芦笋基因组测序——拉开芦笋分子育种的序幕

周劲松^{1,2}, 罗绍春¹, 朱友林^{2*}, 陈光宇^{1*}

1. 江西省农业科学院蔬菜花卉研究所, 南昌 330200;

2. 南昌大学生命科学学院, 南昌 330031

* 联系人, E-mail: genebksh@hotmail.com; ylzhu1999@aliyun.com

收稿日期: 2017-12-06; 接受日期: 2018-03-16; 网络版发表日期: 2018-07-10

江西省农业科学院联合美国乔治亚大学植物生物学系、华大基因、中国科学院植物研究所、中国热带农业科学院生物技术研究所、云南农业大学等国际合作团队(美、中、荷、意、德、日、泰7个国家27个研究机构), 历时7年研究, 合作完成了世界上首个芦笋(*Asparagus officinalis* L.)全基因组序列图谱绘制。采用二代测序(SBS)+三代测序(PacBio)+单分子光学图谱(BioNano)策略, 对芦笋雄性双单倍体植株进行全基因组(1.3 Gb/1C)从头测序、组装、拼接优化, 获得1.205 Gb大小的参考基因组, contig N50高达88.16 kb, scaffold N50达到1.24 Mb, 将93.7%组装序列挂载到10条染色体上, 结合从头、同源和转录组数据预测得到27656个编码基因; 在Y染色体上锁定一个长约1 Mb非重组性别决定区域, 从中鉴定出雌性抑制基因(*SOFF*)和雄性特异基因(*aspTDF1*), 它们独立连锁决定着芦笋性别, 促进天门冬属早期阶段性染色体演化, 从分子水平上直接证实了Westergaard^[1]和Charlesworth^[2]关于植物性染色体起源演化经典的“双基因”假说(two-gene hypothesis)和“双基因座”模型(two-locus model), 雌雄异株植物性别决定和性染色体进化的分子机制取得了重大进展。研究结果发表在*Nature Communications*上^[3]。

1 研究背景

自20世纪以来, 人们基于细胞学和遗传学的研究结果, 发现许多雌雄异体的植物、动物和真菌都具有性染色体, 但是关于性染色体是如何起源的以及它们在早期是如何从常染色体演化而来的, 仍然不清楚。在后生动物中, 性别分化与性染色体伴生起源是很古老的。相比之下, 植物进化出性别是很晚近的事件, 雌雄异株植物在显花植物进化过程中经历了多次起源并独立演变, 进化多样性促使被子植物是研究不同阶段性染色体起源演化的理想系统。雌雄异株植物仅占被子植物的6%, 目前仅在极少数雌雄异株植物中发现了性染色体, 这些植物性染色体上性别决定基因与它们组成结构仍然不很清楚。1958年Westergaard^[1]首先提出, 性染色体是从常染色体演化而来的, 具有活性的Y染色体至少涉及两个基因: 一个显性的雌性抑制基因和一个X染色体上缺失的雄性功能必需基因。1978年Charlesworth^[2]正式明确了性染色体从常染色体演化而来的“双基因座”模型, 常染色体在向性染色体演变同时伴随雌雄同体向雌雄异体性别表型转变(图1)。雌雄异株植物的“双基因座”模型很直观, 具有很好的理论指导意义, 但是, 缺乏直接的实证支持^[3]。或许, 一个独立的性别决定基因能够显性抑制雌性发育并促进雄

引用格式: 周劲松, 罗绍春, 朱友林, 等. 芦笋基因组测序——拉开芦笋分子育种的序幕. 中国科学: 生命科学, 2018, 48: 1128–1132

Zhou J S, Luo S C, Zhu Y L, et al. Genome sequencing of *Asparagus officinalis* L.: the beginning of dioecious garden asparagus molecular breeding (in Chinese). *Sci Sin Vitae*, 2018, 48: 1128–1132, doi: 10.1360/N052017-00292

图 1 芦笋雌雄异株进化途径的“双基因座”模型^[3]

A: 雌雄同株祖先常染色体上存在一对雄性发育必需基因(*aspTDF1*) (绿圈); B: 一个*aspTDF1*等位基因功能缺失突变(红圈); C: *aspTDF1*基因功能缺失的隐性突变体自交分离, 产生纯合的雌性个体; D: 在*aspTDF1*基因附近复制插入一个雌性功能抑制基因*SOFF* (蓝圈). 携带原始Y染色体的个体是雄性的, 但是不育基因座杂交重组, 可产生无性(雌、雄均不育)或雌雄同株的个体; E: 无性或雌雄同株的个体不利于进化选择, 导致Y染色体上的重组抑制(紫色区域)产生; F: 原始常染色体基因座(小灰圈)获得性别拮抗功能, 或转位至性别决定区域(SDR, 小紫罗兰圈); G: 性别拮抗选择可导致非重组区的延伸, 将性别拮抗基因座与SDR连接一起

性功能, 即“单基因座”模型^[4]. 2014年Akagi等人^[5]研究发现, 一个编码sRNA的基因*OGI*控制着雌雄异株的柿子(*Diospyros Kaki* L.f)雌雄性别决定, 上述“双基因座”模型似乎受到挑战.

芦笋是天门冬科(*Asparagaceae*)天门属(*Asparagus*)典型的雌雄异株植物, 具有雄性异配子型的X和Y性染色体, 且X和Y染色体细胞学上是同型的, XY为雄

性, XX为雌性, 但自然群体中偶尔会发现极少数雄性株(XY)会产生两性完全花, 俗称两性株, 引起自花或异花授粉结少量果实, 产生雄雌比例3:1的后代($XY \times XY \rightarrow 1YY + 2XY + 1XX$ 即3♂+1♀), 其中基因型YY (超雄株)是有生活力的, 能够产生可育的花粉粒, 由此可推测, 芦笋X, Y染色体出现处于植物性染色体演化初期. 在芦笋雄性和雌性的花器官发育早期, 两者

花器官都有心皮和雄蕊(雌雄同体时期),然而,在随后的发育进程中,心皮在雄性个体中停止发育,而雄蕊在雌性个体中退化,这也暗示着芦笋中存在双基因座性别决定系统。在XX雌性个体中,花药发育终止在小孢子产生之前,标志是绒毡层降解;在雄性个体中,花粉管和柱头停止发育但不退化,形成正常功能的雄花,并含有退化程度不同的雌性器官,从无花柱到含花柱乃至柱头的雌蕊,芦笋雌雄异体化很可能是经过全雌异株途径。进一步系统发育分析表明,天门冬属雌雄异体化是经过了1~2次进化而形成的,同时伴随着基因组大小和重复序列的增加,其分布区域也从非洲南部扩展到了北部及欧洲。这些研究表明了芦笋及所在的天门冬属雌雄异株与性染色体形成是非常晚近期发生的进化事件。为此,国际合作团队使用芦笋作为天门冬属雌雄异株分支代表,直接检验天门冬属雌雄异株性别决定起源究竟属于“单基因”还是“双基因”遗传系统,探究天门冬属性别决定遗传学基础与性染色体起源进化分子机制。

2 主要研究结果

为了探明芦笋Y染色体非重组区域的结构和基因组成,研究者以一个花药离体培养而来的雄性双单倍体超雄株(YY)为材料,测序组装,首先获得芦笋参考基因组。然后对74个XX或YY双单倍体子代组成的DH群体,以及花药供体XY、双亲XX和YY进行重测序,构建了拥有10条连锁群的芦笋基因组遗传图谱,将组装序列挂载到对应的10条染色体上。根据基因组遗传图谱雌雄性别连锁分离模式与以前定位的Y染色体连锁标记(Asp1-T7)^[6],在132 Mb的Y染色体上鉴定出长为1 Mb非重组性别决定区域,约占Y染色体0.75%,大部分呈现为半合子,反转录转座子分布密度大。从中预测出13个候选基因,包括一个DUF247结构域基因和一个单拷贝拟南芥(*Arabidopsis thaliana*) *TDF1* (defective in tapetum development and function 1)^[7]同源基因,其中12个基因在XX雌性中缺失。

随后,对3个 γ 射线诱变的从雄性向雌性转换的性别突变体进行重测序,结果发现突变体Y染色体整个非重组区都缺失,突变事件表明,非重组区域包含了雄性性别决定所必需的所有基因,即雄性激活基因(male activator)和雌性抑制基因(female suppressor)。同

时,对一个从雄性向雌雄同体转换的性别突变体进行重测序结果发现,突变体Y染色体连锁的DUF247结构域基因发生碱基缺失;另一个雄性XY到雌雄同体转换的自发突变体转录组测序与基因靶向测序结果发现,同样的Y染色体连锁的DUF247结构域基因发生一个移码突变;这两个独立的突变事件表明Y染色体特异的DUF247结构域基因对雌性抑制起关键作用,这个基因被命名为 *SOFF* (suppressor of female function)。这两个具完全花的雌雄同株突变体花药发育正常,表明非重组区域至少有一个Y染色体特异基因在促进雄性功能;基于拟南芥*TDF1*基因缺失突变导致的雄性不育性状,单拷贝的Y染色体*aspTDF1*基因被推断为一个与芦笋性别连锁的促进雄性发育的候选基因。

为了检测芦笋是否像柿子中存在非编码sRNA调控雌雄性别,对芦笋雌、雄植株营养组织和不同生殖组织样品的sRNA表达丰度进行比较分析,结果发现,雌、雄株具有sRNA差异表达模式,6个sRNAs始终在雄性植株嫩茎中有表达偏向性,但没有来自Y染色体非重组区且涉及sRNA转录后调控的基因。根据miRNA表达谱分析,miRNA性别偏向表达可能在花药或雌蕊发育途径临界点受到性别决定基因*aspTDF1*与*SOFF*分别诱导作用,芦笋中未发现类似*OGI*基因编码sRNA控制雌雄性别,即芦笋性别决定不是受“单基因”模式控制。

天门冬属包含超过200个种,其中绝大部分是雌雄同株植物,以前研究表明雌雄异株类群集中在系统发育树的一支当中。为了研究*SOFF*和*aspTDF1*基因是否在天门冬属中均以雄性特异的性别决定基因发挥作用,并导致雌雄异株的产生,研究者对一组XY雄性和XX雌性天门冬(*A. cochinchinensis*)、雌雄同株的扫状天门冬(*A. virgatus*)进行了全基因组重测序。基因同源性分析发现,天门冬和芦笋的常染色体以及扫状天门冬基因组中也分布有DUF247结构域基因,但是天门冬与芦笋Y染色体上DUF247结构域基因1号内含子区域都具有较长的可读框,基于3个物种的DUF247结构域基因与*SOFF*基因的系统发育分析可推测认为天门冬Y染色体上DUF247结构域基因与芦笋的*SOFF*基因是直系同源,天门冬XY雄性个体有*SOFF*直系同源基因,而在XX雌性个体中未发现。鉴于天门冬与芦笋在天门冬属系统发育树上雌雄异体分支中歧化很早,上述研究结果支持了早期DUF247结构域基因复制事件

为天门冬属雌性功能抑制与雌雄异株起源创造了条件. 而天门冬XY雄性和XX雌性个体中都发现有 $aspTDF1$ 基因的外显子, 且没有核苷酸差异, 天门冬和*A. stipularis*雌性基因组中均存在全长的 $TDF1$ 同源基因^[8], 因此, $aspTDF1$ 可能对天门冬属雌雄异体起源没有贡献. 但是, 天门冬XY雄性与XX雌性个体 $aspTDF1$ 基因座存在重组抑制, 这暗示着天门冬属雌雄异株原始Y染色体起源以后, 芦笋Y染色体非重组的雄性特异区域要么是扩张包含了 $aspTDF1$, 要么 $aspTDF1$ 是从基因组其他地方招募而来的. 最后, 共线性和进化树分析结果清晰地表明, 天门冬目两次全基因组复制事件均早于天门冬属出现以及天门冬属雌雄异株物种起源, 芦笋中多倍化事件的发生先于雌雄异株出现.

总而言之, 论文中提供的主要数据支持“双基因座”模型, 天门冬属具有功能活性的Y染色体进化受基因复制和雌性抑制新功能化以及随后的非重组的性别决定区域的扩张所促成的.

3 理论意义与应用价值

基于YY与XX基因组比较分析, 精确定位了Y染色体非重组的性别决定区域, 从中鉴定出雌性抑制基因($SOFF$)和雄性特异基因($aspTDF1$), 它们独立连锁决定着芦笋雌雄性别, 发现经典遗传学中芦笋性别决定基因(M)其实是Y染色体上一个非重组的多基因座, 深入揭示了芦笋性别决定的遗传基础; 分析发现 $SOFF$ 促进天门冬属早期阶段性染色体演化, 从分子水平上证实了Westergaard^[1]和Charlesworth^[2]关于植物性染色体起源进化经典的“双基因座”模型, 植物性染色体进化的分子机制取得了重大进展. 最近Pannell和Gerchen^[3]在*Current Biology*对本文进行了专门评述认为, 研究结果引人入胜, 两个性别决定基因分离鉴定更是令人兴奋的, 但认为雌性抑制基因($SOFF$)和雄性特异基因($aspTDF1$)出现的先后次序并不与雌雄异株植物起源进化的“双基因座模型”之全雌异株途径相符, 需要进一步深入研究.

与此同时, 芦笋基因组测序完成, 为芦笋功能基因组学和分子育种提供高质量的参考蓝图, 拉开了芦笋功能基因组学与分子育种序幕. 芦笋是全球重要的经济作物, 其嫩茎被誉为“蔬菜之王”、“世界十大名菜”之一. 然而, 芦笋长期异交, 基因型高度杂合, 多年生, 且以动态生长的嫩茎为主要收获对象, 许多重要农艺性状受微效多基因控制, 育种周期长、难度大, 育种工作中对基因型选择的依赖尤其突出. 因此, 芦笋全基因组序列的解析不仅有助于理解植物性别染色体的起源和演化模式, 而且为培育具有高产、优质、抗病虫害、抗逆等优良性状的芦笋新品种打下了良好的基础, 将对推动世界芦笋产业科技进步具有十分重大的理论和实践意义.

值得提及的是, 芦笋属于茎菜, 在生产上, 雄株因不结果实消耗养分少, 产量比同期生长条件相同的雌株高出25%以上, 由全部雄株组成的全雄品种很受欢迎, 目前全雄育种是全球芦笋主要育种目标之一. 相比雌株, 芦笋雌株高大, 发茎数较少, 但嫩茎粗壮, 笋品质好, 果实也可用来酿酒和提取多糖、色素. 芦笋是虫媒授粉, 近年来随着我国南方芦笋大棚避雨栽培技术和防虫网的广泛应用, 开花授粉期棚内温度较高, 花粉粒活性受影响, 无虫媒, 雌株基本上能够不结果实, 不多消耗养分, 可充分发挥雌株优势, 产量相应提高, 品质佳, 因此, 本研究组提出“全雌栽培”和“全雌育种”. 但是, 芦笋除了花、果实等繁殖器官外, 雄株和雌株没有显著外观差别, 且它从播种到开花大约需要2年时间, 故早期难以直接区分雌、雄株. 利用雌性抑制基因($SOFF$)和雄性特异基因($aspTDF1$)以及非重组区域其他性别特异基因, 开发与性别共分离的DNA分子标记辅助雌雄性别和超雄株苗期筛选鉴定, 将可极大地提高芦笋性别定向育种效率, 包括全雄育种($XX \times YY \rightarrow XY$)和全雌育种($XX \times XY^S \rightarrow XX$, X花粉粒可育, Y^S 花粉粒致死), 同时雄株(XY) $SOFF$ 基因敲除可创制出雌雄同株突变体, 像水稻(*Oryza sativa*)等雌雄同花作物一样具有完全花, 为芦笋自交系和遗传群体构建打开方便之门, 预计全雌育种将成为芦笋育种研究热点之一.

参考文献

- 1 Westergaard M. The mechanism of sex determination in dioecious plants. *Adv Genet*, 1958, 9: 217–281
- 2 Charlesworth B, Charlesworth D. A model for the evolution of dioecy and gynodioecy. *Am Nat*, 1978, 112: 975–997

- 3 Pannell J R, Gerchen J. Sex determination: sterility genes out of sequence. [Curr Biol](#), 2018, 28: R80–R83
- 4 Harkess A, Zhou J, Xu C, et al. The asparagus genome sheds light on the origin and evolution of a young Y chromosome. [Nat Commun](#), 2017, 8: 1279
- 5 Akagi T, Henry I M, Tao R, et al. A Y-chromosome-encoded small RNA acts as a sex determinant in persimmons. [Science](#), 2014, 346: 646–650
- 6 Jamsari A, Nitz I, Reamon-Büttner S M, et al. BAC-derived diagnostic markers for sex determination in asparagus. [Theor Appl Genet](#), 2004, 108: 1140–1146
- 7 Zhu J, Chen H, Li H, et al. *Defective in Tapetal Development and Function 1* is essential for anther development and tapetal function for microspore maturation in *Arabidopsis*. [Plant J](#), 2008, 55: 266–277
- 8 Tsugama D, Matsuyama K, Ide M, et al. A putative MYB35 ortholog is a candidate for the sex-determining genes in *Asparagus officinalis*. [Sci Rep](#), 2017, 7: 41497