

DOI:10.11798/j.issn.1007-1520.202322142

· 鼻窦疾病专栏 ·

基于随机森林与人工神经网络慢性鼻窦炎伴鼻息肉诊断模型的构建与分析

张浩鹏¹, 王丽华¹, 尹梓名², 岑经途¹, 吉琳¹, 郭裕¹

(1. 上海中医药大学附属市中医医院 耳鼻咽喉科, 上海 200040; 2. 上海理工大学健康科学与工程学院, 上海 200082)

摘要: **目的** 慢性鼻窦炎伴鼻息肉(CRSwNP)是一个全球性的健康问题, 现有的诊断技术存在一定局限性, 因此有必要开发新的诊断模型来补充现有的诊断方法。**方法** 利用CRSwNP患者(GSE23552、GSE36830)的公开基因表达数据来识别潜在的差异基因, 应用随机森林算法和人工神经网络进一步筛选特异性基因, 建立CRSwNP的早期诊断模型。**结果** 共发现78个上调基因和25个下调基因, 随机森林算法筛选了4个特异性基因(HPGDS、ILIRL1、FMO3、DOK3), 人工神经网络构建出基于上述基因的预测模型, 该模型具有良好的预测效果(AUC=0.986), 独立数据集GSE194282进一步验证了其准确性(AUC=0.888)。**结论** 采用机器学习方法建立了一个基于基因表达水平的预测模型, 该模型可以预测早期CRSwNP, 有助于早期诊断和改善临床决策。

关键词: 慢性鼻窦炎伴鼻息肉; 诊断模型; 随机森林; 人工神经网络

中图分类号: R765.4⁺1

Construction and analysis on the diagnostic model of chronic rhinosinuitis with nasal polyps based on random forest and artificial neural network

ZHANG Haopeng¹, WANG Lihua¹, YIN Ziming², CEN Jingtu¹, JI Lin¹, GUO Yu¹

(1. Department of Otolaryngology, Shanghai Municipal Hospital of Traditional Chinese Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai 200040, China; 2. School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200082, China)

Abstract: **Objective** Chronic rhinosinuitis with nasal polyps (CRSwNP) is a global health problem, and the existing diagnostic techniques have some limitations. Therefore, it is necessary to develop new diagnostic models to supplement the existing diagnostic methods. **Methods** The public gene expression data of CRSwNP patients (GSE23552, GSE36830) were used to identify potential differential genes. The random forest algorithm and artificial neural network were used to screen specific genes and establish the early diagnosis model of CRSwNP. **Results** A total of 78 up-regulated genes and 25 down-regulated genes were identified. Four specific genes (HPGDS, ILIRL1, FMO3 and DOK3) were screened by random forest algorithm. The prediction model based on the above genes was constructed by artificial neural network, which had good prediction effect (area under the curve = 0.986). Independent dataset GSE194282 further verified the accuracy (area under the curve = 0.888). **Conclusions** A predictive model based on gene expression level is established by machine learning method. This model can predict early CRSwNP, which is helpful for early diagnosis and clinical decision.

Keywords: Chronic rhinosinuitis with nasal polyps; Diagnostic model; Random forest; Artificial neural network

慢性鼻窦炎伴鼻息肉(chronic rhinosinuitis with nasal polyps, CRSwNP)是一种鼻腔黏膜和鼻窦的慢

性炎症性疾病, 在成人中的患病率大约 2.5% ~ 5.8%, 约占鼻窦炎患者的 20% ~ 30%^[1-2], 由于

基金项目: 国家自然科学基金面上项目(82074581)。

第一作者简介: 张浩鹏, 男, 在读博士研究生。

通信作者: 郭裕, Email: gysyent@163.com

CRSwNP 患者人群常常合并哮喘等呼吸道疾病,或者需要反复使用皮质类固醇和/或鼻腔手术以缓解症状,这严重降低了患者生活质量,带来了极大的身体和心理负担。截至目前研究证实,CRSwNP 的发病机制主要涉及鼻腔黏膜上皮重塑、黏液纤毛功能缺陷、嗜酸性粒细胞增多、免疫功能缺陷等^[2],但尚未完全阐明,有待进一步研究。此外,目前对于 CRSwNP 的诊断存在局限性,CRSwNP 的主要临床症状包括持续 12 周以上的鼻塞、嗅觉丧失、鼻漏、鼻后滴漏、面部充血和疼痛,但临床患者可能仅伴随其中两个或两个以上症状,增加了漏诊的概率。虽然鼻内镜或鼻窦 CT 等诊断工具已广泛运用于临床,但此法受限于个人的操作能力和专家的诊断经验,并且在大规模人群研究中难以应用。此外,基于症状的诊断存在不准确性,有研究表明约 40% 的基于症状的诊断患者没有显示出明确的内镜或放射学证据^[2]。因此,CRSwNP 的诊断通常需要具有丰富临床经验的医生结合内镜或 CT 给出较为准确的结论,特别是在该病的早期诊断中。

对 CRSwNP 的早期诊断、早期切入治疗有助于预测疾病自然病程、逆转疾病进展,改善预后,并有助于减少哮喘等并发症的发生率,因此,亟待开发新的诊断模型以补充现有诊断措施。近年来,第二代测序技术的快速发展为多种疾病相关基因的鉴定提供了基础,在本研究中,我们在 GEO 数据库中筛选了 CRSwNP 和正常鼻黏膜样本之间的差异表达基因,基于这些数据,使用随机森林算法来识别在 CRSwNP 中表达的关键基因,随后通过人工神经网络方法,构建了 CRSwNP 的早期诊断模型(图 1)。

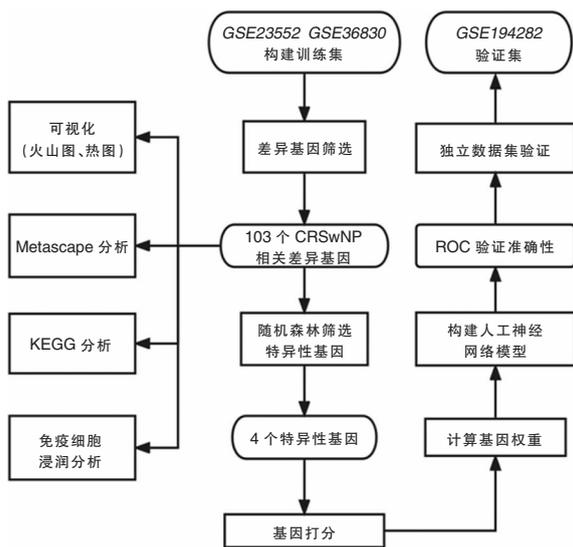


图 1 流程图

1 材料与方法

1.1 数据下载和处理

从 GEO 数据库下载 3 个 CRSwNP 患者相关数据集(*GSE23552*、*GSE36830*、*GSE194282*)的基因表达谱和临床数据(表 1),将 *GSE23552*、*GSE36830* 作为训练集构建诊断模型,将 *GSE194282* 作为验证集以验证准确性。使用 R 软件(4.1.0)对原始数据进行标准化和归一化处理,使用 *sva* 包消除训练集中 2 个数据集的批次效应并进行合并,以供进一步分析。

表 1 训练/验证数据集基本信息

基因	平台文件	样本量	CRSwNP	正常
<i>GSE23552</i>	GPL5175	34	21	13
<i>GSE36830</i>	GPL570	24	18	6
<i>GSE194282</i>	GPL17692	14	7	7

1.2 差异表达和富集分析

使用 *limma* 包筛选训练集中 CRSwNP 样本和正常样本的差异表达基因(筛选标准: $P < 0.05$, $|\log FC| > 1$),使用 *pheatmap* 包和 *ggplot2* 包对差异表达基因进行可视化处理,绘制差异基因热图和火山图。

将差异表达基因导入 Metascape 网站(<http://metascape.org>),进行基因功能注释和通路富集分析,以识别训练集中显著富集的生物学功能和信号通路(筛选标准: $P < 0.05$,富集基因数 ≥ 3),然后使用 R 软件的 *enrichplot*、*ggplot2*、*GOpilot* 包进行 GO 和 KEGG 分析并可视化。

1.3 对差异基因进行随机森林筛选

使用 *randomForest* 包构建差异基因的随机森林模型。首先,设定最大决策树为 500,通过有放回的随机抽样方式构建分类树,并计算误差值,以最低的误差率和最佳的稳定性来确定最优的树数,根据 Gini 系数法选择重要性评分大于 2 的基因作为特异性基因,进行后续模型构建。

将特异性基因的表达数据转换为基因评分表,具体转换规则如下:将基因的表达值与所有样本表达值的中位值进行比较,如果上调基因的表达值大于中位值,则其值为 1,否则为 0;同样,如果下调基因的表达值大于中位值,则其值为 0,否则为 1。

1.4 人工神经网络构建疾病分类模型

使用 *neuralnet* 和 *NeuralNetTools* 包构建人工神经网络模型,将疾病属性设置为 y ,基因得分设置为 x ,神经网络由一个输入层、一个隐藏层和一个输出层组成,隐藏层中包括 5 个隐藏节点,输出层包括

2个节点(疾病和正常),通过获得的基因权重信息构建 CRSwNP 疾病分类模型,该模型中,权重得分乘以疾病特异性基因表达水平的乘积之和作为疾病分类得分。最后,使用 pROC 包绘制接收者操作特征曲线(receiver operating characteristic curve, ROC),并计算 ROC 曲线下面积(area under ROC curve, AUC)以验证模型准确性。

1.5 独立数据集验证

使用独立数据集(*GSE194282*)对构建的 CRSwNP 和正常样本的分类评分模型进行了有效性验证,首先根据人工神经网络模型计算特异性基因的疾病分类得分,然后使用 pROC 包绘制 ROC 曲线,计算 AUC 以验证分类效率,AUC 值大于 0.8 被认为准确度较高^[3]。

1.6 免疫细胞浸润分析

使用 preprocessCore 包对训练集数据进行归一化处理后,采用 CIBERSORT 反卷积算法对浆细胞、活化记忆性 CD4⁺T 细胞、活化 NK 细胞等 22 种免疫细胞在样本中的表达进行模拟计算(设定模拟次数为 1 000 次),然后使用 corplot、vioplot 包进一步进行相关性分析及可视化。

2 结果

2.1 差异基因表达分析

训练集中,*GSE23552* 包含 34 个样本,包括 21 个 CRSwNP 样本和 13 个正常样本,*GSE36830* 数据集包含 24 个样本,包括 18 个 CRSwNP 样本和 6 个正常样本。将训练集数据合并后,通过差异分析得到了 103 个差异基因,包括 78 个上调基因和 25 个下调基因,差异基因的结果显示在火山图(图 2A)和热图(图 2B)中。

2.2 差异基因功能富集分析

为了进一步了解差异基因相关的功能和代谢途径,通过 Metascape 进行了富集分析,并对差异基因显著富集的前 20 个聚类进行了可视化(图 3A、B)。在生物过程中,显著富集的条目为炎症应答、中性粒细胞脱颗粒、外部刺激和免疫效应过程的正向调节等。另外,KEGG 通路富集分析表明差异基因主要参与病毒蛋白与细胞因子和细胞因子受体的相互作用通路、趋化因子信号通路、细胞因子-细胞因子受体的相互作用等通路(图 3C)。鉴于这些生物过程或者通路均在 CRSwNP 中显著富集,因此,抑制相关信号通路可能是未来的补充或替代治疗措施之一。

2.3 差异基因的随机森林筛选

运用随机森林算法对 103 个差异基因进行了分析,为了找到最优的参数,对所有可能的变量都进行循环随机森林分类,并计算平均错误率,选择稳定性最优的模型进行后续分析,最终根据 Gini 系数确定了 4 个重要性 >2 的基因作为后续分析的特异性候选基因(图 4A)。在 4 个基因中,造血前列腺素 D 合成酶(hematopoietic prostaglandin D synthase, *HPGDS*)和白细胞介素 1 受体样 1(interleukin 1 receptor like 1, *IL1RL1*)最为重要,其次是含黄素单氧化酶 3(flavin containing dimethylaniline monooxygenase 3, *FMO3*)和对接蛋白 3(docking protein 3, *DOK3*)。在热图中(图 4B),4 个候选基因可用于区分疾病和正常样本,其中,*HPGDS*、*IL1RL1* 和 *DOK3* 在 CRSwNP 中高表达,在正常组织中低表达,而 *FMO3* 在 CRSwNP 中低表达。

2.4 人工神经网络模型的构建与验证

基于随机森林筛选的特异性基因,构建了一个人工神经网络模型(图 5A)。首先将 4 个特异性基因的表达数据转换为“基因评分”,然后采用神经网络算法计算每个基因的权重值,通过“基因评分”和“基因权重”计算模型分类评分,具体公式如下: CRSwNP 神经网络模型评分 = \sum (基因表达水平 × 基因权重)。然后,将 58 个样本的模型评分设为预测值,将 CRSwNP 和正常样本设为结果值,使用 pROC 包对模型进行分类能力进行验证,结果发现模型的 AUC 为 0.986(图 5B),提示模型具有准确的预测效果。

为了进一步验证模型准确性,使用另一个数据集(*GSE194282*)来验证模型在一个独立的数据集中是否具有同样的预测能力。运用同样的方法计算出了 4 个特异性基因的“基因评分”和“基因权重”,然后通过神经网络模型计算评分,并通过 ROC 曲线进行验证,结果发现,验证集的 AUC 为 0.888(图 5C),进一步证明了模型的有效性和稳定性。

2.5 免疫细胞浸润分析

为了进一步了解差异基因相关免疫细胞浸润情况,采用 CIBERSORT 反卷积算法对差异基因进行了分析。在小提琴图中,未活化的肥大细胞($P = 0.021$)、M2 型巨噬细胞($P = 0.007$)、嗜酸性粒细胞($P = 0.004$)在 CRSwNP 中出现不同程度的高表达,而未活化的记忆 CD4⁺T 细胞($P = 0.002$)、未活化的自然杀伤(natural killer, NK)细胞($P = 0.009$)、单核细胞($P = 0.042$)在 CRSwNP 中低表达

(图 6A),热图(图 6B)进一步展示了 22 种免疫细胞在各样本中的分布。在相关性分析(图 6C)中, Gamma Delta T 细胞和记忆 B 细胞($r = 0.70$)、未活化的肥大细胞和嗜酸性粒细胞($r = 0.46$)、未活化的

NK 细胞和单核细胞($r = 0.45$)呈正相关,而浆细胞和 M2 型巨噬细胞($r = -0.56$)、未活化的肥大细胞和 M0 型巨噬细胞($r = -0.51$)、M0 型巨噬细胞和 M2 型巨噬细胞($r = -0.54$)呈负相关。

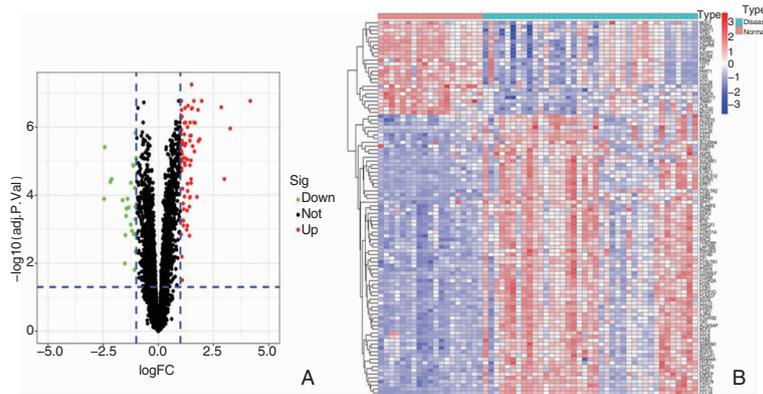


图 2 训练集中差异基因表达分析 A:差异基因表达的火山图,红色和绿色分别代表在 CRSwNP 中显著上调和下调的基因($P < 0.05$, $|\logFC| > 1$); B:差异基因表达热图(红色:高表达;蓝色:低表达。 $P < 0.05$) 注:CRSwNP(慢性鼻窦炎伴鼻息肉)。下同。

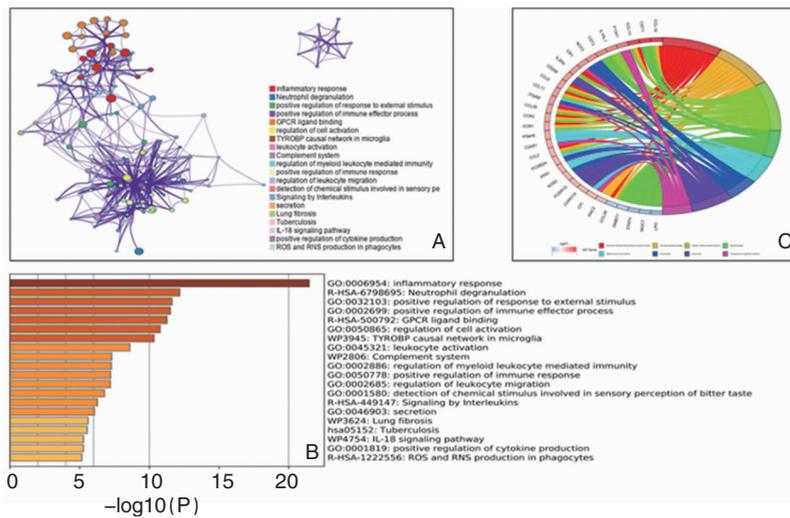


图 3 差异基因功能富集结果 A:前 20 个富集条目聚类网络图,聚类结果用颜色表示,线条粗细代表关联程度大小; B:前 20 个富集条目热图; C:KEGG 通路富集圈图

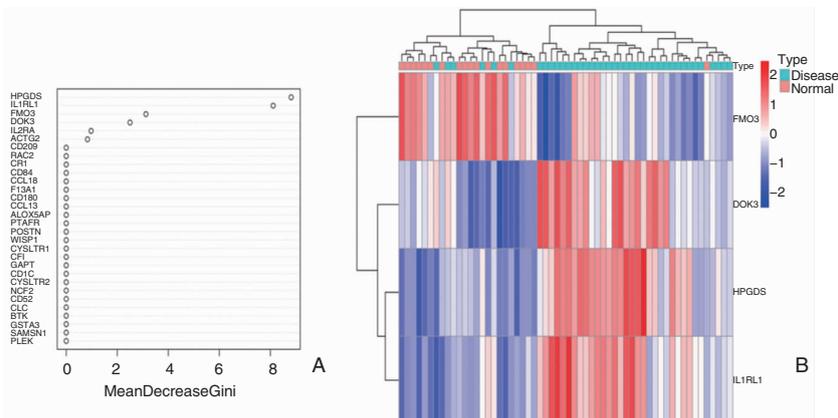


图 4 随机森林筛选结果 A:基因重要性评分(根据 Gini 系数排名); B:候选基因热图(Gini 系数 > 2)

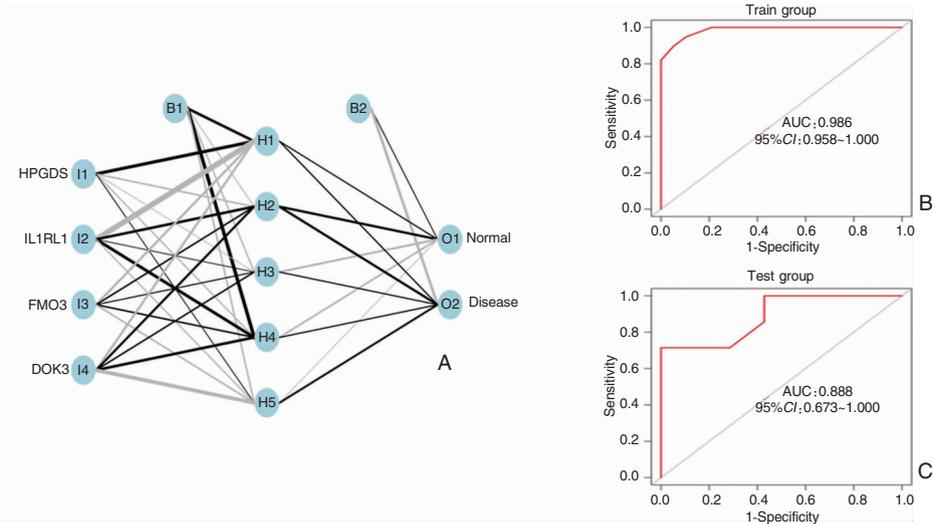


图 5 人工神经网络模型的构建与验证 A:人工神经网络可视化结果; B:训练集 ROC 曲线; C:验证集 ROC 曲线

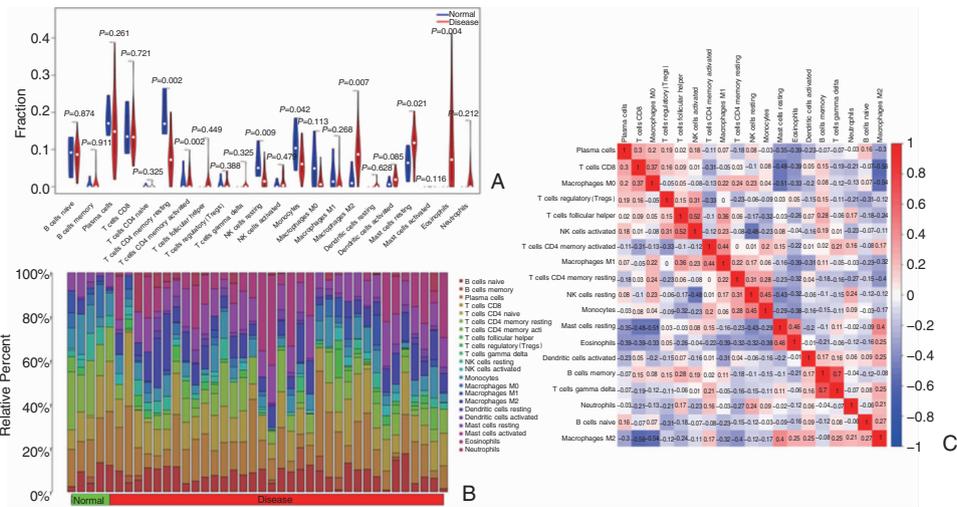


图 6 免疫细胞浸润分析 A:CRSwNP 和正常样本中免疫细胞浸润差异小提琴图; B:CRSwNP 和正常样本中免疫细胞浸润分布; C:免疫细胞浸润差异性分析

3 讨论

近年来,机器学习算法的发展和公共数据库中基因表达数据的应用为研究疾病诊断或预后的生物标志物提供了更多的方法。在 CRSwNP 早期诊断的研究领域中,既往已经通过差异性分析和功能注释确定了部分核心基因^[4-6],这些核心基因的确定有望提高该病的早期诊断率。与这些研究相比,本研究在差异性分析的基础上进一步进行了诊断模型的构建与验证,并成功构建了一个 CRSwNP 的早期诊断模型,进一步提高了筛选基因和预测模型的准确性。该模型将复杂的基因表达值转化为简单的临床评分,便于为早期疑似 CRSwNP 患者提供诊断,

也便于临床医生为患者制定合理、个性化、经济的诊疗方案。

对差异基因的富集结果显示,大多数差异基因参与了炎症应答、中性粒细胞脱颗粒、外部刺激和免疫效应过程的正向调节等。目前已知的在 CRSwNP 的慢性炎症和组织学改变过程中,黏液纤毛功能的改变、鼻窦上皮细胞屏障的异常和组织重塑以及宿主先天性和适应性免疫反应是主要原因^[6],富集结果证明了在 CRSwNP 发病过程中,上述功能的异常或缺陷最终导致了复杂的下游免疫反应和慢性炎症。此外,KEGG 通路分析表明上调基因主要参与病毒蛋白与细胞因子和细胞因子受体的相互作用通路、趋化因子信号通路、金黄色葡萄球菌感染等,金黄色葡萄球菌的定植增加和 Th2 型炎症是 CRSwNP 的

主要特点^[7],作为鼻部的常见定植菌,金黄色葡萄球菌同样也是常见的病原体,与炎症反应和慢性感染高度相关^[2],而趋化因子同样可诱发一系列免疫炎症反应,诱导组织重塑^[8],在 CRSwNP 的免疫炎症反应中亦发挥了重要作用。

本研究确定的 4 个特异性基因中, *HPGDS* 和 *IL1RL1* 已在既往研究中得到了证实。既往大量的临床及实验研究表明, *IL1RL1* 编码的膜受体与白细胞介素 33 (Interleukin 33, IL-33) 结合,参与一系列 2 型免疫、过敏性疾病的发病过程^[9], IL-33 在 CRSwNP 中高度表达,参与中性粒细胞募集和重塑等生物过程,促进 Th2 型炎症反应^[10-11],进一步促进 CRSwNP 的疾病进展。目前,针对 CRSwNP 的 IL-33 靶向生物制剂也正在临床研究之中,有望提高难治性 CRSwNP 患者的临床疗效^[12]。 *HPGDS* 催化前列腺素 H2 转化成前列腺素 D2,参与介导气道炎症,是重要的炎症反应调节因子^[13], *HPGDS* 在 CRSwNP 患者的鼻黏膜中高度表达,并且相对于钩突黏膜, *HPGDS* 在鼻息肉组织中表达水平更高,提示 *HPGDS* 不但具有促炎作用,同时还可能参与鼻息肉的发生与进展^[14-15]。 *DOK3* 是一种酪氨酸残基磷酸化蛋白,在 B 细胞中高度表达,参与细胞增殖、凋亡、生长和迁移等多种过程,目前研究中, *DOK3* 主要参与肿瘤的免疫过程,发挥促癌或者抑癌作用^[16],但有研究表明 *DOK3* 在病毒感染的细胞中上调^[17],推测可能与 CRSwNP 的慢性病毒感染有关,但其在 CRSwNP 中的作用有待进一步研究。 *FMO3* 最近已被证明和多种慢性疾病有关,有研究发现 *FMO3* 的酶产物三甲胺 N-氧化物可能是炎症性疾病的致病介质^[18],但其在 CRSwNP 中的作用仍不清楚,本研究发现 *FMO3* 在 CRSwNP 中低表达,其可能作为一种抑制基因的机制尚需进一步验证。

CIBERSORT 反卷积算法是基于线性支持向量回归的可靠机器学习方法,采用 CIBERSORT 反卷积算法对两组免疫相关差异基因进行筛选,结果提示未活化的肥大细胞、M2 型巨噬细胞、嗜酸性粒细胞在 CRSwNP 中高表达,而未活化的记忆 CD4⁺T 细胞、未活化的 NK 细胞、单核细胞在 CRSwNP 中低表达。肥大细胞和嗜酸性粒细胞在 CRSwNP 中的作用已被证实,二者在鼻息肉中的表达呈正相关,嗜酸性粒细胞产生的 IgE 可以激活肥大细胞,分泌大量 2 型细胞因子,促进 CRSwNP 中的 2 型炎症反应和嗜酸性炎症^[19]。与之相反,同样在鼻息肉中高度表达的 M2 型巨噬细胞被认为是一种保护机制, M2 型

巨噬细胞通过产生细胞外基质成分和血管生成因子,参与促进组织重塑,协调适应性免疫应答,发挥抗炎、营养和调节作用^[20],然而,组织中持续存在的巨噬细胞反而不利于结束炎症反应, M2 型巨噬细胞产生的白细胞介素 10 缺乏可能导致嗜酸性粒细胞型 CRSwNP 的持续炎症反应^[21],提示 M2 型巨噬细胞在 CRSwNP 中可能存在抗炎与促炎的双向作用,尚需进一步研究。此外,在 CRSwNP 患者的鼻息肉组织中已经观察到 NK 细胞的低表达, NK 细胞成熟过程的紊乱可能是 CRSwNP 的发病机制中的一个重要因素^[22],有研究发现 CD4⁺T 细胞在鼻息肉中低表达^[23],但低表达的未活化的 CD4⁺T 细胞在 CRSwNP 中的作用仍需进一步研究。尽管本研究发现嗜酸性粒细胞在 CRSwNP 中高表达,但也存在一些不足,就目前研究而言, CRSwNP 的鼻息肉组织常见有嗜酸性粒细胞浸润,可根据嗜酸性粒细胞的浸润程度分为两个亚型:嗜酸性粒细胞型 CRSwNP 和非嗜酸性粒细胞型 CRSwNP。其中,嗜酸性粒细胞型 CRSwNP 表现出更高的炎症水平和复发倾向,预后更差^[24]。若能根据嗜酸性粒细胞浸润程度将患者进行分组,进一步分析其基因差异,以及在更大的样本量中验证,有可能为 CRSwNP 的个体化精准治疗提供更可靠的参考价值,这也是我们的下一步研究方向。

综上所述,本研究建立了一个基于机器学习和基因表达水平的预测模型,该模型能够用于 CRSwNP 的早期诊断,并基于 GEO 数据库的独立样本验证了该模型的准确性。本研究为临床医生提供了一种新的诊断策略,便于提高早期诊断效率,改善治疗决策。

参考文献:

- [1] Bachert C, Bhattacharyya N, Desrosiers M, et al. Burden of disease in chronic rhinosinusitis with nasal polyps[J]. J Asthma Allergy, 2021, 14: 127 - 134.
- [2] Bachert C, Marple B, Schlosser RJ, et al. Adult chronic rhinosinusitis[J]. Nat Rev Dis Primers, 2020, 6(1): 86.
- [3] Feng J, Chen Y, Feng Q, et al. Novel gene signatures predicting primary non-response to infliximab in ulcerative colitis: development and validation combining random forest with artificial neural network[J]. Front Med (Lausanne), 2021, 8: 678424.
- [4] 王平,郝蕴,赵妍,等.基于生物信息学方法对慢性鼻窦炎伴鼻息肉差异基因表达的分析[J].解放军医学院学报,2021,42(11):1180 - 1187.
- [5] 蓝凤,王麒淇,张罗.慢性鼻窦炎伴鼻息肉鼻黏膜上皮细胞基

- 因转录组分析[J]. 中华耳鼻咽喉头颈外科杂志, 2021, 56(10):1066-1072.
- [6] Hao Y, Zhao Y, Wang P, et al. Transcriptomic signatures and functional network analysis of chronic rhinosinusitis with nasal polyps[J]. *Front Genet*, 2021, 12: 609754.
- [7] Feng T, Miao P, Liu B, et al. Sinus microbiota in patients with eosinophilic and non-eosinophilic chronic rhinosinusitis with nasal polyps[J]. *Front Cell Infect Microbiol*, 2021, 11: 672355.
- [8] 高雅丽, 王威清, 朱真真, 等. 趋化因子在嗜酸性慢性鼻-鼻窦炎伴鼻息肉中的价值研究[J]. 临床耳鼻咽喉头颈外科杂志, 2017, 31(18):1458-1461.
- [9] Wu M, Zheng X, Huang J, et al. Association of IL33, IL1RL1, IL1RAP polymorphisms and asthma in Chinese Han children[J]. *Front Cell Dev Biol*, 2021, 9: 759542.
- [10] Kim DK, Jin HR, Eun KM, et al. The role of interleukin-33 in chronic rhinosinusitis[J]. *Thorax*, 2017, 72(7): 635-645.
- [11] Ryu G, Kim DW. Th2 inflammatory responses in the development of nasal polyps and chronic rhinosinusitis[J]. *Curr Opin Allergy Clin Immunol*, 2020, 20(1): 1-8.
- [12] Thamboo A, Kilty S, Witterick I, et al. Canadian rhinology working group consensus statement: biologic therapies for chronic rhinosinusitis[J]. *J Otolaryngol Head Neck Surg*, 2021, 50(1): 15.
- [13] 陈灵修, 冉丹华, 黄华萍, 等. 前列腺素 D2 受体调节 PGD2 诱导的人气道黏液高分泌[J]. 基础医学与临床, 2016, 36(1): 24-29.
- [14] Yamamoto M, Okano M, Fujiwara T, et al. Expression and characterization of PGD2 receptors in chronic rhinosinusitis: modulation of DP and CRTH2 by PGD2[J]. *Int Arch Allergy Immunol*, 2009, 148(2): 127-136.
- [15] Choi YH, Lee S, Aoyagi H, et al. The extracellular signal-regulated kinase mitogen-activated protein kinase/ribosomal S6 protein kinase 1 cascade phosphorylates cAMP response element-binding protein to induce MUC5B gene expression via D-prostanoid receptor signaling[J]. *J Biol Chem*, 2011, 286(39): 34199-34214.
- [16] Guan Y, Li M, Qiu Z, et al. Comprehensive analysis of DOK family genes expression, immune characteristics, and drug sensitivity in human tumors[J]. *J Adv Res*, 2022, 36: 73-87.
- [17] Kim SS, Lee K, Chin C, et al. DOK3 is required for IFN- β production by enabling TRAF3/TBK1 complex formation and IRF3 activation[J]. *J Immunol*, 2014, 193(2): 840-848.
- [18] Yang Z, Stemmer PM, Petriello MC. Proteomics-based identification of interaction partners of the xenobiotic detoxification enzyme FMO3 reveals involvement in urea cycle[J]. *Toxics*, 2022, 10(2): 60.
- [19] Gelardi M, Giancaspro R, Cassano M, et al. The underestimated role of mast cells in the pathogenesis of rhinopathies[J]. *Int Arch Allergy Immunol*, 2022, 183(2): 153-159.
- [20] Kaczmarek M, Banaszewski J, Leszczyńska M, et al. High frequency of macrophages expressing elevated level of CD80, PD-Ls and TLR1 in nasal polyps of CRS patients[J]. *Immunobiology*, 2019, 224(1): 154-162.
- [21] Wang Z, Yao Y, Wang N, et al. Deficiency in interleukin-10 production by M2 macrophages in eosinophilic chronic rhinosinusitis with nasal polyps[J]. *Int Forum Allergy Rhinol*, 2018, 8(11): 1323-1333.
- [22] Kaczmarek M, Wasicka K, Tin-Tsen Chou J, et al. NK cells in patients with chronic rhinosinusitis show decreased maturity and limited expression of functional receptors [J]. *Immunobiology*, 2020, 225(2): 151890.
- [23] Ickrath P, Kleinsasser N, Ding X, et al. Characterization of T-cell Subpopulations in Patients with Chronic Rhinosinusitis with Nasal Polyposis[J]. *Allergy Rhinol (Providence)*, 2017, 8(3): 139-147.
- [24] 吴鸿泉, 刘展, 覃宇铭. 慢性鼻-鼻窦炎伴鼻息肉患者鼻内镜术后复发的危险因素及相关模型构建[J]. 中国耳鼻咽喉颅底外科杂志, 2021, 27(5): 559-564.

(收稿日期: 2022-04-15)

本文引用格式: 张浩鹏, 王丽华, 尹梓名, 等. 基于随机森林与人工神经网络慢性鼻窦炎伴鼻息肉诊断模型的构建与分析[J]. 中国耳鼻咽喉颅底外科杂志, 2023, 29(1): 24-30. DOI: 10.11798/j.issn.1007-1520.202322142

Cite this article as: ZHANG Haopeng, WANG Lihua, YIN Ziming, et al. Construction and analysis on the diagnostic model of chronic rhinosinusitis with nasal polyps based on random forest and artificial neural network[J]. *Chin J Otorhinolaryngol Skull Base Surg*, 2023, 29(1): 24-30. DOI: 10.11798/j.issn.1007-1520.202322142