

# 利用最大熵模型和规则集遗传算法模型预测孑遗植物裸果木的潜在地理分布及格局

马松梅<sup>1,2</sup> 张明理<sup>1,3\*</sup> 张宏祥<sup>1,2</sup> 孟宏虎<sup>1,2</sup> 陈曦<sup>4\*</sup>

<sup>1</sup>中国科学院干旱区生物地理与生物资源重点实验室, 中国科学院新疆生态与地理研究所, 乌鲁木齐 830011; <sup>2</sup>中国科学院研究生院, 北京 100049; <sup>3</sup>中国科学院植物研究所, 北京 100093; <sup>4</sup>新疆遥感与地理信息系统应用实验室, 中国科学院新疆生态与地理研究所, 乌鲁木齐 830011

**摘要** 利用野外调查的16个居群分布点和7个环境因子图层, 选择最大熵模型(MAXENT)和规则集遗传算法模型(GARP), 在地理和环境空间上模拟了第三纪孑遗植物裸果木(*Gymnocarpus przewalskii*)在中国西北地区的潜在分布。结果表明: (1)裸果木的潜在适生区全部集中在西北荒漠区, 其中最佳适生区主要集中在3个区域, 一是河西走廊中部和玉门以西、宁夏北部及内蒙古乌拉特后旗; 二是塔里木盆地西北缘; 三是柴达木盆地西北缘两片极小的高度适生区。裸果木的生态位被确定在一个较广的干旱环境空间: 适生区极端最高气温基本上在29.2–36.8 °C之间, 极端最低气温在-18.3至-13.4 °C之间; 年平均降水量40–200 mm; 潜在蒸发率在3–15之间。(2) MAXENT和GARP模型都较好地预测了裸果木的潜在分布, 但GARP产生了相对较大、较连续的潜在分布区, 部分过预测了破碎化生境; 而MAXENT预测到的潜在分布区, 在不同区域具有不同的环境适生性指数, 而且成功地排除了不合理的破碎化分布, 从而更直观地展示了裸果木的潜在分布格局和生态位要求。

**关键词** 规则集遗传算法模型, 裸果木, 留一法, 最大熵模型, 潜在分布区

## Predicting potential geographical distributions and patterns of the relic plant *Gymnocarpus przewalskii* using Maximum Entropy and Genetic Algorithm for Rule-set Prediction

MA Song-Mei<sup>1,2</sup>, ZHANG Ming-Li<sup>1,3\*</sup>, ZHANG Hong-Xiang<sup>1,2</sup>, MENG Hong-Hu<sup>1,2</sup>, and CHEN Xi<sup>4\*</sup>

<sup>1</sup>Key Laboratory of Biogeography and Bioresource in Arid Land, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Ürümqi 830011, China; <sup>2</sup>Graduate University of Chinese Academy of Sciences, Beijing 100049, China; <sup>3</sup>Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China; and <sup>4</sup>Xinjiang Laboratory of Geographic Information System and Remote Sensing Application, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Ürümqi 830011, China

### Abstract

**Aims** Understanding the geographical distribution and patterns of endemic species is critical to biodiversity conservation and biogeographical history reconstruction of the area occupied by the species. The ecological niche models (ENMs) are useful techniques to explore the links between the species distribution and the environmental data. The displayed potential habitats for the species, in turn can enable the examination of the predictive abilities of various ENMs. We attempt to determine the potential geographic distributions of the Tertiary relic plant *Gymnocarpus przewalskii* based on the model of maximum entropy (MAXENT) and genetic algorithm for rule-set prediction (GARP).

**Methods** Based on sixteen sampled localities and seven environmental layers (isothermality, maximum temperature, minimum temperature, annual precipitation, potential evapotranspiration ratio (PER), altitude and soil types), we conducted predictions of *G. przewalskii* using MAXENT and GARP models. The spatial distribution maps of *G. przewalskii* with the different environmental suitable values (MAXENT) or overlap index (GARP) displayed the distribution patterns clearly.

**Important findings** The potential distributions of *G. przewalskii* with the highest environment suitability are predicted at first in the middle of Hexi corridor and the western Yumen of Gansu Province, the north of Ningxia-Hui Autonomous Region, and the part of the Wulate banner of Inner Mongolia in China. The others are mainly in northwestern Tarim Basin and small isolated areas in northwestern Qaidam Basin. Both MAXENT and GARP produced good predictions for *G. przewalskii*. However, GARP predicted larger and more continuous suitable habitats around the species' locations and some isolated and fragmented spatial predictions where the species has never been found or collected before. MAXENT predicted a distribution that is a logical proportion of the study

收稿日期Received: 2010-05-19 接受日期Accepted: 2010-07-25

\* 通讯作者Author for correspondence (E-mail: zhangml@ibcas.ac.cn, chenxi@ms.xjb.ac.cn)

area and removed most of the unlikely isolated habitats.

**Key words** Genetic Algorithm for Rule-set Prediction (GARP), *Gymnocarpos przewalskii*, leave one out, Maximum Entropy (MAXENT), potential geographic distribution

了解物种的地理分布对于保护和管理生物多样性, 探索物种分布格局, 以及解决各种有关生态、生物地理和生物进化等问题都十分重要 (Margules & Sarkar, 2000; Guisan & Thuiller, 2005)。但对于大多数地区和物种来说, 通常缺乏详细的地理分布信息, 野外数据收集不仅非常耗时费力 (Ottaviani *et al.*, 2004), 而且依靠采样获得的物种分布数据通常不够完整, 很容易受采集路线或记录错误等主观影响, 可能使得分布记录不能够代表物种真正的地理分布界限 (Graham *et al.*, 2004)。所以, 目前获得的物种地理分布信息在很大程度上依赖于生态位模型 (ecological niche models, ENMs) (Guisan & Harrell, 2000)。ENMs 主要利用生态位原理, 将物种的已知分布点, 例如标本馆的标本记录, 转换成物种假定的分布界限 (da Fonseca, 2000), 不仅可以鉴别出一些被以前的采集工作忽略了物种分布区, 而且能够清楚地估计出潜在的分布格局。

ENMs 也叫物种分布模型 (species distribution models), 它主要依赖物种的已知分布数据和一系列环境变量来探索物种的生态位要求和潜在分布, 这些信息对于研究和保护一些珍稀濒危的孑遗物种尤其重要。因为孑遗物种往往具有较小的居群数量, 地理分布通常被限制在狭小和有限的区域内, 因而也更容易受到外界的干扰而灭绝 (Linder, 1995; Peterson & Watson, 1998)。因此, 利用 ENMs 预测特有种的潜在地理分布及格局, 将会对分布区的生物地理历史重建、生物多样性保护和管理等工作提供非常有用的指导 (Linder, 1995; Caldecott *et al.*, 1996)。

目前, ENMs 已经在很多地区被成功地应用于预测物种的潜在分布区。但是, 每种模型都有其自身的优缺点, 很多模型的对比研究都表明, 最大熵模型 (MAXENT; Phillips *et al.*, 2006) 和规则集遗传算法模型 (GARP; Stockwell & Peters, 1999) 具有较好的预测能力 (Stockwell & Peterson, 2002; Tsoar *et al.*, 2007)。更重要的是, MAXENT 和 GARP 也非常适合模拟具有较少分布数据的稀有或濒危物种, MAXENT 可以在样本个数  $\geq 5$  的情况下进行成功的

预测, GARP 可以在样本  $\geq 10$  的情况下取得较好的预测结果 (Pearson *et al.*, 2007)。本研究基于野外调查的有限分布数据, 选择 MAXENT 和 GARP 为预测工具, 在地理和生态空间上模拟裸果木的潜在分布, 分析其分布格局及原因, 并比较两个模型的预测特点和能力。

裸果木 (*Gymnocarpos przewalskii*) 隶属于石竹科裸果木属 (*Gymnocarpos*), 高 20–100 cm, 为超旱生小灌木, 叶片稍肉质, 线形; 聚伞花序, 花小, 不显著; 瘦果, 喜生于荒漠区海拔 1 000–2 600 m 的砾石戈壁、冲积扇和干河床等。裸果木属在全世界有 10 个种, 从西到东分布于佛得角群岛 (大西洋岛国) 和金丝雀群岛 (大西洋岛国), 分布中心在非洲的最东部, 即“非洲之角”。裸果木在地理上与其他种呈隔离分布, 主要分布在我国西北地区, 少量分布在蒙古南部 (Oxelman *et al.*, 2002)。裸果木为亚洲中部荒漠区第三纪孑遗植物, 稀有种, 是 1987 年我国首批公布的第一批 II 级保护植物 (傅立国, 1992)。我们对 16 个裸果木的自然居群进行了野外调查, 发现只有在甘肃阿克塞县和新疆乌恰县的两个居群中, 裸果木为优势种, 个体数量较多。在其他地区的居群中, 裸果木的个体数量稀少, 很多还遭到骆驼的啃食, 并且其生境因采矿、修路等人为干扰遭到了严重的破坏, 亟待保护。

前人对裸果木的研究, 主要体现在分类学 (Bittrich, 1993; Petrusson & Thulin, 1996; Oxelman *et al.*, 2002)、孢粉学 (党荣理和姜彦成, 1996) 和引种 (刘生龙等, 1995), 以及形态解剖学、细胞学、组织培养 (汪之波等, 2004) 等方面。关于裸果木的分布区, 刘永生等 (2008) 做过一些调查和记录。蒋霞和倪健 (2007) 利用标本分布数据和环境因子的关系, 从理论上推测了其最大潜在分布和中心分布范围。但是关于该种在不同地区的适生程度、潜在分布格局等信息并不清楚。本文在对裸果木居群进行野外调查的基础上, 利用 MAXENT 和 GARP 模拟裸果木的潜在地理分布, 旨在解决两个问题: (1) 模拟并分析裸果木的生态位、潜在分布及格局; (2) 比较 MAXENT 和 GARP 基于有限分布数据的预测特点和能力。

# 1 研究区域概况及材料方法

## 1.1 研究区域概况

为了找出裸果木在我国西北荒漠区所有可能的适生区域, 基于其16个居群的实际分布状况, 选择向东和向南各扩展100 km (73°40'–114°40' E, 32°3'–49°11' N)的范围, 作为研究区域。具体地理范围, 在行政区划上主要包括新疆全境、甘肃全境、内蒙古西部地区、宁夏全境、青海北部和山西大部分地区及西藏部分地区。研究区域内地形地貌复杂, 主要由盆地、高原及构造剥蚀山地相间构成。大的盆地有准噶尔盆地、塔里木盆地、东疆盆地和柴达木盆地; 高原有阿拉善高原和鄂尔多斯高原。高原四周为几个巨大的山系所环绕: 南部有青藏高原外围的昆仑山和祁连山; 东部有贺兰山和狼山; 北部有阿尔泰山和戈壁阿尔泰山。研究区域的地理范围包括了党荣理等(2002)定义的整个西北荒漠区——

我国年降水量不足200 mm, 干燥度大于4, 水文网稀疏且多呈内流性, 并以干旱土和荒漠植被为标志特征的广大地区, 包括新疆的准噶尔盆地、塔里木盆地、东疆盆地、甘肃的河西走廊、青海的柴达木盆地和内蒙古西部的阿拉善高原。

## 1.2 数据获取和研究方法

### 1.2.1 裸果木分布数据

根据标本记载, 我们于2009年7–10月对裸果木的大部分自然分布区进行了野外调查, 共调查到16个居群, 其中甘肃6个, 内蒙古1个, 宁夏1个, 新疆8个(表1)。每个分布点都有GPS精确定位的经纬度, 代表一个独立的环境条件。

### 1.2.2 环境图层数据

所用的7个环境因子包括: 等温性(昼夜温差与年温差的比值×100)、极端最高温度、极端最低温度、年平均降水量、可能蒸散率(PER)、海拔和土壤类型。其中海拔数据来自美国地质调查局(USGS,

表1 裸果木16个居群的基本情况

Table 1 Basal information of the 16 populations of *Gymnocarpus przewalskii*

地区 Region	居群地点 Location of population	经度 Longitude	纬度 Latitude	海拔 Altitude (m)	生境 Habitat
甘肃 Gansu	肃南县白银乡公路旁 Beside highway, Baiyin Township, Sunan County	99°34' E	38°52' N	2 513	砾石山坡 Gravel slopes
	金塔县鼎新镇黑河 Heihe River, Dingxin Town, Jinta County	99°30' E	40°18' N	1 180	山前洪积扇 Piedmont diluvial
	玉门市昌马乡水峡村后 Behind Shuixia Village, Changma Township, Yumen City	96°42' E	39°48' N	2 312	干河床 Dry wash
	肃北县水电站后 Behind hydropower station, Subei County	94°53' E	39°29' N	2 213	干河床 Dry wash
宁夏 Ningxia	阿克塞县公路旁 Beside highway, Akesai County	94°16' E	38°28' N	2 636	砾石干河滩 Gravel dry beach
	柳园县车站附近 Nearby the Liuyuan County station	95°24' E	39°6' N	1 789	砾石洪积扇 Gravel diluvial
	中卫市沙坡头 Shapotou, Zhongwei City	105°11' E	37°31' N	1 128	干河床砾石堆 Gravel pile of the dry riverbed
内蒙古 Inner Mongolia	阿拉善左旗公路旁 Beside highway, Alashan Zuoqi	105°42' E	38°48' N	1 625	山坡砾石 Hillside gravel
新疆 Xinjiang	哈密市庙尔沟公路旁 Beside highway, Miao'ergou, Hami City	93°35' E	43°5' N	1 173	砾石干河滩 Gravel dry beach
	奇台县北塔山牧场 Pasture of Beita Mountain, Qitai County	90°28' E	44°59' N	1 134	石质残丘 Rocky remnant dunes
	轮台县塔拉克乡北5 km 5 km to north Taerlake Township, Luntai County	84°15' E	41°46' N	979	砾石戈壁 Gravel gobi
	库车县克孜尔乡千佛洞 Qianfo Caves, Kezier Township, Kuche County	82°24' E	41°48' N	1 179	河滩砾石堆 Gravel pile of the beach
	拜城县北10 km 10 km to north Baicheng County	81°50' E	41°49' N	1 260	砾石堆 Gravel pile
	柯坪县县城北10 km公路旁 Beside highway, 10 km to north Keping County	79°4' E	40°33' N	1 175	水坝旁山坡 Slopes beside the Dam
	疏附县乌帕乡东10 km公路旁 Beside highway, 10 km to east Wupa Township, Shufu County	75°50' E	39°25' N	1 359	山前洪积扇 Piedmont diluvial
	乌恰县康苏镇东5 km公路旁 Beside highway, 5 km to east Kangsu Town, Wuqia County	75°06' E	39°42' N	2 198	砾石戈壁 Gravel gobi

doi: 10.3773/j.issn.1005-264x.2010.11.010

<http://edc.usgs.gov/products/elevation/gtopo30/hydro/asia.html>), 土壤数据来自国际地圈-生物圈计划(IGBP-DIS, 2000)。并计算了一个反映温度和降水的综合指标——PER, PER系由生物温度和年平均降水量两个图层进行栅格计算得到(蒋霞和倪健, 2005)。生物温度是指植物出现营养生长范围内的平均温度(蒋霞和倪健, 2005), 根据Holdridge等(1971)的定义, 介于0–30 °C之间, 所有不在这个范围内, 即小于0 °C和大于30 °C的温度都用0 °C代替。PER等于1, 代表潜在蒸发和降水在一个较长的时期内是平衡的, 大于和小于1分别代表干旱和湿润程度的不断增加。

其他4个气候因子是通过来自环境气象网站WORLDCLIM (<http://www.worldclim.org>)的19个环境因子图层, 提取主成分(PCA)所得。由于各环境因子之间往往具有很强的相关性, 尤其在高维的环境空间中, 很容易导致过拟合, 所以本研究利用主成分分析法提取了3个主成分轴, 总的解释了19个因子中大于80%的变异, 其中贡献率较大的第一主成分主要代表温度变量, 第二主成分代表降水。同时利用皮尔森相关(Pearson correlation)检查了所获得的各主成分因子之间的共线性, 具有较高平均相关性(> 90%)的变量集就用一个变量代替(Ward, 2007), 最终得到了所用的4个环境因子。

### 1.2.3 模型预测原理与数据处理方法

GARP预测基于遗传算法原理, 首先将预测数据随机分成训练数据(用于训练并构建模型)及验证数据(用于模型的质量检验), 并通过不断迭代进行规则集的选择、评价、验证、接受或拒绝。从4种待选规则集(原子规则(atomic)、逻辑回归(logistic regression)、生物气候包络(bioclimatic envelope)和逆生物气候包络(negated bioclimatic envelope))中归纳出一种方法, 应用到训练数据, 生成一个法则, 根据验证数据判断这个法则是否可提高预测的准确性, 这个算法通常反复运行多次或根据收敛条件终止, 最终形成一个由不同规则共同组成的模型, 从而可预测物种的潜在分布区(Stockwell & Peters, 1999)。

MAXENT (Phillips *et al.*, 2006)是一种基于熵最大原理的预测模型。最大熵理论认为: 在已知条件下, 熵最大的事物最接近它的真实状态, 因而预测的风险也越小。MAXENT是从不完整的已知信息

中作出推断或预测, 利用物种分布数据和环境图层, 探索物种已知分布区的环境特征与研究区域的非随机关系, 在满足一定限制条件(能够代表目标分布的不完整信息)的情况下, 找到熵最大的概率分布(即最均匀的分布)作为最优分布, 用于物种的适生区预测。

由于裸果木的分布数据有限(< 25), 不适合被随机分割成独立的训练和评价数据集, 所以选择留一法(leave one out)来建立和评价模型(Fielding & Bell, 1997; Pearson *et al.*, 2007)。即每次有一个发生点从所有的记录点(N)中被去除, 用于评价模型, 剩余的N-1个点用于训练模型, 这样最终可以得到N(16)套发生数据集和N(16)个模型。得到的发生数据集和环境图层, 均按照两个模型的要求输入MAXENT 3.2.19 (<http://www.cs.princeton.edu/~schapire/maxent>)和GARP 1.1.6 (<http://www.lifemapper.org/desktopgarp>)。为便于模型结果的解释和可视化, 选择一个阈值来区分物种的适生区和非适生区是很有必要的(Pearson *et al.*, 2004, 2007)。本文选择最低存在阈值(lowest presence threshold), 保证所有的观察点都被正确地预测为存在, 而且所预测到的生境至少和物种的已知分布区具有相同的适生性(Hernandez *et al.*, 2006; Pearson *et al.*, 2007)。

将MAXENT得到的16个模型, 在ArcGIS 9.2中进行栅格叠加及平均值计算, 并作为裸果木的最终预测结果。为了获得较稳定的预测结果, GARP对每套发生数据集都重复运行了100次, 这样共得到1 600个模型, 再按照Anderson等(2003)选择最佳模型的方法, 依照16套发生数据集, 从每100个模型中选出10个最佳模型进行叠加, 最后同样将16个最佳模型进行栅格叠加及平均值计算, 并作为最终预测结果。分别统计两个模型预测为存在的测试点个数, 计算模型预测的成功率(被成功预测为存在的测试点个数占所有测试点个数的比例), 并计算MAXENT和GARP在随机情况下预测到的潜在适生区面积占研究区域总面积的比例, 以检测模型预测的成功率是否具有显著的统计意义(Pearson *et al.*, 2007)。

## 2 结果

### 2.1 裸果木的潜在地理分布及格局

MAXENT得到的是裸果木在不同区域(即不同的环境条件下), 适生程度按不同适生性指数进行

累加的潜在分布图(图1)。GARP得到的是二进制的结果(图2), 即0和1表示的存在与不存在, 图中叠加指数(overlap index)是指预测为存在的模型个数, 用来表示预测为存在的可能性程度。比较两模型预测的结果, GARP预测的潜在分布总体上和MAXENT一致, 但模型在裸果木的已知分布区周围产生了较大而且连续的合适生境, 在没有该种发生记录的地区也产生了很多过预测的破碎化分布, 比如叠加指数为9的区域: 山西很多地区, 内蒙古草原区, 及新疆准噶尔地区等。所以, 尽管 $t$ 检验表明两个模型预测结果没有明显的不同, 但对于裸果木的潜在分布区, 我们认为MAXENT预测的结果更合理, 而且累加的环境适生程度也更容易了解物种的生态位。所以从图1可以看出, 裸果木的最适生长分布区: 一部分主要围绕干旱中心内蒙古额济纳旗的东部, 集中在甘肃河西走廊中部及玉门以西、宁夏北部, 还有少量分布在内蒙古乌拉特后旗; 另一部分集中分布在另一干旱中心——新疆塔克拉玛干沙漠的西北缘; 另外, 模型灵敏地识别到在柴达木盆地西北缘也存在两片极小的适生区。此外, 预测结果显示, 裸果木在新疆中天山和南天山的部分地区、哈密盆

地、内蒙古阿拉善地区(阿拉善左右旗, 西鄂尔多斯), 也都具有较高的适生性。显然, 裸果木的潜在适生区全部集中在西北荒漠区, 围绕两个干旱中心呈零星分布。

### 2.2 裸果木的生态位

结合MAXENT和GARP预测的结果, 分析裸果木的生态位, 被确定在一个较广的干旱环境空间: 裸果木适生区的极端最高气温基本上在29.2–36.8 °C之间, 极端最低气温在-18.3–-13.4 °C之间; 最大年平均降水量在裸果木最适生区之一的河西走廊中部能达到100–200 mm, 而在塔里木盆地西北缘只有40–80 mm; 分布区内的可能蒸散率在3–15之间, 在塔里木盆地西北缘通常达到10–15, 其他适生区是3–6。

### 2.3 MAXENT和GARP模型

留一法评价结果表明: 对于裸果木, MAXENT表现出了较强的预测能力, 16个测试点全部被成功地预测为存在, 并且不同模型的累加预测到了较大的潜在分布面积,  $p$ 值也表明所有模型的预测都具有较强的统计意义( $p < 0.0112$ ) (表2)。GARP所产生的模型也具有最高的预测成功率和显著的统计意

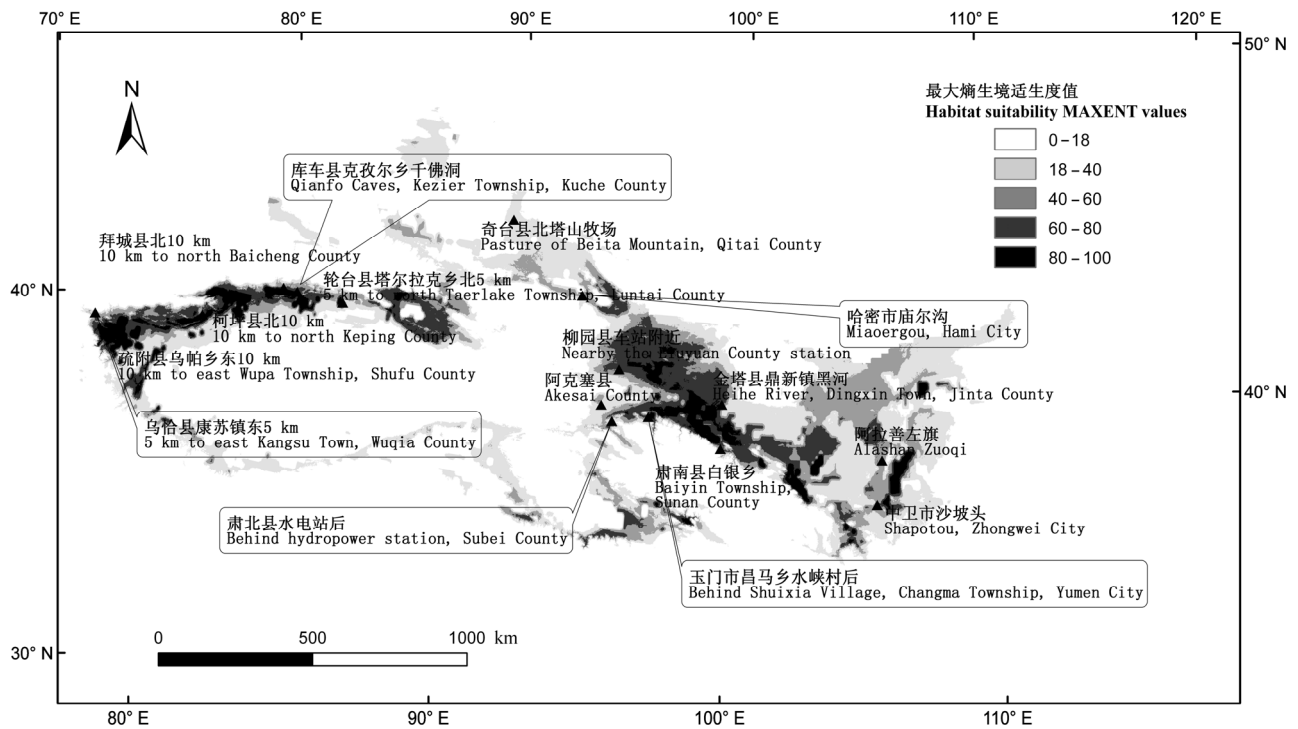


图1 最大熵模型(MAXENT)预测的裸果木的潜在分布区。

Fig. 1 Potential distribution areas of *Gymnocarpus przewalskii* predicted by using Maximum Entropy (MAXENT).

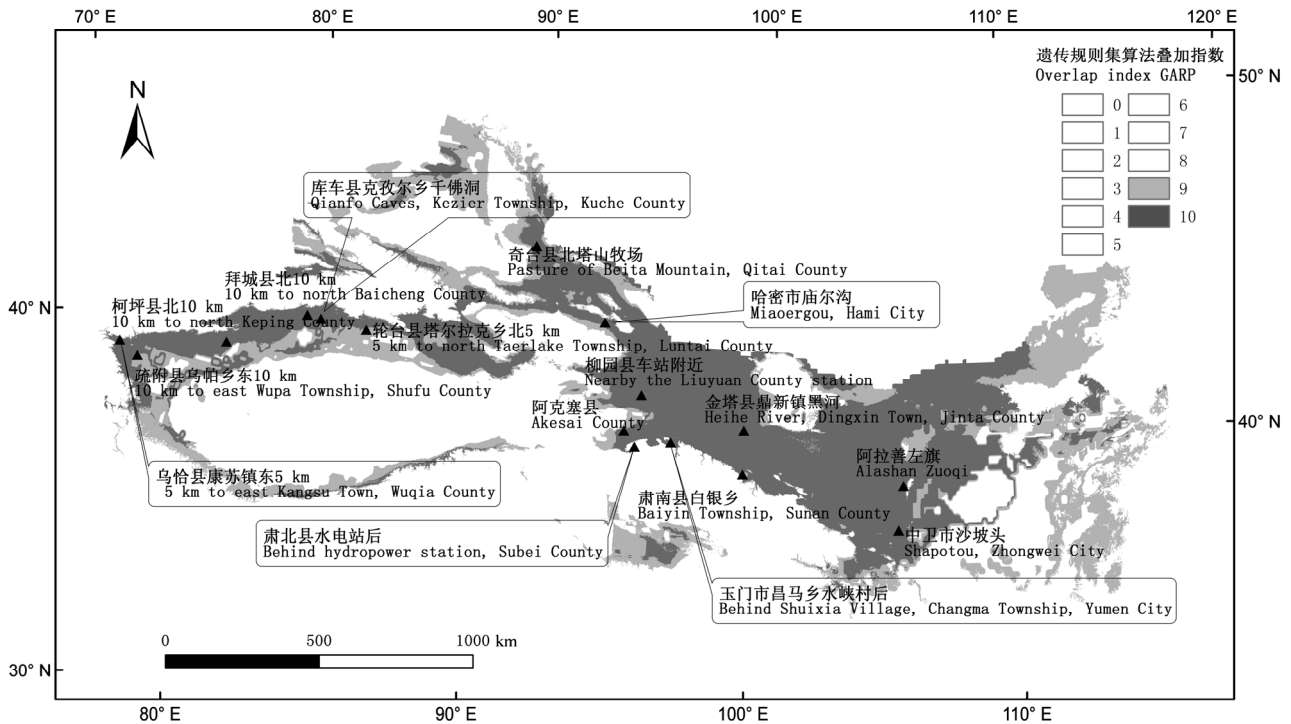


图2 规则集遗传算法模型(GARP)预测的裸果木的潜在分布区。

Fig. 2 Potential distribution areas of *Gymnocarpus przewalskii* predicted by using Genetic Algorithm for Rule-set Prediction (GARP).

表2 最大熵模型(MAXENT)和规则集遗传算法模型(GARP)利用留一法预测的结果

Table 2 Results of Maximum Entropy (MAXENT) and Genetic Algorithm for Rule-set Prediction (GARP) by using the leave one out

模型 Model	最小存在阈值 Lowest predicted value threshold	预测成功的个数 No. of successes predicted	平均预测的面积比例 Mean fractional area predicted	<i>p</i>
MAXENT	18	16	0.159 4	0.011 2
GARP	9	16	0.226 6	0.020 8

义( $p < 0.0208$ ), 所预测到的潜在分布面积是MAXENT的1.4倍左右(表2)。从两种模型在不同阈值下的预测成功率变化趋势可以看出(图3), MAXENT的成功率在阈值5、10和15的时候都达到了最高值1, 在阈值逐渐高于55后, 预测成功率逐渐低于50%, 而在最大阈值为100时预测成功率降为0, 呈现出明显的降低趋势; 而GARP的预测成功率随着阈值的不断增加呈现一个相对平稳的趋势, 在阈值从5到90 (对于GARP相当于叠加指数为0.5到9.0)增加的过程中, 成功率一直保持在1, 当阈值达到95和100时, 预测成功率恒定保持在0.8以上。显然, MAXENT的预测成功率受不同阈值的影响较大。 $p$ 值显著性测验表明: 除了MAXENT模型在LPT (lowest predicted value threshold)为5和10的预测以

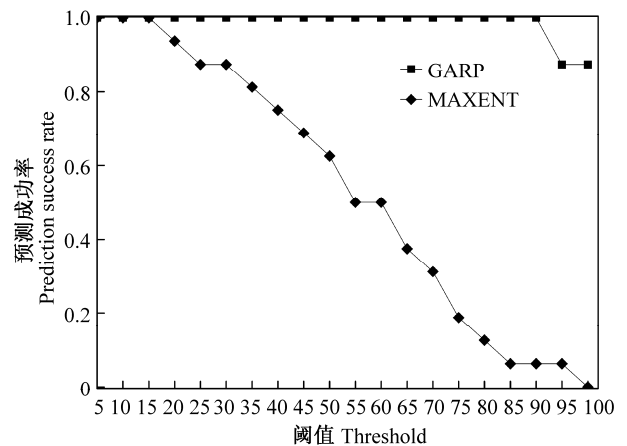


图3 最大熵模型(MAXENT)和规则集遗传算法模型(GARP)在不同最低存在阈值下的预测成功率。

Fig. 3 Prediction success rate predicted by using Maximum Entropy (MAXENT) and Genetic Algorithm for Rule-set Prediction (GARP) at different lowest predicted value threshold.

外, 所有预测成功率大于50%以上的模型都具有显著的统计意义( $p < 0.05$ )。

### 3 讨论

#### 3.1 MAXENT和GARP作为分布预测工具的可行性

模型的预测结果表明: MAXENT和GARP可以利用有限的发生数据, 进行裸果木潜在分布区的较好预测。GARP模型虽然产生了破碎化的过预测生境, 但也可以产生成功率较高、且具有显著统计意义的预测。有研究表明了GARP可以成功地预测到物种大部分的分布点, 同时也会预测到远远大于物种已知分布区的分布范围(Anderson *et al.*, 2003; Hernandez *et al.*, 2006; Peterson *et al.*, 2007)。尽管过预测问题可能是基于生态位原理的模型难以避免的, 但仍然被很多研究者认为是有用的(Anderson *et al.*, 2002a, 2002b, 2003)。MAXENT预测的裸果木潜在分布区更加合理, 主要原因可能是该模型算法明确, 而且其规则化程序可以阻止在小样本的情况下发生过拟合(Phillips *et al.*, 2006; Phillips & Dudik, 2008), 所以, MAXENT更适合模拟分布数据有限、生态位较窄的物种。值得提出的是, 模型利用十分有限的发生数据, 结合留一法预测到的适生区, 通常只能代表与分布区具有相似的环境条件, 而不能作为物种实际的分布界限(Pearson *et al.*, 2007)。

#### 3.2 裸果木的潜在分布格局及原因探讨

裸果木的潜在适生区全部被限制在西北荒漠区, 围绕两个干旱中心, 主要分布在甘肃河西走廊中部及玉门以西、新疆哈密盆地、塔里木盆地西北缘及宁夏北部。探索其分布格局的原因, 应该是和西北荒漠区特殊的气候条件及复杂的地质及气候背景等分不开的(潘晓玲, 2001)。首先, 物种的分布和当前的环境条件有一定的直接关系, 环境条件通常限制了物种生态位(分布区)的建立和发展。西北荒漠区是北半球温带最干旱的地区, 年降水量不足200 mm, 并且降水时空分布不均, 一般从东南向西北方向逐渐减少, 在时间上主要集中在夏季, 再加上强烈的潜在蒸发, 所以形成了独特的荒漠化气候。而裸果木早已形成了长期适应这种夏雨型荒漠气候的形态解剖结构, 如叶片线形、气孔下陷、表皮附有蜡质层、几乎没有海绵组织、贮水组织发达等(汪之波等, 2004)。尽管如此, 水分对裸果木生长的影响还是很大的, 在调查过程中发现, 生长在高

山草原地区(如新疆乌恰县)和高山冰川融雪冲积扇地区(如甘肃阿克塞县)的裸果木高大、多分枝、植株密度高, 分布范围广; 而在沙丘腹地生长的裸果木(如甘肃金塔县)植株矮小, 分布稀疏。同时, 物种的分布不仅依赖于当前的气候条件, 而且和地质气候背景有很大的关系, 裸果木是亚洲中部荒漠区第三纪孑遗植物, 它和裸果木属的其他9种植物在地理上呈隔离分布。由于裸果木属绝大多数种类分布在地中海、南欧和非洲东部, 因此它属于古地中海植物区系和可能的古地中海起源(傅立国, 1992; 尹林克等, 2006)。裸果木的现代分布, 则很有可能是第三纪古地中海西撤后的残留和遗迹, 其在中国西北境内的间断分布, 可能是第三纪干旱化过程所导致, 而西北干旱化过程与青藏高原的隆升密切相关(郑度和姚檀栋, 2006)。裸果木现在集中分布的地区(东部的甘肃河西走廊玉门以西和内蒙古阿拉善地区, 西部的新疆乌恰地区)就是东西间断分布格局的证据。另外, 分布区内多样化的土壤类型, 主要是山前洪积扇戈壁(如甘肃阿克塞、新疆乌恰)、砾石戈壁(如新疆哈密、新疆库车、内蒙古阿拉善左旗)、构造剥蚀山地(如甘肃肃南)等, 都是裸果木适应和偏好的自然生境。

**致谢** 中国科学院新疆生态与地理研究所“百人计划”项目资助。感谢赵一之教授对本文提出的宝贵意见。

#### 参考文献

- Anderson RP, Gomez-Laverde M, Peterson AT (2002a). Geographical distributions of spiny pocket mice in South America: insights from predictive models. *Global Ecology and Biogeography*, 11, 131–141.
- Anderson RP, Lew D, Peterson AT (2003). Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modeling*, 162, 211–232.
- Anderson RP, Peterson AT, Gomez-Laverde M (2002b). Using niche-based GIS modeling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. *Oikos*, 98, 3–16.
- Bittrich V (1993). Caryophyllaceae. In: Kubitzki K ed. *The Families and Genera of Vascular Plants*. Springer-Verlag, Berlin.
- Caldecott JO, Jenkins MD, Johnson TH, Groombridge B (1996). Priorities for conserving global species richness and endemism. *Biodiversity and Conservation*, 5, 699–727.
- da Fonseca GAB (2000). It's time to work together and stop

- duplicating conservation efforts: following Africa's lead in setting priorities. *Nature*, 405, 393–394.
- Dang RL (党荣理), Jiang YC (姜彦成) (1996). The study of pollen morphology of Paronychioideae in Xinjiang. *Acta Botanica Boreali-Occidentalia Sinica* (西北植物学报), 16, 61–64. (in Chinese with English abstract)
- Dang RL (党荣理), Pan XL (潘晓玲), Gu XF (顾雪峰) (2002). Floristic analysis of spermatophyte genera in the arid deserts area in North-West China. *Guihaia* (广西植物), 22, 121–128. (in Chinese with English abstract)
- Fielding AH, Bell JF (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24, 38–49.
- Fu LG (傅立国) (1992). *China Plant Red Data Book* (中国植物红皮书). Science Press, Beijing. (in Chinese)
- Graham CH, Ferrier S, Huettman F, Moritz C, Peterson AT (2004). New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology & Evolution*, 19, 497–503.
- Guisan A, Harrell FE (2000). Ordinal response regression models in ecology. *Journal of Vegetation Science*, 11, 617–626.
- Guisan A, Thuiller W (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8, 993–1009.
- Hernandez PA, Franke I, Herzog SK, Pacheco V, Paniagua L, Quintana HL, Soto A, Swenson JJ, Tovar C, Valqui TH, Vargas J, Young BE (2008). Predicting species distributions in poorly-studied landscapes. *Biodiversity and Conservation*, 17, 1353–1366.
- Hernandez PA, Graham CH, Master LL, Albert DL (2006). The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, 29, 773–785.
- Holdridge LR, Grenke WC, Hatheway WH, Liang T, Tosi JA Jr (1971). *Forest Environments in Tropical Life Zones: a Pilot Study*. Pergamon Press, New York.
- IGBP-DIS (International Geosphere-Biosphere Programme, Data and Information System) (2000). *Global Soil Data Task*. <http://sage.wisc.edu/atlas/>. Cited 3 Jun. 2000.
- Jiang X (蒋霞), Ni J (倪健) (2005). Species-climate relationships of 10 desert plant species and their estimated potential distribution range in the arid lands of northwestern China. *Acta Phytocologica Sinica* (植物生态学报), 29, 98–107. (in Chinese with English abstract)
- Linder HP (1995). Setting conservation priorities: the importance of endemism and phylogeny in the southern African orchid genus *Herschelia*. *Conservation Biology*, 9, 585–595.
- Liu SL (刘生龙), Guo ZZ (郭志中), Wang LD (王理德), Gao ZH (高志海) (1995). Experiment on introduction and cultivation of *Gymnocarpus przewalskii*. *Journal of Desert Research* (中国沙漠), (15), 88–91. (in Chinese with English abstract)
- Liu YS (刘永生), Xu HS (徐怀寿), Sun JZ (孙继周) (2008). The investigation and research on the II protected plants *Gymnocarpus przewalskii* of China. *Ningxia Journal of Agriculture and Forestry Science and Technology* (宁夏农林科技), (1), 16, 26. (in Chinese)
- Margules C, Sarkar S (2000). Systematic conservation planning. *Nature*, 405, 243–253.
- Ottaviani D, Lasinio GJ, Boitani L (2004). Two statistical methods to validate habitat suitability models using presence-only data. *Ecological Modeling*, 179, 417–443.
- Oxelman B, Ahlgren B, Thulin M (2002). Circumscription and phylogenetic relationships of *Gymnocarpus* (Caryophyllaceae-Paronychioideae). *Edinburgh Journal of Botany*, 59, 221–237.
- Pan XL (潘晓玲) (2001). *Plant Flora Geography and Resources Utilization of the Northwest Arid and Desert Area* (西北干旱荒漠区植物区系地理与资源利用). Science Press, Beijing. (in Chinese)
- Pearson RG, Dawson TP, Liu C (2004). Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. *Ecography*, 27, 285–298.
- Pearson RG, Raxworthy CJ, Nakamura M, Peterson AT (2007). Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, 34, 102–117.
- Peterson AT, Papes M, Eaton M (2007). Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. *Ecography*, 30, 550–560.
- Peterson AT, Watson DM (1998). Problems with areal definitions of endemism: the effects of spatial scaling. *Diversity and Distributions*, 4, 189–194.
- Petrusson L, Thulin M (1996). Taxonomy and biogeography of *Gymnocarpus* (Caryophyllaceae). *Edinburgh Journal of Botany*, 53, 1–26.
- Phillips SJ, Anderson RP, Schapire RE (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modeling*, 190, 231–259.
- Phillips SJ, Dudik M (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, 31, 161–175.
- Stockwell D, Peters D (1999). The GARP modeling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science*, 13, 143–158.
- Stockwell DRB, Peterson AT (2002). Effects of sample size on accuracy of species distribution models. *Ecological Modeling*, 148, 1–13.
- Tsoar A, Allouche O, Steinitz O, Rotem D, Kadmon R (2007). A comparative evaluation of presence-only methods for modelling species distribution. *Diversity and Distributions*,



13, 397–405.

Wang ZB (汪之波), Gao QY (高清洋), Sun JZ (孙继周) (2004). Tissue culture and plant regeneration of rare plant *Gymnocarpus przewalskii*. *Acta Botanica Boreali-Occidentalia Sinica* (西北植物学报), 7, 1319–1321. (in Chinese with English abstract)

Wang ZB (汪之波), Gao QY (高清洋), Sun JZ (孙继周), Ma QL (马全林) (2009). Study on biological characteristics of rare endangered plant *Gymnocarpus przewalskii*. *Resource Development & Market* (资源开发与市场), (25), 481–509. (in Chinese with English abstract)

Ward DF (2007). Modelling the potential geographic distribution of invasive ant species in New Zealand. *Biological Invasions*, 9, 723–735.

Yin LK (尹林克), Tan LX (谭丽霞), Wang B (王兵) (2006). *Rare Endangered Endemic High Plants in Xinjiang of China* (新疆珍稀濒危特有高等植物). Xinjiang Science and Technique Press, Ürümqi. (in Chinese)

Zheng D (郑度), Yao TD (姚檀栋) (2006). Uplifting of Tibetan Plateau with its environmental effects. *Advance in Earth Science* (地球科学进展), 21, 451–458. (in Chinese with English abstract)

责任编辑: 倪 健 责任编辑: 王 葳