

两种茶尺蠖近缘种基因组草图比较分析*

程梓淇^{1**} 李佳琦² 叶昕海¹ 李 飞¹ 贺 康^{1***}

(1. 浙江大学昆虫科学研究所, 浙江省作物病虫害生物学重点实验室, 杭州 310058;

2. 浙江大学茶叶科学研究所, 杭州 310058)

摘 要 【目的】灰茶尺蠖 *Ectropis grisescens* 和茶尺蠖 *E. obliqua* 是东南亚茶园中重要的鳞翅目害虫, 二者形态相似, 但茶尺蠖核型多角体病毒 (*Ectropis obliqua* nuclear polyhedrosis virus, *EoNPV*) 对二者的防效存在差异。为深入解析灰茶尺蠖和茶尺蠖基因组水平物种差异特性, 本研究利用全基因组测序, 组装和注释了灰茶尺蠖和茶尺蠖基因组草图, 并开展比较基因组分析。【方法】通过流式细胞术测定灰茶尺蠖和茶尺蠖的基因组大小; 采用 Illumina 二代测序技术获得 2 个近缘种低覆盖度基因组测序数据, 完成基因组组装和注释后对其进行系统进化和基因共线性分析以及 1 018 个单拷贝同源基因的正选择分析。同时分析了 2 种茶尺蠖的病毒抗性基因和化学感受相关基因家族, 并对灰茶尺蠖和茶尺蠖的物种特异性序列进行鉴定和验证。【结果】利用低覆盖度基因组测序数据, 组装构建了灰茶尺蠖和茶尺蠖的基因组草图, 二者基因组大小约为 770 Mb。比较基因组学分析显示, 灰茶尺蠖和茶尺蠖的基因数量相似, 但重复序列存在差异, 分别为 52.62% 和 49.67%。系统发育分析和全基因组共线性分析表明, 二者分化时间为 318 万年前, 但未发生明显的染色体重排事件。灰茶尺蠖中与 *EoNPV* 抗性相关基因载脂蛋白和组织蛋白酶基因发生了扩增, 包括小分子代谢过程通路 (GO: 0044281) 显著正向选择。【结论】本研究获得了灰茶尺蠖和茶尺蠖的基因组草图数据, 分析结果为研究 2 个近缘种的适应性进化策略提供了依据, 同时也为减少农药的使用和开发绿色茶树害虫防治技术提供了参考。

关键词 茶尺蠖; 基因组; 近缘种; 基因组大小; 适应性; 特异性引物

Comparative analysis of the draft genome of two sibling tea geometrids, *Ectropis grisescens* and *Ectropis obliqua*

CHENG Zi-Qi^{1**} LI Jia-Qi² YE Xin-Hai¹ LI Fei¹ HE Kang^{1***}

(1. Key Laboratory of Biology of Crop Pathogens and Insects of Zhejiang Province, Institute of Insect Sciences, Zhejiang University, Hangzhou 310058, China; 2. Department of Tea Science, Zhejiang University, Hangzhou 310058, China)

Abstract 【Objectives】To analyze genome-level species differences between two sibling species of tea geometrids, *Ectropis grisescens* and *E. obliqua*, two major lepidopteran pests in southeast Asian tea plantations with very similar morphology, but different susceptibility to the *E. obliqua* nuclear polyhedrosis virus (*EoNPV*), a common biological pesticide. 【Methods】We assembled and annotated the draft genomes of *E. grisescens* and *E. obliqua* by whole-genome sequencing data and conducted comparative genomic analysis. Flow cytometry was used to estimate the genome size of each species, after which low-coverage genome sequencing data were obtained using Illumina next-generation sequencing technology. After genome assembly and annotation, we conducted phylogenetic and synteny analysis. Positive selection analysis of 1 018 single-copy homologous genes was also conducted. We then analyzed chemosensory genes and those involved in *EoNPV* resistance and identified and verified species-specific sequences. 【Results】We used low-coverage genome survey sequencing to successfully assemble and annotate the draft genomes of *E. grisescens* and *E. obliqua*, which have a similar genome size of 770 megabase (Mb) pairs. Comparative genomics indicates a similar length of contigs, scaffolds and gene

*资助项目 Supported projects: 国家重点研发计划 (2019YFD1002100) 和浙江省基础公益研究计划 (LY22C140005)

**第一作者 First author, E-mail: 22116013@zju.edu.cn

***通讯作者 Corresponding author, E-mail: hekang@zju.edu.cn

收稿日期 Received: 2022-06-10; 接受日期 Accepted: 2022-07-01

number, but a different number of repeat sequences (52.62% for *E. grisea* and 49.67% for *E. obliqua*). Phylogenetic and whole-genome synteny analysis indicates that these sibling species diverged 318 million year ago with no obvious chromosome rearrangement. Gene expansion of apolipoprotein and cathepsin has occurred, including significant positive selection on the GO pathway 0044281 (small molecule metabolic process) in *E. grisea*. **[Conclusion]** We successfully used genome survey sequencing to assemble and annotate draft genomes of *E. grisea* and *E. obliqua*. The results provide fundamental data for genetic comparison between these tea geometrids, thereby facilitating understanding of the adaptive evolution of sibling species, and may also allow reduced pesticide use and promote more environmentally friendly methods of controlling these pests.

Key words tea geometrid; genome; sibling species; genome size; adaption; specific primer

茶树是东南亚最重要的经济作物之一 (Ye *et al.*, 2014)。灰茶尺蠖 *Ectropis grisea* 和茶尺蠖 *E. obliqua* 是茶园中主要的咀嚼式口器害虫 (Hazarika *et al.*, 2008; Sinu *et al.*, 2013)。与茶尺蠖相比, 灰茶尺蠖的体型较大, 繁殖能力较强, 但发育历期较短, 分布范围较广 (姜楠等, 2014; Zhang *et al.*, 2014; 白家赫等, 2018)。因二者生物学形态相似, 生态学特征相仿, 仅通过形态特征难以有效区分。利用微卫星多态性等分子标记, 可以方便地区分和确定茶尺蠖的不同地理种群 (Jing *et al.*, 2016), 但关于灰茶尺蠖与茶尺蠖之间的基因组水平的差异研究较少。

不对称生殖干扰, 即一个物种的适合度比另一个物种受到更多的影响, 会导致劣种被取代, 这一过程也被称为生殖排斥 (Groning *et al.*, 2008)。生殖干扰通常见于近缘物种, 如烟粉虱 *Bemisia tabaci* (生物型 B 和生物型 Q) (Crowder *et al.*, 2010)、蚊子 (*Aedes aegypti* 和 *A. albopictus*) (Bargielowski *et al.*, 2015)、寄生蜂 (*Encarsia szanayi* 和 *E. gennaro*) (Gebiola *et al.*, 2017) 及瓢虫 (*Harmonia yedoensis* 和 *H. axyridis*) (Noriyuki *et al.*, 2012) 等。在实验室条件下, 灰茶尺蠖和茶尺蠖杂交能够产生可存活但不育的 F₁ 后代, 表明二者可能存在不完整的隔离交配系统。这种生殖隔离可能是由昆虫激素生物合成途径中的基因 (*JHDK* 和 *Dib*) 引起, 调控杂交后代的不同步发育进程 (Wang *et al.*, 2019)。两个近缘物种之间微生物群 (如沃尔巴克氏体 *Wolbachia*) 的不同组成也可能影响生殖隔离 (Wang *et al.*, 2020)。此外, 雄性灰茶尺蠖的交配干扰能力强于雄性茶尺蠖, 这可能是由于

其交配识别能力强于茶尺蠖 (Zhang *et al.*, 2016)。近年来, 通过触角电位图 (Electroantennography, EAG) 和风洞实验, Luo 等 (2017) 鉴定了茶尺蠖中特异的性信息素成分 (Z3, epo6, Z9-19:H), 对研究茶尺蠖种间交配分离机制具有重要意义。气味结合蛋白 (Odorant binding protein, OBP) 和化学感觉蛋白 (Chemosensory protein, CSP) 的系统发育关系表明, 在气味识别方面, 茶尺蠖与其它鳞翅目昆虫存在一定的保守性 (毛腾飞等, 2017)。然而, 从基因组水平解释两个物种不对称生殖干扰现象却鲜有报道。

目前, 茶尺蠖核型多角体病毒 (*Ectropis oblique* nuclear polyhedrosis virus, *EoNPV*) 已广泛应用于茶树害虫的生物防治 (殷坤山等, 2000; Hazarika *et al.*, 2001; Ye *et al.*, 2014)。然而, 不同地理种群的茶尺蠖易感性具有显著差异, 研究表明 *EoNPV* 会导致茶尺蠖 82.7%-100% 的死亡率, 但灰茶尺蠖对 *EoNPV* 的敏感程度要弱于茶尺蠖 (席羽, 2011; 唐美君等, 2017)。到目前为止, 已经在家蚕和鳞翅目昆虫当中发现了数十个与核型多角体病毒相关的基因、信号通路和代谢途径 (Nguyen *et al.*, 2013; Jiang and Xia, 2014; Kong *et al.*, 2018; Lü *et al.*, 2018), 两个物种在基因组水平上对 *EoNPV* 的易感性差异仍不清楚。

近缘物种间的比较基因组学分析能够为理解适应性差异的遗传基础提供见解 (Massoumi *et al.*, 2014; Sillo *et al.*, 2015; Buiate *et al.*, 2017; Luo *et al.*, 2017; Vicente *et al.*, 2017)。然而截止目前, 仅有灰茶尺蠖的基因组完成了测序 (Pan *et al.*,

2021)。为解析两个茶尺蠖近缘种物种分化和适应性进化的基因组特征,本研究拟构建茶尺蠖和灰茶尺蠖的基因组草图,利用比较基因组学方法分析茶尺蠖与灰茶尺蠖在基因组水平上的差异,鉴定二者与 *EoNPV* 易感性差异相关的基因家族和基因区域,以及与非对称生殖干扰的相关化学感受基因家族。同时利用物种特异性基因,鉴定并验证能够有效区分两近缘种的特异性序列和引物,为两种害虫的防治策略的研发提供新的思路。

1 材料与方法

1.1 供试昆虫

供试灰茶尺蠖和茶尺蠖来自实验室繁育多代的稳定种群,幼虫均使用新鲜茶树叶片饲喂,饲养环境条件设置为:(25 ± 1) $^{\circ}\text{C}$,相对湿度 70% $\pm$ 5%,光暗比 16L : 8D (Luo *et al.*, 2017)。

1.2 流式细胞术估计基因组大小

细胞核悬浮液制备参考标准程序 (Dolezel *et al.*, 2007; He *et al.*, 2016)。取 10 只成虫头部组织于 600 μL 裂解液中,手术刀片快速剁碎过滤后,依次加入 PI 染液和 RNase A 在避光条件下 4 $^{\circ}\text{C}$ 染色 10 min,然后利在 FACSCaliburTM 流式细胞仪 (BD Biosciences, San Jose, CA) 上机分析样品。通过 FACSComp 软件 (version 4.0; BD Biosciences) 获取样品中 5 000 个细胞核的信号,采用 FLOWJO 软件 (version 7.6.1 <https://www.flowjo.com/>) 分析核峰 (2C-value)。以野生型黑腹果蝇 CantonS 品系细胞核悬浮液作为外参,参数设置相同,所有实验重复 4 次。

1.3 基因组测序、组装和 K-mer 分析

试虫基因组 DNA 提取采用 CTAB 方法 (Doyle, 1987; Porebski *et al.*, 1997)。利用 NEB Next[®] UltraTM DNA 文库准备试剂盒 (New England Biolabs, Ipswich, MA, USA) 构建插入片段为 350 bp 的文库,并在 Illumina HiSeq 2000 平台完成双端 PE150 测序。原始测序数据使用 Trimmomatic 软件 (version 0.36) 过滤接头和引物等低质量和重复的 reads 后 (Bolger *et al.*,

2014),使用 SOAPdenovo2 (version 2.04-r240) 进行 *de novo* 组装 (Luo *et al.*, 2012),原始读取的 contig 和 scaffold 设置为 41-mer。利用 Jellyfish (version 1.1.4) 分析原始测序数据,对基因组大小和基因重复序列进行独立估算,设置 K-mer 大小为 17 以覆盖全部基因组大小 (Kingsford, 2011)。

1.4 基因组注释

采用优化后的 OMIGA 基因组注释流程对基因组进行注释 (Liu *et al.*, 2014b),并进行基因的同源和从头预测。首先,对基因组中的重复序列进行识别和屏蔽。串联重复序列由 Tandem Repeats Finder (TRF) 软件 (version 4.0.4) 识别 (Benson, 1999),识别获得的重复序列通过 RECON (version 1.08) (Bao *et al.*, 2002) 和 RepeatScout (version 1.0.5) (Price *et al.*, 2005) 进行预测。重复序列注释使用 Repeat Modeler (version 1.0.11) 软件对基因组中的重复序列进行从头预测,将构建的物种特异重复序列数据库与 RepBase (version 2017.01.27) 数据库合并,在使用 RepeatMasker 4.0.7 软件基于合并后的重复序列数据库对基因组中的重复序列进行注释和屏蔽。所用程序均使用默认参数。

使用 MAKER (version 2.31.9) 将同源预测结果和从头预测结果相结合 (Holt and Yandell, 2011),对蛋白质编码基因进行结构注释。编码基因注释使用 Augustus (version 2.5.5) (Mario *et al.*, 2006) 和 SNAP (version 2013.11.29) 软件 (Ian, 2004)。此外,利用诗神袖蝶 *Heliconius melpomene*、家蚕 *Bombyx mori*、大红斑蝶 *Danaus plexippus*、小菜蛾 *Plutella xylostella* 和黑腹果蝇 *Drosophila melanogaster* 的蛋白质组数据,以及家蚕和大红斑蝶的表达序列标签 (EST) 数据,作为在 MAKER 中的同源性证据。同源蛋白序列和表达序列标签数据均从 NCBI 无脊椎动物 RefSeq (<https://www.ncbi.nlm.nih.gov/refseq/>) 下载。

对蛋白编码基因进行功能注释,在 UniProtKB/Swiss-Prot 数据库 (<https://www.uniprot.org/>) 上进行 BLASTP, *e* 值为 1×10^{-5} 。使用 ID 映射表从 BLASTP 结果中获得的 GO 注释结果。KEGG 通路注释使用 BlastKOALA web

服务器获得 (Kanehisa *et al.*, 2016)。选择结果中与病毒感染和免疫系统相关的基因进行进一步分析。

1.5 进化分析

采用标准 OrthoMCL 流程 (version 2.0.9) (Li *et al.*, 2003) 对灰茶尺蠖、茶尺蠖、诗神袖蝶、家蚕、大红斑蝶、小菜蛾、冬尺蠖 *Operophtera brumata* 和黑腹果蝇进行同源进化分析。所有的蛋白序列都是从 NCBI 中下载, 并用自定义 perl 脚本删除冗余序列。使用默认参数, 得到 1 018 个单拷贝同源基因以及 15 332 个同源组。使用 MAFFT (version 7.402) (Katoh *et al.*, 2013) 对单拷贝同源基因进行比对, 并用 trimAl 软件 (version 3, latest updated: 29 June 2007) 对保守区进行整理 (Capella-Gutierrez *et al.*, 2009)。系统发育树使用 RAxML 软件构建 (Stamatakis, 2014), 并通过 ProtTest (version 3) (Darriba *et al.*, 2011) 选择合适的 VT 模型, bootstrap 值为 1 000。利用 r8s (version 1.8.1) (Sanderson, 2003) 估算了黑腹果蝇和家蚕的系统发育时间。系统发育树的绘制使用 FigTree (version 1.4.3, <http://tree.bio.ed.ac.uk/software/figtree/>)。利用 EMBOSS 包 (version 6.6.0) 中的 needle 模块统计每对同源基因的成对氨基酸同源性分布 (Olson, 2002)。

1.6 共线性分析

使用 LASTZ (version 1.02.0) (Harris, 2007) 对两个物种的基因组 Scaffolds 进行比对。评分矩阵设为 (1, 5), 置信度为 95%。基因组序列使用最佳比对命中 (Score > 2 300), 然后进行排序和可视化, 用自定义脚本表示 scaffolds 的相似性。通过基因组上相应共线片段的分布情况, 分析基因组重排和扩增的可能性。

1.7 K_a/K_s 进化选择压力分析

使用 Codeml 程序以及 PAML 包 (version 4.9i) (Yang, 2007) 对每个单拷贝同源基因的氨基酸非同义替换 (K_a) 与同义替换 (K_s) 的平均

比值进行估算。 K_a/K_s 值大于 1 的基因被认为是正选择基因, 表明自然选择促进了蛋白质序列当中氨基酸的变化。将这些单拷贝同源基因利用注释信息进一步分配到 GO 功能分类上, 并采用双样本 *t*-检验来识别具有超过 5 个同源基因的 GO 类别。

1.8 基因家族分析

利用 NCBI 蛋白数据库中蛾类蛋白质序列, 分析灰茶尺蠖和茶尺蠖的气味受体 (ORs)、味觉受体 (GRs)、化学感觉蛋白 (CSPs) 和气味结合蛋白 (OBPs) 的化学感受受体相关基因家族。利用 BLASTP 对 *e*-value $< 1 \times 10^{-5}$ 的参考蛾类序列进行比对, 获得了家蚕、灰茶尺蠖和茶尺蠖的蛋白家族候选基因。对所有候选序列针对 Pfam 数据库 (version 32.0) 进行 HMMER 筛选 (version 3.1b2; <http://hmmer.org/>) (El-Gebali *et al.*, 2019)。筛选后的基因使用 MAFFT (version 7.402) (Katoh *et al.*, 2013) 进一步比对, 并使用 trimAl 软件 (version 3, latest updated: 29 June 2007) (Capella-Gutierrez *et al.*, 2009) 对序列进行裁剪。系统发育树由 RAxML (Stamatakis, 2014) 构建, 并通过 ProtTest (version 3) (Darriba *et al.*, 2011) 选择合适的模型, bootstrap 值为 1 000。系统发育树用 FigTree (version 1.4.3, <http://tree.bio.ed.ac.uk/software/figtree/>) 进行可视化。

1.9 物种特异性序列验证

从 OrthoMCL 数据中集中选择物种特异性的单拷贝同源组, 利用 Primer Premier 5.0 软件 (<http://www.premierbiosoft.com/primerdesign/>) 设计引物, 并通过 NCBI Primer-blast 在线服务器 (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>) 检测其特异性 (表 1)。为验证这些引物的特异性, 随机选择 16 个基因进行聚合酶链式反应 (PCR)。引物由擎科生物技术有限公司 (北京) 合成。使用 DNeasy Blood & Tissue kit (QIAGEN, Venlo, Netherlands) 试剂盒对灰茶尺蠖和茶尺蠖的 4 龄幼虫进行 DNA 提取。以家蚕基因组 DNA 为阴性对照。

表 1 灰茶尺蠖和茶尺蠖物种特异性基因以及引物

Table 1 Species specific gene and the primers of *Ectropis grisescens* and *Ectropis obliqua*

Scaffold 名称 Scaffold name	上游引物 (5'-3') Forward primer (5'-3')	下游引物 (5'-3') Reverse primer (5'-3')	产物长度 (bp) Product length (bp)
EG793541*	CTCTGCTGGCTACTAACAACCTGG	TTGACGAAACTGATGCTGGAA	376
EG73659*	CTATCCTTCTGTTGGCTCTTGG	TTGGCGAGTAAATCTGCTATGA	911
EO137625*	CAGCCTATTGTGATAATGGGAGA	GGATTGGGTAAAGAGGGTGGT	398
EO336287*	CCCGTCTCAGTCGTAAGTCAGT	CAAGTCTTTCCAAGCCTCCAGT	413
EO779467*	TTACAGTTGACCGCCATCCTC	GCAAGACCAAACGCTTACCTC	310
EG495723	TTCGGCACTTTTCAGCACTCT	TTACTGGAAGCAGGGCACAT	195
EG795495	ATCCCTCCAACCTCCACGAC	GGCAGCGATTGAGAACACCT	220
EG232423	TGTAAATCAAGAAGAGGGTGGG	ATCTACGGAGGATGCTGTGGT	302
EG2544866	CTCCGAGTCGCATAATCAATAGC	TCTTAGTTTCAGCCCACAGTCTTC	438
EG829377	ACACCGACCGAACAAAGACA	TTATTGGCGATAGAAGGAGACG	509
EO530761	CCCCTCTTCTACTTGTGGCTTGT	ATGATGAGTCACCATTGCCTTAC	325
EO889323	GGCATTGTCTCGGATAAAGTGATT	CAGTGTCTCCTCGGGTGTGT	357
EO736813	TGGTGATGAGGGTATGAACGG	GGGAAGTGAAGTAAAGTATGTGGG	492
EO914419	ATTGTCTTACACTGTCATAGCCCACT	CGAGGTCATCCGTCATCTGC	539
EO807593	TGGCTCGGGACACTTATGCT	GTCGCTAATCCAGGGACAAAC	548

* 表示用特异性引物验证成功的物种特异性基因。

* represents species-specific genes successfully verified with the specific primers.

2 结果与分析

2.1 流式细胞术估计基因组大小

利用流式细胞术测定灰茶尺蠖和茶尺蠖的基因组大小,发现灰茶尺蠖雄虫的基因组大小为 (781.34 ± 98.88) Mb (图 1: A),雌虫为 (766.92 ± 77.29) Mb (图 1: B);茶尺蠖雄虫的基因组大小为 (779.87 ± 168.19) Mb (图 1: C),雌虫为 (735.77 ± 135.69) Mb (图 1: D)。与茶尺蠖基因组大小相比,灰茶尺蠖的略大 (Student's *t*-test, *P* = 0.42),并且雄虫基因组大小大于雌虫个体的 (Student's *t*-test, 灰茶尺蠖 *P* = 0.44, 茶尺蠖 *P* = 0.35)。

2.2 基因组测序和从头组装

利用 Illumina 二代测序,分别获得了 68.72 Gb 灰茶尺蠖的原始数据和 65.27 Gb 茶尺蠖的原始数据。过滤掉低质量和重复的 reads 后,灰茶

尺蠖和茶尺蠖分别保留了 61.84 Gb 和 61.83 Gb 的数据,测序深度约为 90X (表 2)。利用 SOAPdenovo2 (K-mer = 41) 进行基因组从头组装,发现灰茶尺蠖和茶尺蠖的基因组中 GC 含量分别为 37.41%和 37.38%,杂合度分别为 1.11%和 1.15%,灰茶尺蠖的 contig 和 scaffold 的 N50 长度分别为 518 bp 和 976 bp,茶尺蠖为 583 bp 和 1 009 bp。基因组 K-mer 频率分析 (K-mer = 17) 表明,灰茶尺蠖 (图 2: A) 和茶尺蠖 (图 2: B) 基因组组成相似,组装后的基因组大小分别为 743.31 Mb 和 699.07 Mb (图 2,表 2),与流式细胞术估算的基因组大小一致。

2.3 基因组注释

通过改良的 MAKER 注释流程,灰茶尺蠖和茶尺蠖基因组官方基因数据集中 (OGS) 分别注释了 12 021 个基因和 12 540 个基因 (表 2)。使用 RepeatModeler 和 RepeatMasker 进行注释,从灰茶尺蠖基因组中注释了 3 285 675 个重复序

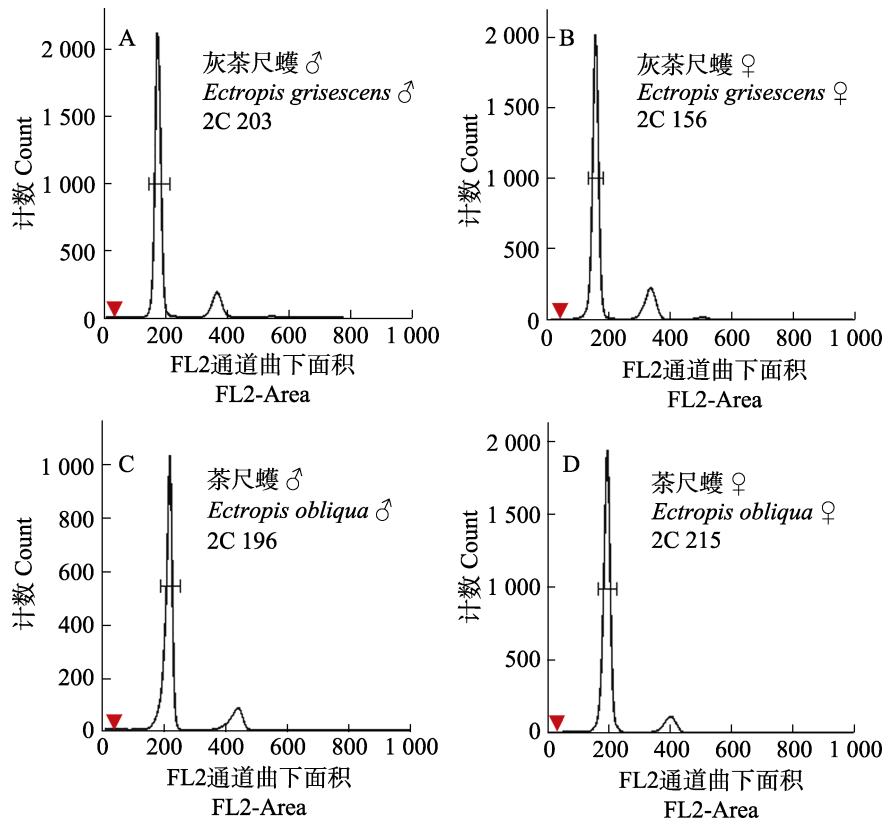


图 1 灰茶尺蠖和茶尺蠖基因组大小

Fig. 1 Genome size of two sibling species of tea geometrids

- A. 雄性灰茶尺蠖 (781.34 Mb, 2C 峰值为 203); B. 雌性灰茶尺蠖 (766.92 Mb, 2C 峰值为 156);
 C. 雄性茶尺蠖 (779.87 Mb, 2C 峰值为 196); D. 雌性茶尺蠖 (735.77 Mb, 2C 峰值为 215).
 A. *E. grisescens* males (781.34 Mb, 2C peak is 203); B. *E. grisescens* females (766.92 Mb, 2C peak is 156);
 C. *E. obliqua* males (779.87 Mb, 2C peak is 196); D. *E. obliqua* females (735.77 Mb, 2C peak is 215).

红色倒三角代表果蝇, 基因组大小为 176.40 Mb, 2C 峰值为 42.8。

The red arrow indicates the *D. melanogaster* (176.40 Mb, 2C peak is 42.8).

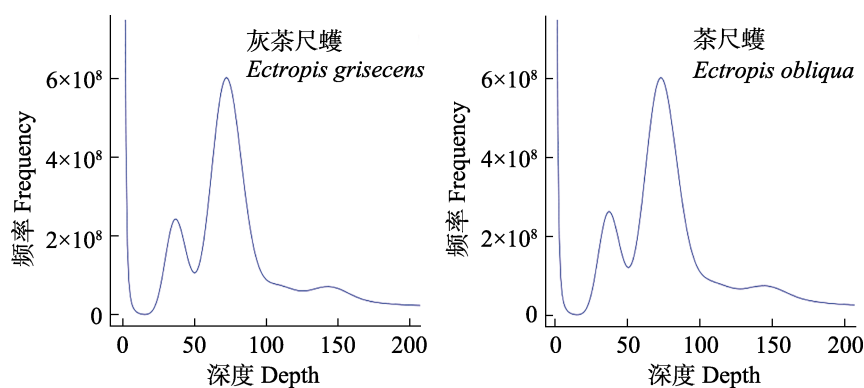


图 2 两种茶尺蠖基因组 K-mer 频率分布图

Fig. 2 Distribution curve of K-mer frequencies in the genome of two sibling species of tea geometrids

- A. 灰茶尺蠖 K-mer 频率分布图; B. 茶尺蠖 K-mer 频率分布图。

A. Distribution curve of K-mer in the genome of *Ectropis grisescens*; B. Distribution curve of K-mer in the genome of *Ectropis obliqua*.

表 2 灰茶尺蠖和茶尺蠖基因组组装
Table 2 Characters summary for the draft genome of *Ectropis grisescens* and *Ectropis obliqua*

种类 Features	测序数据 Sequence statistic			组装数据 Assembly statistic			测序质控 Quality control			基因组特征 Genomic features			基因注释 Gene annotation					
	双端原始数据 (bp)	原始数据 (bp)	Raw base (bp)	Contigs		Scaffolds		Contig		Scaffold		基因组大小 (Mb)	重复序列占比 (%)	杂合度 (%)	G + C含量 (%)	基因数量	GO 功能分类 基因数	KEGG 注释 基因数
				总长度 (bp)	Total length of contigs (bp)	总长度 (bp)	Total length of scaffolds (bp)	N50 长度 (bp)	N50 length of contigs (bp)	N50 长度 (bp)	N50 length of scaffolds (bp)							
灰茶尺蠖 <i>Ectropis grisea</i>	229 099 582	68 729 874 600	680 274 981	2 149 146	753 305 538	1 475 144	518	976	743.31	52.62	1.11	37.41	12 021	7 634	5 472			
茶尺蠖 <i>Ectropis obliqua</i>	217 581 555	65 274 466 500	678 881 363	1 972 958	735 614 840	1 427 772	583	1 009	699.07	49.67	1.15	37.38	12 540	7 787	5 457			

列, 大小约为 270.96 Mb, 占灰茶尺蠖基因组的 35.96%; 茶尺蠖基因组中注释到 2 970 623 个重复序列, 大小约为 262.64 Mb, 占茶尺蠖基因组的 34.79%。在二者当中, 长散在核元件 (LINE) 占 4.45%, 是重复序列当中占比最高的元件类型。同时灰茶尺蠖的长末端重复元件 (LTR) 的重含量是茶尺蠖的 6 倍, DNA 转座子数量是茶尺蠖的 2 倍 (表 3)。

使用 BLASTP 对 UniProt 和 Swiss-Prot 数据库对蛋白质编码基因进行注释, 并通过 GO 功能

分类和 Kyoto Encyclopedia of Genes and Genomes (KEGG) web 服务器基因功能分类。结果表明, 灰茶尺蠖和茶尺蠖分别有 7 634 个基因 (63.51%) 和 7 787 个基因 (62.10%) 被分配到至少 1 个 GO 功能分类, 同时灰茶尺蠖和茶尺蠖分别有 5 472 个基因 (45.52%) 和 5 457 个基因 (43.52%) 在 KEGG 通路上被找到 (表 4)。其中茶尺蠖基因中有 28 基因显著富集于雌激素信号通路 (KO: 04915), 灰茶尺蠖基因中有 20 基因显著富集于谷胱甘肽代谢通路 (KO: 04915)。

表 3 灰茶尺蠖和茶尺蠖基因组重复序列统计

Table 3 Summary for the repeat annotation of *Ectropis grisescens* and *Ectropis obliqua*

重复序列类型 Repeats type	灰茶尺蠖 <i>Ectropis grisescens</i>			茶尺蠖 <i>Ectropis obliqua</i>		
	数量 Number	总长 (bp) Total length (bp)	序列占比 (%) Sequence percentage (%)	数量 Number	总长 (bp) Total length (bp)	序列占比 (%) Sequence percentage (%)
短散在核元件 SINEs	18	1 093	0.00	23	1 070	0.00
长散在核元件 LINEs	308 050	33 545 764	4.45	280 662	32 644 029	4.43
长末端重复元件 LTR elements	1 523	259 408	0.03	220	42 761	0.01
DNA 转座子 DNA elements	14 119	1 668 744	0.22	8 073	1 082 158	0.15
未定义 Unclassified	2 821 053	235 487 366	31.26	2 532 000	222 122 119	30.20
小 RNA Small RNA	680	42 748	0.01	878	59 748	0.01
微卫星 Satellites	3	436	0.00	1	216	0.00
简单重复序列 Simple repeats	123 329	5 442 276	0.72	130 555	5 836 238	0.79
低复杂度序列 Low complexity	16 900	781 438	0.10	18 211	856 541	0.12
总计 Total	3 285 675	270 962 375	35.96	2 970 623	262 644 880	34.79

表 4 灰茶尺蠖和茶尺蠖 KEGG 通路

Table 4 Summary for the KEGG pathway of *Ectropis grisescens* and *Ectropis obliqua*

功能描述 Description	通路 ID Pathway ID	灰茶尺蠖 <i>Ectropis grisescens</i>	茶尺蠖 <i>Ectropis obliqua</i>
外源生物代谢-细胞色素 P450d Metabolism of xenobiotics by cytochrome P450	KO: 00980	8	6
药物代谢-细胞色素 P450 Drug metabolism-cytochrome P450	KO: 00982	5	4
谷胱甘肽代谢 Glutathione metabolism	KO: 00480	20	16
昆虫激素生物合成 Insect hormone biosynthesis	KO: 00981	8	11
雌激素信号通路 Estrogen signaling pathway	KO: 04915	17	28
嗅觉传导 Olfactory transduction	KO: 04740	10	10
味觉传导 Taste transduction	KO: 04742	8	7

2.4 系统发育以及直系同源分析

选取茶尺蠖、灰茶尺蠖、冬尺蛾、家蚕、大白斑蝶、诗神袖蝶和小菜蛾 7 种鳞翅目昆虫以及外群黑腹果蝇的蛋白质序列进行同源分析,共鉴定出所有物种的 15 332 个同源基因和 1 018 个单拷贝同源基因,利用最大似然法进行系统发育树重建,并通过 OrthoMCL 程序计算物种的分化时间节点。结果表明,灰茶尺蠖和茶尺蠖与冬尺蛾亲缘关系密切,在 4 969 万年前开始分化,而灰茶尺蠖和茶尺蠖大约在 318 万年前才开始分化(图 3: A)。对家蚕、冬尺蛾、灰茶尺蠖和茶尺蠖直系同源类群的同源性分析显示,灰茶尺蠖和茶尺蠖共有 6 414 个同源基因,其中二者分别有 243 个和 284 个物种特异基因(图 3: B)。此外,灰茶尺蠖与茶尺蠖的氨基酸同源性比其它 2 个物种都要高(图 3: C)。使用 LASTZ 算法对基因组 scaffold 序列进行了共线性分析,发现两个物种的分化过程中没有发生明显的基因组重排事件,表明染色体重排不活跃(图 3: D)。

为进一步了解灰茶尺蠖和茶尺蠖基因组的全基因组适应性进化,以冬尺蛾和家蚕为外群,使用 CODEML free-ratio 模型,计算茶尺蠖 1 018 个单拷贝同源基因序列的非同义替换(K_a)与同义替换(K_s)比值。灰茶尺蠖 K_a / K_s 平均为 0.132 0 ($SD = 0.436 4$),茶尺蠖 K_a / K_s 平均为 0.120 5 ($SD = 0.384 5$)。灰茶尺蠖正选择基因数($K_a / K_s > 1$)为 25 个,茶尺蠖正选择基因数为 23 个。对 GO 注释集进行双样本 t -检验,鉴定具有 5 个以上同源性的高度分化基因家族,发现信号转导活性(GO: 000 487 1)(t -test; $P = 0.021 0$)和半胱氨酸型肽酶活性(GO: 000 823 4)(t -test; $P = 0.040 4$) 在灰茶尺蠖和茶尺蠖之间存在差异显著。

2.5 参与 EoNPV 抗性的抗病毒基因分析

为分析灰茶尺蠖和茶尺蠖对 EoNPV 的敏感性差异,收集核多角体病毒与鳞翅目寄主相互作用的抗病毒基因和信号通路。使用 BLASTP 和 tBLASTn 对近缘种灰茶尺蠖和茶尺蠖的基因组进行了比对。结果表明,在灰茶尺蠖中,与昆虫免疫系统相关的载脂蛋白基因有 5 个拷贝,而在

茶尺蠖中,只注释得到 1 个载脂蛋白基因。同时预测到在灰茶尺蠖中病毒组织蛋白酶基因的 7 个拷贝,而在茶尺蠖中只有 4 个拷贝,这表明与病毒感染相关的病毒组织蛋白酶基因可能在灰茶尺蠖中扩增(表 5)。此外,由于脂肪酸代谢途径与病毒抗性有关,分析显示灰茶尺蠖和茶尺蠖的小分子代谢过程(GO: 004 428 1)的 K_a/K_s 值分别为 0.046 9 ($SD = 0.092 2$) 和 0.005 5 ($SD = 0.014 9$),具有显著差异。灰茶尺蠖中载脂蛋白和组织蛋白酶基因的扩增以及小分子代谢过程通路的显著正向选择可能与灰茶尺蠖对 EoNPV 敏感性的差异有关。

2.6 化学受体基因分析

在亲缘关系密切的果蝇物种中,化学受体通路分化促进了种间求偶的抑制。为研究化学感受相关基因是否参与了非对称生殖干扰,分别对尺蠖的气味结合蛋白、化学感受蛋白、嗅觉受体和味觉受体进行鉴定和系统发育分析。结果表明,灰茶尺蠖中只有 1 个性信息素结合蛋白亚家族基因(Pheromone-binding, PBP),而茶尺蠖存在 4 个成员,其他化学受体基因家族并没有检测到显著扩增事件,表明性信息素结合蛋白亚家族的增加可能与非对称生殖干扰现象有关(图 4, 表 6)。

2.7 物种特异性基因和引物验证

为了准确区分茶尺蠖与灰茶尺蠖,本研究在获得的物种特异性基因中随机抽取 16 个基因,设计物种特异性引物并评估其识别两个近缘物种的效率(表 1)。结果显示,在两个近缘种中共有 5 对特异性引物,对应的 Scaffold 分别为 EG793541、EG79359、EO137625、EO336287 和 EO779467,可以用于物种鉴定(图 5)。

3 讨论

灰茶尺蠖与茶尺蠖在形态特征和生物学上都难以区分,生产上长期将两者作为同一个物种进行防治,导致了生物防治效果具有明显的差异。本研究利用比较基因组学手段对两种茶

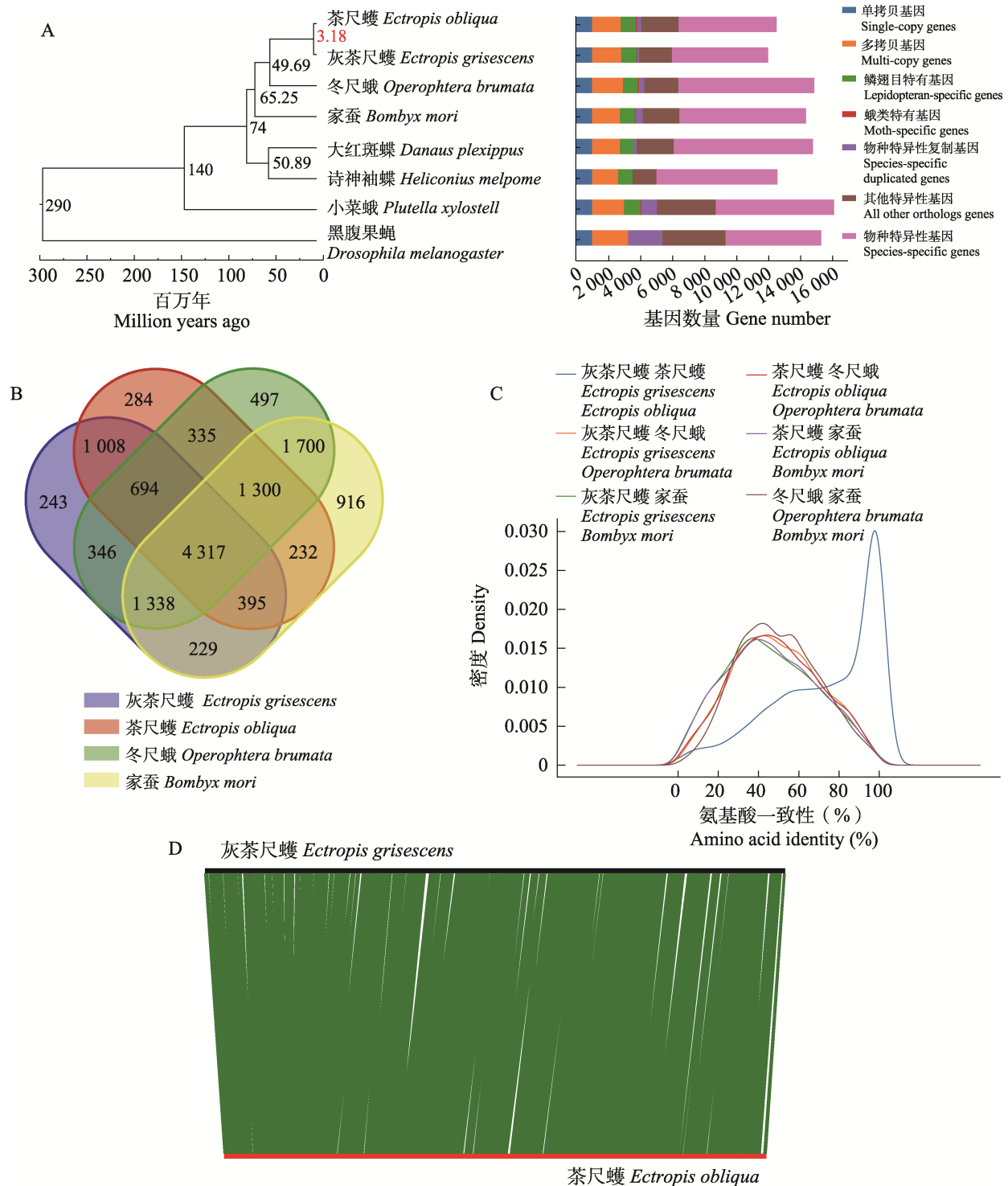


图 3 两种茶尺蠖系统发育和直系同源分析

Fig. 3 Phylogeny comparison and orthologous of *Ectropis grisescens* and *Ectropis obliqua*

A. 8 种昆虫系统发育树及其比较分析; B. 冬尺蛾、茶尺蠖和灰茶尺蠖同源类群分布; C. 灰茶尺蠖、茶尺蠖、冬尺蛾和家蚕的氨基酸同源性分布; D. 灰茶尺蠖 (黑色) 和茶尺蠖 (红色) 基因组共线性分析。

A. Phylogenetic tree and comparative analysis of eight insect species; B. Distribution of orthologous groups of *O. brumata*, *E. grisescens* and *E. obliqua*; C. The distribution of pairwise amino acid identity among *E. grisescens*, *E. obliqua*, *O. brumata* and *B. mori*; D. Collinear analysis of the *E. grisescens* (black) and *E. obliqua* (red) genomes.

表 5 两种茶尺蠖的病毒抗性相关基因

Table 5 Summary for the virus-resistance related genes of *Ectropis grisescens* and *Ectropis obliqua*

功能 Function	注释结果 Annotation	灰茶尺蠖 <i>Ectropis grisescens</i>	茶尺蠖 <i>Ectropis obliqua</i>
信号通路相关 Signal pathway correlation	磷酸肌醇 3-激酶 Phosphoinositide 3-kinase	3	5
	酯酶 E4 Esterase E4	3	6
昆虫免疫相关 Insect immune related	酯酶 B1 Esterase B1	4	2
	载脂蛋白 Apolipoporphins	5	1
病毒感染相关 Virus infection related	病毒组织蛋白酶 Viral cathepsin	7	4

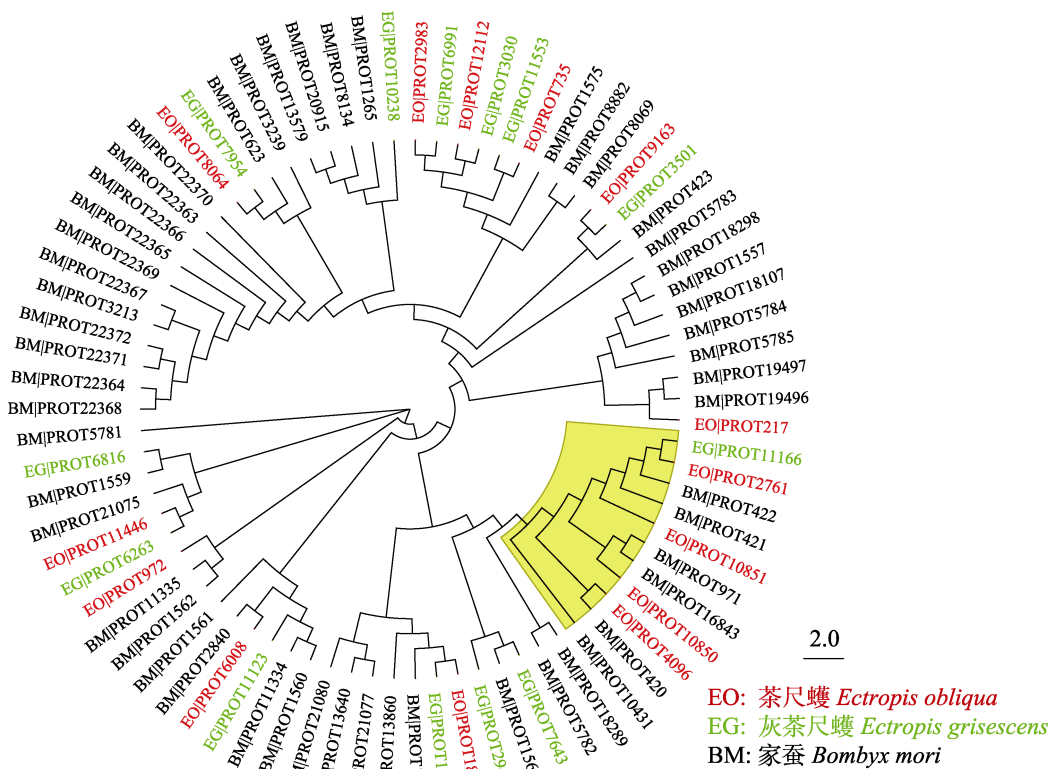


图 4 家蚕、灰茶尺蠖和茶尺蠖气味结合蛋白系统发育树

Fig. 4 Phylogenetic tree of the OBP gene family of *Bombyx mori*, *Ectropis grisescens* and *Ectropis obliqua*

尺蠖进行分析,有助于理解灰茶尺蠖与茶尺蠖的物种分化和进化过程,为理解近源种适应性进化提供新的见解。

通过流式细胞术和 K-mer 分析发现,灰茶尺蠖的基因组大小比茶尺蠖大 30-50 Mb,近缘种之间没有明显的性别和种间偏向。通过基因组测序对灰茶尺蠖和茶尺蠖的基因组草图进行了组

装和注释。结果显示,灰茶尺蠖中已知重复序列的数量和比例略高于茶尺蠖(表 3),这可能是灰茶尺蠖基因组大小扩增的驱动因素。系统发育和共线性分析显示灰茶尺蠖和茶尺蠖在 318 万年前分化,二者基因组组成高度相似,但未发生明显的基因组重排事件(图 3: D)。虽然二者在氨基酸水平和染色体水平上类似,但在同源类群

进化分析中存在差异, 茶尺蠖的种特异性基因多于灰茶尺蠖 (图 3: B)。

表 6 灰茶尺蠖和茶尺蠖化学受体相关基因家族分析结果

Table 6 Summary for the chemosensory related gene family analysis of *Ectropis grisescens* and *Ectropis obliqua*

基因家族 Gene family	灰茶尺蠖 <i>Ectropis grisescens</i>	茶尺蠖 <i>Ectropis obliqua</i>
气味结合蛋白 Odorant-binding protein	13	14
嗅觉受体 Odorant receptor	26	30
味觉受体 Gustatory receptor	15	17
化学感受蛋白 Chemosensory protein	14	13

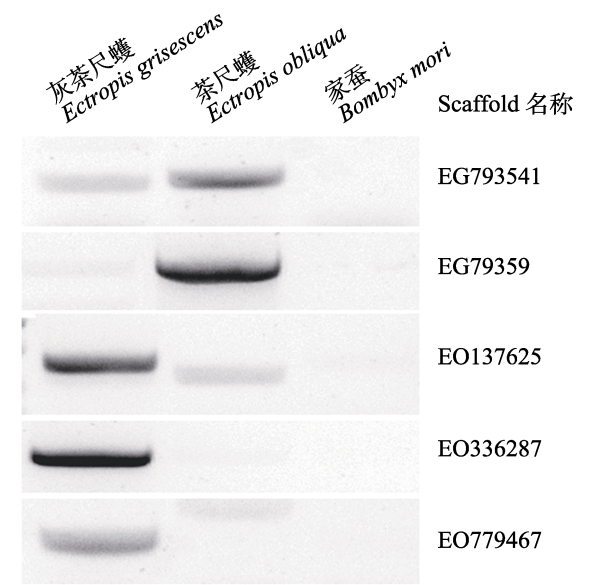


图 5 物种特异性引物 PCR 和凝胶电泳分析结果

Fig. 5 PCR and gel electrophoresis analysis of species specific primers

右侧为筛选出的 Scaffold 序列名称, EG 表示灰茶尺蠖 *Ectropis grisescens*, EO 表示 *Ectropis obliqua*。
The selected Scaffold name is shown on the right, EG represents *Ectropis grisescens*, EO represents *Ectropis obliqua*.

此前, 抗菌素肽基因 (如 *gloverin*、*moricin* 和 *attacin*) 的差异表达可能导致两个近缘种对病毒感染的敏感性产生差异 (毛腾飞等, 2017)。

通过对二者抗病毒基因的比较发现, 灰茶尺蠖有 5 个与昆虫免疫系统有关的载脂蛋白基因拷贝, 而在茶尺蠖基因组中只有 1 个载脂蛋白拷贝。在灰茶尺蠖中预测到了 7 个病毒组织蛋白酶基因, 而在茶尺蠖中预测到了 4 个病毒组织蛋白酶基因。此外, 灰茶尺蠖与抗病毒抗性相关的脂肪酸代谢途径, 如小分子代谢过程 (GO:0044281) 与茶尺蠖相比具有显著的正向选择。研究发现, 载脂蛋白可能对表面角质层脂质运输和沉积以及维持正常角质层屏障功能至关重要 (Zhao *et al.*, 2020), 并且病毒组织蛋白酶是 *EoNPV* 的作用靶标。因此, 推测载脂蛋白和病毒组织蛋白酶基因的扩增以及小分子代谢过程 (GO:0044281) 相关基因的正向进化可能是灰茶尺蠖对 *EoNPV* 产生抗性的原因。

化学感受蛋白 (CSPs) 在昆虫的化学感受器官表达, 调控昆虫繁殖、交配及取食等一系列行为 (Pelosi *et al.*, 2006)。近源种之间化学感受蛋白的表达含量可能有较大差异, 现已证实 CSP1 在烟粉虱 B 生物型和 Q 生物型中的表达量不同 (Liu *et al.*, 2014a)。本研究发现茶尺蠖的气味结合蛋白、嗅觉受体和味觉受体基因数量高于灰茶尺蠖, 表明茶尺蠖的生殖和识别系统比灰茶尺蠖更为复杂。同时, 茶尺蠖信息素结合蛋白亚家族中气味结合蛋白存在扩增事件, 茶尺蠖的性激素合成基因和信息素结合亚家族基因的数量均大于灰茶尺蠖, 这些基因家族的扩增可能为茶尺蠖提供了一个更为复杂的识别系统, 使茶尺蠖在生殖识别上更具特异性, 从而导致两种物种的非对称生殖干扰现象。

本研究基于物种特异的单拷贝同源类群设计了能够有效识别灰茶尺蠖和茶尺蠖物种特异性引物, 能够为茶树害虫的分类和特异防治提供实用的工具和方法。综上所述, 本研究基于比较基因组学方法, 揭示了灰茶尺蠖和茶尺蠖基因组方面的差异, 为茶园防治灰茶尺蠖和茶尺蠖提供了科学依据。

致谢: 感谢中国农业科学院茶叶研究所李兆群副研究员在试虫收集过程中提供的帮助。

参考文献 (References)

- Bai JH, Wang ZB, Xiao Q, 2018. Genetic differentiation and distribution of two sibling species of tea geometrids in tea-growing areas in Zhejiang, eastern China. *Acta Entomologica Sinica*, 61(6): 741–748. [白家赫, 王志博, 肖强, 2018. 浙江茶区茶尺蠖两近缘种的遗传分化及分布. *昆虫学报*, 61(6): 741–748.]
- Bao Z, Eddy SR, 2002. Automated de novo identification of repeat sequence families in sequenced genomes. *Genome Research*, 12(8): 1269–1276.
- Bargielowski I, Blosser E, Lounibos LP, 2015. The effects of interspecific courtship on the mating success of *Aedes aegypti* and *Aedes albopictus* (Diptera: Culicidae) males. *Annals of the Entomological Society of America*, 108(4): 513–518.
- Benson G, 1999. Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Research*, 27(2): 573–580.
- Bolger AM, Marc L, Bjoern U, 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15): 2114–2120.
- Buiate E, Xavier KV, Moore N, Torres MF, Farman ML, Schardl CL, Vaillancourt LJ, 2017. A comparative genomic analysis of putative pathogenicity genes in the host-specific sibling species *Colletotrichum graminicola* and *Colletotrichum sublineola*. *BMC Genomics*, 18(1): 67.
- Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T, 2009. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25(15): 1972–1973.
- Crowder DW, Sitvarin MI, Carriere Y, 2010. Plasticity in mating behaviour drives asymmetric reproductive interference in whiteflies. *Animal Behaviour*, 79(3): 579–587.
- Darriba D, Taboada GL, Doallo R, Posada D, 2011. ProtTest 3: Fast selection of best-fit models of protein evolution. *Bioinformatics*, 27(8): 1164–1165.
- Dolezel J, Greilhuber J, Suda J, 2007. Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols*, 2(9): 2233–2244.
- Doyle JJ, 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, 19: 11–15.
- El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, Sonnhammer E, Hirsh L, Paladin L, Piovesan D, Tosatto S, Finn RD, 2019. The Pfam protein families database in 2019. *Nucleic Acids Research*, 47(D1): D427–D432.
- Gebiola M, Kelly SE, Velten L, Zug R, Hammerstein P, Giorgini M, Hunter MS, 2017. Reproductive interference and fecundity affect competitive interactions of sibling species with low mating barriers: Experimental and theoretical evidence. *Heredity*, 119(6): 438–446.
- Groning J, Hochkirch A, 2008. Reproductive interference between animal species. *Quarterly Review of Biology*, 83(3): 257–282.
- Harris RS, 2007. Improved pairwise alignment of genomic DNA. Doctoral dissertation. Ann Arbor: The Pennsylvania State University.
- Hazarika LK, Bhuyan M, Hazarika BN, 2008. Insect pests of tea and their management. *Annual Review of Entomology*, 54(1): 267–284.
- Hazarika LK, Puzari KC, Wahab S, 2001. Biological Control of Tea Pests. Boston: Springer US. 159–180.
- He K, Lin K, Wang G, Fei L, 2016. Genome sizes of nine insect species determined by flow cytometry and k-mer analysis. *Frontiers in Physiology*, 7: 569.
- Holt C, Yandell M, 2011. MAKER2: An annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics*, 12(1): 491.
- Ian K, 2004. Gene finding in novel genomes. *BMC Bioinformatics*, 5(1): 59.
- Jiang L, Xia Q, 2014. The progress and future of enhancing antiviral capacity by transgenic technology in the silkworm *Bombyx mori*. *Insect Biochemistry and Molecular Biology*, 48(1): 1–7.
- Jiang N, Liu SX, Xue DY, Tang MJ, Xiao Q, Han HX, 2014. External morphology and molecular identification of two tea geometrid moth from southern China. *Chinese Journal of Applied Entomology*, 51(4): 987–1002. [姜楠, 刘淑仙, 薛大勇, 唐美君, 肖强, 韩红香, 2014. 我国华东地区两种茶尺蠖的形态和分子鉴定. *应用昆虫学报*, 51(4): 987–1002.]
- Jing S, Zheng B, Qiao L, Zhang K, Chen G, Guo S, Huang X, Gao Y, Zhang L, Yuan H, 2016. Characterization of eleven microsatellite markers for the tea geometrid *Ectropis obliqua* (Lepidoptera: Geometridae). *Applied Entomology and Zoology*, 51(4): 1–4.
- Kanehisa M, Sato Y, Morishima K, 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *Journal of Molecular Biology*, 428(4): 726–731.
- Katoh K, Standley DM, 2013. MAFFT multiple sequence alignment

- software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30(4): 772–780.
- Kingsford C, 2011. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics*, 27(6): 764.
- Kong M, Zuo H, Zhu F, Hu Z, Chen L, Yang Y, Lv P, Yao Q, Chen K, 2018. The interaction between baculoviruses and their insect hosts. *Developmental & Comparative Immunology*, 83: 114–123.
- Li L, Stoeckert CJ, Roos DS, 2003. OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Research*, 13(9): 2178–2189.
- Liu GX, Xuan N, Chu D, Xie HY, Fan ZX, Bi YP, Picimbon JF, Qin YC, Zhong ST, Li YF, Gao ZL, Pan WL, Wang GY, Rajashekar B, 2014a. Biotype expression and insecticide response of *Bemisia tabaci* chemosensory protein-1. *Archives of Insect Biochemistry and Physiology*, 85(3): 137–151.
- Liu JD, Xiao HM, Huang SQ, Li F, 2014. OMIGA: Optimized maker-based insect genome annotation. *Molecular Genetics and Genomics*, 289(4): 567–573.
- Lü P, Pan Y, Yang Y, Zhu F, Li C, Yao Q, Chen K, 2018. Discovery of anti-viral molecules and their vital functions in *Bombyx mori*. *Journal of Invertebrate Pathology*, 154: 12–18.
- Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung DW, Yiu SM, Peng S, Xiaoqian Z, Liu G, Liao X, Li Y, Yang H, Wang J, Lam TW, Wang J, 2012. SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. *Gigascience*, 1(1): 18.
- Luo ZX, Li ZQ, Cai XM, Bian L, Chen ZM, 2017. Evidence of premating isolation between two sibling moths: *Ectropis grisescens* and *Ectropis obliqua* (Lepidoptera: Geometridae). *Journal of Economic Entomology*, 110(6): 2364–2370.
- Mao TF, Fu JY, Sun L, Zhou XG, Bai JH, Xiao Q, 2017. The expression of the antibacterial peptide genes from two sibling species of tea geometrid was different in resistance to EoNPV infection. *Chinese Journal of Biological Control*, 33(4): 472–480.
- [毛腾飞, 付建玉, 孙亮, 周孝贵, 白家赫, 肖强, 2017. 两近缘种茶尺蠖的抗菌肽基因在抵抗 EoNPV 感染过程中的表达差异. 中国生物防治学报, 33(4): 472–480.]
- Mario S, Oliver K, Irfan G, Alec H, Stephan W, Burkhard M, 2006. AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Research*, 34(Web server issue): W435–W439.
- Massoumi AS, Haridas S, Feau N, Robertson G, Bohlmann J, Breuil C, 2014. Comparative genomics of the pine pathogens and beetle symbionts in the genus *Grossmannia*. *Molecular Biology and Evolution*, 31(6): 1454–1474.
- Nguyen Q, Nielsen L, Reid S, 2013. Genome scale transcriptomics of baculovirus-insect interactions. *Viruses*, 5(11): 2721–2747.
- Noriyuki S, Osawa N, Nishida T, 2012. Asymmetric reproductive interference between specialist and generalist predatory ladybirds. *Journal of Animal Ecology*, 81(5): 1077–1085.
- Olson SA, 2002. Emboss opens up sequence analysis. *Briefings in Bioinformatics*, 3(1): 87–91.
- Pan Y, Fang G, Wang Z, Cao Y, Liu Y, Li G, Liu X, Xiao Q, Zhan S, 2021. Chromosome-level genome reference and genome editing of the tea geometrid. *Molecular Ecology Resources*, 21(6): 2034–2049.
- Pelosi P, Zhou JJ, Ban LP, Calvello M, 2006. Soluble proteins in insect chemical communication. *Cellular and Molecular Life Sciences*, 63(14): 1658–1676.
- Porebski S, Bailey LG, Baum BR, 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Molecular Biology Reporter*, 15(1): 8–15.
- Price AL, Jones NC, Pevzner PA, 2005. De novo identification of repeat families in large genomes. *Bioinformatics*, 21(Suppl.1): i351–i358.
- Sanderson MJ, 2003. r8s: Inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics*, 19(2): 301–302.
- Sillo F, Garbelotto M, Friedman M, Gonthier P, 2015. Comparative genomics of sibling fungal pathogenic taxa identifies adaptive evolution without divergence in pathogenicity genes or genomic structure. *Genome Biology & Evolution*, 7(12): 3190–3206.
- Sinu PA, Mandal P, Banerjee D, Mallick S, Pathak SK, 2013. Moth pests collected in light traps of tea plantations in North East India: Species composition, seasonality and effect of habitat type. *Current Science*, 104(5): 646–651.
- Stamatakis A, 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9): 1312–1313.
- Tang MJ, Guo HW, Ge CM, Yin KS, Xiao Q, 2017. Pathogenic characters of *Ectropis obliqua* nucleopolyhedroviruses on *Ectropis grisescens* Warren and screening of high efficient strain.

- Acta Agriculturae Zhejiangensis*, 29(10): 1686–1691. [唐美君, 郭华伟, 葛超美, 殷坤山, 肖强, 2017. *EoNPV* 对灰茶尺蠖的致病特性及高效毒株筛选. 浙江农业学报, 29(10): 1686–1691.]
- Vicente VA, Weiss VA, Bombassaro A, Moreno LF, Costa FF, Raittz RT, Leão AC, Gomes RR, Bocca AL, Fornari G, de Castro RJA, Sun J, Faoro H, Tadra-Sfeir MZ, Baura V, Balsanelli E, Almeida SR, Dos Santos SS, Teixeira MDM, Soares Felipe MS, Do Nascimento MMF, Pedrosa FO, Steffens MB, Attili-Angelis D, Najafzadeh MJ, Queiroz-Telles F, Souza EM, De Hoog S, 2017. Comparative genomics of sibling species of *fonsecaea* associated with human chromoblastomycosis. *Frontiers in Microbiology*, 8: 1924.
- Wang Z, Bai J, Liu Y, Li H, Zhan S, Xiao Q, 2019. Transcriptomic analysis reveals insect hormone biosynthesis pathway involved in desynchronized development phenomenon in hybridized sibling species of tea geometrids (*Ectropis grisescens* and *Ectropis obliqua*). *Insects*, 10(11): 381.
- Wang Z, Li H, Zhou X, Tang M, Xiao Q, 2020. Comparative characterization of microbiota between the sibling species of tea geometrid moth *Ectropis obliqua* Prout and *E. grisescens* Warren. *Bulletin of Entomological Research*, 110(6): 1–10.
- Yang Z, 2007. PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, 24(8): 1586–1591.
- Ye GY, Xiao Q, Chen M, Chen XX, Yuan ZJ, Stanley DW, Hu C, 2014. Tea: Biological control of insect and mite pests in China. *Biological Control*, 68: 73–91.
- Yin KS, Chen HC, Xiao Q, Tang MJ, 2000. Widespreading and application of *Ectropic obliqua* nuclear polyhedrosis virus (NPV) preparations. *Virologica Sinica*, (S1): 84–87. [殷坤山, 陈华才, 肖强, 唐美君, 2000. 茶尺蠖核型多角体病毒制剂的试制与推广应用. 中国病毒学, (S1): 84–87.]
- Yu XI, Yin KS, 2011. The susceptibility difference against *EoNPV* in different geographic populations of tea geometrid (*Ectropis obliqua* Prout). *Journal of Tea Science*, 31(2): 100–104. [席羽, 殷坤山, 肖强, 2011. 不同地理种群茶尺蠖对 *EoNPV* 的敏感性差异研究. 茶叶科学, 31(2): 100–104.]
- Zhang GH, Yuan ZJ, Zhang CX, Yin KS, Tang MJ, Guo HW, Fu JY, Xiao Q, 2014. Detecting deep divergence in seventeen populations of tea geometrid (*Ectropis obliqua* Prout) in China by COI mtDNA and cross-breeding. *PLoS ONE*, 9(6): e99373.
- Zhang GH, Yuan ZJ, Yin KS, Fu JY, Tang MJ, Xiao Q, 2016. Asymmetrical reproductive interference between two sibling species of tea looper: *Ectropis grisescens* and *Ectropis obliqua*. *Bulletin of Entomological Research*, 11: 1–8.
- Zhao YY, Liu WM, Zhao XM, Yu ZT, Guo HF, Yang Y, Zhang JQ, Moussian B, Zhang JZ, 2020. Apolipoprotein-II/I contributes to cuticular hydrocarbon transport and cuticle barrier construction in *Locusta migratoria*. *Frontiers in Physiology*, 11: 790.