广西科学院学报 2011,27(3):205～210
Journal of Guangxi Academy of Sciences Vol. 27, No. 3 August 2011

Fitting of Precipitation in 39 Beta World Cities from 1901 to 1998 Using Random Walk Model*
用随机漫步模型拟合 39 个贝塔世界城市 1901～1998 年的降水量

YAN Shao-min[1], WU Guang[2]**
严少敏[1]，吴 光[2]

(1. State Key Laboratory of Non-food Biomass Enzyme Technology, National Engineering Research Center for Non-food Biorefinery, Guangxi Key Laboratory of Biorefinery, Guangxi Academy of Sciences, Nanning, Guangxi, 530007, China; 2. DreamSciTech Consulting, Shenzhen, Guangdong, 518054, China)
(1. 广西科学院非粮生物质国家重点实验室，国家非粮生物质能源工程技术研究中心，广西生物炼制重点实验室，广西南宁 530007；2. 深圳市追梦科技咨询有限公司，广东深圳 518054)

Abstract: The random walk model was used to fit the annual precipitation in 39 beta world cities from 1901 to 1998. At first, the recorded precipitation of each city was converted to precipitation walk and random precipitation, then the random walk model was used to fit the converted data and recorded ones. The results show that this model can fit both precipitation walk and recorded precipitation. Thus this study provides a model to describe the precipitation patterns, and suggests the random mechanism may underline in the precipitation change of these cities during this period.

Key words: beta world city, fitting, precipitation, random walk

摘要：用随机漫步模型拟合 1901 年至 1998 年 39 个贝塔世界城市的年降水量。先将每个城市的降水量转换为降水漫步和随机降水，再用随机漫步模型分别对转换数据和实际数据进行拟合。结果显示，该模型能够拟合降水漫步和实际降水，为降水建模提供了一种模型，表明此期间这些城市的降水呈现随机变化。

关键词：贝塔世界城市　拟合　降水量　随机漫步
中图法分类号：P426.61+4　**文献标识码**：A　**文章编号**：1002-7378(2011)03-0205-06

Mathematical description of precipitation is important, not only because it can help us to better understand the precipitation pattern, but also because it can provide a tool to predict the precipitation.

The random walk is an important model to describe natural phenomena, for example, stock pattern[1]. Recently, we used this model to analyze the temperature change indicating that its fitting can be reached in both global level[2] and city level[3]. Thus an interesting question raised here is whether the random walk can describe the precipitation? This hypothesis seems to be likely because the year-to-year precipitation seems to be somewhat fluctuated without clearly visible patterns.

In order to test this hypothesis, we use the random walk model to fit the annual precipitation in 39 beta world cites from 1901 to 1998.

1 Materials and methods

1.1 Data

Thirty-nine beta world cities are chosen according to

收稿日期：2011-03-18
作者简介：严少敏(1958-)，女，博士，研究员，主要从事计算变异学和模型研究。
* This study was partly supported by Guangxi Science Foundation (0907016, 10-046-06, 11-031-11, 2010GXNSFA013003 and 2010GXNSFA013046).
** 通讯作者，E-mail：hongguanglishibahao@yahoo.com。

Wikipedia[4]. However, the number of these cities and their order are changed frequently due to the characteristics of Wikipedia, thus the cities were ordered based on their order in February, 2010. The precipitations recorded in these 39 cities from 1901 to 1998 based on 0.5° by 0.5° latitude and longitude grid-box basis cross globe are obtained from the website of Oak Ridge National Laboratory[5]. The latitudes and longitudes of these 39 beta world cities are determined using Get Lat Lon[6].

1.2 Precipitation walk

The simplest random walk model starts at zero and at each step moves by ±1 with equal probability[1~3] whereas the precipitation is in decimal format, thus we convert the precipitation into ±1 format and call it as precipitation walk. Technically, when the precipitation at certain year is higher than its previous one, we classify it as 1, otherwise we classify it as −1(column 3, Table 1), and then we add them together as the random walk does (column 4, Table 1).

Table 1 Conversion of recorded precipitation into precipitation walk and generation of random walk for precipitation in Melbourne from 1901 to 1998

Year	Precipitation (mm/year)	Precipitation step	Precipitation walk	Generated random number	Random step	Random walk
1901	601		0	0.04879		0
1902	610	1	1	0.74322	1	1
1903	752	1	2	0.05883	−1	0
1904	558	−1	1	0.26845	1	1
1905	634	1	2	−0.56271	−1	0
1906	581	−1	1	0.50705	1	1
1907	501	−1	0	−0.66933	−1	0
1908	382	−1	−1	−0.68492	−1	−1
1909	585	1	0	−0.82691	−1	−2
1910	564	−1	−1	0.09072	1	−1
…	…	…	…	…	…	…
1991	635	1	−6	0.22393	1	−6
1992	788	1	−5	0.20448	−1	−7
1993	701	−1	−6	−0.72712	−1	−8
1994	408	−1	−7	0.34260	1	−7
1995	676	1	−6	0.94707	1	−6
1996	640	−1	−7	−0.21775	−1	−7
1997	384	−1	−8	−0.43780	−1	−8
1998	515	1	−7	0.92797	1	−7

The random number is generated by SigmaPlot with the seed of 4.43052.

1.3 Generation of random walk

We use the SigmaPlot[7] to generate random sequence with different seeds: we generate random number ranged from −1 to 1, and then we classify a random number as 1 if it is larger than its previous one and as −1 if it is smaller than its previous one (column 6, Table 1). Thereafter we add the classified values as random walk (column 7, Table 1).

1.4 Searching for seed

To the best of our knowledge, there is no algorithm available to find the right seed, which produces the best fit between random walk and observed data. However, this could not be problematic with current computational technique, because we can simply search all the seeds in searching space and compare their outcomes.

1.5 Fitting recorded precipitation

Thereafter, we use a more complicated random walk model[8] to fit the recorded precipitation, which is in decimal format. In plain words, the simplest random walk comes from tossing of double-sided coin, while this random walk could be regarded as tossing of dice, which can be not only six-sided but as many as the decimal data. This way, we generate random numbers (column 3, Table 2), and add them to construct the random walk (column 4, Table 2), and the fitting is again to search the best seed that generates best fit.

Table 2 Generation of recorded precipitation into random precipitation in Melbourne from 1901 to 1998

Year	Recorded precipitation (mm/year)	Generated random number	Random precipitation (mm/year)
1901	601		601.00000
1902	610	−149.83324	451.16676
1903	752	167.00549	618.17224
1904	558	15.64558	633.81782
1905	634	−75.30259	558.51523
1906	581	−109.90629	448.60894
1907	501	173.05481	621.66376
1908	382	−37.00038	584.66338
1909	585	−0.51526	584.14812
1910	564	−123.56600	460.58213
…	…	…	…
1991	635	−152.60633	368.06526
1992	788	110.60943	478.67469
1993	701	−6.22320	472.45149
1994	408	6.41809	478.86959
1995	676	101.95389	580.82348
1996	640	−157.55508	423.26839
1997	384	8.94215	432.21054
1998	515	42.18479	474.39533

The random number is generated by SigmaPlot with the seed of 1.34641.

Table 3 Model parameter (seed) and fitted results for fitting precipitation change in 39 beta world cities from 1901 to 1998 using random walk model

City	Fitting of precipitation walk		Fitting of recorded precipitation	
	Seed	Sum of squares	Seed	Sum of squares
Melbourne	4.43052	124	1.34641	2586627.55
Barcelona	0.92618	136	1.51401	2984410.76
Los Angeles	1.09247	132	8.04385	9837002.96
Johannesburg	0.61683	132	0.60545	2257443.47
Manila	0.17978	106	1.19744	17132448.28
Bogota	2.64813	124	1.03233	6722165.76
New Delhi	7.45919	156	0.53868	6485157.38
Atlanta	5.76188	112	1.65197	6444147.79
Washington D. C.	3.70431	120	0.70337	3922775.72
Tel Aviv	4.53547	120	3.70963	4036478.17
Bucharest	2.96749	124	2.82657	1634074.60
San Francisco	0.64649	140	1.46898	3552577.37
Helsinki	1.17561	119	3.88437	1601835.51
Berlin	1.53962	100	8.04386	1267406.51
Dubai	3.94439	108	2.51977	336204.14
Oslo	5.15241	128	0.49518	2762908.96
Geneva	7.85232	124	7.95525	4468964.61
Riyadh	9.00335	124	1.48129	318508.11
Copenhagen	2.67161	136	1.75624	979500.29
Hamburg	2.56624	132	1.68698	1629723.86
Cairo	3.59924	126	1.74148	30729.02
Bangalore	0.93401	112	2.01452	5029654.32
Jeddah	0.41142	118	6.08046	89874.72
Kuwait	1.79627	94	2.71912	178018.85
Luxembourg	1.61462	132	1.46898	2931614.05
Munich	3.34785	109	4.68752	1722156.75
Kiev	1.24473	124	6.08044	1628394.92
Dallas	1.85854	128	1.80649	6494614.91
Boston	2.46612	132	1.70767	3991272.22
Miami	2.2682	168	6.47106	14856926.07
Sofia	1.47405	136	3.50057	1574839.07
Düsseldorf	2.3542	156	1.62457	2051901.84
Houston	2.38592	144	2.43225	12649509.69
Beirut	4.45934	132	9.10630	2714757.59
Guangzhou	0.65443	120	7.29461	12577487.18
Nicosia	0.96966	124	1.68592	1959226.37
Karachi	7.75981	112	0.72414	3769307.76
Montevideo	1.54723	132	3.57487	10120934.97
Rio de Janeiro	4.15844	132	1.64003	7726967.13

1.6 Comparison

For determination of the best seed, we compare the least squares between precipitation walk and random walk, and between recorded precipitation and random precipitation generated from different seeds.

2 Results

Table 1 shows how we construct a precipitation walk and its corresponding random walk. For the precipitation walk, we have the follows: (i) the starting point is the annual precipitation in 1901, 601 (cell 2, column 2), and this starting point corresponds zero in sense of precipitation walk (cell 2, column 4); (ii) the second annual precipitation is 610 (cell 3, column 2), which is larger than the first one, 601 (cell 2, column 2), so we assign 1 as precipitation step (cell 3, column 3); (iii) the precipitation walk is 1 (0 + (1)) (cell 3, column 4); and (iv) the similar computation is applied to all the data in columns 2, 3, and 4.

For the random walk, we have the follows: (i) a good seed we found is 4.43052 (see the legend to Table 1), and this seed generates a series of random numbers (column 5); (ii) the first random number, 0.04879 (cell 2, column 5) is considered as the starting point corresponding to 0 in random walk (cell 2, column 7); (iii) the second random number, 0.74322 (cell 3, column 5), is larger than the first random number, 0.04879 (cell 2, column 5), so we assign 1 (cell 3, column 6); (iv) the random walk is 1 (0 + 1) (cell 3, column 7); and (v) the similar procedure is applied to all the data in columns 5, 6, and 7. In the same manner, we construct the precipitation walk and random walk.

Figure 1 shows the fittings of precipitation walk in 10 cities using random walk model. As can be seen, the random walk can go through precipitation walk so that the random walk can fit the precipitation walk if the precipitation walk can be considered as the precipitation trend.

Table 2 shows how we use a random walk model to fit the recorded annual precipitation, here we only need to construct the random precipitation: (i) the starting point is the first recorded annual precipitation, which is 601 (cell 2 in column 2 and column 4); (ii) the seed for Table 2 is 1.34641 (see the legend to Table 2); (iii) the first random number generated by the seed is −149.83324 (cell 3, column 3); (iv) we add this value to the previous precipitation datum (601) resulting in 451.16676 (cell 3, column 4); and (v) along this procedure, we get the random precipitation in column 4.

Figure 2 displays the fittings of recorded precipitation with random precipitation in 10 cities. As can be seen, the random model can fit the recorded precipitation data because the random precipitation data have a similar pattern with their recorded ones.

Table 3 lists all fitted parameters for the annual precipitation of the 39 cities.

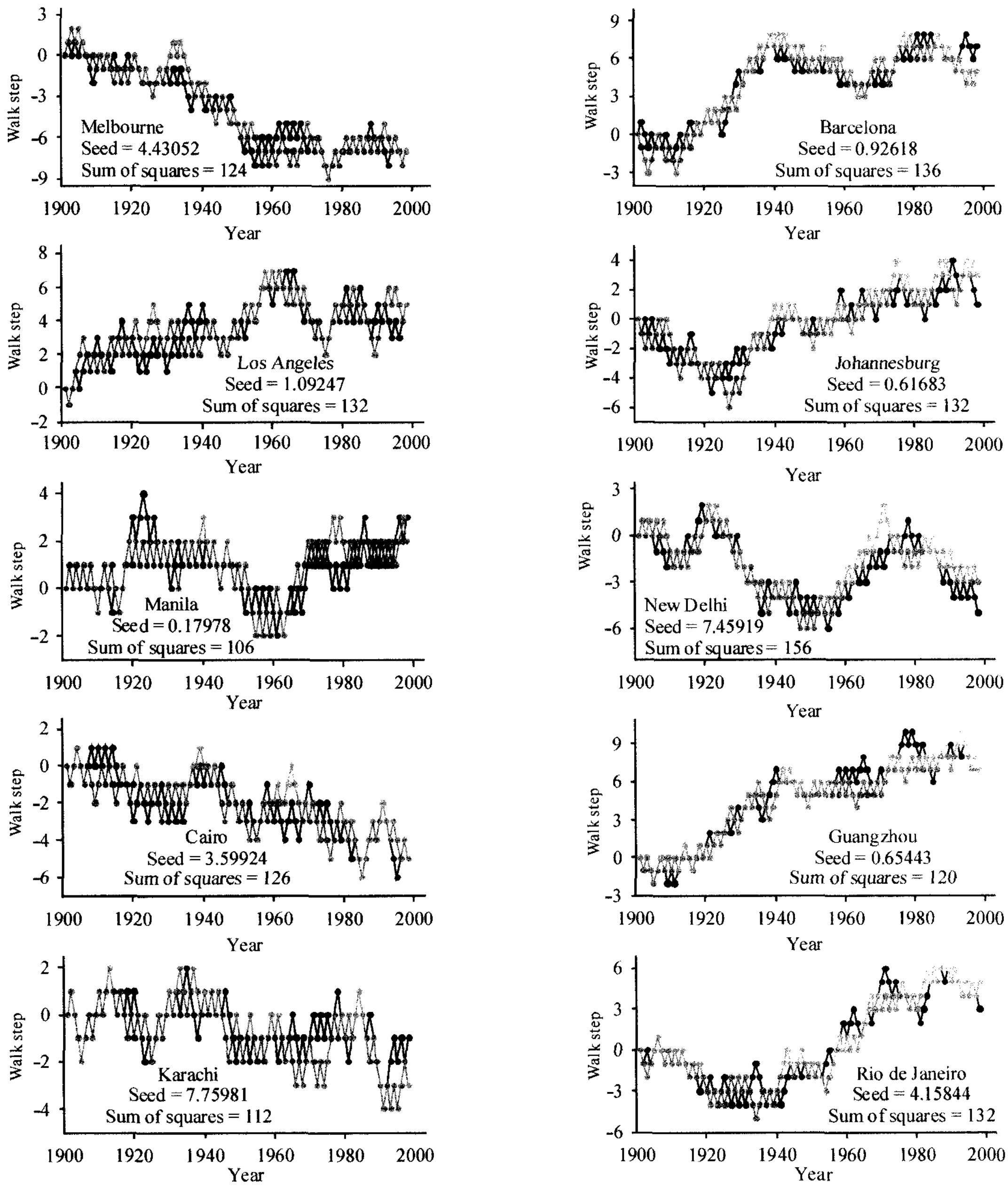

Fig. 1 Comparison of precipitation walk with random walk in 10 cities from 1901 to 1998

—•— Random, ···•··· Precipitation.

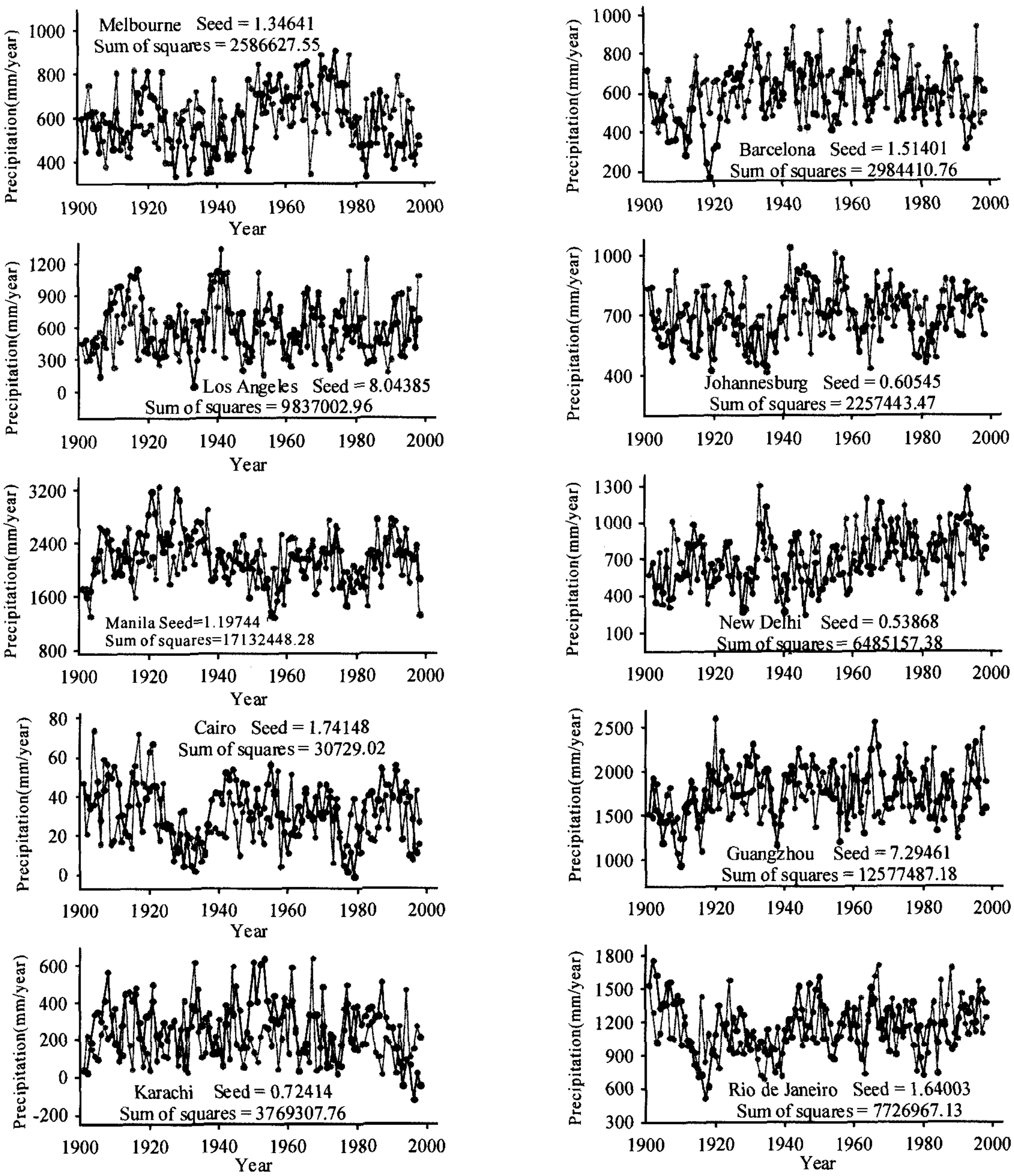

Fig. 2 Comparison of random precipitation with recorded precipitation in 10 cities from 1901 to 1998

—•—: Random, —•—: Recorded.

3 Discussion

The data used in this study spanned for almost a century. With 98 annual precipitations, we may find a good estimate because the seed is the only model parameter for random walk. However, the uncertainty would increase if we use more complicated models that have more model parameters.

The completely perfect fitting of precipitation walk is a rare event. For example, there are 98 annual precipitations, thus the completely perfect fitting has the chance of $(1/2)^{98}$ theoretically, which is extremely small. Clearly this probability is very difficult to achieve in limited time. At this stage, we have no way to compare our results with other models because the output of other models could not be in equal-sized step.

Actually, we can view the precipitation walk, which is the conversation from its annual precipitation, as the trend of recorded precipitation. This is because the precipitation walk answers the very basic question of whether the precipitation in certain year is higher (1) or lower (−1) than in its previous year.

This study demonstrates that the random walk model can fit the precipitation walk and recorded precipitation in city level. Together with our previous studies[2,3], the random walk model provides another model to analyze the changes in both temperature and precipitation, and a piece of evidence that help to understand the true nature.

Acknowledgements

The authors wish to thank Dr Hong Zhang at Biyee SciTech Inc., MA, USA for helpful discussion.

References:

[1] Yan S M, Wu G. Fitting of SSEC index (Shanghai Composite) from January 2000 to July 2010 using random walk model[J]. Guangxi Sci, 2011, 18: 92-96.

[2] Yan S M, Wu G. Fitting of global temperature change from 1850 to 2009 using random walk[J]. Guangxi Sci, 2010, 17(02): 148-150.

[3] Yan S M, Wu G. Application of random walk model to fit temperature in 46 gamma world cities from 1901 to 1998[J]. Natural Sci, 2010, 2: 1425-1431.

[4] Wikimedia Foundation Inc. Wikipedia, the free encyclopedia[EB/OL]. [2010-2-28]. http://en.wikipedia.org/wiki/Global_city.

[5] New M, Hulme M, Jones P. Representing twentieth-century space-time climate variability. Part II: Development of 1901-96 monthly grids of terrestrial surface climate[J]. Journal of Climate, 2000, 13: 2217-2238.

[6] Willison S. Get Lat Lon[EB/OL]. [2010-2-28]. http://www.getlatlon.com/.

[7] SPSS Inc. SigmaPlot for Windows Version 8.02[CP]. 2002.

[8] Borovkov A, Borovkov K. Asymptotic analysis of random walks: heavy-tailed distributions[M]. Cambridge: University Press, 2008.

(责任编辑:陈小玲)

(上接第204页)

3 结束语

本文通过对酒精发酵菌株K6生物学及其生理生化特性的研究,发现菌株K6呈多种形态;繁殖方式为芽殖,有假菌丝,无子囊孢子,革兰氏染色为阴性;能发酵葡萄糖、蔗糖、果糖、半乳糖、甘露糖、麦芽糖、木糖;能同化葡萄糖、蔗糖、果糖、半乳糖、甘露糖、木糖、乳糖、可溶性淀粉(不定)、松三糖、山梨糖(不定);氮源可以同化硫酸胺,不能同化硝酸钾;耐高渗透压;产酯;最适pH值为4.0~4.5,最适温度为28~32℃。同源性分析显示菌株K6与假丝酵母属(*Candida*)的同源性达到98%,比对的E值最小,分值最高,发育关系非常接近,这与生理生化鉴定结果相一致。假丝酵母属是酵母属中最大的一个属,包括近200种,在常规分类所测试的30~40项生理生化形状中,许多种间往往只存在一两项区别[6],所以,对其分类鉴定,我们还应该进一步的深入研究。酒精发酵试验结果表明,菌株K6能同时具有降低挥发酸,缩短发酵时间,提高产酒率,增加酒精产量的作用,其中,在淀粉质原料中的应用特别显著,能缩短时间7~10h,酒精产量平均提高8.8%。菌株K6能广泛应用于不同原料的酒精发酵生产中,是酒精发酵行业技术发展的新突破。

参考文献:

[1] 张纪忠.微生物分类学[M].上海:复旦大学出版社,1990.

[2] J A巴尼特.酵母菌的特性与鉴定手册[M].青岛:青岛海洋出版社,1984.

[3] 周德庆.微生物学实验手册[M].上海:上海科学技术出版社,1986.

[4] J萨母布鲁克,E F弗里奇,T曼尼阿蒂斯.分子克隆实验指南[M].第2版.北京:科学出版社,1992.

[5] 胡嗣明.酒精生产分析检验[M].北京:轻工业出版社,1983.

[6] 刑来君,李明春.普通真菌学[M].北京:高等教育出版社,1999.

(责任编辑:尹 闯)