

DOI: 10.11931/guihaia.gxzw202102026

朱光福, 伊廷双, 刘杰, 2022. 数字植物标本在科学研究中的应用: 进展及挑战 [J]. 广西植物, 42(增刊 1): 127-141.
ZHU GF, YI TS, LIU J, 2022. Application of digitized plant specimens in scientific research: progress and challenges [J].
Guihaia, 42(Suppl. 1): 127-141.



数字植物标本在科学研究中的应用: 进展及挑战

朱光福^{1,3}, 伊廷双¹, 刘杰^{1,2*}

(1. 中国科学院昆明植物研究所 西南野生生物种质资源库, 昆明 650201; 2. 中国科学院昆明植物研究所
中国科学院东亚植物多样性和生物地理学重点实验室, 昆明 650201; 3. 中国科学院大学, 北京 100049)

摘要: 全球自然历史博物馆中保存着约 3.9 亿份植物标本, 这些标本为科学研究提供了重要支撑。近 20 年来, 随着标本数字化的发展, 数字植物标本在科学研究中的应用日益广泛。数字植物标本不仅为标本传统的用途提供了便利, 而且衍生出了新的研究主题和方向。鉴于目前中文文献中缺乏数字植物标本在科学研究中的应用概述及所面临挑战的分析, 该文概述了全球植物标本数字化的现状, 并对数字标本在生物地理学、入侵生物学、气候变化和保护生物学等领域中的应用情况进行了综述。主要集中在以下 5 个方面: (1) 生物地理学研究中物种分布格局及其成因; (2) 编制入侵植物名录、重建入侵历史和预测入侵风险; (3) 气候变化情景下植物分布格局的变迁机制; (4) 生物多样性编目和保护区规划; (5) 其他相关应用(如农业和民族药理学等)。最后, 针对数字植物标本应用中存在的问题进行了讨论, 并提出了应对策略, 同时提出开发将数字标本与其他学科大数据进行整合分析的新理论、新方法和新工具, 为植物学及相关学科提供参考。

关键词: 植物标本, 数字标本, 生物地理学, 入侵生物学, 气候变化, 保护生物学, 生物多样性, 民族药理学
中图分类号: Q94 文献标识码: A 文章编号: 1000-3142(2022)增刊 1-0127-15

Application of digitized plant specimens in scientific research: progress and challenges

ZHU Guangfu^{1,3}, YI Tingshuang¹, LIU Jie^{1,2*}

(1. *Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China;*
2. *CAS Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China;* 3. *University of Chinese Academy of Sciences, Beijing 100049, China*)

Abstract: About 390 million plant specimens are permanently stored in natural history museums around the world, and these specimens provide important support for scientific research. Increasingly, the application digital plant specimens have become widespread in scientific research, especially with the rapid development of the specimen digitization in the

收稿日期: 2021-11-18

基金项目: 国家自然科学基金(41971071, 31770367); 中国科学院“从 0 到 1”原始创新项目(ZDBS-LY-7001); 云南省“万人计划”青年拔尖人才项目(YNWR-QNBJ-2018-146)。

第一作者: 朱光福(1995-), 硕士研究生, 研究方向为物种生态位模型, (E-mail) zhuguangfu@mail.kib.ac.cn。

*通信作者: 刘杰, 博士, 副研究员, 研究方向为植物多样性起源演化及保护, (E-mail) liujie@mail.kib.ac.cn。

last two decades. The digitized plant specimens not only provide convenience for the traditional use of specimens, but also open a new window for other research fields. However, few studies in Chinese literature have reviewed the application of digitized plant specimens in scientific research, and the associated future challenges. In this review, we summarized the current state of specimen digitization across the globe, and provided a brief overview of its application in biogeography, invasion biology, climate change, conservation biology, and other fields. Here, we mainly focused on the following five aspects: (1) Species distribution patterns and determinants of biodiversity in biogeography; (2) the compilation of invasive plant checklist, reconstruction of the invasion history, and prediction of future invasion risk; (3) Range shift mechanisms of plants under climate change scenarios; (4) Biodiversity inventories and protected areas designing; (5) Other fields (Such as agriculture and ethnopharmacology). Finally, we discuss the challenges that digitized plant specimens are facing, and suggest potential solutions such as the adoption of new theories, methods and tools for the integration and analysis of digitized plant specimens with other big data in biological sciences. The present study may offer new insights into botany and other relevant disciplines.

Key words: plant specimens, digitized specimens, biogeography, invasion biology, climate change, conservation biology, biodiversity, ethnopharmacology

源于科学目的进行植物标本采集的历史可追溯到 1523 年 (Lane, 1996)。经过几个世纪的积累,目前在全世界 3 368 个与植物标本相关的自然历史博物馆里,保存有约 3.9 亿份植物标本 (Thiers, 2021)。这些植物标本主要储存在北半球,尤其以北美洲和欧洲的发达国家为主 (图 1)。储存量排名前十的国家包括美国 (7 869 万)、法国 (2 401 万)、英国 (2 346 万)、德国 (2 245 万)、中国 (2 016 万)、俄罗斯 (1 596 万)、日本 (1 233 万)、瑞士 (1 202 万)、瑞典 (1 187 万) 和意大利 (1 159 万) (图 1) (Thiers, 2021)。一直以来,这些植物标本在科学研究中发挥着重要作用,从传统用途的科学凭证 (Funk et al., 2005; Gómez-Bellver et al., 2020)、分类鉴定 (Francisco-Ortega et al., 2008; Jaca et al., 2018; Jukonienė et al., 2018) 和植物区系分析 (Sosef et al., 2017; Cámara-Leret et al., 2020), 到与新工具和新方法相结合的基因组学 (Twyford & Ness, 2017; Mosa et al., 2019) 和植物化学研究 (Oberlies et al., 2019), 再到关系人类粮食安全的植物病原体 (Ristaino et al., 2001; Abd-Elsalam et al., 2010; Bradshaw & Tobin, 2020) 和物种互作的植食关系等 (Lees et al., 2011; Meineke & Davies, 2019), 甚至外延到了科普教育等方面 (Cook et al., 2014; Monfils et al., 2017)。当前,人类活动介导的全球变化致使第六次生物大灭绝即将来临 (Lewis & Maslin, 2015), 故基于标本数据研究全球变化背景下生物多样性的变化机制并寻求保护方案显得

十分迫切 (Meineke et al., 2019a)。

植物标本可以通过自身实体和数字标本 (digitized specimens) 两种形式在科学研究中得以应用。基于标本本身的研究,比如传统的植物分类学,需要到标本馆实地查阅标本,但由于时间和空间限制,研究者终其一生也难以查阅完某一类群在全球各标本馆的所有保藏;而最近损毁性的标本利用,如通过在标本上取样提取遗传信息或植物化学成分等,则会对标本造成不可修复的损伤。得益于全球指数级增长的数字化标本数据,基于数字标本的研究刚好规避了上述缺点。Heberling 等 (2019) 分析了 1923—2017 年间 13 702 篇与植物标本相关的论文分布,发现最近二十年来没有论文数量遥遥领先的研究主题,但一些与数字标本密切相关的新兴研究领域 (如生物多样性和全球变化等),则呈现出逐年增长的趋势。与标本本身密切相关的传统研究如植物分类和植物区系,仍是植物标本的重要应用领域;伴随着数字标本应运而生的学科,则充分利用了数字标本“高通量”的特性,聚焦解决不同时空尺度的生物多样性格局和动态等相关问题 (Heberling et al., 2019)。

随着计算机技术的快速发展,20 世纪 80 年代方兴未艾的标本数字化行动为全球或大洲尺度的研究提供了新的机遇。数字标本作为过去几百年来生物与环境相互作用的记录,已经融到许多学科的研究中,如植物物候学 (Jones & Daehler, 2018)、保护生物学 (Nualart et al., 2017) 和全球变

化生物学 (Lavoie, 2013; Meineke et al., 2018; Lang et al., 2019) 等。随着计算机技术、概率论与统计学的快速发展, 用深度神经卷积网络大规模地鉴定标本 (Carranza-Rojas et al., 2017; Younis et al., 2018) 的技术得到发展, 这有望颠覆传统的依靠专家来鉴定标本的模式; 使用自动或半自动的方法提取植物的功能性状信息 (Gehan & Kellogg, 2017), 也将促进植物标本在生态学和进化生物学等领域的应用。植物标本增加新用途的趋势并不会掩盖或无效化其传统用途; 相反, 随着植物标本数字化的发展, 传统的使用方式将变得更为便捷。Heberling 和 Isaac (2017) 借用进化生物学的术语“扩展适应”来描述标本的使用范围远远超出了采集者们当初预期用途的现象。

关于实体标本在科学研究中应用的综述已有很多, 如 Funk (2004) 就列举了高达 72 个不同的用途, 但针对数字植物标本的应用仍乏善可陈。因此, 鉴于最近二十年出现了很多与数字标本相关的研究主题和方向, 本文综述了数字植物标本在科学研究中的应用概况及挑战。本文简要论述了数字标本的概念和全球标本数字化进展概况; 从生物地理学、入侵生物学、气候变化、保护生物学和其他应用 5 个方面细阐了数字标本在科学研究中的应用; 探讨了数字植物标本存在的问题及对策; 对标本数字化及数字标本应用的新的理论、方法和工具进行了展望。

1 全球标本数字化概况

按 Nelson 和 Ellis (2018) 的观点, 数字化标本 (digitized specimens) 的定义较为宽泛, 是指“把标本数据从模拟信号转换成数字信号, 这包括将标本标签和其他与标本相关的文件中的文本数据转录成数字记录, 但不考虑输入方式 [如语音、键盘、扫描/光学字符识别 (OCR)]; 将物理标本转换为数字图像, 包括二维、三维 (3D)、计算机断层扫描 (CT) 和其他可视化地表示物理标本的数字图像类型; 将模拟的音频和视频转换为数字记录; 将文本记录的采集地信息转换为特定地理坐标系下的 GPS 信息, 以及利用现有或可能具备的技术, 将其他标本相关数据转换为数字格式”。

大规模的标本数字化对于便捷地获取生物多样性数据和平衡全球标本资源的不均匀分布具有

重要作用 (Drew et al., 2017)。为了优化和加速这一过程, 各标本馆已经搭建了可负担得起的数字化工作流程 (Haston et al., 2012; Harris & Marsico, 2017)。植物标本相关的研究从 20 世纪 80 年代初就开始增加, 但与数字标本直接相关的研究直到 2008 年才开始激增 (图 2)。过去二十年, 数字化的生物多样性数据呈指数级增长, 这离不开各国对标本数字化的重视和投入。例如, 2001 年正式建立的全球生物多样性信息服务网络平台 (Global Biodiversity Information Facility, GBIF, <https://www.gbif.org/>), 截至 2021 年 9 月 9 日, 已收集了 3.52 亿份的植物分布记录数据, 其中来源于自然历史博物馆的标本数据有 8 840 万份 (25%)。近年来, 基于 GBIF 数据发表的研究论文呈现指数增长的态势 (图 2)。美国的整合数字化生物标本 (Integrated Digitized Biocollections, iDigBio, <https://www.idigbio.org/>), 已有 5 468 万份植物分布数据, 其中有地理信息的记录有 2 335 万份 (43%)。值得一提的是, iDigBio 已于 2017 年加入 GBIF。澳大利亚生物地图集 (Atlas of Living Australia, ALA, <https://www.ala.org.au/>), 已有 2 045 万份植物分布数据, 来源于自然历史博物馆的标本有 665 万 (32%)。巴西的物种链接库 (speciesLink, <http://www.splink.org.br/index?lang=en>), 已有 1 067 万份植物和真菌分布数据, 有地理信息的记录有 768 万 (72%), 其中部分数据在 GBIF 中共享。

中国的标本数字化, 可以追溯到 20 世纪 80 年代初期, 21 世纪后进入了快速发展时期, 其中以 2006 年初开通的 CVH 为代表, 为国内标本馆的数字化管理提供了范例 (覃海宁和杨志荣, 2011)。中国科学院于 2013 年 10 月正式加入 GBIF, 并通过中国节点向 GBIF 发布数据 (马克平, 2014)。目前, 国家植物标本资源库 (National Plant Specimen Resource Center, NPSRC, <http://www.cvh.ac.cn/>) 是国内最大的数字植物标本整合平台, 收集了 817 万份植物标本数据, 623 万张标本照片, 但遗憾的是尚无 GPS 信息提供 (数据统计于 2021 年 9 月 9 日)。而在 NSII 中, 约有 1 146 万份数字植物标本的记录 (访问时间为 2021 年 9 月 9 日)。而根据纽约植物园 Thiers (2021) 的统计, 中国馆藏有 2 016 万份植物标本。因此, 我国植物标本数字化率应已达 56.84%。

除了上述已有大量数字化标本的在线网站以外,还有正在开展的项目。例如,2015 年开始提议,由欧洲 21 个国家参与的科学采集分布式系统 (Distributed System of Scientific Collections, DiSSCo, <https://www.dissco.eu/>), 预计到 2025 年正式运行。届时,DiSSCo 将包括欧洲各博物馆、标本馆等机构中的 15 亿份生物和地质标本,将覆盖全球 80% 的生物和地质收集保存。尽管标本数字化活动在全球范围内不断增加,但仍有一些重要地区相对比较滞后。俄罗斯拥有全球排名第六 (1 596 万) 的植物标本储存量 (Thiers, 2021), 但目前还只有少量的、对特定研究人员开放的数字化数据,尚无供公众便捷获取的在线门户网站 (Ivanova & Shashkov, 2017)。印度的整体情况和俄罗斯很相似,标本数字化率很低,也还没有对公众开放的数字标本库 (Vattakaven et al., 2016)。其他发展中国家,比如非洲和亚洲的一些国家,标本数字化的工作更为迟缓。

2 数字植物标本相关的科学研究

数字植物标本中,地理、物候性状和文本描述中的药用价值等方面的信息在科学研究中使用频率最高 (图 3)。其中,使用最多的是具有经纬度和海拔的 GPS 数据。一般地,在数字化的过程中都会尽可能地包含地理坐标。在进行大尺度的数据分析时,可以直接从数据库下载分布数据,筛选出包含经纬度的记录,过滤和优化后用于分析。物种地理分布的信息通过与 GIS 技术结合,在生物地理学和保护生物学等领域中得到了广泛应用。其次是具有高分辨率的植物标本照片,除了可以获得用于植物分类学的形态性状外,还可以从标本中提取物候数据 (如花果期),通过不同年代的时间梯度标本数据的物候信息比较,可以探讨气候变化和物候变化的关联。此外,标本文字记录信息也有重要价值,特别是采集自少数民族地区的标本,其中有关本地民族对植物的利用信息 (如药用等经济价值),将为新的资源植物或药物的开发提供线索。下面,我们将从数字标本在生物地理学、入侵生物学、气候变化、保护生物学和其他应用等方面的应用进展进行介绍。

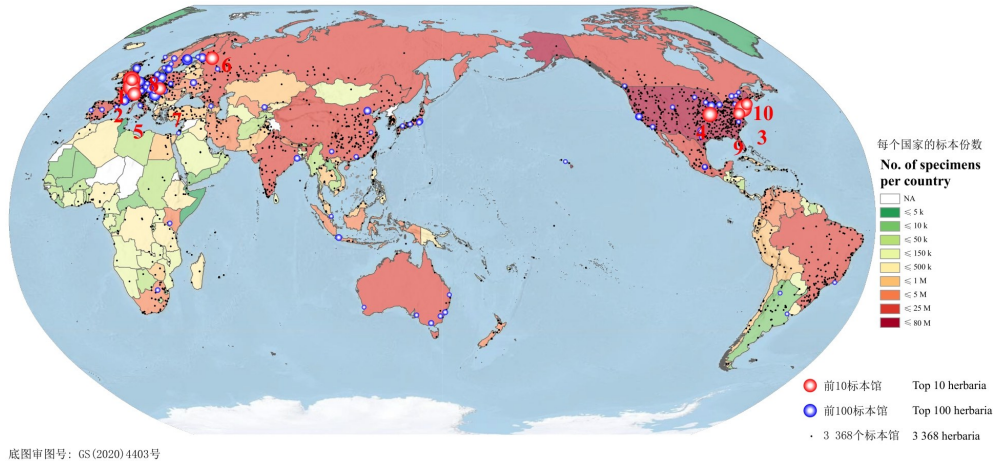
2.1 生物地理学

长期以来,物种多样性的地理分布格局是生

物地理学和宏观生态学研究的重要议题。其中一个核心的科学问题是为什么全球生物多样性呈现出不均衡分布?这不仅是生物多样性研究中的一个基本问题,它对于理解和预测全球变化对生物多样性潜在分布的影响,并最终阻止生物多样性的丧失至关重要 (Brooks et al., 2006)。物种分布格局研究中,数字化的标本数据是最快捷的来源,若辅以专家经验以及来自文献和志书等其他信息的支撑,即可获得比较全面的物种分布数据。近年采集的标本往往有准确的地理信息记录,可用于确定某一特定地点和时刻是否有植物存在的证据,这是物种分布格局分析的关键信息 (Merow et al., 2016)。

数字标本的指数级增长改变了人们认识生物多样性地理分布格局的方式,从过去以文献志书和地图册为主的模糊匹配,过渡到目前将密集样本点和 GIS 技术相结合的精准分析。过去,人们绘制全球植物多样性分布图主要是通过少量的标本记录信息和已发表的植物多样性地图的模糊关联。这方面的典型例子如 Barthlott 等 (2007) 通过 3 300 多幅物种丰富度地图,绘制了全球的维管植物多样性地图。目前,数字标本提供了大量分布数据,在尺度和精度上改变了人们对植物多样性分布格局的认知方式。Kier 等 (2009) 根据有物种分布范围的在线数据库和文献,比较了 315 903 种维管植物在未来土地利用格局变化情景下岛屿与大陆间的生物多样性保护风险,指出岛屿面积小且特有物种丰富的区域应该优先保护。即使是植物多样性和分布数据较为缺乏的非洲地区,也能找到足够多的数据用于生物地理学分析。例如, Sosef 等 (2017) 使用 61 万份有 GPS 数据的 RAINBIO 数据库,回答了非洲有多少个植物物种,并探讨热带非洲植物生活型多样性和分布的问题。然而,在基于数字标本的大数据分析中,标本鉴定错误和书写错误的情况并不罕见,因此仍旧不可忽视专家审核对标本准确程度的重要性。最近有研究发现,新几内亚作为世界上面积最大的热带岛屿,从 GBIF 等数据库获取的物种名录里,经专家审核后减少了 22% (Cámara-Leret et al., 2020)。

基于数字植物标本的生物地理学研究在具体类群上也可以发挥作用。如 Liu 等 (2018) 基于全球红豆杉属 (*Taxus*) 植物的数字标本记录信息,同时结合文献和野外考察记录,利用物种分布模型,

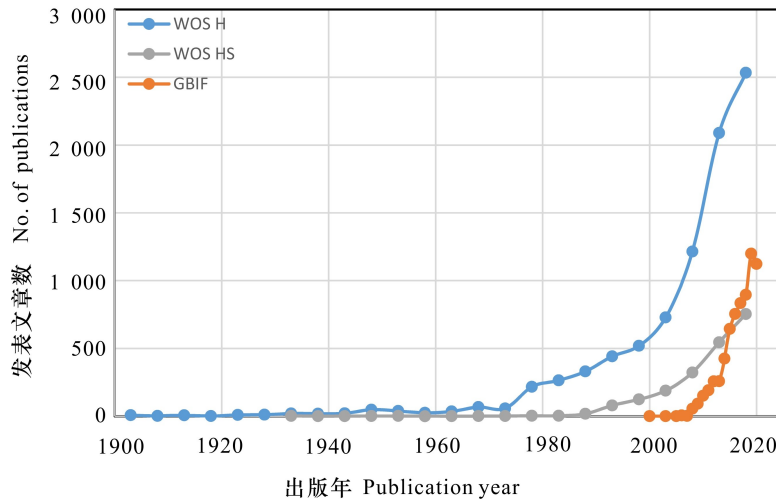


位列前十的标本馆及馆代码: 1. 英国皇家植物园邱园(K); 2. 法国国家自然历史博物馆(P & PC); 3. 美国纽约植物园(NY); 4. 美国密苏里植物园(MO); 5. 瑞士日内瓦植物园(G); 6. 俄罗斯科马罗夫植物研究所(LE); 7. 奥地利维也纳自然历史博物馆(W); 8. 英国伦敦自然历史博物馆(BM); 9. 美国史密森学会(US); 10. 美国哈佛大学(HUH)。分布数据来源于 Thiers (2021)。

List of the top 10 herbaria and their codes: 1. Royal Botanic Gardens, Kew (K), UK; 2. Muséum National d'Histoire Naturelle (P & PC), France; 3. New York Botanical Garden (NY), USA; 4. Missouri Botanical Garden (MO), USA; 5. Conservatoire et Jardin Botaniques de la Ville de Genève (G), Switzerland; 6. Komarov Botanical Institute of RAS (LE), Russia; 7. Naturhistorisches Museum Wien (W), Austria; 8. The Natural History Museum (BM), UK; 9. Smithsonian Institution (US), USA; 10. Harvard University (HUH), USA. Distribution data from Thiers (2021).

图 1 全球植物标本馆的分布和每个国家的植物标本储存量

Fig. 1 Global distribution of herbaria and the number of plant specimens in each county

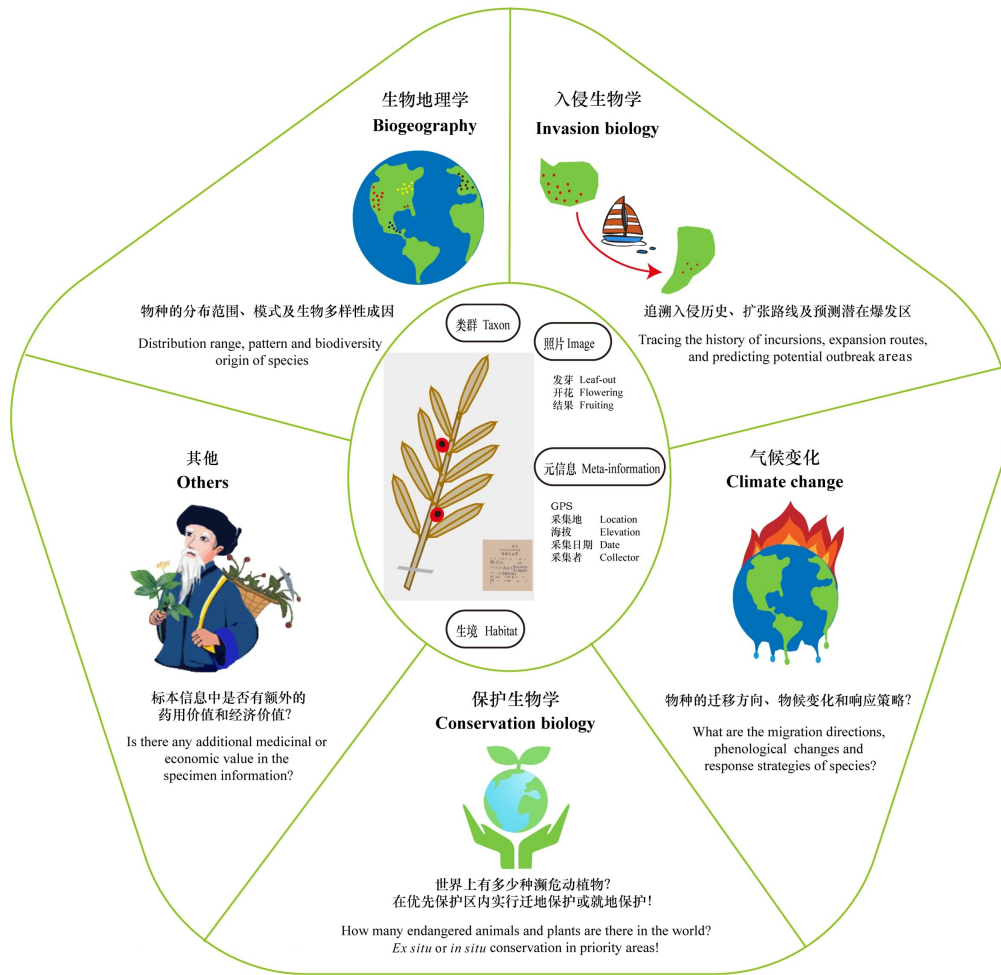


过去一个世纪与植物标本相关的同行评审论文发表数量(1900—2020, Web of Science; 2000—2020, GBIF)。WOSH 指 Web of Science Core Collection 的主题字段中通过(“herbarium”或“herbaria”)检索的文章,最后保留了8 721篇文献;WOSHs 指 Web of Science Core Collection 的主题字段中通过(“herbaria specimens”或“herbarium specimens”)检索的文章,最后保留了2 058篇文献;GBIF 指在 GBIF 的文献数据库里将文献类型限制在了“期刊文章”,相关性限制在了“GBIF 使用的”和“GBIF 引用的”,同行评审选择“是”,最后保留了6 095 篇文献,GBIF 数据库无法直接区分动物标本或植物标本文献(上述信息统计于2020年10月18日)。

Number of peer-reviewed articles using herbaria data published over the past century (1900–2020, Web of Science; 2000–2020, GBIF). WOSH refers to the articles retrieved by the search string “herbarium” or “herbaria” in Web of Science Core Collection (8 721 articles finally retained); WOSHs refers to the articles retrieved by the search string “herbaria specimens” or “herbarium specimens” in Web of Science Core Collection (2 058 articles finally retained); GBIF denotes the articles obtained from the Global Biodiversity Information Facility (GBIF) literature database; the literature type was limited to “Journal articles”, and the relevance was limited to “GBIF used” and “GBIF cited”, peer-review chose “Yes”, and finally 6 095 articles were retained. It was not possible to directly distinguish between animal specimens and herbarium specimens based on the GBIF literature database (The above information is accessed on Oct 18, 2020).

图 2 Web of Science 和 GBIF 中有关标本和数字标本发表的文章数量与年份的关系

Fig. 2 The relationship between the number and year of publications on specimens and digitized specimens in the Web of Science and GBIF



位于图片中心的是可以从植物标本里获取的标本叶片和各种数据类型。外围的插图代表这些数据能解决的科学问题。Specimen sheet in the center surrounded by various information data that can be retrieved from plant specimens. The peripheral illustrations show the scientific questions that these data can be applied to.

图 3 数字植物标本蕴含的科学信息及其主要应用领域
Fig. 3 Scientific information of digitized plant specimens and its main application fields

模拟绘制了红豆杉属植物最为准确全面的全球分布图,为该属物种的鉴定和保护提供了重要依据。此外,利用同样的分布数据,结合历史时期的气候数据,对喜马拉雅红豆杉(*T. wallichiana*)在末次间冰期、末次冰期和当前的潜在分布进行模拟,发现物种的分布格局与基于分子证据的群体动态模型检测的结果相吻合,佐证了物种的群体在冰期扩张的假说(Liu et al., 2013)。数字标本的分布信息还能探讨植物类群分布格局的成因。如 Xu 等(2019)基于数字标本数据,利用约 400 种栎属(*Quercus*)的 20 000 多个分布点数据,比较了北半球栎属的物种丰富度和气候之间的关系,结果表

明亚热带栎属的物种丰富度受水分可获得性的影响最大,而温带则受控于环境温度。基于来自数字标本的分布数据,还可以推断地质时期物种的潜在分布,Dakhil 等(2019)发现中国西南山地的针叶林类群在应对第四纪气候变化时有相反的模式,喜温针叶林第四纪经历了生境适宜区锐减,喜冷针叶林适宜区则向青藏高原方向扩张。

2.2 入侵生物学

入侵植物指外来植物成功占领新的生境并大量繁殖,并损害到本地生态系统。植物的长距离扩散在自然界中较为罕见(Nathan & Muller-Landau, 2000),人类活动剧烈地影响着植物分布

的组成和全球格局(van Kleunen et al., 2015)。伴随人类活动搭乘“顺风车”的外来植物,占领新的生境并改变群落结构、组成和功能,进而导致生物多样性和生态系统服务功能的丧失(Tilman, 1999),外来入侵物种是导致物种灭绝的主要原因之一(Assessment Millennium Ecosystem, 2005)。因此,了解植物入侵的途径和时空动态,有利于保护自然生态系统,防止进一步破坏并可优先采取管理行动(van Kleunen et al., 2015)。

生物入侵会严重威胁当地植物的多样性,造成严重的经济损失和生态破坏,在国家层面需要厘清入侵植物的名录和分布状况。数字植物标本可为上述问题提供重要支撑,如截至 2012 年底,基于标本记录和分类学考证,辅以文献报道和野外调查,发现中国入侵植物有 806 种(马金双, 2013)。加强入侵植物现状调查,有利于当地政府制定管理方向和防治入侵生物。生物入侵是普遍存在的现象,不仅存在于国内,而且国外也存在。例如,在不丹,Dorjee 等(2020)通过数字标本记录和植物区系分析,并结合其他资料,发现 101 种入侵植物中大部分都是无意识引入的,只有极少部分是作为观赏植物和牧草而引入。

虽然现今的野外调查只能描述入侵植物的现状,但是通过历史标本的分析,可以帮助我们确定外来物种的入侵过程,重建它们在时间和空间上的入侵历史。Fuentes 等(2008)通过比较不同时期的 71 764 份数字标本数据,发现智利存在两次植物入侵高峰期,第一次是农业集约化(1910—1940)时期,第二次是麦田机械化(1980—2000)时期。与本土植物相比,外来植物的入侵扩张时间较短,因此占有面积只是气候潜在适宜区的一部分(Bradley et al., 2015)。通过比较外来物种当前占有面积和气候潜在分布区,可以帮助我们辨识成功入侵物种的特征,建立入侵风险等级,并对潜在的入侵物种严加防范(Dodd et al., 2016)。通过入侵植物的数字标本记录,利用物种分布模型预估入侵物种在其原产地之外的潜在分布区,是预防生物入侵的重要手段。Zhu 等(2007)用 GARP 模型预测紫茎泽兰(*Ageratina adenophorum*)未来的分布可能会在我国云贵高原、四川盆地、东南沿海、海南岛和台湾等地扩张。最近,南非也开始担心欧洲桉木(*Alnus glutinosa*)是否会在本国扩张,毕竟该物种有入侵美国和新西兰的先例。Keet 等

(2020)基于来自 GBIF 的数字植物标本信息及其数据,用 1 141 个全球的分布数据,模拟了欧洲桉木未来的潜在分布区,发现水分可获得性是其扩张的制约因素。

2.3 气候变化

全球气候变化是 21 世纪的主要挑战之一。人类活动造成的全球升温比工业化前高了约 1.0 °C,预计到 2030 年至 2052 年将达到 1.5 °C(IPCC, 2018),地球已经进入了以人类为主导的人类世(Anthropocene)(Lewis & Maslin, 2015)。标本采集是未经计划的、长期的和大尺度的实验,研究历史标本可以帮助我们理解全球变化背景下生物的分布和物候响应等过程。植物是固着生长的,生活史中与周围环境变化密切关联。这使得作为物候事件快照而被保存起来的植物标本,成为分析植物应对气候变化的理想数据。

植物对全球变暖等不利条件的可能反应是核心分布区在海拔和纬度梯度上的迁移,这可通过植物标本的时间序列和物种分布模拟分析来追踪和预估。将野外观测与植物标本的数据进行比较,可以发现随着全球气候逐渐变暖,植物将会沿着它们原有的气候生态位,向高海拔和高纬度迁移。Wolf 等(2016)统计了美国加利福尼亚州 4 426 种植物的 681 609 份具有地理信息参考的标本数据,发现 15% 的物种在过去的一个世纪里向高海拔方向迁移。将标本数据和野外重调查数据相结合,可以更准确估计物种迁移速率的大小。通过数字标本的分析发现在台湾的高山地区,1906—2006 年间高山植物向山顶方向的迁移速率为每年 3.6 m(Jump et al., 2012)。向山顶迁移可以应对变暖,但也面临着可利用的土地面积减少和地形限制因素加强等风险(Elsen et al., 2020)。气候变暖不仅直接影响着植物的地理分布,还有可能波及群落中动植物的互作关系(如传粉、取食等)。基于数字标本,Meineke 等(2019b)调查了 112 年间 4 种植物标本在野外被昆虫啃食的情况,发现 21 世纪初采集的植物标本比 20 世纪初采集的标本被昆虫破坏的可能性高 23%,推测这与全球变暖促进昆虫的生长发育有关。

植物应对全球气候变暖的方式除了核心分布区的迁移,还可以通过调节物候期来适应温度的变化。植物物候学主要研究植物对过去和未来气候变化的响应模式(Willis et al., 2017a),物候变

化影响着生物繁殖、种群生存、物种边界和生态系统服务(Chuine, 2010)。数字标本可以为历史物候变化提供证据。首次使用植物标本中的物候信息研究植物物候与气候变化关系的是 Primack 等(2004)。他通过分析 1885—2002 年间采集的 372 份植物标本的物候信息,发现植物花期平均提前了 8 d,与 Willis 等(2017a)统计的中位数(9.5 d)很相近。年平均气温升高是使植物花期提前的主要因素,但冬季和秋季的气温升高对花期的影响完全相反。Hart 等(2014)分析了喜马拉雅地区的 10 295 份杜鹃花属(*Rhododendron*)数字标本数据,发现冬季气温升高通过促进越冬芽形成而使花期提前,而秋季气温升高通过抑制低温需求而使花期延后。当然,基于植物标本得到的物候数据的准确性需要通过与实际观测数据进行验证(Bertin, 2015)。Bolmgren 和 Lönnberg(2005)比较了 77 种共 5 500 份植物标本数据的花期和野外观测数据,发现两者间没有显著的区别。Davis 等(2015)通过比较 20 个物种在 160 年的时间跨度上的标本物候数据和野外实地观测的物候数据,证实植物标本记录可作为物候指标(如温度和光周期)对气候变化响应的可靠来源。

植物物候学为人们认识植物响应过去的物候提供了可能,但同时也有更多的问题亟待解决。在过去的十年中,超过 30 项研究使用植物标本来分析开花物候随时间的变化,但迄今为止对热带地区的研究还普遍缺乏(Jones & Daehler, 2018)。同时,由于没有统一标准的植物标本物候性状测量方法,因此不同作者的研究结果难以进行直接比较(Willis et al., 2017a)。植物物候学未来研究的一个趋势是将数字植物标本数据与其他来源的物候数据进行整合,相互支撑验证,如结合历史时期的照片(Miller-Rushing et al., 2006; Panchen et al., 2012)以及 Landsat 和 MODIS 卫星遥感数据(Park, 2012)。

2.4 保护生物学

气候变化和生物多样性丧失是人类面临的双重危机(Corlett, 2020),两者间有密切联系。生物多样性是全球生态系统健康和人类福祉的基础(Rands et al., 2010)。据世界自然保护联盟(International Union for Conservation of Nature, IUCN, <https://www.iucnredlist.org/>)估计,全球 41% 的两栖类、14% 的鸟类、26% 的哺乳类和 34% 的松柏

类处于濒危状态。然而全世界有高等植物(苔藓、蕨类、裸子植物和被子植物)403 911 种(Lughadha et al., 2016),IUCN 尚未能做出全面的评估。最近,全球植物保护战略(Global Strategy for Plant Conservation)组织评估了 241 919 种植物(包括藻类),其中 11% 的植物处于濒危状态(Bachman et al., 2018)。在中国高等植物红色名录中,有 3 879 种为受威胁物种(覃海宁和赵莉娜,2017)。

保护区是全球生物多样性保护的基础,通过划定禁止人类活动(如伐木和狩猎)的区域来避免生境被干扰。利用数字植物标本的分布数据可开展保护区有效性的评估和优化工作,为决策管理和生态恢复提供理论依据,以采取迁地保护或就地保护措施。随着标本数字化,物种分布的地理坐标可以快速获取,通过将物种的分布点与不同时期(过去、现在和将来)的气候相关联,来模拟物种潜在分布区的变化成为常规手段(Nelson & Ellis, 2018)。标本记录可用于编制某一地区的濒危物种红色名录和确定优先保护区等关键生态问题。Lienert 等(2002)通过分析北温带獐牙菜(*Swertia perennis*)在瑞士 63 个地点的 127 年标本采集史,认为有 15 个地点的种群因剧烈的农业活动和生境破碎化已经灭绝。在 27 个西班牙已灭绝植物中,生境丢失是物种灭绝的主要原因,也与生物入侵和过度放牧有关(Aedo et al., 2015)。Bai 等(2020)分析了 535 种报春花科在中国的分布区,发现国家级自然保护区面积仅占报春花科分布区的 29.5%,保护遗漏区主要集中在我国西南山地。

针对特定类群的珍稀濒危植物保护策略,我们可以采取移栽到植物园等迁地保护措施,但是,基于大尺度的生物多样性保护时,就地保护将会是最好的选择。在确定中国特有种子植物有 14 939 种的基础上(Huang et al., 2011),通过标本数据和文献志书等途径构建了这些物种的分布数据库。Huang 等(2016)分析了我国的物种多样性分布格局,发现种子植物的热点地区为秦岭和横断山地区,现有自然保护区面积仅占种子植物热点地区的 26%,保护的空白区将是未来保护区设定的优先区。保护区和生物多样性地区不重叠的现象并非只发生在特有种子植物里。Zhang 等(2015)评估了《中国生物多样性红色名录·高等植物卷》的 3 244 种濒危物种,通过整合 CVH 在线

标本数据库及文献志书的分布数据,发现保护区(包括国家级自然保护区和省级自然保护区)面积仅占受威胁物种分布区的 28%,其中有 397 个受威胁物种的分布区完全没在保护区内。现有保护区和待保护物种分布区不完全重叠是不争的事实,这为今后保护地体系的优化提供了科学依据。

2.5 其他应用

在农业相关的领域中,数字标本也可用于预测杂草、作物及伴生物种的适宜性分布区及变化。生境适宜性是杂草在全球范围内扩张风险评估的一个有效指标(Richter et al., 2013),因此,基于数字标本的地理分布数据,生态位模型能为预防和控制全球农业杂草扩张提供有力的参考。Wang 和 Wan(2020)使用 GBIF 数据库中 10 种高危杂草分布数据,结合环境因素,预测橄榄和芥末等种植园将有很高的被入侵风险。类似的例子还有黄顶菊(*Flaveria bidentis*),原产于南美洲,在中国定植后影响农业和畜牧业,其潜在分布区将会以河北为中心向西(陕西和山西)和向东北方向(内蒙古和辽宁东部)扩张(Fan et al., 2018)。此外,麻风树(*Jatropha curcas*)可以作为生物能源植物代替化石燃料实现可持续发展,但是在未来气候变化情景下其潜在适生区将会收缩 45%(Hu, 2017)。除了杂草会影响农业以外,传粉者对农业的影响也不能忽视。在拉丁美洲,在未来气候变化情景下,蜜蜂和咖啡的适宜分布区主要表现为共同减少,但在某些区域也表现出两者同时增加,或蜜蜂的分布区减少,但咖啡适宜区却增加的反模式(Imbach et al., 2017)。

数字标本还可用于涉及人类健康福祉的民族药理学。比如,药用植物两千多年来一直被用于传统医药体系中,有效保障了人类的健康。据 Jaiswal 等(2016)统计,传统中药里至少有三分之一的植物被用作药用植物。汇编和分析不同来源的文献有利于增加对药用植物的认识,而数以亿计的数字植物标本数据,为更大规模的搜索研究提供了可能。在从植物标本馆收集药用植物信息的评估中,von Reis(1962)发现哈佛大学植物标本馆的 250 万份标本中,有近 6 800 份标本描述了药用功能。自此,植物标本已成为获取民族植物学信息的途径之一。例如,Lulekal 等(2012)通过研究埃塞俄比亚国家植物标本馆的 293 个物种共 4 717 份标本,发现其中 101 个物种记录有药用信

息,其中 29 种只有 1 份标本记录。孤本标本的出现与药用植物本身的珍贵性和采集困难有关,这也会给评估药用植物的保护状态带来困难。最近,Souza 和 Hawkins(2017)通过比较巴西豆科植物在 105 个出版刊物和 15 个在线标本馆中药用植物的用途,发现 51%的用途仅在文献中记载,17%的用途仅在植物标本的采集记录中出现。这表明标本的记录信息可以有效地补充文献记录。出于植物区系和分类修订等目的而采集的标本,因其采集初期不是为了民族药理学研究,这会导致通过标本记录查找药用信息犹如大海捞针。数字化的标本信息,可以在较大的范围内快速搜索,这在一定程度上克服了搜寻的困难。

3 数字植物标本存在的问题及其对策

尽管数字标本有如上所述的诸多用途,但基于数字标本的数据也可能存在偏差和局限性(Meyer et al., 2016; Daru et al., 2017)。这些缺陷可能源自标本采集、鉴定、数字化以及标本数据的分析和解释等各个环节中。虽然植物标本通常由经验丰富的植物学家鉴定,但标本错误鉴定可能普遍存在,这在一定程度上影响了数字植物标本数据的质量。如 Goodwin 等(2015)发现非洲 21 个国家共 4 500 份的姜科标本中至少有 58%的标本存在学名错误的情况;超过 50%的热带植物标本也存在学名错误。与野外观测数据不同的是,数字植物标本可以通过照片或标本记录对异常标本进行确认。在未来野外采集标本时,配合使用识别率较高的应用程序(如花伴侣和形色等)(许展慧等, 2020),有可能在一定程度上减少入库数字标本的错误鉴定。此外,使用深度卷积神经网络的计算机视觉技术,可以帮助鉴定标本和获取物候性状数据。例如,Younis 等(2018)对 GBIF 中 1 000 种常见植物的标本图片进行了机器学习训练,进而测试了 17 万张数字标本照片,结果发现物种鉴定率可达 82.4%。自然历史博物馆中还有很多只鉴定到了属的标本,神经网络模型能作为初步鉴定供专家再次核查之用,这可以加速标本准确鉴定的过程。但遗憾的是,该方法仍停留在摸索测试阶段,尚无实际的可应用程序。值得注

意的是,将植物标本训练的模型,转移学习用于鉴定野外活体植物效果很差,这可能与标本处理过程中的植物三维结构变化和植物器官色泽的改变有关(Carranza-Rojas et al., 2017)。

植物标本在采集的过程中就会引入额外的偏差。植物学家的采集与个人的兴趣偏好有关,同时也会考虑采集难易性和成本,最终的结果会出现所谓的“大路货”横行的局面。例如,Daru等(2017)基于GBIF的500万标本数据发现标本采集在时间、空间和采集者上存在显著的偏差。采样偏差的另一来源是采集自同一地点的重复标本被送往多个标本馆,随后每一份标本都被当做独立的样本进行分析。数据质量问题也有可能出现在数字化的过程中,例如模糊的手写字迹在转录标本记录信息时有可能出错。此外,不同的国家和个人使用不同的采集记录字段和术语,这给全球数字化标本的整合带来了挑战。虽然数字化植物标本不可避免地存在采样偏差的问题,但在选择物种和训练数据集的过程中也能在一定程度上减少偏差。利用统计学方法纠正偏差是分析数据时的最后手段,例如,Phillips等(2009)开发了一种方法,利用出现记录数据在物种分布模型时处理地理采样偏差。此外,将采集者作为一个变量也许能提高物种分布模型的效果(Daru et al., 2017)。

数字标本使用中的另一个挑战是准确获取GPS地理坐标信息,早期采集的标本尚无GPS记录,而我国的国家植物标本资源库(NPSRC)的数据未开放标本采集的地理学信息字段(经纬度)。因此,需要依据采集地点的信息,获取准确的空间坐标(经纬度)是数字标本使用过程中面临的挑战之一。通过标本标签匹配空间参考的一般流程如下:(1)转录标签文本;(2)处理历史地名变更;(3)提取地名文字与地名数据库匹配。准确识别标签信息是获取空间参考的前提,尤其是转录非母语文本时需要更加注意。例如,伦敦自然历史博物馆的产地为中国的3736份模式标本中,80%的地名文本中包含错误(Lohonya et al., 2020)。所以在录入文本的时候,需要对模糊或不确定的文字进行特殊标记,以便于人工核查。地名文本中有时也会包含多国文字,如汉语、英语、葡萄牙语和法语等,这时借助谷歌翻译工具转为自己熟悉的文字是个不错的办法。针对更为久远的标

本,如20世纪中期以前的标本,探险队(或采集队)一般会写探险游记专著,其中可能包括详细的考察路线和手绘地图,有些地名还包括各种语言版本,甚至还会包含对应的经纬度信息(Lohonya et al., 2020),这些都可以给相关标本提供佐证资料。其次,历史地名变更是标本信息中常见的问题,这种情况下,旅游网站可能会有相关的曾用名的介绍。在我国,针对县市以上地名的变更,国家标本资源共享平台(NSII)开发了在线中国行政区划变迁数据查询匹配工具(<http://www.nsii.org.cn/2017/Placenamesautocheck.php>)。此外,也可以参考维基百科或百度百科,由全球上亿用户共同编辑的百科全书里面也会有曾用地名介绍。最后就是匹配地名数据库(或地名词典)。常见的地名数据库有谷歌地图、百度地图和高德地图等,这些在线网站都提供了应用程序接口(application programming interface, API)供编程用户批量处理使用。也有一些在线数据库针对这些API再封装程序,如MapLocation(<https://maplocation.sjfkai.com/>)。对于编程用户来说,这一步的难点不是搭建查询系统,而是如何从包含描述信息的标本记录文本里提取出只包含地名的字符。以TIOBE排行第三的语言Python为例,jieba包(<https://pypi.org/project/jieba/>)是非常流行的中文分词工具,支持精确模式、全模式和搜索引擎模式。如果想以实现中文标本标签转空间的编程为参考,需要把重点放在处理历史地名变更和通过中文分词剔除描述性的词汇而只留下标准地名文本。

4 数字化植物标本的未来

往更高层级思考,上述的数字植物标本只是一个方面的数据,顺此外延可以拓展到与此相关的DNA数据、物种性状、生态学信息等诸多方面,但这些信息的耦合关联需要在一个更大的顶层框架下实现。达尔文核心标准(Darwin Core Standard, DwC)是一组具有明确定义的用于整合编码数据的专业术语,这些术语可以被理解,也可以被机器译释。总的来说,达尔文核心可以分为9大类,即记录相关的术语、分布数据、事件、地点、鉴定、类群、地质环境、资源关系和度量。达尔文核心标准是一个仍在发展中的、可被扩展的、用于服务生物多样性研究的框架。故达尔文核心标

准为整合不同来源的生物多样性数据, 提供了一个稳定、直接和灵活的框架 (Wieczorek et al., 2012)。

将野外拍摄的植物活体照片关联到数字植物标本中, 这将为鉴定植物和重访提供便利。早在黑白照片时代, 人们就已经将照片作为采集植物标本的补充数据。例如, Gómez-Bellver 等 (2020) 就例证了 1934 年的照片, 其内容包括用于鉴定植物的局部细节照, 以及植物生长的生境照。事实上, 在园艺学中, 有些栽培品种的鉴定就只能通过颜色来识别。Gómez-Bellver 列举了在凭证标本中需要加入植物活体照片的情况: (1) 标本占大体积的植物 (如棕榈和苏铁等); (2) 多肉植物或多刺植物; (3) 有毒或有刺激性乳汁的植物 (如大戟科); (4) 只采集种子的特殊用途时 (如种质资源库); (5) 压制的过程中形状和颜色变化很大的植物 (如兰科); (6) 其他特殊情况如只有单株植物或宗教信仰圣地附近采集的标本。在智能手机高度普及的今天, 野外采集标本时随手获取高分辨的彩色照片已不是问题。通过达尔文核心标准, 将植物活体照片与数字化植物标本结合, 是各大标本馆在数字化的过程中需要考虑的。

在过去的十年里, 已经积累了大量的 DNA 数据。高通量测序和长读测序技术的发展引来了基因组学时代, 进而助推了进化和生态学研究。Sanger 测序技术、基因组浅层测序和全基因组重测序等技术, 都可以将植物标本作为基因组测序的材料使用, 且其测序数据可重复使用 (McKain et al., 2018)。标本馆向研究人员提供基于标本的分子材料时, 就伴随着对标本造成不可修复的损伤。为了最大限度地重复使用这些遗传数据, 标本馆有必要把与标本材料相关的各种遗传数据纳入数字标本体系中。尤其是, 需要将凭证标本信息与相关的 NCBI 序列号进行关联。

数字化过程中公众科学的力量不可小觑, 他们可用于加速标签文本的转录、地名描述转空间参考和依据标本照片标记物候性状数据等 (Ellwood et al., 2015)。让公众充分参与到科学研究实践中, 成功的案例如 CrowdCurio。该网络平台可用于对数字化植物标本的物候信息进行标注。Willis 等 (2017b) 也证明了众包平台 (crowdsourcing platform) 中普通公众提交的结果和专家们统计的物候结果没有显著的差别。国内也有类似的阿里众包

(<https://newjob.taobao.com>), 但目前还没有专门针对科研用途的众包平台, 使用公众科学的力量可在更精细的维度上获取生物多样性数据。值得一提的是, 众包的收入一般都很低, 但如果从获得技能和知识作为回报的角度考虑, 众包将会在科学和公众之间架起桥梁, 积极地引导公众参与科学研究的

活动。

整合散存在世界各地的采集自中国的数字标本, 将为我国生物多样性研究提供更为全面的数据。但是, 这也增加了采集自同一地点的复份标本被送往多个标本馆保存后却被当做独立样本研究的风险, 研究人员在分析数据和解释结果的时候需要注意这一点。汇总某一国家散落在全球的数字标本的实践中, 最为成功的是巴西。由巴西政府在 2010 年建立的重建植物区系数字标本馆 (Reflora Virtual Herbarium, RVH, <http://reflora.jbrj.gov.br/reflora/herbarioVirtual/>), 其宗旨是“提供主要储存在海外植物标本馆的有关巴西植物的图像和信息”。RVH 数据库包含 380 万张植物标本照片, 其中带地理信息的有 161 万张。

数字标本馆与用户之间需要建立起良好的反馈系统, 这将有助于用户将过滤后的数据反馈到数据库中, 提升数字标本的数据质量。植物标本馆通过线上或线下的方式, 来调查研究人员如何使用植物标本的案例有很多 (Carine et al., 2018; Canteiro et al., 2019), 但目前仍缺乏数字标本使用后的数据反馈机制。研究人员在使用数字标本的过程中, 不可避免地需要纠正各种存在的问题, 如经纬度信息缺失、标本鉴定错误或标签信息录入错误等。研究人员花了大量的时间过滤好的数据, 需要及时更新到数据库中, 提升数据库的数据质量。因此, 我们认为, 数字标本平台迫切需要建立一个用户数据反馈通道, 用以提升库存数据的质量。除此之外, 全球的主要在线标本馆之间需要使用统一的术语和字段, 并建立数据交换的机制, 从而避免因害怕数据垄断而不敢数据互换的担忧。数据库共享堪称典范的是 DNA 序列数据库 GenBank、EMBL 和 DDBJ 之间的实时数据交换, 这极大地便利了分子生物学的研究, 实际上也实现了数据的多地备份。数字标本平台之间应该引入类似的机制, 实现不同国家的用户可以在一个数据库中获得全面的数据。

在达尔文核心标准下, 包含活体植物的彩色

照片和 DNA 信息将提供更为全面的多样性数据;充分利用公众科学的力量获取物候性状和信息核查,也能加速数字化的过程;通过数据交换和反馈机制,整合散落在各标本馆的采集自中国的标本信息,也将为国内生物多样性数据的汇集添砖加瓦,最终助力相关领域的研究。鉴于上述提及的数字标本在研究中的价值和显示出的潜力,以及大量的数字植物标本即将上线,以植物为基础的生物多样性研究的潜能是巨大的。使用这些数字标本将需要更严格的标准化方法,以及开发用于大规模数据收集和分析的新工具(Soltis & Soltis, 2016)。大数据时代下需要新的理论、方法和工具让尘封了百年的标本焕发出更多的活力。

5 结 论

数字植物标本在分类和区系等传统领域中依然发挥着重要作用,且在生物地理、入侵生物学、保护生物学等领域衍生出了许多新的用途。通过充分理解这些丰富的数字标本的潜在作用和局限性,结合标本数字化、数据共享和整合新工具等,为解决有关植物多样性分布格局及成因、全球气候变化下的植物入侵、分布格局变迁和植物多样性保护,甚至关乎人类健康与福祉的问题提供了契机,未来,数字标本必将在科学研究中发挥更大作用。

致谢 Moses C. Wambulwa 帮忙修改英文摘要,罗媛女士为图 3 中的其他用途绘的插图,在此一并感谢!

参 考 文 献:

- ABD-ELSALAM KA, YASSIN MA, MOSLEM MA, et al., 2010. Culture collections, the new herbaria for fungal pathogens [J]. *Fungal Divers*, 45(1): 21-32.
- AEDO C, MEDINA L, BARBERÁ P, et al., 2015. Extinctions of vascular plants in Spain [J]. *Nord J Bot*, 33(1): 83-100.
- BACHMAN SP, LUGHADHA EMN, RIVERS MC, 2018. Quantifying progress toward a conservation assessment for all plants [J]. *Conserv Biol*, 32(3): 516-524.
- BAI YH, ZHANG SY, GUO YP, et al., 2020. Conservation status of Primulaceae, a plant family with high endemism, in China [J]. *Biol Conserv*, 248: 108675.
- BARTHOLOTT W, HOSTERT A, KIER G, et al., 2007. Geographic patterns of vascular plant diversity at continental to global scales [J]. *Erdkunde*, 61(4): 305-315.
- BERTIN RI, 2015. Climate change and flowering phenology in Worcester County, Massachusetts [J]. *Int J Plant Sci*, 176(2): 107-119.
- BOLMGREN K, LÖNNBERG K, 2005. Herbarium data reveal an association between fleshy fruit type and earlier flowering time [J]. *Int J Plant Sci*, 166(4): 663-670.
- BRADLEY BA, EARLY R, SORTE CJB, 2015. Space to invade? Comparative range infilling and potential range of invasive and native plants [J]. *Glob Ecol Biogeogr*, 24(3): 348-359.
- BRADSHAW M, TOBIN PC, 2020. Sequencing herbarium specimens of a common detrimental plant disease (powdery mildew) [J]. *Phytopathology*, 110(7): 1248-1254.
- BROOKS TM, MITTERMEIER RA, DA FONSECA GAB, et al., 2006. Global biodiversity conservation priorities [J]. *Science*, 313(5783): 58-61.
- CÁMARA-LERET R, FRODIN DG, ADEMA F, et al., 2020. New Guinea has the world's richest island flora [J]. *Nature*, 584(7822): 579-583.
- CANTEIRO C, BARCELOS L, FILARDI F, et al., 2019. Enhancement of conservation knowledge through increased access to botanical information [J]. *Conserv Biol*, 33(3): 523-533.
- CARINE MA, CESAR A, ELLIS L, et al., 2018. Examining the spectra of herbarium uses and users [J]. *Bot Lett*, 165(3/4): 328-336.
- CARRANZA-ROJAS J, GOEAOU H, BONNET P, et al., 2017. Going deeper in the automated identification of herbarium specimens [J]. *BMC Evol Biol*, 17(1): 181-194.
- CHUINE I, 2010. Why does phenology drive species distribution? [J] *Philos Trans R Soc B*, 365(1555): 3149-3160.
- COOK JA, EDWARDS SV, LACEY EA, et al., 2014. Natural history collections as emerging resources for innovative education [J]. *BioScience*, 64(8): 725-734.
- CORLETT RT, 2020. Safeguarding our future by protecting biodiversity [J]. *Plant Divers*, 42(2): 221-228.
- DAKHIL MA, XIONG QL, FARAHAT EA, et al., 2019. Past and future climatic indicators for distribution patterns and conservation planning of temperate coniferous forests in southwestern China [J]. *Ecol Indic*, 107: 105559.
- DARU BH, PARK DS, PRIMACK RB, et al., 2017. Widespread sampling biases in herbaria revealed from large-scale digitization [J]. *New Phytol*, 217(2): 939-955.
- DAVIS CC, WILLIS CG, CONNOLLY B, et al., 2015. Herbarium records are reliable sources of phenological change driven by climate and provide novel insights into species' phenological cueing mechanisms [J]. *Am J Bot*, 102(10): 1599-1609.
- DING WX, GU JY, CAO L, et al., 2014. Traditional Chinese

- herbs as chemical resource library for drug discovery of anti-infective and anti-inflammatory [J]. *J Ethnopharmacol*, 155(1): 589–598.
- DODD AJ, MCCARTHY MA, AINSWORTH N, et al., 2016. Identifying hotspots of alien plant naturalisation in Australia: approaches and predictions [J]. *Biol Inv*, 18(3): 631–645.
- DORJEE, JOHNSON SB, BUCKMASTER AJ, et al., 2020. Weeds in the land of Gross National Happiness: Knowing what to manage by creating a baseline alien plant inventory for Bhutan [J]. *Biol Inv*, 22(10): 2899–2914.
- DREW JA, MOREAU CS, STIASSNY MLJ, 2017. Digitization of museum collections holds the potential to enhance researcher diversity [J]. *Nat Ecol Evol*, 1(12): 1789–1790.
- ELLWOOD ER, DUNCKEL BA, FLEMONS P, et al., 2015. Accelerating the digitization of biodiversity research specimens through online public participation [J]. *BioScience*, 65(4): 383–396.
- ELSEN PR, MONAHAN WB, MERENLENDER AM, 2020. Topography and human pressure in mountain ranges alter expected species responses to climate change [J]. *Nat Commun*, 11(1): 1–10.
- FAN JY, ZHAO NX, LI M, et al., 2018. What are the best predictors for invasive potential of weeds? Transferability evaluations of model predictions based on diverse environmental data sets for *Flaveria bidentis* [J]. *Weed Res*, 58(2): 141–149.
- FRANCISCO-ORTEGA J, SANTOS-GUERRA A, CARINE MA, et al., 2008. Plant hunting in Macaronesia by Francis Masson: the plants sent to Linnaeus and Linnaeus filius [J]. *Bot J Linn Soc*, 157(3): 393–428.
- FUENTES N, UGARTE E, KÜHN I, et al., 2008. Alien plants in Chile: inferring invasion periods from herbarium records [J]. *Biol Invasions*, 10(5): 649–657.
- FUNK VA, 2004. 100 uses for an herbarium (well at least 72) [J]. *Am Soc Plant Taxon Newsl*, 17(2): 17–19.
- FUNK VA, HOCH PC, PRATHER AL, et al., 2005. The importance of vouchers [J]. *Taxon*, 54(1): 127–129.
- GEHAN MA, KELLOGG EA, 2017. High-throughput phenotyping [J]. *Am J Bot*, 104(4): 505–508.
- GÓMEZ-BELLVER C, IBÁÑEZ N, LÓPEZ-PUJOL J, et al., 2020. How photographs can be a complement of herbarium vouchers: a proposal of standardization [J]. *Taxon*, 68(6): 1321–1326.
- GOODWIN ZA, HARRIS DJ, FILER D, et al., 2015. Widespread mistaken identity in tropical plant collections [J]. *Curr Biol*, 25(22): R1066–R1067.
- HARRIS KM, MARSICO TD, 2017. Digitizing specimens in a small herbarium: a viable workflow for collections working with limited resources [J]. *Appl Plant Sci*, 5(4): 1600125.
- HART R, SALICK J, RANJITKAR S, et al., 2014. Herbarium specimens show contrasting phenological responses to Himalayan climate [J]. *Proc Natl Acad Sci USA*, 111(29): 10615–10619.
- HASTON E, CUBEY R, PULLAN M, et al., 2012. Developing integrated workflows for the digitisation of herbarium specimens using a modular and scalable approach [J]. *Zookeys*, (209): 93–102.
- HEBERLING JM, ISAAC BL, 2017. Herbarium specimens as exaptations: new uses for old collections [J]. *Am J Bot*, 104(7): 963–965.
- HEBERLING JM, PRATHER LA, TONSOR SJ, 2019. The changing uses of herbarium data in an era of global change: an overview using automated content analysis [J]. *BioScience*, 69(10): 812–822.
- HU JH, 2017. Decreasing desired opportunity for energy supply of a globally acclaimed biofuel crop in a changing climate [J]. *Renew Sustain Energy Rev*, 76: 857–864.
- HUANG JH, CHEN JH, YING JS, et al., 2011. Features and distribution patterns of Chinese endemic seed plant species [J]. *J Syst Evol*, 49(2): 81–94.
- HUANG JH, HUANG JH, LIU CR, et al., 2016. Diversity hotspots and conservation gaps for the Chinese endemic seed flora [J]. *Biol Conserv*, 198: 104–112.
- IMBACH P, FUNG E, HANNAH L, et al., 2017. Coupling of pollination services and coffee suitability under climate change [J]. *Proc Natl Acad Sci USA*, 114(39): 10438–10442.
- IPCC, 2018. Global warming of 1.5 °C. An IPCC special report on the impacts of global warming of 1.5 °C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty [M/OL]. https://www.ipcc.ch/site/assets/uploads/sites/2/2019/06/SR15_Full_Report_High_Res.pdf.
- IVANOVA NV, SHASHKOV MP, 2017. Biodiversity databases in Russia: towards a national portal [J]. *Arct Sci*, 3(3): 560–576.
- JACA TP, BOATWRIGHT JS, MOTEETEE AN, 2018. Taxonomic studies of the genus *Rhynchosia* Lour. (Phaseoleae, Fabaceae) in South Africa: a review of section *Chrysoscias* [J]. *S Afr J Bot*, 117: 119–133.
- JAISWAL Y, LIANG ZT, ZHAO ZZ, 2016. Botanical drugs in Ayurveda and Traditional Chinese Medicine [J]. *J Ethnopharmacol*, 194: 245–259.
- JONES CA, DAEHLER CC, 2018. Herbarium specimens can reveal impacts of climate change on plant phenology; a review of methods and applications [J]. *Peer J*, 6: e4576.
- JUMP AS, HUANG TJ, CHOU CH, 2012. Rapid altitudinal migration of mountain plants in Taiwan and its implications for high altitude biodiversity [J]. *Ecography*, 35(3): 204–210.

- JUKONIENĖ I, RASIMAVIČIUS M, RIČKIENĖ A, et al., 2018. S. B. Gorski's bryological collection in the Herbarium of Vilnius University [J]. *Acta Soc Bot Pol*, 87(3): 3588.
- KIER G, KREFT H, LEE TM, et al., 2009. A global assessment of endemism and species richness across island and mainland regions [J]. *Proc Natl Acad Sci USA*, 106(23): 9322–9327.
- LANE MA, 1996. Roles of natural history collections [J]. *Ann Mol Bot Gard*, 83(4): 536–545.
- LANG PLM, WILLEMS FM, SCHEEPENS JF, et al., 2019. Using herbaria to study global environmental change [J]. *New Phytol*, 221(1): 110–122.
- LAVOIE C, 2013. Biological collections in an ever changing world: herbaria as tools for biogeographical and environmental studies [J]. *Perspect Plant Ecol*, 15(1): 68–76.
- LEES DC, LACK HW, ROUGERIE R, et al., 2011. Tracking origins of invasive herbivores through herbaria and archival DNA: the case of the horse-chestnut leaf miner [J]. *Front Ecol Environ*, 9(6): 322–328.
- LEWIS SL, MASLIN MA, 2015. Defining the Anthropocene [J]. *Nature*, 519(7542): 171–180.
- LIENERT J, FISCHER M, DIEMER M, 2002. Local extinctions of the wetland specialist *Swertia perennis* L. (Gentianaceae) in Switzerland: a revisitation study based on herbarium records [J]. *Biol Conserv*, 103(1): 65–76.
- LIU J, MÖLLER M, PROVAN J, et al., 2013. Geological and ecological factors drive cryptic speciation of yews in a biodiversity hotspot [J]. *New Phytol*, 199(4): 1093–1108.
- LIU J, MILNE RI, MÖLLER M, et al., 2018. Integrating a comprehensive DNA barcode reference library with a global map of yews (*Taxus* L.) for forensic identification [J]. *Mol Ecol Resour*, 18(5): 1115–1131.
- LOHONYA K, LIVERMORE L, PENN MG, 2020. Georeferencing the natural history museum's Chinese type collection: of plateaus, pagodas and plants [J]. *Biodivers Data J*, 8: E50503.
- LUGHADHA EN, GOVAERTS R, BELYAEVA I, et al., 2016. Counting counts: revised estimates of numbers of accepted species of flowering plants, seed plants, vascular plants and land plants with a review of other recent estimates [J]. *Phytotaxa*, 272(1): 82–88.
- LULEKAL E, ASFAW Z, KELBESSA E, et al., 2012. Linking ethnobotany, herbaria and flora to conservation: the case of four angiosperm families at the national herbarium of Ethiopia [J]. *J E Afr Nat Hist*, 101(1): 99–125.
- MA JS, 2013. The checklist of the Chinese invasive plants [J]. *Biodivers Sci*, 21(5): 635. [马金双, 2013. 中国入侵植物名录 [J]. *生物多样性*, 21(5): 635.]
- MA KP, 2014. Rapid development of biodiversity informatics in China [J]. *Biodivers Sci*, 22(3): 251–252. [马克平, 2014. 生物多样性信息学在中国快速发展 [J]. *生物多样性*, 22(3): 251–252.]
- MCKAIN MR, JOHNSON MG, URIBE-CONVERS S, et al., 2018. Practical considerations for plant phylogenomics [J]. *Appl Plant Sci*, 6(3): e1038.
- MEINEKE EK, DAVIES TJ, DARU BH, et al., 2019a. Biological collections for understanding biodiversity in the Anthropocene [J]. *Philos Trans R Soc B*, 374(1763): 20170386.
- MEINEKE EK, CLASSEN AT, SANDERS NJ, et al., 2019b. Herbarium specimens reveal increasing herbivory over the past century [J]. *J Ecol*, 107(1): 105–117.
- MEINEKE EK, DAVIES TJ, 2019. Museum specimens provide novel insights into changing plant-herbivore interactions [J]. *Philos Trans R Soc B*, 374(1763): 20170393.
- MEINEKE EK, DAVIS CC, DAVIES TJ, 2018. The unrealized potential of herbaria for global change biology [J]. *Ecol Monogr*, 88(4): 505–525.
- MEROW C, ALLEN JM, AIELLO-LAMMENS M, et al., 2016. Improving niche and range estimates with Maxent and point process models by integrating spatially explicit information [J]. *Global Ecol Biogeogr*, 25(8): 1022–1036.
- MEYER C, WEIGELT P, KREFT H, 2016. Multidimensional biases, gaps and uncertainties in global plant occurrence information [J]. *Ecol Lett*, 19(8): 992–1006.
- MILLENNIUM ECOSYSTEM ASSESSMENT, 2005. Ecosystems and human well-being: biodiversity synthesis [M]. Washington, D.C.: World Resources Ins: 53.
- MILLER-RUSHING AJ, PRIMACK RB, PRIMACK D, et al., 2006. Photographs and herbarium specimens as tools to document phenological changes in response to global warming [J]. *Am J Bot*, 93(11): 1667–1674.
- MONFILS AK, POWERS KE, MARSHALL CJ, et al., 2017. Natural history collections: teaching about biodiversity across time, space, and digital platforms [J]. *SE Nat*, 16(sp10): 47–57.
- MOSA KA, GAIROLA S, JAMDAD R, et al., 2019. The promise of molecular and genomic techniques for biodiversity research and DNA barcoding of the Arabian peninsula flora [J]. *Front Plant Sci*, 9: 1929.
- NATHAN R, MULLER-LANDAU HC, 2000. Spatial patterns of seed dispersal, their determinants and consequences for recruitment [J]. *Trends Ecol Evol*, 15(7): 278–285.
- NELSON G, ELLIS S, 2018. The history and impact of digitization and digital data mobilization on biodiversity research [J]. *Philos Trans R Soc B*, 374(1763): 20170391.
- NUALART N, IBÁÑEZ N, SORIANO I, et al., 2017. Assessing the relevance of herbarium collections as tools for conservation biology [J]. *Bot Rev*, 83(3): 303–325.
- OBERLIES NH, KNOWLES SL, AMRINE CSM, et al., 2019. Droplet probe: coupling chromatography to the *in situ* evaluation of the chemistry of nature [J]. *Nat Prod Rep*, 36(7): 944–959.

- PANCHEN ZA, PRIMACK RB, ANISKO T, et al., 2012. Herbarium specimens, photographs, and field observations show philadelphia area plants are responding to climate change [J]. *Am J Bot*, 99(4): 751–756.
- PARK IW, 2012. Digital herbarium archives as a spatially extensive, taxonomically discriminate phenological record; a comparison to MODIS satellite imagery [J]. *Int J Biometeorol*, 56(5): 1179–1182.
- PHILLIPS SJ, DUDIK M, ELITH J, et al., 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data [J]. *Ecol Appl*, 19(1): 181–197.
- PRIMACK D, IMBRES C, PRIMACK RB, et al., 2004. Herbarium specimens demonstrate earlier flowering times in response to warming in Boston [J]. *Am J Bot*, 91(8): 1260–1264.
- QIN HN, YANG ZR, 2011. The past, present and future of herbaria [J]. *Life World*, 263(9): 4–8. [覃海宁, 杨志荣, 2011. 标本馆的前世今生和未来 [J]. *生命世界*, 263(9): 4–8.]
- QIN HN, ZHAO LN, 2017. Evaluating the threat status of higher plants in China [J]. *Biodivers Sci*, 25(7): 689–695. [覃海宁, 赵莉娜, 2017. 中国高等植物濒危状况评估 [J]. *生物多样性*, 25(7): 689–695.]
- RANDS MRW, ADAMS WM, BENNUN L, et al., 2010. Biodiversity conservation: challenges beyond 2010 [J]. *Science*, 329(5997): 1298–1303.
- RICHTER R, DULLINGER S, ESSL F, et al., 2013. How to account for habitat suitability in weed management programmes? [J] *Biol Invasions*, 15(3): 657–669.
- RISTAINO JB, GROVES CT, PARRA GR, 2001. PCR amplification of the Irish potato famine pathogen from historic specimens [J]. *Nature*, 411(6838): 695–697.
- SOLTIS DE, SOLTIS PS, 2016. Mobilizing and integrating big data in studies of spatial and phylogenetic patterns of biodiversity [J]. *Plant Divers*, 38(6): 264–270.
- SOSEF MSM, DAUBY G, BLACH-OVERGAARD A, et al., 2017. Exploring the floristic diversity of tropical Africa [J]. *BMC Biol*, 15: 15.
- SOUZA ENF, HAWKINS JA, 2017. Comparison of herbarium label data and published medicinal use: herbaria as an underutilized source of ethnobotanical information [J]. *Econ Bot*, 71(1): 1–12.
- THIERS BM, 2021. The world's herbaria 2020: A summary report based on data from Index Herbariorum [J/OL]. [2021-05-07]. <http://sweetgum.nybg.org/science/ih/annual-report/>.
- TILMAN D, 1999. The ecological consequences of changes in biodiversity: a search for general principles [J]. *Ecology*, 80(5): 1455–1474.
- TWYFORD AD, NESS RW, 2017. Strategies for complete plastid genome sequencing [J]. *Mol Ecol Resour*, 17(5): 858–868.
- VAN KLEUNEN M, DAWSON W, ESSL F, et al., 2015. Global exchange and accumulation of non-native plants [J]. *Nature*, 525(7567): 100–103.
- VATTAKAVEN T, GEORGE RM, BALASUBRAMANIAN D, et al., 2016. India biodiversity portal: an integrated, interactive and participatory biodiversity informatics platform [J]. *Biodivers Data J*, 4: e10279.
- VON REIS S, 1962. Herbaria: sources of medicinal folklore [J]. *Econ Bot*, 16(4): 283–287.
- WANG CJ, WAN JZ, 2020. Assessing the habitat suitability of 10 serious weed species in global croplands [J]. *Glob Ecol Conserv*, 23(2020): e01142.
- WIECZOREK J, BLOOM D, GURALNICK R, et al., 2012. Darwin core: an evolving community-developed biodiversity data standard [J]. *PLoS ONE*, 7(1): e29715.
- WILLIS CG, ELLWOOD ER, PRIMACK RB, et al., 2017a. Old plants, new tricks: phenological research using herbarium specimens [J]. *Trends Ecol Evol*, 32(7): 531–546.
- WILLIS CG, LAW E, WILLIAMS AC, et al., 2017b. CrowdCurio: an online crowdsourcing platform to facilitate climate change studies using herbarium specimens [J]. *New Phytol*, 215(1): 479–488.
- WOLF A, ZIMMERMAN NB, ANDEREGG WRL, et al., 2016. Altitudinal shifts of the native and introduced flora of California in the context of 20th-century warming [J]. *Glob Ecol Biogeogr*, 25(4): 418–429.
- XU XT, DIMITROV D, SHRESTHA N, et al., 2019. A consistent species richness-climate relationship for oaks across the northern hemisphere [J]. *Glob Ecol Biogeogr*, 28(8): 1051–1066.
- XU ZH, LIU SY, ZHAO Y, et al., 2020. Evaluation of the identification ability of eight commonly used plant identification application softwares in China [J]. *Biodivers Sci*, 28(4): 524–533. [许展慧, 刘诗尧, 赵莹, 等, 2020. 国内 8 款常用植物识别软件的识别能力评价 [J]. *生物多样性*, 28(4): 524–533.]
- YOUNIS S, WEILAND C, HOEHNDORF R, et al., 2018. Taxon and trait recognition from digitized herbarium specimens using deep convolutional neural networks [J]. *Bot Lett*, 165(3/4): 377–383.
- ZHANG ZJ, HE JS, LI JS, et al., 2015. Distribution and conservation of threatened plants in China [J]. *Biol Conserv*, 192(2015): 454–460.
- ZHU L, SUN OJ, SANG WG, et al., 2007. Predicting the spatial distribution of an invasive plant species (*Eupatorium adenophorum*) in China [J]. *Landscape Ecol*, 22(8): 1143–1154.