

网络出版时间:2013-06-20 15:47

网络出版地址: <http://www.cnki.net/kcms/detail/61.1390.S.20130620.1547.016.html>

基于 MaxEnt 模拟三叶海棠的地理分布

杨俊仙^a, 王雷宏^b, 徐小牛^b

(安徽农业大学 a 理学院, b 林学与园林学院, 安徽 合肥 230036)

【摘要】 **【目的】**用已有的采集标本分布记录,模拟三叶海棠的地理分布。**【方法】**从中国 4 个标本馆获取三叶海棠的 291 个分布数据,从 WorldClim 网站下载 1950—2000 年生物气候数据,用最大熵模型(MaxEnt)模拟其地理分布。**【结果】**①三叶海棠的潜在分布地区有:巴基斯坦东北部;不丹中部、东北部;印度东北部;中国西藏东南部、河北省东北部和北京西部山区。②依据模拟分布值 0.5~0.7 标定的三叶海棠分布区域,排除其潜在分布区域,再结合这些区域的地理间隔性,从宏观景观上将三叶海棠的地理分布格局划分为 4 个区块:中国四川、甘肃、陕西三省交界地区;中国重庆大部分地区,贵州大部分地区,湖北西南部,湖南大部分地区,广西与贵州、湖南交界地区,江西,广东与湖南、江西交界地区,福建东北部,浙江大部分地区,安徽南部;韩国南部沿海地区;日本岛大部分地区。③刀切法检测表明,温暖季节平均降雨量对三叶海棠的分布增益贡献最大,三叶海棠喜生于温暖季节平均降雨量在 450~800 mm 的地区。**【结论】**用 MaxEnt 模拟三叶海棠的地理分布有一定的准确性,反映出了三叶海棠基本的地理分布格局和潜在分布区域,并阐明了主导其地理分布的生物气候因子。

【关键词】 三叶海棠;MaxEnt 模型;地理分布

【中图分类号】 S685.990.19

【文献标志码】 A

【文章编号】 1671-9387(2013)07-0172-05

MaxEnt based geographic distribution pattern of *Malus toringo*

YANG Jun-xian^a, WANG Lei-hong^b, XU Xiao-niu^b

(a School of Science, b School of Forestry & Landscape Architecture, Anhui Agricultural University, Hefei, Anhui 230036, China)

Abstract: **【Objective】** Geographic distribution pattern of *Malus toringo* was simulated using distribution records of collected specimen. **【Method】** Geographic distribution of *M. toringo* was modeled using MaxEnt based on 291 distribution records from four herbariums in China and bioclimatic data (1950—2000) from WorldClim. **【Result】** ① Potential distribution areas of *M. toringo* included the northeast of Pakistan, the central and northeast of Bhutan, the northeast of India, the southeast of Xizang, the northeast of Hebei, and the western mountains of Beijing. ② The geographic distribution pattern of *M. toringo* were divided into four distribution blocks based on distribution values of 0.5—0.7: the border area of Sichuan, Gansu and Shaanxi in China; Chongqing, the most area of Guizhou, the southwest of Hubei, the most area of Hunan, the border area of Guangxi, Guizhou and Hunan, the most area of Jiangxi, the border area of Guangdong, Hunan and Jiangxi, the northeast Fujian, the most area of Zhejiang, the south of Anhui; the southern coastal areas of Korea; and the most area of Japanese islands. ③ Jackknife Test showed that average precipitation in warm seasons had the greatest contribution to the distribution gain of *M. toringo* and it naturally distributed over areas with average precipitation of 450—800 mm. **【Conclusion】** In general, Max-

* [收稿日期] 2013-01-07

[基金项目] 国家自然科学基金项目(31070588);安徽省自然科学基金项目(10040606Q18);安徽农业大学人才引进基金项目(2008010)

[作者简介] 杨俊仙(1976—),女,山西五台人,讲师,硕士,主要从事生物数学研究。E-mail: yangjunxian@ahau.edu.cn

[通信作者] 王雷宏(1977—),男,山西五台人,讲师,博士,主要从事树木学研究。E-mail: leihong_wang@qq.com

Ent accurately simulated the geographical distribution of *M. toringo*. It showed the basic pattern of geographic distribution and the potential distribution areas, and clarified the dominant bioclimatic factors to geographic distribution of *M. toringo*.

Key words: *Malus toringo* Siebold ex De Vriese; maximum entropy modelling (MaxEnt); geographic distribution

三叶海棠 (*Malus toringo* Siebold ex De Vriese)^[1] 在原 Rehder 系统中隶属于花楸苹果组 (Sect. *Sorbomalus* Zabel ex Schneid)^[2], 但有关形态、化学、分子的数据已支持将这个种置于脱萼组 (Sect. *Gymnomeles* Koehne)^[1,3]。鉴定三叶海棠的腊叶标本时, 经常依据该种萌发枝或者嫩枝条上叶有分裂的特征, 但实际上脱萼组湖北海棠 (*Malus hupehensis*) 也有叶分裂的现象, 且不同地域的湖北海棠叶分裂程度有一定的差异; 在我国华东、华中地区, 三叶海棠叶分裂情况与湖北海棠相似, 容易混淆。近年来通过对我国三叶海棠进行大量的野外调查, 发现其与湖北海棠的区别在于花部萼片披针形, 比萼托筒长; 其嫩枝条、萌发枝或者秋发枝叶常常是 2~3 裂, 有浅裂和深裂, 且深裂者比湖北海棠萌发枝上的叶分裂程度要大; 呈灌丛状, 生于溪谷两侧的林下, 也与湖北海棠的生境不同^[1,4-5]。基于以上近缘种识别特征, 整理现有三叶海棠标本记录发现, 该种的分布比较零散, 区域较广, 其地理分布中心、多度中心和潜在分布区域仍未知。查询全球生物多样性信息咨询机构中三叶海棠的分布记录可知, 加拿大 2 个, 中国 16 个, 日本 120 个, 韩国 416 个, 吉尔吉斯斯坦 1 个, 英国 3 个, 美国 57 个, 共 615 个记录。排除人为栽培分布的国家, 只有中国、日本、韩国的分布属于自然分布记录。其中日本、韩国分布记录较多, 这主要是由于, 在有限区域内采集频度、强度都较大, 且这 2 个国家较早加入了全球生物多样性信息咨询机构, 上传网站数据较多。但这些记录不能真实反映三叶海棠的地理分布, 实际情况是三叶海棠广泛地自然分布于中国境内^[1-2]。因此有必要利用中国主要标本馆藏的三叶海棠标本分布记录数据, 探索其分布的详细情况。为此, 本研究采用来自中国 4 个植物标本馆的三叶海棠分布数据, 用最大熵模型 (MaxEnt) 模拟其地理分布, 旨在探明该种的基本分布格局、分布多度、分布中心和潜在分布区, 分析其地理分布与生物气候因子的关系, 明确其基本生态习性, 为开展相关的种质资源和谱系地理研究提供依据。

1 材料与方法

1.1 数据来源

本研究中重新鉴定的标本主要来自于中国科学院北京植物所标本馆 (PE)、江苏省中国科学院植物研究所植物标本馆 (NAS)、南京大学植物标本馆 (N) 和南京林业大学树木标本馆 (NF), 共获得鉴定较为明确的分布记录数据 291 个, 用地名数据库 (<http://v2.cvh.org.cn/>) 和 Google Earth (<http://www.google.com/earth/index.html>) 转换出分布点的经纬度坐标, 输入 Excel 表格, 保存成 *.CSV 格式。气候通用格式数据下载自 WorldClim (<http://www.worldclim.org/>) 的 1950—2000 年生物气候数据 (Bioclim), 分辨率 2.5 arc-minutes, 包括的 19 个生物气候变量为, 年均温 (bio1)、温度日较差 (bio2)、年均温占年温较差百分比 [$\text{bio3} = (\text{bio1} / \text{bio7}) * 100$]、温度季节变化方差 (bio4)、最热月极高温度 (bio5)、最冷月极低温度 (bio6)、年温较差 (bio7)、雨季均温 (bio8)、干季均温 (bio9)、暖季均温 (bio10)、冷季均温 (bio11)、年降雨量 (bio12)、最湿月降雨量 (bio13)、最干月降雨量 (bio14)、季降雨量变异系数 (bio15)、雨季降水量 (bio16)、干季降水量 (bio17)、暖季降雨量 (bio18)、冷季降雨量 (bio19)^[6]。将以上 19 个通用格式的生物气候变量数据用 DIVA-GIS^[7] 分别转换成 ASCII 格式文件, 然后直接导入 MaxEnt 软件。分析矢量底图用 1:4 000 000 的亚洲地图 (<http://www.diva-gis.org>) 和中国地图 (<http://nfgis.nsd.gov.cn>)。

1.2 研究方法

将三叶海棠标本记录数据 (CSV 格式) 和 19 个生物气候变量数据 (ASCII 格式) 导入 MaxEnt^[8-11]。随机选取三叶海棠标本分布数据的 25% 作为测试集 (test data), 其余分布数据作为训练集 (training data), 并从训练集中随机选取 10% 的分布数据来计算分布阈值, 这个分布阈值的统计学意义在于, 只允许训练集中不超过 10% 的分布数据有异常。最大迭代次数为 500 次, 收敛阈值为 10^{-5} , 输出数值为逻辑斯蒂值 (logistic), 表示分布概率 (简称为分布

值)。软件默认分布值为 0.5 的地区与已有标本的分布点吻合程度较高,其余参数也设为软件默认值。将输出结果中 ASCII 格式的文件直接导入 DIVA-GIS 软件,可生成所模拟物种分布的 grd 格式图。各地区分布值和分布范围用软件默认颜色来表示,分布值及相应的分布范围由 MaxEnt 划分,两种颜色相邻时,共有分布值就是它们的分界线。MaxEnt 采用受试者工作特征曲线(Receiver operatin characteristic curve,Roc Curve)检验模拟地理分布的适合度,要求训练集的受试者工作特征曲线下面积(The area under the roc curve,AUC)要大于测试集的受试者工作特征曲线下面积(AUC),而且都远大于随机模型下受试者工作特征曲线下面积(AUC) 0.5。MaxEnt 采用刀切法(Jackknife Test)检测生物气候变量与物种分布值增益之间的关系,建立了

3 种模型进行比对,可识别出对物种分布值增益影响最大的生物气候变量。将这些分析结果以 *.html、数学统计结果以 *.CSV 的格式自动保存于输出结果的文件夹中。

2 结果与分析

2.1 三叶海棠标本记录分布现状

从标本记录来看,三叶海棠分布在北纬 22.81°~43.07°,东经 99.30°~141.35°,分布北边达辽宁沈阳,西边达四川理塘,西南达云南维西,南达广东东莞,东达日本岛。从标本的密集度来看,主要集中在分布于中国的四川、甘肃、陕西、贵州、湖南、江西、浙江等地区及韩国和日本(图 1),但这种分布与采集标本的强度、频度和品质是密切相关的,不能反映其客观地理分布格局和可能的潜在分布区域。

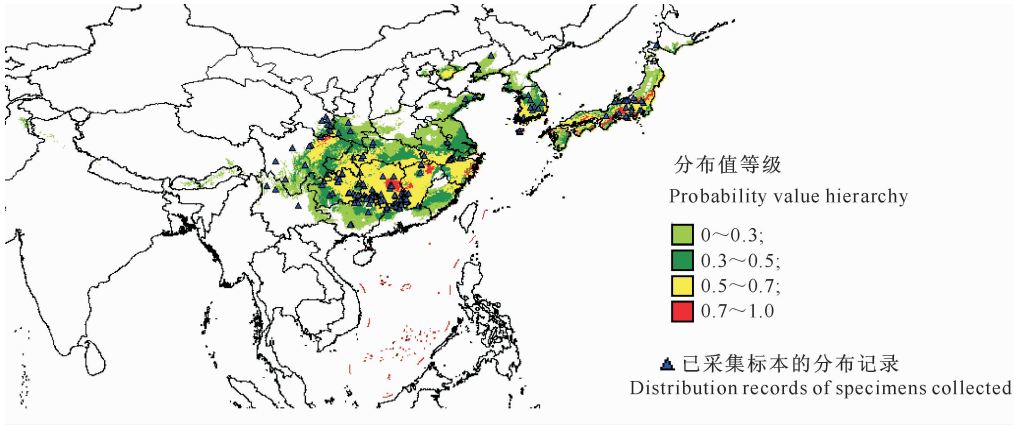


图 1 三叶海棠标本分布记录和最大熵模拟分布

Fig.1 Distribution of *M. toringo* specimen records and modeling results by MaxEnt

2.2 基于 MaxEnt 模拟三叶海棠分布

受试者工作特征曲线显示,训练集的 AUC 值为 0.991,大于测试集的 AUC 值(0.985),且均远大于随机模型预测的 AUC 值 0.5,说明模拟效果良好。分布阈值 0.313 由训练集数据中 10%的随机分布点产生,处于分布阈值(0.313)下的地区可以认为绝对不存在三叶海棠。将三叶海棠的分布数据导入 DIVA-GIS 软件后,软件将其分布值分为 4 个等级,其中,0~0.3 为浅绿色,0.3~0.5 为墨绿色,0.5~0.7 为黄色,0.7~1.0 为红色。各模拟分布区域对应其分布值,颜色自动匹配涂斑,相邻分布区的边界线就是共有分布值。各地区的模拟分布结果(图 1)如下。

(1)分布值为 0.3~0.5 的分布区域主要有,巴基斯坦东北部;不丹中部、东北部;印度北部;中国西藏东南部,横断山区,四川、重庆、甘肃、陕西交界地

区,湖北西北部、东南部,贵州西南部、中东部,广西北部,湖南西部、北部,江西东部、北部,安徽中部,福建大部分地区,浙江西部、南部、东部和北部,江苏中南部,山东东部沿海地区,北京市中南部,河北省东部与辽宁省交界地区;朝鲜西部沿海地区;韩国大部地区;日本岛大部地区。

(2)分布值为 0.5~0.7 的分布区域主要有 5 个区块,中国四川、甘肃、陕西交界地区;中国重庆,贵州大部地区,湖北西南部,湖南大部地区,广西与贵州、湖南交界的北部地区,江西,广东与湖南、江西交界的北部,福建东北部,浙江大部地区,安徽南部;北京西部山区、河北北部与辽宁西南部交界地区;韩国南部沿海地区;日本岛大部分地区。

(3)分布值为 0.7~1.0 的分布区域主要为 7 个相对间隔的区块,从西往东依次为:中国四川省涪江上游,湖南衡山和罗霄山,江西鄱阳湖东北部山区,

安徽霍县地区,浙江天目山和东部山区;韩国济州岛;日本岛大部地区。

以上各地区(分布值 0.3~1.0)仍没有采集标本记录,可定性为潜在分布区域,如巴基斯坦东北部,不丹中部、东北部,印度东北部,中国西藏东南部、河北省东北部、北京西部山区。其中具有高分布值(大于 0.5)的地区,三叶海棠存在的可能性较大。如本研究中福建省未有标本采集记录,但模拟结果显示,福建省东部地区分布值较高(0.5~0.7),查阅《福建植物志》^[11],可知该省宁化县分布有三叶海

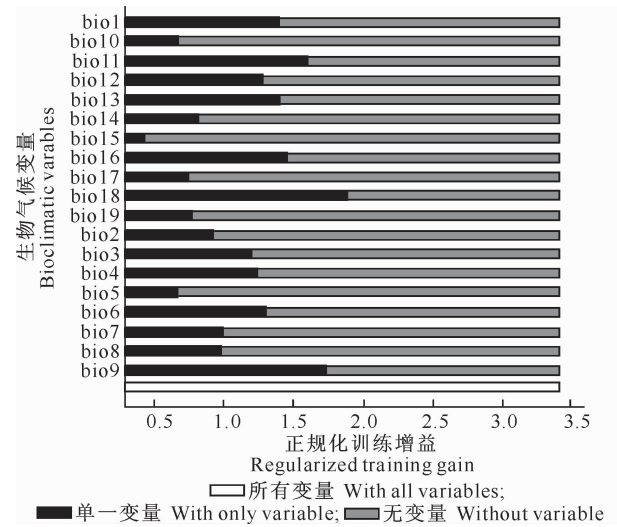


图 2 刀切法检测生物气候变量对分布的重要性
Fig. 2 Effects of bioclimatic variables on gain of distribution using Jackknife Test

3 结论与讨论

3.1 MaxEnt 模拟三叶海棠分布的误差来源及分析

MaxEnt 的工作原理是,利用物种已有分布点所反映的生物气候变量的概率密度来估计一种分布,要求这种分布与景观背景取样为均匀分布时所反映的生物气候变量的概率密度间的相对熵最小。从方法上来看,模拟误差主要来源于物种分布记录与景观背景取样方法的不一致^[10]。虽然受试者工作特征曲线表明本研究模拟效果良好,但由于现有标本分布记录的取样方法是未知的,且记录分布存在明显的偏向性,如中国北部、东部记录偏少,而西南部记录偏多,这可能是造成模拟分布和实际情况不一致的主要原因。本研究模拟结果显示,高分布值(0.7~1.0)地区与标本分布记录数较多的地区也不一致。如中国西南地区标本分布记录较多,但其具高分值(0.7~1.0)的分布范围小。MaxEnt 模型

棠。

2.3 三叶海棠地理分布与生物气候变量的关系

用刀切法(Jackknife Test)检测生物气候变量对于分布增益的贡献,结果(图 2)显示,暖季降雨量(bio18)对三叶海棠分布的增益最大。暖季降雨量与分布值之间的响应曲线(图 3)显示,降雨量为 250~600 mm 时,分布值随降雨量的升高而增大;降雨量为 600~900 mm 时,分布值随降雨量升高而减少。以分布值 0.5 为界,说明三叶海棠喜生于温暖季节降雨量在 450~800 mm 的地区。

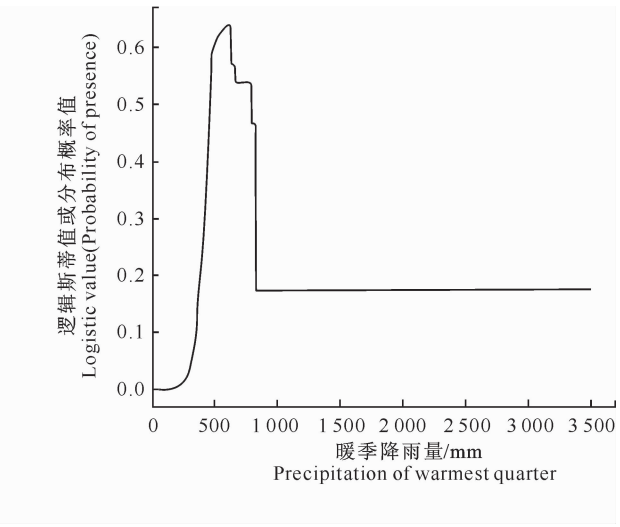


图 3 预测分布值与暖季降雨量的响应曲线
Fig. 3 Response curves between prediction and precipitation in warmest season

本身并不能绝对判断物种在某一地区的存在或不存在,只能给出物种分布于某地点的可能概率,所以一般情况下,软件系统默认分布值为 0.5 的地区应与标本记录丰富的典型分布区吻合。本研究中,虽然中国四川、云南、广西、广东的标本分布记录较多,但模拟结果显示,这些地方的分布值和分布范围偏小。这种偏差也可能是景观背景上均匀取样范围较广、密度较大(10 000 次),加大了与现有标本分布记录取样方法的不一致程度。

3.2 三叶海棠的潜在分布和地理分布格局

本研究发现,一些三叶海棠潜在分布的地区,现仍未见有关三叶海棠的标本和信息记载,如巴基斯坦东北部,不丹中部、东北部,印度东北部,中国西藏东南部、河北省东北部、北京西部山区。其中印度东北部,不丹,中国西藏东南部是苹果属植物多样性较高的地区,种间杂交、多倍化和基因渐渗较为普遍^[1-2],这些地区的苹果属植物标本,仅凭叶特征的分类容易鉴定为花楸苹果组的一些类群,实际可能

存在三叶海棠,有待实地调查。中国河北省东部山区、北京西山地区无三叶海棠的自然分布记录,模拟为潜在分布,说明这些地区的气候条件适合引种栽培三叶海棠。辽宁省只有 2 份标本记录,且大部分区域分布值低于 0.3,因此这一地区三叶海棠是否存在值得怀疑,核查原始标本,由于采集点未详细到具体山名,因此人为栽培可能性较大,本研究将其定为非野生分布。本研究也未见河南的标本记录,但模拟结果显示,河南省中部和东南部具有分布值 0.3~0.7 的微弱点状分布,《河南植物志》记载在伏牛山有三叶海棠分布,其中将三叶海棠的花萼形态描述为三角卵形,从插图中看却是典型的披针形^[12],可见《河南植物志》也主要依据叶有 3~5 分裂来鉴定三叶海棠,但实际上花萼三角卵形,叶有分裂等特征的标本,是湖北海棠。因此,本研究将河南省暂定为无自然分布。以上三叶海棠分布有争议的地区,尚有待实地考察。

从直观表现来看,MaxEnt 模拟的三叶海棠地理分布中,低分布值(0.3~0.5)和高分布值(0.7~1.0)区域都与现有标本的分布记录吻合程度不高,但分布值为 0.5~0.7 时所标定的区域与现有标本记录吻合程度相对较高。因此用分布值 0.5~0.7 的分布区域,排除其潜在分布区域,再结合这些区域地理间隔的特点,从宏观景观上将三叶海棠的现实地理分布格局划分为 4 个区块:中国四川、甘肃、陕西三省交界地区;中国重庆大部分地区,贵州大部地区,湖北西南部,湖南大部地区,广西与贵州、湖南交界地区,江西,广东与湖南、江西交界的北部,福建东北部,浙江大部地区,安徽南部地区;韩国南部沿海地区;日本岛大部分地区。

3.3 三叶海棠地理分布与生物气候变量的关系

刀切法(Jackknife Test)检测表明,暖季降雨量(bio18)对三叶海棠分布的增益最大,考虑到分布值为 0.5 的地区与现有三叶海棠标本分布点吻合程度高,因此以分布值 0.5 为界,推断出三叶海棠喜生于温暖季节降雨量在 450~800 mm 的地区。苹果属脱萼组山荆子(*Malus baccata*)喜生于干旱季节降雨量较大的地区,湖北海棠喜生于干旱季节温度较高的地区^[1,13]。从这些生态习性来看,三叶海棠及其近缘种各自占不同的生态位,它们应该有独立的地理分布区域,形成地理分布替代现象。但实际上随复杂地形的变化,各地区生物气候变化复杂,这些近缘种在地理分布上出现了一定的重叠镶嵌分布现象^[1]。笔者认为,苹果属脱萼组近缘种间的这种镶

嵌分布格局与第四纪冰期气候震荡相关,随气候震荡,它们的分布区发生过反复的变化,基因流的隔离和多次相遇,可能形成一些杂交的地理居群,这也正是无法准确定义这些分类群分类学地位、系统发生关系、分布界限的主要原因。研究三叶海棠地理分布与生物气候变量的关系,也只能从宏观上探明该种的生态学特性。未来基于三叶海棠的地理分布格局,采用分子谱系地理学方法和手段,有望解释该种地理分布格局形成的原因。

此外,本研究未考虑地形、土壤和植被对三叶海棠地理分布的影响。从现有标本记录信息来看,地形的变化明显影响其垂直分布,如在中国四川垂直分布海拔为 3 700~3 800 m,华东、中南地区为 900~1 500 m^[1]。但本研究所用标本多数没有海拔记录数据,因此未将海拔作为一个影响因子进行分析。

[参考文献]

- [1] 钱关泽. 苹果属(*Malus* Mill.)分类学研究 [D]. 南京:南京林业大学,2005:169-176.
Qian G Z. The taxonomic study of the genus *Malus* Mill. [D]. Nanjing: Nanjing Forestry University, 2005: 169-176. (in Chinese)
- [2] 中国科学院中国植物志编辑委员会. 中国植物志(36 卷) [M]. 北京:科学出版社,1974:387-389.
Delectis Florae Reipublicae Popularis Sincae Agendae Academiae Sincae Edita. Flora reipublicae popularis sincae tomus 36 [M]. Beijing: Science Press, 1974: 387-389. (in Chinese)
- [3] Robinson J P, Harris S A, Juniper B E. Taxonomy of the genus *Malus* Mill. (Rosaceae) with emphasis on the cultivated apple, *Malus domestica* Borkh. [J]. Plant Systematics and Evolution, 2001, 226(1): 35-58.
- [4] 陈 曦. 湖北海棠(*Malus hupehensis*)不同居群变异式样及遗传多样性的研究 [D]. 南京:南京林业大学,2009:30-37.
Chen X. A study on variation patterns and genetic diversity of *Malus hupehensis* populations [D]. Nanjing: Nanjing Forestry University, 2009: 30-37. (in Chinese)
- [5] 王雷宏. 山荆子变异式样研究 [D]. 南京:南京林业大学,2008: 24-33.
Wang L H. A study on variation patterns in *Malus baccata* (L.) Borkh. [D]. Nanjing: Nanjing Forestry University, 2008: 24-33. (in Chinese)
- [6] Hijmans R J, Cameron S E, Parra J L, et al. Very high resolution interpolated climate surfaces for global land areas [J]. International Journal of Climatology, 2005, 25: 1965-1978.
- [7] Jarvis A, Guarino L, Williams D, et al. The use of GIS in the spatial analysis of wild peanut distributions and the implications for plant genetic resources conservation [J]. Plant Genetic Resources Newsletter, 2002, 131: 29-35.