ECOLOGICAL CHARACTERISTICS AND SUITABILITY EVALUATION OF *FRITILLARIA CIRRHOSA* D. DON BASED ON MAXENT MODEL

Ming-li Wu[1,2], Qin Zhang[1], Jing-yuan Song[1], Xi-wen Li[3], Cai-xiang Xie[1*], Zhi-gang Hu[2]**

[1]Institute of Medicinal Plant Development, Peking Union Medical College, Chinese Academy of Medical Sciences, Beijing, China, 100193. [2]Pharmacy Faculty, Hubei University of Chinese Medicine, Wuhan, Hubei, China, 430065. [3]Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing, China, 100700.

*Corresponding Author E-mail: caixiangxie@163.com ;
**Corresponding Author E-mail: zghu0608@163.com

Abstract

Background: As a rare and endangered medicinal plant growing in plateaus, *Fritillaria cirrhosa* D. Don is a scant resource in terms of quantity and planting regions. However, there is limited knowledge on predicting the potential ecological suitability of regions for the species with the climate factors. This paper evaluates the ecological suitability of *F. cirrhosa* D. Don on a global scale using Maxent.

Materials and Methods: The ecologically suitable regions for *F. cirrhosa* D. Don around the world were evaluated using the maximum entropy model (Maxent), based on 127 known occurrence records and specific environmental predictor variables.

Results: The Maxent model was highly accurate, with a statistically significant Area Under the Receiving Operator Curve (AUC) value of 0.993, and the most suitable areas and the suitable areas for *F. cirrhosa* D. Don were approximately 450,000 and 700,000 sq. km., respectively, including China, Pakistan, Nepal, and Bhutan. A quantitative study of the climatic characteristics of *F. cirrhosa* D. Don indicated that the period from May to October was critical for plant growth and development. Thus, the stable precipitation-temperature ratios (0.59 to 2.42) during this period could serve as a feature indicator for the geographical distribution of the plant.

Conclusion: This work should be beneficial for the introduction and resource protection of *F. cirrhosa* D. Don, meanwhile, the analytical method could be expanded to predict the potential distribution of other medicinal plants.

Keywords: *Fritillaria cirrhosa* D. Don, maximum entropy model, ecological suitability, climatic characteristics, potential distribution

Introduction

The deterioration of environmental habitat and over-exploitation are two of the most important factors responsible for species loss throughout the world (Barnosky et al., 2011), placing one-fifth of the plant species, with the inclusion of *Fritillaria cirrhosa* D. Don, on the verge of extinction (Brummit et al., 2010). To relieve the problem of resource scarcity, the natural cultivation of *F. cirrhosa* D. Don is one of the effective ways used, whereas the primary task of natural cultivation in ecologically suitable areas has yet to be determined.

Fritillaria cirrhosa D. Don (Liliaceae, *Fritillaria* L.) is a perennial herb, and one of the parent species of *F. cirrhosae bulbus* recorded in Chinese Pharmacopoeia; the dried bulbs of the plant are a valuable Chinese medicine that is effective in treating cough and is characterized by large market demand. Owing to the disordered excavation and distinct growth behavior of *F. cirrhosa* D. Don, this species has declined in recent years. Thus, the suitable areas for growing *F. cirrhosa* D. Don should be determined immediately, as finding such conducive environments will help solve the problem of resource scarcity. Climate factors play a critical role in species distribution in various regions (Engler et al., 2011; Araújo et al., 2011). Previous studies showed that a certain significant relationship exists between fritillaria alkaloid levels and climatic factors (Li, 2009). Moreover, compared with the biophysical factors, the climatic factors are more important in determining the geographic distribution of vegetation. Internationally, predicting the potential ecological suitability of regions for certain species with the climate factors is indispensable.

Maximum Entropy model (Maxent) is a species distribution model derived from statistical mechanics (Schapire, 2004; Phillips and Dudík, 2008; Phillips et al., 2004). The principle of using Maxent is to estimate a target probability distribution by finding its maximum entropy (that is, the most spread out or uniform), which is subjected to a set of constraints that represent the incomplete information on the target distribution. At present, the ecological niche models commonly used are Climex (Sutherst et al., 1995), Domain (Carpenter et al., 1993), Bioclim (Busby, 1991), GARP (Stockwell, 1999), Maxent (Phillips et al., 2004) , and so on. However, many previous studies have shown that Maxent is superior to other ecological models, especially when the species distribution data are incomplete (Wang et al. 2007). Maxent only requires the species presence (or occurrence) data and environmental information, and the data format can be continuous or categorical. Compared with the absence data of species, the presence data of species are more available, which makes Maxent more popular in predicting the species distribution.

Current research on the evaluation of ecological suitability were mainly focused on certain regions or countries. Wang et al. (2014) studied the potential distribution of *F. cirrhosa* D. Don in China and found that isothermality is the key climate factor affecting the distribution of plant species. Chen et al. (2003) investigated the plant community of *F. cirrhosa* D. Don. These studies helped us to understand the influence of climatic factors on the geographic distribution of *F. cirrhosa* D. Don. This paper evaluates the ecological suitability of *F. cirrhosa* D. Don on a global scale using Maxent. Meanwhile, the climate characteristics of *F. cirrhosa* D. Don were quantitatively investigated for the first time, indicating the critical growth period and the corresponding value of the precipitation temperature ratio for *F. cirrhosa* D. Don.

Materials and Methods
Sample distribution data

The distribution points of *F. cirrhosa* D. Don were obtained from the Global Biodiversity Information Facility (GBIF, http://www.gbif.org/) and the Chinese Virtual Herbarium (CVH, http://www.cvh.org.cn/), which integrates the herbarium data of national natural museums from 14 institutes in China. Finally, a total of 565 samples were collected, but some samples were redundant, lacked information on location/coordinates, and incorrect. A total of 127 valid samples were finally retained through strict verification, of which the spatial distribution pattern is shown in Figure 1. The files for latitude/longitude information of the samples were saved in .csv format.

Figure 1: Sample points of *Fritillaria cirrhosa* D. Don.

Environmental variables

The global environmental data were obtained from the WorldClim database (http://www.worldclim.org/), with a spatial resolution of about 1 sq. km. and the coordinate system WGS84, and also included 19 biology climate data, 12 monthly average temperatures, and 12 monthly average precipitation values (Table 1). Interestingly, the temperature value of this database was 10 times that of the actual. The geographic information data were obtained from DIVA-GIS website (http:// www. divagis.org/) and National Science and Technology Platform (http://www.geodata.cn/ Portal/ index.jsp).

Table 1: Environmental factors.

Data Types	Name	Unit
	BIO1: Mean annual temperature	℃×10
	BIO2: Mean of monthly (max temperature−min temperature)	℃×10
	BIO3: Isothermality (Bio2 / Bio7×100)	1
	BIO4: SD of temperature seasonality	1
Climate data	BIO5: Max temperature of warmest month	℃×10
	BIO6: Min temperature of coldest month	℃×10
	BIO7: Temperature annual range (Bio5−Bio6)	℃×10
	BIO8: Mean temperature of wettest quarter	℃×10
	BIO9: Mean temperature of driest quarter	℃×10
	BIO10: Mean temperature of warmest quarter	℃×10

160

BIO11: Mean temperature of coldest quarter		°C×10
BIO12: Annual precipitation		mm
BIO13: Precipitation of wettest month		mm
BIO14: Precipitation of driest month		mm
BIO15: Coefficient of variation of precipitation seasonality		1
BIO16: Precipitation of wettest quarter		mm
BIO17: Precipitation of driest quarter		mm
BIO18: Precipitation of warmest quarter		mm
BIO19: Precipitation of coldest quarter		mm
Prec1- Prec12		mm
Tmean1- Tmean12		°C×10
Terrain data	Altitude	m

Maxent running parameter settings

The presence data for *F. cirrhosa* D. Don was saved in .csv format. Maxent (3.3.3k) (http://www.cs.princeton.edu/~schapire/maxent/) was used to predict the potential distribution of *F. cirrhosa* D. Don. The running parameters were as follows. The random test percentage was 25%, and the number of iterations were 1,000,000. Meanwhile, the boxes for creating response curves were checked, and Do jackknife was used in measuring variable importance. Other options were kept as default. The output value can interpret the probability presence (ranging from 0–1). The jackknife test (by systematically leaving out each variable) was used to evaluate the climatic factors that were most important in determining the potential distribution of the species. The response curves of *F. cirrhosa* D. Don were created using the filtered variables.

ROC curve

Receiver operating characteristic (ROC) curve was constructed by using all the possible thresholds in classifying the scores into confusion matrices, and then obtaining the sensitivity and specificity for each matrix. Plotting sensitivity against the corresponding proportion of false positives is commonly regarded as a good method for assessing the accuracy of ordinal score models. AUC, which is currently recognized as the best evaluation for diagnostic tests (Giedrius, 2004) and is not affected by the decision threshold, was used to evaluate the model, and the model with a higher AUC value was considered as the better performer. The meanings of the obtained AUC values were as follows. Values below 0.7 indicated that the predicting results were unreliable; values ranging from 0.7 to 0.8 indicated fair performance; values from 0.8 to 0.9 indicated good performance; and values from 0.9 to 1.0 indicated very good performance. The ROC curve and the AUC values were applied to evaluate the prediction precision in this study.

Ecological factor optimization

First, to minimize the impact of multicollinearity and overfitting of the model, the test for multicollinearity was conducted by using cross-correlations among environmental variables. Finally, only variables with a Spearman

161

Table 2: Correlation matrix of environmental variables.

| Variables	Bio7	Bio3	Bio2	Bio15	Alt	Prec11	Prec4	Prec5	Prec10	Prec1	Prec7	Tmax1
Bio7	1.000	-.664**	.167	-.077	.356**	-.365**	-.097	-.328**	-.733**	-.214*	-.722**	-.725**
Bio3		1.000	.484**	.495**	-.137	-.110	-.392**	-.059	.406**	-.209*	.519**	.664**
Bio2			1.000	.717**	.338**	-.628**	-.699**	-.447**	-.253**	-.749**	-.003	.064
Bio15				1.000	.247**	-.593**	-.669**	-.372**	-.109	-.685**	.285**	.151
Alt					1.000	-.686**	-.484**	-.715**	-.730**	-.237**	-.527**	-.745**
Prec11						1.000	.617**	.723**	.718**	.500**	.412**	.425**
Prec4							1.000	.778**	.409**	.601**	.204*	.104
Prec5								1.000	.753**	.331**	.543**	.506**
Prec10									1.000	.173*	.794**	.799**
Prec1										1.000	-.034	-.030
Prec7											1.000	.780**
Tmax1												1.000

Global suitability regionalization

A suitable habitat map for *F. cirrhosa* D. Don was successively completed after the normalization processing for Maxent result values. To retain the maximum predictive information, we imported the predicted suitability into 4 habitat class levels; the area with below 20% ecological similarity was classified as unsuitable area, 20% to 40% ecological similarity indicated a sub-suitable area, 40% to 60% ecological similarity signified a suitable area, and above 60% ecological similarity was considered as the most suitable area.

The Minimum Standards of Reporting Checklist contains details of the experimental design, and statistics, and resources used in this study.

Results

The AUC value of the training data was 0.993, and one of the test data presented a 0.9898 AUC value (Figure 2), indicating that the predicted results were highly accurate. Considering that the AUC values tend to be lower for species with broad distribution scope was very important in conducting the study (Mcpherson et al., 2007; Evangelista et al., 2008).
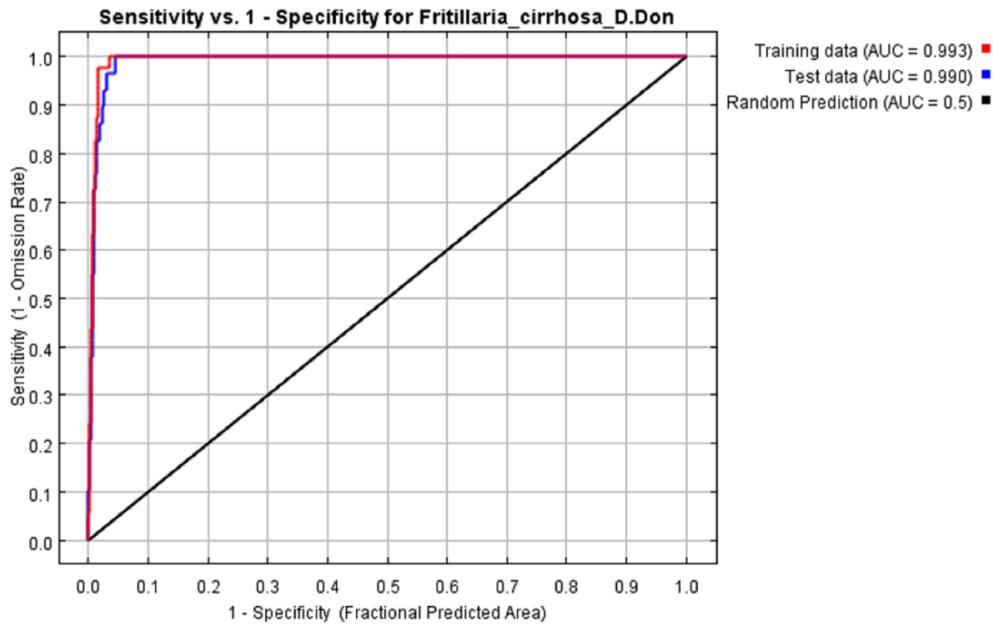
Figure 2: ROC curve of *F. cirrhosa* D. Don.

By selecting the option "Do jackknife to measure important variable" (Jackknife), we obtained the percentage contribution values of each ecological factor (Table 3), which were only heuristically defined and depended on the particular path that the Maxent code used to obtain the optimal solution. Moreover, the model referred to the effect of environmental variables in determining the optimal environmental parameters (Figure 3).

Table 3: Percentage contribution values of each ecological factor.

| Variable	Percentage contribution
alt	43.3
prec_7	16.5
bio_7	13.8
prec_11	12.4
tmax_1	4.3
bio_3	2.7
prec_4	2.2
prec_5	1.5
bio_2	1.4
bio_15	1.1
prec_10	0.6
prec_1	0.2
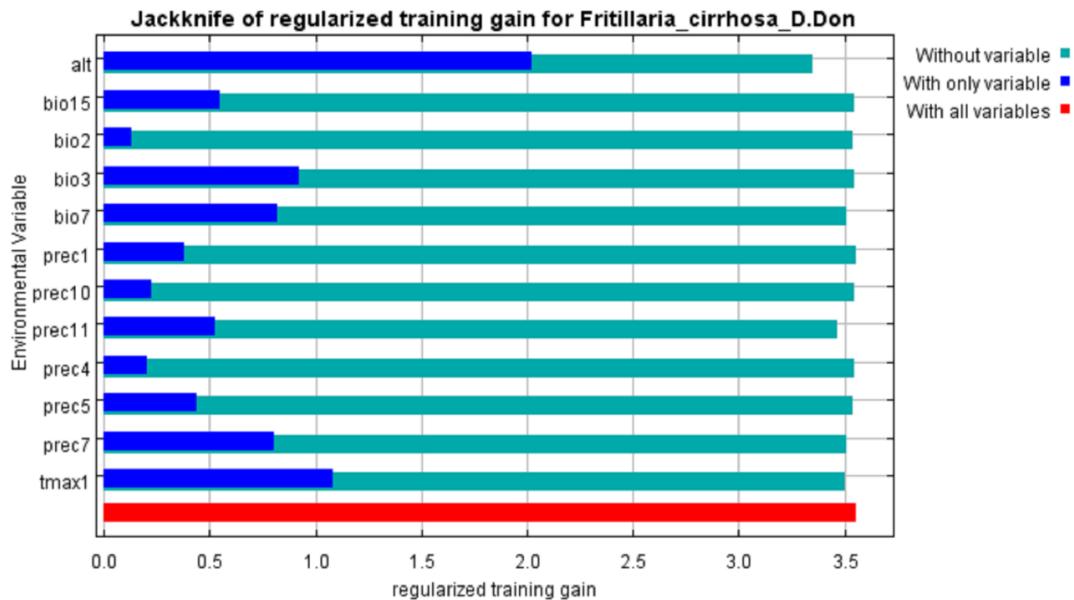
Figure 3: Jackknife of regularized training gain for *F. cirrhosa* D. Don.

Fritillaria cirrhosa D. Don thrives in cold and wet areas with high altitude. Alt, tmax1, and prec7 were finally determined as the most important environmental variables based on the results of the jackknife test and the biological characteristics of *F. cirrhosa* D. Don.

The response curves of each ecological factor were generated based on the modeling results. It is generally known that the probability of presence below 0.3 means that the corresponding value range of the ecological factor is unsuitable for the growth of the plant, the probability of presence from 0.3 to 0.5 for sub-suitable, and the probability of presence more than 0.5 for suitable. So the value ranges of main ecological factors for these grade were calculated as follows：

As a result, the value range of alt in the most suitable area was 843–5356 m, whereas the suitable area was 474–5993 m. The value range of tmax1 in the most suitable area was −16.8 to 25.8 °C, whereas tmax1 in the suitable areas ranged from −19.3 to 29.5 °C. The value range of prec7 in the most suitable area was 20–2012 mm, whereas prec7 in the suitable area was 7–1884 mm.

The results of the evaluation of ecological suitability for *F. cirrhosa* D. Don are shown in Figure 4.
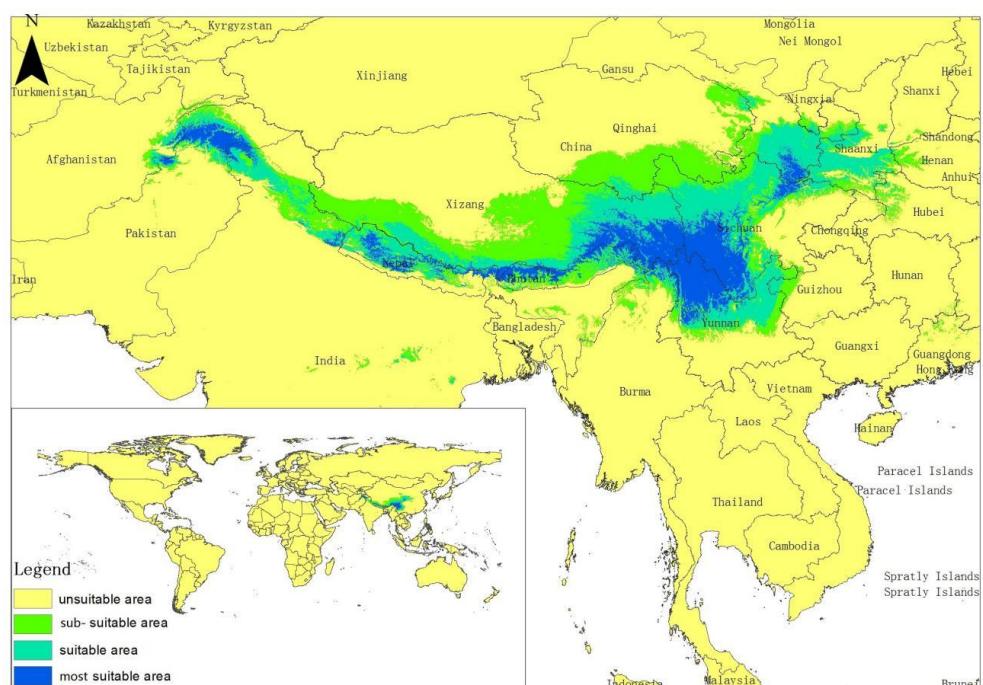
Figure 4: Ecological suitability areas for *F. cirrhosa* D. Don.

Apparently, the most suitable areas were mainly located in southwestern China, north of Pakistan, most areas of Nepal, and Bhutan. In addition, the ecologically suitable regions for *F. cirrhosa* D. Don were mainly located at the edge of Qinghai–Tibet Plateau in China. The acreage of the most suitable area for *F. cirrhosa* was approximately 450,000 sq. km., whereas the suitable area was approximately 700,000 sq. km. Although *F. cirrhosa* D. Don originates from China, the countries of Pakistan, Bhutan, and Nepal are also potential distribution regions for this species.

The ecologically suitable regions for *F. cirrhosa* D. Don are mainly located at the southern edge of the Tibetan Plateau, such as China, Pakistan, Bhutan, and Nepal. Owing to the high altitude of the aforementioned regions, solar radiation is strong, and the temperature is relatively low. At the same time, the Tibetan Plateau is a tall barrier that blocks the warm moist air coming from the southern and northward areas, which is an important factor in South Asia during the rainy season. The climate in these areas fits well with the growth characteristics of *F. cirrhosa* D. Don, which thrives in cold and wet areas with high altitudes.

The quality of Chinese herbal medicine is closely related to various climate factors, and the effects of climate on plants are holistic, dynamic, and complex. This study attempted to discover any relationship among these climate factors by constructing ratio curves, thus to define and establish the climatic characteristics of *F. cirrhosa* D. Don. Results showed that only the precipitation–temperature ratios in the global and ecologically similar suitable regions, obtained from May to October, were stable with variation coefficients of 2.30, 0.60, 0.59, 0.61, 0.76, and 2.42, respectively. May to October was also the main phenological period for *F. cirrhosa* D. Don, which involves the flowering and fruit-bearing stages of the plant. Thus, this study suggests that the climate characteristic values of precipitation–temperature ratios during May to October could serve as the ecological characteristics for *F. cirrhosa* D. Don.

Discussion

The quality of Chinese herbal medicine is closely related to regional environments. Using a maximum entropy model, we first investigated the ecologically suitable regions for *F. cirrhosa* D. Don on a global scale, and pioneered in identifying the ecological characteristics of *F. cirrhosa* D. Don based on the climatic factors during the entire growth period of the plant.

MaxEnt as an excellent species distribution modeling tool, has been extensively applied to predict distributions of more and more species since 2004. However, the application of MaxEnt on medicinal plants is still rare. MaxEnt can model the species distribution merely from presence-only species records, which can be easily received from field survey records, specimen records and literatures. In this study, we collected 127 records for *F. cirrhosa* D. Don. The sample number seems small, but fortunately, the MaxEnt model is independent of simple size and can abstract useful biogeographical information from very small simples, even less than 5 records (Hernandez et al., 2006; Pearson et al., 2007). The ecological factors used in this research involved 19 biology climate datas, 12 monthly average temperatures, 12 monthly average precipitation values and altitude date. Taking the cross-correlations among environmental variables into consideration, the multicollinearity test screened only 12 ecological factors for modeling. The rsults of previous studies (Wang et al., 2014; Zhao et al., 2017) may be unreliable, as they enlarged the impact of multicollinearity and overfitting of the model.

MaxEnt predicts species' probabilities of occurrence in a logistic format, which rangs from 0-1. However, the values are classified into different classes of potential habits in different studies for different species. After comparing different published literatures, we imported the predicted suitability into 4 habitat class levels, and only focus on the suitable habitats with probabilities above 0.4 in this study. The ecologically suitable regions for *F. cirrhosa* D. Don were mainly located in the southwestern areas of China, north of Pakistan, most areas of Nepal, and Bhutan. The acreage of the most suitable area for *F. cirrhosa* was approximately 450,000 sq. km., whereas the suitable area was approximately 700,000 sq. km. According to the jackknife test result, alt, tmax1, and prec7 were finally determined as the most important environmental variables, whose suitability ranges were received based on the response curves.

In addition, this study first identified the precipitation–temperature ratio in different months with some regularity.

The precipitation–temperature ratios from May to October were rather stable (i.e., during the main phenophases of *F. cirrhosa* D. Don). The biological characteristics of *F. cirrhosa* D. Don and the preceding findings indicate that the precipitation–temperature ratios obtained from May to October could serve as the ecological characteristics of *F. cirrhosa* D. Don. The new idea of climatic characteristics was first proposed in this paper; however, this new concept should be further validated and studied in future research.

In summary, this study only considered the impact of climatic factors and ignored the effects of soil factors on plant growth. Thus, expanding and improving the soil database is highly necessary. Moreover, the predicted potential distribution areas by Maxent are always over-estimated compared with the realized niche of the species. As the Maxent model only considers niche-based presence data, it predicts the fundamental niche of the species rather than the realized niche (Pearson, 2007; Kumar et al., 2009). In fact, the species might have failed to disperse because of geographic barriers, human disturbance, or associated competitive species. Thus, the areas predicted were suitable; however, they are currently not occupied by the species, and should be considered as candidate areas for the conservation, prioritization, and propagation of this species. This study aimed to investigate the dynamic equilibrium relationship among the various factors of the plant growth cycle. Considering the limited international studies on the correlation among climate factors, we believe that this study is inadequate and requires further in-depth research.

Conclusion

This paper is the first to present the suitable regions for *F. cirrhosa* D. Don on a global scale using the maximum entropy model and pioneered the study of the required climate for *F. cirrhosa* D. Don growth.

References

1. Araújo, M.B., Alagador, D., Cabeza, M., Nogués-Bravo, D., Thuiller, W. (2011). Climate change threatens European conservation areas. Ecology Letters., 14(5): 484-492.
2. Barnosky, A.D., Matzke, N.O., Tomiya, S., Wogan, G.O., Swartz, B., Quental, T.B., Marshall. C., McGuire, J.L., Lindsey, E.L., Maguire, K.C., Mersey, B., Ferrer, E.A. (2011). Has the earth's sixth mass extinction already arrived? Nature., (471): 51-57.
3. Brummit, N., Bachman, S. (2010). Plants under pressure: a global assessment. The first report of the IUCN sampled red list index for plants. Kew, UK: Royal Botanic Gardens.
4. Busby, J.R. (1991). BIOCLIM: a bioclimate analysis and prediction system. In: C.R.Margules, M.P.Austin, editors. Nature Conservation: Cost Effective Biological Surveys and Data Analysis. CSIRO, Canberra, Australia; pp. 64–68.
5. Carpenter, G., Gillison, A.N., Winter, J. (1993). DOMAIN: a flexible modelling procedure for mapping potential distributions of plants and animals. Biodiversity and Conservation., 2(6): 667-680.
6. Chen, S.L., Jia, M., Wang, Y., Xue, G., Xiao, P.G. (2003). Study on the Plant Community of *Fritillaria cirrhosa*. China Journal of Chinese Materia Medica., 28(5): 398-402.
7. Engler, R., Randin, C.F., Thuiller, W., Dullinger, S., Zimmermann, N.E., Araújo, M.B., Pearman, P.B., Le lay, G., Piedallu, C., Albert, C.H., Choler, P., Coldea, G., deLamo, X., Dirnbock, T., Gegout, J., Gomez-Garcia, D., Grytnes, J., Heegaard, E., Hoistad, F., Nogués-Bravo, D., Normand, S., Puscas, M., Sebastia, M., Stanisci, A., Theurillat, J., Trivedi, M.R., Vittoz, P., Guisan, A. (2011). 21st century climate change threatens mountain flora unequally across Europe. Global Change Biology., 17(7): 2330-2341.

8. Evangelista, P.H., Kumar, S., Stohlgren, T.J., Jarnevich, C.S., Crall, A.W., Norman, J.B., Barnett, D.T. (2008). Modelling invasion for a habitat generalist and a specialist plant species. Diversity and Distributions., 14(5): 808-817.

9. Giedrius, V. (2004). Receiver operating characteristic curves and comparison of cardiac surgery risk stratification systems. Interactive Cardiovascular and Thoracic Surgery., 3(2): 319-322.

10. Hernandez, P.A., Graham, C.H., Master, L.L., Albert, D.L. (2006). The effect of sample size and species characteristics on performance of differnet species distribution modeling methods. Ecography., 29(5): 773-785.

11. Kumar, S., Stohlgren, T.J. (2009). Maxent modeling for predicting suitable habitat for threatened and endangered tree Canacomyrica monticola in New Caledonia. Ecology and Natural Environment., 1(4): 94-98.

12. Mcpherson, J.M., Jetz, W. (2007). Effects of species' ecology on the accuracy of distribution models. Ecography., 30(1): 135-151.

13. Pearson, R.G. (2007). Species' Distribution Modeling for Conservation Educators and Practitioners. Synthesis. http://ncep.amnh.org.

14. Pearson, R.G., Raxworthy, C.J., Nakamura, M., Peterson, A.T. (2007). Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. Journal of Biogeography., 34(1): 102-117.

15. Phillips, S.J., Dudík, M. (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography., 31(2): 161-175.

16. Phillips, S.J., Dudík, M., Schapire, R.E. (2004). A maximum entropy approach to species distribution modeling. Proceedings of the Twenty-First International Conference on Machine Learning., 655-662.

17. Schapire. (2004). A Maximum Entropy Approach to Species Distribution Modeling. Proceedings of the Twenty-First International Conference on Machine Learning., 655-662.

18. Sutherst, R.W., Maywald, G.F., Skarratt, D.B. (1995). Predicting insect distributions in a changed climate. In: Harrington R, Stork NE, editors. In Insects in Changing Environment. London: Academic; p. 59-91.

19. Stockwell, D., Peters, D. (1999). The GARP modelling system: problems nd solutions to automated spatial prediction. International Journal of Geographical Information Systems., 13(2): 143-158.

20. Wang, Y.S., Xie, B.Y., Wan, F.H., Xiao, Q.M., Dai, L.Y. (2007). Application of ROC curve analysis in evaluating the performance of alien species' potential distribution models. Biodiversity Science., 15(14): 365-372.

21. Wang, J.J., Cao, B., Bai, C.K., Zhang, L.L., Che, Le. (2014). Potential Distribution Prediction and Suitability Evaluation of *Fritillaria cirrhosa* D. Don based on Maxent Modeling and GIS. Bulletin of Botanical Research., 34(5): 642-649.

22. Xiwen Li. (2009). Study on conservation biology of *Fritillaria cirrhosa*. Beijing: Peking Union Medical College.

23. Zhao, Q., Li, R., Gao, Y., Yao, Q., Guo, X., Wang, W.G. (2017). Modeling impacts of climate change on the geographic distribution of medicinal plant *Fritillaria cirrhosa* D. Don. Plant Biosystems., 1-7.