中药研究

2 个中药品种常数及多种中药配对的精确分类鉴别

邹华彬

(山东大学化学与化工学院,济南,250100)

摘要 目的:基于双指标信息理论方程-共有遗传与变异信息理论方程最大信息量状态,提出 2 个鉴别生物复杂体系中药的品种常数:共有峰率 $P_g = 61\%$, $P_g = 70\%$ 作为中药品种的绝对定量鉴别标准。在双指标信息量变化相对误差 ±3% 的范围内,可以确定 $P_g \geqslant (61 \pm 3)\%$ 及 $P_g \geqslant (70 \pm 3)\%$ 作为 2 个优化判别中药品种的定量理论标准。方法:基于提出的 2 个定量理论标准,结合最大有效样本数优化法对多个中药样本进行合理模式识别。测试并分析了 4 种 42 个中药复方桂附地黄丸、知柏地黄丸、金匮肾气丸、明目地黄丸样品原药粉末的红外指纹图谱,对 4 种不同复方的 6 个配对进行了鉴别验证。结果:5 个中药复方配对遵从品种常数 $P_g = 61\%$,1 个复方配对桂附地黄丸、知柏地黄丸遵从品种常数 $P_g = 70\%$,样品的正确识别率为 94.2%,类的正确识别率 100%。结论:该法提供了鉴别分类中药的科学的绝对定量标准。

关键词 品种常数;定量鉴别;分类;信息理论;遗传变异;生物模式识别;指纹图谱;桂附地黄丸;金匮肾气丸;明目地黄丸;知柏地黄丸

Accurate Identification of Varieties of Constants and Multiple Pairs Between 2 Chinese Medicine Species

Zou Huabin

(*School of Chemistry and Chemical Engineering of Shandong University*, *Jinan* 250100, *China*)

Abstract **Objective**: The constants of 2 traditional Chinese medicine (TCM) herbs in the complex biological system were analyzed based on the maximum information states of Dual index information theory equation and common heredity and variation information theory. The common peak ratios of the 2 species constants were $P_g = 61\%$ and $P_g = 70\%$, which respectively corresponding to symmetry and asymmetry variation states, were used as 2 absolute quantitative criteria to identify complex biology systems-TCM. Considered the influences of many other factors on components and experiment processes, the really theoretical identification standards should be established $P_g \geqslant (61 \pm 3)\%$ and $P_g \geqslant (70 \pm 3)\%$, within the relative error ±3% of information value around the maximum information states. **Methods**: The infrared (IR) fingerprint spectra (FPS) of their powder in 4 out of 42 compounds including Guifu Dihuang Pills, Jinkui Shenqi Pills, Mingmu Dihuang Pills, Zhibai Dihuang Pills were tested and analyzed. A total of 6 pairs of 4 different compounds were tested. **Results**: The result showed that 5 compounds followed the species constant $P_g = 61\%$, and 1 of them, the pair of Guifu Dihuang Pills and Zhibai Dihuang Pills, obeyed the species constant $P_g = 70\%$. The correct identification ratio of samples was 95.2%, and that of species was 100%. **Conclusion**: The 2 species constants provide the scientifically and absolutely quantitative standards for identifying and classifying traditional Chinese medicines.

Key Words Species constant; Quantitative identification; Classification; Information theory; Heredity and variation; Biological pattern recognition; Fingerprint spectra; Guifu Dihuang Pills; Jinkui Shenqi Pills; Mingmu Dihuang Pills; Zhibai Dihuang Pills

中图分类号:R282 **文献标识码**:A **doi**:10.3969/j.issn.1673-7202.2018.12.050

植物药的品种决定其功效,其正确的鉴别及分类对于传统医学具有根本重要的意义。中药主要是植物药及动物药,其鉴别属于生物物种鉴别,但自古以来关于生物物种却没有一个明确的统一的定义及定性定量判别标准。目前有多达几十种生物物种定义[1-5],主要是根据各类宏观性状[6]、生物功能、地理分布、生态环境等多种性质进行的基于主观标准或推演的表象判别[3-5]。生物学物种的定义,缺乏基于实验科学及数学原理而建立的原理性理论,及基于这些原理的明确的判别物种的定量理论标准,生物及中药的鉴别完全依靠形态、尺寸、颜色、味道等多种性状经验知识[7]。即使引入现代数值分类学,其鉴别标准的建立也必须依赖于大量的经验知识及学习样本。物种的概念也是生物学中的核心概念,缺

基金项目:山东省中医药科技发展计划项目(2013-157)

作者简介:邹华彬(1960.10—),男,博士,副教授,硕士研究生导师,研究方向:中药生物复杂体系鉴别分类及质量评价数理理论方法,E-mail:huabinzou@126.com

乏物种的科学定义,生命科学大部分的内容将会成为没有联系的一盘散沙[8]。同理,缺乏中药品种的科学标准,中医药学也无法成为一门严谨的现代科学。中药作为一种生物复杂体系,是否也像物理学、化学体系那样,存在某些基于生物内禀规律决定的内在绝对的科学标准,即确定的常数标准,无需依赖经验知识对中药品种做出精确鉴别? 这是一个需要探索的重要科学课题。作者长期以来研究基于生物体系内禀特性及基本自然原理的中药鉴别及质量评价数理理论方法,在博士研究期间独立提出并建立了共有峰率及变异峰率双指标序列分析法[9-15],并被许多学者成功应用于植物药、植物、食品及动物等多个领域的鉴别分类,发表了 50 多篇研究论文,如文献[16-28]。以此为基础,作者独立提出了双指标等级序列个性化模式识别理论并与研究小组通过实验研究建立了该模式识别法[29-36],无需任何关于样品的经验知识、学习样本,依据相似尺度 $P_g \geqslant \bar{P}_g + xS$, $-3 \leqslant x \leqslant 3$,连续变化 x,确定每个样本的特征序列,再依据各个样本的核心特征序列构成的独立集合,结合最大有效样本数法[37]可以对中药进行精确分类鉴别。但是该法并没有确立关于植物药的品种鉴别绝对定量标准。作者亦根据生物学特性,建立了描述生物共有遗传及变异信息的理论方程-双指标信息理论方程[38]。根据该方程,生物体系存在两类内禀变异机制,对称变异机制及非对称变异机制。2 种机制的最大信息量所对应 2 个共有峰率(或共有成分),定义为 2 个常数,对称变异状态对应的常数适用于生物体系之间具有对称或近对称变异的状态,这已经被多种生物复杂体系所证明[37-39],采用对称变异常数 $P_g = 61\%$,正确鉴别了如下复杂生物体系:桂附地黄丸、金匮肾气丸无水乙醇提取物红外指纹图谱[38],明目地黄丸、知柏地黄丸、麦味地黄丸 3 种复方中药的粉末红外指纹图谱[34],补中益气丸、十全大补丸氯仿提取物的红外指纹图谱。而非对称变异状态对应的常数是否适合生物体系之间具有非对称变异的状态,未见研究报道。本研究测试并分析了 4 种 42 个中药复方桂附地黄丸、知柏地黄丸、金匮肾气丸、明目地黄丸样品原药粉末-全成分的红外指纹图谱,他们所包含的化学成分与采用不同溶剂的提取物具有显著的差异,因此该实验研究具有更好的代表性。检验了它们全部 6 个品种配对的定量鉴别标准。结果显示应用 $P_g \geqslant 61\%$,5 个不同复方配对可以得到精确的鉴别,只有 1 对配对桂附地黄丸与知柏地黄丸无法区分开。当采用 $P_g \geqslant (61 \pm 2)\%$ 及 $P_g \geqslant (70 \pm 2)\%$,可以将桂附地黄丸与知柏地黄丸精确分开,且其他 5 个配对同样得到准确鉴别,品种鉴别准确率 100%,42 个样品的正确识别率为 95.2%,初步证明自然界生物体系确实存在两类相似度常数。2 个常数可以初步定义为中药品种常数,定量反映生物的内禀品质特征。通过该常数可以对一些不同生物体系精确鉴别/模式识别。

亦采用双指标等级序列个性化模式识别法[29-30,33]对 4 种中药红外指纹图谱进行了分析,当统计相似尺度 $P_g \geqslant P_g + 1.1S \sim P_g \geqslant P_g + 1.3S$ 时,可以精确鉴别 4 种复方,且样品的正确识别率为 90.5%~95.2%。2 种理论方法的分析结果等价,皆可以将 4 种组成相似的生物复杂体系-中药复方进行精确鉴别分类,且相互验证。但 2 个中药品种常数具有确定的绝对定量标准,且具有鉴别判断的简易性、明确的物理意义和生物学意义。

1 仪器与试药

1.1 仪器 美国 NICOLET-5700-FT-IR 傅立叶变换红外光谱仪(光谱范围 4 000~400 cm^{-1},分辨力 4 cm^{-1});高速粉碎机,压片机,分析天平(METTLE TOLEDO),灵敏度 0.1 mg。

1.2 试剂 溴化钾(分析纯,国药集团化学试剂有限公司)。

1.3 分析样品 桂附地黄丸、明目地黄丸、金匮肾气丸及知柏地黄丸(皆为水丸或浓缩丸)样品如下。见表 1。

本论文中所研究的 4 种中药复方组成如下[40]:

桂附地黄丸:地黄、茯苓、牡丹皮、山药、山茱萸、泽泻、附子、肉桂。

金匮肾气丸:地黄、茯苓、牡丹皮、山药、山茱萸、泽泻、附子、牛膝、车前子、桂枝。

明目地黄丸:地黄、茯苓、牡丹皮、山药、山茱萸、泽泻、白芍、当归、枸杞、蔓藜、菊花、石决明。

知柏地黄丸:地黄、茯苓、牡丹皮、山药、山茱萸、泽泻、黄柏、知母。

由上述 4 种中药复方的组成分看,他们皆具有 6 种相同的草药,金匮肾气丸与桂附地黄丸具有 7 种相同的草药。4 种中药复方之间的化学成分具有很高的相似组成,且各自的化学成分具有非常高的复杂性。因此,准确鉴别它们具有极大的困难。

2 方法与结果

2.1 理论方法 2 个中药品种常数的理论推导根据生物体系的双指标信息理论方程-生物遗传与变异信息理论方程[38]。任意2个样品或任意样品2个

表 1 4 种中药复方样品

样品	名称	来源	生产日期	生产批号
S1	桂附地黄丸	合肥神鹿双鹤九华药业有限公司	2003. 04. 14	0303493
S2	桂附地黄丸	合肥神鹿双鹤九华药业有限公司	2003. 09. 01	0308413
S3	桂附地黄丸	河南宛西制药股份有限公司	2003. 12. 21	031204
S4	桂附地黄丸	河南宛西制药股份有限公司	2004. 07. 19	040702
S5	桂附地黄丸	河南宛西制药股份有限公司	2005. 02. 21	050203
S6	桂附地黄丸	河南宛西制药股份有限公司	2006. 01. 05	060101
S7	桂附地黄丸	芜湖张恒春商业有限公司	2004. 08. 12	20040802
S8	桂附地黄丸	芜湖张恒春商业有限公司	2004. 08. 12	20040802
S9	桂附地黄丸	安庆汇丰药业有限公司	2004. 05. 25	040520
S10	桂附地黄丸	苏州长甲药业有限公司	2005. 05. 17	0505171
S11	金匮肾气丸	北京同仁堂科技发展股份有限公司制药厂	2003. 03. 11	3030440
S12	金匮肾气丸	北京同仁堂科技发展股份有限公司制药厂	2003. 03. 11	3030440
S13	金匮肾气丸	北京同仁堂科技发展股份有限公司制药厂	2004. 03. 30	4030045
S14	金匮肾气丸	北京同仁堂科技发展股份有限公司制药厂	2004. 08. 02	4032954
S15	金匮肾气丸	北京同仁堂科技发展股份有限公司制药厂	2004. 08. 23	4032972
S16	金匮肾气丸	北京同仁堂科技发展股份有限公司制药厂	2005. 07. 07	5032371
S17	金匮肾气丸	北京同仁堂科技发展股份有限公司制药厂	2005. 07. 07	5032372
S18	金匮肾气丸	北京同仁堂科技发展股份有限公司制药厂	2005. 08. 10	5032879
S19	明目地黄丸	黄山天目药业有限公司	2006. 05. 26	060526
S20	明目地黄丸	黄山天目药业有限公司	2006. 05. 26	060526
S21	明目地黄丸	黄山天目药业有限公司	2006. 05. 26	060526
S22	明目地黄丸	黄山天目药业有限公司	2006. 11. 27	061127
S23	明目地黄丸	黄山天目药业有限公司	2006. 11. 27	061127
S24	明目地黄丸	河南宛西制药股份有限公司	2006. 05. 29	060507
S25	明目地黄丸	河南宛西制药股份有限公司	2006. 05. 29	060507
S26	明目地黄丸	河南宛西制药股份有限公司	2005. 05. 14	051002
S27	明目地黄丸	河南宛西制药股份有限公司	2005. 05. 14	051002
S28	明目地黄丸	河南宛西制药股份有限公司	2005. 05. 14	051002
S29	明目地黄丸	河南宛西制药股份有限公司	2006. 04. 15	060403
S30	明目地黄丸	河南宛西制药股份有限公司	2006. 04. 15	060403
S31	明目地黄丸	河南宛西制药股份有限公司	2006. 10. 27	061005
S32	明目地黄丸	河南宛西制药股份有限公司	2006. 10. 27	061005
S33	明目地黄丸	河南宛西制药股份有限公司	2006. 10. 27	061005
S34	知柏地黄丸	湖北端药药业有限公司	2005. 08. 11	050801
S35	知柏地黄丸	湖北端药药业有限公司	2005. 11. 09	051101
S36	知柏地黄丸	湖北端药药业有限公司	2005. 11. 09	051101
S37	知柏地黄丸	安徽华佗国药厂	2006. 01. 17	20060101
S38	知柏地黄丸	北京同仁堂科技发展股份有限公司制药厂	2006. 01. 04	6030426
S39	知柏地黄丸	北京同仁堂科技发展股份有限公司制药厂	2006. 01. 04	6030426
S40	知柏地黄丸	湖北清大药业	2005. 10. 20	051002
S41	知柏地黄丸	湖北清大药业	2006. 05. 09	060501
S42	知柏地黄丸	湖北清大药业	2006. 05. 09	060501

不同演化阶段的遗传与变异信息量为：

$$I = -(P_g lnP_g + P_a\ lnP_{va} + P_b\ lnP_{vb}) \quad (1)$$

下面列出各个变量定义见文献[5-9]：

$$P_g = \frac{N_g}{N_g + n_a + n_b} \times 100\% = \frac{N_g}{N_d} \times 100\% \quad (2)$$

共有峰率 P_g：2 个比较的指纹图谱 a、b 中的共有峰数与该 2 个指纹图谱的独立峰数 N_d 的比值。P_g 它只是共有峰 N_g 的函数。简化表示为 P。该共有峰率指标与 Jaccard 及 Sneath、Sokal 系数[41]本质相同。

$$P_{va} = \frac{n_a}{N_g} \times 100\% \quad (3)$$

$$P_{vb} = \frac{n_b}{N_g} \times 100\% \quad (4)$$

P_{va}：图谱 a 的变异峰率。P_{vb}：图谱 b 的变异峰率。

$$P_a = n_a/N_d, P_b = n_b/N_d \quad (5)$$

P_a，P_b 分别是样品 a，b 的变异峰占独立峰的比例。

表 2　最大信息量区间之 P_g 及其该区间对应的信息量变化率

对称度 α	最大信息量区间之 P_g 及其对应的信息量变化率							
	P_g 区间						信息量变化率(%)及 P_g 区间	信息量变化率(%)及 P_g 区间
	0.55	0.61*	0.65	0.66	0.70△	0.75		
0			0.497	0.500	0.504	0.491	1.34(0.65~0.70)	2.58(0.75~0.70)
0.5	0.706	0.724	0.719	0.716			2.49(0.55~0.61)	1.10(0.66~0.61)
1	0.731	0.746	0.739	0.735			2.01(0.55~0.61)	1.47(0.66~0.61)

注：* 对称度为 α=0.5,1 时，最大信息量对应的共有峰率 P_g=0.61；△对称度为 α=0 时，最大信息量对应的供共有峰率 P_g=0.70。根据图 1 信息理论方程曲线，以最大信息量点对应的 P_g 为中心，取其左右各变化 0.05 计算信息量，以直观给出信息量的变化程度。对称度 α=0，对应的最大信息量点为 P_g=0.70=70%；对称度 α=0.5,1，对应的最大信息量点皆为 P_g=0.61=61%. 插入 P_g=0.66 这个点，是为了增加数据点的密度

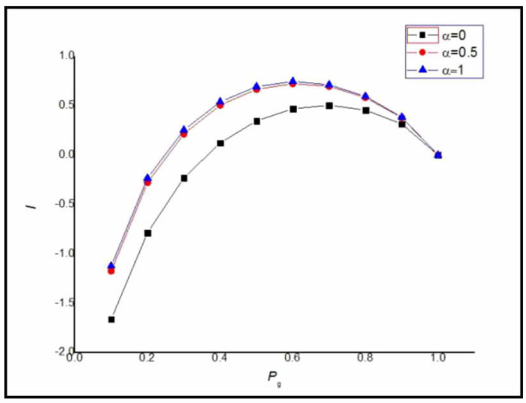

图 1　信息理论方程曲线

注：对称变异曲线对称度 α=1($P_a=P_b$)，非对称变异曲线 α=0.5($P_b=1/2P_a$)，α=0($P_a\neq0$，$P_b=0$)

各类指纹峰之间的关系为：$N_d=N_g+n_a+n_b$，$N_a=N_g+n_a$，$N_b=N_g+n_b$

N_d：2 个相互比较的指纹图谱的独立峰，即指比较的 2 个指纹图谱 a、b 中吸收峰的种类数。N_a：指纹图谱 a 的总指纹峰数。N_b：指纹图谱 b 的总指纹峰数。n_a：图谱 a 的变异峰数；n_b：图谱 b 的变异峰数。

定义一个描述对称程度的参数 $\alpha=\frac{P_b}{P_a}$，$0\leqslant\alpha\leqslant1$，α=1 表示完全对称，α=0，表示完全非对称。

该方程存在着 2 个极大值信息量，分别对应于对称变异 $P_a=P_b$，α=1，共有峰率为 P_g=61.0%，及极端非对称变异，α=0，$P_a\neq0$，$P_b=0$，共有峰率为 69.5%。见图 1。

根据 I~P_g 图，可以发现，对于对称变异 α=1 及非对称变异 α=0.5 曲线的最高点在 P_g=0.610，且在 P_g=0.610 附近存在一个平坦区域，信息量变化不大。而对于极端非对称变异 α=0，曲线的最高点在 P_g=0.695，且在 P_g=0.695 附近存在一个平坦区域，信息量变化很小。

当对称性系数 α=1，α=0.5，最大信息量对应的 P_g 皆为 0.61，最大信息对应的共有峰率 P_g 区间皆对应于 0.55~0.66。

无论在对称变异还是非对称变异，在最大信息量对应的共有峰率 P_g 左右变化 5% 的范围内，信息量的变化皆小于 2.49%，小于 3%。这说明，由于各种因素导致的生物变异的变化相对于最大信息量点之 P_g ±5% 的范围内，最大信息量变化小于 2.49%，该变化范围可以认为处于中药品种特征的临界变化区域。见表 2。

2 个样品的共有峰率越高，相似性越大，品质越相似。因此，根据表 1，考虑到各类因素对中药及实验的随机波动性影响，可以确定 P_g=61% 及 P_g=70% 作为判别生物体系的品种常数及优化判别中药品种的绝对理论标准，其变化区间为 $P_g\geqslant$(61±3)% 及 $P_g\geqslant$(70±3)%，结合最大有效样本数法对样本集进行合理优化分类[37]，同时可以避免为了追求基于经验知识的完全分类而形成的过度分类，如神经网络法中易出现的过度分类。

2.2　实验方法

2.2.1　实验条件　中药样品粉碎并过 80 目筛，60 ℃烘干 2 h，<4 ℃保存。KBr 压片法测试红外指纹图谱。每个样品平行测量指纹图谱 6 次，每个样品的红外指纹图谱峰波数是 6 次测量中出现次数的平均值。红外指纹图谱以约 25 cm^{-1} 平滑处理，最佳灵敏度 80。

2.2.2　重复性实验和稳定性实验　样品 S20 平行测量 6 次，任意 4 次测量指纹图谱中每个吸收峰波数的平均值构成组合数值指纹图谱。他们之间的最低共有峰率为 93%，具有优良的重复性。原药粉末在 <4 ℃下保存，具有良好的稳定性。

2.3　指纹图谱测定与实验数据分析

2.3.1　指纹图谱测定　根据实验条件，测试 4 类 42 个中药复方原药粉末的红外指纹图谱。见图 2~5。由图 2~5 可见，4 种复方中药原药粉末的红外

指纹图谱具有很高的相似性及复杂性，极难凭观察直观鉴别 4 类中药原药粉末的红外指纹图谱。必须用数理分析方法对其指纹图谱数据进行分析，以得到严谨精确的品质鉴别结果。

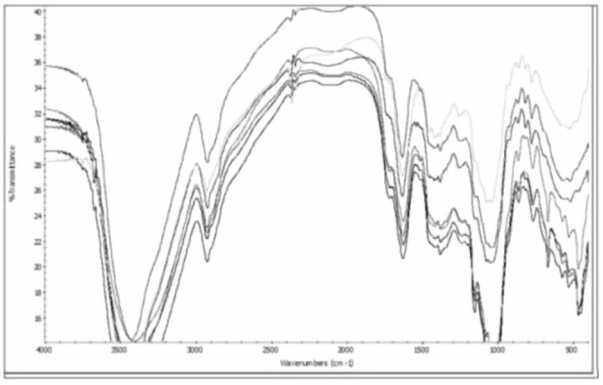

图 2 桂附地黄丸和金匮肾气丸原药粉末红外指纹图谱叠加图

注：在 1 600 cm^{-1} 处由下至上分别为桂附地黄丸 S1，S3，S8，S10，金匮肾气丸 S11，S14，S18

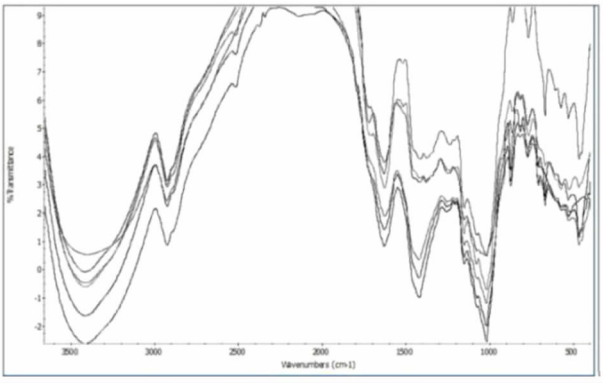

图 3 明目地黄丸和知柏地黄丸原药粉末红外指纹图谱叠加图

注：在 1 600 cm^{-1} 处由下至上明目地黄丸 S24，S26，S32，知柏地黄丸 S34，S38，S42

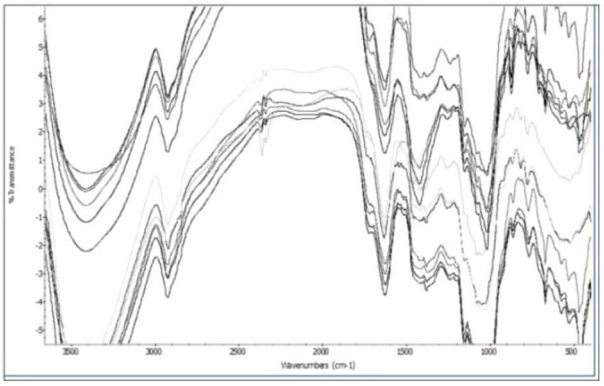

图 4 桂附地黄丸，金匮肾气丸，明目地黄丸，知柏地黄丸原药粉末红外指纹图谱叠加图

注：在 1 600 cm^{-1} 处由下至上分别为桂附地黄丸 S1，S3，S8，S10，金匮肾气丸 S11，S14，S18，明目地黄丸 S24，S26，S32，知柏地黄丸 S34，S38，S42

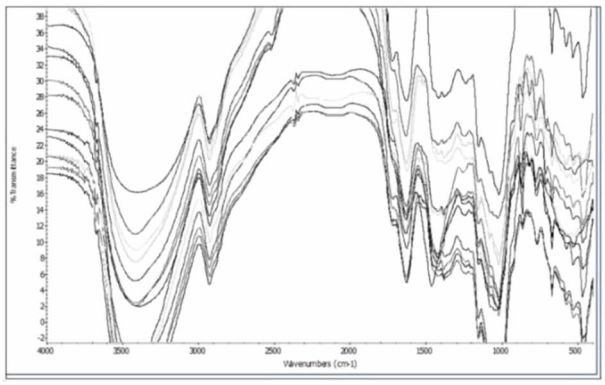

图 5 桂附地黄丸，金匮肾气丸，明目地黄丸，知柏地黄丸原药粉末红外指纹图谱叠加图

注：2 900 cm^{-1} 处由下至上桂附地黄丸 S1，S3，S8，S10，金匮肾气丸 S11，S14，S18 明目地黄丸 S19，S24，S26，S32，知柏地黄丸 S34，S37，S38，S42

2.3.2 实验数据分析 采用 Shapiro-Wilk 检验方法，确定每个样品的组合指纹图谱峰及 42 个样品的共有峰及变异峰。以任一样品为标准，计算其他样品与该样品的共有峰率，并按共有峰率由高到低的方法排序，构成样本及其对应的共有峰率的符号-数值二元序列，即双指标序列。根据得到的 2 个中药品种理论判别标准 $P_g = 61.0\%$，$P_g = 69.5\%$，结合群体特征序列最大有效样本数法[38]：

$$Y = \sum_{i=1}^{M} (N_{ci} - N_{ri}) \quad (6)$$

Y：群体特征序列最大有效样本数；N_{ci}：第 i 个样本的核心特征序列中的样本数；N_{ri}：第 i 个样本的关联序列中的样本数；M：群体样本总数。Y 反映了群体分类中有效分类程度，Y 越大说明品质聚分类越清晰，关联序列中的样本数越少，核心特征序列中的样本数越多，越接近理想聚分类。

分别在 $P_g \geq (61 \pm 3)\%$，$P_g \geq (70 \pm 3)\%$ 优化 P_g，得到 4 个品种的整体优化特征序列，即样品清晰分类且具有最大有效样本数，以实现 4 种中药复方 6 种两两配对关系的精确鉴别。

当 $P_g \geq 61\%$，$P_g \geq 72\%$ 时 4 种中药经过优化的特征序列。见表 3。

采用鉴别标准 $P_g \geq 61\%$，由每个样品的特征序列看，下列 5 个药品配对可以得到精确鉴别：桂附地黄丸-金匮肾气丸；桂附地黄丸-明目地黄丸；金匮肾气丸-明目地黄丸；金匮肾气丸-知柏地黄丸；明目地黄丸-知柏地黄丸。

而桂附地黄丸-知柏地黄丸二者的特征序列交叠严重，无法清晰区分二者。因而无法得到准确鉴别。

表 3　4 种中药经过优化的特征序列

核心特征序列	关联序列
桂附地黄丸	
S1:(S1)S2 S3 S4 S5 S6 S7 S8 S9 S10[a]	S34 S35 S36 S37 S40 S41 S42[b]
S2:S1(S2)S3 S4 S5 S6 S7 S8 S9[c] S10	S34 S35 S36 S37 S40 S41 S42
S3:S1 S2(S3)S4 S5 S6 S7 S8 S9 S10	S37 S40 S41 S42
S4:S1 S2 S3(S4)S5 S6 S7 S8 S9 S10	S34 S35 S36 S37 S40 S42
S5:S1 S2 S3 S4(S5)S6 S7 S8 S9 S10	S34 S35 S36 S37 S40 S41 S42
S6:S1 S2 S3 S4 S5(S6)S7 S8 S9	S34 S35 S36 S40 S41 S42
S7:S1 S2 S3 S4 S5 S6(S7)S8 S9 S10	S34 S35 S36 S40 S41 S42
S8:S1 S2 S3 S4 S5 S6 S7(S8)S9 S10	S34 S35 S36 S37 S40 S42
S9:S1 S2 S3 S4 S5 S6 S7 S8(S9)S10	S35 S34 S37 S41 S42
S10:S1 S2 S3 S4 S5 S7 S8 S9(S10)	S37
金匮肾气丸	
S11:(S11)S12 S13 S14 S15 S16 S17 S18	
S12:S11(S12)S13 S14 S15 S16 S17 S18	
S13:S11 S12(S13)S14 S15 S16 S17 S18	
S14:S11 S12 S13(S14)S15 S16 S17	
S15:S11 S12 S13 S14(S15)S16 S17 S18	
S16:S11 S12 S13 S14 S15(S16)S17 S18	
S17:S11 S12 S13 S14 S15 S16(S17)S18	
S18:S11 S12 S13 S15 S16 S17(S18)	
明目地黄丸	
S19:(S19)S20 S21 S22 S23	S41
S20:S19(S20)S21 S22 S23	
S21:S19 S20(S21)S22 S23	
S22:S19 S20 S21(S22)S23	S40 S41 S42
S23:S19 S20 S21 S22(S23)	
S24:(S24)S25 S26 S27 S28 S29 S30 S31 S32 S33	
S25:S24(S25)S26 S27 S28 S29 S30 S31 S32 S33	
S26:S24 S25(S26)S27 S28 S29 S30 S31 S32 S33	
S27:S24 S25 S26(S27)S28 S29 S30 S31 S32 S33	
S28:S24 S25 S26 S27(S28)S29 S30 S31 S32 S33	
S29:S24 S25 S26 S27 S28(S29)S30 S31 S32 S33	
S30:S24 S25 S26 S27 S28 S29(S30)S31 S32 S33	
S31:S24 S25 S26 S27 S28 S29 S30(S31)S32 S33	
S32:S24 S25 S26 S27 S28 S29 S30 S31(S32)S33	
S33:S24 S25 S26 S27 S28 S29 S30 S31 S32(S33)	
知柏地黄丸	
S34:(S34)S35 S36 S37 S40 S41 S42	S1 S2 S4 S5 S6 S7 S8 S9
S35:S34(S35)S37 S40 S41	S1 S2 S4 S5 S6 S7 S8 S9
S36:S34(S36)S37 S40 S42	S1 S2 S4 S5 S6 S7 S8
S37:S34 S35 S36(S37)S40	S1 S2 S3 S4 S5 S8 S9 S10
S38:(S38)S39	
S39:S38(S39)	
S40:S34 S35 S36 S37(S40)S41 S42	S1 S2 S3 S4 S5 S6 S7 S8 S22
S41:S34 S35 S40(S41)S42	S1 S2 S3 S5 S6 S7 S9 S19 S22
S42:S34 S36 S40 S41(S42)	S1 S2 S3 S4 S5 S6 S7 S8 S9 S22

注:a. 核心特征序列;b,关联序列,该序列部分中的样品不属于核心特征序列样品对应的类。每个样品的特征序列由核心特征序列 + 关联序列构成。c,特征序列中的斜黑体部分,表示当鉴别标准 $P_g \geqslant 72\%$ 时 4 种中药的特征序列,而每个样本的整个序列表示 $P_g \geqslant 61\%$ 时 4 种中药的特征序列

当 $P_g \geqslant 72\%$ 时,S35,S37 的特征序列中,关联序列等于或长于核心特征序列,因而无法正确识别,其余 40 个样品得到正确的品种判别,正确识别率为 40/42 = 95.2%,错误识别率为 2/42 = 4.8%。不仅上述 5 个药品对得到精确鉴别,在此情况下,桂附地黄丸-知柏地黄丸也得到精确鉴别。对于桂附地黄

丸-知柏地黄丸配对，依据非对称变异品种常数 P_g =70(69.5)%，结合群体特征序列最大有效样本数法[37]，在不同 P_g 条件下，优化分类样品，所得结果见表4。

表4 4种复方非对称变异常数 $P_g=(70\pm x^*)\%$ 的优化结果

P_g(%)	优化结果	
	最大有效样本数	正确识别(%)
≥69	249	92.9
≥70	254	92.9
≥71	254	92.9
≥72	256	95.2
≥75	243	100.0

x^*, x = −1,0,1,2,5

由表4结果知，当 P_g≥69%~75%时，桂附地黄丸-知柏地黄丸药品对皆可得到精确鉴别。但 P_g≥72%时，有效样本数最大，分类最合理，而 P_g≥75%时，虽然样本的正确识别了100%，但有效样本数最小，说明过度分类，并不合理。通过对实验测量的指纹图谱数据进行详细分析，可以发现评判这4类中药的品种标准可以确定为 P_g≥61%，P_g≥72%。见表5。

表5 4种复方配对遵从的品种常数 P_g(%)

品种	桂附地黄丸	金匮肾气丸	明目地黄丸	知柏地黄丸
桂附地黄丸	–	61	61	70
金匮肾气丸	61	–	61	61
明目地黄丸	61	61	–	61
知柏地黄丸	70	61	61	–

在上述相似尺度下，6对中药配对皆可得到准确判别分类，样品正确识别率95.2%，具有最大有效样本数。在类的层次上，类的正确识别率为100%。

另一方面，可采用双指标等级序列个性化模式识别法对4种复方粉末红外指纹图谱进行分析，当采用相似尺度 $P_g\geq\bar{P}_g+1.3S$ 时，只有S35错误识别，即关联序列长于核心特征序列。S37关联序列与核心特征序列等长，无法判别，错误识别率2/42=4.8%，正确识别率达95.2%。该结果与采用2个中药品种常数进行鉴别分类结果相同，说明新建立的方法的有效性。

3 讨论

中药品种的科学鉴别是中药现代研究的核心课题之一。只有在科学上绝对保证品种鉴别的可靠性，才能保证疗效的可靠性，为中医药正名，并为中医药的现代深化科学研究奠定不可动摇的基础。目前的中药品种鉴别都基于经验知识，即使是众多的模式识别理论方法，也是基于经验知识或学习样本获得经验参量，这些方法皆无法得到不可怀疑的结论。本研究根据生物体系共有遗传变异信息理论方程确定了2个中药品种常数 P_g=61%，P_g=70%，在科学上提出了2个判别中药品种的绝对理论定量标准。基于它们建立的2个品种判别理论标准 P_g≥(61±3)%及 P_g≥(70±3)%，对4种组成相似的中药复方的6种配对进行了精确鉴别。结合文献[37-39]可以证明，基于生物遗传与变异信息理论方程在2种变异状态取得最大信息量所对应的共有峰率[37]，确实可以反映生物体系的本质相似特征。该理论方法无需任何经验知识、主观设定参数，只需依据实验测量值，结合最大有效样本数法就可实现对某些中药这类生物复杂体系的精确鉴别及模式识别。最大有效样本数法可以防止其他模式识别方法容易产生的过度拟合分类现象，而保证给出最优合理结果。这也证明，在生物体系中同样存在如同物理学、化学中的精确定理定律，有待人们去发现。

依据生物遗传与变异信息理论推导的2个中药品种绝对定量标准而不是依据经验标准对中药精确鉴别，这应是生物体系中药鉴别理论方法的一个重要突破。该理论方法亦具有巨大的简易性，且也应适用于任意生物体系的精确分类鉴别。另一方面，作者多年的系统实验研究及理论研究发现，基于生物体系的红外、液相指纹图谱信息及分子种类信息对生物体系进行鉴别、分类、聚类及模式识别研究，结果较基于成分的含量信息更可靠。因为这些信息主要反映的是生物体系的物质结构信息，结构决定功能，结构决定品种特性。理论上，这2个品种常数适用于基于红外指纹图谱、液相指纹图谱、分子种类信息的生物体系品种的绝对定量鉴别。

目前的中医药研究，主要聚焦于其中的化学成分、功效研究，对中医药数理基础理论研究工作缺乏应有的重视。要使以几千年经验为基础的中医药演化至以数理理论为基础的现代科学，发现其中的精确科学原理，用于更好地指导中医药现代研究，还有漫漫之路有待探索。但这是能够发扬光大祖国中医药学的必由之路。

致谢：本文部分数据取自我的研究生韩智峰、张新玲的毕业论文。

参考文献

[1]汪劲武. 种子植物分类学[M]. 2版. 北京：高等教育出版社，2011：1-394.

[2]周长发. 生物进化与分类原理[M]. 北京:科学出版社,2009:1-211.
[3]恩斯特. 迈尔. 生物学思想发展的历史[M]. 成都:四川教育出版社,2010:1-565.
[4][美]Douglas J. Futuyma. 生物进化[M]. 3 版. 北京:高等教育出版社,2016:443-453.
[5][英]N. H. 巴顿,[美]D. E. G. 布里格斯,J. A. 艾森,D. B. 戈尔茨坦,N. H. 帕特尔. 进化[M]. 北京:科学出版社,2010:648-686.
[6][美]Michael G. Simpson. Plant systematics(second edition)[M]. 北京:科学出版社,2011:3-678.
[7]邹华彬,袁久荣,袁洁. 中药质量控制技术方法的历史发展[J]. 中成药,2003,25(9):附 4-4 附 6.
[8][美]爱德华. 威尔逊. 缤纷的生命[M]. 北京:中信出版集团,2016:59-62.
[9]邹华彬,袁久荣,袁浩. 秘鲁产人参保健品 HPLC 和四维 UV 指纹图谱研究[J]. 中成药,2003,25(4):261-265.
[10]邹华彬,袁久荣,吕青涛,等. 共有峰率和变异率双指标序列分析法分析甘草紫外指纹图谱[J]. 中药材,2003,26(9):625-629.
[11]邹华彬,袁久荣,杜爱琴,等. 甘草氯仿提取物红外指纹图谱双指标序列分析法[J]. 中国中药杂志,2005,30(1):16-20.
[12]Zou Hua-bin,Yuan Jiu-rong,Du Ai-qin,et al. Dual-Index Sequence Analytical Method for IR Fingerprint Spectra of Ethanolic Extract of Various Gylcyrrhizae′s Root Species components[J]. Analytical letters,2005,38(7):1167-1178.
[13]Zou H,Yang G,Du A,et al. Combinational numeral fingerprint spectra of Glycyrrhiza and analysis of common peak ratio invariableness in HPLC[J]. Biomed Chromatogr,2006,20(6-7):642-655.
[14]邹华彬,袁浩,王爱武,等. 白芍紫外指纹图谱共有峰率和变异峰率双指标序列分析[J]. 光谱学与光谱分析,2007,27(9):1815-1819.
[15]Hua-Bin Zou,Guo-Sheng Yang,et al. Recognition of Radix Paeoiae Alba herbs(Fam. Ranunculaceae)with HPLC Fingerprint Spectra[J]. Analytical letters,2008,41(18):3309-3323.
[16]徐国锋,张丽佳,孟宪生,等. 瓜蒌皮药材紫外谱线图谱分析[J]. 辽宁中医学院学报,2006,8(1):87-88.
[17]冀宪领,盖英萍,牟志美,等. 白僵蚕的红外指纹图谱鉴别研究[J]. 光谱学与光谱分析,2007,27(1):66-69.
[18]刘圆,孟庆艳,彭镰心,等. 共有峰率和变异峰率双指标序列法分析藏药材红毛五加特定(指纹)图谱[J]. 药物分析杂志,2007,27(8):1182-1185.
[19]胡爱华,邢世岩,巩其亮. 基于 FTIR 的银杏木材鉴别研究[J]. 中国农学通报,2009,25(4):88-92.
[20]程云清,刘剑锋,刘强,等. 高山红景天红外指纹图谱共有峰率和变异峰率的双指标序列分析[J]. 南京农业大学学报,2011,34(5):155-158.
[21]黄丽萍,吴静. 自然铜远红外指纹图谱共有峰率和变异峰率双指标序列分析法[J]. 中国中药杂志,2011,36(11):1441-1144.
[22]石芸,谈献和,池玉梅,等. 黄蜀葵花红外指纹图谱研究[J]. 中国野生植物资源,2012,31(2):29-32.
[23]王燕,王斌,徐银峰,等. 基于聚类分析法和双指标分析法的淡菜红外指纹图谱比较研究[J]. 中国食品学报,2013,13(1):178-182.
[24]周晔,张庆伟,李佩孚,等. 傅里叶红外光谱法鉴别不同产地中药独活的研究[J]. 中国中药杂志,2013,38(19):3309-3312.
[25]万忠民,安坤,王九峰,等. 太子参红外指纹图谱的双指标序列分析[J]. 天然产物研究与开发,2013,25(5):594-599.
[26]庄琳,黄群,徐燕红. 楠属和润楠属 4 种木材的红外光谱鉴别[J]. 福建林业科技,2014,41(4):21-25.
[27]史静超,刘必旺,李亚妮,等. 不同道地产区黄芪土壤红外指纹图谱比较研究[J]. 山西大学学报:自然科学版,2015,38(2):335-342.
[28]彭舟,赵玉勤,王斌,等. 海洋中药珍珠原材料的红外指纹图谱研究[J]. 安徽农业科学,2015,43(12):32-34.
[29]邹华彬,韩智峰,翟红,等. 桂附地黄丸金匮肾气丸指纹图谱双指标等级序列二次聚类分析及质量评价[J]. 光谱学与光谱分析,2007,27(12):2432-2436.
[30]邹华彬,张新玲,翟红,等. 明目地黄丸知柏地黄丸无水乙醇提取物指纹图谱双指标等级序列模式识别[J]. 中国中药杂志,2008,33(13):1543-1549.
[31]董凤娟,邹华彬,陆玮洁,等. 双指标等级序列模式识别法分析十全大补丸氯仿提取物红外指纹图谱[J]. 山东大学学报:工学版,2010,40(6):113-119.
[32]王丽双,邹华彬,王四顺,等. 维药复方祖卡木的高分辨精细鉴别及质量评价[J]. 中成药,2011,33(7):1198-1201.
[33]邹华彬,王丽双,翟红. N-生物系统双指标等级序列个性化模式识别法及左归丸、右归丸和济生肾气丸核磁共振指纹图谱分析[J]. 山东大学学报:工学版,2012,42(4):114-119,131.
[34]邹华彬,王玲玲,杜爱琴,等. 4 种复方丹参中药的精确分类及药效的理论评价[J]. 世界中医药,2015,10(6):904-909.
[35]邹华彬,董凤娟,张新玲,等. △Sr 等级序列模式识别法分析补中益气丸和十全大补丸 HPLC 指纹图谱[J]. 山东大学学报:工学版,2011,41(6):97-102,114.
[36]邹华彬,阿依古再丽. 阿不力米提,翟红. 矿物药蒙脱石散 EDS 指纹图谱双指标等级序列个性化模式识别分析及快速质量评价[J]. 光谱学与光谱分析,2013,33(12):3321-3325.
[37]邹华彬. 相同品质中药的系统理论判别方法[J]. 世界中医药,2015,10(7):1078-1082.
[38]邹华彬. 双指标信息显著相似序列聚类分析桂附地黄丸、金匮肾气丸无水乙醇提取物红外指纹图谱[J]. 中国中药杂志,2009,34(18):2325-2330.
[39]邹华彬. 基于指纹图谱的生物体系中药内禀品质等级的数理理论判别[J]. 世界中医药,2016,11(9):1876-1881.
[40]宋宪民,郭维加. 新编国家中成药[M]. 北京人民卫生出版社,2002:288,394,524,1086.
[41]徐克学. 生物数学[M]. 北京:科学出版社,2002:91-94.

(2018-03-08 收稿 责任编辑:杨觉雄)