[Article]

doi: 10.3866/PKU.WHXB201112281

www.whxb.pku.edu.cn

MOLMAP 指数及其在变异性预测中的应用

张庆友¹ 龙海林¹ 冯秀林¹ 索净洁¹ 张丹丹¹ 李静亚¹
许力壮² 许 禄^{3,*}

(¹河南大学化学化工学院环境与分析科学研究所, 河南 开封 475004;

²深圳市人民医院, 广东 深圳 518020; ³中国科学院长春应用化学研究所, 长春 130022)

摘要: 分子映射(MOLMAP)指数是以分子的化学键描述符为基础, 通过 Kohonen 自组织映射依据一定的算法而衍生. 化学键描述符是由化学键的物理化学性质, 如两端原子的电荷差和拓扑性质, 键连杂原子数量等所组成. 本文将分子映射指数应用于 4075 个有机物质(Ames 实验结果: 2305 个结构有诱变性, 1770 个结构无诱变性)的变异性预测. 通过随机森林, 分别采用三种类型的指数建立模型: (1) 采用不同维数的分子映射指数; (2) 采用全局分子描述符; (3) 将分子映射指数与全局分子描述符相结合. 整个数据集的集外(out-of-bag)交叉验证的正确预测率达到 85.4%. 为了检验模型的稳定性, 采用所建模型预测源于另一数据库的 472 个化合物, 正确预测率为 86.7%, 与此前的研究相比, 两个预测结果均有所提高.

关键词: 分子映射指数; Kohonen 自组织映射; 随机森林; 诱变性; 结构-活性关系

中图分类号: O641

MOLMAP Descriptor and Its Application to Mutagenicity Prediction

ZHANG Qing-You¹ LONG Hai-Lin¹ FENG Xiu-Lin¹ SUO Jing-Jie¹
ZHANG Dan-Dan¹ LI Jing-Ya¹ XU Li-Zhuang² XU Lu^{3,*}

(¹Institute of Environmental and Analytical Sciences, College of Chemistry and Chemical Engineering, Henan University, Kaifeng 475004, Henan Province, P. R. China; ²Renmin Hospital of Shenzhen, Shenzhen 518020, Guangdong Province, P. R. China;

³Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun 130022, P. R. China)

Abstract: The molecular mapping of atom-level properties (MOLMAP) descriptor was generated on the basis of chemical bond descriptors of a molecule by Kohonen self-organizing map with a specific algorithm. The bond descriptors were composed of the physiochemical properties of the chemical bond, such as the difference of the charges between the two atoms and topological properties, such as the number of hetero-atoms connected to the two atoms. In this paper, the MOLMAP descriptors were used to predict the mutagenicity of 4075 organic substances (2305 mutagens and 1770 nonmutagens in Ames test). Random forests were used to construct mathematical models with three kinds of descriptors: (1) MOLMAP descriptors of different size; (2) global molecular descriptors; (3) the combination of MOLMAP descriptors and global molecular descriptors. The correct prediction percentage of out of bag (OOB) cross-validation of the whole data set reached 85.4%. To test the stability of the prediction model, it was used to predict the properties of a test set that was composed of 472 compounds collected from another database. The percentage of correct prediction of the test set was 86.7%. The prediction results were improved compared with the results of previous work.

Received: October 27, 2011; Revised: December 19, 2011; Published on Web: December 28, 2011.

*Corresponding author. Email: luxu@ciac.jl.cn; Tel: +86-431-85262239.

The project was supported by the National Natural Science Foundation of China (20875022), Scientific Research Foundation for the Returned Overseas Chinese Scholars, Ministry of Education of China (2009(1001)), and International Science and Technology Cooperation of Henan Province, China (114300510009).

国家自然科学基金(20875022), 教育部留学回国人员科研启动基金(2009(1001))及河南省国际科技合作项目(114300510009)资助

© Editorial office of Acta Physico-Chimica Sinica

Key Words: MOLMAP descriptor; Kohonen self-organizing map; Random forest; Mutagenicity; Structure–activity relationship

1 引言

药物研究伴随着药物需求的多样化日益扩展与深入,作为药物研究与开发中不可或缺的过程,早期药物筛选主要以动物实验为主,周期长且实验成本高,而对物质变异性的研究与预测对药物前期的筛选具有参考意义。¹⁻⁴ Ames实验是目前最常用的变异性检测方法之一,⁵已广泛应用于药物开发、饮用水检测等多种领域。^{6,7}但是其也存在着一些缺点,如假阳性高、操作繁琐、检测物质的数量有限等,由此计算机辅助物质变异性预测受到了重视,并已经开展了一些研究。^{8,9} Bursi¹⁰和 Helma¹¹等通过结构片段描述分子结构, Popelier等¹²通过半经验程序计算物理化学性质来描述分子, Pearl等¹³通过拓扑方法描述分子,进而进行变异性预测。总的来讲,相关的研究还较少。

MOLMAP指数在化合物反应性预测和化学反应的分类预测均有应用。¹⁴⁻¹⁸ 本文作者和 Aires-de-Sousa¹⁹合作,采用 Gasteiger等²⁰⁻²²提出的7个经验的化学键物理化学性质,例如共振稳定性和化学键裂解能等,依据一定的算法生成MOLMAP指数,并实现了有机物质的变异性预测。本文在此基础上进一步探索化学键性质的描述,采用27种物理化学性质及41种拓扑性质描述化学键,然后生成分子MOLMAP指数,并通过随机森林应用于一个较大型的数据集(4075个物质),预测物质的变异性。此外还采用该模型预测了源于另一数据库的数据集,检验所建模型的稳定性。

2 数据集

数据集由4083个有机物质所组成,¹⁰由于所用软件 ChemAxon²³的 Marvin不能计算其中的8个分子,因而采用余下4075个有机物质作为研究对象,相应的 Ames实验结果为:2305个结构是诱变物质,1770个结构是非诱变物质。测试集源于另外一个数据库,是由 Young等²⁴收集的472个化合物,其中305个结构为诱变物质,167个结构是非诱变物质。

3 方法

3.1 化学键描述符

本文计算了68种化学键性质,其中包括化学键的电荷差、化学键的轨道电负性差、邻接原子的最大及最小极化率、键端原子的电荷密度等27种物理化学性质和化学键的类型、键端原子的类型、键连杂原子的数目等41种拓扑性质,²³由每个化学键衍生一个68维的矢量。

众所周知,化学反应的基本特征就是化学键的生成和断裂,而物质的变异性与化学反应密切相关,故本文通过Kohonen自组织映射采用化学键的性质衍生MOLMAP指数。

3.2 Kohonen自组织映射

Kohonen自组织映射(SOM)是由Kohonen²⁵在1982年首先提出,它是一种无管理的人工神经网络,SOM是由二维神经元所组成,每一个神经元代表一个向量,其长度与输入向量的维数(本文为化学键描述符)相同,^{26,27}见图1。在应用过程中,首先采用训练集的化学键描述符对SOM进行训练,然后,递交测试集的化学键描述符到被训练的SOM进行预测,其过程简述如下,

SOM训练时,首先进行随机数的初始化,其次,提交化学键描述符矢量至SOM,并计算其与各神经元的欧式距离,找到与该描述符最接近的神经元,称为赢元,然后,调整赢元和邻近神经元的权值,使其与输入向量更相似,调整幅度随着与赢元的距离增大而减小。多次递交训练集中所有的描述符矢量,重复上述操作达到预先指定的次数时停止。经过上述训练,相似的神经元将处于邻近的位置。若递交未知的化学键描述符到被训练的SOM,则所得

图1 Kohonen自组织映射(SOM)模型
Fig.1 Kohonen self-organization mapping (SOM) model

赢元代表该化学键. SOM的特点为相近的输入向量将落入相同的或相近的神经元, 由此描述符相似的化学键将落入相同或者邻近的位置.

3.3 分子 MOLMAP 指数生成

分子 MOLMAP 指数是通过 Kohonen 提出的自组织映射, 由分子中所有化学键描述符所衍生. MOLMAP 指数是一个基于化学键类型的分子指数, 相同类型的化学键是指物理化学性质或拓扑性质相似的化学键. 分子 MOLMAP 指数的生成过程简述如下:²⁸ (1) 计算数据集中所有化合物的化学键描述符; (2) 从中随机选取一定数量的化学键作为训练集, 本文为 4999 个化学键及其描述符; (3) 以此 4999 个化学键训练 SOM, 得到被训练的人工神经网络; (4) 从数据集中提取所有的化学键描述符, 并递交到已经被训练的 SOM, 其赢元代表相应的化学键.

如递交一个分子的所有化学键描述符到上述被训练的 SOM, 然后依据这一分子的所有化学键在 SOM 上的分布, 进一步可将 SOM 上的输出转换为数值形式. 即将每个化学键的赢元(与该化学键描述符最相似神经元)赋值为 1.0, 则考虑到邻近神经元的相似性, 那么赢元周围的 8 个邻接神经元譬如可赋值为 0.3, 由此可转化为数字的形式. 另外, 如果分子中几个化学键落入同一个神经元, 则所得到的数值相加.

图 2 为一个通过 Kohonen SOM 生成 144 (12×12) 维的 MOLMAP 指数示例. 图 2a 所示为一个有机分子的所有化学键经训练后在 SOM 中的映射, 其中有 4 个等价的 Cl—C 键落入第 1 行第 9 列. 若按照前述规则转化为数值形式, 则见图 2b, 其中第 1 行第 9

列包含 4 个 Cl—C 键, 因而其数值为 4.0=4×1.0. 最后以从左向右的顺序一列接一列的裁剪, 并依次把列首尾相接就得到分子 MOLMAP 指数, 本例为 144 (12×12) 维. 可见, 分子 MOLMAP 指数的维数仅仅与二维神经元的维数有关, 而与分子中化学键的数量无关. 该指数的每一个变量对应着 SOM 的一个神经元, 每一个神经元可以看成一种新化学键类型. 本文采用前述 68 维的化学键描述符, 故同一类型的化学键是指具有相似物理化学和拓扑性质因而落入同一神经元的化学键.

为了检验 MOLMAP 指数的维数对模型预测能力的影响, 本文分别采用了维数为 625 (25×25)、900 (30×30) 和 1296 (36×36) 的 SOM 生成 MOLMAP 指数.

3.4 全局分子描述符

除前述 MOLMAP 指数外, 本文还计算出 35 种分子的物理化学性质和拓扑性质,²³ 构成 35 维的全局分子描述符矢量(global molecular descriptor). 其中包括原子数目、化学键的数目、分子质量、芳香原子数目、极化率、NH 键的数目、NH₂ 键的数目、氧原子的数目、氮原子的数目、最小原子电荷、最大原子电荷、氢原子的最小电荷、氢原子的最大电荷、重原子数目、羟基数目、氢键受体个数、氢键给体个数、log*P*、直链数目、脂肪环数目、芳香环数目、最小环的大小、杂环数目、杂芳香环数目、可旋转键数目、分子表面积、分子极性面积、分子最大投影面积、分子最小投影面积、折射率、平均分子极化度、芳香键数目、共振结构的数目、芳香稠环的数目和脂肪稠环的数目. 其中一些性质已经广泛应用于构效关系研究中.^{29,30}

图 2 分子 MOLMAP 指数的生成
Fig.2 Generation of the MOLMAP descriptor for a molecule

3.5 随机森林

随机森林(random forest)是 Breiman 在 Bagging 算法之后,提出的一种利用树的集合进行分类预测和回归预测的组合算法,^{31,32} 随机森林程序来源于 P-program 的 2.10.1 版本,³³ 该算法已经在实践中得到成功的应用。³⁴

随机森林通过随机的方式生成大量的树来建立数学模型,³⁵ 每一棵树类似于一个分类回归树,但不进行修剪. 对于每一棵树,在采用训练集进行训练时首先随机选取一个变量子集,然后从选取的子集中选择变量来划分结点,对输入矢量的每一个变量找到一个最佳的分割点(能够最有效分类的值),其中分类效果最好的变量作为从父结点到子结点的判据,然后根据该结点内的多数票来决定该结点属于哪一类. 每棵树均独立的进行训练,并用于未知样本的预测. 随机森林中树的数量由人为设定(本文为 1000 棵树),每一棵树的预测相当于一票,最终的结果由所有树的多数票决定. 如某化合物在 900 棵树中被预测为变异性物质, 100 棵树中被预测为非变异性物质, 则该化合物被预测为变异性物质. 随机森林提供了变量重要性的量化评价, 评价主要依据两种方式: (1) 随机地交换某变量的值, 由所引起的误分类变化评价该变量; (2) 采用某变量进行结点分类, 由所产生子节点分类效果来量化该变量的重要性. 变量重要性可以作为变量选择的依据, 同时变量重要性结果有助于发现与活性密切相关的化合物结构特征, 从而提供对活性研究有启示意义的信息.

在训练随机森林过程中, 每一棵树均把训练集随机分成两部分, 一部分数据作为训练集, 另一部分则作为测试集, 并综合每棵树所建模型对测试集的预测结果来评价整个训练集, 称为集外(OOB)交

叉验证.

4 结果与讨论

采用训练集的 4075 个有机物质训练随机森林, 并分三种情况建立物质变异性的预测模型: (1) 仅采用全局分子描述符; (2) 仅采用 MOLMAP 指数, 其中 MOLMAP 指数的维数为: 625 (25×25)、900 (30×30) 或 1296 (36×36); (3) MOLMAP 指数和全局分子描述符相结合. 此外, 前述每一个预测模型均应用于独立测试集的 472 个化合物.

首先仅采用全局分子描述符建立预测模型, 所得结果见表 1 第 2 行. 其中训练集的 OOB 交叉验证结果为 81.7%; 独立测试集的预测结果为 77.6%.

本文的 MOLMAP 指数由两类化学键描述符, 即物理化学性质及拓扑性质所衍生. 故分别采用化学键描述符中的 27 个物理化学性质描述符和 41 个拓扑性质描述符建立物质变异性预测模型, 则训练集所有化合物的 OOB 交叉验证结果及测试集预测结果见表 1 与表 2, 其中表 1 所示为由化学键的物理化学性质衍生 MOLMAP 指数的预测结果; 表 2 所示为由化学键的拓扑性质衍生 MOLMAP 指数的预测结果.

由于 Kohonen SOM 采用随机数进行初始化, 为了得到较稳定的预测模型, 采用化学键描述符训练 SOM 三次, 每次均得到一个 MOLMAP 指数, 并分别采用三个 MOLMAP 指数进行预测. 对于一个化合物来说, 若其中两个 MOLMAP 指数的预测结果为变异性, 而一个预测为非变异性, 则预测该化合物为变异性. 即以三个模型的多数票作为最终预测结果.

由表 1 可知: 若仅采用化学键的物理化学性质衍生的 MOLMAP 指数建立预测模型, 则训练集的

表 1 基于 35 维全局分子描述符和由化学键物理化学性质衍生的 MOLMAP 指数的随机森林预测结果
Table 1 Correct predictions of random forest based on 35 global molecular descriptors and MOLMAP descriptors derived from physiochemical properties of chemical bonds

Descriptors ^a	Percentage of correct out-of-bag estimation of training set/%	Percentage of correct prediction of test set/%
35 molecular descriptors	81.7	77.6
MOLMAPs (25×25)	84.7	84.8
MOLMAPs (30×30)	84.3	85.0
MOLMAPs (36×36)	84.6	84.8
MOLMAPs (25×25)+35MDs	84.8	84.3
MOLMAPs (30×30)+35MDs	84.8	85.0
MOLMAPs (36×36)+35MDs	84.9	84.8

^a MDs denote global molecular descriptors.

表2 基于35维全局分子描述符和由化学键的拓扑性质衍生的MOLMAP指数的随机森林预测结果

Table 2 Correct predictions of random forest based on 35 global molecular descriptors and MOLMAP descriptors derived from topological properties of chemical bonds

Topological descriptors	Percentage of correct out-of-bag estimation of training set/%	Percentage of correct prediction of test set/%
MOLMAPs (25×25)	84.7	84.3
MOLMAPs (30×30)	84.9	85.0
MOLMAPs (36×36)	85.4	85.8
MOLMAPs (25×25)+35MDs	85.3	85.0
MOLMAPs (30×30)+35MDs	85.5	84.8
MOLMAPs (36×36)+35MDs	85.4	86.2

表3 基于35维全局分子描述符和由化学键的物理化学性质与拓扑性质衍生的MOLMAP指数的随机森林预测结果

Table 3 Correct predictions of random forest based on 35 global molecular descriptors and MOLMAP descriptors derived from physiochemical properties and topological properties of chemical bonds

Descriptors	Percentage of correct out-of-bag estimation of training set/%	Percentage of correct prediction of test set/%
MOLMAPs (25×25)	84.1	84.3
MOLMAPs (30×30)	84.6	85.6
MOLMAPs (36×36)	84.5	85.6
MOLMAPs (25×25)+35MDs	85.3	85.6
MOLMAPs (30×30)+35MDs	85.6	85.8
MOLMAPs (36×36)+35MDs	85.4	86.7

OOB交叉验证结果为84.3%–84.7%, 独立测试集的预测结果为84.8%–85.0%, 均显著好于仅采用全局分子描述符的结果。若进一步将MOLMAP指数与全局分子描述符相结合, 则训练集的交叉验证结果在84.8%–84.9%之间; 测试集的预测结果在84.3%–85.0%之间。

由表2可知: 若采用化学键的拓扑性质衍生MOLMAP指数与全局分子描述符相结合, 则训练集的交叉验证结果在85.3%–85.5%之间; 测试集的预测结果在84.8%–86.2%之间, 所得结果略优于表1的结果。

在此基础上, 合并27维的物理化学描述符和41维的拓扑性质描述符衍生68维的化学键描述符, 由化学键描述符生成MOLMAP指数, 并采用前述方法通过随机森林构建预测模型和进行预测, 则所得结果见表3。由表3可知: 若MOLMAP指数与35个全局分子描述符相结合, 则OOB的交叉验证结果为85.3%–85.6%, 测试集的预测结果为85.6%–86.7%。在此前研究¹⁹中同一训练集的OOB交叉验证最佳结果为84.1%, 测试集为84.5%。可见, 本文结果好于该结果。

MOLMAP指数的每一个变量均衍生于Kohonen自组织映射的神经元, 而每一个神经元均代表着“化学键类型”。根据随机森林的变量重要性, MOLMAP指数中最重要的两个变量和第四重要的变量均是由N=O双键所衍生, 这些化学键主要

是来源于硝基, 也有少部分来源于C–N=O和O=N–N基团。第三重要的变量由苯并化合物上苯环的碳碳键所衍生。这说明该类基团是影响变异性的的重要基团。在文献¹⁰中指出, 含有上述两个基团的化合物通常具有变异性, 这表明上述模型不仅仅能够建立较好的模型, 还有助于发掘引起变异性的的重要化学键类型。

综上所述, 采用化学键的物理化学描述符衍生MOLMAP指数所建立的模型预测结果最差; 若采用化学键的拓扑性质则预测能力有所提高; 而将二者相结合则预测能力最佳。

5 结 论

对于由4075个化合物所组成的比较大型的数据集, 本文基于Kohonen自组织映射, 采用所提取的68种化学键的物理化学性质和拓扑性质, 在此基础上生成625 (25×25)、900 (30×30)、1296 (36×36)三种维度的MOLMAP指数, 并结合35维全局分子描述符, 通过随机森林建立变异性预测模型, 从而实现了有机物质变异性的自动预测。通过源于另一个数据库的独立数据集进行检验, 模型稳健性较好, 比此前的预测能力有所提高, 说明通过化学键描述符的改进, 能够建立更加稳定的变异性物质预测模型。

References

- (1) Patlewicz, G.; Rodford, R.; Walker, J. D. *Environ. Toxicol.*

- Chem.* **2003**, 22, 1885.
- (2) Benigni, R. *Chem. Rev.* **2005**, 105, 1767.
- (3) Hansen, K.; Mika, S.; Schroeter, T.; Sutter, A.; Laak, A. T.; Steger-Hartmann, T.; Heinrich, N. Muller, K. R. *J. Chem. Inf. Model.* **2009**, 49, 2077.
- (4) Casalegno, M.; Benfenati, E.; Sello, G. *J. Chem. Inf. Model.* **2011**, 51, 1564.
- (5) Ames, B. N.; McCann, J.; Yamasaki, E. *Mutation Res.* **1975**, 3, 347.
- (6) Mortelmans, K.; Zeiger, E. *Mutation Res.* **2000**, 455, 29.
- (7) Meier, J. R. *Mutation Res.* **1988**, 196, 211.
- (8) Zheng, M. Y.; Liu, Z. G.; Xue, C. X. Zhu, W. L.; Chen, K. X.; Luo, X. M.; Jiang, H. L. *Bioinformatics* **2006**, 22, 2099.
- (9) Liao, Q.; Yao, J. H.; Yuan, S. G. *Molecular Diversity* **2007**, 11, 59.
- (10) Kazius, J.; McGuire, R.; Bursi, R. *J. Med. Chem.* **2005**, 48, 312.
- (11) Helma, C.; Cramer, T.; Kramer, S.; Raedt, L. D. *J. Chem. Inf. Comput. Sci.* **2004**, 44, 1402.
- (12) Popelier, P. L. A.; Smith, P. J.; Chaudry, U. A. *J. Comput. -Aided Mol. Des.* **2004**, 18, 709.
- (13) He, L. N.; Jurs, P. C.; Custer, L. L.; Durham, S. K.; Pearl, G. M. *Chem. Res. Toxicol.* **2003**, 16, 1567.
- (14) Gupta, S.; Matthew, S.; Abreu, P. M.; Aires-de-Sousa, J. *Bioorg. Med. Chem.* **2006**, 14, 1199.
- (15) Zhang, Q. Y.; Aires-de-Sousa, J. *J. Chem. Inf. Model.* **2005**, 45, 1775.
- (16) Latino, D. A. R. S.; Aires-de-Sousa, J. *Angew. Chem. Int. Edit.* **2006**, 45, 2066.
- (17) Latino, D. A. R. S.; Zhang, Q. Y.; Aires-De-Sousa, J. *Bioinformatics* **2008**, 24, 2236.
- (18) Latino, D. A. R. S.; Aires-de-Sousa, J. *J. Chem. Inf. Model.* **2009**, 49, 1839.
- (19) Zhang, Q. Y.; Aires-de-Sousa, J. *J. Chem. Inf. Model.* **2007**, 47, 1.
- (20) Gasteiger, J.; Marsili, M.; Hutchings, M. G.; Saller, H.; Löw, P.; Röse, P.; Rafeiner, K. *J. Chem. Inf. Comput. Sci.* **1990**, 30, 467.
- (21) Simon, V.; Gasteiger, J.; Zupan, J. *J. Am. Chem. Soc.* **1993**, 115, 9148.
- (22) Gasteiger, J. *Mini-Rev. Med. Chem.* **2003**, 3, 789.
- (23) <http://www.chemaxon.com/>.
- (24) Young, S. S.; Gombar, V. K.; Emptage, M. R.; Cariello, N. F.; Lambert, C. *Chem. Int. Lab. Syst.* **2002**, 60, 5.
- (25) Kohonen, T. *Biol. Cybern.* **1982**, 43, 59.
- (26) Aires-de-Sousa, J. *Chem. Int. Lab. Syst.* **2002**, 61, 167.
- (27) <http://www.dq.fct.unl.pt/staff/jas/jatoon/>.
- (28) Long, H. L.; Feng, X. L.; Suo, J. J.; Zhang, D. D.; Li, J. Y.; Zhang, Q. Y.; Xu, L. *Computers and Applied Chemistry* accepted. [龙海林, 冯秀林, 索净洁, 张丹丹, 李静亚, 张庆友, 许 禄. 计算机与应用化学, 已接受.]
- (29) Dai, Z. J.; Zhou, W.; Yuan, Z. M. *Acta Phys. -Chim. Sin.* **2011**, 27, 1654. [代志军, 周 玮, 袁哲明. 物理化学学报, **2011**, 27, 1654.]
- (30) Yang, G. B.; Li, Z. R.; Rao, H. B.; Li, X. Y.; Chen, Y. Z. *Acta Phys. -Chim. Sin.* **2010**, 26, 3351. [杨国兵, 李泽荣, 饶含兵, 李象远, 陈宇综. 物理化学学报, **2010**, 26, 3351.]
- (31) Breiman, L. *Machine Learning* **1996**, 24, 123.
- (32) Díaz-Uriarte, R.; Andrés, S. A. D. *BMC Bioinformatics* **2006**, 7, 3.
- (33) <http://www.R-project.org>.
- (34) Svetnik, V.; Liaw, A.; Tong, C.; Culberson, J. C.; Sheridan, R. P.; Feuston, B. P. *J. Chem. Inf. Comput. Sci.* **2003**, 43, 1947.
- (35) Breiman, L. *Machine Learning* **2001**, 45, 5.