

新型苯并呋喃衍生物作为LSD1抑制剂的QSAR研究

张渝琪^{1*}, 郭婧¹, 邱明秀², 刘家妮¹, 王艳¹, 朱春阳¹, 赵淑芬^{1#}, 邱文生^{1#}

¹青岛大学附属医院肿瘤科, 山东 青岛

²青岛市市立医院呼吸与危重症医学科, 山东 青岛

收稿日期: 2023年4月28日; 录用日期: 2023年5月21日; 发布日期: 2023年5月31日

摘要

组蛋白赖氨酸特异性去甲基酶1 (LSD1)是一种黄素腺嘌呤二核苷酸(FAD)依赖性胺氧化酶, 可特异性识别H3K4和H3K9底物, 并去除其单甲基或二甲基修饰。它介导许多细胞内信号通路, 与肿瘤的发生和发展密切相关。因此, 开发高效、特异的LSD1抑制剂不仅有利于研究LSD1的生物学功能, 而且对抗肿瘤药物的开发具有重要的科学意义。建立定量构效关系(QSAR)模型可以预测分子的物理和化学性质。在本研究中, 利用基因表达编程(GEP)建立了一个具有描述符的非线性定量构效关系(QSAR)模型, 并预测了一系列新型苯并呋喃化疗药物的活性。这些描述符是在CODESSA软件中计算的, 并基于启发式算法从描述符池中选择。选择4个描述符来建立多元线性回归模型。获得了训练集和测试集的最佳非线性QSAR模型, 相关系数分别为0.92和0.80, 平均误差分别为0.07和0.60。显然, 基于GEP的QSAR模型具有更好的抑制剂疗效预测稳定性。这些发现对LSD1抑制剂作为高选择性的一流临床候选药物的设计提供了新的价值。

关键词

组蛋白赖氨酸特异性去甲基酶1 (LSD1), 定量构效关系(QSAR), 启发式算法, 基因表达编程(GEP)

QSAR Study of Novel Benzofuran Derivatives as Potent LSD1 Inhibitors

Yuqi Zhang^{1*}, Jing Guo¹, Mingxiu Qiu², Jiani Liu¹, Yan Wang¹, Chunyang Zhu¹, Shufen Zhao^{1#}, Wensheng Qiu^{1#}

¹Oncology Department, Affiliated Hospital of Qingdao University, Qingdao Shandong

²Department of Respiratory and Critical Care Medicine, Qingdao Municipal Hospital, Qingdao Shandong

*第一作者。

#通讯作者。

文章引用: 张渝琪, 郭婧, 邱明秀, 刘家妮, 王艳, 朱春阳, 赵淑芬, 邱文生. 新型苯并呋喃衍生物作为 LSD1 抑制剂的 QSAR 研究[J]. 临床医学进展, 2023, 13(5): 8769-8781. DOI: 10.12677/acm.2023.1351226

Abstract

Histone lysine specific demethylase 1 (LSD1) is a flavin adenine dinucleotide (FAD) dependent amine oxidase, which can specifically recognize H3K4 and H3K9 substrates and remove their monomethyl or dimethyl modifications. It mediates many intracellular signal pathways and is closely related to the occurrence and development of tumors. Therefore, the development of efficient and specific LSD1 inhibitors is not only conducive to the study of the biological function of LSD1, but also has important scientific significance for the development of anti-tumor drugs. Establishing a quantitative structure-activity relationship (QSAR) model can predict the physical and chemical properties of molecules. In this study, gene expression programming (GEP) was used to build a nonlinear quantitative structure activity relationship (QSAR) model with descriptors and to predict the activity of a series of novel DNA-targeted chemotherapeutic agents. These descriptors were calculated in CODESSA software and selected from the descriptor pool based on heuristics. Four descriptors were selected to establish a multiple linear regression model. The best nonlinear QSAR model with a correlation coefficient of 0.92 and 0.80 and mean error of 0.07 and 0.60 for the training and test sets were obtained. It is apparent that the QSAR model based on GEP has better forecasting stability of inhibitor efficacy. These findings should be useful for the design of LSD1 inhibitors as highly selective first-in-class clinical candidate.

Keywords

Lysine Specific Demethylase 1 (LSD1), Quantitative Structure-Activity Relationship (QSAR), Heuristic Method, Gene Expression Programming

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 背景

肿瘤的发生和发展与遗传物质密切相关,即使核苷酸序列没有改变,仍然存在基因表达差异的可能性,从而导致疾病的发生。这种调控机制属于表观遗传学的范畴,主要包括 DNA 甲基化、组蛋白修饰和微小 RNA 的转录后调控等。DNA 附着在组蛋白核小体上,其转录受组蛋白修饰的调节。组蛋白特异性去甲基酶 1 (LSD1)可以调节组蛋白修饰,其异常表达与许多肿瘤疾病的发生和预后密切相关[1]。

组蛋白去甲基化酶 1 (LSD1),又被称作 KDM1A/AOF2,是一种高度保守的腺嘌呤二核苷酸(FAD)依赖性氧化酶[2], LSD1 可以特异性地去除组蛋白 H3 第 4 位赖氨酸(H3K4)和第 9 位赖氨酸(H3K9)的单、二甲基化,从而调节基因转录(包括激活或抑制),而这种表达调控在癌症的发生中起着关键作用,因此 LSD1 的过度表达可能导致抑癌基因的异常沉默[3] [4] [5] [6]。此外,在细胞和动物模型中对 LSD1 的敲除或抑制的研究表明, LSD1 的减少可以导致 H3K4me2 的水平增加,并重新激活肿瘤抑制基因的表达[7] [8] [9]。这些发现表明抑制 LSD1 是治疗癌症的有效策略。目前相关研究已经发现, LSD1 在肺癌、胃癌、结直肠癌、膀胱癌、食管癌、前列腺癌、胶质瘤、急性髓系白血病等癌症疾病中高表达,其表达程度与癌症分级、恶性程度和预后有关[2] [10] [11] [12] [13]。

迄今为止, 在文献中已经发现了许多 LSD1 抑制剂, 并根据其分类作用机制。可以考虑两种不同类型的抑制剂: 1) Tranylcypromine (TCP)及其类似物作为不可逆抑制剂的代表, 其通过与辅因子 FAD 共价结合来抑制 LSD1 的活性; 2) 可逆性抑制剂可以与底物竞争结合活性位点[14]-[26]。据相关研究所知, 苯并咪喃衍生物首次被描述为 LSD1 抑制剂[27]。因此, 可以分析化合物的结构, 为新药的开发提供帮助。

定量构效关系(Quantitative Structure-Activity Relationship, 简称 QSAR)即通过数学方法建立化合物的分子描述符与其生物活性/毒性之间的线性或非线性关系模型, 在分子水平阐明结构与生物学及物理化学特性之间的关系。近十几年来, 随着分子生物学和计算机科学的快速发展, 使 QSAR 研究提高到了一个新的水平, 计算机辅助药物设计(CADD)在新药的研发中起着越来越重要的作用。QSAR 研究对于设计和筛选生物活性显著的药物, 以及阐明药物的作用机理等均具有指导作用[28] [29] [30] [31]。

建立 QSAR 模型的数学算法有很多, 如逐步回归(SMR)、多元线性回归(MLR)、径向基函数神经网络(RBFNN)、人工神经网络(ANN)和支持向量机(SVM)等。近年来的研究表明, 基因表达编程(GEP)算法具有简单、灵活和高效的优点, 因此 GEP 方法被广泛应用于 QSAR 模型的建立。

本研究旨在建立了一个可靠的 QSAR 模型, 使用 GEP 方法揭示靶向 LSD1 的新型苯并咪喃衍生物的结构与其肿瘤抑制活性之间的关系。与此同时, 该模型可以为进一步探索和预测癌症新靶向药物提供重要的理口基口。

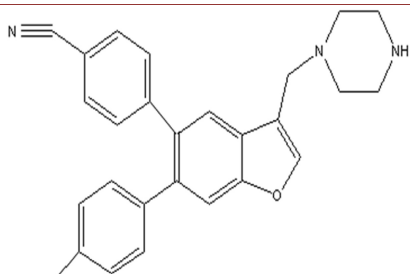
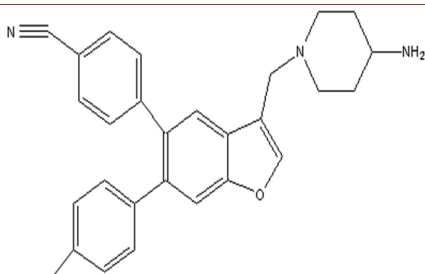
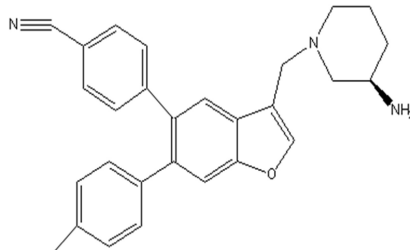
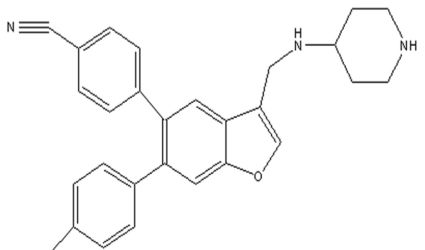
2. 计算与研究方法

2.1. 数据收集及其划分

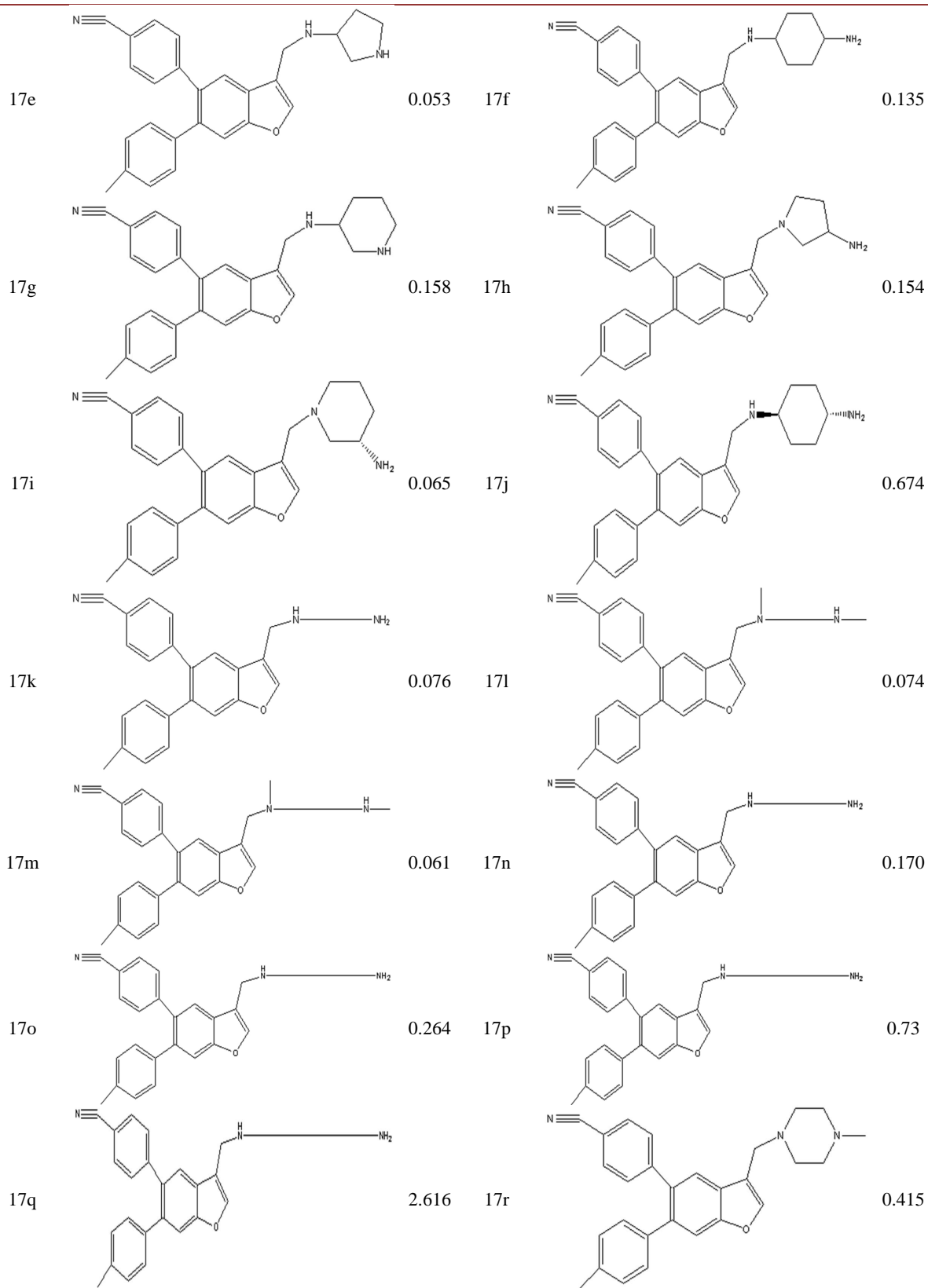
QSAR 研究的第一步是选择合适的数据和建立数据集, 化学物的活性应该在活性范围内均匀分布, 为训练集选择的化学物质应具有足够的结构多样性, 以跨越与所研究的生物活性相关的化学空间范围。从文献[27]获得 38 种化合物和 IC₅₀ 值, 并列于表 1 中。为了获得良好且有意义的结果, 我们使用系统时间作为随机种子; 我们将数据集分为 30 个复合训练集和 8 个复合测试集。训练集用于构建、训练和优化模型, 测试集用于评估模型的预测能力。

Table 1. IC₅₀ of 38 compounds (HM)

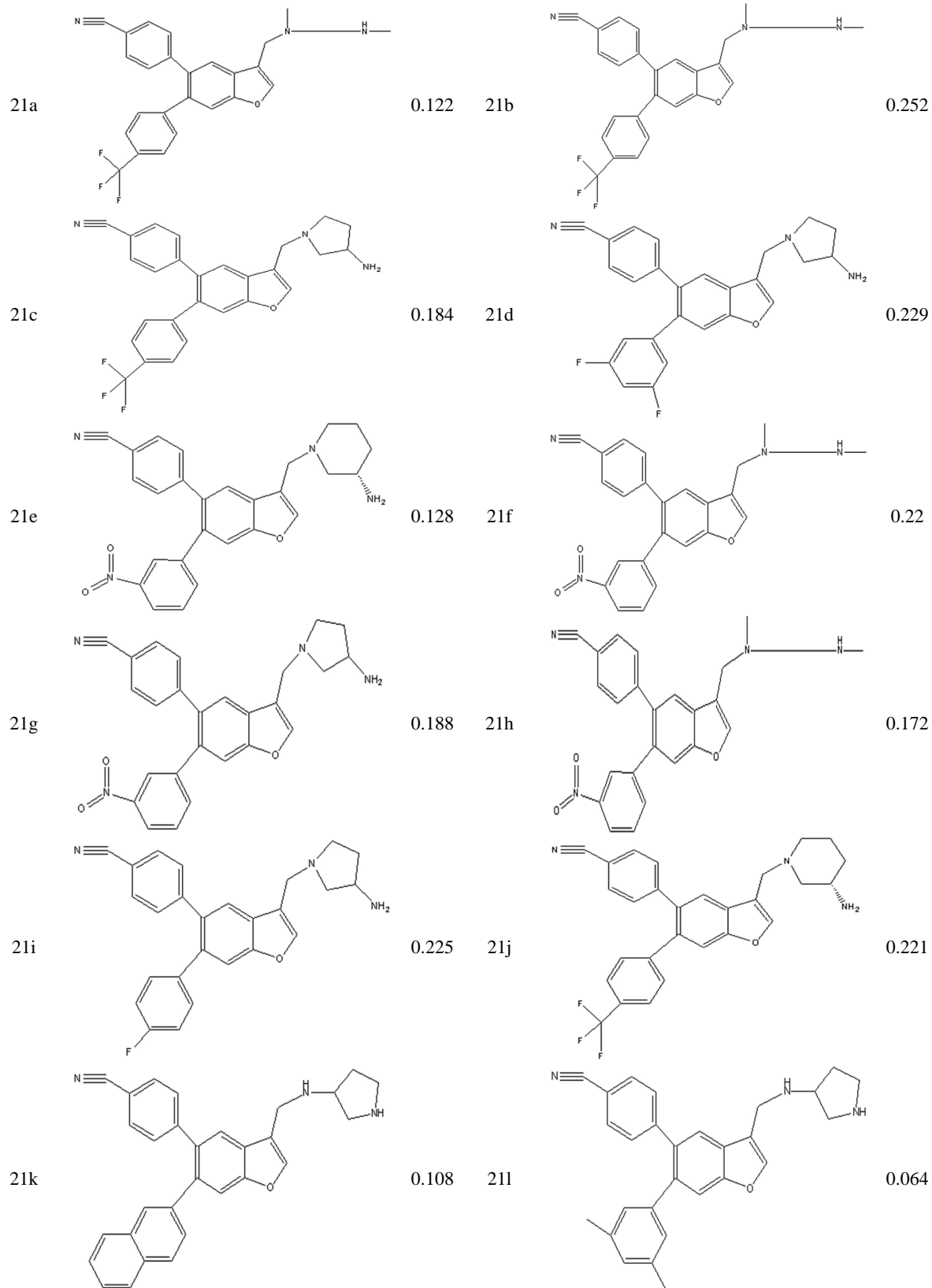
表 1. 38 种化合物的 IC₅₀ 值(HM)

名称	化合物	Ic50	名称	化合物	Ic50
17a		0.135	17b		0.114
17c		0.128	17d		0.42

Continued



Continued



Continued

21m		0.09	21n		1.484
21o		0.5	21p		0.145
21q		2.096	21r		0.147
21s		0.408	21t		9.193

2.2. 描述符的选择与计算

在 QSAR 研究中, 良好的模型的构建必须使用准确合理的分子描述符, 分子描述符测量分子性质的某些方面, 无论是其物理和化学性质, 还是通过基于分子结构的各种算法导出的数值指数。已经报道了各种各样的描述符用于 QSAR 分析, 例如静电、拓扑、几何和量子化学描述符。

在这项研究中, 使用 Chemdraw 软件绘制了化合物的化学结构, 然后将分子结构导入 Hyperchem [32] 软件中进行优化, 实验的最低能量稳态是化合物的最佳状态。目前最流行的半经验量子化学计算方法中使用范围最广泛的方法, 也是精确度、准确度最高的方法是 AM1 (Austin Model 1) 和 PM3 (Parametric Method 3), 为了实现化合物的最佳实验状态, 使用半经验 AM1 或 PM3 方法[33]进行精确优化。用 Polak-Ribiere 算法对分子结构进行优化, 直至均方根梯度达到 0.01, 然后得到三组描述子。

2.3. 启发式线性模型

启发式算法(heuristic method)是二维定量结构 - 活性关系研究常用的方法之一, 在计算机药物辅助设计中也得到广泛应用。通过 CODESSA 软件中的启发式算法可对药物分子结构进行分析, 对计算得到的

大量分子描述符进行筛选,能够快速建立预测药物活性的最佳多元线性方程。对方程中描述符的深入研究,可获得影响药物活性的理化因素,为随后的药物设计提供思路。HM 建立线性关系的步骤模型是:首先,选择单参数描述符,使用 R^2 、F 检验和 T 检验作为标准选择,并删除相关性低的描述符。其次,选择两个参数描述符,并使用 R^2 和 F 检验作为标准选择。最后,选择 n 个参数描述符。在获得具有最佳统计特征的双参数相关系数之后,添加在先前选择过程中未使用的描述符。重复此步骤,直到建立的相关方程包含最多的参数。模型评估的标准为 R^2 、F 检验值、t 检验值和 R^2_{cv} 。因此, HM [34]开发了一个具有 4 个描述符的线性模型。

2.4. GEP 的非线性模型基因表达式编程

遗传表达式编程(Genetic Expression Programming, GEP)是一种用于解决机器学习、数据分析、优化和模式识别等领域问题的进化算法[35]。开发了 GEP 算法来自遗传算法(Genetic algorithm, GAs)和遗传规划算法(Genetic programming, GP) [36],但 GEP 与 GA 和 GP 使用的个体有本质区别。GEP 有两个基本元素,染色体遗传操作子会作用于染色体而不是表达式树,这样保证每一次基因操作都是正确的,不需要检查其是否符合句法[37]。近年来, GEP 算法遗传算法编码具有简单方便的特点操作和灵活性,保留了 GP 表达式搜索能力强大的优势,通过使用不同的基因组和表现型巧妙地避免了 GA 和 GP 的缺点,提高了解决

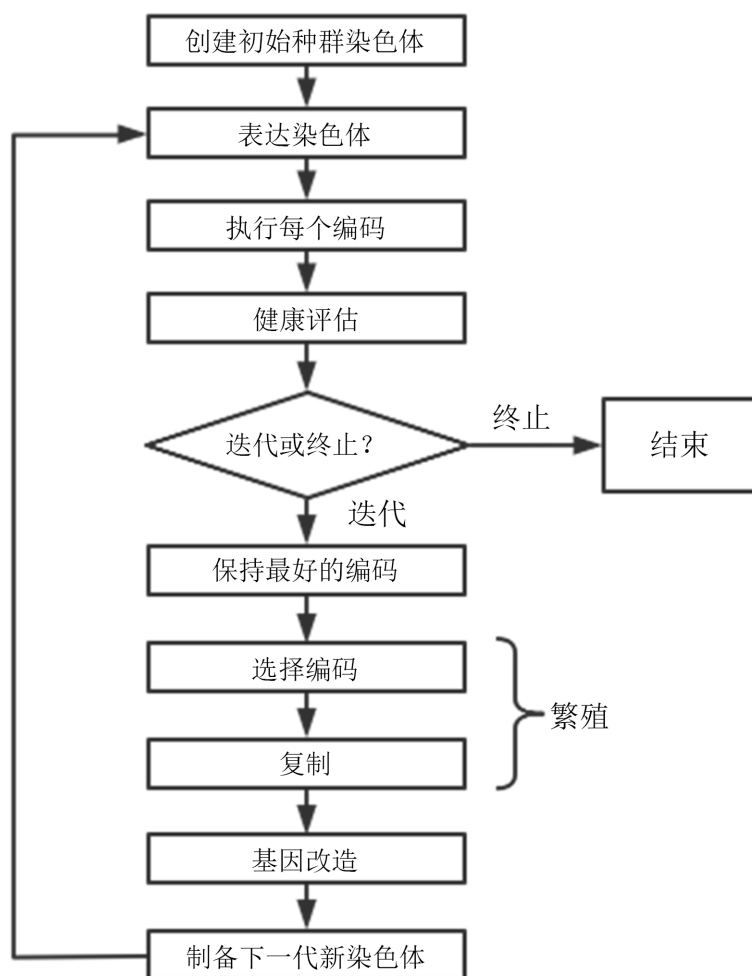


Figure 1. GEP flow chart
图 1. GEP 算法的计算步骤

问题的效率[38] [39]。

在 GEP 中, 基因由头部和尾部组成。标题由终端和功能符号组成, 尾部仅由终端符号组成。假设头部长度为 h , 尾部长度为 t , 关系如下:

$$t = hn - 1 + 1$$

在该方程中, n 表示需要函数符号集中最多变量的函数。显然, 基因的长度可以通过以下公式计算:

$$l = h + t = n * h + 1$$

GEP 算法的主要步骤如图 1 所示。首先, 一定数量的染色体个体是随机的作为初始种群生成并表达。接下来, 从一组适合度样本计算每个个体的适合度。对群体循环执行遗传操作, 直至满足终止准则为止 [40]。否则, 将根据适合度值选择这些个体进行基因操作。最后, 产生具有新特征的后代。重复该过程, 直到出现良好结果。

3. 结果及讨论

3.1. HM 的结果

CODESSA 软件为每个计算了 487 个描述符合物。为了获得与 LSD1 抑制剂活性最相关的一组描述符, 首先建立了描述符数量范围为 1 到 9 的线性回归模型。不同数量的描述符对 R^2 和 S^2 的影响如图 2 所示。结果表明, 随着描述符数量的增加, R^2 增加, S^2 减少。在描述符数量增加到 5 之后, R^2 的增加速率和 S^2 的降低速率略有下降, 并且模型使用过多的描述符的概括能力很差。综合考虑后, 模型具有 4 个描述符被选为预测抑制剂活性的最佳线性模型。

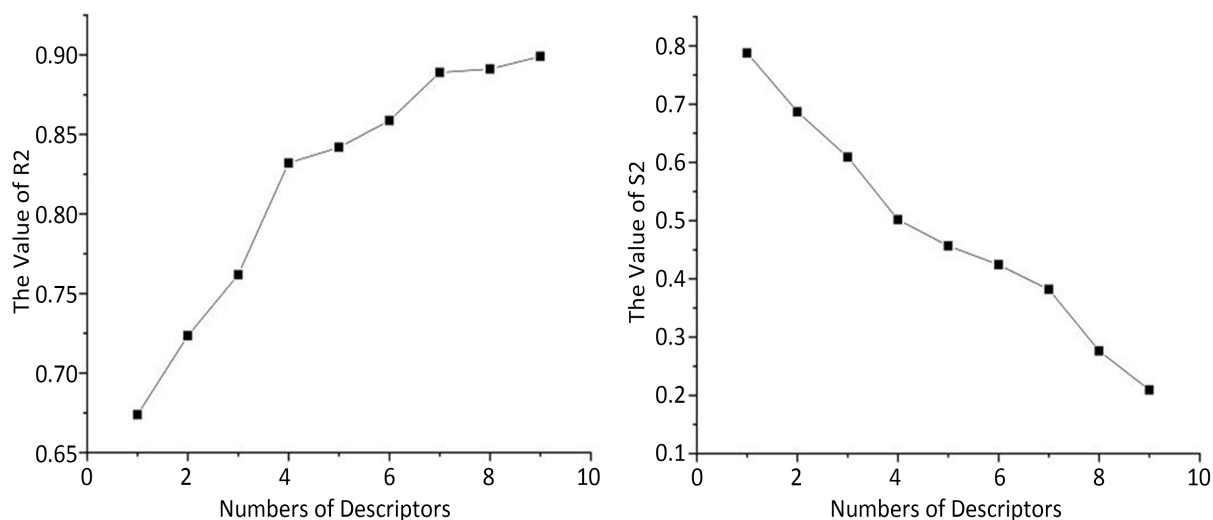


Figure 2. The effects of different numbers of descriptors on R^2 and S^2

图 2. 不同数量的描述符对 R^2 和 S^2 的影响

表 2 显示了这些描述符的详细信息。如前所述, 模型的实用性是最终目的, 但要应用模型就必须保证模型有可靠高效的预测能力。因此本论文从建立 QSAR 模型的各个步骤考虑, 试图解决目前 QSAR 研究中还有待完善的问题, 特别是在化合物活性构象的选择、引入新的建模方法以及新的建模策略等方面, 旨在尽可能考虑 QSAR 的不同角度提高模型的可靠性以及外部预测能力。同时, 把建立的具有很好预测能力的定量构效关系模型用于活性化合物的设计和筛选。

HM 模型的图形如图 3 所示。

Table 2. The selected molecular descriptors and their physical-chemical meaning, coefficient, and T-test
表 2. 选择的分子描述符及其物理 - 化学意义, 系数和 T 检验

分子描述符	物理 - 化学意义	相关系数	T 检验
HACA-2/TMSA	氢键供体原子的表面电荷面积加权表分数	5.8489e+02	10.8596
Relative number of rings	环的相对数量	1.3510e+01	-3.9921
Complementary Information content	补充信息内容	9.8407e-03	3.3387
Min partial charge for a H atom	H 原子的最小部分电荷	5.2309e+01	2.5869

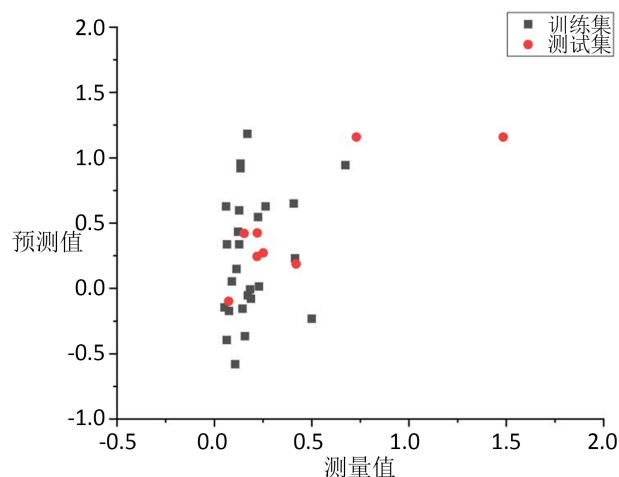


Figure 3. Plot of measured and calculated IC50 by HM
图 3. 通过 HM 测量和计算 IC50 的绘图

线性模型方程如下:

$IC_{50} = 2.3176 + 58.489 \times \text{氢键供体原子的表面电荷面积加权表分数} + 1.351 \times \text{环的相对数量} + 0.0098 \times \text{补充信息内容} + 5.2309 \times \text{H 原子的最小部分电荷}$

根据方程中系数的绝对值, 描述符对 LSD1 抑制剂活性的影响如下: 补充信息内容 > 氢键供体原子的表面电荷面积加权表分数 > H 原子的最小部分电荷 > 环的相对数量。所选的四个描述符是在下文中进行了解释, 以更深入地了解可能影响 LSD1 抑制剂的活性。

HDCA-2/TMSA 表示氢键供体原子 HDCA-2 的面积加权表面电荷; HDCA-2 的计算如下:

$$HDCA2 = \sum_D \frac{q_D \sqrt{S_D}}{\sqrt{S_{tot}}} - D \in H_{H-donor}$$

它对分子的构象变化具有重要意义[41]。最大 HDCA-2/TMSA 系数为正值, 表明 HDCA-2/TSA 值越高, LSD1 抑制剂药物的活性越强。同时, 可以设计具有更高 HDCA-2/TMSA 值的分子来提高 LSD1 抑制剂药物的活性。

Complementary Information content 表示补充信息内容, 计算如下:

$${}^kCIC = \log_2 n - {}^kIC$$

$${}^kIC = -\sum_{i=1}^k \frac{n_i}{n} \log_2 \frac{n_i}{n}$$

Min partial charge for a H atom 和 Relative number of rings 分别表示 H 原子的最小部分电荷, 以及环的

相对数量，他们反映了化合物的分子组成[42]。这三种描述符的正系数表明他们三者值越高，LDS1 抑制剂化疗药物的活性越高。

3.2. GEP 的结果

数据集被随机分成 30 个化合物的训练集和 8 个化合物的测试集，然后是非线性模型是使用软件自动问题求解器(APS)建立的以集成 GEP 算法的实现。为了获得更准确和普遍的模型，我们在 APS 中介绍了六个描述符，并通过 GEP 功能集建立了非线性模型。表 3 列出进化过程中使用的参数。

Table 3. Parameters for the simple symbolic regression problem

表 3. 简单符号回归问题的参数

参数名称	符号	值
乘	*	1
除	/	1
10 ^x	Pow10	1
自然对数	LN	1
绝对值	Abs	1
反函数	Neg	1
余割函数	Csc	1
正割函数	Sec	1

最后，训练集和测试集的相关系数分别为 0.92 和 0.80，并且在 830rd 代中获得了 0.07 和 0.60 的标准误差。图 4 显示了 GEP 模型图。

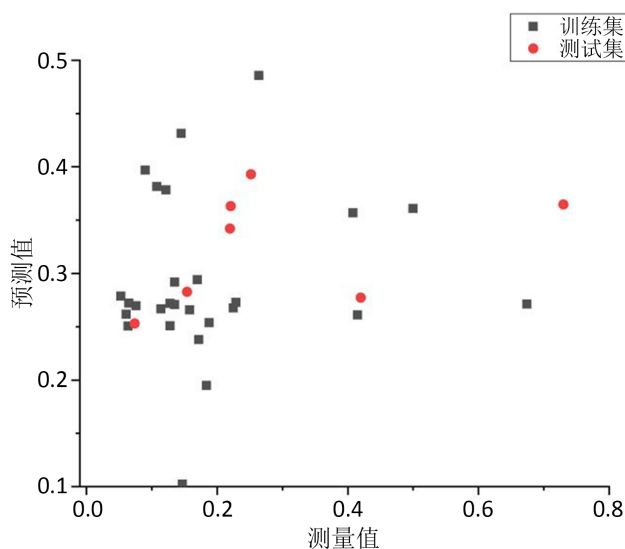


Figure 4. Plot of measured and calculated log (IC50) by GEP

图 4. 通过 GEP 测量和计算的图 IC50

4. 结论

任何 QSAR 建模的主要目标是开发的模型应该足够强大，可以对新化合物的生物活性进行准确和可

靠的预测。在本研究中,由 HM, GEP 建立的非线性模型在预测 38 种 LSD1 抑制剂的活性方面显示出更好的稳定性和预测能力。为了深入了解描述符的重要性,通过基于系数的绝对值对 4 个描述符进行排序,得到描述符对 IC₅₀ 的影响力大小。所建立的模型揭示了影响 LSD1 抑制剂活性的因素,并为进一步设计用于抗肿瘤的高效药物提供了方向和指导。

致 谢

三年时间转瞬即逝,我的肿瘤学硕士研究生阶段的学习即将结束。本论文是在导师邱文生教授和郭婧老师的悉心指导和督促下完成的。导师渊博的专业临床医学知识,雄厚的科研能力,严谨的治学态度,精益求精的工作作风,诲人不倦的高尚师德,严以律己、宽以待人的崇高风范,朴实无华、平易近人的人格魅力对我影响深远。不仅使我树立了远大的学术目标还教会我怎样成为一名优秀的临床医生。

衷心感谢邱文生主任、丁爱萍老师、吕静老师、齐卫卫老师、赵淑芬老师、郭婧老师和我的同门师兄妹三年来对我工作学习上指导与鼓励。

衷心感谢青岛大学附属医院提供的良好科研平台。

基金项目

北京市希思科临床肿瘤学研究基金会(Y-HR2018-185, Y-2019AZZD-0471)。

参考文献

- [1] 倪娜, 袁宜耘, 陈秋荣, 等. 表观遗传学抗肿瘤药物的临床研究进展[J]. 世界临床药物, 2022, 43(7): 831-838.
- [2] Shi, Y., Lan, F., Matson, C., *et al.* (2004) Histone Demethylation Mediated by the Nuclear Amine Oxidase Homolog LSD1. *Cell*, **119**, 941-953. <https://doi.org/10.1016/j.cell.2004.12.012>
- [3] Sorna, V., Theisen, E.R., Stephens, B., Warner, S.L., *et al.* (2013) High-Throughput Virtual Screening Identifies Novel N⁷-(1-Phenylethylidene)-Benzohydrazides as Potent, Specific, and Reversible LSD1 Inhibitors. *Journal of Medicinal Chemistry*, **56**, 9496-9508. <https://doi.org/10.1021/jm400870h>
- [4] Zheng, Y.C., Duan, Y.C., Ma, J.L., *et al.* (2013) Triazole-Dithiocarbamate Based Selective Lysine Specific Demethylase 1 (LSD1) Inactivators Inhibit Gastric Cancer Cell Growth, Invasion, and Migration. *Journal of Medicinal Chemistry*, **56**, 8543-8560. <https://doi.org/10.1021/jm401002r>
- [5] Vianello, P., Botrugno, O.A., Cappa, A., *et al.* (2016) Discovery of a Novel Inhibitor of Histone Lysine-Specific Demethylase 1A (KDM1A/LSD1) as Orally Active Antitumor Agent. *Journal of Medicinal Chemistry*, **59**, 1501-1517. <https://doi.org/10.1021/acs.jmedchem.5b01209>
- [6] Mould, D.P., Alli, C., Bremberg, U., *et al.* (2017) Development of (4-Cyanophenyl) Glycine Derivatives as Reversible Inhibitors of Lysine Specific Demethylase 1. *Journal of Medicinal Chemistry*, **60**, 7984-7999. <https://doi.org/10.1021/acs.jmedchem.7b00462>
- [7] Ma, L.Y., Zheng, Y.C., Wang, S.Q., *et al.* (2015) Design, Synthesis, and Structure—Activity Relationship of Novel LSD1 Inhibitors Based on Pyrimidine—Thiourea Hybrids as Potent, Orally Active Antitumor Agents. *Journal of Medicinal Chemistry*, **58**, 1705-1716. <https://doi.org/10.1021/acs.jmedchem.5b00037>
- [8] Zheng, Y.C., Ma, J., Wang, Z., *et al.* (2015) A Systematic Review of Histone Lysine-Specific Demethylase 1 and Its Inhibitors. *Medicinal Research Reviews*, **35**, 1032-1071. <https://doi.org/10.1002/med.21350>
- [9] Niwa, H., Sato, S., Hashimoto, T., Matsuno, K. and Umehara, T. (2018) Crystal Structure of LSD1 in Complex with 4-[5-(Piperidin-4-Yl-methoxy)-2-(p-Tolyl) Pyridin-3-yl] Benzonitrile. *Molecules*, **23**, Article 1538. <https://doi.org/10.3390/molecules23071538>
- [10] 付佳, 曹宇勃, 刘飒. LSD1 抑制剂抑制非小细胞肺癌 A549 细胞侵袭、迁移及上皮间质转化[J]. 现代肿瘤医学, 2022, 30(11): 1957-1961.
- [11] 王洁, 魏谨, 汪明云. LSD1 和 PTEN 在人卵巢癌组织中的表达及相关性[J]. 现代肿瘤医学, 2021, 29(23): 4181-4184.
- [12] Wang, M., Liu, X., Jiang, G., *et al.* (2015) Relationship between LSD1 Expression and E-Cadherin Expression in Prostate Cancer. *International Urology and Nephrology*, **47**, 485-490. <https://doi.org/10.1007/s11255-015-0915-2>
- [13] Paul, S., Ramalingam, S., Subramaniam, D., *et al.* (2014) Histone Demethylases in Colon Cancer. *Current Colorectal*

- Cancer Reports*, **10**, 417-424. <https://doi.org/10.1007/s11888-014-0253-2>
- [14] Wang, X., Zhang, C., Zhang, X., *et al.* (2020) Design, Synthesis and Biological Evaluation of Tetrahydroquinoline-Based Reversible LSD1 Inhibitors. *European Journal of Medicinal Chemistry*, **194**, Article ID: 112243. <https://doi.org/10.1016/j.ejmech.2020.112243>
- [15] Li, Z., Ding, L., Li, Z., *et al.* (2019) Development of the Triazole-Fused Pyrimidine Derivatives as Highly Potent and Reversible Inhibitors of Histone Lysine Specific Demethylase 1 (LSD1/KDM1A). *Acta Pharmaceutica Sinica B*, **9**, 794-808. <https://doi.org/10.1016/j.apsb.2019.01.001>
- [16] Liu, H.M., Suo, F.Z., Li, X.B., You, Y.H., *et al.* (2019) Discovery and Synthesis of Novel Indole Derivatives-Containing 3-Methylenedihydrofuran-2 (3H)-One as Irreversible LSD1 Inhibitors. *European Journal of Medicinal Chemistry*, **175**, 357-372. <https://doi.org/10.1016/j.ejmech.2019.04.065>
- [17] Li, Z.R., Suo, F.Z., Hu, B., *et al.* (2019) Identification of Osimertinib (AZD9291) as a Lysine Specific Demethylase 1 Inhibitor. *Bioorganic Chemistry*, **84**, 164-169. <https://doi.org/10.1016/j.bioorg.2018.11.018>
- [18] Murray-Stewart, T., Woster, P.M. and Casero Jr, R.A. (2014) The Re-Expression of the Epigenetically Silenced E-Cadherin Gene by a Polyamine Analogue Lysine-Specific Demethylase-1 (LSD1) Inhibitor in Human Acute Myeloid Leukemia Cell Lines. *Amino Acids*, **46**, 585-594. <https://doi.org/10.1007/s00726-013-1485-1>
- [19] Kumarasinghe, I.R. and Woster, P.M. (2014) Synthesis and Evaluation of Novel Cyclic Peptide Inhibitors of Lysine-Specific Demethylase 1. *ACS Medicinal Chemistry Letters*, **5**, 29-33. <https://doi.org/10.1021/ml4002997>
- [20] Nowotarski, S.L., Pachaiyappan, B., Holshouser, S.L., *et al.* (2015) Structure-Activity Study for (Bis) Ureidopropyl- and (Bis) Thioureidopropyldiamine LSD1 Inhibitors with 3-5-3 and 3-6-3 Carbon Backbone Architectures. *Bioorganic & Medicinal Chemistry*, **23**, 1601-1612. <https://doi.org/10.1016/j.bmc.2015.01.049>
- [21] Pachaiyappan, B. and Woster, P.M. (2014) Design of Small Molecule Epigenetic Modulators. *Bioorganic & Medicinal Chemistry Letters*, **24**, 21-32. <https://doi.org/10.1016/j.bmcl.2013.11.001>
- [22] Huang, Y., Stewart, T.M., Wu, Y., *et al.* (2009) Novel Oligoamine Analogues Inhibit Lysine-Specific Demethylase 1 and Induce Reexpression of Epigenetically Silenced Genes. *Clinical Cancer Research*, **15**, 7217-7228. <https://doi.org/10.1158/1078-0432.CCR-09-1293>
- [23] Huang, Y., Marton, L.J., Woster, P.M. and Casero Jr, R.A. (2009) Polyamine Analogues Targeting Epigenetic Gene Regulation. *Essays in Biochemistry*, **46**, 95-110. <https://doi.org/10.1042/bse0460007>
- [24] Kumarasinghe, I.R. and Woster, P.M. (2018) Cyclic Peptide Inhibitors of Lysine-Specific Demethylase 1 with Improved Potency Identified by Alanine Scanning Mutagenesis. *Essays in Biochemistry*, **148**, 210-220. <https://doi.org/10.1016/j.ejmech.2018.01.098>
- [25] Holshouser, S., Dunworth, M., Murray-Stewart, T., *et al.* (2019) Dual Inhibitors of LSD1 and Spermine Oxidase. *MedChemComm*, **10**, 778-790. <https://doi.org/10.1039/C8MD00610E>
- [26] Schenk, T., Chen, W.C., Göllner, S., *et al.* (2012) Inhibition of the LSD1 (KDM1A) Demethylase Reactivates the All-Trans-Retinoic Acid Differentiation Pathway in Acute Myeloid Leukemia. *Nature Medicine*, **18**, 605-611. <https://doi.org/10.1038/nm.2661>
- [27] Zhang, X., Huang, H., Zhang, Z., *et al.* (2021) Design, Synthesis and Biological Evaluation of Novel Benzofuran Derivatives as Potent LSD1 Inhibitors. *European Journal of Medicinal Chemistry*, **220**, Article ID: 113501. <https://doi.org/10.1016/j.ejmech.2021.113501>
- [28] Chen, Y., Yang, Y., Wang, F., *et al.* (2006) Crystal Structure of Human Histone Lysine-Specific Demethylase 1 (LSD1). *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 13956-13961. <https://doi.org/10.1073/pnas.0606381103>
- [29] Hansch, C., Hoekman, D. and Gao, H. (1996) Comparative QSAR: Toward a Deeper Understanding of Chemicobiological Interactions. *Chemical Reviews*, **96**, 1045-1075. <https://doi.org/10.1021/cr9400976>
- [30] Funatsu, K., Miyao, T. and Arakawa, M. (2011) Systematic Generation of Chemical Structures for Rational Drug Design Based on QSAR Models. *Current Computer-Aided Drug Design*, **7**, 1-9. <https://doi.org/10.2174/157340911793743556>
- [31] Roy, K. and Kar, S. (2015) Importance of Applicability Domain of QSAR Models. In: Roy, K., Ed., *Quantitative Structure-Activity Relationships in Drug Design, Predictive Toxicology, and Risk Assessment*, IGI Global, Hershey, 32. <https://doi.org/10.4018/978-1-4666-8136-1.ch005>
- [32] Belka, R. and Piwowarczyk, K. (2010) Some Aspects of Using the Hyperchem Software in Study of Ni-C Nanostructures. *Proceedings of SPIE—The International Society for Optical Engineering*, **7745**, 376-382. <https://doi.org/10.1117/12.871985>
- [33] Karelson, M., Lobanov, V.S. and Katritzky, A.R. (1996) Quantum-Chemical Descriptors in QSAR/QSPR Studies. *Chemical Reviews*, **96**, 1027-1044. <https://doi.org/10.1021/cr950202r>

-
- [34] Katritzky, A.R., Sild, S., Lobanov, V. and Karelson, M. (1998) Quantitative Structure—Property Relationship (QSPR) Correlation of Glass Transition Temperatures of High Molecular Weight Polymers. *Journal of Chemical Information and Modeling*, **38**, 300-304. <https://doi.org/10.1021/ci9700687>
- [35] 陆昕为, 蔡之华. 一种改进的 GEP 方法及其在演化建模预测中的应用[J]. 计算机应用, 2005, 25(12): 2783-2786.
- [36] Parsons, R., Forrest, S. and Burks, C. (1993) Genetic Algorithms for DNA Sequence Assembly. *Proceedings—International Conference on Intelligent Systems for Molecular Biology*, **1**, 310-318.
- [37] Teodorescu, L. and Sherwood, D. (2008) High Energy Physics Event Selection with Gene Expression Programming. *Computer Physics Communications*, **178**, 409-419. <https://doi.org/10.1016/j.cpc.2007.10.003>
- [38] Eftekhari, M.M. (2014) Permeability Estimation in Heterogeneous Oil Reservoirs by Multi-Gene Genetic Programming Algorithm. *Journal of Petroleum Science & Engineering*, **123**, 210-206. <https://doi.org/10.1016/j.petrol.2014.07.035>
- [39] Zhang, L., Chen, J., Gao, C., Liu, C. and Xu, K. (2018) An Efficient Model for Auxiliary Diagnosis of Hepatocellular Carcinoma Based on Gene Expression Programming. *Medical & Biological Engineering & Computing*, **56**, 1771-1779. <https://doi.org/10.1007/s11517-018-1811-6>
- [40] Zhong, J., Feng, L. and Ong, Y.S. (2017) Gene Expression Programming: A Survey. *IEEE Computational Intelligence Magazine*, **12**, 54-72. <https://doi.org/10.1109/MCI.2017.2708618>
- [41] Stanton, D.T. and Jurs, P.C. (1990) Development and Use of Charged Partial Surface Area Structural Descriptors in Computer-Assisted Quantitative Structure-Property Relationship Studies. *Analytical Chemistry*, **62**, 2323-2329. <https://doi.org/10.1021/ac00220a013>
- [42] Csizmadia, I.G. (1976) *Theory and Practice of MO Calculations on Organic Molecules*. Elsevier, Amsterdam.