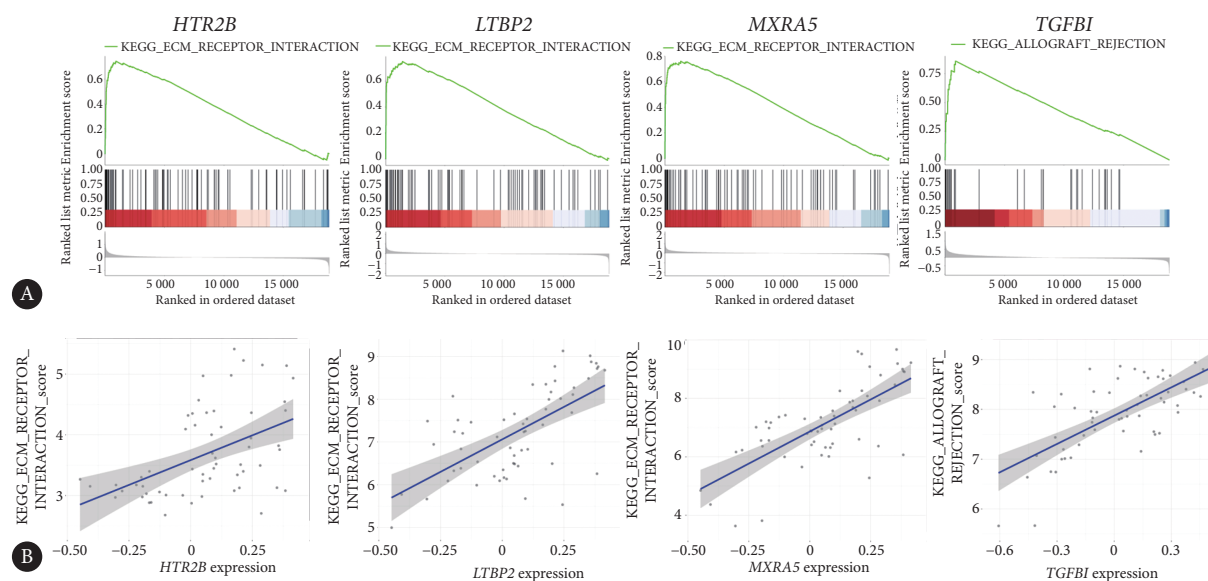


图3 特征基因GSEA和KEGG通路分析

Fig 3 GSEA and KEGG pathway analysis of feature genes

A, GSEA plot of feature genes. B, Scatter plot shows the association between feature genes and KEGG pathways involved genes.

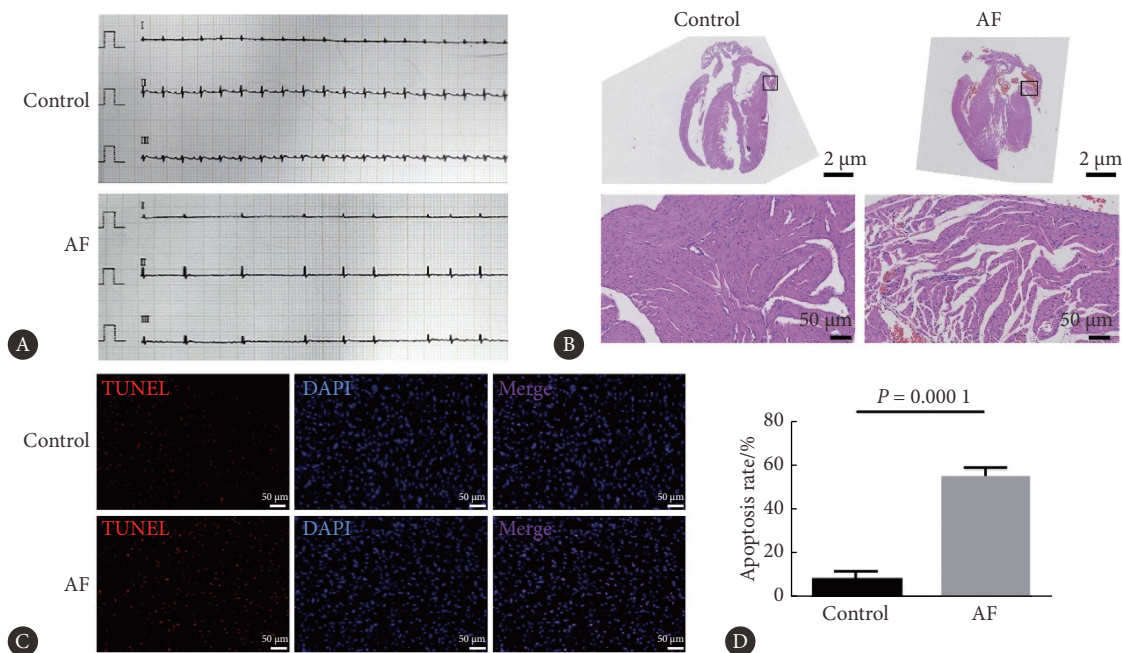


图4 两组大鼠心电图检测和组织病理学改变

Fig 4 Electrocardiogram examinations and histopathological changes in the two groups of rats

AF: atrial fibrillation. A, Electrocardiograph I, II, and III limb leads between two groups. B, HE staining of cardiac tissues between the two groups. C, TUNEL fluorescent labeling detection in the two groups. D, Comparison of apoptosis percentages between the two groups ($n = 6$ per group).

乱的电信号,导致心室收缩无效和充盈减少。房颤的基本特征之一是心房重构,其特点是心肌细胞纤维化,增加细胞外基质和胶原纤维沉积,导致电传导障碍的形成,直接干扰心脏动作电位,并增加维持心房颤动的折返回路的风险^[17-18]。研究表明心房扩张和纤维化增加,如在心力衰竭和AF患者中观察到的,对于心律失常的维持至关重要^[19-20]。AF电重塑通过离子通道功能和表达的变化表现

出来,导致动作电位持续时间和传导速度的改变。激酶的活性,如 Ca^{2+} /钙调蛋白依赖性蛋白激酶II在持续性房颤期间增加能维持心律失常的电学基质^[21]。另外,炎症在房颤和心房重构中的作用也不容忽视。炎症标志物(如 $\text{TNF-}\alpha$ 、IL-6等)水平升高与AF的病理生理学有关,影响了心房组织纤维化和电学特性^[22-24]。本研究将转录组学与炎症相关基因集结合,通过WGCNA共表达分析识

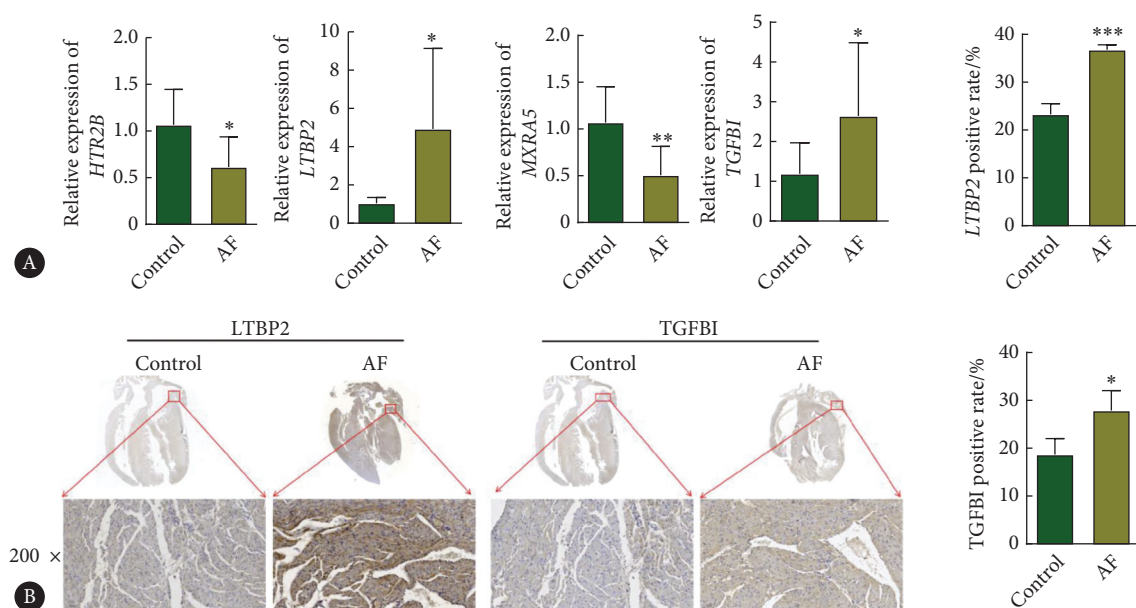


图5 RT-qPCR和免疫组织化学方法验证特征基因表达

Fig 5 Feature genes expression verification by the RT-qPCR and immunohistochemistry

AF: atrial fibrillation. A, Feature genes transcription detected with the RT-qPCR. B, Protein levels of feature genes verified by the immunohistochemistry. $n = 6$ per group. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, vs control.

别出与IRGs高度相关的模块($r=0.6 \sim 0.56$), 揭示了炎症反应在AF电重构中的核心调控作用。相较于传统单一组学分析, 该策略有效缩小了候选基因范围, 提高了生物标志物筛选的精准性。此外, 采用SVM-RFE与随机森林双算法交叉验证, 避免了单一模型过拟合的风险, 确保特征基因(*HTR2B*、*LTBP2*、*MXRA5*、*TGFB1*)的稳健性。

多种生物标志物与房颤的发生、发展及预后密切相关, 涉及神经-体液内分泌(B型利钠肽、肾素血管紧张素)、心肌纤维化(MMP-2、MMP-9、内皮素 I 等); 炎症及氧化应激的标志物包括: C反应蛋白、白介素6、同型半胱氨酸和尿酸等^[25]。但其临床应用仍面临诸多挑战。单一生物标志物的预测价值有限, 通常需要结合多种标志物进行综合评估; 其次, 生物标志物的检测方法和标准化尚未统一, 不同研究之间的结果可能存在差异。吴希泽等对GSE41177和GSE79768数据集进行生物信息学分析, 发现与AF相关的枢纽基因4个, 分别是*TYROBP*、*FCER1G*、*EVI2B*和*SOD2*调控免疫细胞激活及氧化应激等, 可能成为AF通用治疗靶点, 但缺乏有效的实验验证^[26]。我们不仅整合生物信息学和机器学习算法筛选AF的炎症标志物, 还设计AF小鼠模型加以验证。鉴定出2个特征基因*LTBP2*和*TGFB1*作为AF诊断的潜在生物标志物。

*LTBP2*属于LTBP(Latent Transforming Growth Factor Beta Binding Protein)家族成员, 是细胞外基质(extracellular matrix, ECM)的重要组成部分, 它们与微纤

维蛋白微纤维相互作用, 在微纤维生物学中发挥多种作用^[27]。PANG等^[28]研究结果表明TGF- β 的过度激活已经被证实可促进心房纤维化, 增加房颤的易感性, 而*LTBP2*作为TGF- β 的调控蛋白, 其表达异常可能致使信号通路失调, 加剧心房电生理紊乱。WANG等^[29]发现敲除*LTBP2*可减弱NF- κ B信号通路和减少氧化应激, 显著改善小鼠的心脏纤维化和细胞凋亡。*LTBP2*的单核苷酸多态位点(c.2206G>A)突变可影响心肌细胞的收缩性, 对心脏发育产生影响^[30]。此外, *LTBP-1*基因敲除小鼠显示出些小鼠表现出严重的发育性心血管异常并在围产期死亡^[31]。ZHOU等^[32]使用不同数据集发现*LTBP2*在心衰和AF患者中表达均升高, 与本项目AF患者和小鼠*LTBP2*表达增高结果相一致。TGF- β 1作为多效性细胞因子, 通过调控心肌纤维化、炎症反应及电重构等机制参与房颤的发生与发展。TGF- β 1可通过Wnt和Notch信号通路激活心脏成纤维细胞, 促进胶原蛋白 I / III 的合成与沉积, 导致心房结构重构^[33-34]。REN等^[35]发现TGF- β 1以剂量和时间依赖性方式增加电压门控钠通道Nav1.3表达, 进而改变动作电位阈值, 这种效应依赖于ERK1/2-JNK-MAPK通路的激活, 这一电生理异常可能促进房颤的触发和维持。TGF- β 1上调可促进IL-6等促炎因子的释放, 而IL-6本身又可进一步刺激TGF- β 1的表达, 形成正反馈循环^[35-36]。炎症-纤维化网络可能共同加剧心房基质的不均一性, 为房颤提供病理基础。TGFB1作为一种由TGF- β 诱导的细

胞外基质蛋白,在多种组织中参与调控细胞黏附、ECM重塑及信号转导,这些功能与房颤的病理机制存在潜在关联。TGFB1通过与胶原蛋白及整合素家族成员的相互作用,参与ECM的结构维持和动态平衡^[37]。研究表明,TGFB1的缺失可降低炎症因子的表达,如IL-1B、IL-6和JAK2的减少,从而减轻炎症反应^[38]。然而,TGFB1的过表达或功能异常可能通过激活JAK-STAT和PI3K-Akt信号通路,促进炎症因子的释放和炎症细胞的浸润^[39]。TGFB1作为一种多功能的细胞外基质蛋白,可能通过调控ECM重塑、炎症反应和细胞增殖,参与房颤的病理生理过程。

我们采用生物信息学整合机器学习算法筛选炎症相关的AF特征基因,再通过动物模型验证后作为AF生物标记物。但本项研究存在一定的局限性。首先,筛选出4个AF组高表达特征基因(*HTR2B*、*TGFB1*、*MXRA5*、*LTBP2*),而AF大鼠模型验证基因表达和翻译水平均有上调2个基因为*LTBP2*和*TGFB1*。AF模型多样,本实验采用化学试剂诱导的方法建立模型,后续将研究不同AF动物模型4种特征基因表达。其次,物种差异可能导致研究结果出现偏差,收集AF患者心房样本进行生物标记物验证。然后,还需探索类器官、微流控芯片等体外模型中4种基因表达差异。最后,本研究尚未探讨*LTBP2*和*TGFB1*与AF的因果关系及具体分子机制。

综上所述,通过结合生物信息学筛选、机器学习算法过滤和动物实验验证,我们证实AF发生过程中*LTBP2*和*TGFB1*这两个特征基因表达上调,其可作为AF生物标记物的候选基因,有必要深入探索*LTBP2*和*TGFB1*在AF发病机制中的作用,这将为后续研究及临床运用提供理论依据和实验基础。

* * *

作者贡献声明 卜琪负责论文构思、正式分析、调查研究、研究方法和初稿写作,张伟负责数据编、正式分析、研究方法和验证,黄英负责提供资源、软件和可视化,李铮负责调查研究和软件,王杨、樊晓燕和杨媛负责验证,赵璇和贾云莉负责可视化,王志敏负责论文构思、经费获取、研究项目管理、监督指导、初稿写作和审读与编辑写作。所有作者已经同意将文章提交给本刊,且对将要发表版本进行最终定稿,并同意对工作的所有方面负责。

Author Contribution BU Qi is responsible for conceptualization, formal analysis, investigation, methodology, and writing--original draft. ZHANG Wei is responsible for data curation, formal analysis, methodology, and validation. HUANG Ying is responsible for resources, software, and visualization. LI ZHENG is responsible for investigation and software. WANG Yang, FAN Xiaoyan, and YANG Yuan are responsible for validation. ZHAO Xuan and JIA Yunli are responsible for visualization. WANG Zhimin is responsible for conceptualization, funding acquisition, project administration, supervision, writing--original draft, and writing--

review and editing. All authors consented to the submission of the article to the Journal. All authors approved the final version to be published and agreed to take responsibility for all aspects of the work.

利益冲突 所有作者均声明不存在利益冲突

Declaration of Conflicting Interests All authors declare no competing interests.

参 考 文 献

- [1] BRUNDEL B, AI X, HILLS M T, *et al.* Atrial fibrillation. *Nat Rev Dis Primers*, 2022, 8(1): 21. doi: 10.1038/s41572-022-00347-9.
- [2] BORIANI G, LAROCHE C, DIEMBERGER I, *et al.* Asymptomatic atrial fibrillation: clinical correlates, management, and outcomes in the EORP-AF Pilot General Registry. *Am J Med*, 2015, 128(5): 509-518.e2. doi: 10.1016/j.amjmed.2014.11.026.
- [3] WITKOWSKI M, BISSINGER A, GRYCEWICZ T, *et al.* Asymptomatic atrial fibrillation in patients with atrial fibrillation and implanted pacemaker. *Int J Cardiol*, 2017, 227: 583-588. doi: 10.1016/j.ijcard.2016.10.097.
- [4] WEGNER F K, KOCHHÄUSER S, ELLERMANN C, *et al.* Prospective blinded evaluation of the smartphone-based AliveCor Kardia ECG monitor for atrial fibrillation detection: The PEAK-AF study. *Eur J Intern Med*, 2020, 73: 72-75. doi: 10.1016/j.ejim.2019.11.018.
- [5] Van Den BERG M P, MULDER B A, KLAASSEN S H C, *et al.* Heart failure with preserved ejection fraction, atrial fibrillation, and the role of senile amyloidosis. *Eur Heart J*, 2019, 40(16): 1287-1293. doi: 10.1093/eurheartj/ehz057.
- [6] FUHER A N, BORNE R, CUNNINGHAM J. Diagnosing atypical flutter in the post-atrial fibrillation ablation patient: a case report. *Clin Pract Cases Emerg Med*, 2023, 7(2): 106-109. doi: 10.5811/cpcem.1413.
- [7] WU L D, LI F, CHEN J Y, *et al.* Analysis of potential genetic biomarkers using machine learning methods and immune infiltration regulatory mechanisms underlying atrial fibrillation. *BMC Med Genomics*, 2022, 15(1): 64. doi: 10.1186/s12920-022-01212-0.
- [8] DEMERS J, TON A T, HUYNH F, *et al.* Atrial electrical remodeling in mice with cardiac-specific overexpression of angiotensin II type 1 receptor. *J Am Heart Assoc*, 2022, 11(8): e023974. doi: 10.1161/jaha.121.023974.
- [9] ZHENG L, JIA X, ZHANG C, *et al.* Angiotensin II in atrial structural remodeling: the role of Ang II/JAK/STAT3 signaling pathway. *Am J Transl Res*, 2015, 7(6): 1021-1031.
- [10] MURAKATA Y, YAMAGAMI F, MURAKOSHI N, *et al.* Electrical, structural, and autonomic atrial remodeling underlies atrial fibrillation in inflammatory atrial cardiomyopathy. *Front Cardiovasc Med*, 2022, 9: 1075358. doi: 10.3389/fcvm.2022.1075358.
- [11] SMORODINOVA N, LANTOVÁ L, BLÁHA M, *et al.* Biopptic study of left and right atrial interstitium in cardiac patients with and without atrial fibrillation: interatrial but not rhythm-based differences. *PLoS One*, 2015, 10(6): e0129124. doi: 10.1371/journal.pone.0129124.
- [12] ALTIERI D I, ETZION Y, ANDERSON H D. Cannabinoid receptor agonist attenuates angiotensin II-induced enlargement and mitochondrial dysfunction in rat atrial cardiomyocytes. *Front Pharmacol*, 2023, 14: 1142583. doi: 10.3389/fphar.2023.1142583.
- [13] LANGFELDER P, HORVATH S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, 2008, 9(1): 559. doi: 10.1186/1471-2105-9-559.
- [14] HÄNZELMANN S, CASTELO R, GUINNEY J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*, 2013, 14: 7. doi: 10.1186/1471-2105-14-7.
- [15] WU T, HU E, XU S, *et al.* clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation (Camb)*, 2021, 2(3): 100141. doi: 10.1016/j.xinn.2021.100141.
- [16] LIU P, SUN H, ZHOU X, *et al.* CXCL12/CXCR4 axis as a key mediator in atrial fibrillation via bioinformatics analysis and functional identification. *Cell Death Dis*, 2021, 12(9): 813. doi: 10.1038/s41419-021-04109-5.

- [17] SZILÁGYI J, SÁGHY L. Atrial remodeling in atrial fibrillation. Comorbidities and markers of disease progression predict catheter ablation outcome. *Curr Cardiol Rev*, 2021, 17(2): 217-229. doi: [10.2174/1573403x16666200721153620](https://doi.org/10.2174/1573403x16666200721153620).
- [18] LI C Y, ZHANG J R, HU W N, *et al*. Atrial fibrosis underlying atrial fibrillation (review). *Int J Mol Med*, 2021, 47(3): 9. doi: [10.3892/ijmm.2020.4842](https://doi.org/10.3892/ijmm.2020.4842).
- [19] MOLINA C E, ABU-TAHA I H, WANG Q, *et al*. Profibrotic, electrical, and calcium-handling remodeling of the atria in heart failure patients with and without atrial fibrillation. *Front Physiol*, 2018, 9: 1383. doi: [10.3389/fphys.2018.01383](https://doi.org/10.3389/fphys.2018.01383).
- [20] WANG W, ZHANG F L, CHEN J H, *et al*. Telmisartan reduces atrial arrhythmia susceptibility through the regulation of RAS-ERK and PI3K-Akt-eNOS pathways in spontaneously hypertensive rats. *Can J Physiol Pharmacol*, 2015, 93(8): 657-665. doi: [10.1139/cjpp-2014-0416](https://doi.org/10.1139/cjpp-2014-0416).
- [21] LIU Z, FINET J E, WOLFRAM J A, *et al*. Calcium/calmodulin-dependent protein kinase II causes atrial structural remodeling associated with atrial fibrillation and heart failure. *Heart Rhythm*, 2019, 16(7): 1080-1088. doi: [10.1016/j.hrthm.2019.01.013](https://doi.org/10.1016/j.hrthm.2019.01.013).
- [22] YU J F, DONG Q, DU Y M. Interleukin-6: molecular mechanisms and therapeutic perspectives in atrial fibrillation. *Curr Med Sci*, 2025, 45(2): 157-168. doi: [10.1007/s11596-025-00021-7](https://doi.org/10.1007/s11596-025-00021-7).
- [23] LARSEN B C, AL-BADRI K H A, MELLEMEKJ/ER S, *et al*. Inflammatory-mediated atrial cardiomyopathy diagnosed using multimodality imaging and successfully treated with prednisolone: a case report. *Eur Heart J Case Rep*, 2024, 8(4): ytac167. doi: [10.1093/ehjcr/ytac167](https://doi.org/10.1093/ehjcr/ytac167).
- [24] LAKIN R, POLIDOVITCH N, YANG S, *et al*. Cardiomyocyte and endothelial cells play distinct roles in the tumour necrosis factor (TNF)-dependent atrial responses and increased atrial fibrillation vulnerability induced by endurance exercise training in mice. *Cardiovasc Res*, 2023, 119(16): 2607-2622. doi: [10.1093/cvr/cvad144](https://doi.org/10.1093/cvr/cvad144).
- [25] 李春兰, 裴丽娟, 侯鹏. 心房颤动相关的生物学标记物. *中国心血管杂志*, 2020, 25(4): 397-400. doi: [10.3969/j.issn.1007-5410.2020.04.020](https://doi.org/10.3969/j.issn.1007-5410.2020.04.020).
LI C L, PEI L J, HOU P. Biomarkers related to atrial fibrillation. *Chiese Journal of Cardiovascular Medicine*, 2020, 25(4): 397-400. doi: [10.3969/j.issn.1007-5410.2020.04.020](https://doi.org/10.3969/j.issn.1007-5410.2020.04.020).
- [26] 吴希泽, 李越, 潘嘉祥, 等. 基于生物信息学探讨中国人群发生心房颤动的机制及潜在生物标志物. *浙江大学学报(医学版)*, 2024, 53(5): 593-603. doi: [10.3724/zdxbyxb-2024-0027](https://doi.org/10.3724/zdxbyxb-2024-0027).
WU X, LI Y, PAN J, *et al*. Pathogenesis and potential diagnostic biomarkers of atrial fibrillation in Chinese population: a study based on bioinformatics. *Journal of Zhejiang University (Medical Sciences)*, 2024, 53(5): 593-603. doi: [10.3724/zdxbyxb-2024-0027](https://doi.org/10.3724/zdxbyxb-2024-0027).
- [27] ROBERTSON I B, HORIGUCHI M, ZILBERBERG L, *et al*. Latent TGF- β -binding proteins. *Matrix Biol*, 2015, 47: 44-53. doi: [10.1016/j.matbio.2015.05.005](https://doi.org/10.1016/j.matbio.2015.05.005).
- [28] PANG X F, LIN X, DU J J, *et al*. LTBP2 knockdown by siRNA reverses myocardial oxidative stress injury, fibrosis and remodelling during dilated cardiomyopathy. *Acta Physiol (Oxf)*, 2020, 228(3): e13377. doi: [10.1111/apha.13377](https://doi.org/10.1111/apha.13377).
- [29] WANG Y, HUA D, ZHAO K, *et al*. LTBP2 knockdown ameliorates cardiac fibrosis and apoptosis via attenuating NF- κ B signaling pathway and oxidative stress in mice with cardiac hypertrophy. *Advanced Therapeutics*, 2024, 7(8): 2300451. doi: [10.1002/adtp.202300451](https://doi.org/10.1002/adtp.202300451).
- [30] CHEN H X, YANG Z Y, HOU H T, *et al*. Novel mutations of TCTN3/LTBP2 with cellular function changes in congenital heart disease associated with polydactyly. *J Cell Mol Med*, 2020, 24(23): 13751-13762. doi: [10.1111/jcmm.15950](https://doi.org/10.1111/jcmm.15950).
- [31] HORIGUCHI M, TODOROVIC V, HADJIOLOVA K, *et al*. Abrogation of both short and long forms of latent transforming growth factor- β binding protein-1 causes defective cardiovascular development and is perinatally lethal. *Matrix Biol*, 2015, 43: 61-70. doi: [10.1016/j.matbio.2015.03.006](https://doi.org/10.1016/j.matbio.2015.03.006).
- [32] ZHOU M, XU Y, ZHANG L, *et al*. Comprehensive bioinformatics analysis of hub genes in ischemic heart failure and atrial fibrillation. *Front Cardiovasc Med*, 2025, 12: 1499065. doi: [10.3389/fcvm.2025.1499065](https://doi.org/10.3389/fcvm.2025.1499065).
- [33] ZHAO L, XU Y, TAO L, *et al*. Oxymatrine inhibits transforming growth factor β 1 (TGF- β 1)-induced cardiac fibroblast-to-myofibroblast transformation (FMT) by mediating the notch signaling pathway *in vitro*. *Med Sci Monit*, 2018, 24: 6280-6288. doi: [10.12659/msm.910142](https://doi.org/10.12659/msm.910142).
- [34] LIU C, ZHOU D, ZHANG Q, *et al*. Transcription factor EB (TFEB) improves ventricular remodeling after myocardial infarction by inhibiting Wnt/ β -catenin signaling pathway. *PeerJ*, 2023, 11: e15841. doi: [10.7717/peerj.15841](https://doi.org/10.7717/peerj.15841).
- [35] REN Z, LI T, LIU X, *et al*. Transforming growth factor-beta 1 enhances discharge activity of cortical neurons. *Neural Regen Res*, 2025, 20(2): 548-556. doi: [10.4103/nrr.Nrr-d-23-00756](https://doi.org/10.4103/nrr.Nrr-d-23-00756).
- [36] HAN D H, SHIN M K, SUNG J S, *et al*. miR-335-3p attenuates transforming growth factor beta 1-induced fibrosis by suppressing thrombospondin 1. *PLoS One*, 2024, 19(10): e0311594. doi: [10.1371/journal.pone.0311594](https://doi.org/10.1371/journal.pone.0311594).
- [37] POULSEN E T, RUNAGER K, NIELSEN N S, *et al*. Proteomic profiling of TGFBI-null mouse corneas reveals only minor changes in matrix composition supportive of TGFBI knockdown as therapy against TGFBI-linked corneal dystrophies. *FEBS J*, 2018, 285(1): 101-114. doi: [10.1111/febs.14321](https://doi.org/10.1111/febs.14321).
- [38] SCIRIHA G G, BORG J, SULTANA J, *et al*. Transcriptome analysis of TGFBI knockdown vs normal corneal epithelial cells: implications for TGFBI corneal dystrophy treatment. *Biochem Genet*, 2026, 64(3): 3451-3480. doi: [10.1007/s10528-025-11191-3](https://doi.org/10.1007/s10528-025-11191-3).
- [39] CHEN Z, WANG J, PENG P, *et al*. Hypoxia-induced TGFBI maintains glioma stem cells by stabilizing EphA2. *Theranostics*, 2024, 14(15): 5778-5792. doi: [10.7150/thno.95141](https://doi.org/10.7150/thno.95141).

(2025-06-06 收稿, 2026-04-16 修回)

编辑 汤洁



开放获取 本文使用遵循知识共享署名—非商业性使用 4.0 国际许可协议 (CC BY-NC 4.0), 详细信息请访问

<https://creativecommons.org/licenses/by-nc/4.0/>。

OPEN ACCESS This article is licensed for use under Creative Commons Attribution-NonCommercial 4.0 International license (CC BY-NC 4.0). For more information, visit <https://creativecommons.org/licenses/by-nc/4.0/>.

© 2026 《四川大学学报(医学版)》编辑部

Editorial Office of Journal of Sichuan University (Medical Sciences)