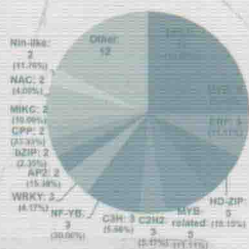
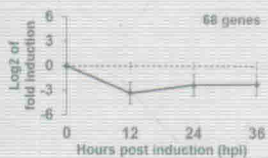


中药资源转录组分析 操作指南

主编 袁媛 黄璐琦



全书论述了中药资源转录组研究的现状和主要技术方法，并重点探讨了基于同源基因功能分化的药用植物活性成分变异机制，以及功能性EST-SSR标记的概念、优势、应用范围等；全面介绍了中药资源转录组的操作流程，包括转录组数据的组装、基因克隆及表达分析、数据库使用及生物信息学分析、SSR标记开发等；同时，附有药用植物与真菌转录组研究的鲜活参考范例。

上海科学技术出版社
SHANGHAI SCIENTIFIC & TECHNICAL PUBLISHERS

中药资源转录组分析 操作指南

主编 袁媛 黄璐琦



上海科学技术出版社
SHANGHAI SCIENTIFIC & TECHNICAL PUBLISHERS

图书在版编目(CIP)数据

中药资源转录组分析操作指南 / 袁媛, 黄璐琦主编.
—上海: 上海科学技术出版社, 2016. 5
ISBN 978-7-5478-3013-0

I. ①中… II. ①袁… ②黄… III. ①药用植物—基
因转录—指南 IV. ①R282.71-62

中国版本图书馆 CIP 数据核字(2016)第 043790 号

中药资源转录组分析操作指南

主编 袁媛 黄璐琦

上海世纪出版股份有限公司 出版
上海科学技术出版社

(上海钦州南路 71 号 邮政编码 200235)

上海世纪出版股份有限公司发行中心发行
200001 上海福建中路 193 号 www.ewen.co
苏州望电印刷有限公司印刷

开本 889×1194 1/32 印张: 10.25

字数 200 千字

2016 年 5 月第 1 版 2016 年 5 月第 1 次印刷

ISBN 978-7-5478-3013-0/R·1095

定价: 38.00 元

本书如有缺页、错装或坏损等严重质量问题, 请向工厂联系调换

内容提要

本书站在学科前沿,以原创的笔触系统介绍了中药资源转录组的理论、技术及研究方法。全书论述了中药资源转录组研究的现状和主要技术方法,并重点探讨了基于同源基因功能分化的药用植物活性成分变异机制,以及功能性 EST-SSR 标记的概念、优势、应用范围等;全面介绍了中药资源转录组的操作流程,包括转录组数据的组装、基因克隆及表达分析、数据库使用及生物信息学分析、SSR 标记开发等;同时,附有药用植物与真菌转录组研究的鲜活参考范例。

本书从理论、技术、实际应用的角度全面介绍了中药资源转录组这一业内关注的热点,同时提供了操作性强的实验指导。本书不仅可以作为医药院校生药专业、中医院校中药专业的本科生、研究生及教师的实验教学用书,也可作为相关科研人员的参考书。

编委会

主 编

袁 媛 (中国中医科学院中药资源中心)

黄璐琦 (中国中医科学院中药资源中心)

编 委

(按姓氏笔画排列)

马小军 (中国医学科学院药用植物研究所)

王绪敏 (中国科学院北京基因组研究所)

刘 越 (中央民族大学生命与环境科学学院)

杨 全 (广东药科大学中药学院)

张重义 (福建农林大学作物科学学院)

周俊辉 (中国中医科学院中药资源中心)

查良平 (中国中医科学院中药资源中心)

高 伟 (首都医科大学中医药学院)

唐金富 (中国中医科学院中药资源中心)

舒少华 (华中农业大学植物科学技术学院)

基因资源作为中药资源的重要组成,已逐渐成为行业研究关注的热点。在后基因组时代,转录组学作为最活跃学科之一,其为基因表达、结构和功能研究提供了技术手段,是保障中药产业现在及未来可持续发展的关键因素之一。

广义的转录组是指生物体细胞或组织在特定状态下所转录出来的所有 RNA 的总和。通过转录组分析,不但可以高通量地获得基因表达的 RNA 水平有关信息,而且可以揭示基因表达与生命现象之间的内在联系,从而表征生命体生理活动规律,确定其代谢特性。由于不同生长时间和生长环境下,同一生物体中基因表达情况并不完全相同,进行多器官、多时期、多维度、多水平的中药资源转录组研究,建立中药资源功能基因库,可为中药资源种质资源评价、新品种选育、中药材质量控制、活性成分的合成生物学等领域提供

分子基础,从而推动中药产业的可持续发展。具体来说,通过转录组可以对药用动植物基因转录的情况,不同个体之间的 SNP 差异、基因拷贝数差异、基因进化、次生代谢合成与调控、药材性状相关分子标记开发等进行研究,其目标包括基因功能注释,检测新的转录本、基因转录水平、非编码区域(ncRNA)功能、转录本结构变异,SSR 和 SNPs 标记开发等。

2009—2014 年,黄花蒿、丹参、西洋参、淫羊藿、蝎子、大麻、紫杉、狼毒、罗汉果、南方红豆杉、三七、柴胡、水蛭、虎杖、大蒜、贯叶连翘、红花、胡黄连、沙冬青、何首乌、白木香、金银花、牡丹、蛹虫草、灵芝、贯叶连翘、红花、沙棘、中华鳖、地黄、毛地黄、土鳖虫、酸浆果、牡丹、黄花蒿、乌拉尔甘草、长春花、沙棘、枸杞、萝卜、茯苓、西洋参、冬瓜、魔芋、博落回、七叶一枝花、铁皮石斛、人参、积雪草、冬虫夏草、鱼腥草等药用植物均已通过 RNA-seq 进行了转录组研究,证明 RNA-seq 在发现药用动植物转录组基因和遗传标记中非常有效。

另一方面,完成转录组测序仅是功能基因组研究的第一步,更大的挑战在于搞清基因序列的遗传信息与其所执行的生物学功能,对直向同源基因和共生同源基因进行正确的功能注释,建立中药资源基因数据库。在基因进化理论和方法的指导下,对药用动植物重复基因功能分化进行分析,并与活性成分种类和数量变化相结合,从分子生药学角度揭示药用动植物物

种、部位、产地间活性成分差异的机制,为指导药材质量控制提供理论依据。

本书是继《中药分子鉴定操作指南》后推出的分子生药学系列第二本实验操作指南,主要包括4个部分:第一部分为转录组数据的组装,主要介绍如何获得基因数据,涉及7个实验操作规程,如样品收集和保存(实验一)、RNA提取(实验二)、mRNA分离纯化(实验三)、Solexa文库制备(实验四)、上机测序(实验五)、转录组数据组装(实验六)、转录组数据功能注释(实验七);第二部分为基因克隆及表达分析,主要介绍如何筛选与目的性状相关的候选基因,涉及3个实验操作规程,如基因克隆(实验八)、基因表达分析(实验九)、共表达分析(实验十);第三部分为数据库使用及生物信息学分析,主要介绍利用生物信息学手段对候选基因的结构和特征进行预测,涉及核酸数据库(实验十一)、蛋白质数据库(实验十二)、代谢数据库(实验十三)、文献数据库(实验十四)及其使用,以及9个实验操作规程,如NCBI序列数据的提交(实验十五)、多序列对位排列分析(实验十六)、系统发育分析(实验十七)、核酸序列分析和转换(实验十八)、基因开放阅读框预测(实验十九)、蛋白质理化性质预测(实验二十)、蛋白质亚细胞定位信号预测(实验二十一)、蛋白质信号肽和跨膜结构域预测(实验二十二)、蛋白质二级结构和三级结构预测(实验二十三);第四部分为SSR标记开发,涉及7个实验操作规程,如SSR位点开发(实

验二十四)、SSR 引物设计(实验二十五)、基因组 DNA 提取(实验二十六)、琼脂糖凝胶电泳(实验二十七)、非变性聚丙烯酰胺凝胶电泳(实验二十八)、SSR 数据统计(实验二十九)、SSR 数据分析(实验三十)。此外,本书还介绍了药用植物和真菌转录组研究方面的实例。

希冀本书的出版,能为从事中药学研究的科研人员和在校师生提供参考。

编者

2016年3月

第一篇 理论篇

一、中药资源转录组概述	3
(一) 转录组的研究目标	3
(二) 转录组的研究方法	5
(三) 中药资源转录组研究现状	5
(四) 中药资源转录组研究展望	12
二、同源基因功能分化与活性成分变异机制	22
(一) 同源基因的正确注释是研究药用植物活性成分 变异的重要前提	22
(二) 重复基因的功能分化是形成药用植物活性成分 变异的基础	24
(三) 药用植物基因功能分化决定其活性成分的变异	26
三、功能性 EST - SSR 标记	31

(一) 功能性 EST - SSR 的基本概念	31
(二) 功能性 EST - SSR 的开发流程	32
(三) 功能 EST - SSR 的优势	35
(四) 功能性 EST - SSR 标记在中药资源中的应用	36
(五) 药用植物功能性 EST - SSR 标记展望	38

第二篇 技术篇

一、转录组数据的组装	43
实验一 样品收集与保存	43
实验二 RNA 提取	44
实验三 mRNA 分离纯化	47
实验四 Solexa 文库制备	49
实验五 上机测序	53
实验六 转录组数据组装	59
实验七 转录组数据功能注释	61
二、基因克隆及表达分析	64
实验八 基因克隆	64
实验九 基因表达分析	71
实验十 共表达分析	80
三、数据库使用及生物信息学分析	84
实验十一 核酸数据库及其使用	84
实验十二 蛋白质数据库及其使用	91
实验十三 代谢数据库及其使用	96
实验十四 文献数据库及其使用	101

实验十五	NCBI 序列数据的提交	104
实验十六	多序列对位排列分析	109
实验十七	系统发育分析	112
实验十八	核酸序列的分析和转换	117
实验十九	基因开放阅读框预测	123
实验二十	蛋白质理化性质预测	129
实验二十一	蛋白质亚细胞定位信号预测	130
实验二十二	蛋白质信号肽和跨膜结构域预测	132
实验二十三	蛋白质二级结构和三级结构预测	135
四、SSR 标记开发		139
实验二十四	SSR 位点开发	139
实验二十五	SSR 引物设计	140
实验二十六	基因组 DNA 提取	143
实验二十七	琼脂糖凝胶电泳	145
实验二十八	非变性聚丙烯酰胺凝胶电泳	147
实验二十九	SSR 数据统计	149
实验三十	SSR 数据分析	155

第三篇 实践篇

一、金银花转录组与质量评价标记筛选	161
二、地黄根部转录组文库的构建及响应连作障碍关键基因的 筛选	174
三、比较转录组学和代谢组学相结合研究丹参酮生物合成分 子机制	186

四、罗汉果果实转录组测序的研究	209
五、基于转录组测序揭示适度干旱胁迫对甘草根基因表达的 调控	222
六、基于黄芩 cDNA 文库的功能 EST - SSR 标记开发	236
七、茯苓转录组 SSR 序列特征及其功能预测	255
八、川西獐牙菜(藏茵陈)转录组 SSR 信息分析	268
九、厚朴 MVA 途径相关基因鉴定及生物信息学分析	288
十、三七香叶基香叶基焦磷酸合酶基因的克隆及表达分析	301

第一篇 理论篇



一、中药资源转录组概述

近年来,中药资源研究进展飞快,为医药产业提供了巨大的新药创制资源。同时,其基因资源研究也逐渐成为行业关注的热点。在后基因组时代,转录组学作为最活跃学科之一,其为基因的结构、表达和功能研究提供了技术手段,是保障中药产业现在及未来可持续发展的关键因素之一。广义的转录组是指生物体细胞或组织在特定状态下所转录出来的所有 RNA 的总和, RNA 包括了编码蛋白质的 RNA(如 mRNA)和非编码蛋白质的 RNA(如 rRNA、tRNA、microRNA 等)。而狭义的转录组通常指所有编码蛋白质的 mRNA 的总和。

作为连接基因组与蛋白质组的纽带,转录组具有时间性和空间性,即转录组反映的是除了异常 mRNA 降解现象以外,特定条件下活跃表达的基因。通过转录组分析,不但可能高通量地获得基因表达的 RNA 水平有关信息,而且可以揭示基因表达与生命现象之间的内在联系,从而表征生命体生理活动规律,确定其代谢特性。由于不同生长时间和生长环境下,同一生物体中基因表达情况并不完全相同,进行多器官、多时期、多维度、多水平的中药资源转录组研究,建立中药资源功能基因库,可为中药资源种质资源评价、新品种选育、中药材质量控制、活性成分的合成生物学等领域提供分子基础,从而推动中药产业的可持续发展。

(一) 转录组的研究目标

对于中药资源而言,通过转录组可以对中药资源基因转录的情况,不同个体之间的 SNP 差异、基因拷贝数差异,基因进化,次生代谢合成与调控,药材性状相关分子标记开发等进行研究,具体

目标包括如下几方面。

(1) 基因功能注释: 将所测 reads 与已有数据库, 如 GO、KEGG 等已注释功能的基因相比对分析, 从而揭示特定转录状态下的基因的功能和生物通路等。

(2) 检测新的转录本: 利用单碱基分辨率的转录组测序技术可极大地丰富基因注释的多个方面, 包括编码区边界鉴定、UTRs 区域鉴定以及新的转录区域鉴定等, 且在低丰度转录本发现中具有很强的技术优越性。

(3) 基因转录水平: 由于外界刺激或环境变化时生物体中基因表达水平的变化是微小的, 转录组测序技术可以定量、准确地确定 RNA 的表达水平, 有可能确定细胞群中每一个分子的绝对数量, 并对实验之间的结果进行直接比较。

(4) 非编码区域(ncRNA)功能: 如转录组学中 microRNA 的研究, 经历了从首个 microRNA——*lin-4* 的发现^[1], 到大规模 microRNA 转录组的测定, 再到目前 microRNA 通过基因沉默方式调节靶基因表达的机制研究^[2,3], 尤其是大规模转录组测序后, 大量的 microRNA 被发现, 与之相关的功能性研究正如火如荼地展开。多年来一直被认为是基因组中垃圾成分的非编码 RNA 终于越来越得到人们的重视。研究发现, microRNA 对基因表达和生长发育起到了重要的调节作用。

(5) 转录本结构变异: 转录本结构的变异能揭示基因转录后表达的多样性。可变剪接使一个基因产生多个 mRNA 转录本, 从而翻译成不同的蛋白。RNA 编辑主要通过碱基的替换或转换等使基因序列发生改变。

(6) SSR 和 SNPs 标记开发: 通过比对转录本和参考基因组间的序列, 寻找潜在的 SNPs 或 SSRs。基于转录组数据开发微卫星标记无需建库、克隆及筛选, 不仅可以避免克隆偏差以及无效克隆造成微卫星丢失而导致微卫星位点分离较少的问题, 而且还提