

· 研究报告 ·

## 石刁柏核质体DNA的生物信息学分析及染色体定位

程广前<sup>1†</sup>, 贾克利<sup>1,2†</sup>, 李娜<sup>1</sup>, 邓传良<sup>1</sup>, 李书粉<sup>1</sup>, 高武军<sup>1\*</sup>

<sup>1</sup>河南师范大学生命科学学院, 新乡 453007; <sup>2</sup>新乡医学院三全学院, 新乡 453003

**摘要** 在植物基因组中, 叶绿体DNA (cpDNA)序列可以向核基因组转移成为核质体DNA (NUPT)。NUPTs在植物染色体(包括性染色体)的演化过程中具有重要作用, 但目前相关研究比较缺乏。以雌雄异株植物石刁柏(*Asparagus officinalis*)为材料, 采用生物信息学方法对其核基因组NUPTs进行注释及分析, 并选取叶绿体基因组反向重复区(IR) 2个片段进行染色体定位。结果表明, 石刁柏核基因组中有2 239个NUPTs序列的插入, 总长度为565 970 bp, 占核基因组的0.047%。不同染色体上插入的NUPTs数量存在较大差异, Y染色体上的NUPTs数量、密度及总长度均高于其它染色体, 表明NUPTs在石刁柏性(Y)染色体上累积的更多。石刁柏叶绿体基因组中的IR区、大单拷贝区(LSC)和小单拷贝区(SSC)序列均能够向核基因组转移, 但IR区序列转移频率更高。此外, 对2个IR区的叶绿体序列进行荧光原位杂交, 其中AocpIR1主要分布在所有染色体的着丝粒部位, 而AocpIR2特异性分布在Y染色体上。研究结果为深入揭示石刁柏基因组的结构及其性染色体的演化奠定了坚实的基础。

**关键词** 石刁柏, 生物信息学分析, 核质体DNA, NUPTs, 性染色体

程广前, 贾克利, 李娜, 邓传良, 李书粉, 高武军 (2019). 石刁柏核质体DNA的生物信息学分析及染色体定位. 植物学报 54, 328–334.

植物细胞除了含有细胞核基因组之外, 还包含细胞器基因组, 主要包括线粒体基因组和叶绿体基因组(Michalovova et al., 2013)。在基因组演化过程中, 细胞器基因组可以向核基因组中转移, 其中由叶绿体DNA转移到核基因组中的序列称为核质体DNA (nuclear integrants of plastid DNAs, NUPTs), 而由线粒体DNA转移到核基因组中的序列称为核线粒体DNA (nuclear integrants of mitochondria DNAs, NUMTs) (Richly and Leister, 2004; Timmis et al., 2004; Kleine et al., 2009; Yoshida et al., 2014)。

目前, 已在多种植物的核基因组中发现了NUPTs (Kejnovsky et al., 2006; Guo et al., 2008; Steflova et al., 2014)。NUPTs是组成植物基因组的一种特殊类型的重复序列, 包括NUPTs和转座子在内的重复序列的插入会引起基因组不稳定, 如引起染色体结构重排、体积膨胀和异染色质化, 在基因组和染色体进化过程中具有重要作用(高东迎等, 2007;

Feschotte and Pritham, 2007; Li et al., 2016)。细胞学研究表明, NUPTs多集中在染色体异染色质区域, 如着丝粒及其侧翼等。在水稻(*Oryza sativa*)中, 大片段的NUPTs主要集中在着丝粒及侧翼区, 随着演化的进行NUPTs长度逐渐衰减, 并且伴随着拷贝数的增加(Guo et al., 2008)。有研究表明, NUPTs在雌雄异株植物的性染色体上累积较多。Kejnovsky等(2006)在白麦瓶草(*Silene latifolia*) Y染色体上发现了NUPTs, 这些NUPTs也多位于Y染色体上的着丝粒及侧翼区, 且NUPTs的累积可能是白麦瓶草Y染色体体积膨胀的主要原因之一。在番木瓜(*Carica papaya*)中, Y染色体的雄性特异区上累积的NUPTs比X染色体高12倍, 比其它常染色体高4倍, 这种差异是由于NUPTs的复制所致(VanBuren and Ming, 2013)。上述研究表明, NUPTs的插入与植物性染色体的演化密切相关。

收稿日期: 2018-05-11; 接受日期: 2018-08-06

基金项目: 国家自然科学基金(No.31470334)和河南省高校科技创新团队支持计划(No.17IRTSTHN017)

† 共同第一作者

\* 通讯作者。E-mail: gaowujun@htu.cn

石刁柏(*Asparagus officinalis*)隶属百合科天门冬属, 是重要的雌雄异株植物。其基因组大小约为1.3 Gb, 染色体组成为 $2n=2x=20$ , 具有X和Y同型性染色体(Löptien, 1979), 是研究性染色体起源和早期演化的理想材料之一。本实验室利用基因组消减杂交和PCR获得了一些石刁柏性别差异NUPTs (李书粉等, 2016), 但石刁柏全基因组NUPTs的分布规律和染色体分布模式尚未见报道。本研究对石刁柏基因组的NUPTs进行注释和分析, 并采用荧光原位杂交对2个序列进行染色体定位, 旨在为进一步分析NUPTs在石刁柏染色体上的分布以及在性染色体的起源和演化中的作用提供证据。

## 1 材料与方法

### 1.1 石刁柏基因组NUPT的生物信息学分析

#### 1.1.1 数据来源

本研究所用的生物信息学分析数据包括石刁柏(*Asparagus officinalis* L.)叶绿体基因组(下载自<https://www.ncbi.nlm.nih.gov/nuccore/1197512595/>, 大小为156 699 bp (Sheng et al., 2017))和石刁柏核基因组(下载自<https://www.ncbi.nlm.nih.gov/genome/?term=10978>, 大小为1 187.54 Mb (Harkess et al., 2017))数据。

#### 1.1.2 基因组NUPTs的注释

采用本地Blast (Altschul et al., 1990)软件, 以石刁柏叶绿体基因组作为query序列, 对石刁柏核基因组进行Blast分析, e value设置为 $e^{-5}$ , 一致性 $\geq 80\%$ 。在结果中去除50 bp以下的序列, 即得到注释为石刁柏基因组NUPTs的序列。根据Blast结果, 计算不同染色体上不同位置NUPT的数量, 滑动窗口设为5 M, 采用R语言作图。

#### 1.1.3 基因组NUPTs在叶绿体基因组中的来源分析

NUPTs起源的叶绿体基因组区域通过将每个识别的NUPT分配(mapping)到叶绿体基因组的序列来确定。由于叶绿体基因组的2个IR (inverted repeat region)区序列互为反向互补序列, 故如果NUPT被mapping到IR区, 只计数一半。根据Blast结果, 计算叶绿体基因组不同位置的NUPT数量, 滑动窗口设为500 bp。

### 1.2 石刁柏叶绿体基因组IR区2个序列的染色体定位

#### 1.2.1 基因组DNA的提取

取石刁柏的幼嫩叶片, 采用传统的CTAB法(Doyle and Doyle, 1987)提取基因组DNA。由于所取的叶片材料含有大量的叶绿体, 因此所提取的基因组DNA中除了核基因组外, 也包含叶绿体基因组DNA (李巧丽等, 2018)。

#### 1.2.2 叶绿体DNA序列扩增及回收纯化

根据石刁柏叶绿体IR区序列, 随机选取2个序列区域并设计引物, 引物序列分别为AocplR1F: 5'-AGATA-GGAAGGGCTGTTGC-3', AocplR1R: 5'-AGCAGG-GTCAGGAACAACG-3'; AocplR2F: 5'-GATCTCAA-TTCCAGCTACGG-3', AocplR2R: 5'-GTAGGTATC-TGGTTCCTACTGC-3'。AocplR1和AocplR2扩增序列长度分别为4 118 bp和3 614 bp。利用这2对特异引物进行PCR扩增, 反应体系如下: 2.5  $\mu$ L 10 $\times$ PCR buffer, 正反向引物各1  $\mu$ L, 1.3  $\mu$ L dNTP, 0.2  $\mu$ L Taq DNA聚合酶, 1  $\mu$ L石刁柏模板DNA, 用去离子水补充至25  $\mu$ L。PCR反应条件如下: 95 $^{\circ}$ C2分钟; 94 $^{\circ}$ C30秒, 57 $^{\circ}$ C50秒, 72 $^{\circ}$ C4分钟, 35个循环; 72 $^{\circ}$ C10分钟。

将扩增得到的PCR产物用0.8%琼脂糖凝胶进行电泳, 并用EasyPure Quick Gel Extraction Kit试剂盒进行切胶回收, 再利用EasyPure PCR Purification Kit试剂盒将切胶产物回收纯化(纯化后浓度不低于200 ng $\cdot\mu$ L $^{-1}$ )后保存备用。

#### 1.2.3 探针标记

为了辅助石刁柏染色体核型分析, 采用玉米(*Zea mays*) 45S rDNA (质粒由中国科学院遗传与发育生物学研究所胡赞民实验室馈赠)和NUPT标记探针进行双色荧光原位杂交。45S rDNA质粒和NUPT序列回收纯化产物采用缺口平移法标记探针。NUPT片段标记红色荧光素, 45S rDNA标记绿色荧光素。探针标记体系为: 2  $\mu$ g DNA, 2  $\mu$ L 10 $\times$ Nick translation buffer, 0.5  $\mu$ L 荧光素标记的dNTPs, 2  $\mu$ L 未标记的dNTPs, 5  $\mu$ L 10 U $\cdot\mu$ L $^{-1}$  DNA polymerase I, 0.5  $\mu$ L 100 mU $\cdot\mu$ L $^{-1}$  DNase I, 用去离子水补充至20  $\mu$ L。依次加入上述试剂后充分离心混匀, 置于15 $^{\circ}$ C金属浴中孵育2小时, 取出置于4 $^{\circ}$ C冰箱5小时, 之后加入160  $\mu$ L 5 $\times$ TAE+140 ng $\cdot\mu$ L $^{-1}$ 鲑精DNA和500  $\mu$ L 90%乙醇-

10%乙酸钠,充分混匀后置于 $-20^{\circ}\text{C}$ 过夜。隔日放入预冷的离心机中, $4^{\circ}\text{C}$  16 000  $\times g$ 离心30分钟,弃上清;加入1 000  $\mu\text{L}$  70%乙醇并于 $4^{\circ}\text{C}$  16 000  $\times g$ 离心2分钟,弃上清;加入1 000  $\mu\text{L}$ 无水乙醇并于 $4^{\circ}\text{C}$  16 000  $\times g$ 离心2分钟,室温放置30分钟至1小时,直至无水乙醇完全挥发;加20  $\mu\text{L}$  2 $\times$ SSC+1 $\times$ TE充分溶解, $-20^{\circ}\text{C}$ 保存备用。

#### 1.2.4 中期染色体玻片标本制备

于 $25^{\circ}\text{C}$ 培养石刁柏种子,待种子根尖长至长约1.0–1.5 cm时,剪取根尖放置在1.01 MPa的含 $\text{N}_2\text{O}$ 密闭容器内处理2小时,后用90%的冰乙酸固定10分钟,用70%乙醇清洗4遍,将根尖保存至70%乙醇中,置于 $-20^{\circ}\text{C}$ 备用。

取出根尖,用1 $\times$ 柠檬酸缓冲液清洗根尖10分钟,取根尖分生组织置于纤维素果胶酶液中, $37^{\circ}\text{C}$ 孵育2小时后用1 $\times$ TE处理2遍,然后用无水乙醇清洗至无TE残留,用金属针将其捣碎制成悬浮液,低速离心40–50秒,弃上清。加入35  $\mu\text{L}$ 冰乙酸,混匀,吸取5  $\mu\text{L}$ 滴片并观察,将染色体形态好的片子保存备用。

#### 1.2.5 荧光原位杂交

将待杂交切片置于紫外交联仪中按每平方米0.125 J的能量进行交联。取3.5  $\mu\text{L}$  2 $\times$ SSC+1 $\times$ TE buffer、2  $\mu\text{L}$ 已标记的叶绿体DNA红色探针、0.5  $\mu\text{L}$ 已标记的45S rDNA绿色探针,充分混匀后离心。

将载玻片放在冰上,在有细胞的地方滴加6  $\mu\text{L}$ 杂交液,盖上塑料盖玻片,置于沸水中5分钟,然后放入预热的密闭湿盒中 $55^{\circ}\text{C}$ 杂交过夜。将载玻片从湿盒中取出,放入2 $\times$ SSC中使盖玻片滑落,放入 $55^{\circ}\text{C}$ 水浴锅中15–20分钟,之后用吸水纸把片子背面擦干,滴上1滴含DAPI的抗淬灭剂,盖上24 mm $\times$ 50 mm的盖玻片,使用BX63荧光显微镜进行镜检,Andor CCD拍照,采用Adobe PhotoShop CS6软件对图片进行亮度、对比度的微调 and 剪切处理。

## 2 结果与讨论

### 2.1 石刁柏基因组NUPTs在染色体上的长度分布模式

为了探明石刁柏基因组NUPTs的分布特征,我们利用已发表的石刁柏基因组测序数据进行基因组

NUPTs注释。结果表明,目前发表的石刁柏核基因组中共有2 239个NUPTs序列的插入,总长度为565 970 bp,占核基因组的0.047%。我们进一步对NUPTs在染色体上的分布进行分析(图1),发现所有染色体上均有NUPTs分布,其中1号染色体(Y染色体)的NUPTs插入密度为272个/100 Mb,平均插入长度为778 bp/Mb,而其余常染色体NUPTs平均插入密度和平均插入长度分别为178个/100 Mb和446 bp/Mb,表明Y染色体上NUPTs插入密度和长度均明显高于常染色体。

我们对NUPTs在不同染色体上的分布位置进行分析,结果表明,NUPTs在染色体上的分布不是随机的,而是存在一些插入热点区域。如在1号染色体的近中部和6号染色体的端部均包含1个NUPTs富集的区域,在5 M的序列内分别有37 755 bp和20 498 bp的NUPTs序列(图2)。

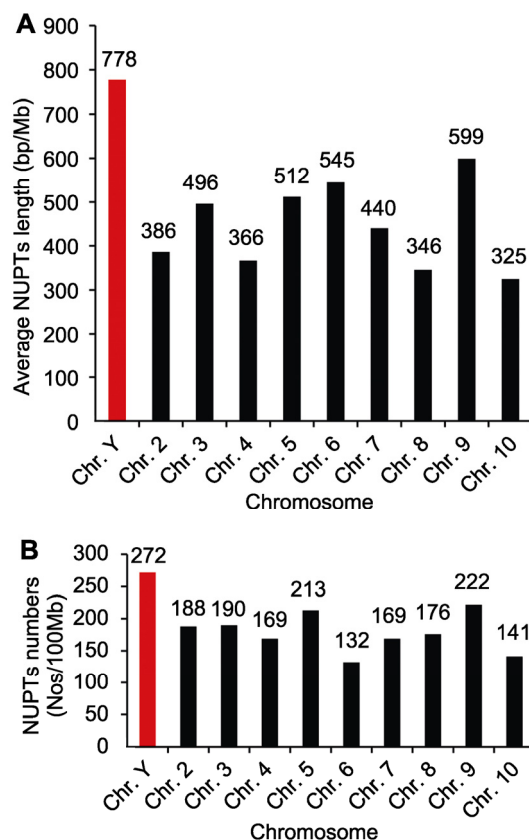


图1 石刁柏不同染色体核质体DNA (NUPTs)长度(A)和数目(B)

Figure 1 The length (A) and number (B) of nuclear intergrants of plastid DNAs (NUPTs) in different chromosomes of *Asparagus officinalis*

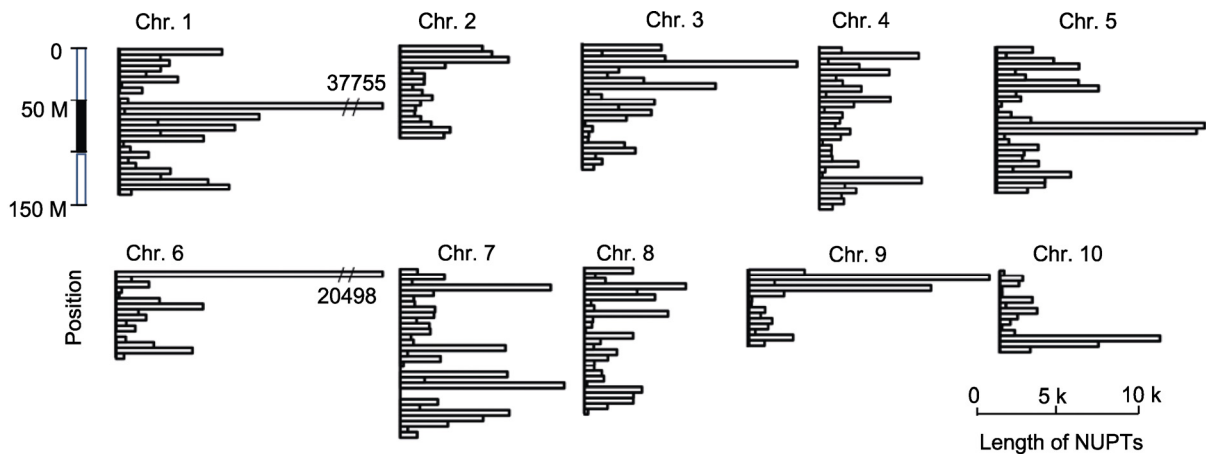


图2 石刁柏基因组核质体DNA (NUPTs)在染色体上的分布模式

Figure 2 The chromosome distribution pattern of nuclear integrants of plastid DNAs (NUPTs) of *Asparagus officinalis* genome

2.2 石刁柏核基因组NUPTs来源分析

为了解叶绿体基因组序列向核基因组的转移特点,我们对石刁柏基因组中的NUPTs数量及来源进行分析,结果表明,在2 239个石刁柏核基因组NUPTs中,来源于IR区、大单拷贝区(large single copy region, LSC)和小单拷贝区(small single copy region, SSC)的NUPTs数量分别为726、1 227和286个。由于石刁柏叶绿体基因组单个IR、LSC和SSC区的长度分别为26 471 bp、83 821 bp和17 901 bp,因此其密度分别为27.9个/kb、14.6个/kb和15.9个/kb(表1)。此外,IR、LSC和SSC区转移片段的平均长度分别为241 bp、258 bp和246 bp,表明三者转移的平均长度没有明

表1 石刁柏核基因组核质体DNA (NUPTs)的叶绿体基因组来源分析

Table 1 Origin analysis of nuclear integrants of plastid DNAs (NUPTs) sequences on the chloroplast genome of *Asparagus officinalis*

	IR	LSC	SSC
Total sequence length (bp)	26471	83821	17904
Number of NUPTs	726	1227	286
Average density (No./kb)	27.9	14.6	15.9
Average length of NUPTs (bp)	241	258	246
The length of NUPTs (bp)	175215	316626	70394

LSC: 大单拷贝区域; SSC: 小单拷贝区域。IR为反向重复区域,又分为IRa区和IRb区。

LSC: Large single copy region; SSC: Small single copy region. IR denotes reverted repeat region, and it can be divided into IRa and IRb.

显差异(表1)。上述研究结果表明,石刁柏叶绿体基因组向核基因组的转移不是随机的,来源于IR区的NUPTs密度明显高于其它2个区域,说明叶绿体基因组中的IR区序列更容易向核基因进行水平转移。

2.3 2个IR区序列的获得

根据IR区序列设计引物,以石刁柏基因组(包含叶绿体DNA)为模板进行PCR扩增。如图3所示,均扩增得到约3–5 kb之间的单一条带,大小与预期一致。

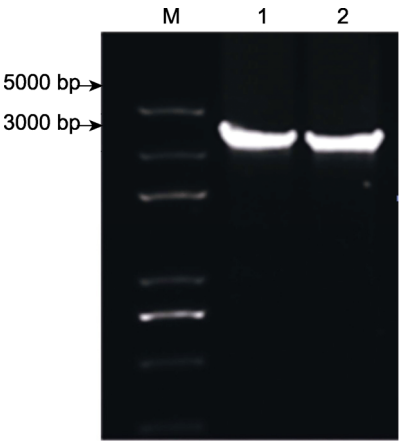
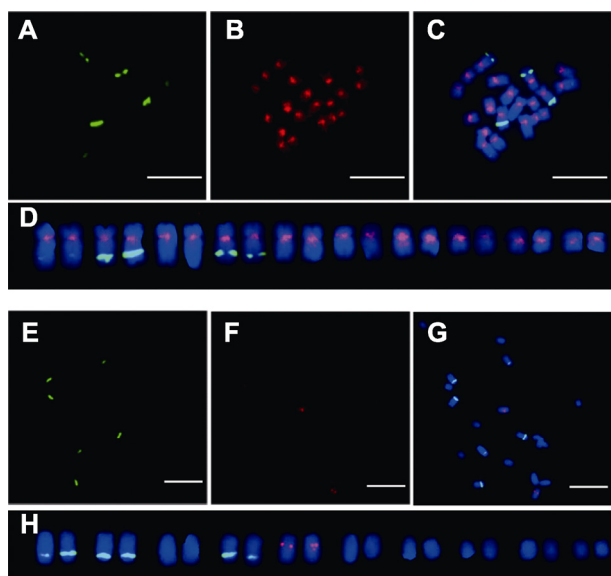


图3 石刁柏2个叶绿体反向重复区(IR)序列的PCR扩增  
1, 2分别代表AocpIR1和AocpIR2的扩增结果。

Figure 3 PCR amplification of inverted repeat region (IR) two sequences of region in *Asparagus officinalis*  
1, 2 respresents the amplification results of AocpIR1 and AocpIR2, respectively.

## 2.4 2个IR区序列的染色体定位

将扩增获得的2个IR序列进行回收纯化后标记探针,与45S rDNA同时进行双色荧光原位杂交。结果表明,45S rDNA共有6个信号,分布在3对同源染色体上(图4A, C, E)。AocpIR1在所有染色体上都有分布,信号主要分布在染色体着丝粒位置,但在着丝粒侧翼位置也有微弱信号(图4B)。AocpIR2特异定位在2条染色体的着丝粒部位(图4F, G),核型分析表明这2条染色体为石刁柏核型的第5对染色体(图4D, H),也正是石刁柏的性染色体(Löptien, 1979)。



**图4** AocpIR1和AocpIR2在石刁柏有丝分裂中期染色体上的定位  
(A), (E) 45S rDNA杂交信号图; (B), (F) AocpIR1和AocpIR2杂交信号图; (C), (G) 合成图; (D), (H) 核型图。Bars=10 μm

**Figure 4** Mitosis metaphase chromosome locations of AocpIR1 and AocpIR2 of *Asparagus officinalis*  
(A), (E) FISH result of 45S rDNA; (B), (F) FISH result of AocpIR1 and AocpIR2; (C), (G) Merged picture; (D), (H) Karyotype analysis. Bars=10 μm

## 2.5 讨论

本研究表明,叶绿体DNA序列不同区域向核基因组转移的频率存在差异,石刁柏叶绿体基因组中IR区序列更易向核基因组中进行水平转移,这与在白麦瓶草(Steflova et al., 2014)和水稻(Matsuo et al., 2005)中的研究结果一致。这种转移偏向性与IR区序列的独特特征关系密切。IR区为反向重复序列,在叶绿体基因

组中以相反的方向重复存在。这种结构可能有助于核基因组的插入重组。Wang等(2008)研究表明,被子植物的IR区包含1个polyA区域,这个区域可能是重组的热点。通过这种机制,其中1个IR区域利用第2个IR区中同源序列的修复/重组过程延伸到相邻的SC区中(Wang et al., 2008)。因此,推测类似的基因重组过程可促进IR区从叶绿体基因组中优先整合到核基因组中。也有观点认为大多数细胞器DNA通过非末端连接修复双链断裂机制转移到细胞核中,但是在其它一些植物,如酸模(*Rumex acetosa*)基因组中,IR区比叶绿体基因组中其它位置的序列更难转移形成NUPTs(Steflova et al., 2014)。这也表明核基因中NUPTs的插入机制更为复杂,仍需深入研究。

研究表明,植物中NUPTs更倾向于插入和累积在着丝粒及其侧翼区等染色体异染色质区域(Michalovova et al., 2013; Steflava et al., 2014)。在黑麦(*Secale cereale*) B染色体中,NUPTs也在基因含量很低的异染色质区域累积(Martis et al., 2012)。本研究定位的2个IR区序列中,其中1个定位于所有染色体的着丝粒及其侧翼区,而另1个特异定位于性染色体的着丝粒区,表明石刁柏的NUPTs也更倾向于插入到染色体的着丝粒及其侧翼区。这可能是由于这些异染色质区重组频率较低,而低重组率更有利于重复序列高度一致性的维持。Matsuo等(2005)研究表明,着丝粒及其侧翼区等异染色质区能更好地接纳大的叶绿体DNA或线粒体DNA序列,具有更适合形成NUPT/NUMT的基因组环境。但是,采用序列比对对NUPTs在核基因组位置的分析未显示NUPTs更倾向于插入着丝粒及其周围的特性。原因主要有2个:(1)基因组组装过程中,对于着丝粒及侧翼区域更不容易组装出来,导致这些区域中的NUPT序列没有在基因组序列中体现出来;(2)本研究只对IR区的2个序列进行FISH定位,可能IR序列的转移位置相比LSC和SSC具有更高的特异性。后续可以全面地对石刁柏叶绿体基因组的各个区域进行FISH定位,从而更全面地获得各个区域在染色体上的分布特征及转移特性。

目前普遍认为,性染色体最初起源于1对常染色体,在出现原始的性别决定基因之后,包括性别决定基因在内的性别决定区的重组抑制是性染色体演化过程的关键步骤(Kejnovsky et al., 2009; Li et al.,



2016; Abbott et al., 2017)。包括NUPTs在内的重复序列的插入造成异染色质化或倒位是重组抑制形成的主要原因, 而重组抑制区还可以招募更多的NUPTs等重复序列, 使重组抑制区进一步扩展, 最终形成异型的性染色体(Li et al., 2016)。本研究定位的2个序列中, 其中1个位于性染色体的着丝粒部位, 而该区域正是染色体不发生重组或很少发生重组的位置, 表明在雌雄异株植物石刁柏中, NUPTs的插入在其性染色体的演化过程中可能发挥作用。期待今后发现更多的性染色体特异的NUPTs, 并利用高分辨率的DNA纤维原位杂交结合免疫荧光来分析NUPTs插入区域附近染色质的DNA和组蛋白的表观修饰模式特征, 以深入揭示NUPTs在植物性染色体演化过程中所起的作用及其作用机制。

## 参考文献

- 高东迎, 何冰, 孙立华 (2007). 水稻转座子研究进展. 植物学通报 **24**, 667–676.
- 李巧丽, 延娜, 宋琼, 郭军战 (2018). 鲁桑叶绿体基因组序列及特征分析. 植物学报 **53**, 94–103.
- 李书粉, 李旭, 王冰肖, 袁金红, 邓传良, 高武军 (2016). 石刁柏雄性偏向核质体DNA的克隆与分析. 西北植物学报 **36**, 2385–2390.
- Abbott JK, Nordén AK, Hansson B (2017). Sex chromosome evolution: historical insights and future perspectives. *Proc Roy Soc B* **284**, 20162806.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403–410.
- Doyle JJ, Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* **19**, 11–15.
- Feschotte C, Pritham EJ (2007). DNA transposons and the evolution of eukaryotic genomes. *Annu Rev Genet* **41**, 331–368.
- Guo XY, Ruan SL, Hu WM, Cai DG, Fan LJ (2008). Chloroplast DNA insertions into the nuclear genome of rice: the genes, sites and ages of insertion involved. *Funct Integr Genomics* **8**, 101–108.
- Harkess A, Zhou JS, Xu CY, Bowers JE, van der Hulst R, Ayyampalayam S, Mercati F, Riccardi P, McKain MR, Kakrana A, Tang HB, Ray J, Groenendijk J, Arikrit S, Mathioni SM, Nakano M, Shan HY, Telgmann-Rauber A, Kanno A, Yue Z, Chen HX, Li WQ, Chen YL, Xu XY, Zhang YP, Luo SC, Chen HL, Gao JM, Mao ZC, Pires JC, Luo MZ, Kudrna D, Wing RA, Meyers BC, Yi KX, Kong HZ, Lavrijsen P, Sunseri F, Falavigna A, Ye Y, Leebens-Mack JH, Chen GY (2017). The asparagus genome sheds light on the origin and evolution of a young Y chromosome. *Nat Commun* **8**, 1279.
- Kejnovsky E, Hobza R, Cermak T, Kubat Z, Vyskot B (2009). The role of repetitive DNA in structure and evolution of sex chromosomes in plants. *Heredity* **102**, 533–541.
- Kejnovsky E, Kubat Z, Hobza R, Lengerova M, Sato S, Tabata S, Fukui K, Matsunaga S, Vyskot B (2006). Accumulation of chloroplast DNA sequences on the Y chromosome of *Silene latifolia*. *Genetica* **128**, 167–175.
- Kleine T, Maier UG, Leister D (2009). DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. *Annu Rev Plant Biol* **60**, 115–138.
- Li SF, Zhang GJ, Yuan JH, Deng CL, Gao WJ (2016). Repetitive sequences and epigenetic modification: inseparable partners play important roles in the evolution of plant sex chromosomes. *Planta* **243**, 1083–1095.
- Löptien H (1979). Identification of the sex chromosome pair in asparagus (*Asparagus officinalis* L.). *Z Pflanzenzücht* **82**, 162–173.
- Martis MM, Klemme S, Banaei-Moghaddam AM, Blattner FR, Macas J, Schmutzer T, Scholz U, Gundlach H, Wicker T, Šimková H, Novák P, Neumann P, Kubaláková M, Bauer E, Haseneyer G, Fuchs J, Doležel J, Stein N, Mayer KFX, Houben A (2012). Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. *Proc Natl Acad Sci USA* **109**, 13343–13346.
- Matsuo M, Ito Y, Yamauchi R, Obokata J (2005). The rice nuclear genome continuously integrates, shuffles, and eliminates the chloroplast genome to cause chloroplast-nuclear DNA flux. *Plant Cell* **17**, 665–675.
- Michalovova M, Vyskot B, Kejnovsky E (2013). Analysis of plastid and mitochondrial DNA insertions in the nucleus (NUPTs and NUMTs) of six plant species: size, relative age and chromosomal localization. *Heredity* **111**, 314–320.
- Richly E, Leister D (2004). NUMTs in sequenced eukaryotic genomes. *Mol Biol Evol* **21**, 1081–1084.
- Sheng WT, Chai XW, Rao YS, Tu XT, Du SG (2017). Complete chloroplast genome sequence of asparagus (*Asparagus officinalis* L.) and its phylogenetic position within asparagales. *J Plant Breed Genet* **5**, 121–128.
- Steflova P, Hobza R, Vyskot B, Kejnovsky E (2014).

- Strong accumulation of chloroplast DNA in the Y chromosomes of *Rumex acetosa* and *Silene latifolia*. *Cytogenet Genome Res* **142**, 59–65.
- Timmis JN, Ayliffe MA, Huang CY, Martin W** (2004). Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat Rev Genet* **5**, 123–135.
- VanBuren R, Ming R** (2013). Organelle DNA accumulation in the recently evolved papaya sex chromosomes. *Mol Genet Genomics* **288**, 277–284.
- Wang RJ, Cheng CL, Chang CC, Wu CL, Su TM, Chaw SM** (2008). Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genomes of monocots. *BMC Evol Biol* **8**, 36.
- Yoshida T, Furihata HY, Kawabe A** (2014). Patterns of genomic integration of nuclear chloroplast DNA fragments in plant species. *DNA Res* **21**, 127–140.

## Bioinformatics Analysis and Chromosome Location of Nuclear Integrants of Plastid DNA in *Asparagus officinalis*

Guangqian Cheng<sup>1†</sup>, Keli Jia<sup>1,2†</sup>, Na Li<sup>1</sup>, Chuanliang Deng<sup>1</sup>, Shufen Li<sup>1</sup>, Wujun Gao<sup>1\*</sup>

<sup>1</sup>College of Life Sciences, Henan Normal University, Xinxiang 453007, China; <sup>2</sup>SanQuan College, Xinxiang Medical University, Xinxiang 453003, China

**Abstract** Chloroplast DNA (cpDNA) can transfer into the plant nuclear genome to form nuclear integrants of plastid DNA (NUPTs). NUPTs may play a role in plant sex chromosome evolution. However, few studies have focused on this research area. In this study, we annotated and analyzed NUPTs in the genome of dioecious *Asparagus officinalis* and located two cpDNA fragments on chromosomes. The nuclear genome of *A. officinalis* contains 2 239 NUPT insertions, with total length 565 970 bp, accounting for 0.047% of the genome. The amount of NUPTs differs among chromosomes; the number, density, and the average length of NUPTs on the Y chromosome were all higher than those on the other chromosomes, which indicates more accumulation of NUPTs on sex chromosomes. All regions of inverted repeats (IRs) and small and large single copy regions of cpDNA could transfer into nuclear DNA; however, the IR region showed the highest transfer frequency. Furthermore, FISH analysis of two cpDNA sequences from IR regions showed that AocpIR1 distributed mainly on the centromeres of all chromosomes, whereas AocpIR2 specifically located on sex chromosomes of *A. officinalis*. The data provide an important foundation for determining the genome structure and sex chromosome evolution of *A. officinalis*.

**Key words** *Asparagus officinalis*, bioinformatic analysis, nuclear integrants of plastid DNA, NUPTs, sex chromosome

**Cheng GQ, Jia KL, Li N, Deng CL, Li SF, Gao WJ** (2019). Bioinformatics analysis and chromosome location of nuclear integrants of plastid DNA in *Asparagus officinalis*. *Chin Bull Bot* **54**, 328–334.

† These authors contributed equally to this paper

\* Author for correspondence. E-mail: gaowujun@htu.cn

(责任编辑: 朱亚娜)