

蜱螨线粒体基因组串联重复序列 DNA 的分布与进化特征分析*

祝 银^{1,2**} 李 一^{1,2} 蒋诗蝶^{1,2} 王进军^{1,2} 魏丹丹^{1,2***}

(1. 昆虫学与害虫控制工程重庆市重点实验室, 西南大学植物保护学院, 重庆 400715;

2. 长江上游农业生物安全与绿色生产教育部重点实验室, 西南大学农业科学研究院, 重庆 400715)

摘 要 【目的】 本研究旨在明确蜱螨线粒体基因组 (mtDNA) 串联重复序列 (微卫星和长串联重复 DNA) 的分布规律和进化特征, 为开发蜱螨分子标记和研究线粒体微卫星的功能提供基础数据。【方法】 采用 Microsatellite identification tool (MISA) 和 Tandem repeat finder 等本地软件发掘蜱螨线粒体基因组串联重复序列 DNA, 分析其组成类型、分布规律和进化特征。【结果】 本研究基于 37 科 217 种蜱螨线粒体基因组, 鉴定获得 1 454 个完美型线粒体微卫星 (mtSSRs) 和 96 个复合型微卫星。其中, 完美型 mtSSRs 包含 271 个单碱基微卫星 (SSR1)、127 个双碱基微卫星 (SSR2)、517 个三碱基微卫星 (SSR3)、323 个四碱基微卫星 (SSR4)、75 个五碱基微卫星 (SSR5) 以及 46 个六碱基微卫星 (SSR6)。分析发现, 随着微卫星核心重复碱基数目的增加, 其重复次数范围逐渐变小。蜱螨大多数 mtSSRs 倾向分布于线粒体电子传递链复合体 I 所编码的 *nad4* 和 *nad5* 基因之中。同时, 本文鉴定获得 167 个长串联重复序列 (核心重复碱基数>6 bp), 且 50.3% 分布于蛋白编码区。但核心重复基序较长的片段类型主要分布在特定物种的 AT 控制区。此外, 革蜱属 *Dermacentor* 多数物种的线粒体基因组含有 2 个长串联重复序列区域。【结论】 蜱螨线粒体基因组中三碱基微卫星和四碱基微卫星的数量最为丰富, 且多分布于线粒体蛋白编码区域。同时, 蜱螨线粒体微卫星存在长度多态性, 且微卫星的变异与某些蜱虫的生物学特性密切相关。革蜱属近缘种间的线粒体基因组近期可能经历了复制、删除或重组事件。总之, 本研究发掘的线粒体串联重复序列将为蜱螨物种鉴定、种群遗传多样性以及分子系统进化研究提供丰富的分子标记。

关键词 真螨目; 寄螨目; 微卫星; 基因组进化; 线粒体复合体

Distribution and evolutionary characteristics of tandem repeat DNA in the mitochondrial genome of Acari

ZHU Yin^{1,2**} LI Yi^{1,2} JIANG Shi-Die^{1,2} WANG Jin-Jun^{1,2} WEI Dan-Dan^{1,2***}

(1. Key Laboratory of Entomology and Pest Control Engineering, College of Plant Protection, Southwest University, Chongqing 400715, China; 2. Key Laboratory of Agricultural Biosafety and Green Production of Upper Yangtze River, Academy of Agricultural Sciences, Southwest University, Chongqing 400715, China)

Abstract [Aim] This study aimed to investigate the distribution and evolutionary characteristics of mitochondrial DNA (mtDNA) microsatellite (SSR) and long tandem repeats (LTR) within the Acari group. This investigation will provide fundamental data for the development of molecular markers and functional research involving mitochondrial microsatellites. [Methods] The mitochondrial microsatellite (mtSSRs) and long tandem repeats of mites and ticks were mined using microsatellite identification tool (MISA) and tandem repeat finder, respectively. A comprehensive analysis was then carried out on composition, distribution patterns, and evolutionary characteristics of the tandem repeats. [Results] In total, 1 454 perfect mtSSRs and 96 complex mtSSRs were identified in 217 species of Acari, representing 37 families. Among these, there were

*资助项目 Supported projects: 国家自然科学基金 (32272525; 31972276); “十四五” 柑橘产业体系-树体虫害防控

**第一作者 First author, E-mail: zhuyinh@163.com

***通讯作者 Corresponding author, E-mail: weidandand@163.com

收稿日期 Received: 2023-08-11; 接受日期 Accepted: 2023-09-15

271 mononucleotide repeats (SSR1), 127 dinucleotide repeats (SSR2), 517 trinucleotide repeats (SSR3), 323 tetranucleotide repeats (SSR4), 75 pentanucleotide repeats (SSR5), and 46 hexanucleotide repeats (SSR6). The results showed that as the motif length of SSR increased, the number of repeat times decreased. The majority of mtSSRs in Acari were found in the *nad4* and *nad5* protein domains, which are encoded by mitochondrial complex I. Additionally, 167 long tandem repeats (LTR, core repeat base number > 6 bp) were found, with 50.3% of them occurring in the coding region. Notably, the LTR with long core repeat units were found predominantly in the AT control region of specific species. Furthermore, most species in the *Dermacentor* genus exhibited two long repetitive regions in their mitochondrial genomes. **[Conclusion]** This study revealed that SSR3 and SSR4 are the most prevalent tandem repeats in the mitochondrial genome of Acari, and that they mainly appear in the mitochondrial protein-coding region. Additionally, SSR length polymorphism was observed in ticks and mites, and this variation in SSRs is strongly associated with various biological traits in ticks. The mitochondrial genomes of related species of *Dermacentor* may have recently undergone replication, deletion, or recombination events. In conclusion, the mitochondrial tandem repeats identified in this study will provide abundant molecular markers for the identification, population genetic diversity assessment, and molecular phylogenetic and evolutionary analysis of ticks and mites.

Key words Acariformes; Parasitiformes; microsatellite; genome evolution; mitochondrial complex

线粒体(Mitochondrion)是双膜结合细胞器,在氧化磷酸化和 ATP 合成中起着重要作用,有“细胞动力工厂”之称(Kumar *et al.*, 2022)。线粒体基因组编码电子传递链复合体的 13 个蛋白,分别为 *nad1-nad6* 和 *nad4L*(复合体 I)、*cytb*(复合体 III)、*cox1-cox3*(复合体 IV)以及 *atp6* 和 *atp8*(复合体 V)(李红智等, 2011),而在复合体 II 中则不含有线粒体基因组编码的蛋白。事实上,线粒体电子传递链中大部分原本由线粒体基因组编码的基因已经转移至核基因组中,仅有一小部分基因由线粒体基因组所编码(Haig, 2016; Miyazawa *et al.*, 2021)。相较于核基因组,线粒体基因组具有母系遗传、进化速率快、引物高度保守且极少发生重组等特点,因而被广泛应用于物种鉴定、系统发育和基因组进化等方面的研究(Yang *et al.*, 2014; Kern *et al.*, 2020)。

简单重复序列(Simple sequence repeats, SSRs)亦被称为微卫星或短串联重复序列(Short tandem repeats, STRs),由长度为 1-6 bp 的核心重复序列串联而成(Tóth *et al.*, 2000),广泛分布于核基因组和细胞器基因组中(Kapil *et al.*, 2014; Kabra *et al.*, 2016; Kumar *et al.*, 2022)。微卫星因其多态性高、选择中性、共显性遗传和易于检测等特点(石仕福, 2009),被广泛用于种群遗传结构分析(Jin *et al.*, 2019; Wu *et al.*, 2019; Wirgin *et al.*, 2022)、QTL 定位(Sugimoto

et al., 2020)、物种系统发育与进化(Sunnucks *et al.*, 1996; Xiong *et al.*, 2022)、昆虫行为与习性(Hacker *et al.*, 2005; Trontti *et al.*, 2005)以及功能基因定位等研究(Thiel *et al.*, 2003),特别是在基因组进化研究中具有重要作用(Kumar *et al.*, 2020; Xiong *et al.*, 2022)。线粒体基因组包含不同核心碱基长度的重复序列(Habano *et al.*, 1998),根据核心碱基序列的长度,本文将分为微卫星和长串联重复序列(Long tandem repeats, LTRs, 核心碱基数> 6 bp)两大类。

蜱螨(蜱螨亚纲),通常被分为蜱虫(Ticks)和螨虫(Mites),是节肢动物中物种多样性最为丰富的类群之一(袁明龙和王进军, 2012)。蜱螨主要分为真螨目(Acariformes)和寄螨目(Parasitiformes)两个支系。真螨目可分为恙螨亚目(Trombidiformes)、无气门亚目(Astigmata)和甲螨亚目(Oribatida);寄螨目包括中气门亚目(Mesostigmata)、蜱亚目(Ixodida)和巨螨亚目(Holothyrida)(Van Dam *et al.*, 2019)。蜱螨因其不同食性而占据着广泛的生态位,如蜱虫绝大多数吸食动物血液,寄生在人类和各种动物上;螨有植食螨、捕食螨、尘螨和水螨等,从农作物、果树到人类居所都有着它们的分布(陈军, 2006)。目前,线粒体基因组 SSRs 的鉴定与分析研究已在多种动、植物中被报道(Zhao *et al.*, 2011; Nagpure *et al.*, 2015; Heo *et al.*, 2019),

但蜚蠊线粒体基因组微卫星和长串联重复序列的相关研究仍未见报道, 且蜚蠊 SSRs 鉴定相关的研究主要集中在核基因组和转录组中。本文利用公共信息数据库, 系统鉴定蜚蠊线粒体基因组中的 SSRs 和 LTRs, 分析不同类型核心重复序列在蛋白编码区和非编码区的分布规律及进化特征, 并探讨串联重复序列在蜚蠊线粒体基因组进化中的潜在作用, 为进一步筛选多态性线粒体微卫星位点和研究串联重复序列在线粒体基因中的功能提供理论基础和数据参考。

1 材料与方法

1.1 蜚蠊线粒体基因组全序列

截止时间 2022 年 10 月 9 日, 本研究从 NCBI 数据库 (www.ncbi.nlm.nih.gov) 中下载蜚蠊线粒体基因组, 检索时的关键词为“mitochondrion+complete genome+arthropods+mites and ticks”。去冗余后, 共获得线粒体基因组序列 540 条, 包含隶属于真螨目和寄螨目的 37 科 217 种 (表 1), 数据大小为 7.81 Mb。将下载在线粒体基因组全序列保存为 FASTA 格式, 同时下载注释文件和蛋白编码基因序列 (Coding sequences) 备用。

1.2 微卫星序列鉴定

微卫星序列的鉴定采用 MISA (Microsatellite identification tool) 本地搜索软件 (<http://pgrc.ipk-gatersleben.de/misa/misa.html>) (Thiel *et al.*, 2003)。MISA 以 FASTA 格式的核苷酸序列数据作为输入文件, 在运行 Perl 脚本之前, 将 misa.pl 的搜索参数设置为 1-12 2-6 3-4 4-3 5-3 6-3, 即搜索 SSR 的标准: 单碱基核心重复次数 ≥ 12 次, 两碱基核心重复次数 ≥ 6 次, 三碱基核心重复次数 ≥ 4 次, 四碱基、五碱基和六碱基核心重复次数 ≥ 3 次。同时, 将其设置为 true 状态下保存。软件运行后输出的文件分别以 “misa” 和 “statistics” 结尾, 包含了微卫星序列和统计信息。本文主要研究完美型微卫星 (Perfect microsatellite)。

1.3 长串联重复 DNA 序列鉴定

针对核心重复碱基数 >6 bp 的长串联重复序

列, 采用 Tandem repeat finder 本地搜索软件 (Benson, 1999), 软件参数设置为: Match weight=2; Mismatching penalty=7; Indel penalty=7; Match probability=80; Indel probability=10; Minimum alignment score=50; Maximum period size=500; - m (Masked sequence file); - f (Flanking sequence); - d (Data file)。搜索重复序列的标准是核心重复次数 ≥ 2 次, 数据分析时选取核心重复次数 3 次以上的。

1.4 数据分析

用 Excel 统计微卫星序列的碱基类型和重复次数, 并以科为单位计算微卫星平均距离 (公式 1); 将统计的每个物种不同核心碱基类型的重复次数归类到真螨目和寄螨目。此外, 分析串联重复序列在线粒体基因组蛋白质编码区和非编码区的分布, 主要参考线粒体基因组的注释文件, 部分没有注释信息的序列利用线粒体在线注释网站 MITOS (<http://mitos2.bioinf.uni-leipzig.de/index.py>) 进行人工注释。为了比较分析不同大小线粒体基因组之间的 SSR 分布情况, 将 SSRs 总数均一化为每序列 1 kb 上的 SSRs 数量 (相对丰度, Relative abundance, RA) (公式 2) (de Freitas *et al.*, 2022)。长串联重复序列数据分析与 SSR 类似, 但同时统计了其位于编码-非编码区之间的分布数据。多序列比对采用 MEGA11 软件进行分析 (Tamura *et al.*, 2021), 作图使用 GraphPad Prism 8.0.1.244 软件。

$$\text{平均距离 (kb/个)} = \frac{\text{总的基因组序列长度}}{\text{总的微卫星数量}} \quad (1)$$

$$\text{相对丰度 (个/kb)} = \frac{\text{各类型微卫星数量}}{\text{基因组序列长度}} \times 1000 \quad (2)$$

2 结果与分析

2.1 蜚蠊线粒体基因组微卫星的类型及数目

本研究涉及了真螨目和寄螨目 217 个物种, 分属于 37 科。其中, 硬蜱科 Ixodidae 拥有物种最多, 共计 108 种。其次是软蜱科 Argasidae, 有 38 个物种; 叶螨科 Tetranychidae 则包含了 13

个物种,而其他大部分科(占比 56.76%)仅包含了 1 个物种。除地里纤恙螨 *Leptotrombidium deliense*、苍白纤恙螨 *L. pallidum* 和 1 种钝缘蜱 *Ornithodoros tholozani* 外,在 214 个蜱螨物种中共检测到 1 454 个完美型线粒体微卫星(mtSSRs)和 96 个复合型微卫星。在完美型微卫星中,各类型微卫星的数量分别为:单碱基微卫星(SSR1) 271 个、双碱基微卫星(SSR2) 127 个、三碱基微卫星(SSR3) 517 个、四碱基微卫星(SSR4) 323 个、五碱基微卫星(SSR5) 75 个以及六碱基微卫星(SSR6) 46 个。在涵盖的 37 个蜱螨科中,统计结果显示,40.54%的科以 SSR3 为主,27.03%的科以 SSR4 为主,而 16.22%的科以 SSR1 为主。上述结果表明,在科级水平上,蜱螨对于三碱基重复序列具有一定的偏好性。在本研究中,中华甲虫蒲螨 *Pyemotes zhonghuajia* 的线粒体基因组最小,仅为 10 555.0 bp,而西方盲走螨 *Metaseiulus occidentalis* 的线粒体基因组最大,为 24 961.0 bp,蜱螨线粒体基因组的平均长度为 14 615.5 bp。蜱螨 37 个科的微卫星平均距离为 3.89 kb,其中脂螨科 Lardoglyphidae 的平均距离最大,为 14.27 kb,而叶螨科的平均距离最小,为 0.90 kb(表 1)。总体而言,蜱中最为丰富的微卫星类型是三碱基和四碱基重复,而螨中最丰富的则是单碱基和四碱基重复类型。此外,蜱和螨中出现频率最高的

微卫星类型相似,且绝大部分由 A/T 组成,说明蜱螨微卫星和节肢动物线粒体基因组的碱基组成类似,存在 AT 偏好性。

2.2 完美型微卫星在蜱螨线粒体基因组中的分布

2.2.1 真螨目和寄螨目 mtSSRs 重复次数分布

不同类型微卫星的重复次数变化范围差异较大,SSR4-SSR6 的重复次数变化范围明显小于其他类型。真螨目 SSR1 重复次数主要分布在 12-20 次,最大重复次数为 150 次;SSR2 重复次数主要分布在 6-11 次,最大重复次数为 27 次;SSR3 重复次数集中分布在 4-6 次,SSR4-SSR6 的重复次数主要集中在 3 次(图 1: A)。寄螨目 SSR1 重复次数主要分布在 12-21 次,最大重复次数为 35 次;SSR2 重复次数主要分布在 6-10 次,最大重复次数为 16 次;SSR3-SSR6 的重复次数分布均较为集中,其中 SSR3 重复次数主要集中在 4-5 次,最大重复次数为 11 次;SSR4-SSR6 则以 3 次重复为主(图 1: B)。总之,在真螨目和寄螨目中,SSR1 和 SSR2 相较 SSR3-SSR6 而言,其重复次数分布范围更加分散。但与寄螨目相比,真螨目 SSR1 和 SSR2 的重复次数变化范围更广泛。上述结果表明,随着微卫星核心碱基长度的增加,微卫星的重复次数分布范围逐渐减小,其数量也随之减少。

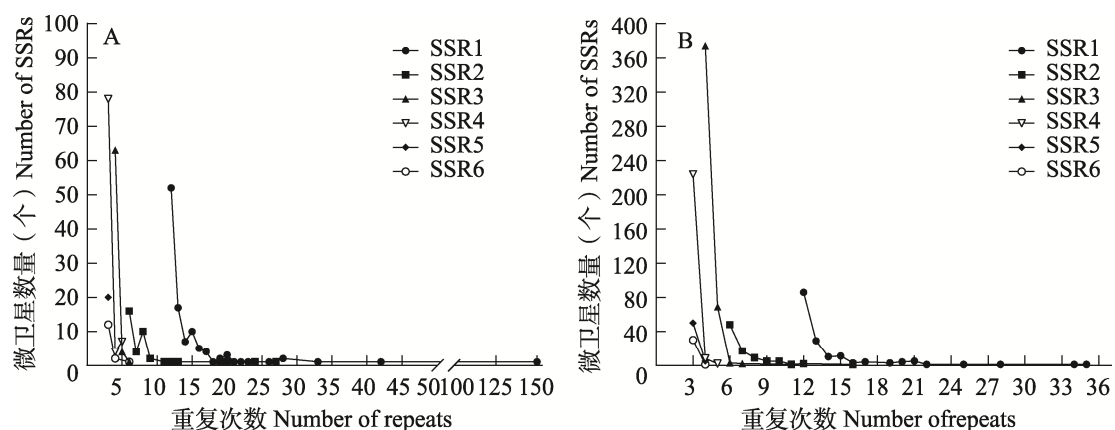


图 1 真螨目 (A) 和寄螨目 (B) 6 种碱基类型的微卫星重复次数分布图

Fig. 1 Distribution of repeated numbers in the perfect SSRs of Acariformes (A) and Parasitiformes (B)

SSR1-SSR6 分别表示单碱基、双碱基、三碱基、四碱基、五碱基以及六碱基重复微卫星。下图同。

SSR1-SSR6 indicate mononucleotide, dinucleotide, trinucleotide, tetranucleotide, pentanucleotide and hexanucleotide microsatellite, respectively. The same below.

表 1 蜚蠊线粒体微卫星的类型和数目分布信息
Table 1 Summary of identified SSRs in mitochondrial genome of Acari

目 Order	亚目 Suborder	科 Family	物种数 Species number	单碱基 Mononuc- leotide	双碱基 Dinuc- leotide	三碱基 Trinuc- leotide	四碱基 Tetranuc- leotide	五碱基 Pentanuc- leotide	六碱基 Hexanuc- leotide	完美型总		平均距离 (kb/个) Mean distance (kb/loci)	
										复合型 Complex	计 Perfect total		
真螨目 Acariformes	无气门亚目 Astigmata	粉螨科 Acaridae	7	1	1	4	6	1	4	4	17	5.80	
		食甜螨科 Glycyphagidae	1	0	1	0	2	0	0	0	3	4.88	
		嗜渣螨科 Chortoglyphidae	1	2	1	5	0	0	0	2	3	10	1.49
		垫螨科 Echimyopodidae	1	1	0	2	1	0	0	1	4	3.78	
		脂螨科 Lardoglyphidae	1	0	0	0	1	0	0	1	1	14.27	
		麦食螨科 Pyroglyphidae	2	0	3	0	3	1	0	0	8	4.07	
		果螨科 Carpoglyphidae	1	0	0	1	1	1	0	1	3	4.92	
		薄口螨科 Histiotomatidae	3	10	1	1	1	3	1	3	2	19	2.34
		特鲁螨科 Trouessartiidae	1	1	1	0	0	0	0	0	3	7.06	
		疥螨科 Sarcoptidae	1	1	2	4	1	2	0	1	10	1.38	
		痒螨科 Psoroptidae	1	0	0	2	1	0	0	1	3	4.75	
		若甲螨科 Oribatulidae	2	1	0	0	4	1	0	0	6	4.95	
		盖卷甲螨科 Steganacaridae	1	1	0	2	1	0	0	0	5	3.45	
		吸螨科 Bdellidae	1	1	0	4	1	2	0	0	7	1.77	
		Trombidiformes	腺水螨科 Lebertiidae	1	0	1	0	0	1	0	1	2	7.17
			叶爪螨科 Penthalpidae	1	0	0	0	2	0	0	0	3	7.35
			蒲螨科 Pyemotidae	1	2	0	2	2	0	0	0	7	1.76
			叶螨科 Tetranychidae	13	75	28	29	49	4	6	32	191	0.90
			刺须螨科 Sperchontidae	1	0	0	1	1	0	0	0	3	7.32
			恙螨科 Trombiculidae	4	0	0	5	1	2	0	0	8	3.57
			羽爪囊螨科 Diptilomiopidae	1	3	0	0	1	0	0	0	4	3.41
瘿螨科 Eriophyidae	3		5	0	2	2	3	0	0	12	3.44		
	蚌螨科 Unionicolidae	2	0	0	2	5	1	0	0	8	3.68		
	蠕形螨科 Demodicidae	2	7	0	2	1	0	0	0	10	2.84		

续表 1 (Table 1 continued)

目 Order	亚目 Suborder	科 Family	物种数 Species number	单碱基 Mononuc- leotide	双碱基 Dinuc- leotide	三碱基 Trinuc- leotide	四碱基 Tetranuc- leotide	五碱基 Pentanuc- leotide	六碱基 Hexanuc- leotide	复合型 Complex	完美型总 计 Perfect total	平均距离 (kb/个) Mean distance (kb/loci)
寄螨目 Parasitiformes	中气门亚目 Mesostigmata	寄螨科 Parasitidae	1	0	0	3	0	0	0	0	3	4.87
		植绥螨科 Phytoseiidae	4	2	3	8	17	4	2	2	36	2.03
		瓦螨科 Varroidae	1	0	0	6	4	0	0	1	10	1.53
		鼻刺螨科 Rhinonyssidae	2	1	3	9	3	0	0	4	16	1.82
		皮刺螨科 Dermanyssidae	1	1	1	7	3	0	0	2	12	1.35
		厉螨科 Laelapidae	2	2	1	3	4	2	0	0	12	2.42
		蟪螨科 Blattisociidae	2	1	2	2	3	0	0	2	8	4.22
		巨螯螨科 Macrochelidae	2	5	1	1	1	1	0	1	9	3.25
		土革螨科 Ologamasidae	1	0	0	1	1	0	0	0	2	7.45
	蜱亚目 Ixodida	硬蜱科 Ixodidae	108	115	50	349	155	42	26	32	733	2.18
		软蜱科 Argasidae	38	33	27	45	39	5	3	5	152	3.51
		纳蜱科 Nuttalliellidae	1	0	0	6	2	0	0	0	8	1.80
	巨螨亚目 Holothyrida	异螨科 Allothyridae	1	0	0	9	2	1	0	0	12	1.21
总计 Total			217	271	127	517	323	75	46	96	1 454	

2.2.2 蜱螨 mtSSRs 在编码区和非编码区的相对丰度分析 基于完美型微卫星在线粒体基因组的定位分析发现, mtSSRs 在蛋白编码区和非蛋白编码区并非均匀分布, 大量微卫星位于线粒体编码区。具体而言, 各亚目中微卫星的相对丰度 (RA) 大小排序如下: 巨螨亚目>恙螨亚目>蜱亚目>中气门亚目>甲螨亚目>无气门亚目。另一方面, 在非编码区, 微卫星的 RA 排序为: 巨螨亚目>恙螨亚目>中气门亚目>无气门亚目>蜱亚目>甲螨亚目。通过对不同类型微卫星的分布统计发现, 除无气门亚目 SSR2、恙螨亚目 SSR4、蜱亚目和巨螨亚目 SSR5 在非编码区的

RA 高于蛋白质编码区, 以及无气门亚目 SSR5 在编码区和非编码区 RA 相同外, 蜱螨 6 个亚目微卫星序列倾向于分布在蛋白质编码区。然而, 不同类型微卫星在蜱螨不同类群的编码区间相对丰度存在差异 (图 2)。例如, SSR1 和 SSR2 在恙螨亚目中的 RA 明显高于其他亚目; SSR3 和 SSR4 在巨螨亚目中分布最多, 且 SSR5 在巨螨亚目非编码区的 RA 明显高于其他亚目编码区和非编码区。SSR5 在甲螨亚目、恙螨亚目和中气门亚目的 RA 相差不大, 但明显高于无气门亚目和蜱亚目。此外, 无气门亚目 SSR6 的 RA 最高。

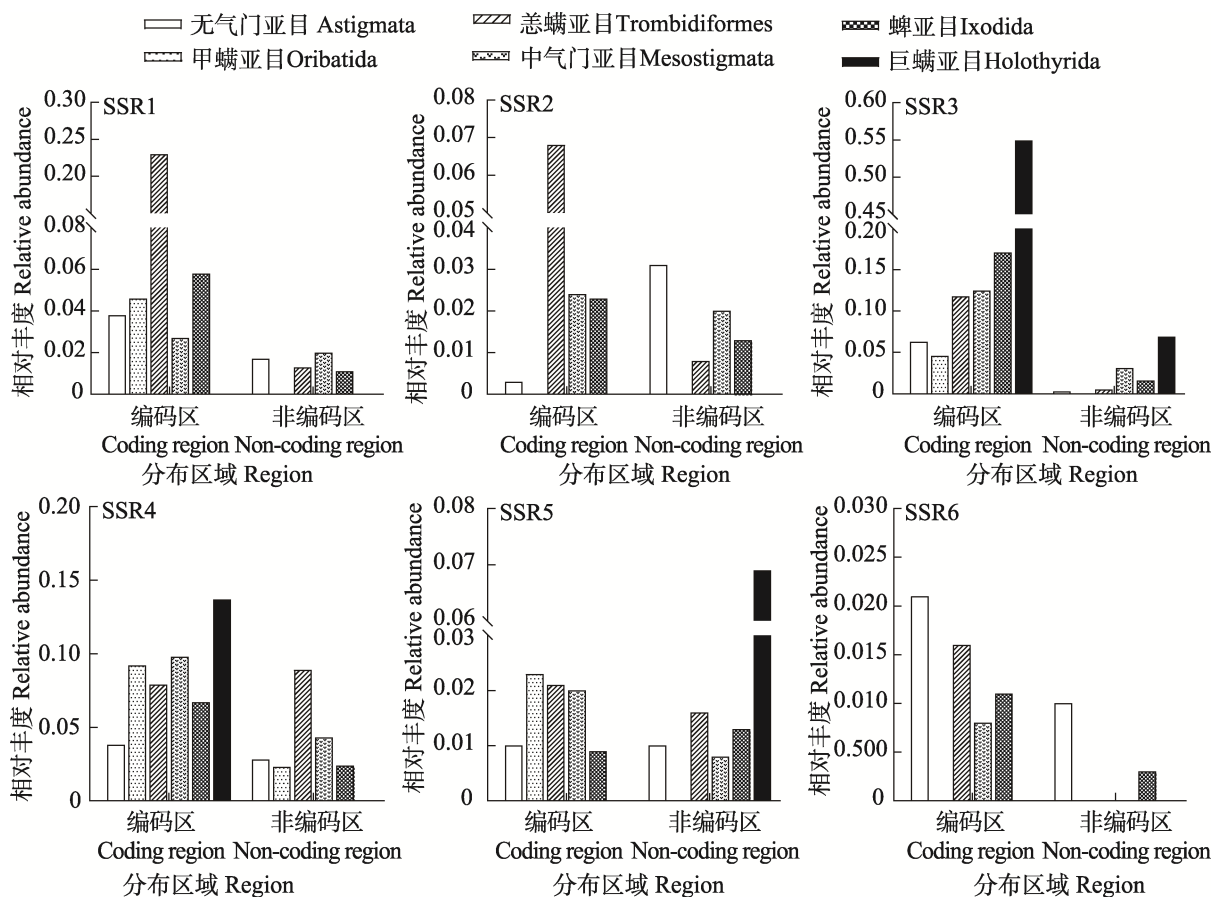


图 2 蜱螨各亚目微卫星在线粒体基因组编码区和非编码区的相对丰度

Fig. 2 The relative abundance of SSR in the protein coding regions and non-coding regions of six suborders of Acari

2.2.3 蜱螨 mtSSRs 在线粒体电子传递链复合体中的分布 通过分析蜱螨 mtSSRs (1 035 个) 在线粒体编码基因中的分布情况发现, 13 个蛋白编码基因均含有微卫星, 其中在 *nad4*

(18.45%) 和 *nad5* (17.20%) 中分布最多。真螨目 260 个 mtSSRs 主要分布在 *nad5* (16.92%)、*nad4* (15%)、*nad1* (13.08%) 和 *nad2* (13.08%) 中, 在 *cox2* (1.54%) 中分布最少; 而寄螨目 775

个 mtSSRs 主要分布在 *nad4* (19.61%)、*nad5* (17.29%) 和 *nad1* (11.74%) 中, 同样在 *cox2* (1.16%) 中分布最少。上述结果说明, 蜱螨 mtSSRs 在 13 个蛋白编码基因中的分布存在一定偏好性, 且都倾向于分布在复合体 I 所编码的基因中 (图 3)。值得一提的是, SSR5 和 SSR6 在真螨目和寄螨目 4 个线粒体复合体中的占比都很少; SSR3 在蜱螨线粒体复合体 III (*cytb* 基因) 中的占比最高, 主要是 (ATT)₄ 和 (ATT)₅。除了上述共性特征外, 真螨目和寄螨目 mtSSRs

还存在一些不同之处: (1) 6 种类型微卫星在真螨目复合体 I 和 IV 中均有分布, 但在寄螨目中, 6 种类型的微卫星仅在复合体 I 中有分布; (2) 寄螨目中 SSR2 仅在复合体 I 中存在, 而在真螨目中, SSR2 不仅在复合体 I 中分布, 还在复合体 IV 中有所分布, 尽管占比较低 (图 3: A); (3) 在寄螨目中, SSR3 在 4 个复合体中的占比都最高, 而真螨目中, SSR3 只在复合体 III 和 V 中占比最多, 而在复合体 I 和 IV 中 SSR1 占比最高 (图 3: B)。

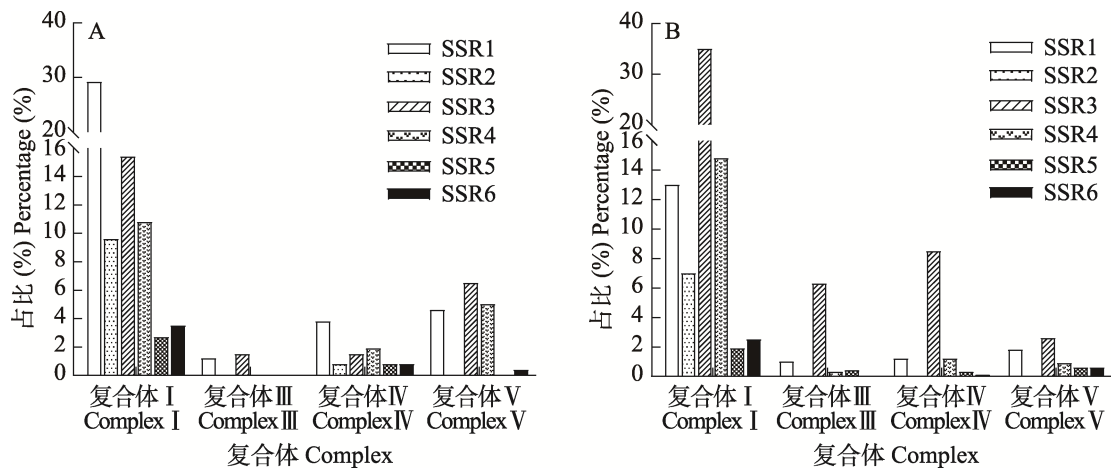


图 3 真螨目 (A) 和寄螨目 (B) 中线粒体微卫星在 4 个复合体中的占比
Fig. 3 Distribution of mitochondrial microsatellites in four respiratory complexes of Acariformes (A) and Parasitiformes (B)

2.3 蜱螨种内不同地理种群 mtSSRs 的序列比较

本研究针对具有不同地理种群线粒体基因组单倍型的序列 (地理种群 > 5), 比较分析 13 个蛋白质编码基因中的微卫星发现, 微小扇头蜱 *Rhipicephalus microplus* 的 *nad2* 基因含有 1 个多态性 SSR 位点 (图 4: A), 即 (TA)₆₋₈。而在长角血蜱 *Haemaphysalis longicornis* 的 *nad4* 基因中, 存在由于碱基突变导致不符合 SSR 搜索参数设置而无法检测到 (TA)₆ 的情况 (图 4: B)。上述情况在白纹革蜱 *Dermacentor albipictus* 中也同样存在, 如在 *nad4L* 基因中, 由于 2 个碱基突变从而导致 (AT)₇ 微卫星位点的出现; 另外, 在 *nad6* 中因插入了 2 个 AT 重复导致 (AT)₇ 微卫星位点的出现 (图 4: C)。

2.4 蜱螨线粒体基因组中长串联重复序列的分析

在 25 科 89 种蜱螨中共鉴定获得 167 个核心碱基数 > 6 bp (重复次数在 3 次以上) 的长串联重复序列, 其中 50.30% 分布在编码区, 且主要位于 *nad5* 中 (35.71%), 37.72% 分布在非编码区, 11.98% 分布于编码区和非编码区之间的间隔区。如在叶螨科 (13 种) 中, 长串联重复序列在编码区的分布频率 (59.38%) 大于非编码区和间隔区 (40.62%)。然而, 在硬蜱科中, 长串联重复序列在非编码区和间隔区的分布频率 (59.09%) 大于编码区 (40.91%)。此外, 长串联重复序列中核心重复基序较长的序列均分布在 AT 控制区内, 如在芒草裂爪蜱 *Stigmaeopsis miscanthi* 线粒

体基因组 AT 控制区分布有核心碱基长度为 273 bp 重复了 8 次的长串联重复序列; 狄斯瓦蜱 *Varroa destructor* 线粒体基因组 AT 控制区含有核心碱基长度为 157 bp 重复了 10 次的长串联重复序列。值得一提的是, 在革蜱属 *Dermacentor* 的物种中, 线粒体基因组拥有 2 个重复区域, 即 Repeat region 1 (RR1) 和 Repeat region 2 (RR2), 且属内不同种间的长串联重复序列相似度较高,

但种内 2 个重复区域之间的相似度较低 (图 5: A, B)。有趣的是, 草原革蜱 *D. nuttalli* 重复区域 2 (RR2) 中的两段长串联重复序列分别与中华革蜱 *D. sinicus* 和森林革蜱 *D. silvarum* 的 RR2 中一段重复序列相同 (图 5: A)。同时, 中华革蜱在 *trnE* 和 *nad1* 之间并不存在 RR1。与蜱不同的是, 3 种蜱的长串联核心重复序列间的相似度较低 (图 5: C, D)。

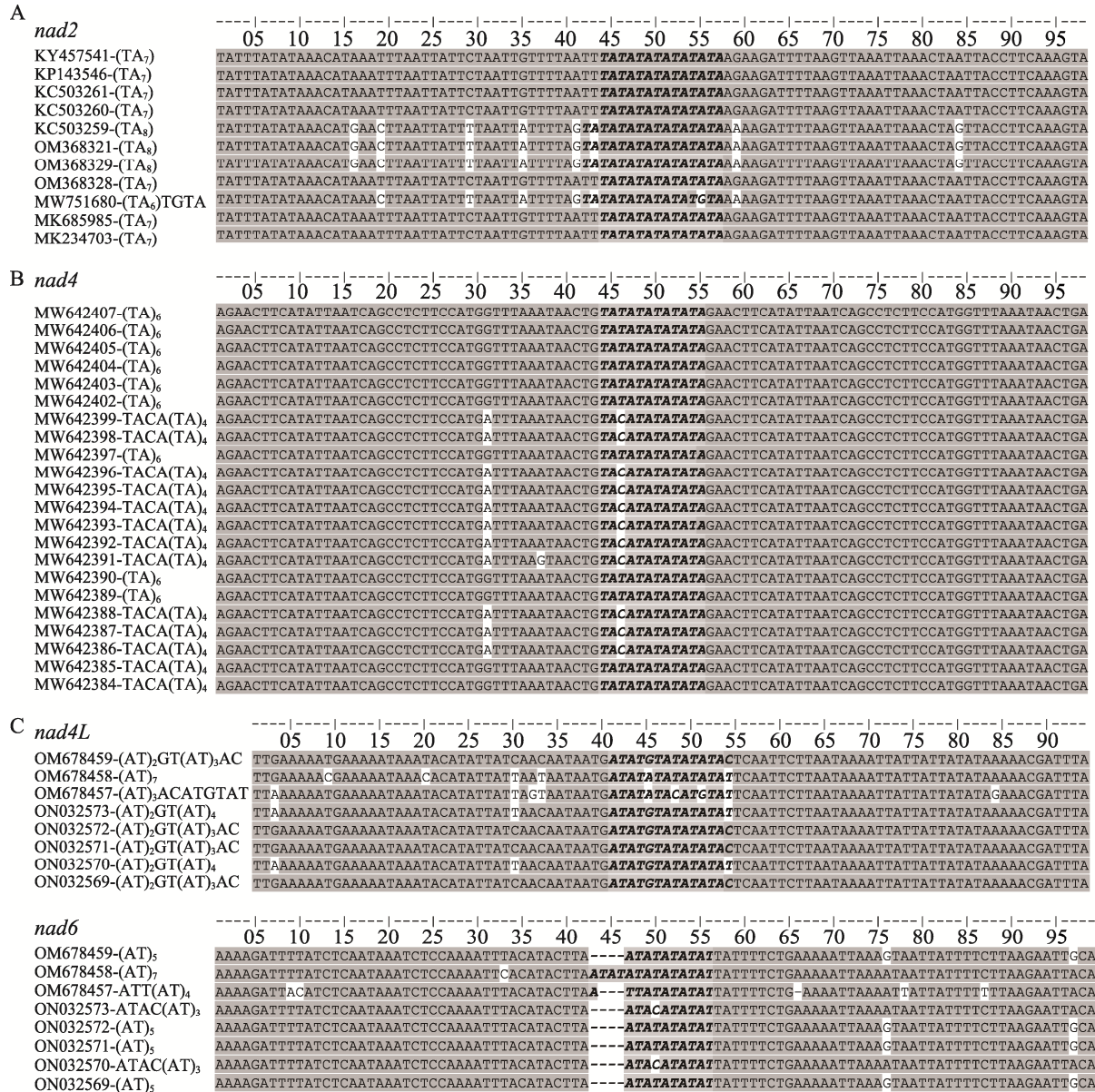


图 4 3 种蜱的不同地理种群线粒体微卫星序列比对

Fig. 4 Sequence alignment of different populations of mtSSRs in three ticks

- A. 微小扇头蜱 *Rhipicephalus microplus*; B. 长角血蜱 *Haemaphysalis longicornis*;
C. 白纹革蜱 *Dermacentor albipictus*. 加粗的碱基为核心重复序列。The coarsed base is the core repeat.

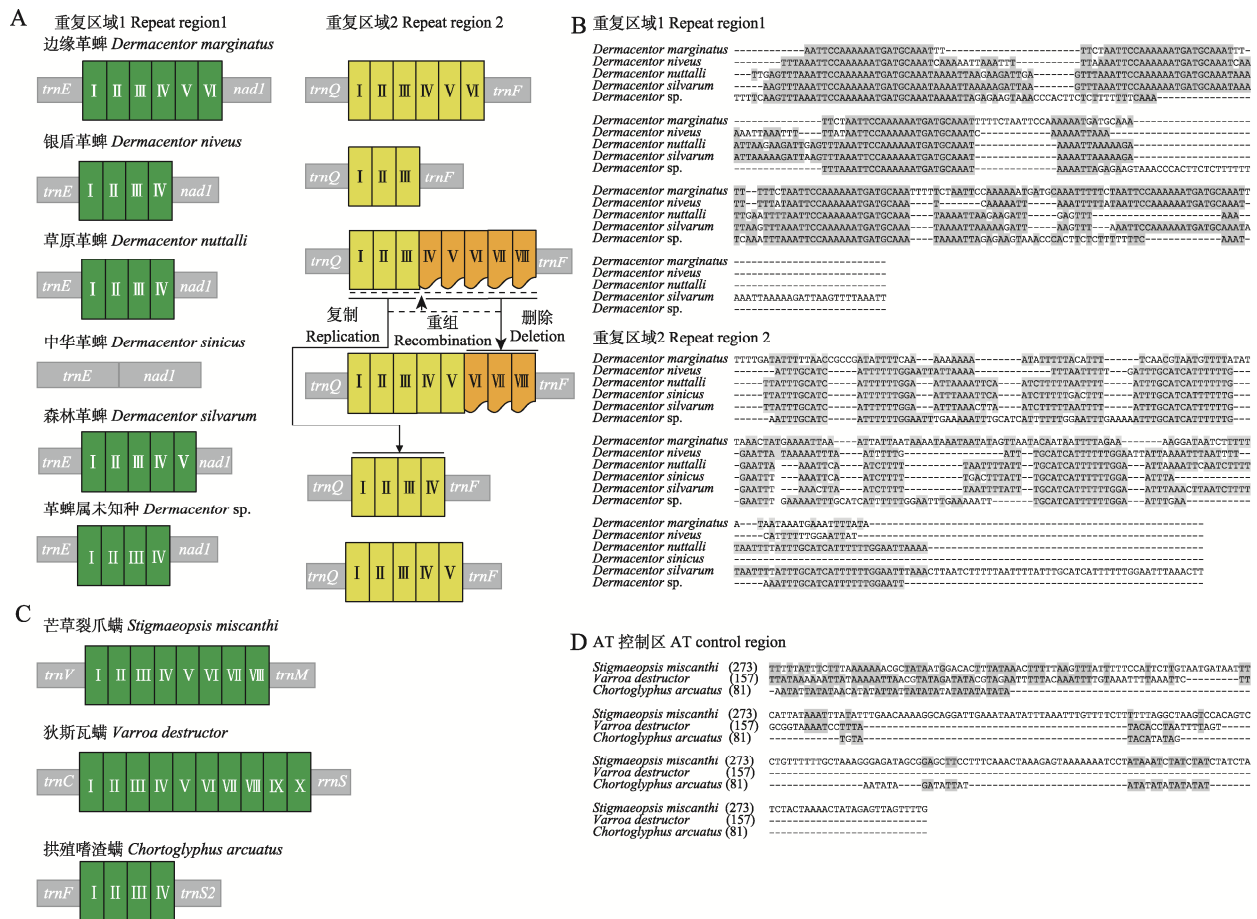


图 5 蜱螨长串联重复序列的结构与核心重复序列比对

Fig. 5 The structural organization and repeat unit alignments of long tandem repeat sequences in Acari

A. 革蜱属的两个重复序列区；B. 6 种革蜱属物种长串联核心重复基序序列比对；C. 3 种螨 AT 控制区内的长串联重复序列；D. 3 种螨的长串联序列核心重复基序序列比对。连线表示核心重复序列相同，橙色的非长方形表示其核心序列是前面黄色长方形核心序列的一部分。

A. The two repeat regions in *Dermacentor* species; B. The alignment of the long tandem repeat units of six *Dermacentor* species; C. The long tandem repeats located in the AT control region of three mites; D. Alignment of the motif of long tandem repeats in three mites. The lines indicate that the motifs are identical between the two species and the orange non-rectangular shapes indicate that the core sequence within the long tandem repeat is a component of the core sequence found in the preceding long tandem repeat represented by the yellow rectangle.

3 讨论

蜱螨不同类型 mtSSRs 在不同科间的分布具有一定的偏倚。如食甜螨科 Glycyphagidae 中仅存在 SSR2 和 SSR4，羽爪瘿螨科 Diptilomiopidae 中仅存在 SSR1 和 SSR4，瓦螨科 Varroidae 仅含有 SSR3 和 SSR4，瘿螨科 Eriophyidae 中则含有 SSR1 和 SSR3-SSR5，而在粉螨科 Acaridae、叶螨科、植绥螨科 Phytoseiidae、硬蜱科和软蜱科中，6 种类型的微卫星均有分布。mtSSRs 在蜱

螨科间水平的分布特征各异，这可能与样本量、物种进化速率以及线粒体基因组大小有关。某些类型的微卫星在特定类群中更丰富的原因可能与该物种线粒体基因组进化过程有关，基因重排、基因获得/丢失事件以及 DNA 复制和修复机制的差异均会导致部分类型微卫星的积累 (Cole *et al.*, 2018)。此外，蜱螨 SSR 分布的核心重复类型在核基因组和细胞器基因组中有所差异。如粉尘螨 *Dermatophagoides farina* 核基因组中 SSR3 数量最多 (Li *et al.*, 2022b)，但在其线粒

体基因组中 SSR2 数量最多。二斑叶螨 *T. urticae* 和肩突硬蜱 *Ixodes scapularis* 也存在类似现象 (汪自立等, 2013), 如二斑叶螨核基因组中 SSR3 占比最高 (72.83%), 其线粒体基因组中 SSR1 数量最多, 其次是 SSR2; 肩突硬蜱则与二斑叶螨正好相反, 即核基因组中 SSR1 数量最多 (73.74%), 其次是 SSR2 (13.73%), 而在其线粒体基因组中则 SSR2 和 SSR3 的数量最为丰富。微卫星在核基因组和细胞器基因组中核心重复类型分布的偏好性, 可能与不同类型基因组间的碱基组成、基因组结构、编码序列占比以及编码基因功能的差异等密切相关。

了解 SSR 在编码区和非编码区的分布特征, 不仅为深入理解蜱螨线粒体基因组的功能和进化提供有价值的信息, 也有助于阐明 SSRs 在基因调控和基因组中的潜在功能。本研究发现, 蜱螨 6 个亚目中大多数 mtSSRs 分布于线粒体蛋白编码区, 这种分布模式的偏好性可能与微卫星参与线粒体基因 RNA 转录后修饰有关 (Zhao *et al.*, 2014)。蜱螨线粒体编码区 mtSSRs 可能承受着较高水平的选择压力, 这与其他动物或植物 mtSSRs 类似。这种选择压力可能导致 mtSSRs 在蛋白编码区域被选择性保留, 而不被随机变异或删除 (Padma Raju *et al.*, 2015)。事实上, 当 SSRs 位于蛋白编码区时, 包括复制滑移和突变在内的变化均可导致基因功能变异, 选择性保留这些位点, 有利于维持线粒体的功能稳定性 (Li *et al.*, 2004)。因此, 蜱螨 mtSSRs 的分布特征和演化与其物种进化之间必然存在一定的关联性。然而, 无气门亚目 SSR2、恙螨亚目 SSR4、蜱亚目和巨螨亚目 SSR5 在非编码区的相对丰度明显高于蛋白质编码区。尽管不同类型微卫星在蛋白编码区和非编码区相对丰度差异的根本原因尚不明确, 但有研究表明, 植物线粒体基因组的结构、基因含量、基因排序、内含子数量以及基因组大小等因素的变化, 会影响 mtSSRs 在编码区和非编码区的分布频率和丰度 (de Freitas *et al.*, 2022)。蜱螨 mtSSRs 在线粒体基因组不同区域的非随机分布可能与某些线粒体基因的功能有关, 并在有效的适应性进化中发挥着重要

作用, 反映了线粒体基因组的进化历程和功能需求差异。

蜱螨 mtSSRs 主要富集在复合体 I 编码的基因中, 尤其是在 *nad5* 和 *nad4* 基因中。这表明复合体 I 中的上述 2 个基因可能受到较强的进化选择压力的影响。在蜱类中, 除 SSR2 仅分布于复合体 I 以及 SSR6 未在复合体 III 中分布外, 所有类型微卫星在 4 个复合体中均有分布, 且主要由 A 和 T 碱基组成。在螨类中, 只有 SSR1、SSR3 和 SSR4 在 4 个复合体中都有分布, 而且常见微卫星类型不仅包括 A 和 T 碱基, 还包括一定比例的 G 和 C 碱基。例如, AGG 类型微卫星在螨类所有 SSR3 中占 8.55%, 而在蜱类线粒体 SSR3 中却没有包含 G 碱基的微卫星。这些差异可能反映了螨类和蜱类的线粒体基因组在进化中的不同特征和适应性。因此, 蜱螨 mtSSRs 在不同复合体中的分布差异可能与不同物种的生物学特征以及进化速率有关。

蜱螨线粒体基因组中包含一定数量的长串联重复序列, 这种现象主要存在于蜱类, 特别是硬蜱科内的物种。有趣的是, 长串联重复序列也多分布于线粒体蛋白编码区, 但那些核心重复基序长度较长的片段更倾向于分布在非编码区, 尤其是 AT 控制区。这与先前的研究结果一致, 如哺乳动物线粒体 DNA 中的串联重复序列除在编码基因内有分布外, 更集中地出现在与线粒体 DNA 复制密切相关的 D-loop 区 (AT 富集区或控制区) (危金普等, 2011)。线粒体基因组 AT 控制区不编码功能基因产物, 这也可能是其可以容忍并累积长串联重复序列的原因之一。值得一提的是, 革蜱属物种线粒体基因组存在两个重复序列区域的现象, 且其中一个物种长串联重复序列与另外两个物种长串联重复序列各有一半相同的现象, 表明近缘种的线粒体基因组在近期可能经历了核心重复序列的复制、删除或重组事件。此外, 在某些情况下, 线粒体重复序列还可能与平行基因转移 (Horizontal gene transfer, HGT) 事件有关 (Sanchez-Puerta *et al.*, 2019)。

线粒体串联重复序列, 已被广泛用于动、植

物鉴定和入侵物种扩散途径预测等方面的研究,如草地贪夜蛾 *Spodoptera frugiperda* 线粒体基因 *nad6* 中存在 2 种微卫星类型,有助于鉴别不同入侵种群并推断其在我国的迁徙路线及分布情况 (Li *et al.*, 2022a)。此外,蜚蠊 mtSSRs 在对于鉴定物种的生殖方式以及复合种同样具有潜在应用价值。如长角血蜚存在孤雌生殖和两性生殖两种生殖方式 (Zhang *et al.*, 2022)。本研究发现,孤雌生殖种群线粒体 *nad4* 基因中包含微卫星 (TA)₆ 位点,但在两性生殖种群中由于碱基突变导致这一位点无法被检测到。白纹革蜚因其具有 2 种色型而导致其分类地位一直存在争议,近期研究将彩色型白纹革蜚划分为 2 个谱系 (谱系 1 和谱系 2) (Reynolds *et al.*, 2022)。本文发现, *nad4L* 和 *nad6* 基因中的 (AT)₇ 微卫星位点仅存在于谱系 2 的个体中,而谱系 1 中未能检测到这两个位点,结果进一步支持了彩色型白纹革蜚可划分为 2 个谱系。上述结果有助于揭示物种之间的亲缘关系和进化历史,也暗示着蜚类不同地理种群微卫星的数量及多态性在一定程度上可能与物种适应性进化相关。总之,蜚蠊线粒体 SSRs 标记在未来蜚蠊种群遗传结构、线粒体基因组进化、物种鉴定等领域的研究中具有广阔的应用前景。

参考文献 (References)

- Benson G, 1999. Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Research*, 27(2): 573–580.
- Cheng J, 2006. An introduction to Acari. *Bulletin of Biology*, 41(7): 17–20. [陈军, 2006. 蜚蠊概述. 生物学通报, 41(7): 17–20.]
- Cole LW, Guo WH, Mower JP, Palmer JD, 2018. High and variable rates of repeat-mediated mitochondrial genome rearrangement in a genus of plants. *Molecular Biology and Evolution*, 35(11): 2773–2785.
- de Freitas KEJ, Busanello C, Viana VE, Pegoraro C, de Carvalho Victoria F, da Maia LC, Costa de Oliveira A, 2022. An empirical analysis of mtSSRs: Could microsatellite distribution patterns explain the evolution of mitogenomes in plants? *Functional & Integrative Genomics*, 22(1): 35–53.
- Habano W, Nakamura S, Sugai T, 1998. Microsatellite instability in the mitochondrial DNA of colorectal carcinomas: Evidence for mismatch repair systems in mitochondrial genome. *Oncogene*, 17(15): 1931–1937.
- Hacker M, Kaib M, Bagine RKN, Epplen JT, Brandl R, 2005. Unrelated queens coexist in colonies of the termite *Macrotermes michaelseni*. *Molecular Ecology*, 14(5): 1527–1532.
- Haig D, 2016. Intracellular evolution of mitochondrial DNA (mtDNA) and the tragedy of the cytoplasmic commons. *BioEssays*, 38(6): 549–555.
- Heo S, Kim C, Chung YS, 2019. High-resolution melting analysis for identification of apple cultivars using simple sequence repeat markers. *Plant Biotechnology Reports*, 13(4): 337–344.
- Jin PY, Tian L, Chen L, Hong XY, 2019. High genetic diversity in a 'recent outbreak' spider mite, *Tetranychus pueraricola*, in mainland China. *Experimental and Applied Acarology*, 78(1): 15–27.
- Kabra R, Kapil A, Attarwala K, Rai PK, Shanker A, 2016. Identification of common, unique and polymorphic microsatellites among 73 cyanobacterial genomes. *World Journal of Microbiology & Biotechnology*, 32(4): 71.
- Kapil A, Rai PK, Shanker A, 2014. ChloroSSRdb: A repository of perfect and imperfect chloroplastic simple sequence repeats (cpSSRs) of green plants. *Database*, 2014: bau107.
- Kern EMA, Kim T, Park JK, 2020. The mitochondrial genome in Nematode phylogenetics. *Frontiers in Ecology and Evolution*, 8: 250.
- Kumar S, Kumari S, Shanker A, 2020. In silico mining of simple sequence repeats in mitochondrial genomes of genus *Orthotrichum*. *Journal of Scientific Research*, 64(2): 179–182.
- Kumar S, Singh A, Shanker A, 2022. pSATdb: A database of mitochondrial common, polymorphic, and unique microsatellites. *Life Science Alliance*, 5(6): e202101307.
- Li C, Song YY, Li LL, Tessnow AE, Zhu JS, Guan XM, Guo WX, Cui HY, Lu ZB, Lv SH, Yu Y, Men XY, 2022a. Two microsatellite types within *NAD6* gene help to distinguish populations and infer the migratory route of the invasive fall armyworm, *Spodoptera frugiperda*, (Lepidoptera, Noctuidae) in China. *Journal of Economic Entomology*, 115(5): 1409–1416.
- Li HZ, Liu DH, Liu JX, Bai YD, 2011. The assembly of the mitochondrial electron transport chain system and its biological implications. *Chinese Bulletin of Life Sciences*, 23(11): 1042–1051. [李红智, 刘丹惠, 吕建新, 白益东, 2011. 线粒体电子传递系统的组装及其生物学意义. 生命科学, 23(11): 1042–1051.]
- Li MN, Fang WX, Wang HH, Chen ZJ, Li XR, Liu YQ, Li Y, Wang D, Huang QY, Lu WY, Han RR, Hong L, Sun ET, 2022b. Genetic diversity and population structure of *Dermatophagoides farinae* based on microsatellite markers. *International Journal of Acarology*, 48(4/5): 287–294.
- Li YC, Korol AB, Fahima T, Nevo E, 2004. Microsatellites within genes: Structure, function, and evolution. *Molecular Biology and*

- Evolution*, 21(6): 991–1007.
- Miyazawa H, Osigus HJ, Rolfes S, Kamm K, Schierwater B, Nakano H, 2021. Mitochondrial genome evolution of Placozoans: Gene rearrangements and repeat expansions. *Genome Biology and Evolution*, 13(1): evaa213.
- Nagpure NS, Rashid I, Pathak AK, Singh M, Singh SP, Sarkar UK, 2015. In silico analysis of SSRs in mitochondrial genomes of fishes. *Mitochondrial DNA*, 26(2): 195–201.
- Padma Raju GV, Rao PS, Rao CS, Sekhar VC, Mudunuri SB, 2015. Microsatellite repeats in mitochondrial genomes: A bioinformatic analysis. Proceedings of the 2015 International Conference on Advanced Research in Computer Science Engineering & Technology (ICARCSET 2015). Unnao, India: 1–5.
- Reynolds S, Hedberg M, Herrin B, Chelladurai JRJJ, 2022. Analysis of the complete mitochondrial genomes of *Dermacentor albipictus* suggests a species complex. *Ticks and Tick-Borne Diseases*, 13(6): 102038.
- Sanchez-Puerta MV, Edera A, Gandini CL, Williams AV, Howell KA, Nevill PG, Small I, 2019. Genome-scale transfer of mitochondrial DNA from legume hosts to the holoparasite *Lophophytum mirabile* (Balanophoraceae). *Molecular Phylogenetics and Evolution*, 132: 243–250.
- Shi SF, 2009. Genetic diversity, population structure and gene flow of *Helicoverpa armigera* (Hübner) in China. Msater dissertation. Nanjing: Nanjing Agricultural University. [石仕福, 2009. 棉铃虫种群的遗传多样性与基因流研究. 硕士学位论文论文. 南京: 南京农业大学.]
- Sugimoto N, Takahashi A, Ihara R, Itoh Y, Jouraku A, Van Leeuwen T, Osakabe M, 2020. QTL mapping using microsatellite linkage reveals target-site mutations associated with high levels of resistance against three mitochondrial complex II inhibitors in *Tetranychus urticae*. *Insect Biochemistry and Molecular Biology*, 123: 103410.
- Sunnucks P, England PR, Taylor AC, Hales DF, 1996. Microsatellite and chromosome evolution of parthenogenetic sitobion aphids in Australia. *Genetics*, 144(2): 747–756.
- Tamura K, Stecher G, Kumar S, 2021. MEGA11: Molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution*, 38(7): 3022–3027.
- Thiel T, Michalek W, Varshney RK, Graner A, 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics*, 106(3): 411–422.
- Tóth G, Gáspári Z, Jurka J, 2000. Microsatellites in different eukaryotic genomes: Survey and analysis. *Genome Research*, 10(7): 967–981.
- Trontti K, Aron S, Sundström L, 2005. Inbreeding and kinship in the ant *Plagiolepis pygmaea*. *Molecular Ecology*, 14(7): 2007–2015.
- Van Dam MH, Trautwein M, Spicer GS, Esposito L, 2019. Advancing mite phylogenomics: Designing ultraconserved elements for Acari phylogeny. *Molecular Ecology Resources*, 19(2): 465–475.
- Wang ZL, Huang J, Du LM, Li WJ, Yue BS, Zhang XY, 2013. Comparison of microsatellites between the genomes of *Tetranychus urticae* and *Ixodes scapularis*. *Sichuan Journal of Zoology*, 32(4): 481–486. [汪自立, 黄杰, 杜联明, 李午俊, 岳碧松, 张修月, 2013. 二斑叶螨和肩突硬蜱基因组微卫星分布规律研究. 四川动物, 32(4): 481–486.]
- Wei JP, Pan XF, Li HQ, Duan F, 2011. Distribution and evolution of simple repeats in the mtDNA D-loop in mammalian. *Hereditas (Beijing)*, 33(1): 67–74. [危金普, 潘学峰, 李红权, 段斐, 2011. 简单重复 DNA 序列在哺乳动物 mtDNA D-loop 区的分布及进化特征. 遗传(北京), 33(1): 67–74.]
- Wirgin I, Maceda L, Stabile J, Waldman J, 2022. Genetic population structure of summer flounder *Paralichthys dentatus* using microsatellite DNA analysis. *Fisheries Research*, 250: 106270.
- Wu HH, Wan P, Huang MS, Lei CL, 2019. Microsatellites reveal strong genetic structure in the common cutworm, *Spodoptera litura*. *Journal of Integrative Agriculture*, 18(3): 636–643.
- Xiong YL, Yu QQ, Xiong Y, Zhao JM, Lei X, Liu L, Liu W, Peng Y, Zhang JB, Li DX, Bai SQ, Ma X, 2022. The complete mitogenome of *Elymus sibiricus* and insights into its evolutionary pattern based on simple repeat sequences of seed plant mitogenomes. *Frontiers in Plant Science*, 12: 802321.
- Yang L, Tan ZQ, Wang DR, Xue L, Guan MX, Huang TS, Li RH, 2014. Species identification through mitochondrial rRNA genetic analysis. *Scientific Reports*, 4: 4089.
- Yuan ML, Wang JJ, 2012. Progress in the complete mitochondrial genomes of the Acari. *Acta Entomologica Sinica*, 55(4): 472–481. [袁明龙, 王进军, 2012. 蜱螨线粒体基因组研究进展. 昆虫学报, 55(4): 472–481.]
- Zhang X, Zhao CY, Cheng CY, Zhang GG, Yu T, Lawrence K, Li HY, Sun JM, Yang ZY, Ye L, Chu HL, Wang Y, Han XH, Jia YC, Fan SZ, Kanuka H, Tanaka T, Jenkins C, Gedye K, Chandra S, Price DC, Liu QY, Choi YK, Zhan XJ, Zhang ZB, Zheng AH, 2022. Rapid spread of severe fever with thrombocytopenia syndrome virus by parthenogenetic Asian longhorned ticks. *Emerging Infectious Diseases*, 28(2): 363–372.
- Zhao L, Zheng ZM, Huang Y, Zhou ZJ, Wang L, 2011. Comparative analysis of the mitochondrial control region in Orthoptera. *Zoological Studies*, 50(3): 385–393.
- Zhao ZX, Guo C, Sutharzan S, Li P, Echt CS, Zhang J, Liang C, 2014. Genome-wide analysis of tandem repeats in plants and green algae. *G3-Genes Genomes Genetics*, 4(1): 67–78.