DOI: 10. 11686/cyxb2026028 http://cyxb. magtech. com. cn
田甜, 王灵洁, 王志军, 等. 小黑麦叶绿体基因组生物信息学解析及分子标记位点挖掘. 草业学报, 2026, 35(12): 183－196.
TIAN Tian, WANG Ling-jie, WANG Zhi-jun, *et al*. Bioinformatic analysis of the triticale chloroplast genome and identification of molecular marker loci. Acta Prataculturae Sinica, 2026, 35(12): 183－196.

小黑麦叶绿体基因组生物信息学解析及分子标记位点挖掘

田甜[1,2,3,4], 王灵洁[1,2,3,4], 王志军[1,2,3,4], 张银霞[1,2,3,4], 杜文华[1,2,3,4*]
(1. 甘肃农业大学草业学院, 甘肃 兰州 730070; 2. 草业生态系统教育部重点实验室, 甘肃 兰州 730070; 3. 农业农村部饲草种质创新与新品种选育重点实验室, 甘肃 兰州 730070; 4. 中-美草地畜牧业可持续发展研究中心, 甘肃 兰州 730070)

摘要: 小黑麦(×*Triticosecale*)是小麦属(*Triticum*)与黑麦属(*Secale*)人工杂交形成的多倍体饲草作物, 兼具高产与抗逆优势。叶绿体基因组相对保守, 是研究杂交作物的理想工具, 以NCBI公共数据库中的六倍体小黑麦(登录号: ON422331. 1)为研究对象, 通过生物信息学方法系统解析了其叶绿体基因组结构、反向重复序列(IR)边界动态变化及与近缘属的比较基因组学特征。结果显示, 该基因组全长135899 bp, 呈典型四分体结构, 共注释了112个独特基因(其中19个基因含内含子)。共鉴定了114个叶绿体简单序列重复(SSR)和47个长重复序列, 发现IR区存在849 bp特异性扩张, 伴随大单拷贝区344 bp相应收缩。比较基因组分析表明编码区高度保守, 非编码区呈序列分化; 系统发育分析显示, 小黑麦以100%bootstrap支持率聚类于小麦属分支, 与硬粒小麦亲缘关系最近。研究解析了小黑麦叶绿体基因组, 为后续细胞器基因功能及分子标记开发提供了基础。

关键词: 小黑麦; 杂交种; 叶绿体基因组; 序列特征; 基因组比较

Bioinformatic analysis of the triticale chloroplast genome and identification of molecular marker loci

TIAN Tian[1,2,3,4], WANG Ling-jie[1,2,3,4], WANG Zhi-jun[1,2,3,4], ZHANG Yin-xia[1,2,3,4], DU Wen-hua[1,2,3,4*]
1. *College of Grassland Science, Gansu Agricultural University, Lanzhou* 730070, *China*; 2. *Key Laboratory of Grassland Ecosystem, Ministry of Education, Lanzhou* 730070, *China*; 3. *Key Laboratory of Forage Germplasm Innovation and New Variety Breeding, Ministry of Agriculture and Rural Affairs, Lanzhou* 730070, *China*; 4. *Sino-U. S. Centers for Grazingland Ecosystem Sustainability, Lanzhou* 730070, *China*

Abstract: Triticale (×*Triticosecale*) is a polyploid forage crop derived from artificial hybridization between *Triticum* and *Secale*, combining desirable agronomic traits such as high productivity and enhanced stress tolerance. The chloroplast genome, being relatively conserved, is an ideal tool for investigating cytoplasmic inheritance in hybrid crops. In this study, the complete chloroplast genome of hexaploid triticale (Accession number: ON422331. 1) was analyzed using bioinformatic methods based on data retrieved from the NCBI. The results show that the chloroplast genome is 135899 bp in length with a typical quadripartite organization. A total of 112 distinct genes were annotated, including 19 with introns. Furthermore, 114 chloroplast simple sequence repeats (SSRs) and 47 long repeats were identified. Comparative boundary analysis revealed an 849 bp expansion in the inverted repeat region and a 344 bp

收稿日期: 2026-01-20; 改回日期: 2026-03-30
基金项目: 国家自然科学基金(32260339)和甘肃省2025年教育科技创新项目(2025CXZX-847)资助。
作者简介: 田甜(1997－), 女, 贵州铜仁人, 在读博士。E-mail: ygsn2705@126. com
*通信作者 Corresponding author. E-mail: duwh@gsau. edu. cn

contraction in the large single copy region. Comparative genomic analyses revealed high conservation in coding regions but divergence in non-coding regions. Phylogenetic reconstruction clustered triticale within the *Triticum* lineage with 100% bootstrap support, showing the closest relationship with *T. turgidum*. The results of this study advance our understanding of triticale chloroplast genome architecture and evolution, and provide a foundation for further research on chloroplast gene function and molecular marker development.

Key words: triticale; hybrid; chloroplast genome; sequence characteristics; genome comparison

小黑麦(×*Triticosecale*,triticale)是一种人工创制的饲草作物,由小麦属(*Triticum*)与黑麦属(*Secale*)杂交后经染色体加倍获得[1-2]。小黑麦遗传小麦的高产和黑麦的强抗逆性以及高赖氨酸含量的特性,表现出显著的杂种优势,已在我国西北地区农牧交错带成为仅次于苜蓿(*Medicago sativa*)和玉米(*Zea mays*)的优质饲草作物[3-4]。目前,小黑麦主要包含两种倍性类型(图1)[5-6]:六倍体(AABBRR)和八倍体(AABBDDRR)[7-8]。六倍体小黑麦($2n=6x=42$, AABBRR)可通过两条途径合成[9]:1)四倍体小麦($2n=4x=28$, AABB)与黑麦($2n=2x=14$,RR)杂交,随后对不育的F_1代杂种($2n=21$, ABR)进行染色体加倍,获得含有21对染色体的新双二倍体,其中7对来源于黑麦;2)将六倍体小麦($2n=6x=42$, AABBDD)与黑麦(RR)杂交,初始杂交种具有完整的AABBDDRR染色体组,但经多代选择,D基因组逐渐被淘汰,而后为六倍体小黑麦。八倍体小黑麦(AABBDDRR)由六倍体小麦(AABBDD)与黑麦的F_1代杂种经染色体加倍后形成。小黑麦因其高产和高含量粗蛋白、粗脂肪及淀粉而多作为饲草使用。由于其具有良好的抗寒性,因此可将其与季节性农作物进行错季种植,即小黑麦在秋冬季节进行播种和生长,次年夏季收割,收割后种植适季作物。既保障适季作物的正常生长,又充分利用了冬春闲田,同时减少了冬春季节土地裸露带来的风沙等问题,兼具经济与生态效应。近年来,随着我国对盐碱地的利用和改良问题的深入研究,小黑麦的耐盐碱性也被挖掘。研究表明,小黑麦具有良好的耐盐碱能力,可在盐度为4‰~12‰的盐碱土壤中生长[10],因此,小黑麦具有良好的耐盐碱性挖掘的潜力,可培育出高产抗逆的新品种。然而,这一目标受限于小黑麦复杂的遗传背景、多倍体特性以及传统育种的漫长周期。

随着叶绿体基因工程技术的成熟,利用叶绿体基因组(chloroplast genome, cp genome)作为生物反应器已成为解决上述问题的新途径。与核基因组转化相比,叶绿体基因工程具有外源基因表达量高、无基因沉默现象、母系遗传避免基因漂移等独特优势[11]。然而,叶绿体基因工程技术的实现需依赖目标物种叶绿体基因组完整信息。目前,小黑麦的遗传改良研究多集中于核基因组层面,对其叶绿体基因组资源的研究较少。尽管已有研究证实了小黑麦的母系遗传特性[12],但其细胞质基因组的精确结构、变异特征及系统发育起源研究仍然较少。叶绿体基因组严格的母系遗传模式使其成为追溯作物驯化历史、鉴定细胞质来源的理想工具[13]。此外,叶绿体简单重复序列(chloroplast simple sequence repeats, cpSSR)标记因具有多态性高、共显性遗传、操作简便等特点,在作物细胞质遗传多样性评价、分子身份证构建及种子纯度鉴定中展现出巨大潜力[14]。对于小黑麦这一人工合成作物而言,解析其叶绿体基因组特征,对于区分不同细胞质来源的育种系、保护创制品种的知识产权具有迫切的实践需求。

本研究以六倍体小黑麦(NCBI登录号:ON422331)[14]为对象,利用生物信息学方法分析其叶绿体基因组特征、重复序列、反向重复序列(inverted repeat, IR)边界及进行基因组比较,旨在填补小黑麦细胞质基因组信息的空白。研究系统鉴定了基因组中的重复序列与SSR位点分布特征,为理解多倍体化过程中的叶绿体基因组变异规律提供了理论参考,为小黑麦叶绿体基因工程改造及后续分子标记开发奠定了数据基础。

1 材料与方法

1.1 叶绿体基因组获取

小黑麦叶绿体基因组在National Center Biotechnplogy Information(NCBI)数据库下载,登录号为ON422331(https://www.ncbi.nlm.nih.gov/nuccore/ON422331.1/)[15]。使用OGDRAW(https://chlorobox.mpimp-golm.mpg.de/OGDraw.html)[16]构建小黑麦叶绿体基因组图谱。

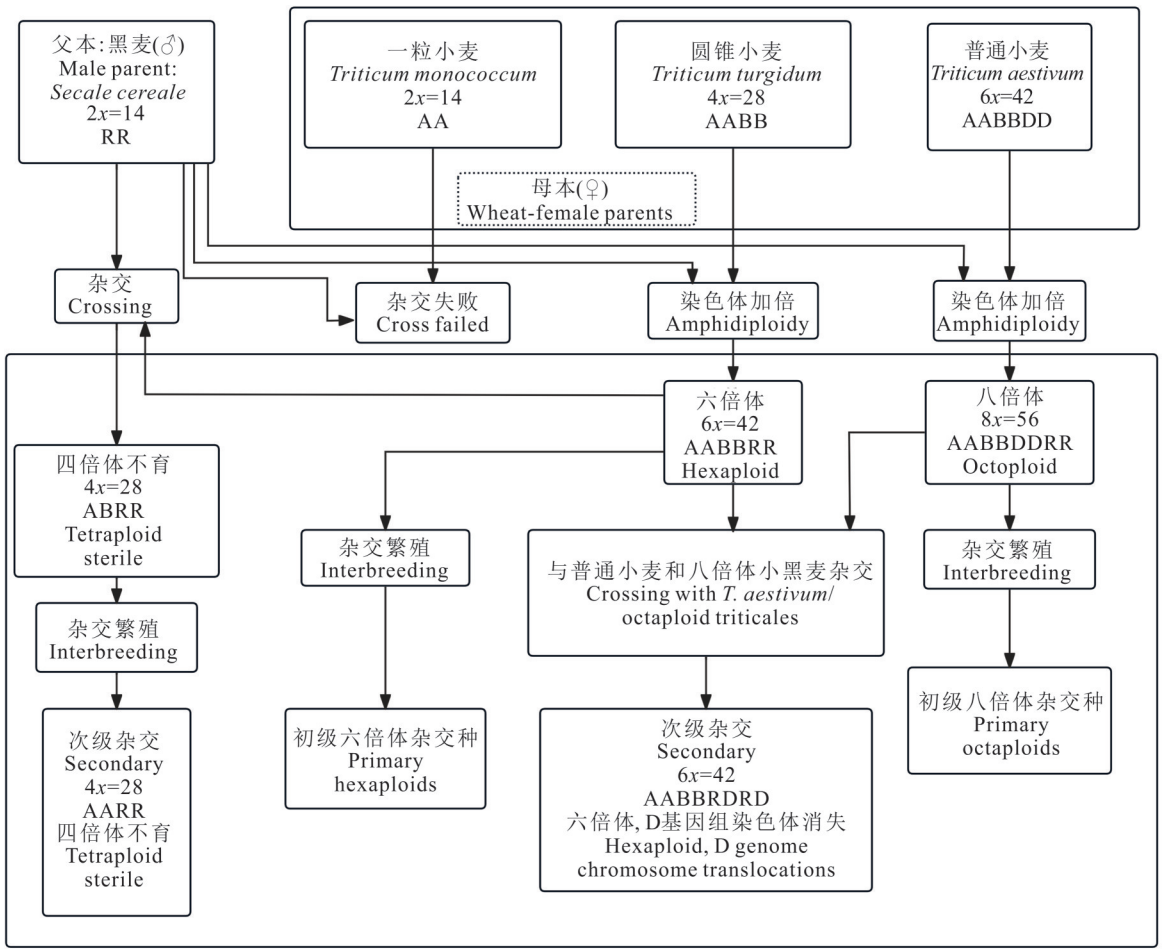


图 1　小黑麦起源概要(译自 Simmonds[6])

Fig. 1　Summary chart of the origins of different types of triticale (Source: Simmonds[6])

1.2　重复序列分析

利用在线工具 GeSeq(采用 REPuter 程序)[17](https://chlorobox.mpimp-golm.mpg.de/geseq.html)分析小黑麦叶绿体基因组的重复序列,筛选阈值设置为 Hamming distance≤3 且 E-value≤1×10^{-5},鉴定所有正向重复(direct repeat, D)和回文重复(palindromic repeat, P),随后利用 MISA 在线软件(https://webblast.ipk-gatersleben.de/misa/index.php? action=1)对叶绿体基因组中的简单重复序列(simple sequence repeats, SSRs)进行分析,分析参数设置如下:单核苷酸重复次数不少于 10 次,二核苷酸重复次数不少于 5 次,三核苷酸重复次数不少于 4 次,四、五、六核苷酸重复次数不少于 3 次。此外,两个 SSR 位点之间的最小间距设定为 100 bp[18]。

1.3　IR 边界分析及基因组比较

为比较小黑麦叶绿体基因组变异,从 NCBI 数据库下载小麦属和黑麦属物种叶绿体全基因组数据,具体包括:普通小麦(登录号:NC_002762.1)、硬粒小麦(*Triticum turgidum*,登录号:NC_024814.1)、马卡小麦(*Triticum macha*,登录号:NC_025955.1)、乌拉尔图小麦 PI428335 品种(*Triticum urartu* cv. PI428335,登录号:KJ614411.1)、脆轴黑麦(*Secale segetale*,登录号:LC645358.1)、野黑麦 6047 品种(*Secale sylvestre* 6047,登录号:MW557517.1)、瓦维洛夫黑麦(*Secale vavilovii*,登录号:OQ700975.1)和非洲黑麦(*Secale africanum*,登录号:OQ700974.1)。在基因组进化过程中,反向重复序列(inverted repeat,IR)边界易发生扩张或收缩,这可能导致特定基因在 IR 区域与单拷贝区域之间的转移[19],利用 mVISTA(https://genome.lbl.gov/vista/mvista/submit.shtml)在线软件,以注释的小黑麦叶绿体基因组为参照,绘制下载物种的序列差异图,并进行人工调整。使用 IRscope 软件[20]对下载物种的叶绿体基因组进行边界分析。

1.4 系统发育分析

从NCBI数据库下载16种禾本科(Poaceae)植物的叶绿体基因组序列,具体包括:芒状山羊草(*Aegilops comosa*,登录号:NC_046697.1)、膝曲山羊草(*Aegilops geniculata*,登录号:NC_023097.1)、科奇山羊草TA1980品种(*Aegilops kotschyi* cv. TA1980,登录号:NC_024832.1)、硬粒小麦(登录号:NC_024814.1)、长穗山羊草TA1924品种(*Aegilops longissima* cv. TA1924,登录号:NC_024830.1)、一粒小麦(登录号:NC_021760.1)、西尔斯山羊草TA1837品种(*Aegilops searsii* cv. TA1837,登录号:NC_024815.1)、粗山羊草(*Aegilops tauschii*,登录号:NC_022133.1)、脆轴黑麦(登录号:LC645358.1)、栽培黑麦(*Secale cereale*,登录号:NC_021761.1)、山地黑麦(*Secale strictum*,登录号:KY636137.1)、野黑麦6047品种(登录号:MW557517.1)、普通小麦(登录号:NC_002762.1)、马卡小麦(登录号:NC_025955.1)、乌拉尔图小麦PI428335品种(登录号:KJ614411.1)、小黑麦(×*Triticosecale*,登录号:ON422331.1)。以玉米(登录号:NC_001666.2)为外类群,因其隶属于黍亚科(Panicoideae),与内类群物种所在的早熟禾亚科(Pooideae)存在亚科间分化,既可保证根节点可靠定位所需的充分序列分歧度,又可在禾本科范围内维持系统发育的关联性。系统发育分析基于完整叶绿体基因组序列,利用MAFFT软件[21]进行多序列比对并辅以人工校正,采用MEGA 7的最大似然法(maximum likelihood),选用General Time Reversible(GTR)[22]模型,设置Bootstrap检验1000次重复。

1.5 数据分析

利用SPSS 25.0和Excel进行序列统计分析和绘制图,利用mVISTA、IRscope及MEGA 7软件绘制小黑麦叶绿体基因组IR边界和系统发育树。

2 结果与分析

2.1 小黑麦叶绿体基因组特征分析

小黑麦叶绿体基因组全长135899 bp,呈典型的四分体架构,其中大单拷贝区(large single copy region, LSC)长为80004 bp,小单拷贝区(small single copy region, SSC)长为12791 bp,包含一对反向重复区,一个区长为21552 bp(表1和图2)。小黑麦叶绿体基因组的GC含量为38.32%,IR区的GC含量(43.91%)高于LSC区(36.28%)和SSC区(32.18%)(表1)。另外,在小黑麦叶绿体基因组中,第一、第二和第三密码子位点的GC含量相差不大,分别为38.26%、38.54%和38.15%(表1)。

表1 小黑麦叶绿体基因组中密码子的碱基组成特征

Table 1 Base composition characteristics of codons in chloroplast genome of triticale

项目 Item	碱基占比 Percentage of bases (%)				长度 Length (bp)
	GC	GC1	GC2	GC3	
全长 Full length	38.32	38.26	38.54	38.15	135899
大单拷贝区 Large single copy (LSC)	36.28	35.85	36.37	36.63	80004
小单拷贝区 Small single copy (SSC)	32.18	33.19	32.89	30.47	12791
重复拷贝区域A Inverted repeat A (IRA)	43.91	44.06	44.43	43.25	21552
重复拷贝区域B Inverted repeat B (IRB)	43.91	43.25	44.43	44.06	21552

注:GC1、GC2和GC3分别指基因编码区中所有密码子第1位、第2位、第3位碱基的G和C所占比例。

Note: GC1, GC2, and GC3 refer to the proportions of G and C at the first, second, and third positions of all codons in a coding region, respectively.

2.2 小黑麦叶绿体基因组基因组成分析

小黑麦叶绿体基因组共有133个基因,其中21个基因在IR区被重复(表2)。LSC区包含81个基因(包括59个蛋白编码基因和22个tRNA基因),SSC区包含10个基因(均为蛋白编码基因),各IR区均含21个基因(包括9个蛋白编码基因、8个tRNA基因及4个rRNA基因)。这些基因功能涉及光合作用、自我复制、其他基因及未知功能基因(表2)。

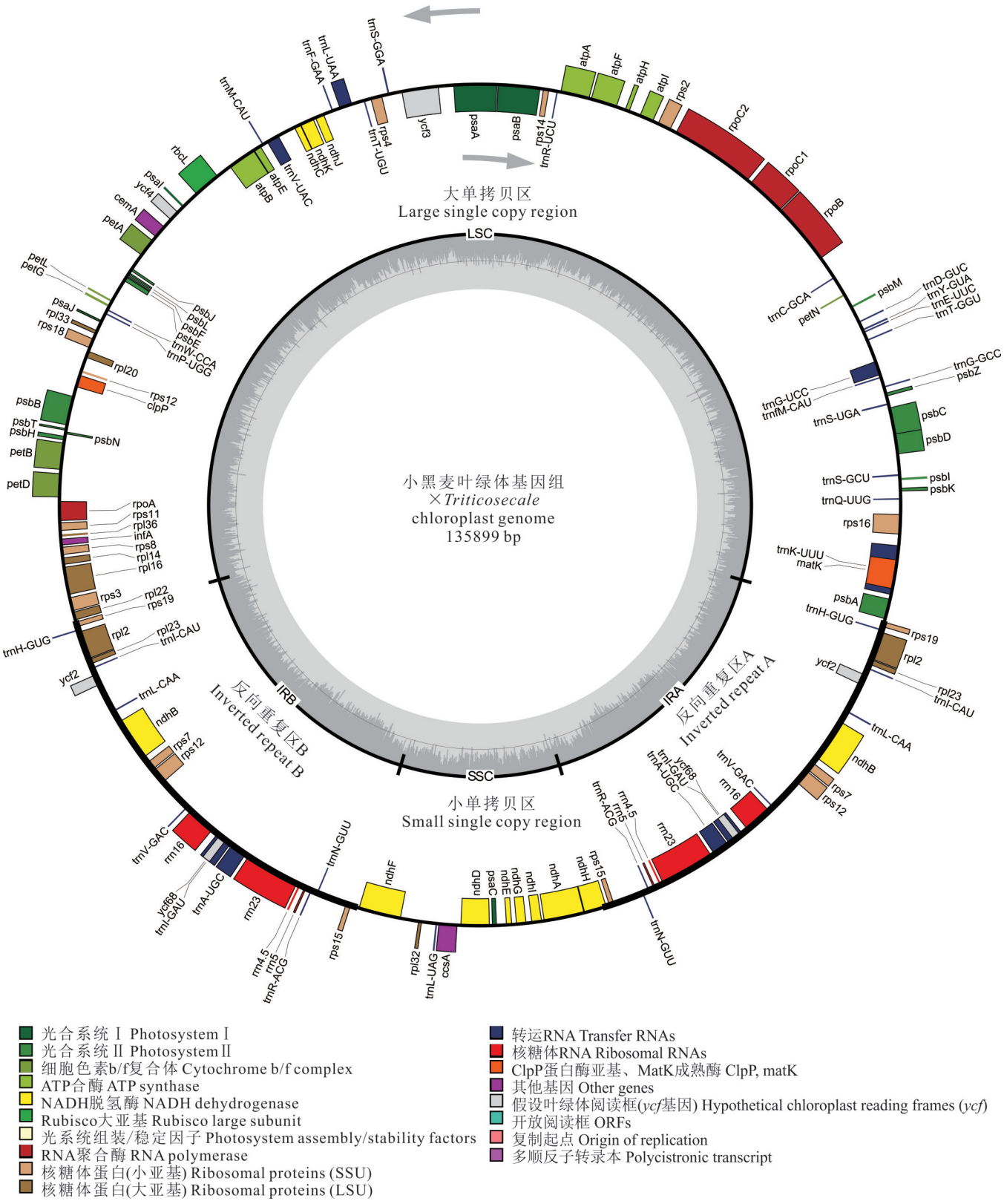


图 2　小黑麦叶绿体基因组

Fig. 2　The complete chloroplast genome of triticale

2.3　小黑麦叶绿体基因组基因的内含子结构分析

叶绿体基因组内含子长度为 581～2484 bp，最小为 *trnV-UAC*，最大为 *trnK-UUU*。在此长度范围内，共计 19 个基因含有内含子（表 3），其中仅 *ycf3* 基因含有 2 个内含子，其余 18 个基因只含有 1 个内含子。

表 2 小黑麦叶绿体基因功能分类

Table 2 Functional classification of chloroplast genes in triticale

功能 Function	基因类别 Gene group	基因名称 Gene name
光合作用 Photosynthesis	光系统Ⅰ的亚基 Subunits of photosystem Ⅰ	*psaA*, *psaB*, *psaC*, *psaI*, *psaJ*
	光系统Ⅱ的亚基 Subunits of photosystem Ⅱ	*psbA*, *psbB*, *psbC*, *psbD*, *psbE*, *psbF*, *psbH*, *psbI*, *psbJ*, *psbK*, *psbL*, *psbM*, *psbN*, *psbT*, *psbZ*
	NADH脱氢酶的亚基 Subunits of NADH dehydrogenase	*ndhA**, *ndhB**(2), *ndhC*, *ndhD*, *ndhE*, *ndhF*, *ndhG*, *ndhH*, *ndhI*, *ndhJ*, *ndhK*
	细胞色素b/f复合物的亚基 Subunits of cytochrome b/f complex	*petA*, *petB**, *petD**, *petG*, *petL*, *petN*
	ATP合酶的亚基 Subunits of ATP synthase	*atpA*, *atpB*, *atpE*, *atpF**, *atpH*, *atpI*
	Rubisco的大亚基 Large subunit of Rubisco	*rbcL*
	光叶绿素酸酯还原酶的亚基 Subunits of photochlorophyllide reductase	—
自我复制 Self-replication	大核糖体亚基的蛋白质 Proteins of large ribosomal subunit	*rpl14*, *rpl16**, *rpl2**(2), *rpl20*, *rpl22*, *rpl23*(2), *rpl32*, *rpl33*, *rpl36*
	小核糖体亚基的蛋白质 Proteins of small ribosomal subunit	*rps11*, *rps12***(2), *rps14*, *rps15*(2), *rps16**, *rps18*, *rps19*(2), *rps2*, *rps3*, *rps4*, *rps7*(2), *rps8*
	RNA聚合酶的亚基 Subunits of RNA polymerase	*rpoA*, *rpoB*, *rpoC1*, *rpoC2*
	核糖体RNA Ribosomal RNAs	*rrn16*(2), *rrn23*(2), *rrn4.5*(2), *rrn5*(2)
	转运RNA Transfer RNAs	*trnA-UGC**(2), *trnC-GCA*, *trnD-GUC*, *trnE-UUC*, *trnF-GAA*, *trnG-GCC*, *trnG-UCC**, *trnH-GUG*(2), *trnI-CAU*(2), *trnI-GAU**(2), *trnK-UUU**, *trnL-CAA*(2), *trnL-UAA**, *trnL-UAG*, *trnM-CAU*, *trnN-GUU*(2), *trnP-UGG*, *trnQ-UUG*, *trnR-ACG*(2), *trnR-UCU*, *trnS-GCU*, *trnS-GGA*, *trnS-UGA*, *trnT-GGU*, *trnT-UGU*, *trnV-GAC*(2), *trnV-UAC**, *trnW-CCA*, *trnY-GUA*, *trnfM-CAU*
其他基因 Other genes	成熟酶 Maturase	*matK*
	蛋白酶 Protease	*clpP*
	包膜蛋白 Envelope membrane protein	*cemA*
	乙酰辅酶A羧化酶 Acetyl-CoA carboxylase	—
	c型细胞色素合成基因 c-type cytochrome synthesis gene	*ccsA*
	翻译起始因子 Translation initiation factor	*infA*
	其他 Other	—
未知功能基因 Genes of unknown function	保守的假设叶绿体开放阅读框 Conserved hypothetical chloroplast open reading frame	*ycf2*(2), *ycf3***, *ycf4*, *ycf68*(2)

*：含一个内含子的基因 Gene with one introns；**：含两个内含子的基因 Gene with two introns；♯：假基因 Pseudo gene；(2)：多拷贝基因的拷贝数 Number of copies of multi-copy genes.

2.4 小黑麦叶绿体基因组重复序列分析

小黑麦叶绿体基因组重复序列长度的最大值和最小值分别为21552与30 bp，前者为跨越IR边界的大型回文重复，后者则为最短同源片段。在此范围内共鉴定出47个重复位点（表4），其中正向重复占主导地位（32个，68.1%），回文重复次之（15个，31.9%）。功能位点中，*rps18*与*rpoC2*为重复热点，表明核糖体蛋白基因及RNA聚合酶亚基基因为复制滑移的易感位点。值得注意的是，43%的重复序列分布于基因间隔区（intergenic region，IGS），其中LSC区最多，表明非编码区是小黑麦叶绿体基因组结构动态变化的主要部位。

表 3　小黑麦叶绿体基因组中含内含子基因的外显子与内含子长度

Table 3　The length of exons and introns in genes with introns in triticale chloroplast genome (bp)

基因 Gene	外显子 1 Exon I	内含子 1 Intron I	外显子 2 Exon Ⅱ	内含子 2 Intron Ⅱ	外显子 3 Exon Ⅲ	基因 Gene	外显子 1 Exon I	内含子 1 Intron I	外显子 2 Exon Ⅱ	内含子 2 Intron Ⅱ	外显子 3 Exon Ⅲ
trnK-UUU	38	2484	36			*rpl2*	388	663	431		
rps16	40	789	218			*ndhB*	775	712	758		
trnS-CGA	31	658	62			*trnT-CGU*	32	787	59		
atpF	145	826	407			*trnA-UGC*	37	805	36		
ycf3	124	756	230	731	159	*ndhA*	550	1036	539		
trnL-UAA	36	588	51			*trnA-UGC*	37	805	36		
trnV-UAC	38	581	53			*trnT-CGU*	32	787	59		
petB	6	748	642			*ndhB*	775	712	758		
petD	8	749	475			*rpl2*	388	663	431		
rpl16	9	1043	402								

2.5　小黑麦叶绿体基因组SSR位点分析

小黑麦叶绿体基因组共含 114 个 SSR 位点，其中 82 个（71.93%）分布于 LSC 区，其余 32 个位点在 IR 区与 SSC 区大致均分，分别为 19 和 13 个（图 3）。这些 SSR 包括单核苷酸重复 73 个（64.04%）、二核苷酸重复 3 个（2.63%）、三核苷酸重复 30 个（26.32%）、四核苷酸重复 5 个（4.39%）及五核苷酸重复 3 个（2.63%）。此外，cpSSR 标记在揭示野生植物物种鉴定、遗传多样性及进化过程方面具有巨大潜力[23]。本研究的 cpSSR 标记可在后期研究中用于小黑麦及其近缘物种的遗传多样性检测，并为抗逆种质的筛选提供有效手段。

2.6　叶绿体基因组比较分析与 IR 边界分析

包括小黑麦在内的 9 个禾本科叶绿体基因组呈现出 2571 bp 的大小变异（134545～137116 bp，图 4）。这种程度的差异较小，不足平均基因组长度的 2%，表明该谱系中 IR 边界的动态变化在进化上受到约束，叶绿体基因组结构相对保守。

小黑麦与小麦属的马卡小麦叶绿体基因组一致，与普通小麦和硬粒小麦的叶绿体基因组高度相似，与普通小麦的差异在于小黑麦的 IR 区进行了 849 bp 扩张，在 LSC 区收缩了 344 bp（图 5）；与硬粒小麦相比，其差异极小，在 LSC 区和 IR 区分别扩张 43 和 11 bp。小黑麦与黑麦属植物野黑麦 6047 品种、瓦维洛夫黑麦和非洲黑麦相比，最大的差异在于小单拷贝区与反向重复区 B 的边界（junction SSC/IRB，JSB）的基因的不同，在黑麦中以 *ndhH* 为边界基因，小黑麦则以 *rps15* 为边界基因（与小麦一致）。*rpl19* 基因作为在叶绿体中转录最丰富的基因[24]，在小黑麦中和 JLB 边界（大单拷贝区与反向重复区 B 的边界 junction LSC/IRB，JLB）的距离与脆轴黑麦和野黑麦 6047 品种在 JLB 边界的距离相比分别收缩了 4 和 6 bp，而 *ndhF* 基因在小黑麦中和 JSB 边界的距离与两者在 JSB 边界的距离相比都收缩了 10 bp，*rps19* 基因在小黑麦和 JLA 边界的距离与两者在 JLA 边界的距离相比分别收缩了 4 bp 和 6 bp，*psbA* 基因在小黑麦中与 JLA 边界（大单拷贝区与反向重复区 A 的边界 junction LSC/IRA，JLA）的距离与两者在 JLA 边界的距离相比分别扩张了 16 和 13 bp。总体而言，黑麦属在 JLB 边界呈独特的 *ndhH* 基因插入（213 bp），而小黑麦保留母本小麦属的 *rps15* 基因结构。

表 4　小黑麦叶绿体基因组中的重复序列

Table 4　Repeat sequences in triticale chloroplast genome

基因 Gene	长度 1 Length 1 (bp)	起始 1 Start 1 (bp)	类型 Type	长度 2 Length 2 (bp)	起始 2 Start 2 (bp)	Hamming 值 Hamming	E 值 E-value	区域 Region
—	21552	80004	P	21552	114347	0	0	LSC-IRb; SSC-IRa
IGS; *rpl23*	286	56298	P	286	82092	−1	2.88×10^{-160}	LSC; IRb
IGS; *rpl23*	286	56298	D	286	133525	−1	2.88×10^{-160}	LSC; IRa

续表 Continued Table

基因 Gene	长度1 Length 1 (bp)	起始1 Start 1 (bp)	类型 Type	长度2 Length 2 (bp)	起始2 Start 2 (bp)	Hamming值 Hamming	E值 E-value	区域 Region
IGS; *rpl23*	111	56473	P	111	82092	0	7.71×10^{-58}	LSC; IRb
IGS; *rpl23-2*	111	56473	D	111	133700	0	7.71×10^{-58}	LSC; IRa
IGS; *IGS*	60	100722	P	60	100722	−2	6.23×10^{-23}	IRb; IRb
IGS; *IGS*	60	100722	D	60	115121	−2	6.23×10^{-23}	IRb; IRa
IGS; *IGS*	60	115121	P	60	115121	−2	6.23×10^{-23}	IRa; IRa
IGS; *rps12*	49	14368	D	49	88910	−3	8.15×10^{-15}	LSC; IRb
IGS; *rps12*	49	14368	P	49	126944	−3	8.15×10^{-15}	LSC; IRa
IGS; *IGS*	43	12606	D	43	12774	−3	2.24×10^{-11}	LSC; LSC
rps18; *rps18*	40	65224	D	40	65266	−2	3.02×10^{-11}	LSC; LSC
trnfM-CAU; *IGS*	41	12682	D	41	36224	−3	3.09×10^{-10}	LSC; LSC
psaB; *psaA*	38	38253	D	38	40477	−2	4.35×10^{-10}	LSC; LSC
IGS; *IGS*	38	41450	D	38	41464	−2	4.35×10^{-10}	LSC; LSC
ycf3; *IGS*	39	43222	D	39	89574	−3	4.24×10^{-9}	LSC; IRb
ycf3; *IGS*	39	43222	P	39	126290	−3	4.24×10^{-9}	LSC; IRa
ndhB; *ndhB*	30	86577	P	30	86577	0	4.51×10^{-9}	IRb; IRb
ndhB; *ndhB*	30	86577	D	30	129296	0	4.51×10^{-9}	IRb; IRa
ndhB; *ndhB*	30	129296	P	30	129296	0	4.51×10^{-9}	IRa; IRa
trnS-GCU; *trnS-GGA*	36	7527	P	36	44746	−2	6.24×10^{-9}	LSC; LSC
IGS; *IGS*	33	60479	P	33	60479	−1	6.97×10^{-9}	LSC; LSC
IGS; *IGS*	38	12577	D	38	36289	−3	1.57×10^{-8}	LSC; LSC
infA; *infA*	38	75735	D	38	75753	−3	1.57×10^{-8}	LSC; LSC
rps18; *rps18*	37	65244	D	37	65286	−3	5.77×10^{-8}	LSC; LSC
IGS; *IGS*	31	100753	D	31	115152	−1	1.05×10^{-7}	IRb; IRa
IGS; *IGS*	36	12785	D	36	36330	−3	2.12×10^{-7}	LSC; LSC
rpoC2; *rpoC2*	36	27000	D	36	27051	−3	2.12×10^{-7}	LSC; LSC
trnfM-CAU; *IGS*	35	12666	D	35	36208	−3	7.77×10^{-7}	LSC; LSC
rpoC2; *rpoC2*	34	27081	D	34	27102	−3	2.84×10^{-6}	LSC; LSC
rpoC2; *rpoC2*	34	27143	D	34	27218	−3	2.84×10^{-6}	LSC; LSC
rpoC2; *rpoC2*	34	27206	D	34	27272	−3	2.84×10^{-6}	LSC; LSC
rps18; *rps18*	34	65266	D	34	65287	−3	2.84×10^{-6}	LSC; LSC
rps3; *rps3*	34	79232	D	34	79277	−3	2.84×10^{-6}	LSC; LSC
rps18; *rps18*	31	65282	D	31	65303	−2	4.71×10^{-6}	LSC; LSC
trnG-UCC; *IGS*	33	12814	D	33	36191	−3	1.04×10^{-5}	LSC; LSC
psbM; *psbM*	30	16660	P	30	16660	−2	1.76×10^{-5}	LSC; LSC
rps18; *rps18*	30	65680	P	30	65680	−2	1.76×10^{-5}	LSC; LSC
trnS-UGA; *trnS-GGA*	32	11223	P	32	44753	−3	3.77×10^{-5}	LSC; LSC
trnT-GGU; *trnT-UGU*	32	14733	P	32	45996	−3	3.77×10^{-5}	LSC; LSC
rpoC2; *rpoC2*	32	26962	D	32	27061	−3	3.77×10^{-5}	LSC; LSC
rpoC2; *rpoC2*	32	27109	D	32	27163	−3	3.77×10^{-5}	LSC; LSC
rpoC2; *rpoC2*	31	26954	D	31	27149	−3	1.37×10^{-4}	LSC; LSC
rpoC2; *rpoC2*	31	27131	D	31	27272	−3	1.37×10^{-4}	LSC; LSC
rps18; *rps18*	30	65213	D	30	65318	−3	4.94×10^{-4}	LSC; LSC
rps18; *rps18*	30	65266	D	30	65308	−3	4.94×10^{-4}	LSC; LSC
rps18; *rps18*	30	65301	D	30	65322	−3	4.94×10^{-4}	LSC; LSC

D：直接重复序列 Direct repeat；P：回文重复序列 Palindromic repeat；IGS：基因间隔区 Intergenic spacer region；LSC：大单拷贝区 Large single copy region；SSC：小单拷贝区 Small single copy region；IRA：反向重复区 A Inverted repeat A；IRB：反向重区 B Inverted repeat B.

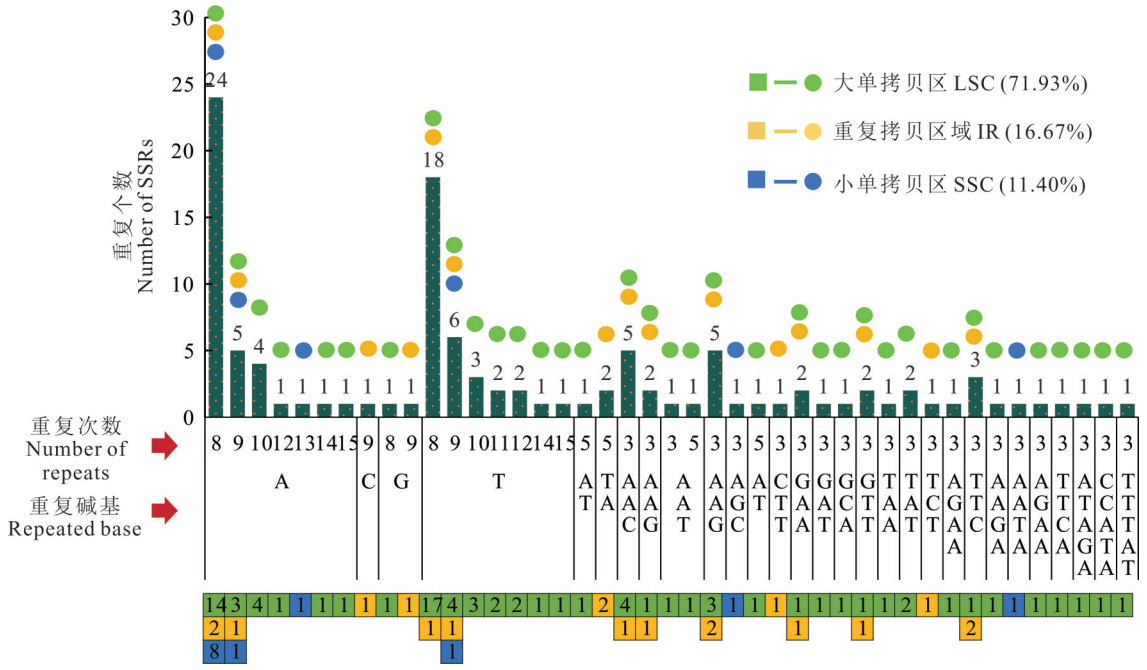


图 3　小黑麦叶绿体基因组中重复序列长度、重复序列及区域

Fig. 3　Length of repeat, repeate sequences and region in the triticale chloroplast genome

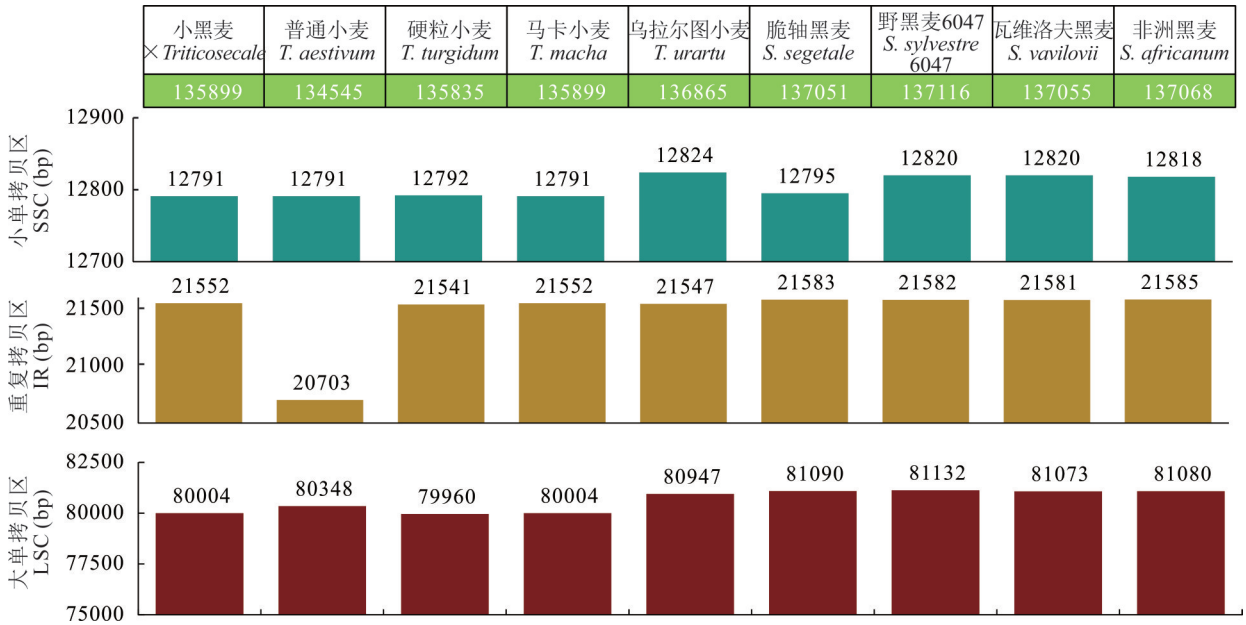


图 4　小麦属、黑麦属与小黑麦的叶绿体基因组比较

Fig. 4　Comparison of chloroplast genomes in the genus *Triticum*, *Secale* and triticale

比较基因组分析揭示了小麦、黑麦及其杂交种小黑麦间的DNA序列分歧特征。基于mVISTA程序，以注释的小黑麦叶绿体基因组为参照，绘制了9个物种叶绿体基因组的全序列一致性图谱（图6）。结果显示，编码区保守性显著高于非编码区，高度分歧区域主要位于基因间隔区，如LSC区的*psbM-petN*、*petN-rpoB*及*ndhC-ycf4*等位点；而边界基因*rps19*、*ndhF*和*psbA*在所有5个基因组中的排列方式均保持一致。

2.7　小黑麦系统发育分析

小黑麦与硬粒小麦TA2836品种及马卡小麦在小麦类群内形成一个高支持率分支（图7）。小黑麦与这些小麦物种间的遗传距离极短（<0.001），指示其作为近期人工合成杂交种的起源特性，与其可追溯至140年前的育

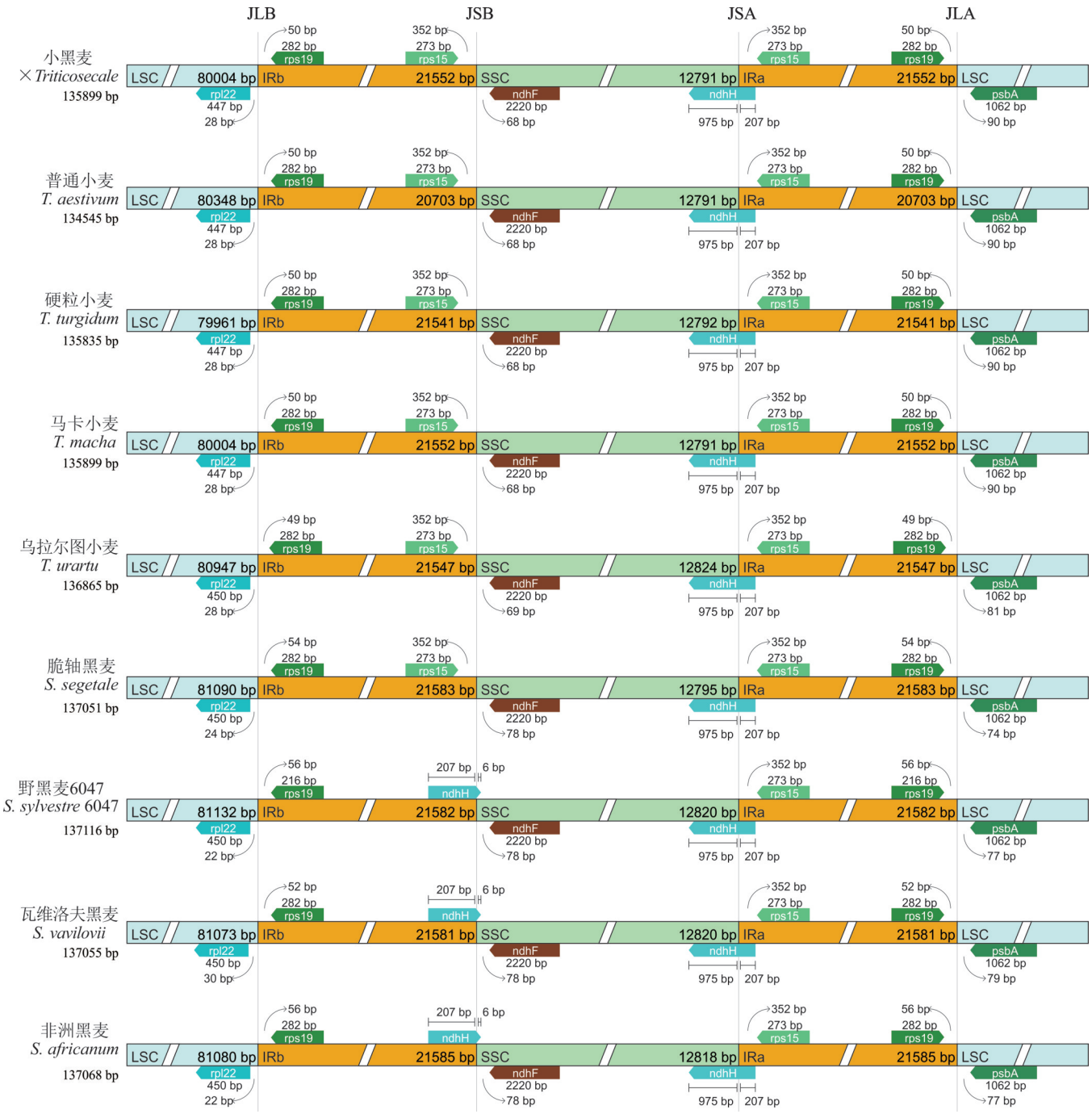


图5 小麦属、黑麦属及小黑麦的LSC、SSC及IR区域边界比较

Fig. 5 Comparison of the borders of the LSC, SSC, and IR regions of the genus *Triticum*, *Secale* and triticale

种历史相符[5],小黑麦系通过小麦与黑麦杂交后经染色体加倍而创制。因此叶绿体基因组系统发育分析将小黑麦置于小麦谱系之中,而小麦类群连同多数山羊草属物种构成黑麦类群的姐妹群。

3 讨论

小黑麦是我国西北农牧交错带的优质饲草,兼具高产和抗寒,能够充分利用冬春闲田,并充分实现草田轮作,在全国多个地区实现了“饲用小黑麦+青贮玉米”一年两作的轮作模式[25]。小黑麦具有较好的耐盐性,可在盐碱地资源的修复改良中发挥较好的作用[10]。作为一种优质饲草,提高产量和培育抗逆品种是首要目标,分子育种手段在当前可以节约传统育种的年限,而分子育种的基础是对质体基因组的充分挖掘和解析。小黑麦是杂交种,具

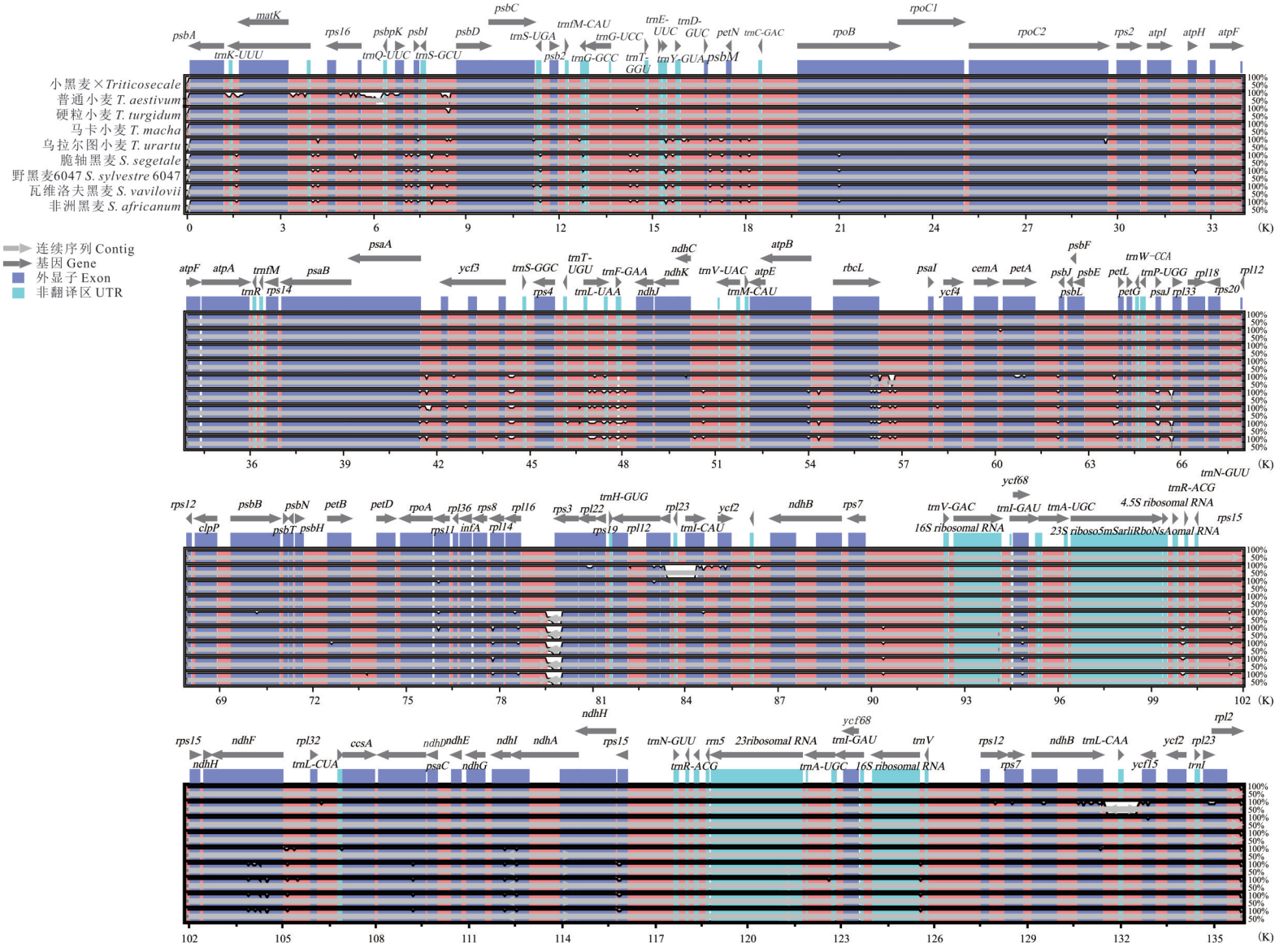


图 6　小麦属、黑麦属及小黑麦的叶绿体基因组序列分析

Fig. 6　Sequence analysis of chloroplast genome with the genus *Triticum*, *Secale* and triticale

比对结果上方的灰色箭头和粗黑线表示基因方向 Above the alignment result, gray arrows and thick black lines indicate gene orientation; 紫色条带 Purple band: 外显子 Exons; 蓝色条带 Blue band: 非翻译区 Untranslated regions; 粉色条带 Pink band: 保守非编码序列 Conserved non-coding sequences; Y 轴刻度表示相似度百分比(显示范围:50%～100%)The Y-axis scale indicates the percentage of similarity (display range: 50%－100%).

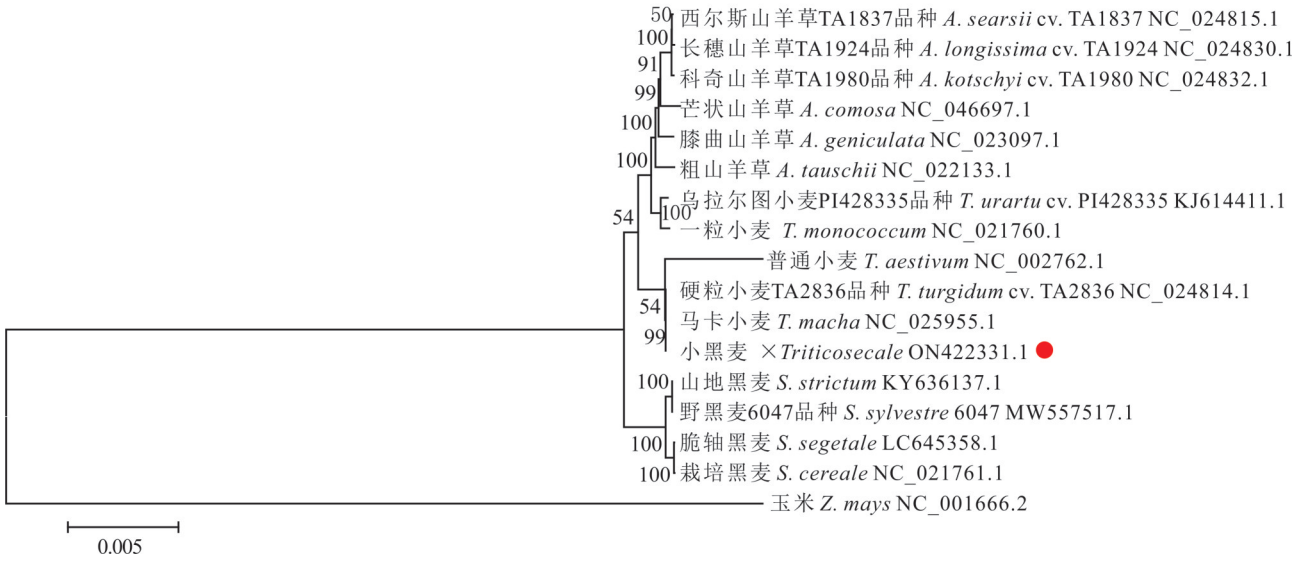


图7 17个禾本科植物叶绿体基因组的系统发育树

Fig. 7 Phylogenetic tree of chloroplast genomes from 17 Poaceae species

有典型的母系遗传特性，而叶绿体基因组具有母系遗传特征，是研究母系遗传的最佳质体基因组，因此，通过解析叶绿体基因组研究小黑麦遗传具有重要意义。

研究结果显示，小黑麦叶绿体基因组全长135899 bp，与大多数陆生植物叶绿体基因组一致[26]。小黑麦叶绿体基因组中共含有112个基因，其中19个基因含有内含子，而*ycf3*基因含有2个内含子。*ycf3*基因是光合作用系统（尤其是PSI复合组装过程中）发挥功能不可或缺的重要元件，直接与PSI的PsaA和PsaB核心亚基相互作用，维持光合生命活动[27]。研究表明，内含子长度的变化可以反映光合功能状态变化[28]。在小黑麦中，*ycf3*基因的2个内含子，长度分别为756和731 bp，这种长内含子长度的增加可能促进光合系统的功能，进而反映在其叶片蛋白含量等营养成分上。因此，内含子中可能包含在进化过程中失去功能的基因片段，如单细胞真核生物由于没有复杂的细胞分化，不需要内含子参与的精细调控，从而在进化过程中倾向于失去无用的内含子[29]。然而，内含子在植物中仍然存在，有研究认为是内含子从祖先基因组（内含子贫乏）开始，在核基因组中随机、逐步积累，具有一定的进化优势。在小黑麦中，多个基因含有内含子，这可能是小黑麦的母系遗传导致，其母本小麦在进化过程中对不同环境的适应性而使内含子选择性变异，在杂交中具有进化优势的内含子通过遗传保留在小黑麦中。

叶绿体SSR表现为单亲遗传，广泛用于群体遗传学、多样性评估和母系分析[11,14]。小黑麦叶绿体基因组含有114个SSR，主要分布于LSC区，且大多数为单核苷酸和三核苷酸重复，这一模式与大多数植物一致[30]。单核苷酸SSR主要为A/T类型，可能反映了A-T碱基配对的热力学稳定性。由于cpSSR可作为研究物种遗传进化的宝贵标记，因此在小黑麦叶绿体基因组中鉴定到的SSR位点，可以为小黑麦的群体遗传学和种质保护提供分子育种基础。

本研究揭示的IR边界变异模式为理解人工多倍体细胞器基因组重塑机制提供了量化证据。小黑麦与普通小麦有849 bp的IR区特异性扩张及其LSC区相应收缩（图4），远超天然多倍体小麦在驯化过程中的边界漂移幅度（通常＜100 bp），这一结构性差异凸显了近期（约140年）人工杂交多倍化事件对叶绿体基因组稳定性的瞬时冲击。该变异并非中性进化痕迹，而是细胞器基因组对核基因组剂量倍增的代偿性响应，暗示人工选择可能通过核－质互作反馈回路加速了IR边界的重排速率。从功能进化角度看，JSB边界基因不保留黑麦属的*ndhH*而保留小麦属的*rps15*，揭示了人工合成物种在保留母本转录特征与适应异源核背景间的权衡策略：*rps15*作为核糖体小亚基蛋白基因，其双拷贝定位可优化多倍体细胞对蛋白质合成能力的需求，而黑麦*ndhH*插入片段的丢失虽可能削弱部分胁迫响应通路，但已在六倍体小麦核基因组中获得功能补偿。这种边界基因替换模式表明，在小黑麦

叶绿体基因工程中，直接移植黑麦调控元件未必有效，必须考虑核－质协同进化背景[31]。研究中精确定位差异（如*psbA*扩张 16 bp、*rps19*位移 4～6 bp）具有开发成分子标记的潜力，可用于鉴别不同小麦母本来源的小黑麦品种，为细胞质指纹图谱构建提供结构变异靶点。但本研究的局限在于仅依赖单一基因组序列，未来需扩大不同小黑麦育种系的样本量，结合转录组数据验证边界变异对基因表达的影响。

系统发育树以 bootstrap 高支持率将小黑麦嵌合于小麦属分支内部，与硬粒小麦 TA2836 品种和马卡小麦形成姐妹群，这从全基因组层面确证了小黑麦叶绿体的母系遗传特性。相较于六倍体普通小麦，小黑麦与四倍体硬粒小麦 TA2836 品种的超短距离聚类关系表明该育种系采用四倍体硬粒小麦（AABB）×黑麦（RR）的杂交途径，而非通过淘汰 D 基因组获得，这一分子诊断信息可有效区分不同育种系谱来源。更关键的是，尽管小黑麦核基因组已整合黑麦 R 组染色体，其细胞质基因组仍完整保留母本小麦结构，其杂种优势可能源于核－质基因组间的特异性互作平衡，后续叶绿体基因工程改造必须考量这一背景。未来研究可基于该拓扑结构开发细胞质系谱特异性标记，为小黑麦细胞质指纹图谱构建提供系统发育学支撑。

4　结论

研究利用公共数据资源（NCBI 登录号：ON422331.1）完成了小黑麦叶绿体基因组的深度解析，明确了其基因组结构、重复序列特征及母系遗传来源。鉴定的 114 个 cpSSR 位点与 IR 边界结构性变异（*psbA*、*rps19*等基因精确定位差异）可转化为分子标记，可用于区分不同细胞质来源育种系。IR 区 849 bp 扩张模式揭示了人工多倍化对叶绿体基因组稳定性的瞬时冲击，为理解核－质协同进化提供了定量参考。然而，本研究局限于单一基因组序列，未来需扩大小黑麦育种系群体样本，结合转录组与功能试验验证重复序列及边界变异的生物学效应。总体而言，研究为叶绿体定向基因工程及分子标记辅助选择奠定了理论基础。

参考文献 References：

[1] Mergoum M, Singh P K, Peña R J, *et al*. Triticale: A "new" crop with old challenges//Handbook of plant breeding. New York: Springer US, 2009: 267－287.

[2] Zhu F. Triticale: Nutritional composition and food uses. Food Chemistry, 2018, 241: 468－479.

[3] Mcgoverin C M, Snyders F, Muller N, *et al*. A review of triticale uses and the effect of growth environment on grain quality. Journal of the Science of Food and Agriculture, 2011, 91(7): 1155－1165.

[4] Wang W Q, Liu J, Tian X H, *et al*. Evaluations on the adaptability of ×*Triticosecale* Wittmack 'Gannong No. 4' in different regions of Qinghai Province. Acta Agrestia Sinica, 2020, 28(6): 1626－1634.
王伟强, 刘晶, 田新会, 等. '甘农 4 号'小黑麦品种在青海省不同区域的适应性评价. 草地学报, 2020, 28(6): 1626－1634.

[5] Mergoum M, Sapkota S, ElDoliefy A E F A, *et al*. Triticale (×*Triticosecale* Wittmack) breeding//Advances in plant breeding strategies: Cereals. Cham: Springer International Publishing, 2019: 405－451.

[6] Simmonds N W. Evolution of crop plants. London: Longman Group Ltd., 1976.

[7] Hammouda D, Baaziz N, Khalfallah N. Genetic characterization of octoploid (AABBDDRR) and hexaploid (AABBRR) triticales. European Scientific Journal, 2015, 11(9): 284－297.

[8] Wang B L, Chen Q F, Men W Q, *et al*. Analysis of chromosome composition of hexaploid triticale of different plant heights. Journal of Henan Agricultural University, 2021, 55(4): 631－638.
王贝麟, 陈麒帆, 门文强, 等. 六倍体小黑麦株高分离的染色体构成分析. 河南农业大学学报, 2021, 55(4): 631－638.

[9] Yang T, Chen X Z, Sui J S, *et al*. Chromosomic analysis of hexaploid triticale by fluorescence *in situ* hybridization. Journal of Southwest University (Natural Science Edition), 2019, 41(4): 42－48.
杨婷, 陈星灼, 隋建枢, 等. 六倍体小黑麦染色体的荧光原位杂交分析. 西南大学学报(自然科学版), 2019, 41(4): 42－48.

[10] Wang S X. A "contest" with saline-alkali land with triticale. Lanzhou: Gansu Daily Publishing House, 2025.
王思璇. "以麦治碱"一场与盐碱地的"较量". 兰州: 甘肃日报出版社, 2025.

[11] Moore M J, Soltis P S, Bell C D, *et al*. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. Proceedings of the National Academy of Sciences of the United States of America, 2010, 107(10): 4623－4628.

[12] Tian T, Zhang Y X, Du W H, *et al*. Comparative genomics of *Triticum*, *Secale*, and triticale: Codon usage bias in chloroplast genomes and its implications for evolution and genetic engineering. International Journal of Molecular Sciences, 2025, 26(21):

10266.

[13] Zhang Y Z. Evolutionary patterns analysis of the main morphology and karyotypic ancestral characteristics of the triticeae. Ya'an: Sichuan Agricultural University, 2023.
张瑜卓. 小麦族植物主要形态及核型祖先特征的演化式样分析. 雅安: 四川农业大学, 2023.

[14] Tian T, Wang H X. Characterization analysis of microsatellite in chloroplast of *Syringa reticulata* subsp. *amurensis*. Molecular Plant Breeding, 2024, 22(8): 2501-2507.
田甜, 王宏霞. 暴马丁香的叶绿体微卫星特征分析. 分子植物育种, 2024, 22(8): 2501-2507.

[15] Cao D. Triticale chloroplast genome. GenBank, ON422331.1, 2022.

[16] Greiner S, Lehwark P, Bock R. Organellar Genome DRAW (OGDRAW) version 1.3.1: Expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Research, 2019, 47(W1): W59-W64.

[17] Kurtz S, Choudhuri J V, Ohlebusch E, *et al*. REPuter: The manifold applications of repeat analysis on a genomic scale. Nucleic Acids Research, 2001, 29(22): 4633-4642.

[18] Beier S, Thiel T, MÜnch T, *et al*. MISA-web: A web server for microsatellite prediction. Bioinformatics, 2017, 33(16): 2583-2585.

[19] Raubeson L A, Peery R, Chumley T W, *et al*. Comparative chloroplast genomics: Analyses including new sequences from the angiosperms *Nuphar advena* and *Ranunculus macranthus*. BMC Genomics, 2007, 8(1): 174.

[20] Amiryousefi A, Hyvönen J, Poczai P. IRscope: An online program to visualize the junction sites of chloroplast genomes. Bioinformatics, 2018, 34(17): 3030-3031.

[21] Katoh K, Standley D M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Molecular Biology and Evolution, 2013, 30(4): 772-780.

[22] Nguyen L T, Schmidt H A, Von Haeseler A, *et al*. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Molecular Biology and Evolution, 2015, 32(1): 268-274.

[23] Ebert D, Peakall R O D. Chloroplast simple sequence repeats (cpSSRs): Technical resources and recommendations for expanding cpSSR discovery and applications to a wide array of plant species. Molecular Ecology Resources, 2009, 9(3): 673-690.

[24] Lee J, Kang Y, Shin S C, *et al*. Combined analysis of the chloroplast genome and transcriptome of the *Antarctic vascular* plant *Deschampsia antarctica* Desv. PLoS One, 2014, 9(3): e92501.

[25] Du W H, Tian X H, Liu H C, *et al*. Industrialization demonstration of high efficient double cropping pattern of Gannong No. 2 triticale in alpine pastoral area and agro-pastoral ecotone of Gansu Province. Lanzhou: Gansu Agricultural University, 2023.
杜文华, 田新会, 刘汉成, 等. 甘肃省高寒牧区和农牧交错区甘农2号小黑麦一年两作高效种植模式产业化示范. 兰州: 甘肃农业大学, 2023.

[26] Wicke S, Schneeweiss G M, Depamphilis C W, *et al*. The evolution of the plastid chromosome in land plants: Gene content, gene order, gene function. Plant Molecular Biology, 2011, 76(3): 273-297.

[27] Naver H, Boudreau E, Rochaix J D. Functional studies of *Ycf3*: Its role in assembly of photosystem Ⅰ and interactions with some of its subunits. The Plant Cell, 2001, 13(12): 2731-2745.

[28] Boudreau E, Takahashi Y, Lemieux C, *et al*. The chloroplast *ycf3* and *ycf4* open reading frames of *Chlamydomonas reinhardtii* are required for the accumulation of the photosystem Ⅰ complex. The EMBO Journal, 1997, 16(20): 6095-6104.

[29] Irimia M, Roy S W. Spliceosomal introns as tools for genomic and evolutionary analysis. Nucleic Acids Research, 2008, 36(5): 1703-1712.

[30] Guo S, Guo L L, Zhao W, *et al*. Complete chloroplast genome sequence and phylogenetic analysis of *Paeonia ostii*. Molecules, 2018, 23(2): 246.

[31] Goldbecker E S, de Vries J. Systems biology of streptophyte cell evolution. Annual Review of Plant Biology, 2025, 76(1): 493-522.