

基于马铃薯腐烂茎线虫基因组的 SSR 信息分析

马居奎, 孙厚俊, 张成玲, 杨冬静, 谢逸萍*

(江苏徐淮地区徐州农业科学研究所, 农业部甘薯生物学与遗传育种重点实验室, 江苏 徐州 221131)

摘 要: 为了研究马铃薯腐烂茎线虫在我国的遗传关系, 获得稳定的马铃薯腐烂茎线虫分子标记, 使用 MISA 软件对马铃薯腐烂茎线虫全基因组 1 761 个 scaffolds 进行搜索, 共获得 9 745 个 SSR 位点, 平均 11.40 kb 出现一个 SSR。在重复基元中, 单核苷酸基元出现频率最高(83.06%), 其次为二核苷酸重复基元(10.62%); 除单核苷酸重复基元外出现次数最多的重复基元是 AT/AT(5.36%)。马铃薯腐烂茎线虫基因组 SSR 位点长度小于 12 bp 分布最为广泛, 占总 SSR 的 59.67%, 长度大于 30 bp 的最少, 共有 114 个, 仅占 1.17%。研究表明, 利用马铃薯腐烂茎线虫基因组数据可作为开发 SSR 标记的有效来源, 为进一步开发马铃薯腐烂茎线虫 SSR 标记奠定了基础。

关键词: 马铃薯腐烂茎线虫; 全基因组序列; SSR 标记

中图分类号: S532

文献标识码: A

文章编号: 1672-755X(2019)02-0070-05

Analysis of SSR Information in Genome Resource of *Ditylenchus destructor* Thorne

MA Ju-kui, SUN Hou-jun, ZHANG Cheng-ling, YANG Dong-jing, XIE Yi-ping*

(Xuzhou Institute of Agricultural Sciences in Jiangsu Xuhuai Area, Key Laboratory of Biology and Genetic Improvement of Sweet Potato, Xuzhou 221131, China)

Abstract: In order to study the genetic relation of *Ditylenchus destructor* Thorne in China, and to obtain their stable molecular marker, the software MISA was used to search and analyze a total of 1 761 scaffolds of whole genome sequences of 9 745 SSR of *D. destructor* Thorne. This was equivalent to 1 SSR per 11.40 kb genome sequence. In the SSRs, the mononucleotide repeat motifs were the most abundant (83.06%), followed by dinucleotide repeat motifs (10.62%). In addition to mononucleotide repeats, AT/AT was the most frequent repeat unit (5.36%). The SSR which had a length less than 12 bp accounted for 59.67% of total SSR, while the number of SSR over 30 bp was 114, only accounting for 1.17% at the lowest proportion. The results indicated that the genomic data of *D. destructor* Thorne could be used as an effective source for the development of SSR markers, which could serve as a foundation for the further development of SSR markers.

Key words: *Ditylenchus destructor* Thorne; whole genome sequences; simple sequence repeat (SSR) marker

甘薯茎线虫病又称糠心病、空心病、糠榔病等, 是一种重要的植物寄生性病害, 其病原线虫为马铃薯腐烂茎线虫(*Ditylenchus destructor* Thorne)。马铃薯腐烂茎线虫在世界范围内均有分布, 现有报道的发生区域包

收稿日期: 2019-05-15

基金项目: 国家甘薯产业技术体系项目(CARS-10-B15)

作者简介: 马居奎(1991—), 男, 山东寿光人, 研究实习员, 硕士, 主要从事甘薯病虫害研究。

通信作者: 谢逸萍(1962—), 女, 广东罗定人, 研究员, 主要从事甘薯病虫害研究。

括亚洲、非洲、北美洲、南美洲、大洋洲和欧洲。该线虫已被亚太植物保护组织及许多国家和地区列为重要的植物检疫性有害生物^[1]。我国自首次发现马铃薯腐烂茎线虫以来,已对其开展形态学、生物学、抗病品种筛选和病害防治技术等研究,然而关于我国马铃薯腐烂茎线虫种群遗传多样性和种群遗传结构的研究较少。

分子标记技术是研究群体遗传学与种群遗传变异的主要工具之一。利用群体遗传学与生态因子的关系来估计物种的传播扩散方式,对于了解不同物种的繁殖方式、遗传结构及基因流等具有重要意义^[2]。常用的分子标记有很多种,包括限制性片段长度多态性(Restriction Fragment Length Polymorphism, RFLP)、随机扩增片段长度多态性(Random Amplified Polymorphie DNA, RAPD)、扩增片段长度多态性(Amplified Fragment Length Polymorphism, AFLP)、微卫星标记又称简单重复序列(Simple Sequence Repeat, SSR)、单核苷酸多态性(Single Nucleotide Polymorphisms, SNP)等^[3]。分子标记在生物种群遗传多样性研究中具有广泛的应用^[4-7]。SSR 标记具有分辨率高、重复性好、稳定性强、共显性、反应速度快和操作简便等优点,得到广泛的应用。

近年来,第二代测序技术(Next Generation Sequencing, NGS)的发展,极大促进了微卫星标记的开发^[8]。基于 NGS 构建微卫星文库,能显著提高微卫星开发的效率,加速物种遗传多样性的研究进程^[9-10]。

1 材料与方法

马铃薯腐烂茎线虫(*D. destructor* Thorne)的基因组数据来自 NCBI(美国国家生物技术信息中心)数据库(<https://www.ncbi.nlm.nih.gov/>)。

运用 MISA software (<http://pgrc.ipk-gatersleben.de/misa/>)对基因组序列进行微卫星查找,设置参数为 1-10、2-6、3-5、4-5、5-5 和 6-5,两个微卫星之间间隔最少 100 bp。以 SSR 出现的频率和 SSR 平均分布距离描述 SSR,同时认为出现频率前 2 位的重复基元为优势重复基元。SSR 出现频率 $f_c = c/n \times 100\%$, c 为检索到的 SSR 数量, n 为 scaffolds 序列数量。SSR 平均分布距离 $f_N = N/c$, N 为基因组的总碱基数。

2 结果与分析

2.1 基因组数据

在 NCBI 数据库平台公布了马铃薯腐烂茎线虫的全基因组信息,其大小为 111.14 Mb,由 1 761 个 scaffolds 组成,预测有 13 938 个蛋白质编码基因。马铃薯腐烂茎线虫基因组 G+C 含量为 44%,高于南方根结线虫(*Meloidogyne incognita*)、北方根结线虫(*M. hapla*)、马铃薯白线虫(*Globodera pallida*)、秀丽隐杆线虫(*Caenorhabditis elegans*)、松材线虫(*Bursaphelenchus xylophilu*)和马来丝虫(*Brugia malayi*)的 G+C 含量(表 1)。

表 1 马铃薯腐烂茎线虫基因组与其他六种线虫基因组的比较

线虫种类	Estimated genome size/Mb	Assembly length/Mb	Scaffolds/n	Scaffold N50/kb	GC content/%	Gene models/n	Number of proteins/n
马铃薯腐烂茎线虫 <i>D. destructos</i>	—	111.14	1761	47	44	18 074	13 938
南方根结线虫 <i>M. incognita</i>	47~51	86	2 187	83	31.4	19 212	20 365
北方根结线虫 <i>M. hapla</i>	54	53	1 523	84	27.4	14 420	13 072
马铃薯白线虫 <i>G. pallida</i>	100	124.7	6 873	122	36.7	16 419	16 419
秀丽隐杆线虫 <i>C. elegans</i>	100	100	—	—	35.4	20 056	24 890
松材线虫 <i>B. xylophilus</i>	63~75	74.6	1 231	1158	40.4	18 074	18 074
马来丝虫 <i>B. malayi</i>	90~95	95.8	8 180	94	30.5	18 348	21 252

2.2 基因组 SSR 分布频率

在搜索的 1 761 个 scaffolds 序列中共查找到 SSR 位点 9 745 个,分布在 602 个 scaffolds 序列中,发生频率为 34.19%,单碱基~六碱基重复序列均有存在;其中 420 个 scaffolds 序列中含有多个 SSR 基序;复合 SSRs 数量为 2 338 个,占总 SSR 数量的 24.0%。从分布情况来看,马铃薯腐烂茎线虫基因组中 SSR 的分布概率为 1 SSR/11.40 kb。单核苷酸重复基元出现频率最高,出现频率(f_c)为 83.06%,平均分布距离(f_N)为 13.73 kb;其次为二核苷酸重复基元, f_c 为 10.62%, f_N 为 107.37 kb,其余核苷酸重复基元出现的频率都较低(表 2)。

表 2 马铃薯腐烂茎线虫 SSR 类型分布

重复基元	不同重复次数时的出现次数									SSR 位点数	出现频率 (f_c)/%	分布距离 (f_N)/kb
	5	6	7	8	9	10	11	12	>12			
单核苷酸	—	—	—	—	—	3 995	1 425	1 042	1 632	8 094	83.06	13.73
二核苷酸	—	554	195	109	58	44	29	16	30	1 035	10.62	107.37
三核苷酸	276	53	16	9	1	1	4	1	4	365	3.75	304.49
四核苷酸	98	42	13	8	5	2	1	1	8	178	1.82	624.37
五核苷酸	46	14	2	2	3	—	1	—	—	68	0.70	1 634.38
六核苷酸	3	—	—	1	—	—	—	—	1	5	0.05	22 227.64
总计	423	663	226	129	67	4 042	1 460	1 060	1 675	9 745	—	11.40

2.3 基因组 SSR 的特性

在检索出的马铃薯腐烂茎线虫的基因组中,总共观察到 161 种重复基元,其中单、二、三、四、五、六核苷酸重复基元的种类分别有 4、12、49、69、22、5 种(表 3)。其中,单核苷酸重复基元在基因组中出现频率最高,达到 83.06%,六核苷酸重复基元出现频率最低,仅为 0.05%。

表 3 重复基元类型的优势重复基元的比较分析

重复基元	重复基元种类	优势重复基元	SSR 数量	频率/%
单核苷酸	4	A/T	4 199	43.09
		C/G	3 895	39.97
二核苷酸	12	AT/AT	522	5.36
		AG/CT	250	2.57
三核苷酸	49	AAT/ATT	246	2.52
		AAG/CTT	36	0.36
四核苷酸	69	AAGT/ACTT	52	0.53
		ACAG/CTGT	17	0.17
五核苷酸	22	ACCGG/CCGGT	43	0.44
		AAATT/AATTT	5	0.05
六核苷酸	5	AAGCCT/AGGCTT	2	0.02
		AGGCGG/CCGCCT	2	0.02
总计	161		9 269	

从图 1 看出,大部分马铃薯腐烂茎线虫基因组 SSR 位点长度小于 12 bp,共有 5 815 个,占总 SSR 的 59.67%;位点长度分布在 12~19 bp 的共有 3 275 个,占总 SSR 的 33.61%;位点长度分布在 20~30 bp 的共有 541 个,占总 SSR 的 5.55%;位点长度大于 30 bp 的最少,共有 114 个,仅占 1.17%。

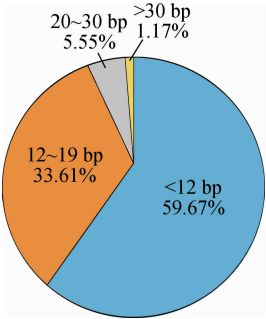


图 1 马铃薯腐烂茎线虫基因组 SSR 位点长度分布

3 结论与讨论

微卫星标记是分析种内遗传的重要工具,目前针对马铃薯腐烂茎线虫遗传多样性研究的报道相对较少,且都聚焦于少量地理群体的 ITS-RFLP 图谱和 ISSR 标记分析^[11-12],然而由于 ITS-RFLP 图谱只针对于核糖体 DNA-ITS 区,其多态性无法反映腐烂茎线虫种群遗传分化情况;ISSR 技术尽管引物设计比 SSR 简单,产物多态性丰富,但由于其采用通用引物,而重复序列和锚定碱基的随机性,可能会造成遗传信息的丢失,此外呈显性遗传标记的特性,且不能区分显性纯合基因型和杂合基因型^[13],尚不能全面揭示马铃薯腐烂茎线虫不同群体的遗传结构差异。

本文对数据库中下载的马铃薯腐烂茎线虫基因组数据进行微卫星位点检索,共获得微卫星位点 9 745 个,SSR 的分布概率为 1 SSR/11.4 kb。马铃薯腐烂茎线虫基因组 SSR 主要为单、二核苷酸重复基元类型,这与其他植物寄生线虫基因组 SSR 的报道相类似^[10,14],而与植物寄生线虫转录组 SSR 主要为三核苷酸重复基元类型的报道有所不同^[15-16]。可能是由于非三碱基重复会引起移码突变而导致的定向选择,三核苷酸重复序列通常出现在编码区,而二核苷酸序列通常在非编码区出现的频率更高^[17]。

已知的研究表明,SSR 基元长度不小于 20 bp 时,在不同群体中具有较高的多态性,长度在 12~19 bp 的 SSR 具有中度多态性^[18],马铃薯腐烂茎线虫基因组中不小于 12 bp 的 SSR 位点共 3 930 个,占总 SSR 位点的 40.33%,这部分 SSR 位点多态性潜能较高,这在预测马铃薯腐烂茎线虫的遗传多样性分析方面具有较高的应用价值。本文表明利用马铃薯腐烂茎线虫基因组数据库进行 SSR 标记具有可行性,为进一步开发马铃薯腐烂茎线虫 SSR 标记奠定了基础。

参考文献:

- [1] 周忠,马代夫.甘薯茎线虫病的研究现状和展望[J].园艺与种苗,2003,23(5):288-290
- [2] Picard D,Plantard O,Scurrah M,et al. Inbreeding and population structure of the potato cyst nematode(*Globodera pallida*) in its native area(Peru)[J]. Molecular Ecology,2004,13(10):2899-2908
- [3] 周延清. DNA 分子标记技术在植物研究中的应用[J]. 生命科学,2005,17(3):15-17
- [4] Panwar P,Nath M,Yadav V K,et al. Comparative evaluation of genetic diversity using RAPD,SSR and cytochrome P450 gene based markers with respect to calcium content in finger millet(*Eleusine coracana* L. Gaertn.)[J]. Journal of Genetics,2010,89(2):121-133
- [5] Nascimento W,Frodrigues J F,Koehler S,et al. Spatially structured genetic diversity of the Amerindian yam(*Dioscorea trifida* L.)assessed by SSR and ISSR markers in Southern Brazil[J]. Genetic Resources & Crop Evolution,2013,60(8):2405-2420
- [6] Muys C,Thienpont C N,Dauchot N,et al. Integration of AFLPs,SSRs and SNPs markers into a new genetic map of industrial chicory(*Cichorium intybus* L. var. *sativum*)[J]. Plant Breeding,2014,133(1):130-137
- [7] Eoche-Bosy D,Gauthier J,Juhel A S,et al. Experimentally evolved populations of the potato cyst nematode *Globodera pallida* allow the targeting of genomic footprints of selection due to host adaptation[J]. Plant Pathology,2017,66(6):1022-1030
- [8] Metzker M L. Sequencing technologies—the next generation[J]. Nature Reviews Genetics,2010,11(1):1-13
- [9] Takeuchi T,Yamaguchi M,Tanaka R,et al. Development and validation of SSR markers for the plant-parasitic nematode *Subanguina moxaeusing* genome assembly of illumina pair-end reads[J]. Nematology,2015,17(5):515-522
- [10] Wang G,Li E F,Mao Z C,et al. Development of polymorphic microsatellites for *Meloidogyne incognita*, through screening predicted microsatellite loci based on genome sequence[J]. Genetika,2015,51(1):128-132
- [11] 季镭,王金成,杨秀丽,等. 3 种茎线虫 rDNA 区的 PCR-RFLP 分析[J]. 南京农业大学学报,2006,29(3):39-43
- [12] 黄文坤,何旭峰,丁中,等. 马铃薯腐烂茎线虫的遗传变异及其与生态因子的相关性[J]. 农业生物技术学报,2012,20(1):62-69
- [13] Zietkiewicz E,Rafalski A,Labuda D. Genome fingerprinting by simple sequence repeat (SSR)—anchored polymerase chain reaction amplification[J]. Genomics,1994,20(2):176-183

- [14] 许峻荣,吴小芹,刘云,等. 基于松材线虫全基因组序列的 SSR 标记开发[J]. 南京林业大学学报,2014,38(2):36 – 42
- [15] Wang H M,Zhao H H,Zhao C Z, et al. EST-SSR markers from *Heterodera glycines* Ichinohe[J]. Genetika,2014,50(10):1259 – 1262
- [16] Wang X,Ma J,Liu H,et al. Development and characterization of EST-derived SSR markers in the cereal cyst nematode *Heterodera avenae*[J]. European Journal of Plant Pathology,2018,150(1):105 – 113
- [17] Metzgar D,Bytof J,Wills C. Selection against frameshift mutations limits microsatellite expansion in coding DNA[J]. Genome Research,2000,10(1):72 – 80
- [18] Temnykh S,DeClerck G,Lukashova A,et al. Computational and experimental analysis of microsatellites in rice(*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential[J]. Genome research, 2001,11(8):1441 – 1452

(责任编辑:湛 江)