

# 基于最大熵模型预测青海省莱姆病的地理分布

张琳 侯学霞 刘慧鑫 刘炜 万康林 郝琴

102206 北京, 中国疾病预防控制中心传染病预防控制所 传染病预防控制国家重点实验室

通信作者: 郝琴, Email: haoqin@icdc.cn

DOI: 10.3760/cma.j.issn.0254-6450.2016.01.020

**【摘要】** 目的 应用最大熵模型(MaxEnt)预测青海省莱姆病的分布。方法 查阅1990年以来青海省人群莱姆病的血清检测数据,共收集到6个县(互助、泽库、同德、大通、祁连、循化)的血清学结果,将互助、泽库、同德县的血清学检测结果以及青海省环境和人为活动数据[包括海拔、人口足迹、归一化植被指数(NDVI)、温度等]导入MaxEnt软件,分析环境以及人为条件适宜的莱姆病潜在地理分布,然后以大通县、祁连县、循化县血清学数据作为验证数据,与模型预测结果进行比较。结果 MaxEnt预测结果显示,青海省存在3个莱姆病的热点区域,主要分布在青海省东部林区。在相关影响因子当中,NDVI对于模型的贡献最大;其次是人口足迹。用于验证模型的大通县、祁连县和循化县均分布于青海省东部。其中,循化县位于热点区域Ⅱ中,而大通县紧邻热点区域Ⅲ,位于热点区域Ⅲ的北部地区,祁连县不在预测的热点区域中,且莱姆病血清阳性率在3个调查县中为最低。模型运行良好[曲线下面积(AUC)=0.980]。结论 实际的人群莱姆病血清学数据与模型预测结果基本吻合。MaxEnt模型可用于莱姆病风险分布的预测。植被和人为活动可能与莱姆病传播有关。

**【关键词】** 莱姆病; 最大熵模型; 预测分布; 热点区域

**基金项目:** 国家科技重大专项(2012ZX10004219-007; 2013ZX10004-0101)

## Prediction of potential geographic distribution of Lyme disease in Qinghai province with Maximum Entropy model

Zhang Lin, Hou Xuexia, Liu Huixin, Liu Wei, Wan Kanglin, Hao Qin  
State Key Laboratory for Communicable Disease Prevention and Control, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China

Corresponding author: Hao Qin, Email: haoqin@icdc.cn

**【Abstract】 Objective** To predict the potential geographic distribution of Lyme disease in Qinghai by using Maximum Entropy model (MaxEnt). **Methods** The sero-diagnosis data of Lyme disease in 6 counties (Huzhu, Zeku, Tongde, Datong, Qilian and Xunhua) and the environmental and anthropogenic data including altitude, human footprint, normalized difference vegetation index (NDVI) and temperature in Qinghai province since 1990 were collected. By using the data of Huzhu Zeku and Tongde, the prediction of potential distribution of Lyme disease in Qinghai was conducted with MaxEnt. The prediction results were compared with the human sero-prevalence of Lyme disease in Datong, Qilian and Xunhua counties in Qinghai. **Results** Three hot spots of Lyme disease were predicted in Qinghai, which were all in the east forest areas. Furthermore, the NDVI showed the most important role in the model prediction, followed by human footprint. Datong, Qilian and Xunhua counties were all in eastern Qinghai. Xunhua was in hot spot area II, Datong was close to the north of hot spot area III, while Qilian with lowest sero-prevalence of Lyme disease was not in the hot spot areas. The data were well modeled in MaxEnt (Area Under Curve=0.980). **Conclusions** The actual distribution of Lyme disease in Qinghai was in consistent with the results of the model prediction. MaxEnt could be used in predicting the potential distribution patterns of Lyme disease. The distribution of vegetation and the range and intensity of human activity might be related with Lyme disease distribution.

**【Key words】** Lyme disease; Maximum Entropy models; Prediction; Hot spot area

**Fund programs:** National Science and Technology Major Project of China (2012ZX10004219-007; 2013ZX10004-0101)

莱姆病的感染过程主要涉及到宿主动物、媒介蜱及人<sup>[1]</sup>。不同地区的莱姆病螺旋体、宿主动物和媒介蜱的种类可能不同,因此感染率以及途径和方式也不相同<sup>[2-3]</sup>。青海省林区面积广袤,鼠蜱带菌率高<sup>[4-5]</sup>,高危人群多<sup>[6]</sup>。在青海省建立莱姆病分布模型既有代表性,又十分有必要。由于青海省莱姆病调查数据较少,因此直接利用传播路径“宿主动物-蜱-人”建立模型不可行。莱姆病螺旋体的传播、宿主动物分布以及媒介蜱的分布受到环境因素以及人为因素的影响<sup>[7]</sup>,莱姆病的感染具有季节性特点<sup>[8]</sup>,土地利用与人口足迹又反映了在感染过程中人为因素的影响<sup>[9]</sup>。因此应用空间生态学的方法,整合环境因子及人为活动影响评估莱姆病的分布是可行的。最大熵模型(MaxEnt)软件根据物种现实分布点和分布地区的环境变量运算得出预测模型,再利用模型模拟目标物种在目标地区的可能分布<sup>[10]</sup>。目前,MaxEnt已用于预测疾病发生的地理分布<sup>[11]</sup>。本研究利用MaxEnt软件对青海省莱姆病的地理分布进行预测。

### 资料与方法

1. 数据来源:收集并整理了我国自1990年以来青海省人群莱姆病血清检测的数据,包括6个县的9个调查点<sup>[12-14]</sup>,见表1,同时收集环境变量和人为活动影响因子[最高温的月平均温度、最低温的月平均温度、月平均温度、月平均降水量、归一化植被指数(NDVI)、土地利用、人口足迹、19个生物温度及海拔],见表2。利用互助县、泽库县以及同德县的秀麻、谷芒和河北乡的莱姆病血清学数据建立模型;利用大通县、祁连县和循化县的数据对模型进行验证。

**表1** 青海省人群莱姆病血清阳性调查数据

县名	地点(乡)	调查人数	阳性数	阳性率(%)
大通县		302	10	3.31 <sup>[12]</sup>
祁连县		142	3	2.11 <sup>[12]</sup>
循化县		71	6	8.45 <sup>[12]</sup>
互助县		598	126	21.07 <sup>[14]</sup>
泽库县		519	41	7.90 <sup>[14]</sup>
同德县	秀麻	54	4	7.41 <sup>[13]</sup>
	唐干	64	0	0 <sup>[13]</sup>
	谷芒	103	9	8.74 <sup>[13]</sup>
	河北	72	7	9.72 <sup>[13]</sup>

2. 研究方法:将互助县、泽库县以及同德县的秀麻、谷芒和河北3乡的共5个点的莱姆病血清学数据

**表2** 用于模型预测莱姆病分布的环境变量

类别	变量	变量描述
气候因子	Tmax	最高温的月平均温度(°C)(1950-2010), <a href="http://www.worldclim.org">http://www.worldclim.org</a>
	Tmean	月平均温度(°C)(1950-2010), <a href="http://www.worldclim.org">http://www.worldclim.org</a>
	Tmin	最低温的月平均温度(°C)(1950-2010), <a href="http://www.worldclim.org">http://www.worldclim.org</a>
	Prec	月平均降水(mm)(1950-2010), <a href="http://www.worldclim.org">http://www.worldclim.org</a>
	BioT1	年平均温度
	BioT2	月均温差
	BioT3	等温性(BioT2/BioT7)(×100)
	BioT4	温度季节性变化的标准差(standard deviation ×100)
	BioT5	最暖月份的最高温度
	BioT6	最冷月份的最低温度
	BioT7	年均温差(BioT5-BioT6)
	BioT8	最湿润季度的平均温度
	BioT9	最干燥季度的平均温度
	BioT10	最暖季度的平均温度
	BioT11	最冷季度的平均温度
	BioT12	年均降水量
	BioT13	最湿润月份的降水量
	BioT14	最干燥月份的降水量
	BioT15	降水量变异系数
BioT16	最湿润季度的降水量	
BioT17	最干燥季度的降水量	
BioT18	最暖季度的降水量	
BioT19	最冷季度的降水量	
地形因子	海拔	海拔高度(m), <a href="http://www.worldclim.org">http://www.worldclim.org</a>
	适宜生境	土地利用
人为因子	人口足迹	土地利用类型,包括:城市、农田、湿地等, <a href="http://bioval.jrc.ec.europa.eu/products/glc2000/glc2000.php">http://bioval.jrc.ec.europa.eu/products/glc2000/glc2000.php</a>
		NDVI

作为样本,按照MaxEnt 3.3.3k软件要求整理格式,保存为.csv格式,放入模型。应用ArcGIS软件的Conversion工具,将环境变量数据Raster格式转换为ASC II格式,导入MaxEnt,参数设置为软件默认,运行得出青海省环境以及人为条件适宜的莱姆病传播潜在地理分布。最后与大通县、祁连县、循化县人群血清学数据进行比较。适宜值越高代表自然与人文条件越适宜莱姆病的传播和保存。

### 结 果

1. 莱姆病分布预测:MaxEnt预测结果显示,一年12个月中莱姆病分布的适宜区域主要分布于青海省东部(图1)。1-4月,莱姆病适宜区域呈散发状态,5-7月,适宜区域分布较集中,8-12月,适宜

区域又逐渐分散开来。由于5月游离蚰的密度较高,具有代表性,因此使用5月的预测数据对青海省的热点区域进行划分。结果显示,热点区域(风险值>0.75)主要有3个:热点 I 包括同德县的中部、玛沁县东北部、兴海县东南部以及泽库县的西部;热点 II 位于同仁县和循化县中部以及尖扎县的东南部;热点 III 位于湟中县中部(图2)。受试者工作特征(ROC)曲线表明模型预测良好[曲线下面积(AUC)=0.980],见图3。

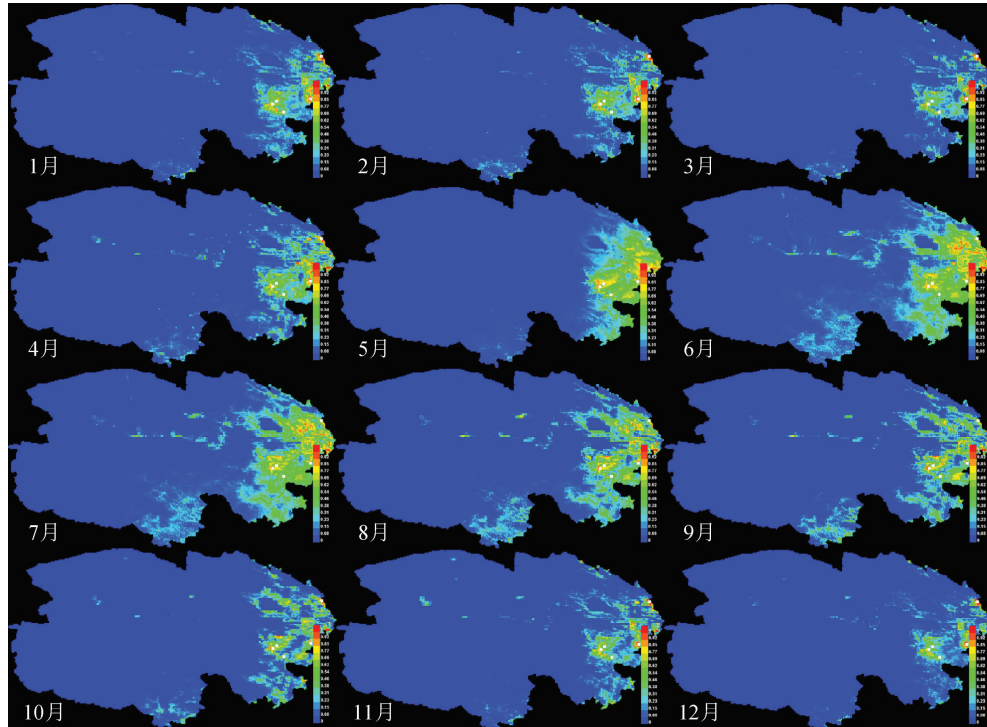


图1 MaxEnt风险预测1—12月青海省莱姆病地理分布

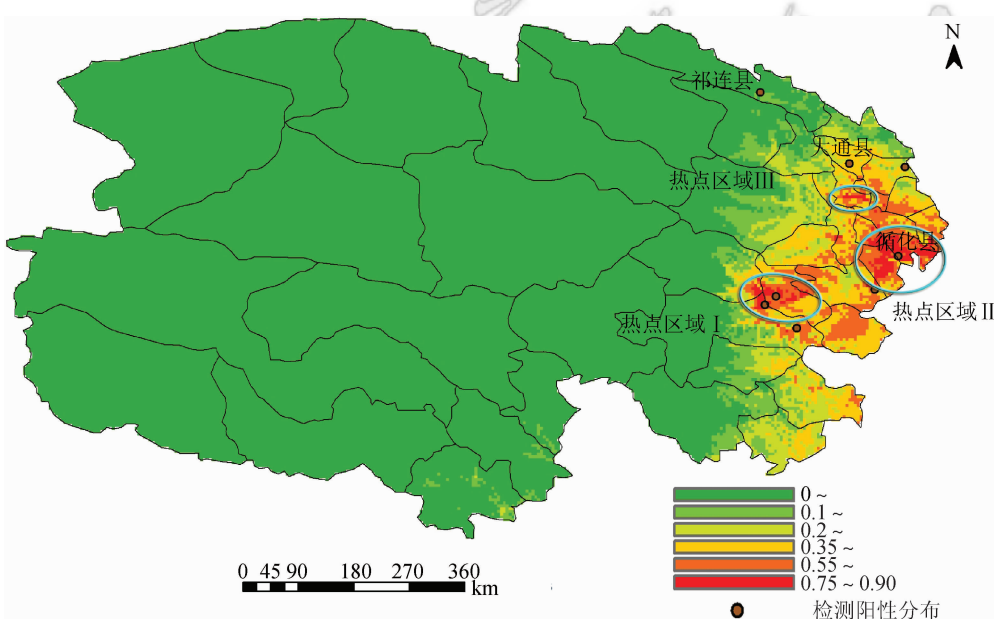


图2 青海省3个莱姆病高风险区域分布

Jackknife 过程的结果表明, 27个变量当中, 1—4月与8—12月, NDVI对于模型的贡献最大, 贡献率为32.6%~58.6%, 其次是人口足迹(footprint), 贡献率为28.2%~37.1%。5月平均降水(Prec)贡献最大(62.2%), 人口足迹(18.3%)次之。而6、7月份不同, 贡献率最大的变量为人口足迹(45.1%~49.1%), 其次为NDVI(32.1%~39.5%)。

2. 莱姆病预测及实际数据比较: 用于验证模型的大通县、祁连县和循化县均分布于青海省东部, 其中, 循化县位于热点区域 II 中, 大通县则位于热点区域 III 的北部地区。祁连县地处青海省东北部, 不在预测的热点区域中, 且该地区的人群莱姆病血清阳性率在3个用于验证的县中最低。调查数据与模型结果基本吻合。

### 讨论

莱姆病为蜱传疾病, 传播的过程涉及到宿主动物、媒介蜱以及人, 因此自然因素对于莱姆病传播的影响至关重要<sup>[15-16]</sup>。本研究也印证了这一点, NDVI在模型的建立当中起到了最重要的作用。由于人为因素对莱姆病的传播也具有重要影响<sup>[9]</sup>, 本研究整合了自然因子与人为因子, Jackknife 过程的结果也显示, 人为因子的影响仅次于NDVI。总体来讲, 在全部27个变量当中, NDVI对于模型的贡献最大, 提示植被种类与分布对莱姆病的传播影响



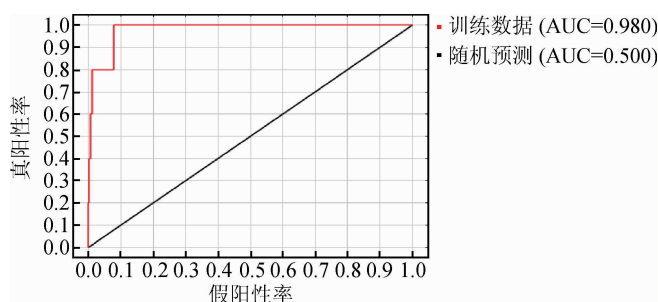


图3 莱姆病分布预测的ROC曲线

最大,这一结果对于莱姆病的预防和控制有着积极的作用,通过宏观植被分布数据,对热点区域作业的工作人员实施必要的防护措施,可以有效地保护人群免于蜱叮咬,进而防止莱姆病传播。同时,人类的行为及活动对莱姆病传播有一定影响。应用空间流行病学研究方法,根据各个因子的影响比重,对自然因子以及人为因子进行调控,从而降低疾病发生概率。然而,如何进行有效的调控,还有待进一步研究和探索。

本研究利用大通、祁连和循化3县的人群血清学调查数据对模型进行验证,结果显示,与模型预测结果基本吻合。本研究中整合了青海省流行病学调查数据,以及相关人为因素与自然因素,应用MaxEnt预测青海省莱姆病的分布,结果显示的3个热点区域可以作为流行病学调查的重点区域。莱姆病的病原学以及流行病学研究可以这3个区域作为重点。

综上所述,MaxEnt作为空间生态学的研究工具,可应用于莱姆病的空间生态学的研究,为莱姆病的流行病学及病原学提供宏观支持。然而,由于人群血清阳性率调查数据的匮乏,本研究结果存在不足,有待进一步的研究提供相关数据进行验证。

利益冲突 无

### 参 考 文 献

[1] Stanek G, Wormser GP, Gray J, et al. Lyme borreliosis [J]. *Lancet*, 2012, 379 (9814): 461-473. DOI: 10.1016/S0140-6736(11)60103-7.

[2] Ostfeld RS, Glass GE, Keasing F. Spatial epidemiology: an emerging (or re-emerging) discipline [J]. *Trends Ecol Evol*, 2005, 20(6): 328-336. DOI: 10.1016/j.tree.2005.03.009.

[3] LoGiudice K, Ostfeld RS, Schmidt KA, et al. The ecology of infectious disease: Effects of host diversity and community composition on Lyme disease risk [J]. *Proc Natl Acad Sci USA*, 2003, 100(2): 567-571. DOI: 10.1073/pnas.0233733100.

[4] 侯学霞, 耿震, 郝琴, 等. 中国6省莱姆病螺旋体主要宿主动物的初步调查[J]. *中国人兽共患病学报*, 2010, 26(11): 1034-1036. DOI: 10.3969/j.issn.1002-2694.2010.11.012.

Hou XX, Geng Z, Hao Q, et al. Rats, the primary reservoir hosts of *Borrelia burgdorferi*, in six representative provinces, China

[J]. *Chin J Zoon*, 2010, 26 (11): 1034-1036. DOI: 10.3969/j.issn.1002-2694.2010.11.012.

- [5] 张琳, 侯学霞, 耿震, 等. 巢式PCR与LAMP方法在蜱的莱姆病螺旋体检测中的应用[J]. *中国人兽共患病学报*, 2014, 30(12): 1192-1195. DOI: 10.3969/cjz.j.issn.1002-2694.2014.12.04.
- Zhang L, Hou XX, Geng Z, et al. Nested-PCR and LAMP in detection of *B. burgdorferi* in ticks [J]. *Clin J Zoon*, 2014, 30 (12): 1192-1195. DOI: 10.3969/cjz.j.issn.1002-2694.2014.12.04.
- [6] Wu XB, Na RH, Wei SS, et al. Distribution of tick-borne diseases in China [J]. *Parasit Vectors*, 2013, 6: 119. DOI: 10.1186/1756-3305-6-119.
- [7] Lane RS, Piesman J, Burgdorfer W. Lyme borreliosis: relation of its causative agent to its vectors and hosts in North America and Europe [J]. *Annu Rev Entomol*, 1991, 36(1): 587-609. DOI: 10.1146/annurev.en.36.010191.003103.
- [8] Ciesielski CA, Markowitz LE, Horsley R, et al. Lyme disease surveillance in the United States, 1983-1986 [J]. *Rev Infect Dis*, 1989, 11 Suppl 6: S1435-1441. DOI: 10.1093/clinids/11.Supplement\_6.S1.
- [9] de Thoisy B, Richard-Hansen C, Goguillon B, et al. Rapid evaluation of threats to biodiversity: human footprint score and large vertebrate species responses in French Guiana [J]. *Biodivers Conserv*, 2010, 19 (6): 1567-1584. DOI: 10.1007/s10531-010-9787-z.
- [10] Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions [J]. *Ecol model*, 2006, 190 (3/4): 231-259. DOI: 10.1016/j.ecolmodel.2005.03.026.
- [11] 曾辉, 黄冠胜, 林伟, 等. 利用MaxEnt预测橡胶南美叶疫病菌在全球的潜在地理分布[J]. *植物保护*, 2008, 34(3): 88-92. DOI: 10.3969/j.issn.0529-1542.2008.03.023.
- Zeng H, Huang GS, Lin W, et al. Prediction of potential geographic distribution of *Microcyclus ulei* in the world using MaxEnt [J]. *Plant Prot*, 2008, 34 (3): 88-92. DOI: 10.3969/j.issn.0529-1542.2008.03.023.
- [12] 刘光润, 马秀兰, 曲月照, 等. 青海省三县林区莱姆病血清流行病学调查[J]. *中国公共卫生学报*, 1995, 14(4): 202.
- Liu GR, Ma XL, Qu YZ, et al. The sero-investigation of prevalence of Lyme disease in three counties in Qinghai [J]. *Chin J Public Health*, 1995, 14(4): 202.
- [13] 姜双应, 易虎, 巩天祥, 等. 青海省同德县莱姆病血清流行病学调查分析[J]. *青海医药杂志*, 2001, 31(12): 51-52. DOI: 10.3969/j.issn.1007-3795.2001.12.031.
- Jiang SY, Yi H, Gong TX, et al. The prevalence investigation of Lyme disease in Tongde, Qinghai province [J]. *Qinghai Med J*, 2001, 31 (12): 51-52. DOI: 10.3969/j.issn.1007-3795.2001.12.031.
- [14] Hao Q, Geng Z, Hou XX, et al. Seroepidemiological investigation of lyme disease and human granulocytic anaplasmosis among people living in forest areas of eight provinces in China [J]. *Biomed Environ Sci*, 2013, 26 (3): 185-189. DOI: 10.3967/0895-3988.2013.03.005.
- [15] Steere AC, Coburn J, Glickstein L. The emergence of Lyme disease [J]. *J Clin Invest*, 2004, 113 (8): 1093-1101. DOI: 10.1172/JCI200421681.
- [16] Eisen L, Eisen RJ, Mun J, et al. Transmission cycles of *Borrelia burgdorferi* and *B. bissettii* in relation to habitat type in northwestern California [J]. *J Vector Ecol*, 2009, 34(1): 81-91. DOI: 10.3376/038.034.0110.

(收稿日期: 2015-06-03)

(本文编辑: 万玉立)