

中国黄花蒿中青蒿素含量空间分布特征分析

张小波^{1,2}, 郭兰萍¹, 邱智东², 曲晓波^{2*}, 王慧¹, 景志贤¹, 黄璐琦^{1*}

(1. 中国中医科学院 中药资源中心 道地药材国家重点实验室培育基地, 北京 100700;

2. 长春中医药大学 药学院, 吉林 长春 130117)

[摘要] 道地药材指经过中医临床长期应用优选出来的, 产在特定地域, 受到特定生产加工方式影响, 较其他地区所产同种药材品质佳、疗效好且质量稳定, 具有较高的知名度的药材。道地药材是我国几千年悠久文明史、中医中药发展史形成的特有概念。根据道地药材的定义, 可以看出不同地域之间的同种药材在质量和疗效等方面存在一定的差异性, 在特定区域还存在一定的相似性。该文应用空间统计分析方法, 对中国各地黄花蒿中青蒿素含量空间分布特征进行分析。①采用“*I* 系数”对各地青蒿素含量的空间自相关性进行分析, 结果显示, 各省青蒿素含量存在着显著的、正的空间自相关; 各省青蒿素含量高低并非随机分布, 而是具有明显的空间聚集特征。②采用“*G* 统计量”对各地青蒿素含量的空间差异性进行分析, 结果显示, 总体上青蒿素含量高的区域分布在西南地区, 在空间上呈现连片分布, 青蒿素含量趋于高值空间聚集; 天津市和辽宁省青蒿素含量较低, 青蒿素含量趋于低值空间聚集; 其他省份青蒿素含量属于高值和低值过渡区域。③基于各省青蒿素含量一个指标, 采用“Moran 散点图”对各区域青蒿素含量空间关联关系进行分析, 结果显示: 广西、重庆等 9 个省市的青蒿素含量属于高值聚集区。④采用“地理探测器”基于采样点的青蒿素含量和生态环境因子指标, 同时对连续型的定量数据、非连续型的定性数据进行分析, 结果显示: 日照、气温、降水量是影响青蒿素含量的主要因素。

[关键词] 空间统计分析; 黄花蒿; 青蒿素; 空间分布特征

Analysis of spatial distribution of artemisinin in *Artemisia annua* in China

ZHANG Xiao-bo^{1,2}, GUO Lan-ping¹, QIU Zhi-dong², QU Xiao-bo^{2*}, WANG Hui¹,
JING Zhi-xian¹, HUANG Lu-qi^{1*}

(1. State Key Laboratory Breeding Base of Dao-di Herbs, National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China;

2. College of Pharmacy, Changchun University of Chinese Medicine, Changchun 130117, China)

[Abstract] Geo-herbals are preferred by long-term clinical application of Chinese medicine, they are produced in a specific area and affected by the specific production and processing methods. Geo-herbals have a high reputation, compared with other regions produced by the same kind of herbs with good quality, efficacy and stable quality characteristics. Geo-herbal is a unique concept that has been formed by the history of Chinese civilization for thousands of years and the history of Chinese medicine. According to the definition of geo-herbal, it can be seen that there are some differences in the quality and efficacy of the same kind of medicinal herbs in different regions, and there is some similarity in the specific area. In this study, the spatial distribution characteristics of artemisinin in *Artemisia annua* were analyzed by spatial statistical analysis. ①The spatial autocorrelation of artemisinin content was analyzed by "*I* coefficient". The results showed that the content of artemisinin in each province was significant and positive spatial autocorrelation. The content of artemisinin in each province was not randomly distributed, but with a clear spatial aggregation characteristics. ②The spatial variability of artemisinin content was analyzed by "*G* statistic". The results showed that the area with high content of artemisinin was distributed in the southwest region, and the content of artemisinin was concentrated in the space, which tends to accumulate at high value. The content of

[收稿日期] 2017-03-13

[基金项目] 中央本级重大增减支项目(2060302); 科技基础性工作专项(2013FY114500); 中医药全国性专项(ZZYK2013-科技司 A-004)

[通信作者] * 黄璐琦, Tel: (010) 84044340, E-mail: huangluqi01@126.com; * 曲晓波, E-mail: quxiaobo0504@hotmail.com

[作者简介] 张小波, Tel: (010) 64081790, E-mail: jack110007@163.com

artemisinin in Tianjin and Liaoning province is low, and it tends to accumulate at low value, while the content of artemisinin in other provinces is the transition zone of high and low value. ③Based on the index of artemisinin in each province, the spatial correlation of artemisinin content in each region was analyzed by "moran scatter plot". The results showed that the content of artemisinin in nine provinces such as Guangxi and Chongqing belonged to high value gathering area. ④Based on the artemisinin content and ecological environment factor of sampling point, the "geophysical detector" was used to analyze the continuous quantitative data and discontinuous qualitative data. The results showed that the sunshine, temperature and precipitation were the main factors affecting the content of artemisinin. [Key words] spatial statistical analysis; *Artemisia annua*; artemisinin; spatial distribution characteristics

空间分异是地理现象的基本特点之一,随着全球定位系统等空间信息技术的发展,无论是精细或者大范围的调查研究,还是大数据分析结果,中药资源在空间分布和质量等方面均具有较大的空间差异性。中药青蒿为菊科蒿属植物黄花蒿 *Artemisia annua* L. 的干燥地上部分^[1],具有悠久的历史。黄花蒿在我国广泛分布,我国各地所产青蒿中青蒿素含量差异较大,北部高纬度地区青蒿中青蒿素含量较低,南部青蒿中青蒿素含量较高^[2]。植物生长既受区域内的气候、地形等因素影响,同时也受区域内土地利用类型、植被类型的影响,由于经典统计学分析方法,只对空间连续的定量化变量进行分析研究,对于定性数据无法进行统计分析。目前针对气候、地形等方面对青蒿素含量的影响因子进行分析^[2-3],这些研究多偏重于经典统计学的方法,主要使用定量化数据。黄花蒿为广布种分布区域范围大,一般调查工作难以获取所有区域的实际数据,存在基础数据不完整或数据缺失的问题,研究所需数据和实际获取的数据在空间粒度不一致等问题。因此,目前关于非连续型数值变量方面的指标对青蒿素含量影响研究成果较少。

地理空间信息与一般信息从数据结构来讲是不同的,前者是时空坐标数据;从数据分析与处理来看,空间数据是相互关联的,时间序列空间数据具有不可重复性,空间信息维的加入使数据量大大丰富,能够揭示数据表面背后的空间格局机制,在优化运筹领域加入空间维能大大优化结果。空间统计分析是通过空间位置建立数据间的统计关系,进而明确与地理位置相关数据的空间依赖、空间关联、空间系相关性。对于那些与空间数据的结构性和随机性,或空间相关性和依赖性,或空间格局与变异有关的空间现象,均可应用空间分析方法进行研究。空间统计学与经典统计学的共同之处在于它们都是在大量采样的基础上,通过对样本属性值的频率分布或

均值、方差关系及其相应规则的分析,确定其空间分布格局与相关关系。但空间统计学区别于经典统计学的最大特点是空间统计学研究的变量在采样前是经典随机变量,而采样后是特定的空间三维函数值,且表现出空间结构性,具有某种程度的空间自相关性。空间统计学既考虑到样本值的大小,又重视样本空间位置及样本间的距离,弥补了经典统计学无法对空间方位进行分析的缺陷,空间统计能够较精确地反映变量的空间的分布与变化情况。

本研究基于文献和实地调查结果,运用探索性空间数据分析、趋势面分析、空间变异函数等空间统计分析技术^[4],对中国各省黄花蒿中青蒿素含量进行研究,分析中国各地黄花蒿中青蒿素含量的空间差异特性,为我国黄花蒿种植及与青蒿产业发展相关的政策规划制定提供参考依据。

1 材料

1.1 文献数据

遵循“数据已在核心级期刊公开发表,野生青蒿采集时间在7—9月,青蒿素含量测定部位为叶,具有经纬度或县域等位置信息”的原则,通过查阅文献共获得13个省202个采样点青蒿中青蒿素含量数据。

1.2 实测数据

在全国19个省份选择250个采样点,采样过程中用GPS测定采样点的经度、纬度等位置信息。

2 方法

2.1 我国黄花蒿中青蒿素含量空间自相关性分析

通过ArcGIS,根据采样点的经纬度,生成采样点分布图。通过ArcGIS根据采样点的经纬度,基于矢量的行政区划数据,对采样点所在的省和县进行分类汇总。使用R语言,分别计算各省和县域内采样点青蒿素含量的均值。

根据地理学第一定律“任何东西与别的东西之间都是相关的,但近处的东西比远处的东西相关

性更强”,即距离越近的空间单元越相似,距离越远的空间单元越不相似^[5]。青蒿素含量空间分布总体特征,采用全局空间自相关分析方法。全局空间自相关主要是对属性在整个区域空间上特征的描述,反映了观测变量在整个研究区域内空间相关性的整体趋势。主要用 Moran's I 系数表示,其取值范围在 $-1 \sim 1$ 。当 Moran's $I > 0$,代表空间事物在空间分布上呈正相关性,空间上呈聚集分布。当 Moran's $I < 0$,代表空间事物在空间分布上呈负相关性,空间上呈离散分布。当 Moran's $I = 0$ (当 Moran's I 接近 0 时),代表空间事物在空间分布上不具有相关性,空间上呈随机分布。

2.2 我国黄花蒿中青蒿素含量空间差异性分析

采用局域自相关分析方法,分析研究县域间中药资源种类差异性空间分布特征。主要有 LISA 指数和 G_i^* 指数表示。LISA 指数是全局空间自相关指数的分解形式,反映局域空间相关性的显著性水平;还可识别空间异常值;可用于具体度量每个区域与周边区域之间的局部空间关联和空间差异程度。 G_i^* 指数用于识别属性要素高值或低值的空间聚类情况,如果 G_i^* 指数值为正且显著,表明该县域周围中药资源种类丰富,属于高值空间集聚(热点区);如果 G_i^* 指数值为负且显著,表明该县域周围中药资源种类匮乏,属于低值空间集聚(冷点区)。

2.3 我国黄花蒿中青蒿素含量差异性分布影响因素分析

地理探测器^[6]是探测空间分异性,以及揭示其背后驱动因子的一种新的统计学方法,是探测和利用空间分异性的工具。基本思想是:假设研究区分为若干子区域,如果子区域的方差之和小于区域总方差,则存在空间分异性;如果两变量的空间分布趋于一致,则两者存在统计关联性。地理探测器 q -统计量,可用以度量空间分异性、探测解释因子、分析变量之间交互关系,已经在自然和社会科学多领域应用。地理探测器的优势:地理探测器既可以探测数值型数据,也可以探测定性数据,分析定性数据(类型量)是地理探测器的一大优势。而对于顺序量、比值量或间隔量,只要进行适当的离散化,也可以利用地理探测器对其进行统计分析。探测两因子交互作用于因变量,这是地理探测器的另一个独特优势。地理探测器通过分别计算和比较各单因子 q 及两因子叠加后的 q ,可以判断两因子是否存在交互

作用,以及交互作用的强弱、方向、线性还是非线性等。两因子叠加既包括相乘关系,也包括其他关系,只要有关系,就能检验出来。地理探测器已被运用于从自然到社会十分广泛的领域,在实际应用中,地理探测器主要用来分析各种现象的影响因子和驱动力,以及多因子交互作用。应用地理探测器进行因子和驱动力分析时,因变量是数值量,而自变量是类型量。

3 结果

3.1 我国黄花蒿中青蒿素含量空间自相关性分析

应用 ArcGIS 根据收集到各个样点的经纬度,基于矢量的行政区划数据,生成采样点分布图,结果见图 1。除海南、福建和浙江等少部分省份没有数据外,中国 80% 以上的省份均有采样点数据。

使用 R 语言,分别计算省和县域内各采样点青蒿素含量的均值。根据中国各省(区、市)青蒿素含量的平均值,应用 GIS 制图功能,生成 26 个省青蒿素含量的柱状图,结果见图 2。从图 2 中可以看出各地青蒿素含量存在一定的差异。

选取中国各省青蒿素含量,计算全局空间自相关 Moran 指数 I ,并计算其检验的标准化统计量 Z ,结果见图 3。其中, Moran 指数值 $I = 0.65$, 标准化统计量 $Z = 126.67$, $P = 0$,置信度为 99%;在正态分布假设条件下, Moran 指数检验结果高度显著。表明:各省青蒿素含量存在着显著的、正的空间自相关;各省青蒿素含量高低并非随机分布,而是具有明显的空间聚集特征。

趋势图分析结果:用 GIS 软件绘制各省青蒿素含量的趋势图,见图 4。从整体研究区域来看,自北向南有逐渐增加的趋势。总体上西南部地区青蒿中的青蒿素较高,北部地区青蒿中的青蒿素较低。

3.2 我国黄花蒿中青蒿素含量空间差异性分析

3.2.1 空间自相关分析 总体空间格局从空间自相关分析整体上表明了全国青蒿素含量分布态势,并不代表每个地区都是如此,为了进一步分析各省青蒿素含量的空间相关性,需结合局部空间自相关模型进一步分析。通过计算 LISA 指数,通过 5% 显著水平检验的地区见图 5。其中,全国青蒿素含量的空间正相关性较强,总体上青蒿素含量高的省域仍然占少数。

3.2.2 G_i^* 指数分析结果 根据到各省青蒿素含量、用 GIS 软件计算 G_i^* 统计量,并计算其检验的标准化统计量 Z 。用自然断裂法(Jenks)进一步可



图 1 青蒿采样点分布 [审图号: GS2016(1580) 号]

Fig. 1 The distribution of sampling points of *Artemisia annua* [trail No. GS2016(1580)]

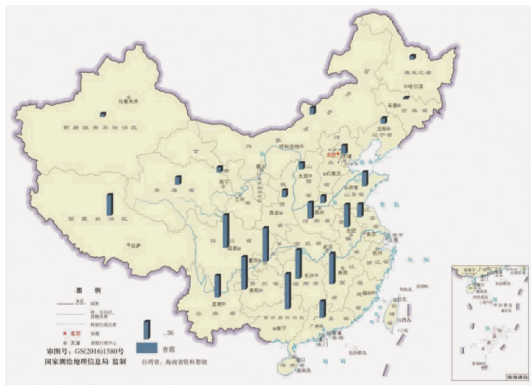


图 2 各省青蒿素含量 [审图号: GS2016(1580) 号]

Fig. 2 Artemisinin content of each province [trail No. GS2016(1580)]

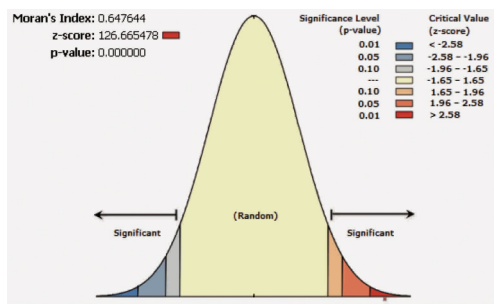


图 3 Moran's I 指数图

Fig. 3 Moran's I index

对数据进行可视化处理,将数值由低到高划分为 3 类,分别为:冷点地区、温点地区、热点地区,结果见图 6。总体上热点省份占总数的 25%;这些省的 Z 在 0.05 的显著性水平下显著,青蒿素含量高的区域在空间上相连成片分布,从统计学意义上说,其青蒿素含量趋于高值空间聚集;总体上分布在西南地区。

• 4280 •

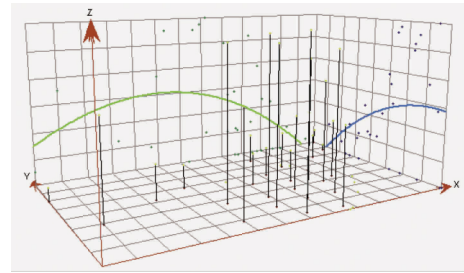


图 4 中国各省青蒿素含量高低分布趋势图

Fig. 4 Distribution trend of artemisinin in different provinces of China



图 5 青蒿素含量空间自相关显著性检验 [审图号: GS2016(1580) 号]

Fig. 5 Spatial autocorrelation significance of artemisinin content [trail No. GS2016(1580)]

冷点省份占总数的 6.2%;这些县的 Z 在 0.05 的显著性水平下显著,青蒿素含量较为低,从统计学意义上说,青蒿素含量趋于低值空间聚集;总体上冷点地区分布在天津市和辽宁省。温点省份个数最多,温点省域占总数的 68.8%,青蒿素含量属于高值和低值过渡区域。

3.3 我国黄花蒿中青蒿素含量空间差异性影响因素分析

3.3.1 Moran 散点图 运用 R 语言绘制 Moran 散点图,进一步分析全国各省青蒿素的局部空间相关性,结果见图 7。其中,第一象限(HH)代表高值区域被高值邻居包围,为青蒿素含量最高的区域;第二象限(LH)代表低值区域被高值邻居包围,第三象限(LL)代表低值区域被低值邻居包围,第四象限(HL)代表高值区被低值邻居包围。结果显示,广西、重庆等 9 个省市(自治区)的青蒿素含量属于高值聚集区;辽宁和天津属于低值聚集区。



图6 青蒿素含量 G_i^* 指数分析 [审图号: GS2016(1580) 号]
Fig.6 G_i^* index analysis of artemisinin content [trail No. GS2016(1580)]

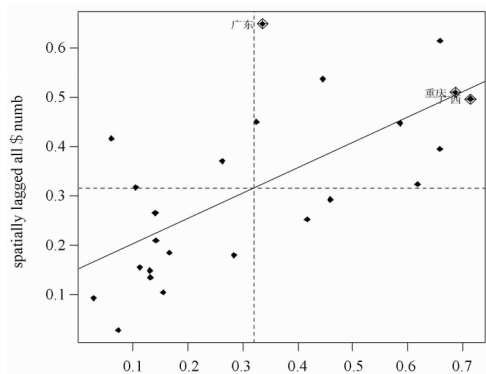


图7 各省青蒿素含量空间自相关 Moran 散点图
Fig.7 The spatial self-correlation Moran scatter plot of artemisinin content in each province

3.3.2 青蒿素含量空间分布差异影响因子 基于全国各采样点的生态环境数据,包括7个指标:气温、降水量、日照、相对湿度、酸碱度、坡度、海拔。应用地理探测器,对影响青蒿素含量空间差异性分布的因子进行分析结果见表1。

表1 青蒿素含量影响因素探测

Table 1 Detection of factors affecting artemisinin content

项目	气温	降水量	日照	相对湿度	酸碱度	坡度	海拔
气温	0.49						
降水量	0.53	0.43					
日照	0.59	0.56	0.52				
相对湿度	0.52	0.45	0.58	0.40			
酸碱度	0.58	0.52	0.58	0.51	0.24		
坡度	0.56	0.47	0.58	0.46	0.33	0.06	
海拔	0.58	0.47	0.58	0.46	0.41	0.23	0.12

由地理探测器计算得到的各种风险因子对于青蒿素含量的相对重要性和具体 q 值如下:日照(0.52) > 气温(0.49) > 降水量(0.43) > 相对湿度(0.40) > 酸碱度(0.24) > 海拔(0.12) > 坡度(0.06)。这一结果显示日照、气温、降水量是青蒿素含量的主要因素。

对因子之间交互作用的分析显示,日照、气温的交互作用解释了青蒿素含量的59%,酸碱度、海拔与气温的交互作用都解释了58%。日照与大部分变量的解释力也为58%。需要特别指出的是,这一发现表明,这些风险因子的交互作用均大于单一因子的单独作用,由此可见不同环境因子之间交互作用的重要性。

4 结论和讨论

道地药材指经过中医临床长期应用优选出来的,产在特定地域,受到特定生产加工方式影响,其他地区所产同种药材品质佳、疗效好且质量稳定,具有较高的知名度的药材^[7]。道地药材是我国几千年悠久文明史、中医中药发展史形成的特有概念,根据道地药材的定义,可以看出不同地域之间的同种药材在质量和疗效等方面存在一定的差异性,在特定区域还存在一定的相似性。我国各地青蒿中青蒿素含量差异较大,通过对各地青蒿素含量的空间自相关分析研究发现,各省青蒿素含量存在着显著的、正的空间自相关;各省青蒿素含量高低并非随机分布,而是具有明显的空间聚集特征。总体上青蒿素含量高的区域分布在西南地区,在空间上呈现连片分布,青蒿素含量趋于高值空间聚集;天津市和辽宁省青蒿素含量较低,青蒿素含量趋于低值空间聚集;其他省份青蒿素含量属于高值和低值过渡区域。[参考文献]

- [1] 中国药典. 一部[S]. 2015: 198.
- [2] 张小波, 郭兰萍, 黄璐琦. 我国黄花蒿中青蒿素含量的气候适宜性等级划分[J]. 药学报 2011, 46 (4): 472.
- [3] 张小波, 王利红, 郭兰萍, 等. 广西地形对青蒿中青蒿素含量的影响分析[J]. 生态学报 2009 29(2): 688.
- [4] 王劲峰, 李连发, 葛咏, 等. 地理信息空间分析的理论体系探讨[J]. 地理学报 2000(1): 92.
- [5] 王劲峰. 空间分析[M]. 北京: 科学出版社 2006.
- [6] 王劲峰, 徐成东. 地理探测器: 原理与展望[J]. 地理学报, 2017(1): 116.
- [7] 张小波, 赵宇平, 黄晓巍, 等. 青蒿道地药材研究综述[J]. 中国中药杂志 2016 41(11): 2015.

[责任编辑 吕冬梅]