

• 药材与资源 •

基于流式细胞术和基因组 survey 分析的地黄基因组研究

赵 乐^{1,2}, 朱昀昊^{1,2}, 王 敏³, 韩永光¹, 马利刚^{1,2}, 冯卫生^{1,2}, 郑晓珂^{1,2*}

1. 河南中医药大学药学院, 河南 郑州 450046

2. 呼吸疾病诊疗与新药研发河南省协同创新中心, 河南 郑州 450046

3. 北京工商大学 化学与材料工程学院, 北京 100048

摘要:目的 地黄 *Rehmannia glutinosa* 是我国传统中药, 具有较高的药用价值和经济价值。利用流式细胞技术和高通量测序技术, 分析了地黄基因组大小和特征等信息, 为绘制地黄全基因组精细图谱提供依据。方法 将已知基因组大小的大豆 *Glycine max* 和辣椒 *Capsicum annuum* 作为对照, 用流式细胞技术估算地黄基因组大小。然后利用高通量测序技术对地黄基因组进行 survey 分析, 通过生物信息学分析获得了地黄全基因组重复序列比例、基因组杂合度以及 GC 含量等信息。结果 根据流式细胞实验结果, 地黄基因组大小介于大豆和辣椒之间, 估算地黄基因组大小应在 2.00~2.12 Gb。通过高通量测序, 获得了 132.79 Gb 高质量的数据, 总测序深度约为 66.40×, 估算地黄基因组大小为 2.03 Gb。根据 Kmer 分布情况, 估算地黄基因组重复序列比例为 78.48%, 杂合度为 1.93%, GC 含量为 37.27%。从基因组基本结构特征上看, 地黄基因组属于高重复、高杂合、大基因组的复杂基因组。结论 获得了地黄基因组大小和特征等信息, 这些信息将为绘制地黄全基因组的精细图谱奠定基础, 也为阐明地黄有效成分的生物合成和调控途径提供依据。

关键词: 地黄; 流式细胞术; 高通量测序; 基因组大小; 重复序列; 杂合度

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2021)03-0821-06

DOI: 10.7501/j.issn.0253-2670.2021.03.025

Estimation of *Rehmannia glutinosa* genome size based on flow cytometry and genome survey analysis

ZHAO Le^{1,2}, ZHU Yun-hao^{1,2}, WANG Min³, HAN Yong-guang¹, MA Li-gang^{1,2}, FENG Wei-sheng^{1,2}, ZHENG Xiao-ke^{1,2}

1. School of Pharmacy, Henan University of Chinese Medicine, Zhengzhou 450046, China

2. Collaborative Innovation Center for Respiratory Disease Diagnosis and Treatment & Chinese Medicine Development of Henan Province, Zhengzhou 450046, China

3. College of Chemistry and Material Engineering, Beijing Technology and Business University, Beijing 100048, China

Abstract: Objective *Rehmannia glutinosa* is a traditional Chinese medicine with high medicinal and economic value. In this study, flow cytometry and high-throughput sequencing technology were used to analyze the size and characteristics of *R. glutinosa* genome so as to provide a basis for the detailed mapping of the whole genome of *R. glutinosa* in the future. **Methods** The genome size of *R. glutinosa* was estimated by flow cytometry using soybean (*Glycine max*) and pepper (*Capsicum annuum*) of known genome size as controls. Then, high-throughput sequencing technology was used to carry out survey analysis on the genome of *R. glutinosa*, and information such as the proportion of repetitive sequences in the whole genome of *R. glutinosa*, genomic heterozygosity, and GC content were obtained through bioinformatics analysis. **Results** According to the results of flow cytometry, the genome size of *R. glutinosa* was between soybean and pepper, and it was estimated that the genome size of *R. glutinosa* was between 2.00 and 2.12 Gb. Through high-throughput sequencing, 132.79 Gb high-quality data were obtained, with a total sequencing depth of about 66.40×, and

收稿日期: 2020-09-02

基金项目: 国家重点研发计划 (2017YFC1702800); 国家自然科学基金资助项目 (81603232); 河南省重大科技专项 (171100310500); 中央引导地方科技发展专项资金项目; 河南省科技攻关项目 (172102110099); 国家级大学生创新创业训练计划项目 (201710471022)

作者简介: 赵 乐 (1983—), 男, 博士, 副教授, 主要从事药用植物分子生物学研究。E-mail: zhaole1983@126.com

*通信作者: 郑晓珂, 女, 博士, 教授, 主要从事中药活性成分及作用机制研究。E-mail: zhengxk.2006@163.com

the genome size of *R. glutinosa* was estimated to be 2.03 Gb. According to the distribution of Kmer, the proportion of repetitive sequences, the heterozygosity and the GC content in the whole genome of *R. glutinosa* were estimated to be 78.48%, 1.93% and 37.27%, respectively. From the perspective of the basic structural features of the genome, the genome of *R. glutinosa* was a large and complex genome with high heterozygosity and high repetitiveness. **Conclusion** In this study, the size and characteristics of *R. glutinosa* genome were obtained, which will lay an important foundation for detailed mapping of the whole genome of *R. glutinosa* in the future, and also provide a basis for elucidating the biosynthesis and regulatory pathways of active components in *R. glutinosa*.

Key words: *Rehmannia glutinosa* Libosch; flow cytometry; high-throughput sequencing; genome size; repetitive sequences; heterozygosity

地黄 *Rehmannia glutinosa* Libosch 是玄参科地黄属多年生草本植物, 为《中国药典》2015 年版药材地黄的唯一来源植物^[1], 是我国传统的大宗中药材, 具有很高的药用价值和经济价值。地黄以其块根入药, 块根中含有环烯醚萜、黄酮、木脂素及多糖等多种化合物^[2]。地黄药理活性显著, 对心脑血管系统、神经系统、免疫系统均有一定保护作用^[3]。梓醇是地黄的主要活性成分, 从化学结构分析, 梓醇属于环烯醚萜苷类化合物, 具有降血糖、神经保护、抗炎、抗氧化等多种生物学活性^[4]。Yan 等^[5]发现梓醇能够缓解 2 型糖尿病中的胰岛素抵抗, 起到降血糖的作用, 还能减缓糖尿病晚期的并发症。此外, 梓醇还可以改善小鼠阿尔茨海默症的病理特征, 同时改善相关的行为障碍^[6-7]。

目前关于地黄的研究, 主要集中于化学成分的分离、梓醇和地黄其他活性成分的药理作用、以及地黄连作障碍等方面, 关于地黄基因组方面的研究报道^[8]较少。Zeng 等^[9]对地黄属 6 种地黄植物的叶绿体基因组进行了测序, 同时进行了系统发育关系分析, 发现地黄叶绿体基因组大小为 153 622 bp, 而且地黄属 6 种地黄的叶绿体基因组在基因结构方面高度一致。闫坤等^[10]通过观察地黄属 6 种地黄的小孢子母细胞减数分裂, 发现地黄和茄叶地黄属于四倍体植物 ($2n=56$), 天目地黄、高地黄、裂叶地黄和湖北地黄属于二倍体植物 ($2n=28$), 说明地黄属植物存在多倍体现象。这种多倍体现象可能是由于全基因组复制形成的, 王姣等^[11]利用系统发育生物学方法, 通过分析地黄属 5 种地黄的转录组数据, 发现地黄属植物经历了 2 次全基因组复制, 加速了地黄属植物的物种分化, 通过二倍体祖先的基因组加倍形成了四倍体植物。闫婧^[12]用二代高通量测序对地黄属的天目地黄基因组进行了 survey 分析, 发现天目地黄基因组大小估算为 1.14 Gb, GC 含量为

40.64%, 重复序列比例为 76.94%, 杂合度为 0.34%, 从基因组基本结构特征分析, 天目地黄基因组属于高重复、低杂合的大型基因组。

地黄的许多栽培种在长期的栽培过程中受到了选择, 这些栽培品种之间的遗传关系不清楚, 而且这些栽培品种在叶和花的颜色和形态、以及根的性状方面变异较大^[13]。由于地黄栽培品种之间的遗传关系不清楚, 因此获得地黄的全基因组精细图谱对于研究其主要药用成分—梓醇的生物合成及代谢调控途径具有重要的意义, 同时也为研究地黄株高、叶形、块根膨大等性状变异奠定了基础。根据报道地黄为四倍体 ($2n=56$)^[10,13], 本研究以地黄栽培品种“沁怀 1 号”为材料, 通过二代高通量测序技术对地黄基因组进行 survey 分析, 同时结合流式细胞技术来评估地黄基因组大小, 最终获得地黄全基因组重复序列比例、基因组杂合度以及 GC 含量等信息。这些信息将为今后绘制地黄全基因组的精细图谱提供依据, 也为阐明梓醇的生物合成及代谢调控途径奠定基础。

1 材料

地黄所用材料为地黄栽培品种“沁怀 1 号”组培苗, 经河南中医药大学董诚明教授鉴定为地黄 *Rehmannia glutinosa* Libosch, 以已知基因组大小的大豆组培苗(中国农业科学院)和辣椒组培苗(中国农业科学院)作为流式细胞技术测定的对照样品, 采集这 3 种植物组培苗新鲜的叶片, 经液氮速冻后, 放置于 -80°C 超低温冰箱保存。

BD FACSCalibur 流式细胞仪(美国 BD 公司), 离心机 5810R(德国 Eppendorf 公司); 碘化丙啶(propidium iodide, PI)批号 340242, 购自美国 BD 公司; 吐温 20(批号 P2287)购自美国 Sigma-Aldrich 公司。Otto I 缓冲液: 0.1 mol/L 柠檬酸、0.5%聚山梨酯 20, 用 0.22 μm 滤器滤过后 4°C 保存。Otto II 缓冲液: 0.4 mol/L $\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$, 用 0.22 μm 滤器滤过后室温 ($18\sim 25^{\circ}\text{C}$) 保存。

2 方法

2.1 细胞核悬液制备及基因组大小测定

根据 Doležal 等^[14]的方法, 各取地黄、大豆和辣椒组培苗新鲜叶片 100 mg, 放入培养皿中, 加入 1 mL 预冷的 Otto I 缓冲液, 用锋利的刀片将组织切碎, 用移液器上下吹打几次混合均匀 (注意避免气泡), 然后将匀浆液通过 42 μm 尼龙筛网滤过到离心管中, 4 ℃、1000 r/min 离心 5 min, 小心的弃去上清, 收集细胞核沉淀。再用 1 mL Otto II 缓冲液重悬细胞, 然后加入 DNA 荧光染料碘化丙啶 (PI) 和 RNA 酶, 使二者的终浓度均在 50 μmol/mL, 置于冰上, 避光染色 20 min, 再将地黄、大豆和辣椒的染液混合形成混合样品, 然后上机测定这 3 种植物的单独样品和混合样品。

将已知基因组大小的大豆和辣椒作为对照样品, 首先用流式细胞仪单独测定地黄、大豆和辣椒基因组 DNA 的相对荧光强度, 然后检测混合样品的相对荧光强度, 并根据不同样品相对荧光强度峰值的大小, 同时参考对照样品的基因组, 计算地黄的基因组大小。用流式细胞仪测定时, 每个样品收集 10 000 个颗粒, 检测 PI 的荧光强度, 收集 FL2 通道的荧光, 使用 FACSCComp 软件对数据进行分析, 并通过 FL2-A 通道计算得到每个样品 G1 期的荧光强度^[15]。

2.2 地黄基因组 survey 分析

采用改良 CTAB 法提取地黄叶片的基因组 DNA, 用紫外分光光度计和琼脂糖凝胶电泳分别检测地黄基因组 DNA 的浓度、纯度和完整性。

检测合格后, 用超声波震荡方法将地黄基因组 DNA 破碎至目的片段 (350 bp), 然后经过末端修复、加 A、加接头、目标片段选择和 PCR 等步骤构建小片段测序文库; 用安捷伦 2100 和定量 PCR

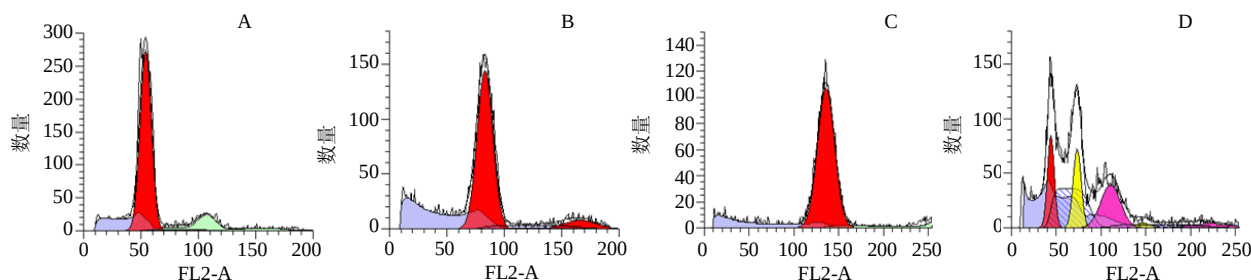
检测文库片段大小和文库定量, 确定文库是否符合测序标准; 通过桥式 PCR 将文库固定到测序芯片上; 利用北京百迈客 Illumina Hiseq 4000 测序仪对构建好的文库进行双端 150 bp (PE 150) 测序, 测序所产生的数据经过质控后用于下一步信息分析。双端测序数据经过评估 (GC 分布统计、质量值 Q₂₀、Q₃₀ 评估)、过滤后得到高质量的数据, 用于地黄基因组大小的评估、基因组的组装、GC 含量的统计、杂合度的统计 (以及组装后的评估)。根据地黄基因组大小, 结合流式细胞技术测定结果来估算测序深度。

Kmer 是从测序数据中提取出的长度为 K 的寡聚核苷酸序列, 本研究中取 K=19 进行分析, 在测序 reads 均匀分布的前提下, 根据公式, 计算基因组大小 (基因组大小=总碱基数/平均测序深度=总 Kmer 数/平均 Kmer 深度)。由于测序片段是随机打断的, 标准的 Kmer 深度分布曲线呈正态分布, 根据实际曲线偏离正态分布的程度, 通过 Jellyfish 和 GenomeScope 软件^[16]对测序数据进行分析, 统计产生 Kmer 频数分布数据, 再使用该数据进行拟合作图, 得到 Kmer 分布图, 进一步评估地黄基因组杂合度和重复序列比例。

3 结果与分析

3.1 流式细胞技术测定地黄基因组大小

以已知基因组大小的物种为对照品: 大豆 (基因组大小为 1.1 Gb)^[17-18]和辣椒 (基因组大小为 3.2 Gb)^[19], 进行细胞流式实验, 来估算地黄基因组的大小。首先用流式细胞技术将大豆、地黄和辣椒单独测定, 检测每个样品的相对荧光强度 (图 1-A~C), 然后采用内标法, 以大豆和辣椒为对照品, 用流式细胞技术对混合样品进行相对荧光强度的检测, 大豆、地黄和辣椒混合样品的基因组



A-大豆单独检测 B-地黄单独检测 C-辣椒单独检测 D-大豆、辣椒、地黄混合样品 (红色为大豆, 黄色为地黄, 粉色为辣椒)

A-Individual detection of *G. max* B-Individual detection of *R. glutinosa* C-Individual detection of *C. annuum* D-Mixed detection of *G. max*, *R. glutinosa*, *C. annuum* (red represents *G. max*, yellow represents for *R. glutinosa*, and pink represents *C. annuum*)

图 1 不同样品流式细胞术检测结果

Fig. 1 Peak value image of different species

DNA 相对荧光强度峰值分别为 41.99、73.25、110.61 (图 1-D)。根据流式细胞实验结果,地黄基因组大小介于大豆和辣椒之间,估算地黄基因组大小应在 2.00~2.12 Gb。

3.2 测序结果统计

使用地黄“沁怀 1 号”样品的基因组 DNA 构建 350 bp 文库,在 Illumina Hiseq 4000 测序平台测序并过滤得到 132.79 Gb 高质量的数据,总测序深度约为 66.40×,测序数据 Q₂₀ 比例均在 97.15% 以上, Q₃₀ 比例均在 92.86% 以上,满足 60× 以上的测序数据量。

3.3 样品污染评估

如果地黄基因组 DNA 样品存在污染不仅会降低有效数据量,同时还会影响基因组 survey 分析结果的准确性,导致基因组大小、重复序列比例、杂合度和 GC 含量等基因组特征评估结果出现较大偏差,使得基因组组装建库策略出现偏差,最终影响后续的基因组组装效果。为了判断提取的地黄基因组 DNA 是否受到污染,从测序得到的 350 bp 文库中,随机取 10 000 条单端 reads,与 NT 库 (Nucleotide Sequence Database) 进行 BLAST 比对。BLAST 使用 ncbi-blast+2.2.29 版本,参数设置为“-num_descriptions 100-num_alignments 100-evalue 1×10⁻⁵”。能够比对上 NT 库的 reads 占提取 reads 数的 8.74%,其中比对到芝麻 *Sesamum indicum* L.、宽叶沟酸浆 *Erythranthe guttatus* DC.、可可 *Theobroma cacao* L.、番茄 *Solanum lycopersicum* L. 和地黄的 reads 数分别占比对上 NT 库 reads 数的 12.47%、4.57%、4.11%、3.08% 和 2.17%。芝麻是地黄的近缘物种,于 2014 年完成芝麻基因组测序^[20],在 NT 库中基因注释信息丰富,所以在比对物种所占比例中最高为 12.47%,在前期地黄基因的进化分析中也发现地黄和芝麻的亲缘关系较近^[21-24],而地黄基因组信息未知,在 NT 库中基因注释极少,所以在比对物种所占比例中仅为 2.17%。比对结果中未发现动物、微生物等异常比对,因此该地黄基因组 DNA 样品测序数据不存在污染,可用于基因组 survey 分析。如果有一定比例的 reads 比对到进化距离较远的物种如动物、微生物等,则判断样品可能存在污染,需要进一步检查原因,不能用于基因组 survey 分析。

3.4 核外 DNA 含量评估

由于细胞核外也存在核酸序列,如果测序文

库中核外 DNA 含量过高,会加大后期基因组组装的难度,甚至产生错误。因此评估文库中核外 DNA 含量对判断数据能否用于后续基因组组装非常必要。为了评估测序数据中核外 DNA 的含量,利用 Illumina Hiseq 测序得到的 350 bp 文库与地黄的叶绿体序列 (NC_034308.1, 153 622 bp) 进行 SOAP 比对^[25]。SOAP 使用 2.21 版本,参数设置为“-m 260-x 440”。双端比上的 reads 数为 17 415 302,占总 reads 的 1.96%,单端比上的 reads 数为 11 313 201,占总 reads 的 1.28%,这 2 个的比例都低于经验值 5%。由此判断测序数据的核外 DNA 含量很低,不影响基因组 survey 分析的准确性和后期基因组的组装。

3.5 地黄基因组大小、重复序列比例和杂合度评估

使用 350 bp 文库数据构建 K=19 的 Kmer 分布图 (图 2),进行地黄基因组大小、重复序列比例和杂合度的评估。由图 2 可观察到明显的 3 个峰,第 1 个峰的 Kmer 深度分布为 27,第 2 个峰的 Kmer 深度分布为 56,第 3 个峰的 Kmer 深度分布为 102,这 3 个峰的 Kmer 深度分布符合 1:2:4 的趋势,推测地黄为四倍体。假设地黄基因组为“AABB”,则代表“AA”一致,“BB”一致,但“AA”“BB”不同,根据地黄基因组呈现的 Kmer 分布,从测序数据中得到的总 Kmer 数为 116 809 179 227 个,去除深度异常的 Kmer 后,共 113 026 782 538 个 Kmer 用于基因组大小估计,用 Jellyfish 和 GenomeScope 软件计算得到的基因组大小约 2.03 Gb,与流式细胞技术测定结果一致。根据 Kmer 分布情况,估计地黄基因组重复序列比例约 78.48%,杂合度约 1.93%。

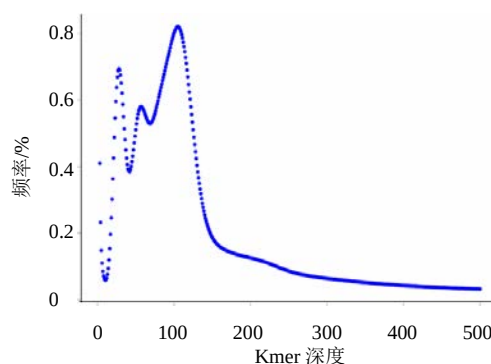


图 2 Kmer 分布图预估地黄基因组大小

Fig. 2 Kmer analysis for estimating genome size of *R. glutinosa*

基因组 GC 含量对二代基因组测序的随机性有较大影响, 过高 (>65%) 或过低 (<25%) 的 GC 含量会导致测序偏向性, 严重影响基因组分析结果。物种 GC 含量是评估基因组 survey 分析准确性和后续基因组组装难度的重要指标之一, 通过对 350 bp 文库测序数据分析, 地黄基因组的 GC 含量约 37.27%, 较为适中, 不会影响分析的准确性。

综上所述, 地黄基因组大小约 2.03 Gb, 重复序列比例约 78.48%, 杂合度约 1.93%, 基因组的 GC 含量约 37.27%, 从基因组基本结构特征分析, 地黄基因组属于高重复、高杂合、大基因组的复杂基因组。

4 讨论

基因组是一个生物体中全部的 DNA 信息, 包括其所有基因, 通过对药用植物的全基因组进行测序, 获得药用植物基因组的 DNA 序列信息, 就能够从分子水平研究活性成分的生物合成和代谢调控途径, 以及药用植物之间的进化关系。随着高通量测序技术的迅速发展, 目前已有几十种药用植物的全基因组被解析, 如大麻、灵芝、铁皮石斛、丹参、三七、罂粟、黄花蒿等重要的药用植物^[26]。这些药用植物全基因组序列信息的发布, 为药用植物活性成分生物合成和代谢调控途径的研究奠定了坚实的基础, 极大的推动了药用植物分子生物学的研究水平, 也为药用植物的遗传育种提供了重要依据^[27]。

由于流式细胞技术方便、快捷而且结果可靠, 是测定植物基因大小的首选方法^[14], 目前应用流式细胞技术, 测定了黄芪^[28]、柴胡^[29]和人参^[30]等药用植物基因组的大小。除了流式细胞技术以外, 利用高通量测序技术对植物基因组进行 survey 分析, 也是检测植物基因组大小和特征的有效方法。通过对基因组进行 survey 分析, 提前获得基因组的大小、重复序列、GC 含量等信息, 为后续进行全基因组深度测序时选择合适的测序策略提供依据。目前利用高通量测序方法, 已对罗汉果^[31]、马鞍藤^[32]、三岛柴胡^[33]等药用植物的基因组进行 survey 分析, 获得了这些植物基因组大小、复杂程度以及基因组的其他特征等信息。

玄参科有 4500 种植物, 包括多种药用植物, 如地黄、玄参、洋地黄、胡黄连、阴行草等, 具有重要的药用价值和经济价值^[34], 但是对于这些药用植物进行遗传育种和基础研究的主要障碍在于缺乏参

考基因组序列, 截至目前还没有玄参科植物的参考基因组发布。准确测定地黄基因组大小是后续进行地黄全基因组深度测序和遗传分析的基础, 目前在许多基因组测序项目中都使用高通量测序的 Kmer 分析估算基因组大小^[35]。本研究应用流式细胞技术估算地黄基因组大小应在 2.00~2.12 Gb, 同时结合高通量测序的 Kmer 分析, 地黄基因组大小修正为 2.03 Gb, 而且通过 Kmer 分析还能获得地黄基因组的重复序列 78.48%、杂合度 1.93%、GC 含量 37.27% 等基因组特征信息。通过将流式细胞技术和基因组 survey 分析 2 种方法结合在一起, 对地黄基因组大小和特征进行分析, 提高了实验结果的可靠性。从基因组基本结构特征上看, 地黄基因组属于高重复、高杂合、大基因组的复杂基因组。由于二代高通量测序技术限制, 测序读长太短, 仅有 200~500 bp, 对于地黄这种大型复杂基因组的组装难度大、花费多, 而从头组装 *de novo* 组装对算法的要求更高, 所以针对地黄基因组杂合度高、重复序列比例高、基因组大的特点, 可采用 3 代高通量测序技术 (PacBio Sequel) 同时结合染色质区域捕获 (chromosome conformation capture, Hi-C) 技术来完善三代高通量测序基因组组装结果, 将其组装到染色体水平^[36]。PacBio Sequel 测序平均长度达 12~15 kb, 能够跨越基因组高重复区域和高复杂区域, 减少拼接成本, 同时结合 Hi-C 技术, 能够获得高质量染色体水平的基因组序列, 可以阐明物种间基因组的演化历程, 揭示物种间的进化关系。本研究通过流式细胞技术和基因组 survey 分析, 获得的地黄基因组大小和特征等信息, 为今后绘制地黄全基因组的精细图谱奠定了基础, 也为研究地黄属植物的多倍体现象以及遗传进化关系提供依据。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 中国药典 [S]. 一部. 2015: 124-125.
- [2] Zhang R X, Li M X, Jia Z P. *Rehmannia glutinosa*: Review of botany, chemistry and pharmacology [J]. *J Ethnopharmacol*, 2008, 117(2): 199-214.
- [3] Wang F Q, Zhi J Y, Zhang Z Y, et al. Transcriptome analysis of salicylic acid treatment in *Rehmannia glutinosa* hairy roots using RNA-seq technique for identification of genes involved in acteoside biosynthesis [J]. *Front Plant Sci*, 2017, 8: 787.
- [4] Zhi J Y, Li Y J, Zhang Z Y, et al. Molecular regulation of catalpol and acteoside accumulation in radial striation and

- non-radial striation of *Rehmannia glutinosa* tuberous root [J]. *Int J Mol Sci*, 2018, 19(12): E3751.
- [5] Yan J T, Wang C Y, Jin Y, *et al.* Catalpol ameliorates hepatic insulin resistance in type 2 diabetes through acting on AMPK/NOX4/PI3K/AKT pathway [J]. *Pharmacol Res*, 2018, 130: 466-480.
- [6] Wan D, Xue L J, Zhu H F, *et al.* Catalpol induces neuroprotection and prevents memory dysfunction through the cholinergic system and BDNF [J]. *Evid Based Complement Alternat Med*, 2013, 2013: 134852.
- [7] Howes M R, Fang R, Houghton P J. Effect of Chinese herbal medicine on Alzheimer's disease [J]. *Int Rev Neurobiol*, 2017, 135: 29-56.
- [8] Tian Y H, Feng F J, Zhang B, *et al.* Transcriptome analysis reveals metabolic alteration due to consecutive monoculture and abiotic stress stimuli in *Rehmannia glutinosa* Libosch [J]. *Plant Cell Rep*, 2017, 36(6): 859-875.
- [9] Zeng S Y, Zhou T, Han K, *et al.* The complete chloroplast genome sequences of six *Rehmannia* species [J]. *Genes (Basel)*, 2017, 8(3): E103.
- [10] 闫坤, 赵楠, 李宏庆. 地黄属种间亲缘关系研究 [J]. 西北植物学报, 2007, 27(6): 1112-1120.
- [11] 王姣. 地黄属植物全基因组复制与物种分化研究 [D]. 西安: 西北大学, 2019.
- [12] 闫婧. 珍稀植物天目地黄 (*Rehmannia chingii*) 的基因组调查研究 [D]. 西安: 西北大学, 2018.
- [13] Qi J J, Li X N, Song J Y, *et al.* Genetic relationships among *Rehmannia glutinosa* cultivars and varieties [J]. *Planta Med*, 2008, 74(15): 1846-1852.
- [14] Doležal J, Greilhuber J, Suda J. Estimation of nuclear DNA content in plants using flow cytometry [J]. *Nat Protoc*, 2007, 2(9): 2233-2244.
- [15] 林丹, 李冰冰, 赵振利, 等. 基于流式细胞仪对不同品种泡桐倍性及白花泡桐基因组大小的测定 [J]. 河南农业大学学报, 2019, 53(3): 337-342.
- [16] Vurture G W, Sedlazeck F J, Nattestad M, *et al.* GenomeScope: fast reference-free genome profiling from short reads [J]. *Bioinformatics*, 2017, 33(14): 2202-2204.
- [17] Shen Y T, Liu J, Geng H Y, *et al.* *De novo* assembly of a Chinese soybean genome [J]. *Sci China Life Sci*, 2018, 61(8): 871-884.
- [18] Schmutz J, Cannon S B, Schlueter J, *et al.* Genome sequence of the palaeopolyploid soybean [J]. *Nature*, 2010, 463(7278): 178-183.
- [19] Qin C, Yu C S, Shen Y O, *et al.* Whole-genome sequencing of cultivated and wild peppers provides insights into *Capsicum* domestication and specialization [J]. *Proc Natl Acad Sci USA*, 2014, 111(14): 5135-5140.
- [20] Wang L H, Yu S, Tong C B, *et al.* Genome sequencing of the high oil crop sesame provides insight into oil biosynthesis [J]. *Genome Biol*, 2014, 15(2): R39.
- [21] 赵乐, 马利刚, 俎梦航, 等. 地黄 *RgGGPPS2* 基因克隆、生物信息学分析及表达分析 [J]. 中草药, 2017, 48(11): 2269-2278.
- [22] 朱昀昊, 董诚明, 俎梦航, 等. 地黄 *RgDXR* 基因的克隆及表达分析 [J]. 植物生理学报, 2017, 53(4): 563-571.
- [23] 赵乐, 史晶晶, 马利刚, 等. 地黄 *GGPPS1* 基因克隆及表达分析 [J]. 西北植物学报, 2016, 36(5): 888-895.
- [24] 杨超飞, 李欣容, 智惊宇, 等. 地黄环烯醚萜合酶基因的克隆、序列特征与表达分析 [J]. 中国中药杂志, 2019, 44(12): 2472-2479.
- [25] Li R, Li Y, Kristiansen K, *et al.* SOAP: Short oligonucleotide alignment program [J]. *Bioinformatics*, 2008, 24(5): 713-714.
- [26] 陈士林, 吴问广, 王彩霞, 等. 药用植物分子遗传学研究 [J]. 中国中药杂志, 2019, 44(12): 2421-2432.
- [27] 尉广飞, 董林林, 陈士林, 等. 本草基因组学在中药材新品种选育中的应用 [J]. 中国实验方剂学杂志, 2018, 24(23): 18-28.
- [28] 樊慧杰, 柴智, 黄浩楹, 等. 药用黄芪基因组大小的研究 [J]. 中国中医基础医学杂志, 2019, 25(9): 1299-1302.
- [29] 都明理, 徐娇, 朱楚然, 等. 3 种柴胡染色体数目测定及基因组大小估测 [J]. 江苏农业科学, 2019, 47(11): 191-193.
- [30] 张小燕, 刘志香, 廖保生, 等. 基于本草基因组学应用流式细胞术和高通量测序技术检测人参基因组大小 [J]. 世界科学技术—中医药现代化, 2017, 19(10): 1724-1728.
- [31] 唐其, 马小军, 莫长明, 等. 罗汉果全基因组 Survey 分析 [J]. 广西植物, 2015, 35(6): 786-791.
- [32] 霍恺森, 赵冬兰, 陈艳丽, 等. 甘薯属耐盐植物马鞍藤基因组大小及特征分析 [J]. 植物遗传资源学报, 2019, 20(3): 728-735.
- [33] 朱楚然, 徐娇, 都明理, 等. 三岛柴胡基因组 Survey 分析及 SSR 位点挖掘 [J]. 中国中药杂志, 2019, 44(18): 3960-3966.
- [34] 中国科学院中国植物志编辑委员会. 中国植物志 [M]. 北京: 科学出版社, 1979: 214-216.
- [35] Liu B, Shi Y, Yuan J, *et al.* Estimation of genomic characteristics by analyzing k-mer frequency in *de novo* genome projects [J]. *arXiv: Genomics*, 2013: 1308.
- [36] Xie T, Zheng J F, Liu S, *et al.* *De novo* plant genome assembly based on chromatin interactions: A case study of *Arabidopsis thaliana* [J]. *Mol Plant*, 2015, 8(3): 489-492.

[责任编辑 时圣明]