

‘惠水金橘’的叶绿体基因组特征分析

王小柯, 郑乾明, 罗 悱, 柏自琴, 李文云, 李金强, 蔡永强*

(贵州省农业科学院果树科学研究所, 贵阳 550006)

摘要:【目的】分析贵州地方柑橘品种‘惠水金橘’的叶绿体基因组特征及分子标记,为研究该品种的系统进化关系、丰富柑橘类果树叶绿体分子标记奠定基础。【方法】提取‘惠水金橘’成熟叶片总DNA,构建小片段插入文库并测序,以甜橙叶绿体基因组为参考,筛选叶绿体基因组相关序列组装,开展SSR(simple sequence repeat,简单重复序列)预测、SNP(single nucleotide polymorphisms,单核苷酸多态性)/InDel(insertion-deletion,插入缺失)分析以及聚类分析。【结果】组装后获得叶绿体基因组全长为160 698 bp,GC含量为38.42%,共有117种基因被注释;检测到92个SSR位点,大部分的SSR位点都偏向于A/T组成;检测到417个SNPs位点、149个InDels位点、InDel的标记密度约为每kb 0.927个。进化树显示该品种与其他宽皮柑橘聚为一大类。【结论】利用高通量测序技术挖掘柑橘叶绿体基因组DNA分子标记是可行的。

关键词:柑橘;叶绿体基因组;高通量测序;SSRs;SNP/InDels

中图分类号:S666.2 文献标志码:A 文章编号:1009-9980(2019)03-0257-09

Characterization of the complete chloroplast genome of the ‘Huishui Jin Ju’ (*Citrus erythrosa* Hort. ex Tan.)

WANG Xiaoke, ZHENG Qianming, LUO Yi, BAI Ziqin, LI Wenyun, LI Jinqiang, CAI Yongqiang*

(Institute of Pomology Science, Guizhou Academy of Agricultural Sciences, Guiyang 550006, Guizhou, China)

Abstract: 【Objective】 In this study, we reported a complete chloroplast genome of ‘Huishui Jin Ju’ (*Citrus erythrosa* Hort. ex Tan.) with assistance from the next-generation sequencing, and the objective is to ascertain the chloroplast genetic difference between it and other citrus species, and to develop molecular markers for citrus varieties identification. 【Methods】 The total genomic DNA was extracted from the leaves of ‘Huishui Jin Ju’, and then the sequencing library was constructed with fragmented and qualified DNA samples. The Illumina HiSeq 4000 PE150 was used to sequence the library. The high quality short reads of sequencing were assembled with the assistance from *Citrus sinensis* chloroplast reference genome (NCBI Reference Sequence: NC_008334.1). The coding genes were annotated by DOGMA (<http://dogma.ccbb.utexas.edu/>). Using MISA (<http://pgrc.ipk-gatersleben.de/misa/misa.html>), it was undertaken to conduct SSRs chloroplast genome analysis. The genetic variation analysis was performed by MUMmer (version 3.23) nucmer. The neighbor-joining (NJ) methods from the MEGA (version 5.2) program were used to create phylogenetic trees. 【Results】 The complete chloroplast genome of ‘Huishui Jin Ju’ was a double-stranded circular DNA of 160 698 bp in length, consisting of a pair of inverted repeats (IR) regions of 26 990 bp, large single copy (LSC) region of 87 919 bp and a small single copy (SSC) region of 18 799 bp. The overall GC content accounted for 38.42% of the whole chloroplast genome. A total of 117 genes were predicated, including 83 protein-coding genes, 30 tRNA genes and 4 ribosomal RNA genes. In addition, 92 simple sequence repeats (SSRs) loci (≥ 5 bp) were identified including 90 single nucleotide repeats (97.8%), one trinucleotide repeat and one tetranu-

收稿日期:2018-08-30 接受日期:2018-11-01

基金项目:国家现代农业(柑桔)产业技术体系建设专项(CARS-26);贵州省科技攻关项目(黔科合NZ[2015]3001)

作者简介:王小柯,女,硕士,研究方向为果树分子生物学。Tel:18375638987,E-mail:wangxiaoke9191@163.com

*通信作者 Author for correspondence. Tel:0851-83762643,E-mail:Caiyq2015@sohu.com

cleotide repeat. In these chloroplast SSRs (cpSSR), the most of single nucleotide repeat sequences were A and T. The pairwise sequence alignment between the sweet orange and ‘Huishui Jin Ju’ chloroplasts genomes revealed approximately 0.71% (relative to the length of ‘Huishui Jin Ju’) sequence divergence, including 417 SNPs (0.45%) and 149 InDels (0.26%). The labeled density of InDel was 0.927 kb. The analysis of phylogenetic tree showed that ‘Huishui Jin Ju’ and some mandarin species were grouped into one branch in this study. **【Conclusion】**In this study, the high-throughput sequencing technology was used to sequence the complete chloroplast genome of ‘Huishui Jin Ju’, and we obtained the complete chloroplast sequence of 160 698 bp in length, and a total of 117 genes were predicated and annotated. The analysis of phylogenetic tree showed that ‘Huishui Jin Ju’ was closely related to mandarins. As for the highly conserved feature of chloroplast genome, we can develop molecular markers with chloroplast genomic to identify citrus varieties and genetic breeding materials.

Key words: Citrus; Chloroplast genome; High-throughput sequencing; SSRs; SNP/InDels

叶绿体起源于蓝藻菌,是具有半自主遗传体系的细胞器,是绿色植物和藻类进行光合作用的重要场所^[1]。叶绿体基因组是独立于核基因组外的细胞器基因组,具有单独的转录和转运系统,编码一些叶绿体特有构件,并且涉及多种主要代谢途径,包括糖合成、淀粉储存、多种氨基酸、脂质、维生素和色素的生成以及主要的硫和氮代谢途径等^[2]。1986年Shinozaki等^[3]和Ohyama等^[4]分别将烟草和地钱的叶绿体完整基因组拼接出来,自此之后越来越多植物的完整叶绿体基因组得到注释。叶绿体基因组多为环状双链结构,在细胞中是多拷贝,大小介于120~190 kb,序列特征具有相对保守性^[5]。随着进一步的研究发现,叶绿体基因组还存在许多结构重组、IR(inverted repeats,反向重复区)区段增长或缩短以及基因缺失等现象^[6-7]。因此对植物叶绿体基因组结构进行注释,并对其序列特征进行分析,可以为不同物种之间的亲缘关系鉴定提供重要信息,同时也将促进叶绿体基因组控制的重要农艺性状相关基因的功能分析。

‘惠水金橘’属于柑橘属中的朱橘类型,又名‘金钱橘’‘朱红橘’,果实较小,果皮油胞明显,结果率较高,易剥皮,果实含糖量高,可鲜食,也可用于加工制成蜜饯和果汁等,具有较高的经济价值^[8]。其主要分布于贵州省惠水县,是传统的地方柑橘品种,作为一种重要的种质资源,在柑橘新品种选育中发挥着关键作用。叶绿体基因组通常为母系遗传,保守程度较高,叶绿体分子标记具有很好的种属通用性。因此,开发一些叶绿体基因组DNA分子标记,有利于柑橘的系统发育研究、遗传多样性保护、亲缘关系鉴定以及分子育种等^[9-10]。笔者对‘惠水金橘’叶绿

体基因组序列特征进行分析,目的在于开发出一些叶绿体基因组DNA分子标记,以期促进该品种的系统发育研究及亲缘关系鉴定。

1 材料和方法

1.1 试验材料

供试品种为贵州省惠水县的实生‘惠水金橘’植株,树龄为15 a(年)以上。从长势良好、无病虫害的植株上采集成熟叶片,置于冰盒带回实验室。用无菌水冲洗叶片并擦干,液氮速冻后-80 ℃保存备用。

1.2 试验方法

1.2.1 DNA提取、建库测序 用植物基因组DNA提取试剂盒(TIANGEN Beijing China)提取‘惠水金橘’的总基因组DNA,并送样检测。样品经检测合格后,用机械打断的方法(超声波)将DNA片段化,然后对片段化的DNA进行片段纯化、末端修复、3’端加A、连接测序接头,富集大小为400 bp的DNA片段,再进行PCR扩增形成测序文库。建好的文库先进行文库质检,质检合格的文库用Illumina Hiseq 4000进行双末端各150 bp测序。

1.2.2 序列拼接组装 利用SOAPnuke软件对测序得到的序列进行低质量数据过滤得到有效数据,从而得到clean reads。过滤标准:去除N碱基含量超过5%的reads;去除低质量(质量值小于等于5)碱基数目达到50%的reads;去除有adapter污染的reads。使用SPAdes软件对上述clean reads进行基因组拼接,将拼接结果与甜橙叶绿体基因组(NCBI登录号: NC_008334.1)进行blastn比对,保留比对上且测序深度高于100的序列,基于序列测序深度、PE reads比对情况以及与参考基因组的比对顺序

等,确定候选序列的连接关系。

1.2.3 基因注释、结构预测 使用针对叶绿体和线粒体的注释软件 DOGMA (<http://dogma.cccb.utexas.edu/>) 进行编码基因注释,根据要预测基因组类型进行参数设置,最后的注释结果利用 OGDRAW 软件进行作图分析。

1.2.4 重复序列结构分析 利用 MISA 软件 (<http://pgrc.ipk-gatersleben.de/misa/misa.html>) 对叶绿体全基因组进行 SSRs 搜索,搜索条件设定为:含有完全重复的单碱基最小重复数为 10,二碱基最小重复数为 6,三、四、五、六碱基最小重复数为 5,2 个 SSR 之间的最小距离设置为 100 bp。

1.2.5 变异检测 利用 MUMmer 软件 (version: 3.23) 的 nucmer 对叶绿体全基因组进行 SNP 和 InDels 筛选。

1.2.6 聚类分析 从 NCBI 中下载 19 个柑橘属及其近缘属品种的 *matK* 基因序列,品种分别是 ‘Yoshida Ponkan’ (*Citrus reticulata* 1, AB626762.1)、‘Trifoliolate Orange’ (*Citrus trifoliata* 1, AB762360.1)、‘Sasebo Unshiu’ (*Citrus unshiu*, AB626749.1)、‘Clementine de Nules’ (*Citrus clementina*, AB626750.1)、‘Kabishi’ (*Citrus depressa*, AB626777.1)、‘Yamane’ (*Citrus junos*, AB626776.1)、‘Fukuhara’ (*Citrus sinensis*, AB626793.1)、‘Mato Buntan’ (*Citrus maxima*, AB626794.1)、‘Bergamot’ (*Citrus bergamia*, LC102229.1)、‘Sweet Lemon’ (*Citrus limetta*, LC102230.1)、‘Eureka’ (*Citrus limon*, AB626784.1)、‘Ichang papeda’ (*Citrus ichangensis*, AB626781.1)、‘Kabusu’ (*Citrus aurantium*, AB626785.1)、‘Maru busshukan’ (*Citrus medica*, AB626783.1)、‘Atlantia’ (*Atalantia spinose*, AB762379.1)、‘Orange Jessamine’ (*Murraya paniculata*, AB762389.1)、‘Wampee’ (*Clausena harmandiana*, AB762394.1)、*Citrus reticulata* 2 (FJ716729.1)、*Citrus trifoliata* 2 (FJ716730.1), 将这些序列与 ‘惠水金橘’ 的 *matK* 基因序列基于 MEGA 5.0 软件构建进化树,进行多序列比对后,将两端序列截齐,根据 Neighbor-Joining 算法构建进化树,Bootstrap 置信度重复 1 000 次。

2 结果与分析

2.1 叶绿体基因组组装及结构特征

基于高通量测序技术得到 ‘惠水金橘’ 全基因组

原始数据 30.93 G, 去掉低质量 reads 后得到 clean reads 26.97 G, Q30 为 93.91%, GC 含量为 36.86%。使用 SPAdes 软件进行基因组拼接,将拼接结果与甜橙 (*Citrus sinensis*) 叶绿体基因组进行 blastn 比对,保留比对上且测序深度高于 100 的序列,共得到 9 条 Scaffolds。基于序列测序深度、PE reads 比对情况以及参考基因组的比对顺序等,确定候选序列的连接关系,获得完整的叶绿体基因组环状拼接图(图 1)。将抽提的 ‘惠水金橘’ 叶绿体 clean data 比对到所有组装出的序列上,作出的 GC-depth 图表明不存在样品污染,组装情况较好。提取到的叶绿体基因组测序数据量为 630 M,有效数据的测序深度为 300×,有效数据的 Q20 和 Q30 分别为 98.18% 和 94.54%。

被子植物叶绿体基因组相对保守,绝大多数为双链环状结构,包括 2 个反向重复区 (inverted repeats, IRs)、大单拷贝区 (large single copy, LSC) 和小单拷贝区 (small single copy, SSC) 4 个部分^[7]。本研究中拼接组装出的叶绿体完整基因组大小为 160 698 bp,其中 2 个反向重复区 IRs 长度为 26 990 bp (包括 IRA 和 IRB),大单拷贝区 LSC 长度为 87 919 bp,小单拷贝区 SSC 长度为 18 799 bp。叶绿体基因组总 GC 含量为 38.42%,其中的 IR 区 GC 含量为 42.93%,高于 LSC 区 (36.78%) 和 SSC 区 (33.1%)。

2.2 基因注释情况

‘惠水金橘’ 叶绿体基因组有 117 种不同的基因类型,共 136 个基因被注释,包含 83 种 (102 个) 蛋白质编码基因、30 种 tRNA 编码基因和 4 种 rRNA 编码基因。其中有 12 个蛋白质编码基因 (*rpl2*、*rps7*、*rps19*、*rpl22*、*rpl23*、*ycf2*、*ycf15*、*ycf68*、*rps12*、*orf42*、*orf56*、*ndhB*) 和 7 个 tRNA 编码基因 (*trnA-UGC*、*trnL-CAA*、*trnI-CAU*、*trnI-GAU*、*trnN-GUU*、*trnR-ACG*、*trnV-GAC*) 在 IRs 区域被重复。大多数蛋白质编码基因由 1 个外显子组成,10 个基因 (*trnA-UGC*、*trnI-GAU*、*trnL-UAA*、*trnV-UAC*、*rpl2*、*ycf15*、*ndhA*、*ndhB*、*atpF*、*rpoC1*) 包含 1 个内含子,2 个基因 (*ycf3*、*alpP*) 包含 2 个内含子。按功能类型大致分为 4 类 (表 1): (1) 光合作用相关基因; (2) 自身翻译相关基因; (3) 生物合成相关基因; (4) 未知功能基因。

2.3 SSR 位点分析

本研究中 ‘惠水金橘’ 叶绿体基因组共发现 92 个 SSRs 位点 (≥ 5 bp), SSR 标记密度为每 kb 0.573

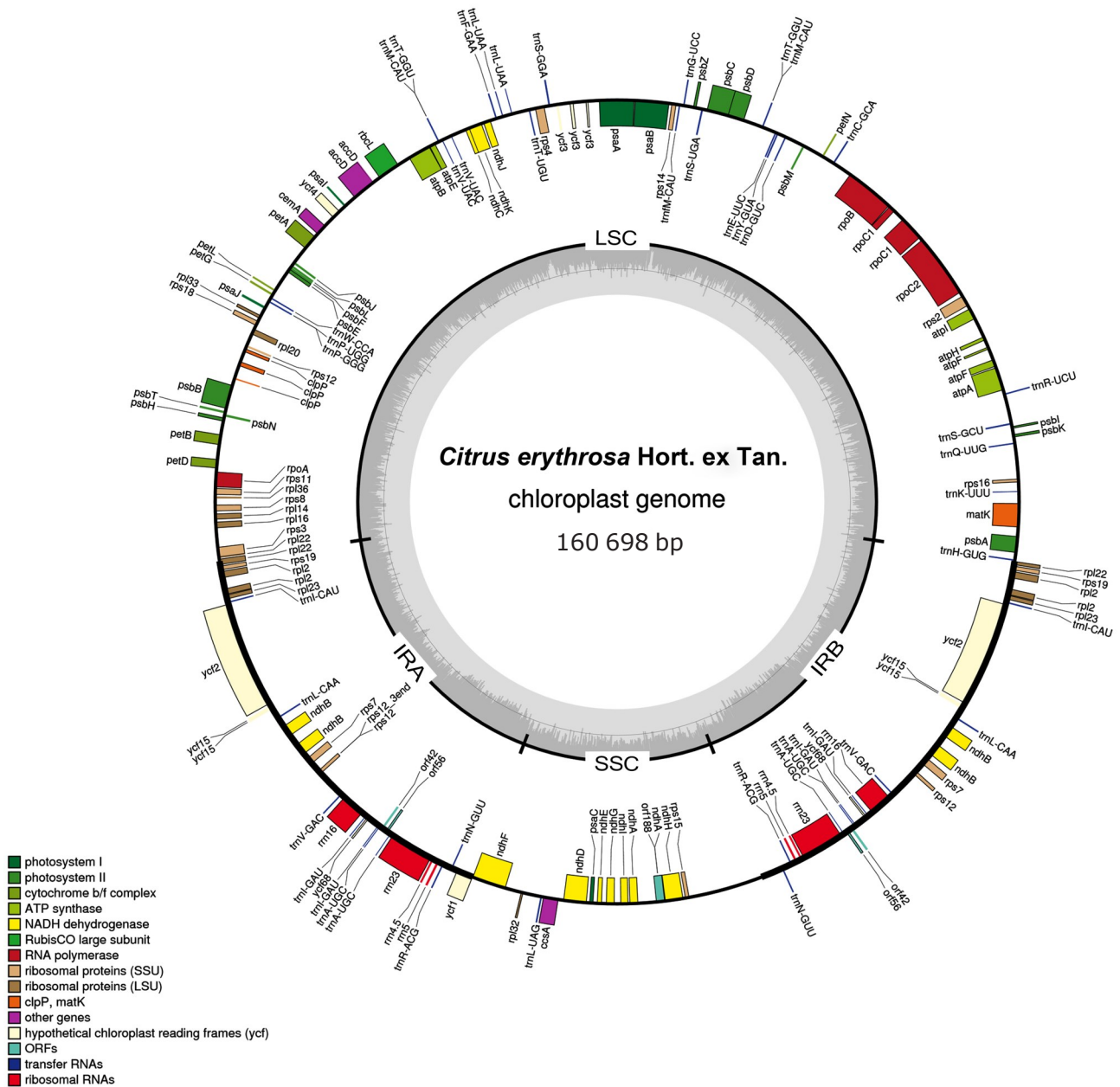


图 1 ‘惠水金橘’的完整叶绿体基因组

Fig. 1 The complete chloroplast genome of ‘HuiShui Jin Ju’ (*Citrus erythrosa* Hort. ex Tan.)

个。其中有 90 个单核苷酸重复, 约占 97.8%, 1 个三核苷酸重复和 1 个四核苷酸重复(表 2)。在这些叶绿体 SSRs 中, 大多数的单核苷酸重复序列的碱基组成偏向于 A(39) 和 T(50) 碱基, 仅有 1 个为 C, 最长的 1 个重复序列为 poly T(22 bp), 大多数重复单元的长度为 11 bp 左右(20 个 10 bp、18 个 11 bp、19 个 12 bp), 占有所有叶绿体基因组 SSRs 的 61.96%(57/92)。

2.4 SNP 和 InDel 分析

与甜橙叶绿体参考基因组比较后发现, 两者之

间大约有 0.71% 的序列差异(相对于‘惠水金橘’叶绿体基因组的长度), 包括 417 个 SNP 位点(0.45%) 和 149 个 InDel 位点(0.26%)(图 2)。在 SNP 变异中, 发生碱基转换的比例是 32.13%(134/417), 发生碱基颠换的比例是 67.87%(283/417)。在 InDel 变异中, 标记密度为每 kb 0.927 个, 大部分的插入或缺失片段大小都是 20 bp 以内, 其中碱基插入约占 64.43%(96/149), 最长的插入片段为 48 bp; 碱基缺失约占 35.57%(53/149), 最长的缺失片段为 19 bp。

表1 ‘惠水金橘’叶绿体基因组基因列表
Table 1 A list of genes predicate in the chroloplast genome of ‘HuiShui Jin Ju’ (*Citrus erythrosa* Hort. ex Tan.)

基因类别 Gene type		基因名称 Gene name
光合作用相关基因 Photosynthetic genes	光系统 I Photosystem I	<i>psaA, psaB, psaC, psal, psaJ</i>
	光系统 II Photosystem II	<i>psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ</i>
	细胞色素 b/f 复合体 Cytochrome b/f complex	<i>petA, petB, petD, petG, petL, petN</i>
	ATP 合酶 ATP synthase	<i>atpA, atpB, atpE, atpF^c, atpH, atpI</i>
	NADH 脱氢酶 NADH dehydrogenase	<i>ndhA^b, ndhB^b, ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK</i>
	二磷酸核酮糖羧化酶大亚基 RubisCO large subunit	<i>rbcl</i>
	ATP 蛋白酶基因	<i>ClpP^c</i>
自身翻译相关基因 Self-translation related genes	RNA 聚合酶 RNA polymerase	<i>rpoA, rpoB, rpoC1^b, rpoC2</i>
	核糖体蛋白(SSU) Ribosomal proteins(SSU)	<i>rps2, rps3, rps4, rps7^c, rps8, rps11, rps12^a, rps14, rps15, rps16, rps18, rps19^a</i>
	核糖体蛋白(LSU) Ribosomal proteins(LSU)	<i>rpl2^{a,b}, rpl14, rpl16, rpl20, rpl22^a, rpl23^a, rpl32, rpl33, rpl36</i>
	转运 RNAs Transfer RNAs	<i>trnA-UGC^{a,b}, trnC-GCA, trnD-GUC, trnE-UUC, trnF-GAA, trnJ-M-CAU, trnG-UCC, trnH-GUG, trnK-UUU, trnL-CAA^a, trnL-CAU^a, trnL-GAU^{a,b}, trnL-UAA^b, trnL-UAG, trnM-CAU(×2), trnN-GUU^b, trnP-GGG, trnP-UGG, trnQ-UUG, trnR-ACG^a, trnR-UCU, trnS-GCU, trnS-GGA, trnS-UGA, trnT-GGU(×2), trnT-UGU, trnV-GAC^c, trnV-UAC^c, trnW-CCA, trnY-GUA,</i>
生物合成相关基因 Biosynthesis related genes	核糖体 RNAs Ribosomal RNAs	<i>rrn4.5^a, rrn5^a, rrn16^a, rrn23^a</i>
	成熟酶基因 matK	<i>matK</i>
	C 型细胞色素合成基因 ccsA	<i>ccsA</i>
	乙酰辅酶 A 羧化酶亚基 accD	<i>accD</i>
未知功能基因 Unknown functional genes	囊膜蛋白基因 cemA	<i>cemA</i>
	Hypothetical chloroplast reading frames (ycf) ORFs	<i>ycf1, ycf2^a, ycf3^c, ycf4, ycf15^{a,b}, ycf68^a orf42^a, orf56^a, orf188</i>

注:a 表示在 IRs 有 2 个重复基因;b 表示该基因包含 1 个内含子;c 表示该基因包含 2 个内含子。

Note: a means two gene copies due to IRs; b means gene that contains one intron, c means gene that contains two intron.

表2 ‘惠水金橘’叶绿体基因组 SSRs 位点分布
Table 2 Distribution of simple sequence repeats (SSRs) loci in the ‘HuiShui Jin Ju’ (*Citrus erythrosa* Hort. ex Tan.) chloroplast genome

长度 Length/bp	SSRs 的位置和数量 Location and number of SSRs					
	A	C	T	G	ATT	ATTT
5	0	0	0	0	0	1 (117 935)
6	0	0	0	0	1 (1 771)	0
10	10 (5 329, 8 155, 32 179, 39 189, 48 514, 54 580, 118 692, 119 053, 135 597, 160 088)	1 (29 319)	10 (10 937, 17 560, 27 514, 32 836, 46 660, 64 352, 74 560, 78 446, 88 521, 113 012)	0	0	0
11	8 (3 951, 4 513, 72 606, 73 320, 74 877, 85 328, 122 482, 145 033)	0	10 (2 428, 19 694, 46 709, 49 912, 81 109, 84 256, 103 575, 119 583, 126 762, 131 984)	0	0	0
12	8 (4 915, 13 943, 64 641, 112 600, 114 772, 118 463, 119 105, 141 097)	0	11 (1 787, 29 574, 31 683, 64 248, 74 648, 86 049, 107 510, 118 613, 122 436, 133 835, 136 007)	0	0	0
13	3 (397, 13 437, 86 860)	0	5 (54 925, 82 165, 87 726, 126 128, 128 632)	0	0	0
14	2 (6 882, 12 542)	0	2 (24 253, 118 479)	0	0	0
15	2 (10 644, 71 010)	0	1 (13 967)	0	0	0
16	1 (48 005)	0	3 (57 901, 69 058, 118 734)	0	0	0
17	2 (33 139, 85 328)	0	1 (118 975)	0	0	0
18	1 (70 684)	0	1 (82 096)	0	0	0
19	0	0	0	0	0	0
20	2 (7 401, 70 198)	0	1 (47 626)	0	0	0
21	0	0	3 (9 929, 32 140, 53 454)	0	0	0
22	0	0	2 (15 271, 34 054)	0	0	0
合计 Total	39	1	50	0	1	1

注:括号中表示位置。

Note: The loci are indicated in parentheses.

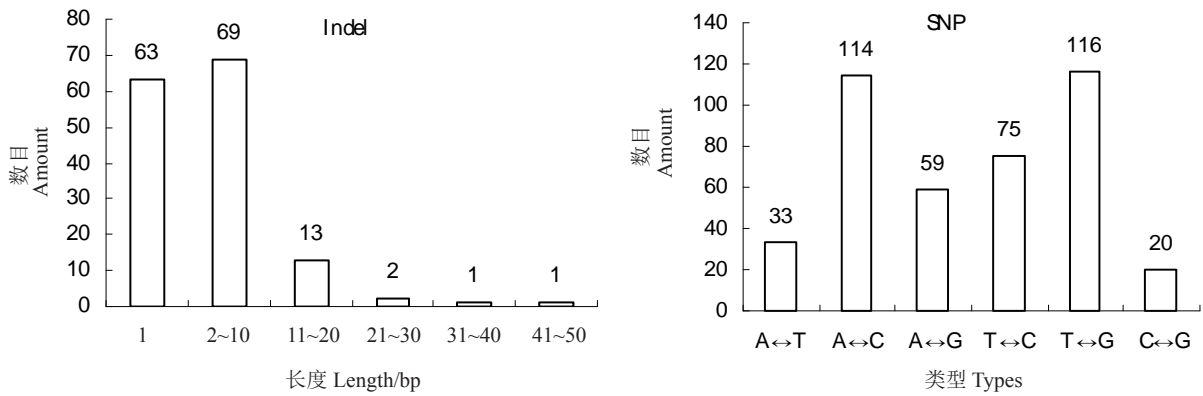


图 2 ‘惠水金橘’叶绿体基因组的 InDel 和 SNP 统计分析
 Fig. 2 The InDel and SNP of ‘HuiShui Jin Ju’ (*Citrus erythroa* Hort. ex Tan.) chloroplast genomes

2.5 聚类分析

有文献报道表明, *matK* 基因的进化速度相对较快^[27], 因此本研究基于不同品种的 *matK* 基因序列构建系统进化树(图3), 主要包含柑橘属的 17 个品种, 并加入九里香(*Murraya paniculata*)、酒饼簕(*Atalantia spinosa*)和黄皮(*Clausena harmandiana*)作为

外类群。聚类结果显示, ‘惠水金橘’与其他宽皮橘聚在一起形成一个大类。

3 讨 论

传统的叶绿体基因组获取方法为提取植物叶绿体基因组 DNA 后测序, 该方法关键点在于分离纯化

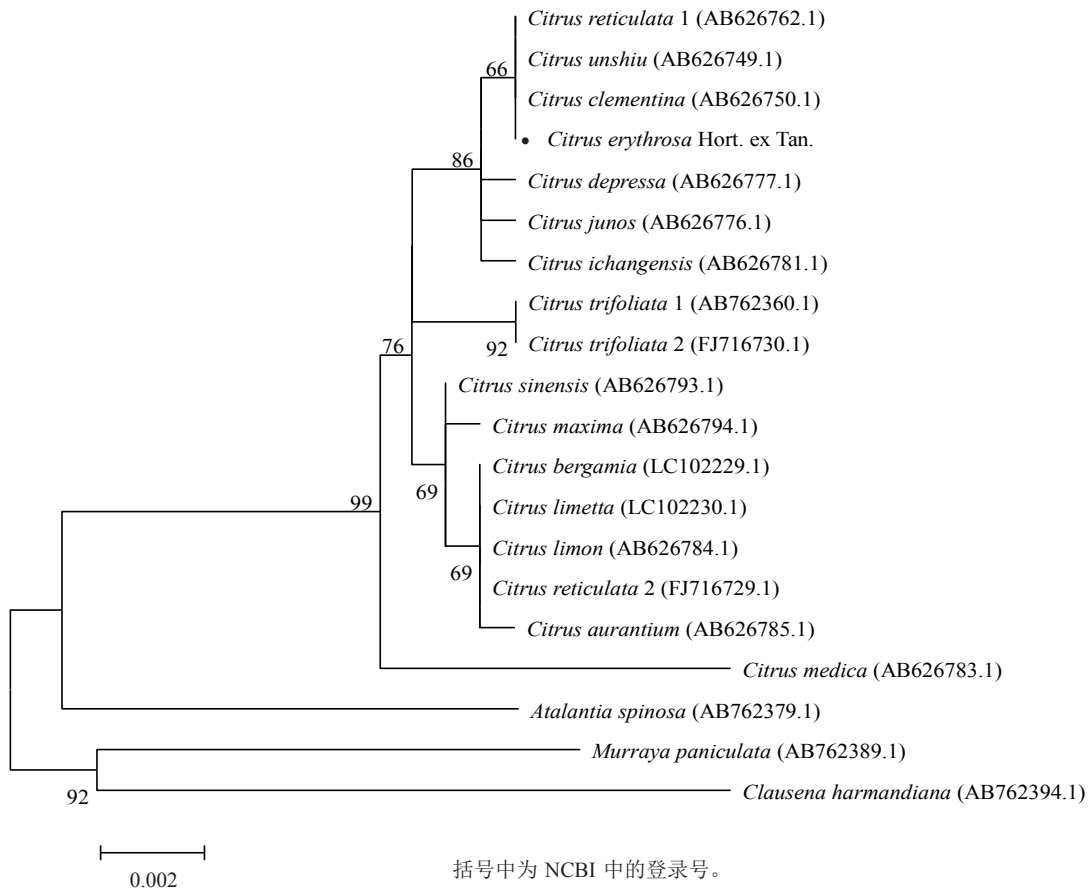


图 3 基于邻接法构建 20 份柑橘种质资源 *matK* 基因的系统发育树
 Fig. 3 Neighbor-joining tree of the *matK* genes from 20 *Citrus* germplasm resources

出高质量的叶绿体DNA。Bausher等^[11]通过对甜橙(*Citrus sinensis*)叶片进行低温黑暗处理、调整蔗糖密度梯度等一系列的步骤改良,分离出较纯的甜橙叶绿体,但该方法操作繁琐、难度大且成本较高。2009年,台湾科学家首次利用一种基于PCR的方法,不用分离纯化叶绿体DNA,获得了麻竹(*Dendrocalamus latiflorus*)和绿竹(*Bambusa oldhamii*)的叶绿体基因组^[12]。随着第二代测序技术应用的越来越普遍,许多物种的叶绿体基因组完整序列被公布,其中不乏果树类,如苹果^[13]、梨^[14]、葡萄^[15]、猕猴桃^[16]、樱桃^[17]、菠萝^[18]以及香蕉^[19-20]等,柑橘类有甜橙^[11]、柠檬^[21]、扁实柠檬^[22]和柚子^[23]等。笔者利用高通量测序技术对贵州地方柑橘品种‘惠水金橘’进行重测序,通过提取植物总DNA建库测序,以已知近缘物种的叶绿体基因组作为参考序列进行BLASTN比对,提取出相关的叶绿体reads,再利用SPAdes软件对这些reads进行拼接组装。与传统的叶绿体基因组测序方法相比,本方法避免了分离纯化叶绿体DNA的步骤,采用直接提取植物总DNA的方法,操作简单,节约成本。同时,与三代测序技术相比,本研究中所使用的二代测序技术具有测序流程较为成熟、准确率高、成本低等特点。

根据已知报道,甜橙叶绿体基因组大小160 129 bp,其IR区、SSC区以及LSC区的长度分别为26 996 bp、18 393 bp以及87 744 bp,有79种蛋白编码基因、4种rRNA编码基因,30种tRNA编码基因,总GC含量为38.48%^[11]。本研究中‘惠水金橘’叶绿体基因组大小为160 698 bp,其中IR区、SSC区以及LSC区长度分别为26 990 bp、18 799 bp以及87 919 bp,总GC含量为38.42%,有83种蛋白编码基因、4种rRNA编码基因,30种tRNA编码基因。与甜橙相比,‘惠水金橘’叶绿体基因组大小、各区段长度以及GC含量差异不大;基因注释结果显示,‘惠水金橘’与甜橙拥有相同数目的tRNAs和rRNAs;蛋白质编码基因种类和数目与甜橙相比增加4种未知功能基因(*orf42*、*orf56*、*orf188*、*ycf68*),其他基因一致。表明二者叶绿体基因组序列具有高度保守性。

在被子植物中,叶绿体DNA通常为母系遗传,在进化过程中不经历基因重组,可以很好地反应物种系统发育过程^[24]。叶绿体基因组相较于核基因组更小,便于测序,且叶绿体基因组保守性较高,含有

大量的遗传信息,其编码区进化速度较慢,非编码区进化速度较快,能提供较多的遗传变异位点^[5]。因此可利用高通量测序的手段,开发叶绿体SSR、SNP和InDel标记位点,对柑橘群体遗传学研究具有较高的利用价值^[25]。本研究共检测到92个SSRs位点,有3种重复类型,分别为单核苷酸重复、三核苷酸重复以及四核苷酸重复,SSR标记密度为每kb 0.573个,大部分的碱基组成偏向于A/T。叶绿体基因组差异分析共检测到417个SNP位点和149个InDel位点,差异类型多为单核苷酸多态性变异以及短片段的插入/缺失变异,大片的缺失或插入较少,最长的一个插入片段为48 bp,其中InDel的标记密度约为每kb 0.927个,远远高于SSR标记的密度。关于柑橘的分类至今仍存在诸多争议,本研究中发掘的叶绿体SSR、SNP和InDel分子标记,可以为柑橘品种的分类鉴定提供参考依据。

*matK*基因位于叶绿体基因组DNA上,编码的成熟酶参与RNA转录本形成过程中II型内含子的剪切,*matK*的突变率相对较快,它的进化速度也比*rbcL*基因要快^[26]。因此可用于研究柑橘类品种之间的种系发生规律。Penjor等^[27]对柑橘亚科22个属135个品种(包括柑橘属的宽皮橘、橙、柚、柠檬以及莱檬等)的*matK*基因序列进行测序,并构建不同类型的遗传进化树,对柑橘属的分类进行全面地探讨。于杰等^[28]通过对柑橘及其近缘属植物59份样品进行*matK*、*rpoB*、*rpoCl*、*rbcL*编码基因序列测序、比对与校正,计算属间、种内种间的遗传距离,通过wilcoxon秩和检验得出*matK*序列的差异率最大,可作为柑橘及其近缘属植物DNA条形码研究中重要的特征序列之一。本研究基于*matK*编码基因序列构建不同品种之间的系统进化树,聚类结果显示‘惠水金橘’与宽皮橘聚在一起形成一个大类。

从形态学特征上观察,‘惠水金橘’果实扁圆形,偏小,果皮较薄,橙黄色,易剥皮,具特殊气味,接近于古老栽培种中的乳橘类型,推测其在系统演化过程中可能与南丰蜜橘亲缘关系较近^[29]。本研究中并没有将‘惠水金橘’、克里曼丁、椪柑及温州蜜柑等宽皮柑橘类品种区分开来,究其原因一方面由于叶绿体基因组是母系遗传,相较于核基因组来说具有高度保守性^[7];另一方面本研究是基于*matK*编码基因构建进化树,尽管研究表明不同品种间*matK*序列碱基差异较大,但单一序列仍不能很好地展示不同品

种的遗传差异。Carbonell-Caballero 等^[30]对柑橘属 34 个品种的叶绿体全基因组完整序列进行遗传多样性分析和变异分析,绘制出柑橘属不同种类叶绿体基因组遗传变异图谱,更好地研究了柑橘属的种系发生和系统演化规律,但上述序列尚未完全公开。未来的研究中,有必要利用完整的叶绿体基因组序列,结合柑橘属其他品种的叶绿体全基因组对‘惠水金橘’的遗传地位进行鉴别。

4 结 论

利用高通量测序技术对贵州地方柑橘品种‘惠水金橘’的叶绿体全基因组进行测序,通过生物信息学分析,挖掘出一些 SSR、SNP 和 Indel 标记,将为后续品种鉴定以及遗传育种研究提供参考依据。

参考文献 References:

- [1] 李继耕. 叶绿体遗传与细胞质雄性不育性[J]. 中国农业科学, 1983, 16(1):49-55.
LI Jigeng. Chloroplast heredity and cytoplasmic male sterility [J]. Scientia Agricultura Sinica, 1983, 16(1):49-55.
- [2] 林张翔, 王营营, 付菲, 叶楚玉, 樊龙江. 东乡野生稻叶绿体基因组拼接及系统进化分析[J]. 浙江大学学报(农业与生命科学版), 2014, 40(4):397-403.
LIN Zhangxiang, WANG Yingying, FU Fei, YE Chuyu, FAN Longjiang. Assembly and phylogenetic analysis of Dongxiang wild rice chloroplast genome[J]. Journal of Zhejiang University (Agriculture and Life Sciences), 2014, 40(4):397-403.
- [3] SHINOZAKI K, OHME M, TANAKA M, WAKASUGI T, HAYSHIDA N, MATSUBAYASHA T, ZATITA N, CHUNWONGSE J, OBOKATA J, YAMAGUCHI-SHINOZAKI K, OHTO C, TORAZAWA K, MENG B Y, SUGITA M, DENO H, KAMOGASHIRA T, YAMADA K, KUSUDA J, TAKAIWA F, KATA A, TOHDOH N, SHIMADA H, SUGIURA M. The complete nucleotide sequence of the tobacco chloroplast genome[J]. Plant Molecular Biology Reporter, 1986, 5(9):2043-2049.
- [4] OHYAMA K, FUKUZAWA H, KOHCHI T, SHIRAI H, SANO T, SANO S, UMESONO K, SHIKI Y, TAKEUCHI M, CHANG Z, AOTA S, INOKUCHI H, OZEKI H. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA[J]. Nature, 1986, 322(6079):572-574.
- [5] 燕安, 朱登云. 叶绿体基因组在系统发育学及基因工程领域的应用[J]. 中国细胞生物学学报, 2004, 26(2):153-156.
YAN An, ZHU Dengyun. The application of chloroplast genome in the studies of systematics and bioengineering[J]. Chinese Journal of Cell Biology, 2004, 26(2):153-156.
- [6] CHUMLEY T W, PALMER J D, MOWER J P, FOURCADE H M, CALIE P J, BOORE J L, JANSEN R K. The complete chloroplast genome sequence of *Pelargonium × hortorum*: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants[J]. Molecular Biology and Evolution, 2006, 23(11):2175-2190.
- [7] 邢少辰, CLARKE JIHONG LIU. 叶绿体基因组研究进展[J]. 生物化学与生物物理进展, 2008, 35(1):21-28.
XING Shaochen, LIU J H. Progress in chloroplast genome analysis[J]. Progress in Biochemistry and Biophysics, 2008, 35(1):21-28.
- [8] 周开隆, 叶荫民. 中国果树志·柑橘卷[M]. 北京: 中国林业出版社, 2010.
ZHOU Kailong, YE Yinmin. China fruit's monograph·Citrus volume[M]. Beijing: China Forestry Publishing Press, 2010.
- [9] 付涛, 王志龙, 钱萍仙, 李文, 袁冬明, 严春风. 高等植物 DNA 条形码最新研究进展及其应用[J]. 核农学报, 2016, 30(5):887-896.
FU Tao, WANG Zhilong, QIAN Pingxian, LI Wen, YUAN Dongming, YAN Chunfeng. The latest research progress and application of the DNA barcode in higher plants[J]. Journal of Nuclear Agricultural Sciences, 2016, 30(5):887-896.
- [10] 程慧, 葛春峰, 张红, 乔玉山. 果树叶绿体基因组测序及系统发育研究进展[J]. 核农学报, 2018, 32(1):58-69.
CHENG Hui, GE Chunfeng, ZHANG Hong, QIAO Yushan. Advances on chloroplast genome sequencing and phylogenetic analysis in fruit trees[J]. Journal of Nuclear Agricultural Sciences, 2018, 32(1):58-69.
- [11] BAUSHER M G, SINGH N D, LEE S B, JANSEN R K, DANIELL H. The complete chloroplast genome sequence of *Citrus sinensis* (L.) Osbeck var ‘Ridge Pineapple’: organization and phylogenetic relationships to other angiosperms[J]. BMC Plant Biology, 2006, 6(1):21.
- [12] WU F H, KAN D P, LEE S B, DANIELL H, LEE Y W, LIN C C, LIN N S, LIN C S. Complete nucleotide sequence of *Dendrocalamus latiflorus* and *Bambusa oldhamii* chloroplast genomes [J]. Tree Physiology, 2009, 29(6):847-856.
- [13] 金桂花, 陈斯云, 伊廷双, 张书东. 苹果叶绿体基因组特征分析[J]. 植物分类与资源学报, 2014, 36(4):468-484.
JIN Guihua, CHEN Siyun, YI Tingshuang, ZHANG Shudong. Characterization of the complete chloroplast genome of apple (*Malus × domestica*, Rosaceae)[J]. Plant Diversity and Resources, 2014, 36(4):468-484.
- [14] TERAOKAMI S, MATSUMURA Y, KURITA K, KANAMORI H, KATAYOSE Y, YAMAMOTO T, KATAYAMA H. Complete sequence of the chloroplast genome from pear (*Pyrus pyrifolia*): genome structure and comparative analysis[J]. Tree Genetics & Genomes, 2012, 8(4):841-854.
- [15] 谢海坤, 焦健, 樊秀彩, 张颖, 姜建福, 孙海生, 刘崇怀. 基于高通量测序组装‘赤霞珠’叶绿体基因组及其特征分析[J]. 中国农业科学, 2017, 50(9):1655-1665.

- XIE Haikun, JIAO Jian, FAN Xiucui, ZHANG Ying, JIANG Jianfu, SUN Haisheng, LIU Chonghuai. Assembling and characteristic analysis of the complete chloroplast genome of *Vitis vinifera* cv. Cabernet Sauvignon from high-throughput sequencing data[J]. *Scientia Agricultura Sinica*, 2017, 50(9): 1655-1665.
- [16] YAO X, TANG P, LI Z, LI D, LIU Y, HUANG H. The first complete chloroplast genome sequences in Actinidiaceae: genome structure and comparative analysis[J]. *PLoS One*, 2015, 10(6): e0129347.
- [17] FENG Y, LIU T, WANG X, LI B, LIANG C, CAI Y. Characterization of the complete chloroplast genome of the Chinese cherry *Prunus pseudocerasus* (Rosaceae)[J]. *Conservation Genetics Resources*, 2017, 10(1): 85-88.
- [18] NASHIMA K, TERAOKAMI S, NISHITANI C, KUNIHISA M, SHODA M, TAKEUCHI M, URASAKI N, TARORA K, YAMAMOTO T, KATAYAMA H. Complete chloroplast genome sequence of pineapple (*Ananas comosus*)[J]. *Tree Genetics & Genomes*, 2015, 11(3): 60.
- [19] MARTIN G, BAURENS F C, CARDI C, AURY J M, HONT A D. The complete chloroplast genome of banana (*Musa acuminata*, Zingiberales): Insight into plastid monocotyledon evolution [J]. *PLoS One*, 2013, 8(6): e67350.
- [20] SHETTY S M, MD S M, MAKALE K, MOHD- YUSUF Y, KHALID N, OTHMAN R Y. Complete chloroplast genome sequence of *Musa balbisiana* corroborates structural heterogeneity of inverted repeats in wild progenitors of cultivated bananas and plantains[J]. *The Plant Genome*, 2016, 9(2): 1-14.
- [21] SU H J, HOGENHOUT S A, AL-SADI A M, KUO C H. Complete chloroplast genome sequence of Omani Lime (*Citrus aurantiifolia*) and comparative analysis within the rosids[J]. *PLoS One*, 2014, 9(11): e113049.
- [22] RYUJI I, NATHAN B, KUNIMASA M, KANAME M, TOSHIKI O, MASAYUKI O. Multi-lineages of Shiikuwasha (*Citrus depressa* Hayata) evaluated by using whole chloroplast genome sequences and its bio-diversity in Okinawa, Japan[J]. *Breeding Science*, 2016, 66(4): 490-498.
- [23] LIU J, SHI C. The complete chloroplast genome of wild shaddock, *Citrus maxima* (Burm.) Merr.[J]. *Conservation Genetics Resources*, 2017, 9(4): 599-601.
- [24] 张庆滢, 陈璇, 郭孟璧, 郭蓉, 许艳萍, 杨明, 郭鸿彦. 野生大麻叶绿体基因组分子多态标记的筛选与开发[J]. *分子植物育种*, 2017, 15(3): 979-985.
- ZHANG Qingying, CHEN Xuan, GUO Mengbi, GUO Rong, XU Yanping, YANG Ming, GUO Hongyan. Screening and development of chloroplast polymorphic molecular markers on wild hemp (*Cannabis sativa* L.)[J]. *Molecular Plant Breeding*, 2017, 15(3): 979-985.
- [25] KUANG D Y, WU H, WANG Y L, GAO L M, ZHANG S Z, LU L, BONEN L. Complete chloroplast genome sequence of *Magnolia kwangsiensis* (Magnoliaceae): implication for DNA barcoding and population genetics[J]. *Genome*, 2011, 54(8): 663-673.
- [26] HILU K, LIANG H. The *matK* gene: sequence variation and application in plant systematics[J]. *American Journal of Botany*, 1997, 84(6): 830-839.
- [27] PENJOR T, YAMAMOTO M, UEHARA M, IDE M, MATSUMOTO N, MATSUMOTO R, NAGANO Y. Phylogenetic relationships of citrus and its relatives based on *matK* gene sequences [J]. *PLoS One*, 2013, 8(4): e62574.
- [28] 于杰, 闫化学, 鲁振华, 周志钦. 基于柑橘及其近缘属植物DNA条形码的叶绿体编码序列筛选[J]. *中国农业科学*, 2011, 44(2): 341-348.
- YU Jie, YAN Huaxue, LU Zhenhua, ZHOU Zhiqin. Screening potential DNA barcode regions of chloroplast coding genome for *Citrus* and its related genera[J]. *Scientia Agricultura Sinica*, 2011, 44(2): 341-348.
- [29] 王小柯, 江东, 孙珍珠. 利用GBS技术研究240份宽皮柑橘的系统演化[J]. *中国农业科学*, 2017, 50(9): 1666-1673.
- WANG Xiaoke, JIANG Dong, SUN Zhenzhu. Study on phylogeny of 240 Mandarin accessions with genotyping-by-sequencing technology[J]. *Scientia Agricultura Sinica*, 2017, 50(9): 1666-1673.
- [30] CARBONELL- CABALLERO J, ALONSO R, IBAÑEZ V, TEROL J, TALON M, DOPAZO J. A Phylogenetic analysis of 34 chloroplast genomes elucidates the relationships between wild and domestic species within the genus *Citrus*[J]. *Molecular Biology and Evolution*, 2015, 32(8): 2015-2035.