

基于人工智能的艾叶活性成分多维药理活性及分子机制系统研究

张洪榕¹, 许奇², 马忠民³, 方朝晖^{4*}

(¹ 安徽中医药大学医药信息工程学院, 合肥 230012; ² 安徽中医药大学中医学学院, 合肥 230038; ³ 安徽守艾人中医药科技有限公司, 合肥 230031; ⁴ 合肥综合性国家科学中心大健康研究院新安医学与中医药现代化研究所, 合肥 230000)

摘要 利用人工智能技术探究艾叶中的活性成分的药理活性及潜在作用机制。检索 HIT、TCMSP、TCMIO 数据库获取 199 种艾叶的活性成分。综合 K 最近邻、多层感知机、随机森林、支持向量机算法及基于 Lipinski 规则及 Veber 规则的算法模型对艾叶化合物成分进行毒性和口服利用度预测, 发现其中 14 种成分满足无毒性且口服利用度较好。采用合成可行性分数 (SAscore) 模型对以上 14 种成分进行可合成性分析, 并通过 BRICS 及 RECAP 算法分割分子片段。STP 和 PM 数据库挖掘得到 406 个艾叶核心成分的作用靶点, 利用 Cytoscape 筛出 5 个核心靶点为 SRC、EGFR、PTPN11、HRAS 和 PDGFRB。GO 与 KEGG 富集分析提示核心靶点涉及 808 项 GO 富集分析条目和 EGFR 酪氨酸激酶抑制剂抵抗、间隙连接、磷脂酶 D、JAK/STAT 等 71 条信号通路。分子对接结果显示艾叶活性化合物与 SRC、EGFR、PTPN11、HRAS 蛋白具有良好的结合能力。细胞实验证实了艾叶活性成分杜香醇在一定浓度范围内能够促进 HUVEC 细胞增殖, 并可促进 EGFR 蛋白的表达。本研究揭示了艾叶活性成分的药理特性及其潜在的分子机制, 为其药用开发奠定了坚实的科学基础。

关键词 人工智能; 艾叶; 活性成分; 药理特性; 分子机制

中图分类号 TP181;R914 **文献标志码** A **文章编号** 1000-5048(2025)03-0358-10

doi: 10.11665/j.issn.1000-5048.2024080701

引用本文 张洪榕, 许奇, 马忠民, 等. 基于人工智能的艾叶活性成分多维药理活性及分子机制系统研究 [J]. 中国药科大学学报, 2025, 56(3): 358–367.

Cite this article as: ZHANG Hongrong, XU Qi, MA Zhongmin, *et al.* Artificial intelligence-based systematic study on the multidimensional pharmacological activity and molecular mechanism of the active ingredients of *Artemisia argyi*[J]. *J China Pharm Univ*, 2025, 56(3): 358–367.

Artificial intelligence-based systematic study on the multidimensional pharmacological activity and molecular mechanism of the active ingredients of *Artemisia argyi*

ZHANG Hongrong¹, XU Qi², MA Zhongmin³, FANG Zhaohui^{4*}

¹School of Medical Informatics Engineering, Anhui University of Chinese Medicine, Hefei 230012; ²School of Chinese Medicine, Anhui University of Chinese Medicine, Hefei 230038; ³Anhui Shou Airen Traditional Chinese Medicine Science and Technology Co., Ltd., Hefei 230031; ⁴Xin'an Institute of medicine and modernization of Chinese medicine, National Science Center's Institute of Comprehensive Health, Hefei 230000, China

Abstract To investigate the pharmacological activities and potential mechanisms of action of the active components in *Artemisia argyi* with artificial intelligence technology, a search was conducted in the HIT, TCMSP, and TCMIO databases, obtaining 199 active components of *A. argyi*. A comprehensive set of algorithms, including KNN, MLP, RF, SVM, and models based on Lipinski's and Veber's rules, was employed to

收稿日期 2024-08-07 * 通信作者 Tel: 17355066323 E-mail: fangzhaohui1111@163.com

基金项目 安徽省高校协同创新项目 (GXXT-2020-025); 合肥综合性国家科学中心大健康研究院新安医学与中医药现代化研究所“揭榜挂帅”项目 (2023CXMMTCM003)

predict the toxicity and oral bioavailability of *A. argyi* compounds, identifying 14 components that are non-toxic and have good oral bioavailability. The synthetic accessibility score (SAscore) model was used to analyze the synthetic accessibility of the 14 components mentioned above, and molecular segments were fragmented using BRICS and RECAP algorithms. Mining of the STP and PM databases yielded 406 target proteins for the core components of *A. argyi*, and Cytoscape was used to screen out 5 core targets: SRC, EGFR, PTPN11, HRAS, and PDGFRB. GO and KEGG enrichment analyses indicated that the core targets were involved in 808 GO enrichment analysis entries and 71 signaling pathways, including EGFR tyrosine kinase inhibitor resistance, gap junction, phospholipase D, and JAK/STAT. Molecular docking results showed that active compounds of *A. argyi* have a good binding affinity with proteins SRC, EGFR, PTPN11, and HRAS. Cellular experiments have confirmed that ledol, an active component of *A. argyi*, can promote the proliferation of HUVEC cells within a certain concentration range and can increase the expression of EGFR protein. This study reveals the pharmacological characteristics and potential molecular mechanisms of the active components of *A. argyi* and lays a solid scientific foundation for its medicinal development.

Key words artificial intelligence; *Artemisia argyi*; active ingredients; pharmacological properties; molecular mechanisms

This study was supported by the University Synergy Innovation Program of Anhui Province (GXXT-2020-025), and the “Open Bidding for Selecting the Best Candidates” Project of Xin’an Medicine and Modernization of Traditional Chinese Medicine of Institute of Health and Medicine of Hefei Comprehensive National Science Center(2023CXMMTCM003)

艾叶, 菊科植物艾 *Artemisia argyi* Lévl.et Vant. 的干燥叶, 作为我国传统民俗植物和中药材应用历史悠久。早在东汉时期, 医圣张仲景所著《金匱要略》中就记载了使用艾叶的中药复方芎归胶艾汤和柏叶汤并为众医家沿用至今^[1]。艾叶性温, 味辛、苦, 归肝、脾、肾经, 具有温经止血, 散寒止痛, 外用祛湿止痒的功效, 多用于治疗吐血、衄血等出血证和崩漏、胎漏下血、月经不调、少腹冷痛等妇科疾病; 外用可治疗皮肤瘙痒, 在临床上应用广泛。现代药理学研究亦表明, 艾叶含有包括挥发油、黄酮类、多糖、鞣酸、萜类等多种化学成分, 在抗病毒、抗氧化、抗菌、抗炎及降血脂、免疫调节等方面均表现出极大的药用价值^[2]。基于此, 艾叶被认为具有作为天然药物开发与研究的巨大潜力。

中药化合物中活性成分的多样性与中医特色的整体观念与辨证论治理论使得阐释中医药治疗疾病的科学内涵尤为困难。如何从海量的数据中快速挖掘有效的数据资源, 探索中药的整体物质基础与其在分子和系统水平上的复杂关系, 是揭示中药独特机制与治疗潜力的关键。在传统中医药领域的研究中, 由于中药成分的多样性加之提取工艺的复杂性使其在国际学术界缺乏认可。人工智能技术的高速发展为解开中医之谜提供了一条快速、高效、廉验的有效途径^[3]。本文利用人工

智能技术对艾叶的疗效特征 (中药质量、化合物毒性及口服性、可合成性、分子片段分割及药物靶点等) 进行评估, 通过使用计算机多模型算法协助探索艾叶中的先导化合物和药物作用机制, 为传统中药材中物质基础的研究及天然药物活性成分的分析与开发提供新思路。

1 材 料

1.1 试 剂

DMEM 高糖培养基、胎牛血清 (FBS, 美国 Gibco 公司); CCK-8 试剂盒 (南京诺唯赞公司); 杜香醇 (ledol, 纯度 ≥95%, 上海阿拉丁公司); 二甲基亚砜 (DMSO, 北京索莱宝公司); 抗 GAPDH 单克隆抗体、抗 EGFR 单克隆抗体 (武汉 Proteintech 公司); 辣根过氧化物酶 (HRP) 标记山羊抗小鼠 IgG (美国 Seracare 公司)。

1.2 仪 器

PP-1105 电泳仪 (北京凯元信瑞仪器有限公司); Multifuge X1R 台式冷冻离心机、Multiskan Sky 酶标仪 (美国 Thermo 公司); JS-1070P 化学发光仪 (上海培清科技有限公司)。

1.3 细 胞

永生化人脐静脉内皮细胞 (HUVEC) 购自镇江爱必梦科技有限公司, 编号 T0014。

2 方法

2.1 活性成分的收集

以“艾”“艾叶”或“艾草”为关键字,在 Herbal Ingredients' Targets Platform(HIT: <http://hit2.baddcao.net/>)数据库、Traditional Chinese Medicine Systems Pharmacology (TCMSP: <https://tcmisp-e.com/>)数据库以及 Traditional Chinese Medicine on Immuno-Oncology (TCMIO: <http://tcmio.xielab.net/>)数据库中进行检索,合并数据库结果后删除重复数据得到艾叶的活性成分。同时,通过 PubChem 数据库(<https://pubchem.ncbi.nlm.nih.gov/>)获取艾叶活性成分的 3D sdf 结构格式文件,对于部分缺少 3D sdf 结构格式文件的成分使用 2D sdf 结构格式文件替代。

2.2 药物毒性及口服利用度分析

基于 VenomPred(<https://www.mmvsl.it/wp/venompred2/>)中的 K 最近邻(k-nearest neighbor, KNN)、多层感知机(multilayer perceptron, MLP)、支持向量机(support vector machine, SVM)、随机森林(random forest, RF)算法分别对艾叶活性成分的肝毒性、口服毒性、诱变毒性、致癌性、雌激素活性及雄激素活性计算毒性得分。当毒性得分<50,该活性成分无毒,毒性得分>50,该活性成分有毒^[4]。基于 Lipinski 规则及 Veber 规则构建模型对艾叶活性成分的口服利用度进行计算。两种规则均满足的情况下,将该成分视为具有良好的口服利用度^[5]。使用“UpSet”R 包将满足无毒与口服利用度较好的活性成分取交集,作为核心化合物成分进行后续分析。

2.3 艾叶核心化合物成分的可合成性分析

基于合成可行性分数(SAScore)计算模型对核心化合物成分的 Smiles 值进行计算^[6]。SAScore 模型通过寻找分子环状系统的复杂性对分子进行评分。当分子中存在如中环,螺旋骨架和许多不对称中心的复杂结构时,可合成性评分将会降低,同时多个立体中心分子的存在也会导致合成复杂性显著变大,降低可合成性评分。SAScore 模型计算结果在 0~1 之间,分数越高代表可合成性越高^[7-8]。

2.4 艾叶核心成分的分子片段分割

使用基于化学子结构的分子断裂(breaking retaining informed chemical substructures, BRICS)算法及逆合成化学分析(retrosynthetic chemical analysis procedure, RECAP)算法对核心化合物成分进行分子片段分割,评估是否具有分解成小分子碎片作为

新药开发的可能。在 Python 环境下使用 RECAP 算法遍历核心化合物的所有叶子节点化合物,同时,使用 BRICS 算法对核心化合物进行分割,获取小分子片段。

2.5 艾叶核心化合物成分的靶点预测及核心基因筛选

使用 Swiss Target Prediction(STP, <http://www.swisstargetprediction.ch/>)数据库对核心化合物成分进行靶点预测,对于无法预测的成分使用 PharmMapper(PM, <http://lilab-ecust.cn/pharmmapper/>)数据库进行辅助预测加以补充。通过 Uniprot(<https://www.uniprot.org/>)数据库对预测得到的所有蛋白质靶标进行基因名标准化注释。去除重复后,将靶点导入 String(<https://version-12-0.string-db.org/>)数据库进行 PPI 网络构建,物种设置为“Homo sapiens”,设置最低相互作用得分为最高置信度(0.9),得到基因集群。利用 Cytoscape 中的 MCODE 插件对基因集群进行拓扑属性分析,筛选出核心蛋白质靶点,进行后续分析。

2.6 核心靶点的 GO 功能与 KEGG 通路富集分析

使用 R 软件中的“org.Hs.eg.db”包对筛选出的艾叶核心靶点进行基因本体(Gene Ontology, GO)功能以及京都基因与基因组百科全书(Kyoto Encyclopedia of Genes and Genomes, KEGG)通路富集分析,设置筛选阈值 $P < 0.05$ 。使用 R 软件的“circlize”包绘制 GO 富集分析结果,“ggplot2”绘制 KEGG 富集分析结果。

2.7 分子对接

从 RCSB(<https://www.rcsb.org/>)数据库中下载 SRC(id: 4mxo)、EGFR(id: 4zse)、HRAS(id: 5b2z)、PTPN11(id: 8t6g)以及 PDGFRB(id: 3mjg)的蛋白质 pdb 结构。将艾叶核心化合物成分的 sdf 结构转换为 mol2 结构。在 Linux 环境下使用 MGTools 进行去水加氢,并对靶点与艾叶核心成分之间进行虚拟筛选计算结合能。一般而言,结合能小于-29.3 kJ/mol 表明具有较好的亲和力^[9]。使用 Pymol 显示分子对接情况、分子之间相互作用分析及预测受体与配体之间的结合模式及亲和力。

2.8 细胞实验

2.8.1 细胞培养 HUVEC 细胞在含 10%FBS 的高糖 DMEM 培养基中 37 °C, 5%CO₂ 培养箱中培养,待细胞密度长至 90% 以上时按 1:2 进行传代,取传至第 3~5 代细胞进行后续实验。

2.8.2 CCK-8 检测细胞增殖率 取 HUVEC 细胞制成每毫升 2×10^4 个细胞的悬液接种于 96 孔板中, 每孔 100 μL , 待 24 h 细胞贴壁后将其分为对照组 (0 $\mu\text{mol/L}$ 杜香醇) 以及实验组 (杜香醇终浓度分别为 5、10、15、20、30、50 $\mu\text{mol/L}$)。空白组仅加入培养基。每组设置 5 个复孔, 继续培养 24 h 后吸弃培养基后加入 PBS 100 μL 和 CCK-8 溶液 10 μL , 培养箱中孵育 2 h 后酶标仪检测 450 nm 处的吸收度。

2.8.3 Western Blot 检测蛋白表达 取 HUVEC 细胞, 将其分为对照组与杜香醇组, 分别接种至 10 cm 培养皿中, 待细胞贴壁后, 对照组原培养条件下继续培养, 杜香醇组加入终浓度为 5 $\mu\text{mol/L}$ 杜香醇的培养基干预, 两组均干预 24 h 后胰酶消化收集两组细胞沉淀, 按照文献 [10] 报道方法检测杜香醇对细胞中 EGFR 蛋白表达的影响。实验中使用的抗体稀释度分别为: 抗 EGFR 抗体 (1:2000)、抗 GAPDH 抗体 (1:10000) 和 HRP 标记的山羊抗小鼠 (1:50000)。使用 Image J 1.8.0 对条带进行灰度分析, 以 GAPDH 为内参, 进行统计分析。

2.9 统计分析

实验数据均使用 SPSS 26.0 进行统计分析, 每组数据以均值 $\pm$ 标准差表示, 两组独立样本之间使用 t 检验, 多样本之间使用单因素方差分析, 以 $P <$

0.05 为差异具有统计学意义。

3 结果

3.1 艾叶活性成分的收集

检索 HIT 数据库共获取 53 种活性成分, TCMSP 数据库中获取 135 种活性成分, TCMIO 数据库中获取 192 种活性成分。3 个数据库取并集后去除重复结果, 一共获得了 199 种活性成分 (见图 1)。

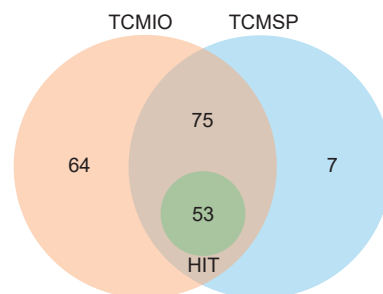


Figure 1 Venn diagram showing the union of *Artemisia argyi* ingredients in TCMIO, TCMSP, and HIT databases (Different colours indicate different databases)

3.2 艾叶活性成分的药物毒性及口服利用度分析

毒性预测模型的结果显示, 满足肝毒性、口服毒性、诱变毒性、致癌性、雌激素活性及雄激素活性预测无毒标准的活性成分个数分别为 30、184、191、119、174 和 194 (如图 2)。综合毒性预测模型及口服利用度模型计算结果, 二者取交集后共得到 14 种核心成分 (如图 3)。详见表 1。

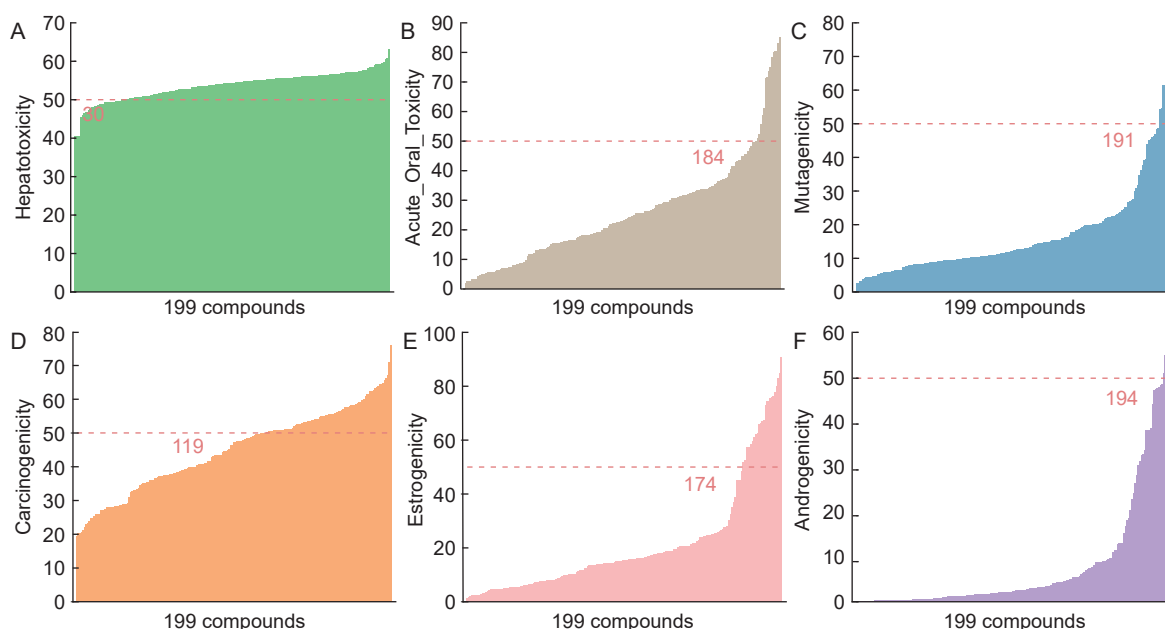


Figure 2 Six toxicity scores of the ingredients of *Artemisia argyi*

A: Hepatotoxicity; B: Acute oral toxicity; C: Mutagenicity; D: Carcinogenicity; E: Estrogenicity; F: Androgenicity. The dotted line marks the toxicity threshold; numbers below indicate predicted non-toxic ingredients.

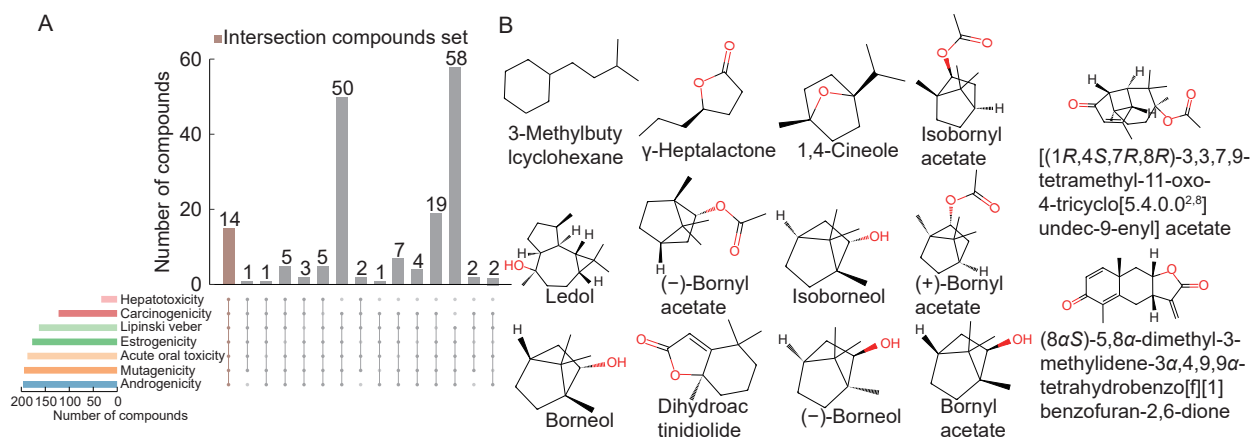


Figure 3 14 core ingredients of *Artemisia argyi* satisfied both toxicity prediction models and oral utilization models

A: Intersection of toxicity prediction models and oral utilization models for active ingredients; B: Chemical structural formulae of the 14 core ingredients

Table 1 Toxicity prediction scores for 14 ingredients in *Artemisia argyi*

Compd.	Molecular formula	Toxicological prediction					
		Mutagenicity	Carcinogenicity	Hepatotoxicity	Estrogenicity	Androgenicity	Acute_Oral_Toxicity
Bornyl acetate	C ₁₂ H ₂₀ O ₂	6.27	27.96	40.48	14.21	2.09	25.42
(-)-Bornyl acetate	C ₁₂ H ₂₀ O ₂	6.27	27.96	40.48	14.21	2.09	25.42
Isobornyl acetate	C ₁₂ H ₂₀ O ₂	6.27	27.96	40.48	14.21	2.09	25.42
(+)-Bornyl acetate	C ₁₂ H ₂₀ O ₂	6.27	27.96	40.48	14.21	2.09	25.42
[(1 <i>R</i> ,4 <i>S</i> ,7 <i>R</i> ,8 <i>R</i>)-3,3,7,9-tetramethyl-11-oxo-4-tricyclo[5.4.0.0 ^{2,8}]undec-9-enyl]acetate	C ₁₇ H ₂₄ O ₃	12.59	39.92	46.69	5.66	18.22	35.56
3-Methylbutylcyclohexane	C ₁₁ H ₂₂	8.15	39.09	48.02	7.67	0.46	14.99
Dihydroactinidiolide	C ₁₁ H ₁₆ O ₂	12.87	40.79	48.40	3.28	6.36	28.68
γ -Heptalactone	C ₇ H ₁₂ O ₂	10.43	36.91	48.84	9.88	0.79	5.54
(8 <i>αS</i>)-5,8 <i>α</i> -dimethyl-3-methylidene-3 <i>α</i> ,4,9,9 <i>α</i> -tetrahydrobenzo[<i>f</i>][1]benzofuran-2,6-dione	C ₁₅ H ₁₆ O ₃	14.29	40.72	48.88	6.49	16.75	49.69
(-)-Borneol	C ₁₀ H ₁₈ O	4.35	28.78	49.21	44.99	13.12	33.69
Isoborneol	C ₁₀ H ₁₈ O	4.35	28.78	49.21	44.99	13.12	33.69
Borneol	C ₁₀ H ₁₈ O	4.35	28.78	49.21	44.99	13.12	33.69
Ledol	C ₁₅ H ₂₆ O	15.05	33.30	49.36	35.04	6.53	28.30
1,4-Cineole	C ₁₀ H ₁₈ O	11.42	36.57	49.53	20.06	1.16	8.76

3.3 艾叶核心成分的可合成性分析

SAscore 模型的结果显示 14 种化合物的可合成性评分均大于 0.5, 说明均具有一定的可合成性。其中 3-甲基丁基环己烷的可合成性评分大于 0.9, 表明具有较高的可合成性。[(1*R*,4*S*,7*R*,8*R*)-3,3,7,9-四甲基-11-氧代-4-三环 [5.4.0.0^{2,8}] 十一碳-9-烯基] 乙酸酯因其化学结构的复杂性导致其可合成性最低, 评分为 0.52, 见表 2。

3.4 艾叶核心成分的分子片段分割结果

通过 BRICS 及 RECAP 两种算法模型对 14 种化合物进行片段分割。结果如图 4 所示, 两种算法显示乙酸异龙脑酯、[(1*R*,4*S*,7*R*,8*R*)-3,3,7,9-四甲基-11-氧代-4-三环 [5.4.0.0^{2,8}] 十一碳-9-烯基] 乙酸酯、(-)-乙酸龙脑酯以及 (+)-乙酸龙脑酯 4 种化合物的分子片段分割的结果表现一致, γ -庚内酯、1,4-桉叶素以及 3-甲基丁基环己烷的两种算法的分割结果

Table 2 SAScores for 14 ingredients in *Artemisia argyi*

Compd.	SAScore	Compd.	SAScore
γ -Heptalactone	0.79	(-)-Borneol	0.64
1,4-Cineole	0.63	Isoborneol	0.64
Ledol	0.64	Dihydroactinidiolide	0.69
(-)-Bornyl acetate	0.66	Isobornyl acetate	0.66
3-Methylbutylcyclohexane	0.91	(+)-Bornyl acetate	0.66
Bornyl acetate	0.66	Borneol	0.64
[(1 <i>R</i> ,4 <i>S</i> ,7 <i>R</i> ,8 <i>R</i>)-3,3,7,9-tetramethyl-11-oxo-4-tricyclo[5.4.0.0 ^{2,8}]undec-9-enyl]acetate	0.52	(8 <i>α</i> S)-5,8 <i>α</i> -dimethyl-3-methylidene-3 <i>α</i> ,4,9,9 <i>α</i> -tetrahydrobenzo[<i>f</i>][1]benzofuran-2,6-dione	0.60

不一致, 这可能是 RECAP 算法基于逆合成, 而 BRICS 算法更注重减少碎片空间的策略所致。杜香醇、龙脑、(-)-龙脑、异冰片、二氢猕猴桃内酯、(8 *α* S)-5,8 *α* -二甲基-3-亚甲基-3 *α* ,4,9,9 *α* -四氢苯并[*f*][1]苯并呋喃-2,6-二酮以及乙酸冰片酯 7 种化合物无法进行片段分割。

3.5 艾叶核心成分的靶点预测及核心靶点筛选

使用 STP 以及 PM 数据库对 14 种艾叶活性化合物进行靶点预测, 其中 STP 数据库发现 393 个靶点, PM 数据库发现 13 个靶点, 共 406 个靶点。Uniprot 对靶点进行标准化注释后去除重复的 211 个靶点, 一共获取 195 个艾叶的作用靶点。将

195 个靶点导入 String 在线数据库中构建 PPI 网络, 筛出 114 个重要靶点。进一步利用 Cytoscape 根据 114 个靶点的连接子图与靶点间密度筛出核心的 5 个作用靶点, 分别为 SRC、EGFR、PTPN11、HRAS、PDGFRB(图 5-A)。

3.6 艾叶核心靶点的 GO 功能与 KEGG 通路富集分析

为了进一步探究艾叶活性成分中核心靶点的功能和作用机制, 对上述 5 个靶点进行了 GO 和 KEGG 富集分析。通过生物过程、细胞组分、分子功能 3 方面进行 GO 富集分析, 共获得 808 项条目, 其中生物过程 727 项, 主要集中于 ERK1 和 ERK2 级联的正向调控、MAPK 级联的正向调控、蛋白质丝氨酸/苏氨酸激酶活性的正向调节和磷脂酶 C 活性的正向调节等; 细胞成分 30 项, 包括膜筏、前缘膜和细胞-基质接点等; 分子功能 51 项, 主要集中于蛋白酪氨酸激酶活化、胰岛素受体结合、钙黏蛋白结合与跨膜受体蛋白酪氨酸激酶活性等, 详见图 5-B。KEGG 富集通路显示共有 71 条具有统计学意义的富集信号通路, 基因计数最高的前 5 条通路分别是 EGFR 酪氨酸激酶抑制剂抵抗、间隙连接、磷脂酶 D 信号通路、JAK/STAT 信号通路和膀胱癌(图 5-C)。

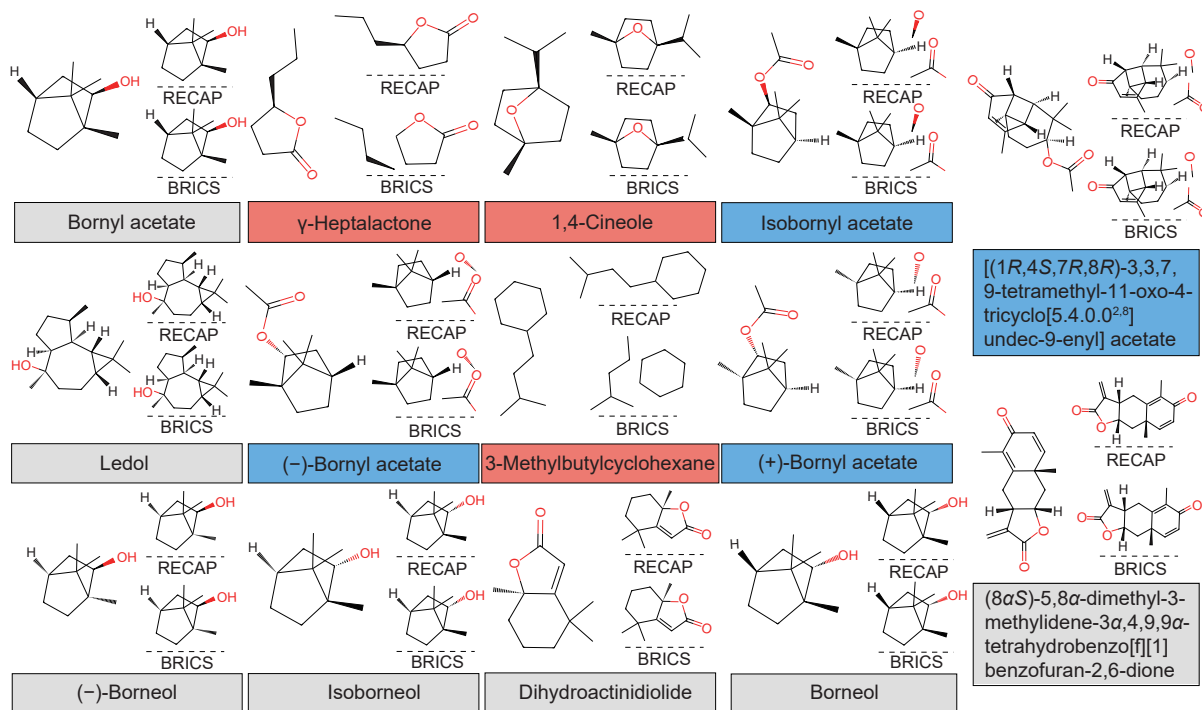


Figure 4 Segmentation results of the molecular fragments in *Artemisia argyi* core ingredients by RECAP and BRICS algorithms. The blue module indicates that both algorithms yield identical segmentation results; the red module indicates that the segmentation results differ between the two algorithms; and the grey module indicates that segmentation could not be performed

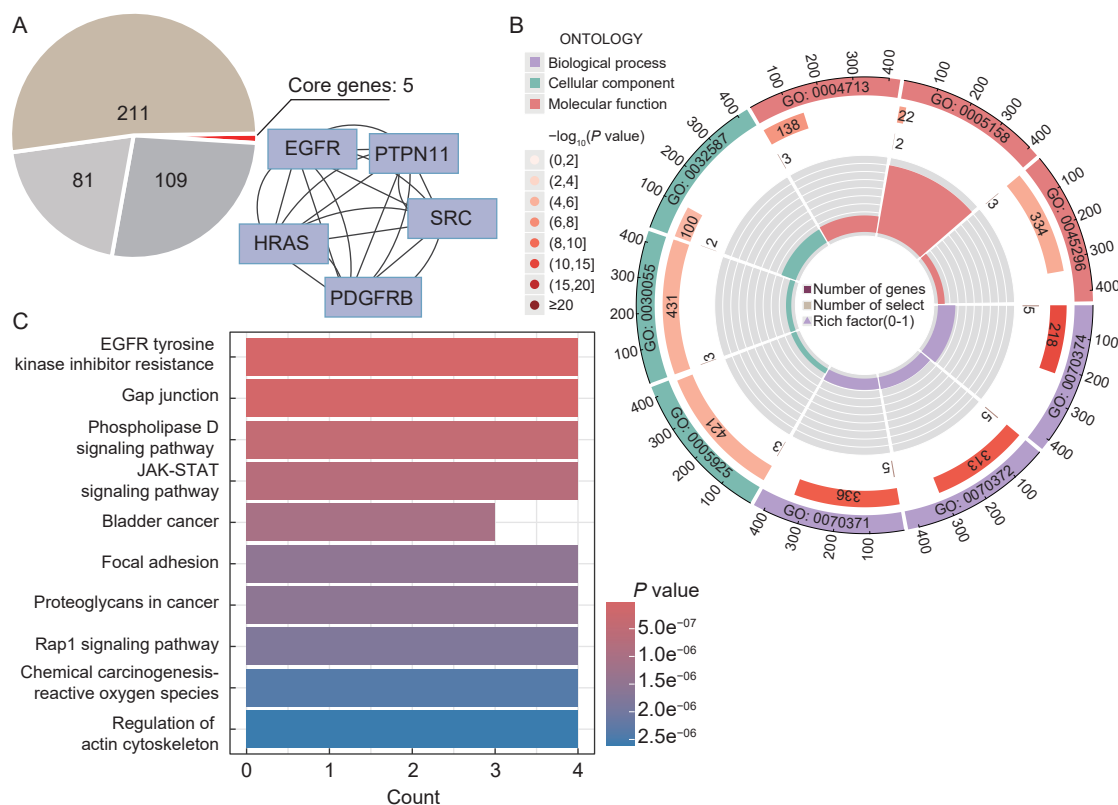


Figure 5 Predicting the target gene and potential mechanism of core *Artemisia argyi* ingredients

A: Core gene screening; B: GO enrichment analysis of core genes; C: KEGG enrichment analysis of core genes

3.7 艾叶核心成分与靶点的分子对接结果

将 SRC、EGFR、HRAS、PTPN11、PDGFRB 与 14 种化合物依次进行分子对接,结果见表 3。将结合能低于 -29.3 kJ/mol 的结果进行可视化,如图 6 所示。EGFR 受体与配体杜香醇通过 ASN-808 形成了氢键维持,键长为 2.0 Å,特定键长有助于受体和配体间的结合,同时 EGFR 与杜香醇的结合能较低,为 -30.98 kJ/mol。HRAS 与 1,4-桉叶素之间通过 TYR-141 形成氢键维持。PTPN11 与杜香醇通过 PRO-101 和 ASN-103 之间形成了不同的

键长,分别是 1.7 Å 和 1.9 Å。SRC 受体与其相关配体均具备稳定构象。

3.8 杜香醇对 HUVEC 的增殖与 EGFR 蛋白表达的影响

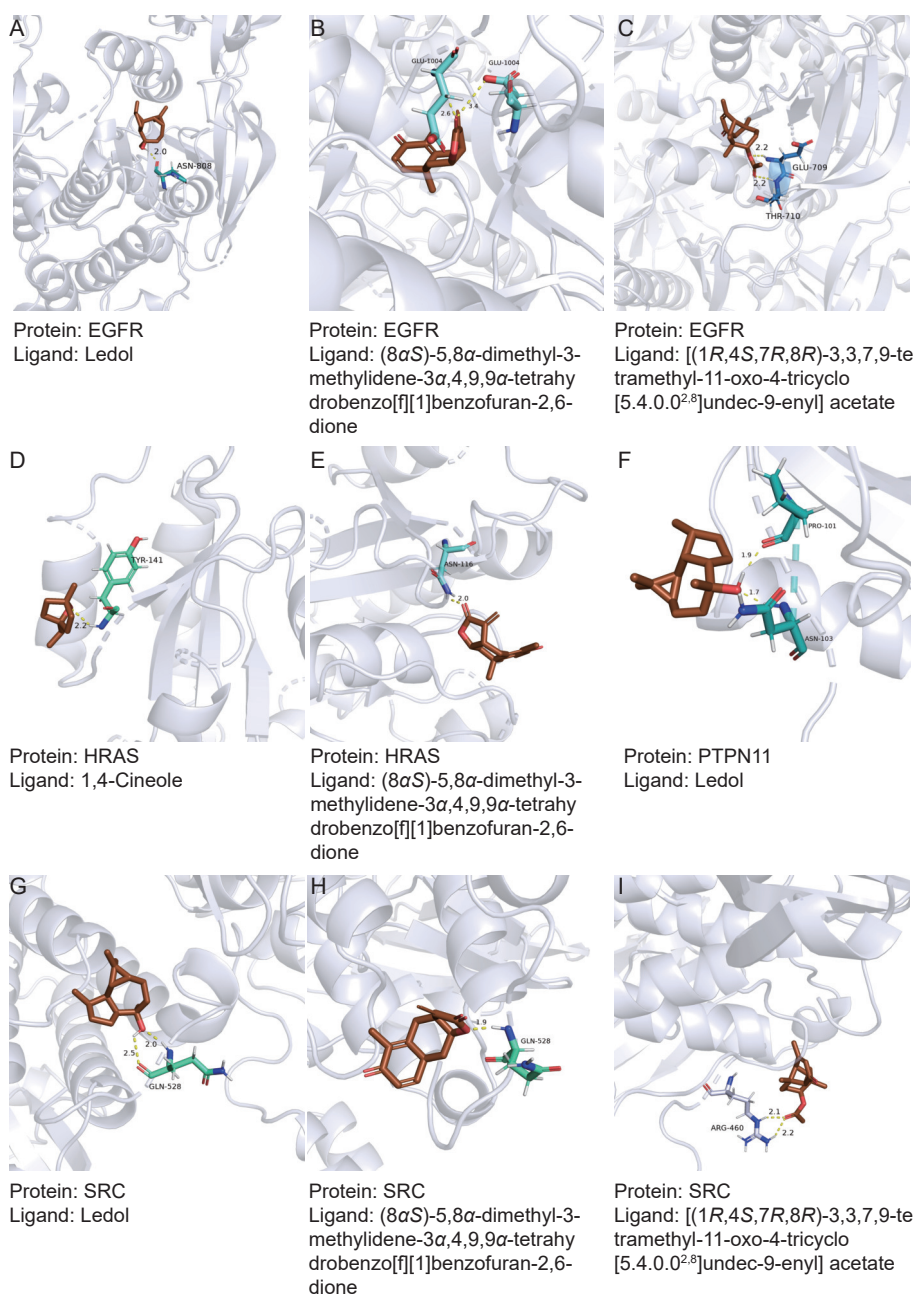
CCK-8 实验结果表明, 5 μmol/L 杜香醇对 HUVEC 的增殖具有促进作用 ($P < 0.001$);当浓度逐渐升高后,这种促增殖作用逐渐减弱,除了 10 μmol/L 组外,余下浓度 (15 、 20 、 30 、 50 μmol/L) 杜香醇均对 HUVEC 的增殖表现出抑制作用 ($P < 0.001$),并随着药物剂量增加而增加(图 7-A)。

Table 3 Molecular docking results of core ingredients and target sites

Component	Binding energy/(kJ/mol)				
	EGFR	HRAS	PDGFRB	PTPN11	SRC
γ-Heptalactone	-25.53	-28.46	-21.77	-22.19	-20.09
1,4-Cineole	-26.79	-29.30	-24.70	-24.70	-24.70
Ledol	-30.98	-28.88	-26.79	-30.98	-31.81
(-)-Bornyl acetate	-23.86	-20.93	-21.77	-22.60	-25.95
3-Methylbutylcyclohexane	-23.44	-27.63	-22.60	-24.70	-23.02
(-)-Borneol	-21.35	-19.25	-26.79	-21.35	-23.86
Isoborneol	-22.60	-25.95	-23.02	-20.93	-23.44
Dihydroactinidiolide	-25.53	-24.70	-21.35	-25.53	-28.88
Isobornyl acetate	-23.86	-25.12	-21.35	-24.28	-26.79

(Continued)

Component	Binding energy/(kJ/mol)				
	EGFR	HRAS	PDGFRB	PTPN11	SRC
(+)-Bornyl acetate	-23.44	-21.35	-22.19	-24.28	-25.12
(8 α S)-5,8 α -Dimethyl-3-methylidene-3 α ,4,9,9 α -tetrahydrobenzo[f][1]benzofuran-2,6-dione	-32.65	-32.23	-24.7	-27.63	-32.23
[(1R,4S,7R,8R)-3,3,7,9-tetramethyl-11-oxo-4-tricyclo[5.4.0.0 ^{2,8}]undec-9-enyl] acetate	-29.30	-25.53	-24.28	-26.37	-29.72
Borneol	-23.02	-20.09	-20.93	-21.77	-22.60
Bornyl acetate	-23.44	-25.95	-23.86	-23.86	-23.86

**Figure 6** Molecular docking results of high binding between 14 ingredients of *Artemisia argyi* and target sites

A-C: EGFR binding to core ingredients; D-E: HRAS binding to core ingredients; F: PTPN11 binding to core ingredients; G-I: SRC binding to core ingredients

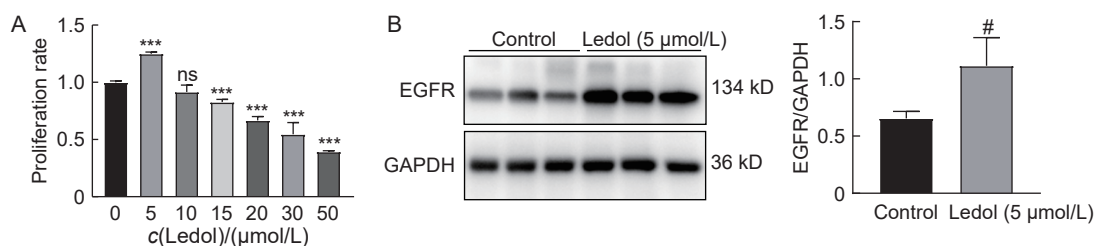


Figure 7 Effect of ledol on proliferation(A) and EGFR expression(B) of HUVEC after incubation for 24 h($\bar{x} \pm s$, $n=3$)
 *** $P<0.001$ vs 0 $\mu\text{mol/L}$ group; # $P<0.05$ vs control group. ns: no significance

进一步采用 5 $\mu\text{mol/L}$ 杜香醇干预 HUVEC 细胞 24 h 后发现, 杜香醇干预能够显著上调 EGFR 的蛋白表达, 差异具有统计学意义($P<0.05$) (图 7-B)。

4 讨论

近年来, 人工智能技术在先导化合物的优化、开发及新药合成等方面研究中取得了巨大进展^[11-12]。艾叶作为中国传统中草药的一种, 在临床治疗中应用广泛, 效果明显^[13-14]。因此, 本研究基于人工智能技术结合多种特征化方法对艾叶中化合物进行全面预测及性能评估, 探究其核心成分在现实世界中的转化意义与价值。

药物毒性的合理预测在药物研发过程中至关重要, 能够降低药物临床试验的高风险^[15]。因此, 本研究利用 4 种算法模型对艾叶的化合物毒性进行预测, 综合筛选出了 17 种无毒性的化合物成分。由于药物口服利用度在新药转化中的重要性, 构建基于 Lipinski 规则及 Veber 规则的预测模型对化合物进行评估^[16]。当两种规则同时满足时被认为该成分具有良好的口服利用度^[5]。最终, 本研究获取了包括异冰片、龙脑、杜香醇、1,4-桉叶素等在内的 14 个艾叶核心化合物成分。

合成性分子评分是判断一个分子是否可以用于药物发现的原则^[17]。本研究中通过模型计算探讨 14 种核心化合物成分是否具有药物设计合成的可能, 发现 14 种核心成分均具有一定的可合成性, 具备药物成分开发的可能^[18]。鉴于模型预测在化合物合成中的可行性往往受限于现有数据库的广度和深度, 因此本研究中模型所提供的计算结果应被视为初步的理论推导, 可用作辅助性的参考框架, 对于这些化合物的具体合成潜力仍需要进行实证性验证以确保其合成过程中的可行性和效率。此外, 基于片段的药物发现是一种新型药物研究策略。本研究通过 BRICS 和 RECAP 两种针对分子

片段分割的算法, 在最大限度地保留分子活性的前提下对化合物进行片段分割, 获取了乙酸异龙脑酯、(+)-乙酸冰片酯以及 1,4-桉叶素在内的 7 种成分的小分子片段。

在本研究中, EGFR 与 EGFR 酪氨酸激酶抑制剂抵抗信号通路是 GO 与 KEGG 富集所得关键靶点及通路, 分子对接结果显示 EGFR 与艾叶的核心活性成分杜香醇具有良好的结合活性。EGFR 作为表皮生长因子受体家族成员之一, 可与其配体 AREG 在细胞膜结合, 发挥调控细胞存活、增殖、运动的作用, 提高人绒毛外滋养层细胞的侵袭能力^[19]。缺氧环境下, EGFR 可促进 HIF-1 α 表达, 从而调控血管生成和子宫内膜容受性, 在不孕症中发挥重要作用。血管内皮细胞是循环血液及管壁内皮下之间的一种单层细胞, 是研究生理病理性血管生成体外实验的主要载体^[20]。人脐静脉内皮细胞(HUVECs)是体外研究血管内皮细胞的重要来源, 具有来源充足、获取方便, 与来源于动物的内皮细胞相比, 实验条件和结果更符合人体情况等优势^[21]。因此本研究中研究了杜香醇对 HUVEC 增殖的影响, 发现低浓度的杜香醇能够发挥促进细胞增殖的作用。Xue 等^[22]通过构建胚胎着床性不孕症的小鼠模型, 观察了调经促孕方可通过 AREG/EGFR/HIF-1 α 通路改善子宫内膜血流灌注, 恢复子宫内膜相对缺氧环境, 从而促进小鼠子宫内膜腺体和血管发育, 增加胞饮突数量, 改善子宫内膜容受性, 利于胚胎黏附及植入来发挥治疗不孕症的作用。因此, 本研究进一步检测了杜香醇对 EGFR 蛋白表达的影响, 发现杜香醇能够促进 EGFR 的蛋白表达, 说明 EGFR 可能是艾叶活性成分杜香醇发挥促血管生成作用的重要靶点。

综上所述, 本研究通过多种人工智能算法对艾叶的成分、药物毒性、口服利用度、可合成性和作用靶点及通路进行了全方位的预测与评估, 为后期

艾叶中物质基础的研究及药物活性成分的开发及转化提供了充足的前期依据。但本研究仍存在一定的局限性, 算法的泛化能力不足以及生物学系统的多样性均可能导致计算结果出现偏差, 因此想要实现新药的开发与再利用后续仍需要更多的体内外实验进一步验证结果的可靠性。

References

- [1] Wang JL, Guo QT, Zhang J, *et al.* *Artemisia argyi* and its application in traditional Chinese medicine agriculture[J]. *Mod Horticult*(现代园艺), 2024, **47**(17): 80-82.
- [2] Erdenebileg S, Kim M, Nam Y, *et al.* *Artemisia argyi* ethanol extract ameliorates nonalcoholic steatohepatitis-induced liver fibrosis by modulating gut microbiota and hepatic signaling[J]. *J Ethnopharmacol*, 2024, **333**: 118415.
- [3] Zhang XJ. Artificial intelligence-driven pharmaceutical industry upgrade, capital rushes to AI-powered drug discovery and development[N]. *Econic Inform Daily*(经济参考报), 2024-06-05(003).
- [4] Di Stefano M, Galati S, Piazza L, *et al.* VenomPred 2.0: a novel *in silico* platform for an extended and human interpretable toxicological profiling of small molecules[J]. *J Chem Inf Model*, 2024, **64**(7): 2275-2289.
- [5] Sharma S, Pandey KM. Computational bioprospecting of phyto-constituents as potential inhibitors for peptide deformylase from *Streptococcus oralis*: an opportunistic pathogen[J]. *Arch Biochem Biophys*, 2024, **758**: 110079.
- [6] Kerstjens A, De Winter H. Molecule auto-correction to facilitate molecular design[J]. *J Comput Aided Mol Des*, 2024, **38**(1): 10.
- [7] Calvi A, Gaudin T, Miketa D, *et al.* Leap: molecular synthesis-ability scoring with intermediates[EB/OL]. 2024: 2403.13005. <https://arxiv.org/abs/2403.13005v2>.
- [8] Skoraczynski G, Kitlas M, Miasojedow B, *et al.* Critical assessment of synthetic accessibility scores in computer-assisted synthesis planning[J]. *J Cheminform*, 2023, **15**(1): 6.
- [9] Li JY, Liu T, Sun F, *et al.* Therapeutic effect and mechanism of hordenine on ovalbumin-induced allergic rhinitis in rats[J]. *J China Pharm Univ*(中国药科大学学报), 2025, **56**(1): 80-90.
- [10] Xu MM, Ren L, Fan JH, *et al.* Berberine inhibits gastric cancer development and progression by regulating the JAK2/STAT3 pathway and downregulating IL-6[J]. *Life Sci*, 2022, **290**: 120266.
- [11] Jiang ZY, You QD. Assistance of artificial intelligence in new drug development[J]. *J China Pharm Univ*(中国药科大学学报), 2024, **55**(3): 281-283.
- [12] Zhu JQ, Zhou X, Hou F, *et al.* DLGCN: Prediction of drug-lncRNA associations based on graph convolution network[J]. *Chin J Bioinform*(生物信息学), 2024, **22**(2): 93-100.
- [13] Jiang XJ, Chen YH, Song Z, *et al.* Research progress in the pharmacology and mechanism of *Artemisia argyi*[J]. *Acta Medicinæ Sinica*(华夏医学), 2023, **36**(6): 182-188.
- [14] Zhang XF, Zeng YX, Zhou LP, *et al.* Preliminary study on optimization of extraction process and antioxidant activity of total flavonoids from *Artemisia argyi* leaves[J]. *Ind Microbiol*(工业微生物), 2024, **54**(1): 134-136.
- [15] Zhang XY, Chen H, Wang Z, *et al.* Status of application of machine learning algorithm in drug toxicity prediction[J]. *Chin J New Drug*(中国新药杂志), 2024, **33**(2): 110-117.
- [16] Haritha M, Sreerag M, Suresh CH. Quantifying the hydrogen-bond propensity of drugs and its relationship with Lipinski's rule of five[J]. *New J Chem*, 2024, **48**(11): 4896-4908.
- [17] Zeng H, Wu GZ, Zou WX, *et al.* Optimization of Menin inhibitors based on artificial intelligence-driven molecular factory technology[J]. *J China Pharm Univ*(中国药科大学学报), 2024, **55**(3): 326-334.
- [18] Wen HQ, Sun QH, Shen WF. Targeted intelligent molecular generation framework based on fragments chemical space[J]. *Ciesc J*(化工学报), 2024, **75**(4): 1655-1667.
- [19] Wang JJ, Cui WF, Dou XW, *et al.* *Euonymus alatus* delays progression of diabetic kidney disease in mice by regulating EGFR tyrosine kinase inhibitor resistance signaling pathway[J]. *J S Med Univ*(南方医科大学学报), 2024, **44**(7): 1243-1255.
- [20] Cheng XD, Zhou H, Zhou Y, *et al.* M2 macrophage-derived exosomes inhibit apoptosis of HUVEC cell through regulating miR-221-3p expression[J]. *Biomed Res Int*, 2022, **2022**: 1609244.
- [21] Xu WM, Qian JJ, Xu Q, *et al.* Proteins and signaling pathways response to *Wenjingtongluo* drug-contained serum in IHU-VECs: an explorative proteomic study[J]. *Cell Mol Biol (Noisy-le-grand)*, 2023, **69**(6): 116-124.
- [22] Xue H, Dong L, Yang PS, *et al.* Effects of *Tiaojing Cuyun* prescription on endometrial receptivity and AREG/EGFR/HIF-1 α signaling pathway in embryo implantation dysfunction infertility mice[J]. *Chin J Inform TCM*(中国中医药信息杂志), 2024, **31**(10): 74-80.