

DOI: 10.11931/guihaia.gxzw201404041

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

Tang Q, Ma XJ, Mo CM, et al. Genome survey analysis in *Siraitia grosvenorii*[J]. *Guihaia*, 2015, 35(6):786—791

罗汉果全基因组 Survey 分析

唐 其^{2,3}, 马小军^{1*}, 莫长明³, 潘丽梅³, 韦荣昌³, 赵 欢¹

(1. 中国医学科学院 药用植物研究所, 北京 100193; 2. 湖南农业大学 园艺园林学院, 长沙 410128; 3. 广西药用植物园, 南宁 530023)

摘 要: 罗汉果是广西特有药用及甜料植物, 其主要成分之一甜苷 V 作为天然、非糖甜味剂, 具有广阔的开发前景, 但罗汉果目前完全来自于栽培, 适生区狭窄, 连作障碍严重, 加之含量低导致甜苷 V 生产成本居高不下, 严重限制了其应用。为了减少盲目性, 在大规模全基因组深度测序之前, 先做低覆盖度的基因组 Survey 测序, 评价基因组的大小及复杂程度, 以确定适合该植物全基因组的测序研究策略。该研究采用第二代高通量测序技术(Illumina HiseqTM2000)首次测定了罗汉果基因组大小, 并利用生物信息学方法估计罗汉果杂合率、重复序列和 GC 含量等基因组信息。结果表明:(1)获得了 18.1 Gb 罗汉果基因组测序数据, 基因组大小估计为 344.95 Mb 左右, 测序深度为 52×;(2)从 K-mer 分布曲线发现罗汉果基因组有明显的杂合峰, 杂合率达 1.5%, 基因组高杂合导致组装的结果中 Contig N50 和 Scaffold N50 的长度比预期的要短很多, 还造成 GC 平均深度及含量分布明显异常, 存在一个低深度分布区域。基因组主峰后面有微弱的重复峰, 说明罗汉果存在较多的重复序列;(3)由于罗汉果存在高杂合率和重复序列较多的特点, 该基因组测序分析仅采用全基因组鸟枪法(WGS)策略不合适, 为了更好地对全基因组进行序列拼接和组装, 可尝试结合采用 Fosmid-to-Fosmid 或 BAC-to-BAC 策略。该研究结果对于揭示罗汉果产量、有效成分含量、发育及抗病虫的分子机制, 以及通过分子育种来提高甜苷 V 含量和降低生产成本具有重要意义, 为全基因组测序策略的选择提供了依据。

关键词: 罗汉果; 基因组测序; 杂合率; GC 含量; 鸟枪法测序策略

中图分类号: Q943.2 **文献标识码:** A **文章编号:** 1000-3142(2015)06-0786-06

Genome survey analysis in *Siraitia grosvenorii*

TANG Qi^{2,3}, MA Xiao-Jun^{1*}, MO Chang-Ming³,
PAN Li-Mei³, WEI Rong-Chang³, ZHAO Huan¹(1. *Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Beijing 100193, China;*
2. *Horticulture and Landscape College, Hunan Agricultural University, Changsha 410128, China;*
3. *Guangxi Botanical Garden of Medicinal Plants, Nanning 530023, China*)

Abstract: *Siraitia grosvenorii* (Luohanguo) is a herbaceous perennial medicinal and sweetener plant native in Guangxi of China. It has long been used in traditional Chinese medicine as a natural sweetener and also as a folk medicine for the treatment of lung congestion, colds and sore throats. Many cucurbitane-type triterpene glycosides have been isolated and characterized from *S. grosvenorii*. The active components responsible for the sweetness are the mogrosides, which are members of the family of triterpene glycosides. Mogroside V has an important prospect as natural and low calorie sweetener, which is nearly 425 times sweeter than sucrose. *S. grosvenorii* currently depends totally

收稿日期: 2014-05-28 修回日期: 2014-07-22

基金项目: 国家自然科学基金(81373914, 31400275); 国家科技支撑计划项目(2011BAI01B03); 广西农业科技成果转化项目(桂科转 1123013-12); 广西自然科学基金(2013GXNSFBA019170); 湖南省科技计划重点项目(2014SK2005); 广西卫生厅中医药科技专项(GZPT1235)。

作者简介: 唐其(1981-), 男, 湖南株洲县人, 博士, 助理研究员, 研究方向为药用植物分子生物学, (E-mail)tangqi423@sina.com。

*通讯作者: 马小军, 博士, 教授, 从事药用植物生物技术研究, (E-mail)mayixuan10@163.com。

on cultivation in China. It's limitedly applied because the narrow distribution, serious continuous cropping obstacle, low content and high extraction costs of mogroside V. In order to reduce blindness research and determine the appropriate sequencing strategy, the genome survey before large-scale genome sequencing is needed. This survey can provide information about the size and complexity of the whole genome of the *S. grosvenorii*. The next generation sequencing technology which has been emerged as a cost effective approach for high-through-put sequence determination has dramatically improved the efficiency and speed of genes discovery and genome research. Genome sequencing of *S. grosvenorii* has the vital significance to reveal the molecular mechanism of yield, content, growth, pest and disease resistance, and provides an efficient approach to improve content and reduce cost of mogroside V by molecular breeding. In this study, the genome size of *S. grosvenorii* was determined by next-generation sequencing technologies (NGS, Illumina Hiseq™ 2000). The hybridity percentage, repeats, and GC depth were also estimated by bioinformatics analysis. The results were as follows: (1) Two DNA libraries of 170 bp and 500 bp are constructed. After cleaning and quality checks, more than 18.1 Gb high quality data from the genome is generated, which were assembled into 943 296 contigs and 433 325 scaffolds by SOAP denovo software. The contig and scaffold numbers of the length more than 2 kb were 17 855 and 27 993 separately. The longest length of contig and scaffold were 29 kb and 268 kb. The N50 length of contig and scaffold were 484 bp and 2 331 bp. The average genome size and sequencing coverage depth of *S. grosvenorii* was about 344.95 Mb and 52 times respectively; (2) The genome of *S. grosvenorii* had obvious hybridity peak by K-mer method, the hybridity percentage as high as 1.5%. The assembly results showed that the length of contig N50 and scaffold N50 are much shorter than expected. High hybridity percentage of the genome leads to apparently unusual phenomenon between average depth and GC content, and had a low depth distribution area. There was a weak repeat peak behind the main peak, which demonstrated that *S. grosvenorii* has more repetitive sequences; (3) Whole-genome shotgun sequencing (WGS) should not be used to *S. grosvenorii* genome sequencing separately, and the Fosmid-to-Fosmid or BAC-to-BAC library could be combinational used for better results. This study would not only obtain the basic resources of genome, but also provide a theoretical basis and target genes for *S. grosvenorii* in transgenic breeding and genetic engineering.

Key words: *Siraitia grosvenorii* (Luohanguo); genome sequencing; hybridity percentage; GC depth; whole-genome shotgun sequencing

罗汉果甜苷 V 目前无法通过化学合成而完全依赖于栽培后提取获得,栽培品种含量为 1%左右,甜苷 V 仅存在于占果重不到 15%的果肉中(苏小建等,2006)。此外,罗汉果适栽区狭小且存在严重的连作障碍,导致了甜苷 V 的生产成本居高不下,作为甜味剂无法与蔗糖、木糖醇和甜菊糖 RA95 竞争,从而限制了其广泛应用。由于常规杂交育种对于甜苷 V 含量的提高潜力有限且效率低,因此全面解析罗汉果基因组,阐明甜苷 V 生物合成途径对于通过分子育种手段提高甜苷 V 含量、降低生产成本具有重要的研究价值。

罗汉果染色体数目和核型前人已有研究(庄伟建等,1997;李琦等,2007),罗汉果染色体数目为 $2n = 2x = 28$,由 $6m + 12sm + 6st + 4t$ 构成,按核型分类属于 3A 型(Fu *et al.*, 2012)。罗汉果甜苷 V 生物合成途径已初获阐明(Tang *et al.*, 2011),部分基因如法呢基焦磷酸合酶(FPP)和葡萄糖基转移酶(UDPG)已进行了克隆(蒙娇荣等,2011;邢爱佳等,

2013),罗汉果功能基因组研究处于起步阶段。然而,有关罗汉果基因组大小和全基因组测序研究均未见报道。常用的测定基因组大小的方法有 Feulgen 分光光度法(Du *et al.*, 2006)和流式细胞法(王亚之等,2010;李秋实等,2013)。新一代测序技术(New-Generation Sequencing technologies, NGS)的迅猛发展加速了植物全基因组的研究进程,成为当前最新的测定基因组大小的手段(伍艳芳等,2014)。目前,有报道已完成 80 余种植物的全基因组测序,还有大量的植物正在测序,这为研究植物的全基因组测序提供了大量的参考信息,特别是同属葫芦科的黄瓜(Huang *et al.*, 2009; Qi *et al.*, 2013)、西瓜(Guo *et al.*, 2012)和甜瓜(Garcia-Mas *et al.*, 2012)全基因组草图的绘制完成,对于罗汉果基因组的解析具有借鉴作用。

罗汉果栽培虽有 400 多年历史,但由于罗汉果属于异交作物,杂合性较高,基因组大小不明,而且遗传背景复杂。为减少盲目性,在大规模全基因组

深度测序之前,可以先做低覆盖度的初步测序,了解基因组的大小及复杂程度,以确定适合该植物全基因组的测序研究策略和拼接软件及技术(施季森等, 2012)。本研究采用第二代 Solexa 高通量测序技术(Illumina HiSeq™ 2000),对罗汉果全基因组大小进行测定和评估,旨在为罗汉果全基因组测序方案的制定提供依据。罗汉果的全基因组测序,将为掌握和利用罗汉果潜在的基因资源,阐明甜苷 V 生物合成途径及其调控机制,从而利用分子生物学手段对罗汉果进行靶向遗传改良奠定基础。同时,对于提高甜苷 V 含量,降低生产成本和拓宽应用范围也具有促进作用。

1 材料与方法

1.1 材料

所用材料为罗汉果青皮果品种“农院 B6”组培苗。黄化处理,剪取顶端幼嫩叶片,液氮速冻后置-70 °C 超低温冰箱保存备用。

1.2 方法

1.2.1 样品提取及检测 采用改良 CTAB 法(Liu *et al.*, 2011)提取罗汉果叶片基因组 DNA,紫外分光光度计检测浓度,琼脂糖凝胶电泳检测完整性,检测参数为胶浓度 0.5%,电压 3.5 V,电泳时间 960 min,以 1 kb DNA Ladder (Takara) 和 λ -Hind III digest (Takara) 作为 Marker。

1.2.2 测序 将提取的 DNA 样品进行测序分析。先将罗汉果的 DNA 样品进行随机打断,构建 170 bp 和 500 bp 的小片段测序文库,再用 Illumina HiSeq™ 2000 平台进行双末端(Pair-End)测序,过滤掉低质量数据,得到的高质量数据用于后续的基因组大小、杂合度和 GC 含量等信息分析。结合罗汉果的自身特性,按照 345 Mb 左右的罗汉果基因组大小来估算测序覆盖度。

1.2.3 17-mer 分析及基因组大小估计 在基因组组装前,采用基于 K-mer 的分析方法来估计基因组大小和杂合率等,取 K 为 17 来进行分析(Huang *et al.*, 2009)。假设从 reads 中逐碱基取出的所有 K-mer 能够遍历整个基因组,且 K-mer 深度频率分布服从泊松分布,即可从所有测序 reads 中统计 K-mer 频数分布,计算获得 K-mer 深度估计值,用于估计基因组大小。K-mer 分布图用来判断基因组的重复序列多少,如果这个基因组含有高比例的重复,

那么其分布图将显示出粗的拖尾现象。

1.2.4 杂合率估计 采用模拟数据拟合的方式进行基因组杂合率评估。选用酵母基因组序列,随机生成 $57 \times, 100$ bp 读长的 reads,读长错误率为 0.35%。因为二倍体复杂基因组进一步分为微杂合基因组($0.5\% \leq \text{杂合率} < 0.8\%$)、高杂合基因组(杂合率 $\geq 0.8\%$)以及高重复基因组(重复序列比例 $\geq 50\%$),针对不同类型的基因组,采用的测序方法及组装软件不同,因此分别加入杂合率为 1.4%、1.5% 和 1.6% 的模拟数据进行拟合,将所得模拟数据分别进行 17-mer 分析。测序物种真实曲线的主峰和杂合峰与哪条模拟数据最接近,就可大致认为该物种的杂合率处于相应的水平。

1.2.5 GC 含量及分布分析 利用高质量数据进行 SOAP denovo 组装(Li *et al.*, 2009),采用 K=41 构建 Contig 和 Scaffold,得到 Scaffold 序列后没有补洞,直接拼接组装获得原始基因组序列,这是一个最初的组装版本。用 SOAP 将过滤后的 reads 比对到该组装序列上,获得碱基深度。以 10 kb 为窗口,在序列上无重复前进,计算每个窗口的平均深度与 GC 含量,做出 GC_depth 点图。根据 GC_depth 分布图测序是否有明显的 GC 偏向,可以判断是否存在细菌污染等情况。同时,还可根据 GC 聚成块的分层来判断基因组的杂合率和重复的分布情况。

2 结果与分析

2.1 测序数据量统计

采用高通量测序技术进行本次测序,过滤掉低质量数据后,得到的总测序量为 18.1 Gb 用于后续分析,若罗汉果基因组大小如预计的 345 Mb,那么测序深度为 $52 \times$ (表 1)。

2.2 17-mer 分析以及基因组大小估计

使用罗汉果 18.1 Gb 的数据用于 17-mer 分析,其频率分布如图 1。横坐标表示 17-mer 出现的次数,纵坐标表示出现的频率。图 1 显示,17-mer 分布曲线成峰情况较好。在 44 附近有个主峰值,即 K-mer 的期望深度。表 2 显示,K-mer 的总数是 15 177 Mb(约 15 Gb),从而可以通过公式(基因组大小 = K-mer 的总数 / K-mer 的期望深度)估算基因组大小为 344.95 Mb。从图 1 还可看出,在期望深度 1/2 处有一个明显的凸峰。据此可以判断罗汉果基因组具有较高杂合率的可能性。K-mer 曲线呈

现明显拖尾,说明罗汉果基因组重复序列含量较高。

表 1 过滤数据量统计
Table 1 Data statistics

文库 ID Lib ID	读长 Read length (bp)	插入片段大小 Insert size (bp)	数据量 Data (Mb)	测序深度 Sequence depth (\times)
SZAXPI007823-39	100	170	11 860.36	34
SZAXPI007822-41	100	500	6 226.53	18
合计 Total			18 086.90	52

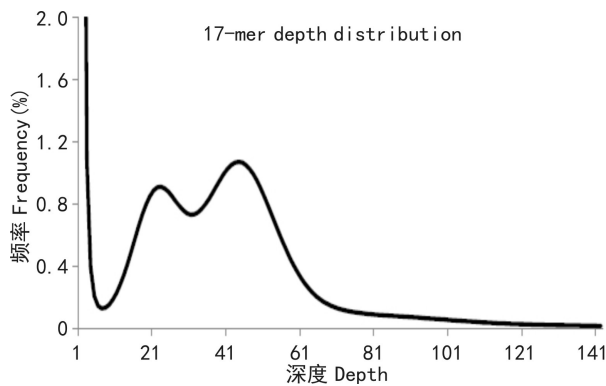


图 1 17-mer 分布曲线

Fig. 1 Distribution curve of 17-mer

表 2 17-mer 分析数据统计

Table 2 Data statistics and analysis of 17-mer

K-mer 总数 K-mer total number (Mb)	K-mer 深度 K-mer depth	基因组大小 Genome size (Mb)	所用碱基数 Used bases (Mb)	所用读长 Used reads (Mb)	测序深度 Sequence depth (\times)
15 178	44	344.95	18 069	181	52.38

2.3 杂合率估计

选用酵母基因组序列作为模拟数据估计罗汉果基因组的杂合率。由图 2 可以看出,罗汉果真实曲线的主峰和杂合峰在杂合率 1.5% 时形成的峰与酵母的模拟数据最接近,因此可大致认为罗汉果的杂合率处于 1.5% 水平,说明罗汉果杂合率高,用普通的建库方法和拼接软件可能影响组装效果。

2.4 基因组初步组装结果

利用 18.1 Gb 高质量数据进行罗汉果基因组的初步组装, $K=41$ 构建 Contig 和 Scaffold, 没有补洞, 获得原始基因组序列(表 3)。从表 3 可以看出, 由于罗汉果杂合率高达 1.5%, 组装的结果中 Contig N50 和 Scaffold N50 的长度比预期的要短很多。因此, 进一步说明全基因组鸟枪法不适合复杂的罗汉果基因组的组装。

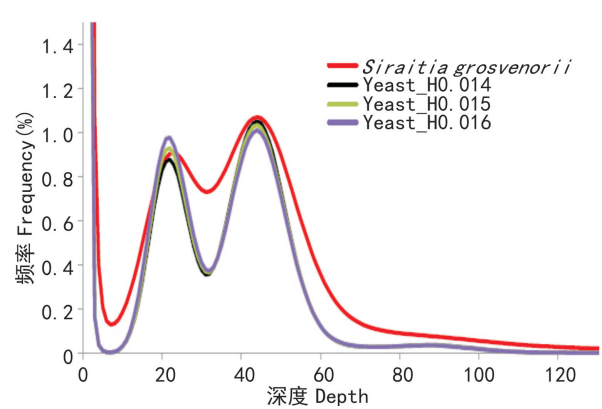


图 2 17-mer 杂合率估计图

Fig. 2 Statement of hybridity percentage of 17-mer

表 3 组装结果统计表

Table 3 Statistics of assembly results
in *S. grosvenorii* Genome

项目 Item	Contig		Scaffold	
	大小 Size (bp)	数目 Number	大小 Size (bp)	数目 Number
N90	144	683 011	232	214 362
N80	174	483 215	541	123 898
N70	225	323 175	853	73 813
N60	316	203 914	1 369	42 618
N50	484	122 378	2 331	23 690
最长 The longest (bp)	29 948		267 878	
总大小 Total size (bp)	314 519 645		332 386 660	
总数 Total number (>100 bp)	943 296		433 325	
总数 Total number (>2 kb)	17 855		27 933	

2.5 GC 含量及分布分析

GC_depth 分析显示(图 3), GC 测序无明显偏向, 但罗汉果平均深度与 GC 含量分布明显异常, GC_depth 深度分布被分成了 2 层, 存在一个低深度分布区域。取出低深度区域的序列与 Nt 核酸库进行 Blastn 比对, 未比对上细菌序列, 而且细菌序列一般 GC 含量较高, 因此可以排除此区域为细菌污染的可能性。GC 聚成的块分成了 2 层, 平均测序深度为 52 \times , 低深度分布区域的深度约为正常深度分布的 1/2, 结合杂合模拟显示罗汉果高度杂合的结果, 由此我们推测是由杂合引起。因为杂合会使两条同源染色体杂合的部位只装出了一条, 或 2 条都没有装出, 同时该部位以上的 read 乘数是整个基因组乘数的一半, 导致 GC 含量图中出现较低的一层。于是取出该区域对应窗口的序列, 与正常深度分布区域序列进行了 Blastn 比对。如果是由于杂

合造成的,会比对出较多的高相似度区域,结果证实了此推测。从低深度 B 区域取出来的 21.42 Mb 序列中,有 16.82 Mb(78.5%)的序列相互间存在高相似度。因此低深度分布 B 区域是由于罗汉果基因组高杂合造成的。这不仅解释了图 3 中 GC_depth 分布图的异常,也进一步验证了罗汉果的高杂合率的特性。说明单纯的使用全基因组鸟枪法测序后不能很好的进行基因组组装。

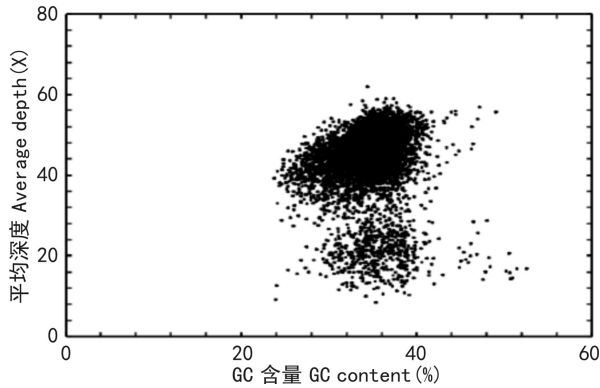


图 3 GC_depth 分布图

Fig. 3 Distribution figure of GC_depth

3 讨论与结论

基因组大小又称基因组含量或 DNA 1C 值,是指物种配子染色体组所含 DNA 的量。基因组大小是比较和进化基因组学研究的基础,对不同物种基因组大小进行比较分析,可以掌握基因组大小变化规律。全基因组复制及基因组染色体翻倍现象已成为热门的研究方向。本文对罗汉果基因组大小进行测定,为葫芦科植物基因组大小变化规律提供参考。据报道,葫芦科中的黄瓜(*Cucumis sativus*)基因组为 345 Mb,西瓜(*Citrullus lanatus*)的基因组为 425 Mb,甜瓜(*Cucumis melo*)的基因组为 450 Mb。基于第二代高通量测序技术的罗汉果基因组大小为 344.95 Mb,与黄瓜的基因组大小相近,比西瓜和甜瓜略小。其他 3 个物种由于经过长期的人工驯化,杂合率比罗汉果的低。

流式细胞仪是目前应用较多的基因组大小测定方法,在毛竹(李潞滨等,2008)、五节芒(邓果特等,2013)等植物和药用真菌如茯苓(李秋实等,2013)、灵芝(王亚之等,2010)中都有广泛的应用。除此之外,还有 Feulgen 分光光度法和脉冲场凝胶电泳法

等,高通量测序技术的快速发展为基因组评估提供了更迅捷的方法。本文采用的 K-mer 分析法是基于全基因组测序片段的 K-mer 深度分布估计物种基因组大小的方法,得到了罗汉果基因组大小、杂合度、GC 含量等结果,为该物种的进一步精细全基因组测序提供详细的遗传背景。测序技术在近几十年得到飞跃发展,其中第一代如 Sanger 测序比较成熟,但是通量低。第二代测序如 Illumina 公司的 Solexa、ABI 公司的 SOLiD 和 Roche 公司的 454 测序仪由于高通量、速度快等优点目前应用极为广泛,成为测序主流。第三代测序技术-单分子测序正在研制之中,虽已推出了第三代测序仪并有少量应用,但是进入大规模应用还需时日。由于许多植物属于高杂合、高重复的基因组,因此采用不同的测序技术相结合,可以获得较好的结果,如 Sanger+Illumina(黄瓜)、Sanger+454(甜瓜)、Illumina(西瓜)等。两种或多种测序方法的互补和联用进行测序研究,不失为当前一种可取的策略。通过全基因组测序,一般 Contig N50 能达到 20 kb 以上,Scaffold N50 能在 300 kb 以上,这样才能满足后续的组装要求,拼接出比较满意的基因组序列。二代测序技术中常用的组装软件主要有 SOAP、Velvet、ABYSS、GS Assembler 等 20 余种。实践证明,单个拼接软件在不同物种测序结果的拼接中存在一定的局限性,多种序列拼接软件的综合研究能够提高序列拼接的精确度,对于非模式植物的全基因组测序研究具有重要的指导意义。经过全基因组测序和生物信息学后获得的大量全基因组测序结果,可以为该物种的分子育种或分子标记辅助育种选择奠定坚实的基础。

本研究首次测定了罗汉果基因组大小,并对相应的参数进行了初步评估,主要结论如下:(1)罗汉果基因组大小粗略估计为 345 Mb 左右;(2)罗汉果基因组有较高的杂合度和一定的重复,属于复杂基因组范畴,杂合率约为 1.5%,对组装的效果影响较大。用 WGS 策略进行组装有一定的风险和难度;(3)由于罗汉果基因组杂合率较高,全基因组鸟枪法不适合该基因组分析。可尝试使用 BAC-to-BAC 策略或 Fosmid-to-Fosmid 策略对 WGS 策略进行补充,联合进行组装,这种组装策略对杂合度较高的复杂基因组拼接帮助较大。上述罗汉果全基因组 survey 分析结果将对罗汉果全基因组精细图谱绘制方案的制定提供重要的依据。

参考文献:

- Deng GT(邓果特), Liu QB(刘清波), Jiang JX(蒋建雄), et al. 2013. Estimation of genome size of *Miscanthus floridulus* (五节芒基因组大小测定)[J]. *J Plant Gen Resour* (植物遗传资源学报), **14**(2): 339-341
- Du B, Wang D. 2006. C-values of seven marine mammal species determined by flow cytometry[J]. *Zool Sci*, **23**(11): 1 017-1 020
- Fu W, Ma XJ, Tang Q, et al. 2012. Karyotype analysis and genetic variation of a mutant in *Siraitia grosvenorii* [J]. *Mol Biol Rep*, **39**(2): 1 247-1 252
- Garcia-Mas J, Benjak A, Sanseverino W, et al. 2012. The genome of melon (*Cucumis melo* L.) [J]. *Proc Nat Acad Sci USA*, **109**(29): 11 872-11 877
- Guo SG, Zhang JG, Sun HH, et al. 2013. The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions[J]. *Nat Genet*, **45**(1): 51-58
- Huang SS, Li RQ, Zhang ZH, et al. 2009. The genome of the cucumber, *Cucumis sativus* L. [J]. *Nat Genet*, **41**(12): 1 275-1 281
- Li LB(李潞滨), Wu JY(武静宇), Hu T(胡陶), et al. 2008. Estimation of genome size of Moso bamboo (*Phyllostachys edulis*) (毛竹基因组大小测定)[J]. *Chin Bull Bot* (植物学通报), **25**(5): 574-578
- Li Q(李琦), Ma L(马璐), Huang J(黄婧), et al. 2007. Chromosomal localization of ribosomal DNA Sites and karyotype analysis in three species of Cucurbitaceae (西瓜、苦瓜与罗汉果染色体的 rDNA 定位及其核型分析)[J]. *J Wuhan Univ: Nat Sci Ed* (武汉大学学报·理学版), **53**(4): 449-456
- Li QS(李秋实), Xu J(徐江), Zhu YJ(朱英杰), et al. 2013. Estimation of the genome size of *Ganoderma Lucidum* based on a flow cytometric analysis (基于流式细胞技术的灵芝基因组大小估测)[J]. *Mycosystema* (菌物学报), **32**(5): 899-906
- Li RQ, Yu C, Li YR, et al. 2009. SOAP2: an improved ultrafast tool for short read alignment [J]. *Bioinformatic*, **25**(15): 1 966-1 967
- Liu LH, Ma XJ, Mo CM, et al. 2011. The first genetic linkage map of Luohanguo (*Siraitia grosvenorii*) based on ISSR and SRAP markers[J]. *Genome*, **54**(1): 19-25
- Meng JR(蒙姣荣), Chen BY(陈本勇), Li QQ(黎起秦), et al. 2011. Cloning and sequence analysis of Farnesyl pyrophosphate synthase gene in *Siraitia grosvenorii* (罗汉果法呢基焦磷酸合成酶基因的克隆及其序列分析)[J]. *Chin Trad Herb Drugs* (中草药), **42**(12): 2 512-2 517
- Qi JJ, Liu X, Shen D, et al. 2013. A genomic variation map provides insights into the genetic basis of cucumber domestication and diversity[J]. *Nat genet*, **45**(12): 1 510-1 515
- Qi XY(戚向阳), Chen WJ(陈维军), Zhang LQ(张俐勤), et al. 2006. Study on the inhibitory effects of natural sweetener mogrosides on radical and lipid peroxidation (罗汉果皂甙清除自由基及抗脂质过氧化作用的研究)[J]. *Sci Agric Sin* (中国农业科学), **39**(2): 382-388
- Shi JS(施季森), Wang ZJ(王占军), Chen JH(陈金慧). 2012. Progress on whole genome sequencing in woody plants (木本植物全基因组测序研究进展)[J]. *Hereditas* (遗传), **34**(2): 145-156
- Su XJ(苏小建), Liu GX(刘国雄), Nie X(聂晓), et al. 2006. Study on the content distribution of Mogroside V in each part of *Grosvenor momordica* (罗汉果甜甙 V 在各部位的含量分布)[J]. *Food Sci & Technol* (食品科技), **5**: 76-78
- Tang Q, Ma XJ, Mo CM, et al. 2011. An efficient approach to finding *Siraitia grosvenorii* triterpene biosynthetic genes by RNA-seq and digital gene expression analysis[J]. *BMC Genom*, **12**: 343
- Wang YZ(王亚之), Li QS(李秋实), Chen SL(陈士林), et al. 2010. Flow cytometric analysis of nuclear DNA content in *Poria cocos* (基于流式细胞分析技术的茯苓基因组大小测定)[J]. *World Sci Technol / Mod Trad Chin Med Mat Med* (世界科学技术—中医药现代化), **12**(3): 452-456
- Wu YF(伍艳芳), Xiao FM(肖复明), Xu HN(徐海宁), et al. 2014. Genome survey in *Cinnamomum camphora* L. presl (樟树全基因组调查)[J]. *J Plant Gen Res* (植物遗传资源学报), **15**(1): 149-152
- Xing AJ(邢爱佳), Ma XJ(马小军), Mo CM(莫长明), et al. 2013. Cloning and prokaryotic expression of UDP-glycosyltransferase in *Siraitia grosvenorii* (罗汉果葡萄糖基转移酶基因的克隆及原核表达)[J]. *Acta Horti Sin* (园艺学报), **40**(6): 1 195-1 204
- Zhang WJ(张伟建), Lin ZL(林治良), Zheng SK(郑伸坤). 1997. The chromosomes of *Siraitia grosvenorii* (罗汉果染色体组型的研究)[J]. *J Trop & Subtrop Bot* (热带亚热带植物学报), **5**(2): 23-25

(上接第 801 页 Continue from page 801)

- Wei RC(韦荣昌), Zhao H(赵欢), Ma XJ(马小军), et al. 2014. Extraction of total RNA and cloning of *SgDHAR* gene from *Siraitia grosvenorii* (罗汉果果实 RNA 的提取及 *SgDHAR* 基因的克隆与表达)[J]. *Acta Pharm Sin* (药学报), **49**(1): 115-123
- Xu YH(徐艳红), Yang X(杨欣), Zhang Z(张争), et al. 2013. Cloning and expression analysis of HMG-CoA reductase from *Aquilaria sinensis* (Lour.) Gilg (白森香 3-羟基-3-甲基戊二酰辅酶 A 还原酶 ASHMGR₂ 的克隆及表达分析)[J]. *Acta Pharm Sin* (药学报), **48**(6): 953-959
- Zeng LH(曾黎辉), Wu JS(吴金寿), Ke SS(柯石山), et al. 2005. Establishment of the explant regeneration system for genetic transformation and transformation by *Agrobacterium rhizogene* of *Siraitia grosvenorii* (Swingle) C. Jeffrey (罗汉果遗传转化受体再生体系的建立及发根农杆菌转化初探)[J]. *Chin Agric Sci Bull* (中国农学通报), **21**(12): 403-406