



沈族恒,周雁.异色瓢虫基因组及免疫系统分析[J].江西农业大学学报,2023,45(1):45-52.

SHEN Z H,ZHOU Y.Genome for *Harmonia axyridis* and immune system analysis[J].Acta agriculturae universitatis Jiangxiensis, 2023,45(1):45-52.

异色瓢虫基因组及免疫系统分析

沈族恒,周 雁*

(复旦大学 生命科学学院,上海 200438)

摘要:【目的】异色瓢虫是在全世界广泛分布的一种捕食性甲虫,由于其强大的适应和繁殖能力,异色瓢虫种群数量已经大幅增加并且威胁到当地的生态环境。基于基因组数据和分析结果,可以推动对异色瓢虫资源的利用和其生物入侵的治理。【方法】对中国本地经过八代近交的成虫进行培养,进行DNA抽提测序,利用Velvet和SSPACE-STANDARD进行组装拼接,利用RepeatMasker、GeneMark、Augustus和Fgenesh等软件进行基因预测与标记,利用OrthoDB数据库及ClustalW进行比较基因组分析,利用Blast鉴定异色瓢虫基因组中免疫相关基因。【结果】最终拼接的异色瓢虫基因组草图大小为382 Mb,其中contig的N50大小是12.1 kb,scaffold的N50大小是47.5 kb。预测得到18 413个完整基因。根据已发表的昆虫基因组和异色瓢虫基因组的比较基因组分析表明,鞘翅目中含有最多的物种特异性基因,物种进化树显示,所有鞘翅目的物种都聚在了一起,异色瓢虫的位置在赤拟谷盗和蜚蠊之间。通过对异色瓢虫免疫通路相关基因并且进行了进化分析,发现异色瓢虫和其他两种瓢虫在PRR通路、Toll通路和IMD通路中的免疫系统基因上没有显著差异。在异色瓢虫中找到了9个Defensin,7个attacin和10个Coleoptericin,在基因组水平上证实了异色瓢虫抗菌肽家族对比其他两种瓢虫相比存在明显的扩张。【结论】异色瓢虫的AMP家族有了显著的扩张,与其他瓢虫相比,基因数量成倍增长。这一结果也从基因组数据水平上验证了Vilcinskas从转录组数据上推测的结果。更多的AMP基因数量是异色瓢虫得以适应复杂环境的重要保证,是异色瓢虫成为生物入侵物种的免疫基础,表明异色瓢虫在可诱导性天然免疫方面比其他瓢虫有着明显的优势。

关键词:异色瓢虫;基因组;比较基因组;免疫

中图分类号:S433.5

文献标志码:A

开放科学(资源服务)标识码(OSID):

文章编号:1000-2286(2023)01-0045-08



Genome for *Harmonia axyridis* and Immune System Analysis

SHEN Zuhuan,ZHOU Yan*

(School of Life Sciences,Fudan University,Shanghai 200438,China)

Abstract: [Objective] *Harmonia axyridis* is a coccinellid beetle and widely distributed around the world, it was once introduced to the world to control aphids and small insects. However, due to its strong adaptability and reproductive ability, the number of ladybirds has increased dramatically and threatens the local environment. Its genomic data and further analysis, can promote the utilization of *H. axyridis* resources and the control of its biological invasion. [Method] Eight-generation inbred adults in China were cultured and their DNA was

收稿日期:2022-10-30 修回日期:2022-11-10

基金项目:国家重点研发计划项目(2017YFC0907503)

Project supported by National Key Research Program of China(2017YFC0907503)

作者简介:沈族恒,硕士生,orcid.org/0000-0003-2425-0664,shenzuhuan@gmail.com; *通信作者:周雁,副教授,博士,主要从事生物信息学研究,orcid.org/0000-0002-7712-0030,15210700101@fudan.edu.cn。

extracted and sequenced. Velvet and SSPACE-STANDARD were used for assembly. Software RepeatMasker, GeneMark, Augustus and Fgenesh were used for gene prediction and marking. OrthoDB database and ClustalW were used for comparative genomic analysis, and Blast was used to identify immune-related genes in the genome of *H. axyridis*. [Result] The final genome was 382Mb. The N50 of contig and the N50 of the scaffold were 12.1 kb and 47.5 kb respectively. 18 413 complete genes were predicted. Comparative genomic analyses of published insect genomes and *H. axyridis* genomes showed that Coleoptera contained the most species-specific genes, and the species evolutionary tree showed that all Coleoptera species clustered together, with *H. axyridis* lying between *Tribolium castaneum* and *Onthophagus taurus*. It was found that there were no significant differences in the immune system genes of the PRR pathway, Toll pathway and IMD pathway between *H. axyridis* and other two species of ladybugs. The identification of 9 Defensin, 7 attacin, and 10 Coleopteracin in *H. axyridis* confirmed a significant expansion of the *H. axyridis* antimicrobial peptide family at the genomic level compared with the other two species. [Conclusion] AMP family of *H. axyridis* has expanded significantly. Compared with other ladybugs, its number of genes multiplied. This result also confirms Vilcinskis' prediction from transcriptome data at the genomic data level. More AMP gene is an important guarantee for *H. axyridis* to adapt to the complex environment, and it is the immune basis for *H. axyridis* to become a biological invasive species. It shows that *H. axyridis* has obvious advantages over other *H. axyridis* in natural immunity.

Keywords: *Harmonia axyridis*; genome; comparative genomics; immune

【研究意义】异色瓢虫(*Harmonia axyridis*, Pallas)是一种鞘翅目瓢虫科昆虫,大小从0.8 mm到18 mm不等^[1]。异色瓢虫在幼虫时期和成虫时期都可以捕食蚜虫,高龄幼虫每天可捕食100头蚜虫以上,成虫可捕食180头蚜虫以上^[2]。【前人研究进展】Teddners等^[3]报道利用异色瓢捕食蚜虫的生物学特性可以有效控制蚜虫对山核桃树的侵害。异色瓢虫的这一捕食特性也被广泛地运用到农业、园艺、林业生产等诸多方面^[4]。异色瓢虫除了在美国作为生物防治的天敌昆虫被大量利用,在其他国家也被广泛用于生物防治,1982年异色瓢虫被引入到法国^[5],1986年又被引入到南美洲^[6]。异色瓢虫在进入新的栖息地后表现出强烈的侵略性并超越本地其他瓢虫群体成为优势种的特性使其成为生物入侵研究的模式生物^[7]。这种强适应性可能与其拥有区别于其他瓢虫尤其是本地瓢虫的免疫机制有关,异色瓢虫免疫系统表现出对新环境抗原的强抵抗能力^[8]。【本研究切入点】昆虫的主要免疫系统是天然免疫防御系统。天然免疫防御系统受到一定抗原刺激后通过自身系统识别进行反应^[9]。系统的启动需要识别抗原,由PRRs(pattern recognition receptors)激活^[10]。在PRRs蛋白家族中,PGRP(peptidoglycan recognition gene)、 β GRP(β -1,3-glucan recognition gene)主要参与体液免疫识别,Tep(thioester-containing gene)主要参与细胞免疫识别^[11]。由于PRRs的不同蛋白家族具有丰富的遗传多样性,其功能也不尽相同。一些PRR直接参与的免疫效应过程,例如吞噬和黑化,可以激活其他细胞内的信号通路和免疫效应基因的转录,还有一些则会同时激活效应器和信号通路^[12]。信号通路主要可分为以下3种:Toll通路、IMD通路和Jak/Stat通路。Toll信号通路是一种进化上比较保守的信号通路,在昆虫的生长发育和免疫方面都起了非常重要的作用,同时可以有效抵御革兰氏阳性(Gram⁺)细菌、病毒和真菌^[13]。在生物体免疫阶段,PRR会激活胞外细胞因子Spatzle并与细胞受体Toll结合,导致Myd88、Pelle、Tube在胞内聚集^[14]。之后抑制因子Cactus会发生磷酸化,通过水解后与NF- κ B转录因子Dif/Dorsal分开,进而诱导AMP(Anti microbial peptide)及其他效应分子基因的表达^[15]。IMD信号通路对抵御革兰氏阴性(Gram⁻)细菌和病毒起着比较重要的作用,同时也对AMP表达起到调控作用^[16]。在IMD信号通路中,当抗原结合物结合与胞外受体PGRP-LC结合后,会激活Imd通路进行细胞内信号传导。通过IMD/Fadd/Dredd等诱导胞内信号转导,Fadd和Dredd共同将NF- κ B转录因子Relish激活,最后Relish转运到细胞核内,诱导其他效应基因和AMP的表达^[17]。昆虫抗菌肽是一类小分子蛋白质,大多数成员长度在100个氨基酸左右^[18]。有研究^[19]发现异色瓢虫相比于其他两种瓢虫,抗菌肽无论从种类还是数量来说都得到了一定的扩张。而且高度多样化的抗菌肽种类可以使异色瓢虫抵抗各种不同的真菌、细菌和病毒,也是异色瓢虫适应不同环境的重要因素。【拟解决的关键问题】通过

对异色瓢虫基因组测序,对异色瓢虫免疫相关基因进行筛选鉴定,基因组数据和分析结果对关注异色瓢虫领域的学者具有重要的基础意义,通过进一步分析,可以推动对异色瓢虫资源的利用和其生物入侵的治理。

1 材料与方法

1.1 异色瓢虫培养

用于基因组和转录组测序的异色瓢虫样本均取自科云生物河南济源白云实业有限公司,样本均是经过8代近交系培养并喂食以当地蚜虫,在养虫室内集中培养。异色瓢虫卵、1龄幼虫、3龄幼虫、蛹和成虫样本经塑料培养皿保存并运送到上海的实验室,经过光照和室温48 h饥饿处理后,将各个时期的异色瓢虫样本置于-80℃冰箱保存。

1.2 异色瓢虫DNA抽提

基因组测序样本选用8只异色瓢虫成虫,DNA用E.Z.N.A.[®]Insect DNA Extraction Kit进行抽提,需要先除去异色瓢虫身体较为坚硬的外壳,取身体较为柔软的部分在研钵中加液氮进行研磨,加入25 μL蛋白酶和350 μL buffer CTL,混匀之后再用60℃水浴振荡孵育30 min,随后水温调至55℃,过夜。后续步骤按照该试剂盒操作步骤进行。

1.3 文库构建与测序

使用DNA Seq NGS Library Preparation Kit(Illumina)构建350 bp和500 bp文库。之后用Agencourt AMPure XP beads(Beckman)纯化文库。根据Illumina提供的操作步骤构建2 kb和8 kb的长片段双端文库。通过Illumina MiSeq平台对2个大小为350 bp和500 bp的文库进行测序,基因组覆盖度分别为22.72×和49.85×。通过Illumina HiSeq2000平台对3个大小为350 bp,2 kb和8 kb的文库进行测序,基因组覆盖度分别为32.98×,39.82×和26.94×。Ion Torrent平台也被用来测序并得到了4 G数据,基因组覆盖度是9.39×。

1.4 基因组拼接

异色瓢虫基因组通过Velvet^[20]和SSPACE-STANDARD^[21]进行组装拼接。为了降低低质量序列对基因组拼接的影响,首先对测序原始数据通过测序得分进行了过滤筛选,保留的高质量数据作为拼接原始数据。通过Velvet软件将短片段序列切割成k-mer并建立de Bruijn表,之后将所有k-mer连接合并成contigs。根据de Bruijn表来完善所有contigs。

通过SSPACE-STANDARD将长片段双端文库(2 kb、8 kb)从contigs进一步组装到scaffolds。利用短片段数据进行补洞,并根据转录组数据协助完善所有scaffolds。异色瓢虫基因组拼接结果已经被上传到NCBI数据库中,编号是PRJNA437351。

1.5 基因预测与注释

通过RepeatMasker^[22]软件对异色瓢虫基因组进行基因预测与标记。完整的基因预测通过GeneMark^[23]、Augustus^[24]、Fgenesh^[25]3种不同的预测软件分别进行,所有结果使用GLAD程序合并并去除冗余。

将预测好的基因翻译成蛋白质序列,分别用NCBI蛋白质数据库、KEGG蛋白质数据库、Uniprot数据库进行序列比对注释,结果用E值筛选,之后取3个结果中E值最低的作为该基因注释。

通过BUSCO^[26]软件来评价基因组组装和拼接的完整性,选择节肢动物数据库(Arthropoda set),包含了1 066个节肢动物中非常核心并且保守的蛋白序列,该软件可以通过比对基因组与核心蛋白序列的相似性,直接给出报告来作为评价基因组质量的重要指标。

1.6 比较基因组分析

通过OrthoDB^[27]数据库确定了其他12个节肢动物和异色瓢虫做比较基因组分析,其中选了黑脉金斑蝶 *Danaus plexippus*,1种黑腹果蝇 *Drosophila melanogaster*,6种鞘翅目甲虫亚洲天牛 *Anoplophora glabripennis*,马铃薯甲虫 *Leptinotarsa decemlineata*,高山甲虫 *Dendroctonus ponderosae*,赤拟谷盗 *Tribolium castaneum*,蜚螂 *Onthophagus taurus*,白蜡窄吉丁 *Agilus planipennis*,西方蜜蜂 *Apis mellifera*,人虱 *Pediculus humanus*,温带臭虫 *Cimex lectularius*,内华达古白蚁 *Zootermopsis nevadensis*。下载各个物种蛋白质序列后利用OrthoDB软件来计算比较聚类分析。

通过计算得到各个物种的单拷贝基因,将所有物种的单拷贝基因按物种顺序排好之后使用这些单拷贝基因利用 ClustalW^[28]进行多序列比对,比对的结果通过 RAxML^[29]软件构建物种进化树,选用默认参数,自展值(bootstraps)设置为 100。

1.7 免疫相关蛋白注释

下载与异色瓢虫亲缘较近物种的免疫相关蛋白序列,将异色瓢虫基因组蛋白序列,二星瓢虫蛋白序列和七星瓢虫蛋白序列分别建库,利用 Blastp 分别进行相似性比对,排除 positive 小于 40 的序列,其中将比对到的长度(alignment length)与输入序列长度(query length)的比例称之为输入占比。排除输入占比小于 0.5 的序列,剩下保留的序列为候选基因。

将候选基因分别上传到 NCBI 数据库和 InterPro 数据库中,在 NCBI 中选择与非冗余数据库(Nr)比对,*E* 值小于 10⁻⁴⁰,InterPro 预测序列结构域,如果该候选基因在这两个数据库中都确认为对应的功能蛋白,则确认该序列为相应的免疫相关基因。

2 结果与分析

2.1 基因组拼接结果

最终基因组草图大小为 382 Mb(表 1),其中 contig 的 N50 大小是 12.1 kb, scaffold 的 N50 大小是 47.5 kb。异色瓢虫基因组比大多数已经公布的甲壳类基因组大,大多数甲壳类基因组的范围在 163~208 Mb。但异色瓢虫基因组比亚洲天牛基因组小得多,亚洲天牛的基因组大小为 710 Mb。GC 含量为 35%,与其他已测序报道的相关昆虫含量相似,赤拟谷盗的 GC 含量为 33%,黑脉金斑蝶为 31.6%,家蚕为 37.7%。黑腹果蝇相对较高,有 42%。

通过 fgenesh 软件进行基因预测,分别选用 15 个物种当作参考基因组,将每个版本的预测基因分别通过 BUSCO 软件来评价预测基因质量,选取节肢动物核心序列作为数据库,选择评分最高的一个版本,占比 82.1%,之后通过 blast 找出哪些 scaffolds 未预测出基因,把这些 scaffold 再重新预测,得到的预测基因放进之前的版本中,得到最终版本,最终预测基因成功比对到 928(87.1%)条节肢动物核心序列。总共预测得到 18 413 个基因,之后用 blastp 以非冗余蛋白质序列(nr)作为数据库对所有预测基因进行了注释。

表 1 基因组拼接结果
Tab.1 Genomic DNA assembly statistics

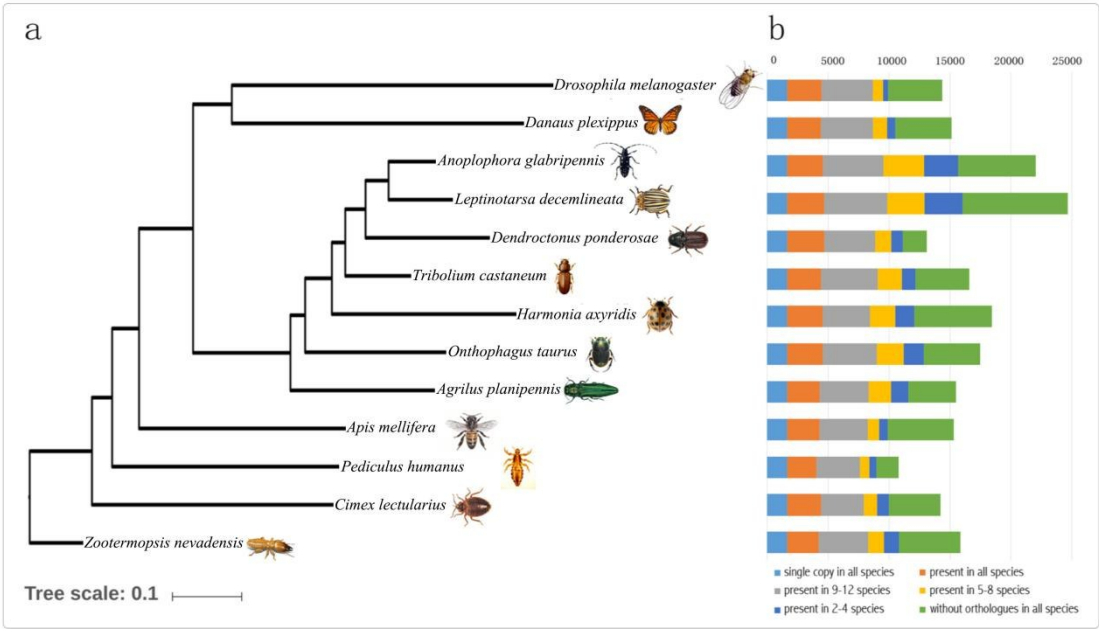
项目 Item	数值 Value
重叠群数量 Contig number	59 118
重叠群 N50/bp Contig N50	12 089
长重叠群数量 Scaffolds number	31 635
长重叠群 N50/bp Scaffold N50	47 483
GC 含量/% GC content	35.00
预测基因个数 Predicted gene number	18 413
基因组组装完整性/% Genome BUSCO	96.0%
预测基因完整性/% Predicted gene BUSCO	87.1%
拼接总长/bp Total assembled length	381 892 561

2.2 比较基因组分析

为研究异色瓢虫与其他节肢动物的差异,使用 OrthoDB Pipeline 软件计算异色瓢虫基因组和其他 12 个节肢动物基因组之间在基因层面的差异(图 1)。其中,一个物种的基因如果可以和其他物种的基因聚类在一起,称之为直系同源组,在这 13 个物种中,一共得到了 15 443 个直系同源组。将其分为两类,将可以和其它 9 种以上物种基因聚类在一起的基因归为一类,在 13 个物种中,马铃薯甲虫得到了最多的直系同源组,共 9 865 个,人虱直系同源组则最少,为 7 632 个,其中有 1 644 个是单拷贝直系同源组(单拷贝基因指基因组中只有一个拷贝,在这里指 13 个物种都能聚在一起的单拷贝基因,总共有 1 644 组)。另一部分定义为特殊的直系同源组,也就是出现在少于 9 个物种的直系同源组,这里面各个物种差异较大,尤其是鞘翅目的几个物种,其中马铃薯甲虫有最多(8 629)的物种特别基因(没有与任何其他物种的基因聚类),

而异色瓢虫有次多(6 377)的物种特别基因。将异色瓢虫的这些物种特别基因与异色瓢虫转录组进行 Blast 比对, e -value 标准选择 10^{-5} 。总共有 4 034(63.3%)条物种特别基因可以在异色瓢虫转录组中找到, 这也说明了这些物种特别基因大部分是在异色瓢虫中真实存在的。

同时将这 13 个物种的 1 644 个单拷贝基因按顺序合并计算构建物种进化树, 利用了 RAxML 软件和最大似然估计方法。进化树以内华达古白蚁为根。通过进化树发现, 所有鞘翅目的物种都聚在了一起。异色瓢虫的位置在赤拟谷盗和蜚螂之间。进化关系依次为白蚁, 臭虫, 人虱, 蜜蜂, 白蜡窄吉丁, 蜚螂, 异色瓢虫, 赤拟谷盗, 高山甲虫, 马铃薯甲虫, 亚洲天牛, 蝴蝶和果蝇。同时也在 NCBI Taxonomy 数据库验证了这一进化关系。



a. 物种进化树是通过 RAxML 软件利用各个物种的 1 644 个单拷贝基因计算得到的, 进化树以白蚁为根。b. 每个物种中都有 1 644 个单拷贝基因, 异色瓢虫有 4 561 个基因出现在其他 12 个物种中, 有 3 886 个基因出现在大多数物种中(9~12), 有 2 070 个基因出现在部分物种中(5~8), 有 1 553 个基因出现在少量物种中(2~4), 同时有 6 377 个基因是异色瓢虫基因组所独有的。

a. The phylogeny was built using RAxML and based on the 1 644 single-copy orthologues that were present in all 13 species. All nodes have 100% ML bootstrap support. The tree was rooted with *Z. nevadensis*. b. There were 1 644 genes that were present as single copies in all 13 species. Another 4 561 of *H. axyridis* genes were found in the other 12 species, while 3 886 were found in the majority of species (in 9–12 species), 2 070 were found in some of the species (in 5–8 species) and 1 553 genes were found in a few species (in 2–4 species). Moreover, 6 377 genes are specific from other 12 species.

图 1 包括异色瓢虫在内的 13 个物种的进化树分析以及比较基因组分析

Fig.1 Phylogenetic placement and orthology comparison among *H. axyridis* and other arthropod species

2.3 天然免疫相关基因注释分析

以 NCBI 中已经注释好的所有昆虫物种的各个 PRR 免疫基因家族成员蛋白作为起始序列, 通过相似性比对验证, 3 种瓢虫基因组的免疫相关基因数量见表 2。PRR 处于免疫应答的启动阶段, 能够识别入侵的病原体。根据基因组比较分析, PRR 家族的主要 7 种基因均可以在 3 种瓢虫基因组中找到, 其中 PGRP、 β GRP/GNBP 及 Galectin 数量完全一致, 其余 4 种基因的数量也大致相等, 表明 3 种瓢虫在免疫识别上没有明显差异。

Toll 通路和 IMD 通路都是通过 PRR 免疫识别后开始激活, 最终合成抗菌肽以抵御细菌侵染。根据基因组比较分析, 除 Toll 通路中 Toll 基因异色瓢虫与其他两种瓢虫有一定差异以外, 其余基因数量均无差异。

通过对 3 种瓢虫 PRR、Toll、IMD 家族基因组注释分析发现, 3 种瓢虫在 PRR 进行免疫, Toll 信号通路激活并合成抗菌肽, IMD 激活并合成抗菌肽方面均为明显差异。认为这不是异色瓢虫能够取得种间优势的原因。

表 2 3 种瓢虫免疫系统相关基因比较
Tab.2 Gene counts for immune-related gene families in three ladybugs

通路 Pathway	家族成员蛋白 Gene family	异色瓢虫 <i>H. axyridis</i>	二星瓢虫 <i>A. bipunctata</i>	七星瓢虫 <i>C. septempunctata</i>
PRR 通路 PRR Pathway	PGRP	8	8	8
	FREP	2	1	3
	Tep	5	3	4
	βGRP/GNBP	5	5	5
	Galectin	4	4	4
	CTL	21	20	19
	DSCAM	7	5	6
TOLL 通路 TOLL Pathway	Spätzle	8	8	8
	Toll	24	22	21
	MyD88	1	1	1
	Tube	1	1	1
	Pelle	1	1	1
	TRAF2	1	1	1
	Cactus	1	1	1
	Dorsal	2	2	2
	IMD	1	1	1
	Dredd	1	1	1
	Tak1	1	1	1
	FADD	1	1	1
	IKK	2	2	2
IMD 通路 IMD Pathway	relish	1	1	1
	Defensin	9	2	3
	attacin	7	4	3
	Coleopteracin	10	3	4

2.4 AMP 家族注释分析

抗菌肽可以在昆虫受到微生物感染或意外伤害时,由血淋巴级机体组织产生的一类多肽,在昆虫体液免疫防御发挥了重要作用。Vilcinskas 等^[19]在转录组水平上发现异色瓢虫抗菌肽数量和其他两种瓢虫相比,异色瓢虫中每一种抗菌肽数量都有显著提高。在基因组水平上证实了这一观点,在异色瓢虫中找到了 9 个 Defensin, 7 个 attacin 和 10 个 Coleopteracin, 3 种抗菌肽在异色瓢虫中跟其他两种瓢虫相比有显著扩张(表 2)。这也表明了异色瓢虫 AMP 的扩张是异色瓢虫能够与其他瓢虫相比有种间优势的重要原因。

3 结论与讨论

3.1 建库策略优化

由于异色瓢虫个体较小,一只瓢虫的 DNA 无法满足长片段文库的测序需要,所以需要多只瓢虫的 DNA 混合测序,而不同个体之间遗传性会有比较大的差异,这就给基因组的拼接组装带来了较大难度。针对这一情况,建立了异色瓢虫近交系,通过蚜虫喂养,让异色瓢虫通过亲代与子代重复杂交若干代,保证了其遗传型相对纯合,由于瓢虫生命周期较短且近交瓢虫存活率较低,为了保证测序样本的质量,最后二代测序试验采用的是异色瓢虫近交 8 代的样本。由于目前三、四代测序技术的更新,在建库方面可以再利用三、四代测序平台补一些测序数据,使基因组拼接结果更加完整。

3.2 基因组拼接优化

由于多只异色瓢虫不可避免地遗传多样性,拼接结果受到了比较大的影响,进行了多次重复组装,最后选择了拼接结果较好的一个版本。首先用短片段原始测序数据 350 bp 和 500 bp reads 进行 contig 拼接。之后利用长片段原始测序数据(2 kb、8 kb)将 contig 连接成 scaffold,删掉含其他物种的污染序列,将

基因组自身利用Blastn比对,删除序列相似度95%以上的scaffold中较短的一条,之后得到最初的拼接版本,但是拼接结果并不理想,scaffold N50值为29.6 kb。其主要原因是多只瓢虫本身的遗传多样性让拼接很难得到比较好的结果。之后对基因组进行了更为严格的质控,删除了大量的短reads序列,同时依次利用其他平台的测序数据对基因组进行补洞优化拼接,成功地将scaffold N50提升到31.5 kb,之后再利用拼接好的转录组数据补洞优化拼接以获得更完整的基因组草图,最终获得了相对较好的基因组拼接结果,异色瓢虫基因组草图scaffold N50为47.5 kb。

3.3 异色瓢虫比较基因组分析

通过对异色瓢虫基因组与其他12个物种进行比较基因组分析,通过进化树发现所选物种进化关系依次为白蚁、臭虫、人虱、蜜蜂、白蜡窄吉丁、蛱螂、异色瓢虫、赤拟谷盗、高山甲虫、马铃薯甲虫、亚洲天牛、蝴蝶和果蝇。所有鞘翅目昆虫全都聚类在一起,其中异色瓢虫在物种进化过程中与赤拟谷盗和蛱螂亲缘关系较近。

而根据直系同源基因聚类分析,本来物种选择上选择了最多种类的鞘翅目,其物种特有基因反而整体较高,这说明鞘翅目昆虫与其他昆虫相比,有着更多的属于自己的特殊基因,通过转录组对这些物种特殊基因进行注释,发现这些基因确实存在,而通过KEGG通路对这些基因进行注释,大多数基因都没有注释结果。从目前来看,对这些物种特殊基因的相关研究非常匮乏,这些基因可能对深入研究鞘翅目昆虫有比较大的意义。

由于异色瓢虫属于鞘翅目,在比较分析时加入了所有已被测序的鞘翅目昆虫,而对于其他昆虫,每个种类只选择了有代表性的一种,在之后随着实验室服务器的升级,可以将其他种类的昆虫增加到2种或3种,进而再次分析物种特殊基因,获得更准确可靠的结果。

3.4 瓢虫免疫通路相关基因研究

通过对异色瓢虫和其他两种瓢虫免疫通路PRR、Toll、IMD的所有相关基因注释和进化分析研究,发现了3种瓢虫的免疫通路相关基因在数量、功能、进化关系上都比较类似,在基因组层面上没有发现显著差异。这说明了异色瓢虫比其他瓢虫更强的免疫能力可能不是主这由免疫通路上的相关基因所致。在之后可以通过基因组比对所得到的相关基因序列,通过试验来论证这一观点。

3.5 异色瓢虫AMP家族显著扩张

异色瓢虫的AMP家族是其抵御外界细菌的重要免疫活性物质,AMP家族种类繁多,功能各异,但是都对个体免疫系统起到了重要的作用。通过注释异色瓢虫AMP家族和其他瓢虫AMP家族,发现异色瓢虫的AMP家族有了显著的扩张,与其他瓢虫相比,基因数量成倍增长。这一结果也从基因组数据水平上验证了Vilcinskis等^[19]从转录组数据上推测的结果。更多的AMP基因数量是异色瓢虫得以适应复杂环境的重要保证,是异色瓢虫成为生物入侵物种的免疫基础,暗示着异色瓢虫在可诱导性天然免疫方面比其他瓢虫有着明显的优势。该结果可为研究异色瓢虫AMP的学者提供相关的序列参考,对于异色瓢虫AMP家族的研究具有重要意义。

致谢:上海市科学技术委员会基金(11JC1400900)同时对本研究给予了资助,谨致谢意!

参考文献 References:

- [1] SEAGO A E, GIORGI J A, LI J, et al. Phylogeny, classification and evolution of ladybird beetles (Coleoptera: Coccinellidae) based on simultaneous analysis of molecular and morphological data[J]. Molecular phylogenetics & evolution, 2011, 60(1): 137-151.
- [2] KOCH R L. The multicolored Asian lady beetle, *Harmonia axyridis*: a review of its biology, uses in biological control, and non-target impacts[J]. Journal of insect science, 2003(1): 32.
- [3] TEDDERS W, SCHAEFER P. Release and establishment of *Harmonia axyridis* (Coleoptera: Coccinellidae) in the southeastern United States[J]. Entomological news, 2004, 105(4): 228-43.
- [4] EVANS E W, DIXON A. Insect predator-prey dynamics: ladybird beetles and biological control[J]. Quarterly review of biology, 2000, 82(3): 244.
- [5] BROWN P M J, THOMAS C E, LOMBAERT E, et al. The global spread of *Harmonia axyridis* (Coleoptera: Coccinellidae):

- distribution, dispersal and routes of invasion[J]. *Biocontrol*, 2011, 56(4): 623-641.
- [6] KOCH R L, GALVAN T L. Bad side of a good beetle: the North American experience with *Harmonia axyridis*[J]. *Biocontrol*, 2008, 53(1): 23-35.
- [7] ROY H, WAJNBERG E. From biological control to invasion: the ladybird *Harmonia axyridis* as a model species[J]. *Biocontrol*, 2008, 53(1): 1-4.
- [8] LAURENT P, BRAEKMAN J C, DALOZE D, et al. In vitro production of adaline and coccinelline, two defensive alkaloids from ladybird beetles (coleoptera: coccinellidae)[J]. *Insect biochemistry and molecular biology*, 2002, 32(9): 1017-1023.
- [9] HILLYER J F. Insect immunology and hematopoiesis[J]. *Developmental & comparative immunology*, 2016, 58: 102-118.
- [10] KEEHNEN N L P, ROLFF J, THEOPOLD U, et al. Insect antimicrobial defences: a brief history, recent findings, biases, and a way forward in evolutionary studies[M]//LIGOXYGAKIS P. *Advances in insect physiology*, London: Academic Press, 2017, 52: 1-33.
- [11] PALMER W J, JIGGINS F M. Comparative genomics reveals the origins and diversity of arthropod immune systems[J]. *Molecular biology & evolution*, 2015, 32(8): 2111-2129.
- [12] CAO X, HE Y, HU Y, et al. The immune signaling pathways of *Manduca sexta*[J]. *Insect biochemistry and molecular biology*, 2015, 62: 64-74.
- [13] CLAYTON A M, DONG Y, DIMOPOULOS G. The Anopheles innate immune system in the defense against malaria infection[J]. *Journal of innate immunity*, 2014, 6(2): 169-181.
- [14] EVANS J D, ARONSTEIN K, CHEN Y P, et al. Immune pathways and defence mechanisms in honey bees *Apis mellifera*. [J]. *Insect molecular biology*, 2010, 15(5): 645-656.
- [15] KINGSOLVER M B, HUANG Z, HARDY R W. Insect antiviral innate immunity: pathways, effectors, and connections[J]. *Journal of molecular biology*, 2013, 425(24): 4921-4936.
- [16] LEMAITRE B, HOFFMANN J. The host defense of *Drosophila melanogaster*[J]. *Annual review of immunology*, 2007, 25(1): 697-743.
- [17] KLEINO A, SILVERMAN N S. The *Drosophila* IMD pathway in the activation of the humoral immune response[J]. *Developmental & comparative immunology*, 2014, 42(1): 25-35.
- [18] JIN F, SUN Q, XU X, et al. cDNA cloning and characterization of the antibacterial peptide cecropin 1 from the diamondback moth, *Plutella xylostella* L. [J]. *Protein expression & purification*, 2012, 85(2): 230-238.
- [19] VILCINSKAS A, MUKHERJEE K, VOGEL H. Expansion of the antimicrobial peptide repertoire in the invasive ladybird *Harmonia axyridis*[J]. *Proceedings of the royal society B: biological sciences*, 2013, 280(1750): 20122113.
- [20] ZERBINO D R, BIRNEY E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs[J]. *Genome research*, 2008, 18(5): 821-829.
- [21] BOETZER M, PIROVANO W. SSPACE-LongRead: scaffolding bacterial draft genomes using long read sequence information[J]. *BMC bioinformatics*, 2014, 15(1): 1-9.
- [22] CHEN N. Using RepeatMasker to identify repetitive elements in genomic sequences[J]. *Current protocols in bioinformatics*, 2004, 5(1): 4-10.
- [23] BESEMER J, BORODOVSKY M. GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses[J]. *Nucleic acids research*, 2005, 33(S2): 451-454.
- [24] STANKE M, TZVETKOVA A, MORGENSTERN B. AUGUSTUS at EGASP: using EST, protein and genomic alignments for improved gene prediction in the human genome[J]. *Genome biology*, 2006, 7(1): 1-8.
- [25] SOLOVYEV V, KOSAREV P, SELEDOV I, et al. Automatic annotation of eukaryotic genes, pseudogenes and promoters[J]. *Genome biology*, 2006, 7(1): 1-12.
- [26] SIMÃO F A, WATERHOUSE R M, IOANNIDIS P, et al. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs[J]. *Bioinformatics*, 2015, 31(19): 3210-3212.
- [27] ZDOBNOV E M, TEGENFELDT F, KUZNETSOV D, et al. OrthoDB v9.1: cataloging evolutionary and functional annotations for animal, fungal, plant, archaeal, bacterial and viral orthologs[J]. *Nucleic acids research*, 2017, 45(1): 744-749.
- [28] THOMPSON J D, GIBSON T J, HIGGINS D G. Multiple sequence alignment using Clustal W and Clustal X[J]. *Current protocols in bioinformatics*, 2002(1): 2-3.
- [29] STAMATAKIS A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies[J]. *Bioinformatics*, 2014, 30(9): 1312-1313.