

药用模式生物研究策略

徐江^①, 孙超^①, 徐志超^①, 季爱加^①, 胡鸢雷^②, 孙伟^③, 王丽芝^①, 汪波^①, 杨培^①,
张鑫^①, 宋经元^①, 陈士林^{①③*}

① 中国医学科学院/北京协和医学院药用植物研究所, 北京 100193;

② 北京大学生命科学学院, 蛋白质工程及植物基因工程国家重点实验室, 北京 100871;

③ 中国中医科学院中药研究所, 北京 100700

* 联系人, E-mail: slchen@implad.ac.cn

2013-05-12 收稿, 2013-07-09 接受, 2014-02-26 网络版发表

国家自然科学基金重点项目(81130069)和长江学者和创新团队发展计划(IRT1150)资助

摘要 模式生物是用于研究某种特定的生物学现象而被选定的物种. 基于模式生物的研究策略已经在多个生物学领域取得巨大的成功, 揭示了众多生命科学的机理, 发展了一系列生命科学研究技术. 近年来, 现代生物学发展迅速, 其相关的概念和技术正不断渗入并影响着药用生物的研究领域, 包括基于模式生物的研究策略. 丹参(*Salvia miltiorrhiza*)、灵芝(*Ganoderma lucidum*)等模式药用生物已经提出了多年, 受到国内外的广泛关注. 但是由于缺乏成熟的模式药用生物研究体系, 使得当前药用生物的研究成果比较分散, 难以在理论水平上汇总提高, 极大地消耗了药用生物界的研究资源和研究力量, 阻碍了药用生物学的发展. 本文结合近期药用生物学的研究成果, 从药用模式生物研究体系建立的必要性出发, 提出了药用模式生物研究体系的基本概念和建立目的, 分析了应用药用模式生物研究体系研究的可行性, 提出了药用模式生物的选择原则并从遗传信息获得、遗传转化体系建立、突变体库构建和次生代谢物生产体系建立 4 个方面描述了药用模式生物研究体系的建立策略. 本文还介绍了以药用模式生物为对象的次生代谢产物生源合成及调控研究策略, 同时认为, 药用模式生物体系在阐释药材道地性等中医药传统问题及发展合成生物学等现代科学前沿中发挥重要作用.

关键词

模式生物
药用生物学
次生代谢
道地药材
合成生物学

灵芝^[1,2](*Ganoderma lucidum*)和丹参^[3](*Salvia miltiorrhiza*)被提出作为药用模式真菌和药用模式植物, 受到国内外的广泛关注. 本文介绍模式生物与药用模式生物的概念和源起, 并从筛选原则、研究体系建立及应用等方面探讨药用模式生物的研究策略.

1 模式生物与药用模式生物

模式生物是用于研究某种特定的生物学现象而被选定的物种, 帮助人们从错综复杂的生物学现象中抽取通用的生物学本质^[4]. 以模式生物为研究对象的概念已经提出了 100 多年, 其研究策略在多个生命

科学领域获得了巨大的成就. 据统计, 发表在 *Nature*, *Science* 和 *Cell* 等高水平期刊上有关生命过程和机理的文章, 以模式生物为材料完成的占到了 80% 以上^[5]. 随着生命科学研究的进一步深入和细化, 模式生物的范畴正在不断扩大(图 1), 在一些特定的生物学分支中, 针对不同研究方向和科学问题, 出现了一些新的模式生物, 如 20 世纪 80 年代起用于神经发育研究的模式动物——斑马鱼(*Danio rerio*)^[6]和最近兴起的用于抗盐研究的模式植物——盐芥(*Thellungiella halophila*)^[7].

药用生物是对可用于疾病治疗的所有生物的统一称. 次生代谢产物是药用生物的主要有效成分, 因此

引用格式: 徐江, 孙超, 徐志超, 等. 药用模式生物研究策略. 科学通报, 2014, 59: 733-742

Xu J, Sun C, Xu Z C, et al. Research strategy for model medicinal species (in Chinese). *Chin Sci Bull (Chin Ver)*, 2014, 59: 733-742, doi: 10.1360/972013-565

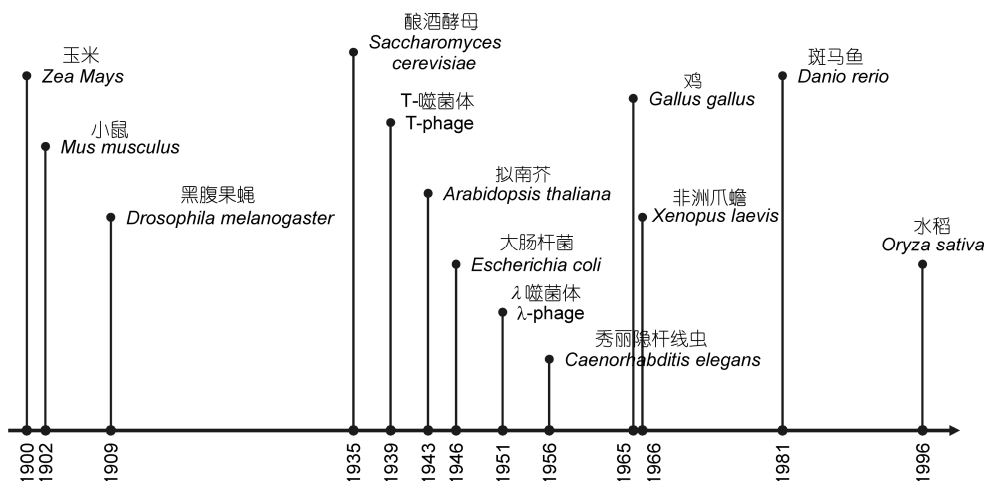


图1 部分模式生物及其提出年份

次生代谢产物的合成与调控机理的阐释是药用生物研究的核心问题^[8]。自然界中已经发现数十万种次生代谢产物，它们来源于有限的前体或模块，生源合成这些前体或模块的途径仅有 10 余种，这表明次生代谢途径具有一定的保守性^[9]。次生代谢是生物对外界环境信号的响应，研究发现，一定范围内，这种响应机制也是保守的，如高等植物中茉莉素是很多次生代谢途径响应外界信号的共同诱导物^[10,11]。因此，从进化的角度看，不同生物之间次生代谢产物的合成与调控必然存在着共有的规律和机制，以模式生物为对象的研究策略可以在药用生物研究中发挥重要作用。

药用生物研究还缺少成熟的模式生物研究体系，这也是药用生物研究与其他生物学领域相比还相对落后的一个重要原因。在未发现良好的药用模式生物的情况下，拟南芥(*Arabidopsis thaliana*)、酿酒酵母(*Saccharomyces cerevisiae*)等经典模式生物曾作为次生代谢的研究模型，但是由于缺乏特有的次生代谢产物富集器官如腺毛^[12]或者一些重要的药用活性次生代谢产物合成的关键酶，很多次生代谢产物合成和调控的过程在经典的模式生物研究中仍未找到答案，所以亟待建立以药用生物为对象的模式生物研究系统暨药用模式生物研究系统。

2 药用模式生物的一般特征及选择策略

药用生物资源丰富、种类繁多，仅药用植物就超过 10000 种^[13]。如何从众多的药用生物中选择少数合适的物种作为模式生物，是药用模式生物研究的首要

问题。药用模式生物的选择需要遵循一定的原则。

2.1 拥有典型的次生代谢途径和代表性次生代谢产物

次生代谢产物是大多数药用生物药效的物质基础，因此药用模式生物应该具有代表性的药用活性成分及典型的次生代谢途径。按照合成的起始分子不同，次生代谢产物可以分为萜类、生物碱、脂肪酸和苯丙烷类等；合成途径主要包括丙二酸途径、莽草酸途径、甲戊二羟酸途径、2-甲基-D-赤藻糖醇-4-磷酸途径、氨基酸途径等^[14](图2)。同一次生代谢产物可能由不同途径独立生成，也可能由经多条途径共同合成。不同生物合成次生代谢产物的种类和能力不同，因此，针对不同途径有必要推选不同的药用模式生物。研究的次生代谢产物应具有代表性，且易于提取和检测。

2.2 具有模式生物的一般特征

药用模式生物应该具有模式生物的共同特征。从模式生物的一般生物学属性上看，通常具有世代周期较短、子代多，表型稳定等特征^[15]。世代短可以节省实验观察周期，子代多有利于突变表型的发现。基因组序列是开展分子生物学研究的基础和捷径，因此，基因组相对较小、易于进行全基因组测序是目前新的模式生物筛选的一个重要标准。

2.3 具有良好的前期研究基础

与其他生命科学研究领域相比，药用生物学研

细图^[2]和丹参基因组框架图已完成^[13]. 基因组的测序完成将极大地推进该物种的分子生物学研究, 进而为其成为药用模式生物、发挥模式作用奠定基础.

3 药用模式生物研究体系的建立策略

模式生物研究体系包括4个基本方面: (i) 高精度的遗传信息; (ii) 高效的遗传转化体系; (iii) 高覆盖度的突变体库; (iv) 适合的次生代谢产物生产研究系统. 本文从这4个方面对完善现有药用模式生物研究体系及建立新的模式药用生物研究体系的策略和技术进行简要阐述(图3).

3.1 基因组图谱绘制策略

模式物种的研究需要较清晰的遗传背景信息, 最直接的方法就是全基因组图谱的绘制^[23]. 基因组图谱绘制包括材料获取、遗传图谱或物理图谱构建、测序文库构建、序列测定、序列组装、基因注释和后期分析等. 材料获取是基因组图谱绘制的第一步, 纯合体是基因组测序的最优选择. 单倍体加倍或者多代自交是纯合系获得的主要方式. 获得单倍体的方法有花粉花药培养、基于种间杂交的染色体消除、诱导孤雌生殖等^[24,25], 最近, *Nature* 还报道了一种基于着丝粒改造的单倍体获得技术, 为植物单倍体获取提供了新的思路^[26]. 遗传图谱和物理图谱可以辅助

序列的拼接、定位. 当前高通量测序技术和光学图谱技术的应用有效地缩短了遗传图谱^[27,28]及物理图谱^[29,30]的建立周期. 高通量测序技术也是当下基因组序列测定的主要技术, 其中二代测序技术通量高^[31]、三代技术侧重序列的长度^[32], 多种测序平台和建库策略混合使用是目前基因组测序策略的主流. 新的测序技术的引入和海量数据的涌现对生物信息提出了更高的要求, 序列拼接是很多基因组项目的瓶颈. 主流的拼接程序主要基于图论, 包括应用于OLC(overlap/layout/consensus)方法的重叠图^[33]和基于贪婪算法的 de Bruijn 图^[34]. 序列拼接、组装和定位需借助遗传图谱和物理图谱, 有时也可参考近缘物种. 完成的基因组图谱应进行系统的验证.

3.2 遗传转化体系的建立策略

自1972年, Berg 实验室报道将噬菌体和大肠杆菌(*Escherichia coli*)半乳糖操纵子插入到病毒 SV40 中以来^[35,36], 遗传转化技术已成为当前生物技术领域的重要组成部分, 也是研究基因功能的重要手段. 已经报道的外源基因转化方法多达十几种(表1), 其中农杆菌(*Agrobacterium tumefaciens*)介导的遗传转化由于其受体类型多样、转化效率高、单拷贝随机插入和转化子稳定等特点在植物和真菌转化中广泛使用^[37-39]. 最近, 也有一些新的转化方法出现, 如叶绿

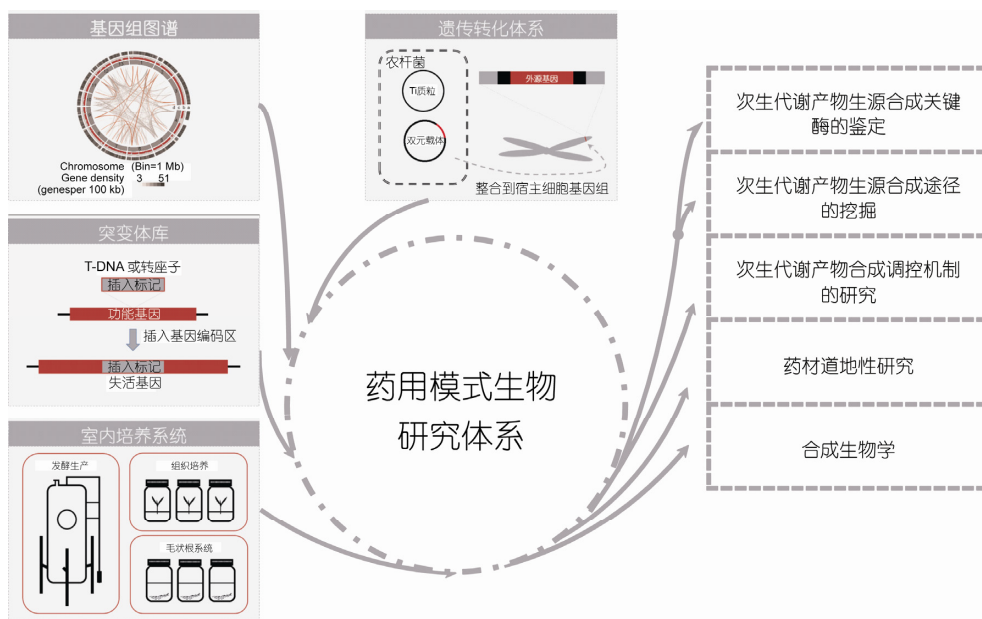


图3 药用模式生物研究体系建立策略及应用方向

体转化、人工微小染色体转化方法等。这些新的转化手段可以一次性转入多个基因，并避免插入位点导致的基因沉默，尤其适用于次生代谢途径的研究和改造^[40-42]。转化事件的检测借助于标记基因，包括选择标记基因和报告基因。很多标记基因兼具选择标记和报告两种功能。常用的选择标记基因主要有抗生素抗性标记、抗代谢物标记、除草剂抗性标记、激素代谢标记、氨基酸代谢标记及糖代谢标记等。常用的报告基因有 beta-葡萄糖苷酸酶(beta-glucuronidase, *GUS*)、萤火虫荧光素酶(luciferase, *LUC*)、绿色荧光蛋白(green fluorescent protein, *GFP*)和花青素生物合成调节基因等^[43,44]。

3.3 突变体库的构建策略

观察功能缺失突变体的表型是一种研究基因功能的直接方式。早期人工突变体主要由理化诱变方法如化学诱变^[45]、电离辐射^[46]等生成。目前，插入突变是人工突变库建立主要方式。插入突变是利用转移 DNA(transferred DNA, T-DNA)或转座子标签插入到受体基因组中，插入位点的基因功能受到抑制最终产生基因敲除突变体。T-DNA 标签突变是一种以

农杆菌介导的遗传转化为基础的插入突变研究方法，利用根癌农杆菌的 Ti 质粒上可以整合到植物基因组中稳定表达的 T-DNA，将外源基因随机插入植物基因组中。在模式生物酿酒酵母^[47]、拟南芥^[48]及水稻(*Oryza sativa*)^[49]中已建立了较为完善的 T-DNA 标签插入突变体库。转座子诱导的突变体库利用转座子激活子(activator, Ac)/分离(dissociation, Ds)系统，也是常用的一种突变体库的建立方法。通过对插入标签的分离可以确定插入位点在基因组中的位置，并通过表型鉴定相应基因功能^[50]。表 2 为部分模式生物及其突变体库。

3.4 辅助研究系统的建立

很多药用生物自身生长周期较长，活性成分也多在特定发育阶段或特定的部位富集，因此在药用模式生物研究中，推荐建立除全植株外的辅助研究系统，选择恰当的发育阶段或合适的培养方式，建立优化的研究系统，使目的产物产量最优，也可有效稳定实验条件，缩短实验周期，降低实验背景噪音。发酵和毛状根培养是常用的次生代谢产物研究系统。灵芝中应用二阶段发酵法，其三萜酸总含量可占到

表 1 常用原核/真核生物转化/转染方法

方法名称	方法简述	优点	缺点	作用对象
农杆菌介导	农杆菌侵染植物伤口, T-DNA 插入植物基因组中	拷贝数低, 转化效率高, 稳定性高	受宿主范围限制	植物、真菌
PEG 介导	用聚乙二醇(polyethylene glycol, PEG)促进原生质体吸收外源质粒 DNA	不受宿主范围限制, 操作简便	原生质体再生困难, 转化率低	细菌、真菌、植物、动物
基因枪法	外源 DNA 吸附到金属颗粒表面, 高压下轰击射进受体细胞	不受宿主范围限制, 操作简便, 可控性高	轰击细胞损伤较大, 拷贝数高, 会导致基因重排现象及转化后代不稳定	细菌、真菌、植物、动物
电激法	高压脉冲作用在原生质体膜上电激形成瞬间通道, 将质粒 DNA 导入受体细胞中	不受宿主范围限制	原生质体再生困难, 转化率低	细菌、真菌、植物、动物
超声法	低声强脉冲超声波造成细胞膜出现可逆小孔, 外源 DNA 进入细胞	不受宿主范围限制, 操作简便	损伤细胞, 易使外源 DNA 分子发生断裂	细菌、真菌、植物、动物
显微注射法	将外源 DNA 直接注射到受体细胞质或细胞核中	转化效率高	操作困难, 细胞不易固定, 表达不稳定	细菌、真菌、植物、动物
花粉管通道法	授粉后, 外源 DNA 经花粉管注入子房, 转化入尚未分化的卵、合子或早期胚胎细胞	无需继代培养, 操作简便	只适用开花植物, 转化率低	植物
脂质体法	包含外源 DNA 的脂质体与原生质体融合	操作简便	原生质体再生困难, 转化率低	细菌、真菌、植物、动物
叶绿体转化	基因枪法介导的叶绿体转化	不通过花粉传播, 安全性高	缺少有效的选择标记基因, 外源基因不稳定	植物
纳米基因载体转化法	外源基因吸附在纳米微粒表面或包埋于内部形成纳米基因复合物, 与细胞表面受体结合, 导入细胞	不受宿主范围限制, 操作简便	应用较少, 转化率低	细菌、真菌、植物、动物

表2 部分模式生物突变体库

物种	数据库	网络地址	文献
大肠杆菌(<i>E. coli</i>)	CGSC	http://cgsc.biology.yale.edu/	[51]
酿酒酵母(<i>S. cerevisiae</i>)	PhenoM	http://phenom.cabr.utoronto.ca/	[52]
秀丽隐杆线虫(<i>Caenorhabditis elegans</i>)	NBRP::C. elegans	http://www.shigen.nig.ac.jp/c.elegans/index.jsp	[53]
果蝇(<i>Drosophila melanogaster</i>)	FlyBase	http://flybase.org/	[54]
小鼠(<i>Mus musculus</i>)	PBmice	http://idm.fudan.edu.cn/PBmice/	[55]
拟南芥(<i>A. thaliana</i>)	CSHL Trapper DB	http://genetrapp.cshl.org/TrPhenotypes.html	[56]
水稻(<i>O. sativa</i>)	Rice Mutant Database	http://rmd.ncpgr.cn/	[49]

菌丝干重的4%，单一化合物7-乙氧基灵芝酸O的获得率达到1.5 g/100 g^[57]。毛状根作为生物反应器合成次生代谢产物在植物中应用广泛^[58]。野生型发根农杆菌诱导后的毛状根中次生代谢产物明显提高，毛状根相比细胞悬浮培养具有生长快、易转化、稳定性高、次生代谢产物产量高等优势，在长春花、丹参、人参、黄芩(*Scutellaria baicalensis*)、甘草(*Glycyrrhiza uralensis*)等药用植物中已比较成熟，并且在毛状根基础上利用基因工程等手段来研究药用植物关键酶基因及转录因子功能等也被广泛关注，如过表达、基因沉默等手段。

4 药用模式生物的应用策略

4.1 单基因的功能研究策略

超表达、基因沉默、基因敲除等技术是模式生物研究基因功能的通用策略^[59]。在药用生物次生代谢产物关键基因的功能研究上有着巨大的应用潜力。

超表达是利用高活性的组成型启动子或特异型启动子，通过转化使目的基因获得高水平的表达从而观察基因功能的方法。诱导型超表达可部分实现基因在时间、空间和数量上的控制，非诱导条件下，基因处于沉默状态，不会影响生物的正常功能，也不会导致多重效应^[59]。

基因沉默技术是基于反义技术人工合成与DNA或RNA互补的寡核苷酸，使其抑制目的基因的表达。它可在转录水平和翻译水平调节基因的表达^[60]。在转录水平上，反义寡核苷酸可与基因组目标区域形成三螺旋或D环抑制靶基因的表达；在翻译水平上，可与目标mRNA形成双链，诱导RNaseH降解目标mRNA，阻滞mRNA的翻译。病毒介导的基因沉默技术(virus induced gene silencing, VIGS)^[61]是一种新的

基因沉默技术，能够快速观察基因沉默效果且不需要稳定的转化体系。该方法可以同时干涉一个特定基因或者多个同源的基因。VIGS的过程主要包括整合目的基因片段到病毒载体、侵染以及植物防御过程。近年来应用VIGS技术进行药用植物次生代谢产物调控的研究取得诸多重要成果，包括那可汀^[62]和甜菜红素^[63]的生物合成等。病毒介导的应用微小RNA(microRNA, miRNA)进行基因沉默技术(VIGS using miRNAs, MIR-VIGS)是VIGS的一个发展，以病毒载体介导的人工miRNA进行基因沉默^[64]。该方法不同于传统的VIGS技术，它可以通过计算设计出人工miRNA用来调节靶标基因的表达量，从而达到基因沉默的目的。该技术也可以用来研究物种中本身miRNA的功能。

基因敲除通过构建目标基因突变或缺失的同源序列，通过重组方法获得嵌合体，进一步通过交配获得突变纯合子^[65]。类转录激活因子效应物核酸酶(transcription activator-like effector nuclease, TALEN)技术是一种新的敲除技术。TALEN技术通过表达一个重组核酸酶，在靶点识别结构域的作用下，识别靶点序列，发挥内切酶作用，由于在DNA双链修复中会引入错误，因而目标基因有一定的概率被失活，从而达到基因敲除的目的^[66]。目前，TALEN技术已经在人^[67,68]及小鼠^[69]、斑马鱼^[70]、爪蟾(*Xenopus laevis*)^[71]、拟南芥^[72]和水稻^[73]等模式生物中广泛应用。

4.2 次生代谢途径的挖掘及形成机制研究

次生代谢生源合成途径的挖掘及形成机制在经典的模式生物研究中已取得阶段性成果。比如燕麦素是三萜类衍生物，其合关键酶为β-香树精合酶。燕麦素及其衍生物由多步修饰完成，包括乙酰化、糖基化等。β-香树精合酶及相关的修饰酶在燕麦(*Avena*

spp.)中以基因簇的形式存在^[74],这种以基因簇的形式合成和调控次生代谢产物的形成机制在植物中一直没有得到答案.Field和Osborn^[75]发现,拟南芥中去饱和 thaliandiol 合成的关键酶 At5g47980 是燕麦 β -香树精合酶的近缘基因.去饱和 thaliandiol 是拟南芥根特异表达的一种三萜化合物,其下游途径主要由 3 步反应完成.在拟南芥中,催化这 3 步反应的 3 个酶位于一个 30 kb 左右的区间内,成簇存在.尽管 At5g47980 与燕麦 β -香树精合酶近缘,去饱和 thaliandiol 基因簇和燕麦素基因簇并非来自同一祖先,也不是水平转移的产物,因此,进化中的选择压力是这 2 个基因簇形成的原因.去饱和 thaliandiol 基因簇的发现是利用模式生物研究策略发现和鉴定次生代谢产物生源合成途径的一个成功范例,同时它首次揭示了植物中以基因簇的形式合成次生代谢产物的产生机制.

4.3 次生代谢产物调控机制的研究

药用生物次生代谢途径是一个复杂的动态过程,受到外界诱导因子的刺激和内在调控机制的影响.代谢调控机制包括转录因子层面的调控也包括基因组甲基化、组蛋白修饰、非编码 RNA 等表观遗传学层面的调控.模式生物研究系统同样在次生代谢途径调控的研究中起了不可替代的作用.以拟南芥花青素的生物合成为例,通过突变体库筛查,发现转录因子 AtMYB113, AtMYB114, AtMYB75 和 AtMYB90 可激活苯丙烷合成途径,调控拟南芥中花青素的含量.miRNA 在花青素合成过程中也起着重要的作用,如 miRNA156 可降解 SPL(squamosa promoter binding like)转录本,稳定拟南芥中 MYB-bHLH-WD40 复合体从而促进花青素的合成^[76,77].

5 对我国中药学研究的意义

许多中药学问题有着复杂的研究背景和独特的历史渊源,药材的道地性问题便是其中之一.道地药材是在长期中医药的创新和传承中优胜劣汰、道地自成的优质药材,与药材种质、生态环境和人文传统有着紧密联系^[78,79].次生代谢产物是道地药材药性的物质基础,因此次生代谢产物的组成、含量及其生成规律是道地药材这一复杂问题的核心部分.道地药材的代谢表型由遗传基础和环境因素共同决定,所以在道地药材的研究中,往往同时存在着基因型、表型和环境等多个变量,使道地药材的研究难以系统

地进行.药用模式生物学体系在道地药材研究上有着独到的优势.(i)模式清晰的生物遗传背景有助于固定道地药材研究中基因型这一变量;(ii)模式生物的培养和观察在室内,为逐一改变环境因素和系统定量表型变化创造了条件;(iii)模式生物体系技术丰富,理论坚实,有助于将道地药材地域-道地的直观感觉抽象为环境/信号/转录调控/次生代谢产物合成的多维模型.需要指出的是,道地药材涉及一定的人文因素^[78],对于这一问题,当前模式生物研究系统并不适用.

药用模式生物研究系统也有助于中药材替代资源的开发.当前,我国经济高速发展,由此带来的城镇化规模迅速扩大,生态环境破坏严重,可用耕地持续紧缺,中药材的野生资源供给和人工生产都面临着巨大的压力.合成生物学为中药材短缺的问题提供一条可行的途径.基因元件的挖掘和标准化、合成途径的装配和底盘系统的优化是天然药物的合成生物学最关注的 3 个问题.当前合成生物学应用的大部分底盘系统都来源于模式生物,元件的标准化工作也基于相应的底盘系统展开.药用模式生物是很好的天然药物合成系统,次生代谢产物合成酶一般高效,合成过程中代谢流分配合理,在相同物质输入的前提下更易获得大量的目的产物,在元件发掘、改造和适配性等方面对合成生物学具有重要的借鉴意义;药用模式生物遗传信息清晰,遗传操作方便,有条件进行系统的基因组简化和改造,且较其他生物而言,对自身高产的天然药物具有较强的耐受性,有机会为天然药物合成生物学提供新的底盘系统.药用模式生物研究体系和合成生物学的结合有利于天然药物的人工合成,为中药材提供替代资源.

药用模式生物研究策略利用现代生命科学的新技术新方法对药用生物次生代谢领域的一般规律进行研究和阐释,有助于汇集中医药行业内不同地域和单位的研究力量,统一对中药研究方向的认识,完成对关键问题的重点突破,从而提升整个领域的研究水平.不可忽视的是,新技术、新方法对研究策略影响巨大,在突破性的创新技术出现时,研究策略也会相应发生变化,药用模式生物研究策略也不例外.也应该看到,现有的药用模式生物体系还不足以覆盖天然药物生源合成的全部问题,因此,新的药用模式生物也会继续涌现,不同物种的研究基础存在差别,具体到特定的物种,药用模式生物研究策略

也会有所微调。当前,药用生物已经进入到系统研究的阶段,在一段时间内,药用模式生物及其研究体系将是我国药用生物研究中新的学科制高点和科学生长点。

参考文献

- 1 孙超, 胡鸾雷, 徐江, 等. 灵芝: 一种研究天然药物合成的模式真菌. 中国科学: 生命科学, 2013, 43: 1–10
- 2 Chen S L, Xu J, Liu C, et al. Genome sequence of the model medicinal mushroom *Ganoderma lucidum*. Nat Commun, 2012, 3: 913
- 3 宋经元, 罗红梅, 李春芳, 等. 丹参药用模式植物研究探讨. 药学学报, 2013, 48: 1099–1106
- 4 Davis R H. The age of model organisms. Nat Rev Genet, 2004, 5: 69–76
- 5 朱作言. 模式生物研究. 生命科学, 2006, 18: 419
- 6 Traver D, Herbomel P, Patton E E, et al. The zebrafish as a model organism to study development of the immune system. Adv Immunol, 2003, 81: 253–330
- 7 Wu H J, Zhang Z, Wang J Y, et al. Insights into salt tolerance from the genome of *Thellungiella salsuginea*. Proc Natl Acad Sci USA, 2012, 109: 12219–12224
- 8 Wink M. Medicinal plants: A source of anti-parasitic secondary metabolites. Molecules, 2012, 17: 12771–12791
- 9 Kliebenstein D J, Osbourn A. Making new molecules—evolution of pathways for novel metabolites in plants. Curr Opin Plant Biol, 2012, 15: 415–423
- 10 Meldau S, Erb M, Baldwin I T. Defence on demand: Mechanisms behind optimal defence patterns. Ann Bot, 2012, 110: 1503–1514
- 11 Vanstraelen M, Benkova E. Hormonal interactions in the regulation of plant development. Annu Rev Cell Dev Biol, 2012, 28: 463–487
- 12 Tissier A. Glandular trichomes: What comes after expressed sequence tags? Plant J, 2012, 70: 51–68
- 13 陈士林, 孙永珍, 徐江, 等. 本草基因组计划研究策略. 药学学报, 2010, 45: 807–812
- 14 Julsing M K, Koulman A, Woerdenbag H J, et al. Combinatorial biosynthesis of medicinal plant secondary metabolites. Biomol Eng, 2006, 23: 265–279
- 15 Fields S, Johnston M. Whither model organism research? Science, 2005, 307: 1885–1886
- 16 Yan Y P, Wang Z Z. Genetic transformation of the medicinal plant *Salvia miltiorrhiza* by *Agrobacterium tumefaciens*-mediated method. Plant Cell Tiss Organ Cult, 2007, 88: 175–184
- 17 Sun L, Cai H, Xu W, et al. CaMV 35S promoter directs beta-glucuronidase expression in *Ganoderma lucidum* and *Pleurotus citrinopileatus*. Mol Biotechnol, 2002, 20: 239–244
- 18 Kiselev K V, Kusaykin M I, Dubrovina A S, et al. The *rolC* gene induces expression of a pathogenesis-related beta-1,3-glucanase in transformed ginseng cells. Phytochemistry, 2006, 67: 2225–2231
- 19 Hwang S J. Catapol production in Chinese foxglove (*Rehmannia glutinosa* Libos.) hairy roots transformed with *Agrobacterium rhizogenes* ATCC15834. Methods Mol Biol, 2009, 547: 263–273
- 20 Geu-Flores F, Sherden N H, Courdavault V, et al. An alternative route to cyclic terpenes by reductive cyclization in iridoid biosynthesis. Nature, 2012, 492: 138–142
- 21 Graham I A, Besser K, Blumer S, et al. The genetic map of *Artemisia annua* L. identifies loci affecting yield of the antimalarial drug artemisinin. Science, 2010, 327: 328–331
- 22 Paddon C J, Westfall P J, Pitera D J, et al. High-level semi-synthetic production of the potent antimalarial artemisinin. Nature, 2013, 496: 528–532
- 23 Joyce A R, Palsson B O. The model organism as a system: Integrating ‘omics’ data sets. Nat Rev Mol Cell Biol, 2006, 7: 198–210
- 24 Dunwell J M. Haploids in flowering plants: Origins and exploitation. Plant Biotechnol J, 2010, 8: 377–424
- 25 Germana M A. Gametic embryogenesis and haploid technology as valuable support to plant breeding. Plant Cell Rep, 2011, 30: 839–857
- 26 Ravi M, Chan S W. Haploid plants produced by centromere-mediated genome elimination. Nature, 2010, 464: 615–618
- 27 Kump K L, Bradbury P J, Wisser R J, et al. Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. Nat Genet, 2011, 43: 163–168
- 28 Tian F, Bradbury P J, Brown P J, et al. Genome-wide association study of leaf architecture in the maize nested association mapping population. Nat Genet, 2011, 43: 159–162
- 29 Lai Z, Jing J, Aston C, et al. A shotgun optical map of the entire *Plasmodium falciparum* genome. Nat Genet, 1999, 23: 309–313
- 30 Lin J, Qi R, Aston C, et al. Whole-genome shotgun optical mapping of *Deinococcus radiodurans*. Science, 1999, 285: 1558–1562
- 31 Shendure J, Lieberman A E. The expanding scope of DNA sequencing. Nat Biotechnol, 2012, 30: 1084–1094
- 32 Eid J, Fehr A, Gray J, et al. Real-time DNA sequencing from single polymerase molecules. Science, 2009, 323: 133–138

- 33 Koren S, Miller J R, Walenz B P, et al. An algorithm for automated closure during assembly. *BMC Bioinformatics*, 2010, 11: 457
- 34 Li R, Yu C, Li Y, et al. SOAP2: An improved ultrafast tool for short read alignment. *Bioinformatics*, 2009, 25: 1966–1967
- 35 Morrow J F, Berg P. Cleavage of Simian virus 40 DNA at a unique site by a bacterial restriction enzyme. *Proc Natl Acad Sci USA*, 1972, 69: 3365–3369
- 36 Jackson D A, Symons R H, Berg P. Biochemical method for inserting new genetic information into DNA of Simian Virus 40: Circular SV40 DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli*. *Proc Natl Acad Sci USA*, 1972, 69: 2904–2909
- 37 Frandsen R J. A guide to binary vectors and strategies for targeted genome modification in fungi using *Agrobacterium tumefaciens*-mediated transformation. *J Microbiol Methods*, 2011, 87: 247–262
- 38 Shrawat A K, Lorz H. *Agrobacterium*-mediated transformation of cereals: A promising approach crossing barriers. *Plant Biotechnol J*, 2006, 4: 575–603
- 39 Sheludko Y V. *Agrobacterium*-mediated transient expression as an approach to production of recombinant proteins in plants. *Recent Pat Biotechnol*, 2008, 2: 198–208
- 40 Sidorov V A, Kasten D, Pang S Z, et al. Stable chloroplast transformation in potato: Use of green fluorescent protein as a plastid marker. *Plant J*, 1999, 19: 209–216
- 41 Day A, Goldschmidt-Clermont M. The chloroplast transformation toolbox: Selectable markers and marker removal. *Plant Biotechnol J*, 2011, 9: 540–553
- 42 Krichevsky A, Zaltsman A, King L, et al. Expression of complete metabolic pathways in transgenic plants. *Biotechnol Genet Eng Rev*, 2012, 28: 1–13
- 43 Sundar I K, Sakthivel N. Advances in selectable marker genes for plant transformation. *J Plant Physiol*, 2008, 165: 1698–1716
- 44 Tuteja N, Verma S, Sahoo R K, et al. Recent advances in development of marker-free transgenic plants: Regulation and biosafety concern. *J Biosci*, 2012, 37: 167–197
- 45 Lee J H, Lee S Y. Selection of stable mutants from cultured rice anthers treated with ethyl methane sulfonic acid. *Plant Cell Tiss Organ Cult*, 2002, 71: 165–171
- 46 Proano V A, Greene G L. Endogenous gibberellins of a radiation induced single gene dwarf mutant of bean. *Plant Physiol*, 1968, 43: 613–618
- 47 Kumar A, Snyder M. Emerging technologies in yeast genomics. *Nat Rev Genet*, 2001, 2: 302–312
- 48 Alonso J M, Stepanova A N, Leisse T J, et al. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science*, 2003, 301: 653–657
- 49 Zhang J, Li C, Wu C, et al. RMD: A rice mutant database for functional analysis of the rice genome. *Nucleic Acids Res*, 2006, 34 (Suppl 1): D745–D748
- 50 Kuromori T, Takahashi S, Kondou Y, et al. Phenome analysis in plant species using loss-of-function and gain-of-function mutants. *Plant Cell Physiol*, 2009, 50: 1215–1231
- 51 Maloy S R, Hughes K T. Strain collections and genetic nomenclature. *Methods Enzymol*, 2007, 421: 3–8
- 52 Jin K, Li J, Vizeacoumar F S, et al. PhenoM: A database of morphological phenotypes caused by mutation of essential genes in *Saccharomyces cerevisiae*. *Nucleic Acids Res*, 2012, 40: D687–D694
- 53 Takeshita H, Sawa H. Asymmetric cortical and nuclear localizations of WRM-1/beta-catenin during asymmetric cell division in *C. elegans*. *Genes Dev*, 2005, 19: 1743–1748
- 54 Tweedie S, Ashburner M, Falls K, et al. FlyBase: Enhancing *Drosophila* Gene Ontology annotations. *Nucleic Acids Res*, 2009, 37 (Suppl 1): D555–D559
- 55 Sun L V, Jin K, Liu Y, et al. PBmice: An integrated database system of piggyBac (PB) insertional mutations and their characterizations in mice. *Nucleic Acids Res*, 2008, 36 (Suppl 1): D729–D734
- 56 Springer P S, McCombie W R, Sundaresan V, et al. Gene trap tagging of PROLIFERA, an essential MCM2-3-5-like gene in *Arabidopsis*. *Science*, 1995, 268: 877–880
- 57 Fang Q H, Zhong J J. Two-stage culture process for improved production of ganoderic acid by liquid fermentation of higher fungus *Ganoderma lucidum*. *Biotechnol Prog*, 2002, 18: 51–54
- 58 Talano M A, Oller A L, Gonzalez P S, et al. Hairy roots, their multiple applications and recent patents. *Recent Pat Biotechnol*, 2012, 6: 115–133
- 59 Lee W S, Hammond-Kosack K E, Kanyuka K. Barley stripe mosaic virus-mediated tools for investigating gene function in cereal plants and their pathogens: Virus-induced gene silencing, host-mediated gene silencing, and virus-mediated overexpression of heterologous protein. *Plant Physiol*, 2012, 160: 582–590
- 60 许亮, 卢向阳, 田云. 后基因组时代基因功能分析的策略. *中国生物工程杂志*, 2003, 23: 29–34
- 61 Senthil-Kumar M, Mysore K S. New dimensions for VIGS in plant functional genomics. *Trends Plant Sci*, 2011, 16: 656–665
- 62 Winzer T, Gazda V, He Z, et al. A *Papaver somniferum* 10-gene cluster for synthesis of the anticancer alkaloid noscapine. *Science*, 2012, 336: 1704–1708

- 63 Hatlestad G J, Sunnadeniya R M, Akhavan N A, et al. The beet *R* locus encodes a new cytochrome P450 required for red betalain production. *Nat Genet*, 2012, 44: 816–820
- 64 Tang Y, Wang F, Zhao J, et al. Virus-based microRNA expression for gene functional analysis in plants. *Plant Physiol*, 2010, 153: 632–641
- 65 Allemand I, Angulo J F. Transgenic and knock-out models for studying DNA repair. *Biochimie*, 1995, 77: 826–832
- 66 Bedell V M, Wang Y, Campbell J M, et al. *In vivo* genome editing using a high-efficiency TALEN system. *Nature*, 2012, 491: 114–118
- 67 Kim Y, Kweon J, Kim A, et al. A library of TAL effector nucleases spanning the human genome. *Nat Biotechnol*, 2013, 31: 251–258
- 68 Ding Q, Lee Y K, Schaefer E A, et al. A TALEN genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell*, 2013, 12: 238–251
- 69 Sung Y H, Baek I J, Kim D H, et al. Knockout mice created by TALEN-mediated gene targeting. *Nat Biotechnol*, 2013, 31: 23–24
- 70 Zu Y, Tong X, Wang Z, et al. TALEN-mediated precise genome modification by homologous recombination in zebrafish. *Nat Methods*, 2013, 10: 329–331
- 71 Lei Y, Guo X, Liu Y, et al. Efficient targeted gene disruption in *Xenopus* embryos using engineered transcription activator-like effector nucleases (TALENs). *Proc Natl Acad Sci USA*, 2012, 109: 17484–17489
- 72 Zhang Y, Zhang F, Li X, et al. Transcription activator-like effector nucleases enable efficient plant genome engineering. *Plant Physiol*, 2013, 161: 20–27
- 73 Li T, Liu B, Spalding M H, et al. High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat Biotechnol*, 2012, 30: 390–392
- 74 Qi X, Bakht S, Leggett M, et al. A gene cluster for secondary metabolism in oat: Implications for the evolution of metabolic diversity in plants. *Proc Natl Acad Sci USA*, 2004, 101: 8233–8238
- 75 Field B, Osbourn A E. Metabolic diversification—-independent assembly of operon-like gene clusters in different plants. *Science*, 2008, 320: 543–547
- 76 Gonzalez A, Zhao M, Leavitt J M, et al. Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in *Arabidopsis* seedlings. *Plant J*, 2008, 53: 814–827
- 77 Gou J Y, Felippes F F, Liu C J, et al. Negative regulation of anthocyanin biosynthesis in *Arabidopsis* by a miR156-targeted SPL transcription factor. *Plant Cell*, 2011, 23: 1512–1522
- 78 肖小河, 陈士林, 黄璐琦, 等. 中国道地药材研究 20 年概论. *中国中药杂志*, 2009, 34: 519–523
- 79 孟祥才, 陈士林, 王喜军. 论道地药材及栽培产地变迁. *中国中药杂志*, 2011, 36: 1687–1692

Research strategy for model medicinal species

XU Jiang¹, SUN Chao¹, XU ZhiChao¹, JI AiJia¹, HU YuanLei², SUN Wei³, WANG LiZhi¹, WANG Bo¹, YANG Pei¹, ZHANG Xin¹, SONG JingYuan¹ & CHEN ShiLin^{1,3}

¹Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100193, China;

²State Key Laboratory of Protein and Plant Gene Research, College of Life Sciences, Peking University, Beijing 100871, China;

³Institute of Chinese Materia Medica, Chinese Academy of Chinese Medical Science, Beijing 100700, China

Use of model organisms in biology research enables understanding of mechanisms and properties of biological systems. The investigation of model organisms has also led to the development of experimental techniques for different areas of biology. Species like *Ganoderma lucidum* and *Salvia miltiorrhiza* have been nominated for candidate model medicinal organisms. In this review, we discuss the concept of model medicinal organisms and define the aims of its research. Then, we analyze the feasibility of applying model organism research strategies to medicinal species. We also summarize the selection criteria for model medicinal organisms. Further, we introduce strategies for creating model medicinal organisms from different aspects including genome sequencing, genetic transformation platforms, and mutagenesis. Finally, we discuss the application of the model medicinal species system in gene function analysis, metabolic pathways, and biosynthetic metabolite analysis. We conclude that the model medicinal organism system will enhance understanding of the genetic basis of synthetic metabolites and accelerate the improvement of medicinal species research.

model organism, medicinal biology, secondary metabolites, genuine medicinal materials, synthetic biology

doi: 10.1360/972013-565