



叶昕海, 杨义, 梅洋, 肖花美, 李飞. 草地贪夜蛾基因组注释及分析 [J]. 环境昆虫学报, 2019, 41 (4): 706–717.

草地贪夜蛾基因组注释及分析

叶昕海¹, 杨 义¹, 梅 洋¹, 肖花美^{1,2*}, 李 飞¹

(1. 浙江大学昆虫科学研究所, 杭州 310058; 2. 宜春学院生命科学与资源环境学院, 江西省作物生长发育调控重点实验室, 江西宜春 336000)

摘要: 草地贪夜蛾 *Spodoptera frugiperda* 近年来在我国迅速扩散, 造成了重大的经济损失, 引起社会关注。草地贪夜蛾基因组序列对深入研究其迁飞、入侵和抗药性等特性具有十分重要的作用。目前, 已有 5 个版本的基因组序列被公开报道, 但 3 个版本无基因组注释信息。除以 Sf9 细胞系为 DNA 来源的基因组版本外, 其他版本的 scaffold N50 过小, 拼接质量偏低。为此, 本研究选取了 scaffold N50 最大的草地贪夜蛾 Sf9 细胞系基因组进行了蛋白编码基因注释。该版本的基因组重复序列占比 28.1%。CEGMA 评估显示该版本基因组可覆盖 93.6% 的核心基因, BUSCO 评估显示可覆盖 90.8% 的核心基因。利用 OMIGA 注释流程预测到 25 699 个蛋白质编码基因, 详细的基因序列可从 InsectBase 网站获得 (http://www.insect-genome.com/FAW/), 其中具有 GO 注释的基因为 15 623 个, 具有 KEGG 注释的基因共有 9 213 个。选取了 12 个鳞翅目昆虫进行比较基因组学分析, 发现草地贪夜蛾与斜纹夜蛾的亲缘关系最近, 两者分化时间大约在 1 284 万年前。对 12 个鳞翅目昆虫蛋白质编码基因进行同源分析, 在草地贪夜蛾中发现了 2 490 个单拷贝基因、891 个鳞翅目特有基因、2 360 个物种特异扩增基因和 4 180 个物种特异基因。GO 富集分析显示, 2 360 个物种特异扩增基因主要参与 DNA 整合、代谢相关的生物过程; 4 180 个物种特异基因主要参与酶活性、光感受、糖代谢等, KEGG 通路富集发现草地贪夜蛾特异基因主要参与氨基酸代谢、糖代谢和 Wnt 信号通路。本研究结果丰富了草地贪夜蛾的基因信息, 对进一步了解其生物学特性、开发新型绿色防控方法具有指导意义。

关键词: 草地贪夜蛾; 基因组注释; 比较基因组学; 基因家族; 进化

中图分类号: Q963; S433.4

文献标识码: A

文章编号: 1674-0858 (2019) 04-0706-12

The genome annotation and comparative genomics analysis of *Spodoptera frugiperda*

YE Xin-Hai¹, YANG Yi¹, MEI Yang¹, XIAO Hua-Mei^{1,2*}, LI Fei¹ (1. Institute of Insect Sciences, Zhejiang University, Hangzhou 310058, China; 2. College of Life Sciences and Resource Environment / Key Laboratory of Crop Growth and Development Regulation, Jiangxi Province, Yichun University, Yichun 336000, China)

Abstract: The fall armyworm *Spodoptera frugiperda* was first found in China in January, 2019, and has rapidly spread to 20 provinces within half a year. This insect has caused the great economic losses and social concerns all over the world. Genome data is important for studying the molecular mechanisms of many important traits such as long-distance migration, invasive and insecticide resistance of this notorious insect pest. Nowadays, five versions of *S. frugiperda* genome assemblies have been reported, but some genome annotation information was still unavailable. In addition, only the genome assembly obtained from Sf9 cell has a high scaffold N50 of 601.1 Kb. 28.1% of the Sf9 cell-line genome is identified as repeat

基金项目: 国家自然科学基金 (31760514); 江西省自然科学基金青年基金 (20181BAB214012)

作者简介: 叶昕海, 男, 博士研究生, 研究方向为昆虫基因组学与生物信息学

* 通讯作者 Author for correspondence: 肖花美, 博士, 副教授, 主要研究方向为昆虫基因组学与生物信息学, E-mail: xiaohuamei625@163.com

收稿日期 Received: 2019-07-18; 接受日期 Accepted: 2019-07-22

sequences. We assessed its genome space using the software Core Eukaryotic Genes Mapping Approach (CEGMA) and Benchmarking Universal Single-Copy Orthologs (BUSCO), showing that 93.6% CEGMA genes and 90.8% BUSCO genes are successfully recovered. To provide gene information of *S. frugiperda*, we annotated the genome of the Sf9 cell line using the pipeline of Optimized Maker-based Insect Genome Annotation (OMIGA), yielding 25 699 protein-coding genes which can be obtained from <http://www.insect-genome.com/FAW/>. Among which, 15 623 genes have GO terms and 9 213 genes have homology in the KEGG database. Phylogenetic analysis shows *S. frugiperda* diverges from the common ancestor of *S. frugiperda* and *Spodoptera litura* approximately 12.84 million years ago. In addition, we identified 2 490 single-copy genes, 891 Lepidoptera specific genes, 2 360 *S. frugiperda*-specific duplication genes and 4 180 species-specific genes in the *S. frugiperda* genome. GO enrichment analyses show that *S. frugiperda*-specific duplication genes are significantly overrepresented for DNA integration and DNA metabolic process, and species-specific genes are mostly related to enzyme activity, photoreceptor and carbohydrate metabolism. KEGG enrichment analyses show that species-specific genes are highly related to amino acids metabolism, carbohydrate metabolism and Wnt signaling pathway. Our data provide the genomic information of *S. frugiperda*, which should be great helpful for further understanding its biological characteristics and developing novel control methods.

Key words: *Spodoptera frugiperda*; genome annotation; comparative genomics; gene family; evolution

草地贪夜蛾 *Spodoptera frugiperda* (J. E. Smith, 1797), 属于鳞翅目 Lepidoptera 夜蛾科 Noctuidae, 原产于美洲热带和亚热带地区, 广泛分布于美国东部各地, 最北可达加拿大南部 (Todd & Poole, 1980), 是一种重要的迁飞性害虫, 严重威胁农业生产 (Martinelli *et al.*, 2006)。草地贪夜蛾繁殖能力强, 成虫产卵量大, 每头雌虫一生平均可产约 1 500 粒卵, 甚至超过 2 000 粒 (Prasanna *et al.*, 2018)。其扩散速度快, 飞行能力强, 成虫可借风力进行远距离迁飞 (Westbrook *et al.*, 2016), 每晚飞行距离长达 100 km (Abrahams *et al.*, 2017), 若风向风速适宜, 迁飞距离会更长, 最长纪录为 30 h 内飞行 1 600 km (Johnson, 1987)。2016 年 1 月, 在非洲尼日利亚西南部热带雨林区发现该虫入侵, 2017 年 4 月便扩散至非洲 12 个国家, 5 个月, 28 个撒哈拉以南的非洲国家证实有草地贪夜蛾的入侵, 至 2018 年 1 月已经扩散至非洲 44 个国家, 发生面积超过 220 万 km² (Goergen *et al.*, 2016; Feldmann *et al.*, 2019)。截止 2019 年 7 月, 全球已有 103 个国家有该虫的分布 (<https://www.cabi.org/ISC/datasheet/29810>)。2018 年 5 月, 草地贪夜蛾扩散至亚洲国家, Sharanabasappa 等通过分子生物学手段, 证明在印度卡纳塔克邦州, 发现了草地贪夜蛾的入侵 (Sharanabasappa *et al.*, 2018)。随后迅速扩散至也门、泰国、斯里兰卡、缅甸、朝鲜、孟加拉国等亚洲国家 (Ganiger

et al., 2018; Li *et al.*, 2019)。

2019 年 1 月确认草地贪夜蛾入侵我国云南省, 截至 5 月 10 日, 该虫已在 13 省 (区) 61 市 (州) 261 个县 (市、区) 发现幼虫危害玉米, 发生面积约 108 万亩 (全国农业技术推广服务中心, 2019); 截止 2019 年 7 月 5 日, 草地贪夜蛾已快速扩散至全国 20 个省 (市、区) 1 128 个县, 累计发生面积约 831 万亩 (央视网, 2019)。我国华南、华中、华东全部和西南、华北部分区域为草地贪夜蛾的适生区 (林伟等, 2019), 7-8 月是西南季风最强的季节, 也是草地贪夜蛾随季风向北迁徙的关键季节, 存在向北扩散的可能 (吴秋林等, 2019)。

草地贪夜蛾食性杂, 其寄主多达 353 种, 多为禾本科、豆科和菊科植物 (Montezano *et al.*, 2018)。由于对寄主植物的偏好性, 草地贪夜蛾在种内分化出玉米型 (C strain, C) (取食玉米、高粱和棉花) 和水稻型 (R strain, R) (取食水稻和各种牧草) 两种品系, 两种品系形态特征无明显差异 (Pashley *et al.*, 1985; Pashley, 1988), 但 Cañas-Hoyos 等发现在南美洲, 这两个品系翅形和翅大小存在显著差异 (Canas-Hoyos *et al.*, 2016)。此外, 两个品系在性信息素成分、交配行为、对农药敏感性 (Adamczyk *et al.*, 1997; Lima&McNeil 2009; Schofl *et al.*, 2009; Cruz-Esteban *et al.*, 2018) 等方面也存在明显差异。基于线粒体 COI

基因序列分析和位于 Z 染色体的 *Tpi* 基因第三和第四个外显子 SNP 位点是目前鉴定草地贪夜蛾品系的主要方法,通过这两种方法,对入侵中国 13 个省(市、自治区) 131 个县(市、区)的 318 份草地贪夜蛾样品进行鉴定,发现入侵中国的为玉米型,可能来自美国佛罗里达州(张磊等, 2019)。

据国际农业和生物科学中心(CABI)报道,若不采取防治措施,草地贪夜蛾每年可造成 830 万至 2 060 万吨玉米的损失,直接经济损失高达 24.81 亿至 61.87 亿美金(Abrahams, *et al.*, 2017)。长期以来对草地贪夜蛾的防治依赖于化学防治和转 Bt 作物,但最新表明草地贪夜蛾已对多种农药及转 Bt 玉米产生抗性(Yu *et al.*, 2003; Carvalho *et al.*, 2013)。寻找新的草地贪夜蛾的防治和管理方法,还需依赖于遗传信息,从分子水平了解其生物学特性。2014 年, Kakumani 等采用罗氏 454 测序平台获得了草地贪夜蛾 *Sf*21 细胞系的基因组信息,并对基因组进行了注释,但注释信息目前无法获取(Kakumani *et al.*, 2014),随后, 2017 年, Gouin 等(Gouin *et al.*, 2017)对草地贪夜蛾的 C 品系和 R 品系分别测序,获得了两种品系的基因组信息,同年, Nandakumar 等获得了草地贪夜蛾 *Sf*9 细胞系的基因组,但未对基因组进行注释(Nandakumar *et al.*, 2017)。此外, GlycoBac 在 NCBI 中提交了 *Sf*-RVN 细胞系基因组,但未提交注释信息(NCBI Genome)。上述可公开获得的草地贪夜蛾基因组中, *Sf*9 细胞系的基因组 scaffold N50 达 601 Kb,其余基因组 scaffold N50 均不超过 60 Kb。本研究对 scaffold N50 最大的草地贪夜蛾基因组进行了注释,共获得 25 699 个蛋白质编码基因。同时选取了其他 14 个昆虫物种,对草地贪夜蛾进行了比较基因组学分析,发现了 2 360 个物种特异扩增基因,4 180 个物种特异基因,获得了草地贪夜蛾基因家族的扩增与收缩情况,以期对草地贪夜蛾生物学遗传特性研究提供理论支撑,并为害虫控制开发新型绿色防控方法提供研究方向。

1 材料与方法

1.1 草地贪夜蛾基因组与转录组数据

草地贪夜蛾基因组数据下载自 NCBI Assembly (assembly ASM221328v1),用于基因组注释的

21 个草地贪夜蛾转录组数据下载自 NCBI SRA (ERR2270787; SRR8449072; SRR8449070; SRR7665875; SRR6060226; SRR6060233; SRR6060235; SRR6060238; SRR6060239; SRR6060241; SRR6060265; SRR6060273; SRR6060285; SRR6060287; SRR6060294; SRR6060315; SRR3406059; ERR986580; ERR986579; ERR986578; ERR986577)。

1.2 草地贪夜蛾基因组评估与注释

1.2.1 基因组组装质量评估

使用 CEGMA (version 2.4) (Parra *et al.*, 2007) 和 BUSCO (version 3.0) (Simao *et al.*, 2015) 流程对基因组组装质量进行评估。CEGMA 的核心基因库为存在于 6 个真核模式生物中的 248 个保守基因。BUSCO 的基因库选取自 insecta_db9 中的 1 658 个昆虫保守基因。

1.2.2 重复序列注释

使用 RepeatModeler (version 1.0.7) (<http://www.repeatmasker.org>) 对基因组中的重复序列进行从头预测,得到结果文库与 RepBase (<http://www.girinst.org/repbase>) 中的节肢动物文库合并,再使用 RepeatMasker (version 4.0.5) (<http://www.repeatmasker.org>) (Tarailo-Graovac, 2009) 注释重复序列。

1.2.3 蛋白质编码基因注释

蛋白质编码基因注释使用昆虫基因组注释流程 (OMIGA) (Liu *et al.*, 2014)。使用 Hisat2-Stringtie 流程对转录组进行有参拼接 (Pertea *et al.*, 2016),获得基因转录证据;使用 GeneWise (Birney *et al.*, 2004) 和 exonerate (version 2.2.0) 软件,将基因组与 Invertebrate Refseq 中的蛋白序列进行比对,获得同源预测证据;使用基因预测软件 Augustus (version 2.5.5)、SNAP 和 GeneMark-ES (version 4.38) 对基因进行从头预测,获得从头预测证据。最后,使用 MAKER (Cantarel *et al.*, 2008) 整合转录组、同源比对和从头预测证据,获得最终的基因结构模型。

1.2.4 GO 与 KEGG 注释

Gene Ontology (GO) 注释使用 BLAST2GO 软件 (Conesa & Gotz, 2008),KEGG 注释使用 KEGG 数据库网站在线注释工具 BlastKOALA (<https://www.kegg.jp/blastkoala/>)。

1.3 比较基因组分析

本研究共选取 14 种昆虫基因组进行同源基因分析,包括 10 种蛾类昆虫、2 种蝶类昆虫、1 种毛

翅目石蛾 *Stenopsyche tienmushanensis* 以及黑腹果蝇 *Drosophila melanogaster*, 分别是: 草地贪夜蛾、斜纹夜蛾 *Spodoptera litura*、粉纹夜蛾 *Trichoplusia ni*、棉铃虫 *Helicoverpa armigera*、烟芽夜蛾 *Heliothis virescens*、家蚕 *Bombyx mori*、日本柞蚕 *Antheraea yamamai*、烟草天蛾 *Manduca sexta*、冬尺蛾 *Operophtera brumata*、帝王蝶 *Danaus plexippus*、柑橘凤蝶 *Papilio xuthus*、小菜蛾 *Plutella xylostella*。以上大多数数据下载自 NCBI Genome, 仅日本柞蚕和石蛾的数据下载自 GigaDB (<http://gigadb.org/>)。所有数据在进行后续分析之前, 均需要去冗余, 每个基因只保留一条最长的蛋白序列。使用 OrthoMCL (Fischer *et al.*, 2011) 流程进行同源基因分析。使用 CAFÉ (version 4.2.1) 软件 (Bie *et al.*, 2006) 对 OrthoMCL 同源基因分析的结果进行基因家族扩增与收缩分析。GO、KEGG 富集分析使用在线工具包完成 (<http://www.omicshare.com/>)。

1.4 鳞翅目系统发育树的构建

使用 OrthoMCL 结果中多个物种保守的 722 个单拷贝基因构建鳞翅目系统发育树, 毛翅目石蛾和黑腹果蝇作为外群。使用 MAFFT 软件 (version 7) (Katoh *et al.*, 2005) 对每个单拷贝基因序列进行多序列比对, 后使用 TrimAl (version 1.4.rev22) 软件 (Capella-Gutierrez *et al.*, 2009) 去除多序列比对结构中的低保守区域, 再将所有裁剪后的单拷贝基因连接成一个超长基因用于系统发育树的构建。使用 iqtree (version 1.6.7) 软件中的 ModelFinder 预测建树最佳模型, 然后使用 iqtree 构建系统发育树 (Nguyen *et al.*, 2015)。物种分化时

间使用 PAML (version 4.9e) 软件包中的 mcmctree 进行估算 (Yang, 2007), 4 个化石校正点信息来源于 paleobioDB (<https://www.paleobioDB.org/>), 分别为: 毛翅目 (*Phryganea solitaria*, 314.6 ~ 311.45 百万年前)、鳞翅目 (化石未命名, 208.5 ~ 201.3 百万年前)、夜蛾总科 (*Noctuities incertissima*, 33.9 ~ 28.1 百万年前) 和蚕蛾总科 (*Attacus fossilis*, 37.2 ~ 33.9 百万年前)。

2 结果与分析

2.1 草地贪夜蛾基因组组装情况

目前已公开可获得的草地贪夜蛾基因组共有 5 个版本 (表 1), *Sf*21 细胞系、C 品系以及 R 品系具有注释信息, *Sf*9 细胞系和 *Sf*-RVN 细胞系无注释信息。*Sf*9 细胞系基因组 scaffold N50 较大, 其余版本基因组 scaffold N50 均低于 60 Kb。为提供更高质量的草地贪夜蛾遗传信息, 对 scaffold N50 值最大的草地贪夜蛾基因组版本 *Sf*9 细胞系基因组 (NCBI assembly ASM221328v1) 进行注释及后续分析。*Sf*9 细胞系基因组基因组组装大小 514.3 Mb (表 1), 包括 2 396 条 scaffold, scaffold N50 为 601.1 Kb (表 1、表 2)。分别使用 CEGMA 和 BUSCO 基因组组装质量评估流程对其进行评估, 结果显示, 该组装版本的基因组可覆盖 93.6% 的 CEGMA 核心基因 (n = 248) 和 90.8% 的昆虫保守基因 (n = 1 658) (表 2), 表明基因组组装质量较好, 满足后续基因组分析的要求。

表 1 已公开发表的草地贪夜蛾基因组信息统计
Table 1 Statistic information for the published *Spodoptera frugiperda* genome

版本信息 Version	ASM75363v2 (Kakumani <i>et al.</i> , 2014)	<i>Sf</i> -RVN WGS	CORN (Gouin <i>et al.</i> , 2017)	RICE (Gouin <i>et al.</i> , 2017)	ASM221328v1 (Nandakumar <i>et al.</i> , 2017)
来源 Isolation	<i>Sf</i> 21 细胞系 <i>Sf</i> 21 cell-line	<i>Sf</i> -RVN 细胞系 <i>Sf</i> -RVN cell-line	C 品系 C strain	R 品系 R strain	<i>Sf</i> 9 细胞系 <i>Sf</i> 9 cell-line
公布年份 Publication time	2014	2016	2017	2017	2017
基因组大小 (Mb) Genome size	358	392	438	371	514.3
Scaffold N50 长度 (Kb) Length of scaffold N50	53.8	12.4	52.7	28.5	601.1
CEGMA (%)	86.24	—	—	—	—
重复序列 (%) Repeats	43.6	—	29.2	29.1	—
蛋白质编码基因 Protein-coding genes	11 595	—	21 700	26 329	—

表 2 草地贪夜蛾基因组组装与注释信息

Table 2 Overview of assembled genome and annotation of *Spodoptera frugiperda*

统计信息 Statistic information	草地贪夜蛾 <i>S. frugiperda</i> (ASM221328v1) (Nandakumar <i>et al.</i> , 2017)
基因组组装信息 Assembly information	
Number of contigs	2 844
Number of scaffolds	2 396
基因组质量评估 Quality control	
Contig N50 (Kb)	516.1
CEGMA genes (%)	93.6
BUSCO genes (%)	90.8
基因组特征 Genomic features	
Total Repeat (%)	28.1
SINEs (%)	0.7
LINEs (%)	10.7
LTR elements (%)	1.1
DNA elements (%)	1.7
G + C (%)	36.5
基因组注释 Annotation	
Number of protein-coding genes	25 699
Genes with GO terms	15 623
Genes with KEGG annotations	9 213

2.2 草地贪夜蛾基因组注释

2.2.1 重复序列注释

在草地贪夜蛾 *Sf*9 细胞系基因组中注释到 731 584 处重复序列, 总长度为 144 280 258 bp, 占整个基因组大小的 28.1%。对基因组中的转座子进行注释, 共鉴定到短散在核元件 (Short interspersed nuclear elements, SINEs) 21 926 个, 总长度为 3 530 068 bp, 占全基因组大小的 0.7%; 长散在核元件 (Long interspersed nuclear elements, LINEs) 269 624 个, 总长度为 54 825 615 bp, 占全基因组大小的 10.7%; LTR 反转录转座子 (Long terminal repeat retrotransposon) 11 041 个, 总长度为 5 804 301 bp, 占全基因组大小的 1.1%; DNA 转座子 (DNA transposon) 33 521 个, 总长度为 8 764 171 bp, 占全基因组大小的 1.7% (表 2)。

2.2.2 蛋白编码基因注释

使用昆虫基因组注释流程 (OMIGA), 结合下载自 NCBI SRA 数据库中的 21 个草地贪夜蛾转录组表达证据, 共注释得到 25 699 个蛋白编码基因 (表 2), 详细基因序列可访问 InsectBase 下载 <http://www.insect-genome.com/FAW/>。

2.2.3 基因功能注释

使用 BLAST2GO 软件对草地贪夜蛾基因组蛋白质编码基因进行 GO 注释, 最终获得 GO 注释的基因为 15 623 个 (60.8%)。使用 KEGG 在线工具 BlastKOALA 进行 KEGG 通路基因注释, 共有 9 213 个 (35.8%) 基因获得 KEGG 通路注释 (表 2)。

2.3 鳞翅目比较基因组学

选取 12 种鳞翅目昆虫进行比较基因组学分析, 包括 2 种蝶类 (帝王蝶、柑橘凤蝶) 和 10 种蛾类 (夜蛾科 Noctuidae: 草地贪夜蛾、斜纹夜蛾、粉纹夜蛾、棉铃虫、烟芽夜蛾, 蚕蛾总科 Bombycoidea: 家蚕、日本柞蚕、烟草天蛾, 尺蛾科 Geometridae: 冬尺蛾, 菜蛾科 Plutellidae: 小菜蛾)。根据昆虫系统发育关系可知毛翅目、双翅目与鳞翅目亲缘关系较近, 毛翅目是鳞翅目的姐妹群, 故本研究使用毛翅目石蛾和黑腹果蝇作为外群。

使用从 OrthoMCL 软件结果中提取的 722 个严格单拷贝基因对 14 种昆虫构建系统发育树。结果显示, 12 种鳞翅目昆虫形成鳞翅目簇, 并与毛翅目石蛾互为姐妹群。而双翅目昆虫与鳞翅目和毛翅目的共同祖先互为姐妹群。菜蛾科昆虫小菜蛾在鳞翅目进化过程中较早地分化出来, 属于鳞翅目进化树较基部的类群。5 种夜蛾科昆虫 (草地贪夜蛾、斜纹夜蛾、粉纹夜蛾、棉铃虫、烟芽夜蛾) 与尺蛾科昆虫 (冬尺蛾) 聚成一支, 后与蚕蛾总科昆虫 (家蚕、日本柞蚕、烟草天蛾) 聚为一大支。两种蝴蝶 (帝王蝶、柑橘凤蝶) 聚为一支, 与 9 种蛾类形成的大支形成姐妹群。在夜蛾科内, 棉铃虫和烟芽夜蛾聚在一起, 而隶属于同属的草地贪夜蛾和斜纹夜蛾聚在一起, 粉纹夜蛾作为夜蛾科较早分化的物种, 位于夜蛾科系统发育树的基部 (图 1)。所有节点的 Bootstrap 值均为 100, 说明各节点的支持率都很高, 本研究构建的鳞翅目昆虫系统发育关系较可靠。

挑选 4 个化石校正点, 对鳞翅目物种分化时间进行校正。结果显示, 鳞翅目起源于距今 2.04 亿年的三叠纪和侏罗纪的过渡时期, 蝶类与

蛾类分化大约在距今 1.67 亿年的侏罗纪时期, 夜蛾科昆虫出现在大约距今 3 000 万年前, 草地贪夜蛾与斜纹夜蛾的分化时间大约在 1 284 万年前 (图 1)。

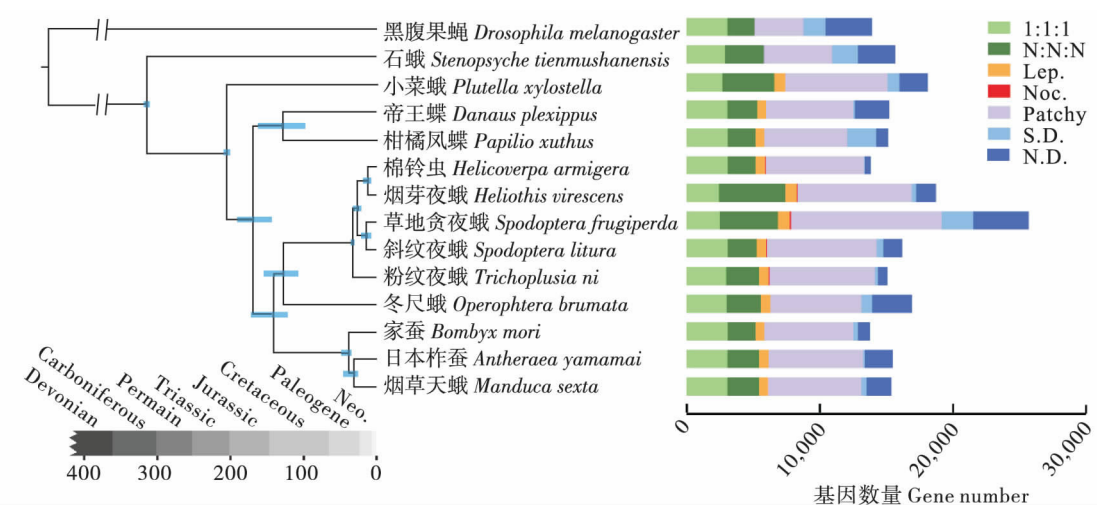


图 1 鳞翅目比较基因组学分析
Fig. 1 Lepidoptera comparative genomics

注: 鳞翅目系统发育树由 722 个单拷贝基因构建, 每个节点的 Bootstrap 值均为 100。进化树各节点上的浅蓝色横条代表分化时间 (95% 置信区间)。右侧柱形图为各物种基因组基因分类图。1:1:1, 单拷贝基因 (universal single-copy gene), 即至少在 13 个物种中为单拷贝的基因; N:N:N, 多拷贝基因 (universal multiple-copy gene); Lep., 鳞翅目昆虫特有的基因; Noc., 夜蛾科昆虫特有基因; Patchy, 其余所有的基因; S. D., 物种特异扩增的基因; N. D., 物种特异基因。Note: Phylogenetic reconstruction of the Lepidoptera based on 722 single-copy genes, and all bootstrap values are equal to 100 for every nodes. Divergence times are indicated by light blue bars at the internodes, the bars indicates the 95% confidence interval of the divergence time. The horizontal bars shows the protein-coding genes categories for each species. 1:1:1, universal single-copy genes represent genes are single copy across all species we analyzed, absence or duplication in a single genome was allowed; N:N:N, universal multiple-copy genes represent other universal genes; Lep., common genes unique to Lepidoptera; Noc., common genes unique to Noctuidae; Patchy, includes all remaining orthologues. S. D., species specific duplication genes; N. D., species specific genes with only one copy in genome.

使用 OrthoMCL 对所选鳞翅目昆虫的蛋白编码基因进行同源分析, 在草地贪夜蛾中共鉴定出 2 490 个物种单拷贝基因、4 347 个多拷贝基因、891 个鳞翅目昆虫特有基因、100 个夜蛾科昆虫特有基因、2 360 个物种特异扩增基因和 4 180 个物种特异基因 (图 1)。单拷贝基因和多拷贝基因在物种中序列相似, 较为保守。而物种特异扩增基因与物种特异基因往往因为序列与其他物种相差较大, 导致其无法聚类, 这些基因主要是一些新基因、进化速率较快的基因。因此, 这些特异基因在物种长期进化过程中往往具有重要作用。

对草地贪夜蛾的 2 360 个物种特异扩增基因进行 GO 富集分析, 结果显示, 这些特异扩增的基因主要参与 DNA 整合、代谢相关的生物进程 (表 3)。使用 KEGG 通路富集分析, 发现草地贪夜蛾的特异扩增基因主要涉及 HIF-1 信号通路、糖代谢通路

和寿命调节通路等 (图 2A)。对草地贪夜蛾的 4 180 个物种特异基因进行 GO 富集分析, 结果显示, 特异基因主要与酶活性、光感受、质膜结构、糖代谢等相关 (表 4)。KEGG 通路富集分析显示, 草地贪夜蛾的特异基因主要参与氨基酸代谢、糖代谢和 Wnt 信号通路 (图 2B)。与糖类的消化和吸收相关的特异扩增基因主要为己糖激酶 (hexokinase) 和磷脂酰肌醇磷脂酶 (phosphatidylinositol phospholipase C, beta)。在淀粉和蔗糖代谢通路中, 鉴定到多个草地贪夜蛾物种特异基因, 包括 1, 4- α -葡聚糖分支酶 (1, 4- α -glucan branching enzyme)、葡萄糖-6-磷酸酯酶 (glucose-6-phosphatase)、 α 淀粉酶 (α -amylase)、 β -呋喃果糖苷酶 (β -fructofuranosidase) 和 α - α 海藻糖酶 (α , α -trehalase)。在果糖和甘露糖代谢通路中, 草地贪夜蛾特异基因包括果糖二磷酸醛缩酶

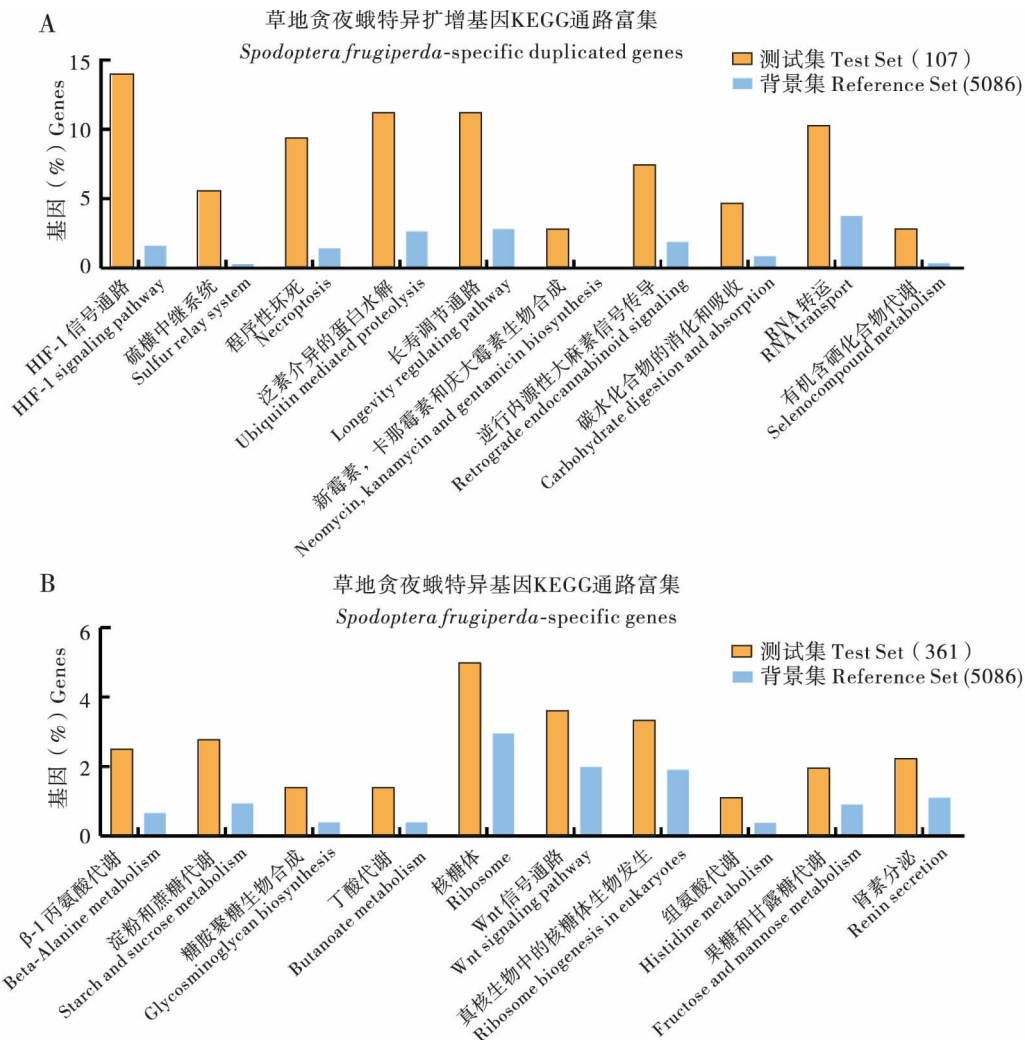


图2 草地贪夜蛾特异扩增基因与特异基因 KEGG 通路富集

Fig. 2 Significantly overrepresented KEGG pathways for specific duplication genes and species specific genes of *Spodoptera frugiperda*

(fructose-bisphosphate aldolase, class I)、木糖异构酶 (xylose isomerase) 和磷酸甘露糖变位酶 (phosphomannomutase)。在草地贪夜蛾的多个氨基酸代谢通路 (包括丙氨酸、组氨酸、精氨酸、脯氨酸、色氨酸、酪氨酸、亮氨酸、异亮氨酸) 中均可发现物种特异基因。草地贪夜蛾的 Wnt 信号通路中共有 12 个特异基因, 分别是 WNT5 (wingless-type MMTV integration site family, member 5)、WNT7 (wingless-type MMTV integration site family, member 7)、酪蛋白激酶 II 亚基 β (casein kinase II subunit beta)、蛋白激酶 A (protein kinase A)、RAC1 (Ras-related C3 botulinum toxin substrate 1)、JNK (c-Jun N-terminal kinase)、类 nemo 激酶 (nemo like kinase)、GROUCHO、vang-like、DAAM (disheveled associated activator of morphogenesis)、

Ras 同源基因家族 A (Ras homolog gene family, member A) 和酪蛋白激酶 (casein kinase 1, alpha)。综上所述, 草地贪夜蛾大量的特异扩增与物种特异基因很可能与其的强大的适应性和快速扩散能力密切相关。

2.4 基因家族的扩增与收缩

使用 CAFE 软件对草地贪夜蛾基因组进行基因家族扩增与收缩分析。结果显示, 草地贪夜蛾相对于其与斜纹夜蛾的共同祖先, 共有 4 090 个基因家族发生了扩增现象, 同时也有 1 974 个基因家族发生了收缩 (图 3)。然而, 斜纹夜蛾相对其与草地贪夜蛾的共同祖先仅有 644 个基因家族发生了扩增, 并有 2 253 个基因家族发生的收缩 (图 3)。

对草地贪夜蛾发生扩增与收缩的基因家族进行 GO 与 KEGG 富集分析。结果显示, 草地贪夜蛾

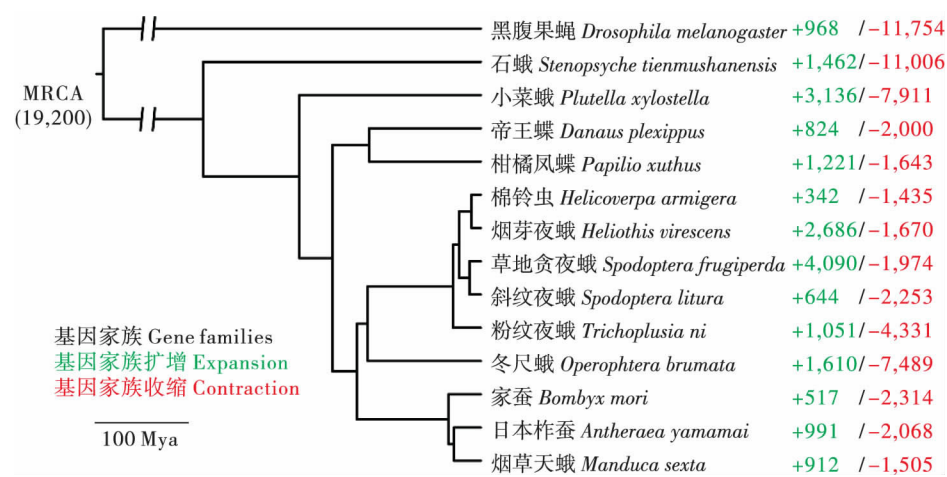


图3 草地贪夜蛾基因组基因家族扩增与收缩分析

Fig. 3 Gene family expansions and contractions of *Spodoptera frugiperda*

注: 鳞翅目系统发育树由 722 个单拷贝基因构建, 同图 1。每个物种名后的数字代表扩增的基因家族 (绿色) 与收缩的基因家族 (红色)。MRCA, 最近共同祖先。括号中的数字代表用于 CAFÉ 软件分析的最近共同祖先基因家族的数量。Note: Phylogenetic reconstruction of the Lepidoptera based on 722 single-copy genes, modified form Fig. 1. Numbers behind the species names indicate the expansion (in green) and contraction (in red) of gene families. MRCA, most recent common ancestor. The number in parentheses means the number of gene families in the MRCA as estimated by CAFÉ.

表3 草地贪夜蛾 2 360 个物种特异扩增基因 GO 富集分析

Table 3 GO enrichment analysis of 2 360 specific duplication genes of *Spodoptera frugiperda*

GO 编号 GO ID	描述 Description	P 值 P-value
细胞组分 Cellular component		
GO: 0005634	核 Nucleus	2.04E-19
GO: 0044815	DNA 包装复合体 DNA packaging complex	3.12E-13
GO: 0000786	核小体 Nucleosome	1.94E-12
GO: 0000788	核小体 Nuclear nucleosome	3.19E-12
GO: 0005704	多烯染色体带 Polytene chromosome band	6.96E-12
分子功能 Molecular function		
GO: 0003964	RNA 介导的 DNA 聚合酶活性 RNA-directed DNA polymerase activity	6.91E-56
GO: 0016779	核苷酸转移酶活性 Nucleotidyltransferase activity	1.23E-54
GO: 0034061	DNA 聚合酶活性 DNA polymerase activity	6.54E-54
GO: 0016772	转移酶活性 (含磷酸基团) Transferase activity, transferring phosphorus-containing groups	4.78E-38
GO: 0016740	转移酶活性 Transferase activity	2.11E-27
生物学过程 Biological process		
GO: 0006313	DNA 介导的转座 Transposition, DNA-mediated	6.70E-55
GO: 0032196	转座 Transposition	7.92E-55
GO: 0006259	DNA 代谢进程 DNA metabolic process	5.25E-54
GO: 0015074	DNA 整合 DNA integration	6.20E-53
GO: 0006310	DNA 重组 DNA recombination	2.94E-46

表 4 草地贪夜蛾 4 180 个物种特异基因 GO 富集分析

Table 4 GO enrichment analysis of 4 180 species specific genes of *Spodoptera frugiperda*

GO 编号 GO ID	描述 Description	P 值 P-value
细胞组分 Cellular component		
GO: 0042597	周质空间 Periplasmic space	0.000813954
GO: 0005581	胶原蛋白三聚体 Collagen trimer	0.001419401
GO: 0030312	外部封装结构 External encapsulating structure	0.001456913
GO: 0001917	光感受器 Photoreceptor inner segment	0.001632073
GO: 0031226	质膜内组分 Intrinsic component of plasma membrane	0.003568763
分子功能 Molecular function		
GO: 0004989	章鱼胺受体活性 Octopamine receptor activity	0.000014978
GO: 0016787	水解酶活性 Hydrolase activity	0.000097204
GO: 0003824	催化活性 Catalytic activity	0.000112899
GO: 0019789	SUMO 转移酶活性 SUMO transferase activity	0.000352665
GO: 0004175	内肽酶活性 Endopeptidase activity	0.000410684
生物学过程 Biological process		
GO: 0050775	树突形态发生的正调节 Positive regulation of dendrite morphogenesis	0.000059171
GO: 0007156	通过质膜粘附分子的嗜同性细胞粘附 Homophilic cell adhesion via plasma membrane adhesion molecules	0.000254421
GO: 0019321	戊糖代谢过程 Pentose metabolic process	0.000481142
GO: 0032211	端粒酶对端粒维持的负调节作用 Negative regulation of telomere maintenance via telomerase	0.000598973
GO: 0006739	NADP 代谢途径 NADP metabolic process	0.000722700

扩增的基因家族主要与 DNA 包装 (GO: 0044815, DNA packaging complex, $P = 7.22\text{E-}15$)、DNA 聚合酶活性 (GO: 0034061, DNA polymerase activity, $P = 3.75\text{E-}37$)、DNA 代谢 (GO: 0006259, DNA metabolic process, $P = 5.20\text{E-}32$)、脂肪酸合成通路 (ko00061, Fatty acid biosynthesis, $P = 1.10\text{E-}05$)、PPAR 信号通路 (ko03320, PPAR signaling pathway, $P = 0.001$) 和柠檬酸循环 (ko00020, Citrate cycle, $P = 0.001$) 等相关。进一步分析发现, 草地贪夜蛾柠檬酸循环通路中的苹果酸脱氢酶 (malate dehydrogenase) 基因、异柠檬酸脱氢酶 (isocitrate dehydrogenase) 基因、丙酮酸脱氢酶 (pyruvate dehydrogenase) 基因、琥珀酸脱氢酶 (succinate dehydrogenase) 基因、二氢硫辛酰胺脱氢酶 (dihydrolipoamide dehydrogenase) 基因、2-氧代戊二酸脱氢酶 (2-oxoglutarate dehydrogenase) 基因、磷酸烯醇丙酮酸羧激酶 (phosphoenolpyruvate carboxykinase) 基因、富马酸水合酶 (fumarate

hydratase) 基因和琥珀酰辅酶 A 合成酶 (succinyl-CoA synthetase) 基因均相较斜纹夜蛾有扩增迹象。另外, 草地贪夜蛾的脂肪酸合成酶 (fatty acid synthase) 基因、长链酰基辅酶 A 合成酶 (long-chain acyl-CoA synthetase) 基因以及硬脂酰辅酶 A 去饱和酶 (stearoyl-CoA desaturase) 等脂肪酸合成、代谢相关通路基因也有扩增现象。有趣的是, 草地贪夜蛾的核心组蛋白基因 Histone H2A、Histone H2B、Histone H3 和 Histone H4 相比斜纹夜蛾也有扩增迹象。

对草地贪夜蛾收缩的基因家族进行分析, 结果发现其主要与过氧化物酶体 (GO: 0005777, Peroxisome, $P = 3.80\text{E-}12$)、铁离子结合 (GO: 0005506, Iron ion binding, $P = 4.19\text{E-}20$)、类固醇激素生物合成 (ko00140, Steroid hormone biosynthesis, $P = 2.73\text{E-}16$) 和视黄醇代谢 (ko00830, Retinol metabolism, $P = 7.97\text{E-}16$) 等生物过程和通路相关。

3 结论与讨论

目前在 NCBI 中可下载的草地贪夜蛾基因组共有 4 个版本, 分别来自于 *Sf21* 细胞系 (Kakumani *et al.*, 2014), *Sf-RVN* 细胞系, C 品系 (Gouin *et al.*, 2017) 以及 *Sf9* 细胞系 (Nandakumar *et al.*, 2017), Gouin 等 (Gouin *et al.*, 2017) 通过二代测序获得了草地贪夜蛾 C 品系和 R 品系基因组, 但 R 品系基因组未上传至 NCBI, 5 个版本获得的草地贪夜蛾基因组大小在 350 ~ 520 Mb 之间, C 品系和 R 品系版本基因组具有注释信息, 但 *Sf21*、*Sf-RVN* 和 *Sf9* 细胞系版本均无法获取注释信息。本研究选取 5 个版本中 scaffold N50 最长的 *Sf9* 细胞系基因组进行注释, 以期获得更高质量的草地贪夜蛾基因信息。*Sf9* 细胞系基因组 scaffold N50 长达 601.1 Kb, CEGMA 评估发现该版本基因组涵盖 93.6% 的 CEGMA 核心基因, 高于 *Sf21* 细胞系基因组的 86.24%, 重复序列占基因组的 28.1%, 低于 *Sf21* 细胞系, 与 C 品系和 R 品系基因组相近。

本研究选取 12 个鳞翅目昆虫物种进行比较基因组学分析, 结果显示草地贪夜蛾与斜纹夜蛾亲缘关系最近, 用化石信息校正物种分化时间后, 进一步分析发现, 鳞翅目起源于 2.04 亿年前, 草地贪夜蛾与斜纹夜蛾的分化时间约在 1 284 万年前。Misof 等对 7 个鳞翅目类群以及 3 个鳞翅目的化石信息进行分析, 发现鳞翅目昆虫起源于 2.45 亿年前 (Misof *et al.*, 2014), 而 Condamine 等基于化石和分子证据分析昆虫起源时间, 发现鳞翅目昆虫起源于 2.7 亿年前 (Condamine *et al.*, 2016)。昆虫分化的估计受多因素的影响, 鳞翅目昆虫的起源时间仍不能确定。

对鳞翅目昆虫的蛋白质编码基因进行同源分析, 在草地贪夜蛾中共发现 2 360 个物种特异扩增基因和 4 180 个物种特异基因, 进一步 KEGG 通路富集发现, 在淀粉和蔗糖代谢通路中, 有多个物种特异基因, 其中包括 β -呋喃果糖苷酶。 β -呋喃果糖苷酶是一种蔗糖水解酶, 广泛存在于植物和微生物中, 在动物体内报道甚少, 2008 年, Daimon 等首次在家蚕中扩增出动物型 β -呋喃果糖苷酶基因 *BmSuc1* 和 *BmSuc2*, *BmSuc1* 在家蚕中肠中高表达, 可编码产生有活性的 β -呋喃果糖苷酶, 且其活性不受桑叶乳汁中高浓度的 DNJ、D-AB1 等糖类生物碱影响, 进而避开生物碱的毒害, 而

另一广泛存在于生物体内的保守水解酶- α 葡萄糖苷酶, 其活性受 DNJ、D-AB1 等生物碱的强烈抑制 (Daimon *et al.*, 2008)。目前, 已在小蔗螟 *Diatraea saccharalis* (Carneiro *et al.*, 2004)、木薯天蛾 *Erinnyis ello* (Santos & Terra, 1986)、烟草天蛾 (Pauchet *et al.*, 2010)、棉铃虫 (Pauchet *et al.*, 2008) 等多个鳞翅目昆虫中发现 β -呋喃果糖苷酶的存在, 而除鞘翅目象鼻科部分物种外, 在黑腹果蝇、赤拟谷盗 *Tribolium castaneum*、冈比亚按蚊 *Anopheles gambiae* 等物种中未发现该基因的存在 (Li, 2016)。鳞翅目昆虫以植食性为主, 而蔗糖是高等植物光合作用的主要产物, 鳞翅目昆虫的取食、营养和代谢依赖于蔗糖的获取, 利用中肠的蔗糖水解酶将蔗糖水解为葡萄糖和果糖, 产生碳源和能源。天敌昆虫的取食是对植物最典型的生物胁迫 (Schuman & Baldwin, 2016), 在长期进化过程中, 植物已进化出结构性防御、稳定态防御、诱导性防御和激活性防御等多种防御机制 (Behmer *et al.*, 2005), 典型的昆虫取食非适应性寄主植物, 植物分泌化学防御物质毒杀昆虫, 鳞翅目昆虫对植物化学防御物质的适应多为诱导性机制, 特别是取食后的消化过程, 进化出多种适应性机制, 如中肠呈强碱性环境 (Gringorten *et al.*, 1993), 草地贪夜蛾中肠 pH 高达 10, 对玉米体内防御性关键酶- β -葡萄糖苷酶的活性具有显著抑制作用 (Pentzold *et al.*, 2014)。而草地贪夜蛾寄主植物多达 76 个科 353 种 (Montezano *et al.*, 2018), 如此广适性寄主范围, 推测其体内具有多种解毒水解酶, 分解被活化的植物化学防御物质。鳞翅目昆虫中肠存在的蔗糖水解酶- β -呋喃果糖苷酶在草地贪夜蛾中特异性存在, 具有 7 个拷贝, 序列与其他鳞翅目物种不相似, orthomcl 分析不聚为一类, 该基因在其他鳞翅目物种如帝王蝶、菜粉蝶 *Pieris rapae*、棉铃虫、家蚕、小菜蛾、斜纹夜蛾基因组中为单拷贝基因, 在脐橙螟 *Amyelois transitella*、烟草天蛾、冬尺蛾基因组中具有 2 个拷贝 (OrthoDB 数据), 推测草地贪夜蛾在进化过程中, 为适应生存, β -呋喃果糖苷酶基因快速扩增, 躲避寄主植物化学防御物质的毒害, 并快速水解蔗糖获得能量进行生存和繁殖。此外, 在草地贪夜蛾淀粉和糖代谢通路, 果糖和甘露糖代谢通路中, 均存在其他特异性基因, 基因组家族扩增收缩分析也发现草地贪夜蛾基因组中存在大量基因家族发生扩增, 这些特异基因以及基因家族扩增, 可能与草地贪夜蛾的快

速繁殖、高适应性、快速扩散密切相关。但本研究分析数据来源于草地贪夜蛾 *Sf9* 细胞系, 其基因组与草地贪夜蛾真实虫体基因组具有一定差异, 而在 bioRxiv 上有文献预先报道了高质量的基因组, 这将极大地推动草地贪夜蛾的有关研究 (Liu *et al.*, 2019)。

参考文献 (References)

- Abrahams P, Bateman M, Beale T, *et al.* Fall armyworm: Impacts and implications for Africa. Evidence note (2), September 2017 [R]. Report to DFID. Wallingford, UK: CAB International.
- Adamczyk JJ, Holloway J, Leonard B, *et al.* Susceptibility of fall armyworm collected from different plant hosts to selected insecticides and transgenic Bt cotton [J]. *The Journal of Cotton Science*, 1997, 1: 21–28.
- Behmer ST, Lloyd CM, Raubenheimer D, *et al.* Metal hyperaccumulation in plants: Mechanisms of defence against insect herbivores [J]. *Functional Ecology*, 2005, 19 (1): 55–66.
- Bie TD, Cristianini N, Demuth JP, *et al.* CAFE: Computational analysis of gene family evolution [J]. *Genome Research*, 2006, 22 (10): 1269–1271.
- Birney E, Clamp M, Durbin R. Genewise and genomewise [J]. *Genome Research*, 2004, 14 (5): 988–995.
- Canas – Hoyos N, Marquez EJ, Saldamando – Benjumea CI. Heritability of wing size and shape of the rice and corn strains of *Spodoptera frugiperda* (JE Smith) (Lepidoptera: Noctuidae) [J]. *Neotropical Entomology*, 2016, 45 (4): 411–419.
- Cantarel BL, Korf I, Robb SM, *et al.* MAKER: An easy – to – use annotation pipeline designed for emerging model organism genomes [J]. *Genome Research*, 2008, 18 (1): 188–196.
- Capella – Gutierrez S, Silla – Martinez JM, Gabaldon T. trimAl: A tool for automated alignment trimming in large – scale phylogenetic analyses [J]. *Bioinformatics*, 2009, 25 (15): 1972–1973.
- Carneiro CNB, Isejima EM, Samuels RI, *et al.* Sucrose hydrolases from the midgut of the sugarcane stalk borer *Diatraea saccharalis* [J]. *Journal of Insect Physiology*, 2004, 50 (11): 1093–1101.
- Carvalho RA, Omoto C, Field LM, *et al.* Investigating the molecular mechanisms of organophosphate and pyrethroid resistance in the fall armyworm *Spodoptera frugiperda* [J]. *PLoS ONE*, 2013, 8 (4): 1–5.
- CCTV-2. *Spodoptera frugiperda* was found in 1, 128 counties of 20 provinces in China [R/OL]. [2019-07-08]. <http://sannong.cctv.com/2019/07/08/ARTIElFvtrkKcSVZQuHjOC0I90708.shtml>. [央视网. 全国 20 个省市区 1, 128 个县发现草地贪夜蛾 [R/OL]. [2019-07-08]. <http://sannong.cctv.com/2019/07/08/ARTIElFvtrkKcSVZQuHjOC0I90708.shtml>]
- Chen N. Using Repeat Masker to identify repetitive elements in genomic sequences [J]. *Current Protocols in Bioinformatics*, 2009, 4. 10.1–4. 10.14.
- Condamine FL, Clapham ME, Kergoat GJ. Global patterns of insect diversification: Towards a reconciliation of fossil and molecular evidence? [J]. *Scientific Reports*, 2016, 6: 19208.
- Conesa A, Gotz S. Blast2GO: A comprehensive suite for functional analysis in plant genomics [J]. *International Journal of Plant Genomics*, 2008, 2008: 619832.
- Cruz-Esteban S, Rojas JC, Sanchez-Guillen D, *et al.* Geographic variation in pheromone component ratio and antennal responses, but not in attraction, to sex pheromones among fall armyworm populations infesting corn in Mexico [J]. *Journal of Pest Science*, 2018, 91 (3): 973–983.
- Daimon T, Taguchi T, Meng Y, *et al.* Beta-fructofuranosidase genes of the silkworm, *Bombyx mori*: Insights into enzymatic adaptation of *B. mori* to toxic alkaloids in mulberry latex [J]. *Journal of Biological Chemistry*, 2008, 283 (22): 15271–15279.
- Feldmann F, Rieckmann U, Winter S. The spread of the fall armyworm *Spodoptera frugiperda* in Africa—what should be done next? [J]. *Journal of Plant Diseases and Protection*, 2019, 126 (2): 97–101.
- Fischer S, Brunk BP, Chen F, *et al.* Using OrthoMCL to assign proteins to OrthoMCL-DB groups or to cluster proteomes into new ortholog groups [J]. *Current Protocols in Bioinformatics*, 2011, Chapter 6: Unit 6.12.11–19.
- Ganiger PC, Yeshwanth HM, Muralimohan K, *et al.* Occurrence of the new invasive pest, fall armyworm, *Spodoptera frugiperda* (JE Smith) (Lepidoptera: Noctuidae), in the maize fields of Karnataka, India [J]. *Current Science*, 2018, 115 (4): 621–623.
- Goergen G, Kumar PL, Sankung SB, *et al.* First report of outbreaks of the fall armyworm *Spodoptera frugiperda* (J. E. Smith) (Lepidoptera, Noctuidae), a new alien invasive pest in west and central Africa [J]. *PLoS ONE*, 2016, 11 (10): e0165632.
- Gouin A, Bretaudeau A, Nam K, *et al.* Two genomes of highly polyphagous lepidopteran pests (*Spodoptera frugiperda*, Noctuidae) with different host-plant ranges [J]. *Scientific Reports*, 2017, 7 (1): 11816. doi: 10.1038/s41598-017-10461-4.
- Gringorten JL, Crawford DN, Harvey WR. High pH in the ectoperitrophic space of the larval Lepidopteran midgut [J]. *Journal of Experimental Biology*, 1993, 183: 353–359.
- Johnson SJ. Migration and the life-history strategy of the fall armyworm, *Spodoptera frugiperda* in the western hemisphere [J]. *Insect Science and Its Application*, 1987, 8 (4–6): 543–549.
- Kakumani PK, Malhotra P, Mukherjee SK, *et al.* A draft genome assembly of the army worm, *Spodoptera frugiperda* [J]. *Genomics*, 2014, 104 (2): 134–143.
- Katoh K, Kuma K, Toh H, *et al.* MAFFT version 5: Improvement in accuracy of multiple sequence alignment [J]. *Nucleic Acids Research*, 2005, 33 (2): 511–518.
- Lima ER, Mcneil JN. Female sex pheromones in the host races and hybrids of the fall armyworm, *Spodoptera frugiperda* (Lepidoptera: Noctuidae) [J]. *Chemoeology*, 2009, 19 (1): 29–36.
- Li X. Site-directed Mutagenesis of β -fructofuranosidase of *Bombyx mori* and Analyses on the Active Residues [D]. Hefei: Anhui Agricultural University, 2016: 4–5. [李鑫. 家蚕 β -呋喃果糖苷酶的定点突变与功能活性位点研究 [D]. 合肥: 安徽农业大学, 2016: 4–5]
- Li X, Wu M, Ma J, *et al.* Prediction of migratory routes of the invasive fall armyworm in eastern China using a trajectory analytical approach [J/OL]. *bioRxiv*, <https://doi.org/10.1101/625632>, 2019.
- Lin W, Xu MF, Quan YB, *et al.* Potential geographic distribution of

- Spodoptera frugiperda* in China based on MaxEnt model [J/OL]. *Plant Quarantine*, 2019, <http://kns.cnki.net/kcms/detail/11.1990.s.20190422.1026.002.html>. [林伟, 徐森锋, 权永兵, 等. 基于 MaxEnt 模型的草地贪夜蛾适生性分析 [J/OL]. 植物检疫, 2019, <http://kns.cnki.net/kcms/detail/11.1990.s.20190422.1026.002.html>]
- Liu H, Lan TM, Fang DM, *et al.* Chromosome level draft genomes of the fall armyworm, *Spodoptera frugiperda* (Lepidoptera: Noctuidae), an alien invasive pest in China [J/OL]. *bioRxiv*. <https://www.biorxiv.org/content/10.1101/671560v1>.
- Liu JD, Xiao HM, Huang SQ, *et al.* OMIGA: Optimized maker-based insect genome annotation [J]. *Molecular Genetics and Genomics*, 2014, 289 (4): 567–573.
- Martinelli S, Barata RM, Zucchi MI, *et al.* Molecular variability of *Spodoptera frugiperda* (Lepidoptera: Noctuidae) populations associated to maize and cotton crops in Brazil [J]. *Journal of Economic Entomology*, 2006, 99 (2): 519–526.
- Misof B, Liu SL, Meusemann K, *et al.* Phylogenomics resolves the timing and pattern of insect evolution [J]. *Science*, 2014, 346 (6210): 763–767.
- Montezano DG, Specht A, Sosa-Gomez DR, *et al.* Host plants of *Spodoptera frugiperda* (Lepidoptera: Noctuidae) in the Americas [J]. *African Entomology*, 2018, 26 (2): 286–300.
- Nandakumar S, Ma H, Khan AS. Whole-genome sequence of the *Spodoptera frugiperda* Sf 9 insect cell line [J]. *Microbiology Resource Announcements*, 2017, 5 (34): e00829–17.
- National Agricultural Technology Extension and Service Center. *Spodoptera frugiperda* invaded 13 provinces and damage the spring maize [R/OL]. Plant disease and insect information, No. 18, 2019. [全国农业技术推广服务中心. 草地贪夜蛾侵入 13 省份为害春玉米 [R/OL]. 植物病虫情报, 第 18 期, 2019. https://www.natesc.org.cn/Html/2019_05_14/28092_151760_2019_05_14_458487.html].
- Nguyen LT, Schmidt HA, Von Haeseler A, *et al.* IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies [J]. *Molecular Biology and Evolution*, 2015, 32 (1): 268–274.
- Parra G, Bradnam K, Korf I. CEGMA: A pipeline to accurately annotate core genes in eukaryotic genomes [J]. *Bioinformatics*, 2007, 23 (9): 1061–1067.
- Pashley DP. Quantitative genetics, development, and physiological adaptation in host strains of fall armyworm [J]. *Evolution*, 1988, 42 (1): 93–102.
- Pashley DP, Johnson SJ, Sparks AN. Genetic population-structure of migratory moths – the fall armyworm (Lepidoptera, Noctuidae) [J]. *Annals of the Entomological Society of America*, 1985, 78 (6): 756–762.
- Pauchet Y, Muck A, Svatos A, *et al.* Mapping the larval midgut lumen proteome of *Helicoverpa armigera*, a generalist herbivorous insect [J]. *Journal of Proteome Research*, 2008, 7 (4): 1629–1639.
- Pauchet Y, Wilkinson P, Vogel H, *et al.* Pyrosequencing the *Manduca sexta* larval midgut transcriptome: Messages for digestion, detoxification and defence [J]. *Insect Molecular Biology*, 2010, 19 (1): 61–75.
- Pentzold S, Zagrobelny M, Rook F, *et al.* How insects overcome two – component plant chemical defence: plant beta – glucosidases as the main target for herbivore adaptation [J]. *Biological Reviews*, 2014, 89 (3): 531–551.
- Pertea M, Kim D, Pertea GM, *et al.* Transcript – level expression analysis of RNA – seq experiments with HISAT, StringTie and Ballgown [J]. *Nature Protocol*, 2016, 11 (9): 1650–1667.
- Prasanna BM, Huesing JE, Eddy R, *et al.* Fall armyworm in Africa: A guide for integrated pest management [J]. CABI, 2018. https://www.usaid.gov/sites/default/files/documents/1867/Fall-Armyworm-IPM-Guide-for-Africa-Jan_30-2018.pdf.
- Santos CD, Terra WR. Midgut alpha-glucosidase and beta-fructosidase from *Erinnyis ello* larvae and imagoes-physical and kinetic-properties [J]. *Insect Biochemistry*, 1986, 16 (5): 819–824.
- Schoff G, Heckel DG, Groot AT. Time-shifted reproductive behaviours among fall armyworm (Noctuidae: *Spodoptera frugiperda*) host strains: Evidence for differing modes of inheritance [J]. *Journal of Evolutionary Biology*, 2009, 22 (7): 1447–1459.
- Schuman MC, Baldwin IT. The layers of plant responses to insect herbivores [J]. *Annual Review of Entomology*, 2016, 61: 373–394.
- Sharanabasappa, Kalleshwaraswamy CM, Asokan R, *et al.* First report of the fall armyworm, *Spodoptera frugiperda* (J. E. Smith) (Lepidoptera Noctuidae), an alien invasive pest on maize in India [J]. *Pest Management in Horticultural Ecosystems*, 2018, 24 (1): 23–29.
- Simao FA, Waterhouse RM, Ioannidis P, *et al.* BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs [J]. *Bioinformatics*, 2015, 31 (19): 3210–3212.
- Tarailo-Graovac M and Chen N. Using RepeatMasker to identify repetitive elements in genomic sequences [J]. *Current Protocols in Bioinformatics*, 2009, 4. 10.1–4. 10.14.
- Todd EL, Poole RW. Keys and illustrations for the armyworm moths of thenoctuid genus *Spodoptera* guenee from the western hemisphere [J]. *Annals of the Entomological Society of America*, 1980, 73 (6): 722–738.
- Westbrook JK, Nagoshi RN, Meagher RL, *et al.* Modeling seasonal migration of fall armyworm moths [J]. *International Journal of Biometeorology*, 2016, 60 (2): 255–267.
- Wu QL, Jiang YY, Hu G, *et al.* Analysis on spring and summer migration routes of fall armyworm (*Spodoptera frugiperda*) from tropical and southern subtropical zones of China [J]. *Plant Protection*, 2019, 45 (3): 1–9. [吴秋林, 姜玉英, 胡高, 等. 中国热带和亚热带地区草地贪夜蛾春夏两季迁飞轨迹的分析 [J]. 植物保护, 2019, 45 (3): 1–9]
- Yang ZH. PAML 4: Phylogenetic analysis by maximum likelihood [J]. *Molecular Biology and Evolution*, 2007, 24 (8): 1586–1591.
- Yu SJ, Nguyen SN, Abo-Elghar GE. Biochemical characteristics of insecticide resistance in the fall armyworm, *Spodoptera frugiperda* (J. E. Smith) [J]. *Pesticide Biochemistry and Physiology*, 2003, 77 (1): 1–11.
- Zhang L, Liu B, Jing YY, *et al.* Molecular characterization analysis of fall armyworm populations in China [J/OL]. *Plant Protection*, 2019, <https://doi.org/10.16688/j.zwbh.2019296>. [张磊, 柳贝, 姜玉英, 等. 中国不同地区草地贪夜蛾种群生物型分子特征分析 [J/OL]. 植物保护, 2019, <https://doi.org/10.16688/j.zwbh.2019296>]