



中国昆虫基因组学的研究进展^{*}

侯 丽¹ 詹 帅² 周 欣³ 李 飞⁴ 王宪辉^{1**}

(1. 中国科学院动物研究所, 农业虫害鼠害综合治理研究国家重点实验室, 北京 100101;

2. 中国科学院上海生命科学研究院, 植物生理生态研究所, 上海 200032;

3. 中国农业大学, 昆虫学系, 北京 100193; 4. 浙江大学, 昆虫科学研究所, 杭州 310058)

摘 要 随着测序技术的革新和多种组学综合分析方法的应用, 昆虫基因组学研究领域得到了迅速的发展。目前, 我国昆虫学家已经发表了家蚕 *Bombyx mori*、小菜蛾 *Plutella xylostella*、飞蝗 *Locusta migratoria*、榕小蜂 *Ceratosolen solmsi*、褐飞虱 *Nilaparvata lugens*、柑橘凤蝶 *Papilio xuthus*、烟粉虱 *Bemisia tabaci* 等多种昆虫基因组图谱, 同时在比较基因组学、功能基因组学、生物信息学以及基因编辑等研究领域也取得了突出的成绩。本文主要综述了近五年我国昆虫基因组学的重要进展, 并提出了未来在数据分析、团队合作和基因操作等领域面临的挑战。

关键词 全基因组测序, 比较和系统基因组学, 功能基因组学, 生物信息学, 数据库

Advances in research on insect genomics in China

HOU Li¹ ZHAN Shuai² ZHOU Xin³ LI Fei⁴ WANG Xian-Hui^{1**}

(1. State Key Laboratory of Integrated Management of Pest Insects and Rodents, Institute of Zoology, Chinese Academy of Sciences,

Beijing 100101, China; 2. Institute of Plant Physiology & Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy

of Sciences, Shanghai 200032, China; 3. Department of Entomology, China Agricultural University, Beijing 100193, China;

4. Institute of Insect Science, Zhejiang University, Hangzhou 310058, China)

Abstract With the advent of next-generation sequencing, insect genomics has undergone rapid development in recent decades triggered by the systematic analysis of insect genomes, transcriptomes, proteomes, and other cellular components. Several of important insect species have recently been sequenced, including *Bombyx mori*, *Plutella xylostella*, *Locusta migratoria*, *Ceratosolen solmsi*, *Nilaparvata lugens*, *Papilio xuthus*, and *Bemisia tabaci*. Meanwhile, great advances have been made in research on comparative genomics, functional genomics, bioinformatics and databases, and genomic editing. This paper summarizes recent progress in insect genomics over the past five years in China. We also discuss the challenges, opportunities, and emerging trends in insect genomics.

Key words whole genome sequence, comparative and systematic genomics, functional genomics, bioinformatics, database

昆虫作为世界上最大的动物群体, 物种多样性丰富多彩。二代测序技术的快速发展, 已经使得 100 多种昆虫完成全基因组测序。目前, 昆虫基因组学已经成为昆虫学科的重要组成部分, 是当前昆虫学领域发展迅速的前沿分支。近 5 年来, 我国多种重要昆虫全基因组测序完成, 功能基因

组、进化基因组、生物信息学和基因组编辑等领域发展迅速。我国昆虫学家于 2014 年正式成立中国昆虫学会昆虫基因组学专业委员会, 并成功举办三届“国际昆虫基因组学研讨会”。本报告主要以我国 2012—2016 年的代表性研究进展做一概述, 并探讨昆虫基因组学未来的战略发展方向。

*资助项目 Supported projects: 国家自然科学基金面上项目 (31472047); 国家自然科学基金青年基金项目 (31601875)

**通讯作者 Corresponding author, E-mail: Wangxh@ioz.ac.cn

收稿日期 Received: 2017-07-16, 接受日期 Accepted: 2017-09-10

1 昆虫 *De novo* 基因组测序和分析

随着以 Illumina 平台为代表的高通量测序技术飞速发展, 生命科学的研究思路已被彻底改变, 各种组学水平的大数据有力地推进了各领域的发展。我国各昆虫物种的传统优势团队也陆续启动、并完成了一系列重要昆虫物种的基因组项目, 极大推动了我国昆虫学的发展。

1.1 小菜蛾基因组研究

小菜蛾 *Plutella xylostella* 属鳞翅目菜蛾科, 是危害十字花科蔬菜的一种重要的世界性害虫。20 世纪 50 年代发现小菜蛾对双对氯苯基三氯乙烷 (DDT) 产生抗性; 同时小菜蛾还是第一种被发现对苏云金芽孢杆菌 (Bt) 毒素产生抗性的昆虫, 目前已经对所有类型的杀虫剂产生耐药性 (Heckel *et al.*, 1999)。福建农林大学尤民生团队克服了小菜蛾基因组高度杂合的天然属性, 公布了约为 343 Mb 的高质量小菜蛾基因组, 其中 scaffold N50 达到 737 kb, 预测编码 18 071 个基因 (You *et al.*, 2013)。转录组分析显示气味化学感应、食物消化和代谢解毒的相关基因倾向于幼虫期表达; 而硫酸酯酶修饰因子 1 (SUMF1) 和硫代葡萄糖苷硫酸酯酶 (GSS) 存在共表达, 这可能有助于小菜蛾取食十字花科植物。此外, 小菜蛾的 ABC 转运载体、细胞色素 P450 等 4 个解毒基因家族存在明显的基因扩增, 可能在其应对寄主的次级代谢产物中发挥关键作用。这些证据揭示了小菜蛾与十字花科植物协同进化及其抗药性的生物学基础。小菜蛾基因组是第一个鳞翅目害虫的基因组, 它的完成奠定了我国在害虫研究领域的国际领先地位。

1.2 榕小蜂基因组研究进展

榕树-榕小蜂是最古老也是最经典的植物-昆虫共生系统之一。在长期的协同进化过程中, 榕小蜂为了适应榕果内黑暗封闭的生活环境, 在形态结构和生理特性上均发生了一定的变化 (Cook and Rasplus, 2003)。中国科学院动物研究所黄大卫研究员团队通过解析榕小蜂 *Ceratosolen solmsi* 的基因组, 揭示该榕小蜂与对叶榕专性共

性、雌雄异型等特性的分子机制 (Xiao *et al.*, 2013)。通过对 *C. solmsi* 进行基因组深度测序和 *de novo* 组装, 研究人员拼接得到 278 Mb 参考基因组序列, 注释出 11 412 个编码基因。与其它昆虫相比, 榕小蜂的大量基因家族发生收缩。其中化学感受和解毒相关基因显著减少。化学受体在昆虫寄主选择过程中具有重要作用, 因此化学感受基因的极度减少可能与该昆虫成虫期快速准确定位到寄主植物有关。此外, 榕小蜂基因组中三大类解毒基因——谷胱甘肽转移酶、细胞色素 P450 和羧酸酯酶均显著减少, 一方面共生关系决定了榕树不必对榕小蜂产生防御物质, 另一方面密闭的榕果也成为隔绝外界不良环境的天然屏障, 因此, 在长期的进化过程中榕小蜂解毒功能发生退化。扩张的基因家族主要涉及脑形态发生、运动、神经肌肉形成等功能, 这可能与成虫期的寄主选择、产卵、交配等行为有关。与雌虫相比, 雄虫的翅、复眼、触角等发生明显退化, 表现出极端的雌雄异型现象。转录因子差异表达分析表明, 雌虫幼虫期至成虫期上调的转录因子大量富集翅、复眼及触角发育相关基因, 成虫期上调的转录因子还涉及器官分化、生殖器发育和生物节律等, 这些结果进一步证明了榕小蜂雌雄形态和生理上的差异是由基因表达差异造成的。榕小蜂基因组的破译对进一步研究榕树-榕小蜂特殊的协同进化关系具有重要意义。

1.3 飞蝗基因组研究

自古以来蝗灾一直被认为是最具破坏性的自然灾害之一。蝗虫的最显著特征一是长距离飞行; 二是具有密度依赖型的非遗传多型性, 通过群居型与散居型的转变来适应环境, 涉及体色、形态、行为、生理、免疫应答等多个方面 (Pener and Simpson, 2009)。蝗虫的型变由群体密度的增高触发, 可快速发生且可逆转 (Guo *et al.*, 2011)。飞蝗 *Locusta migratoria* 是分布最为广泛的一类迁飞型蝗虫, 长期以来被用作昆虫形态学、行为学和生理学研究的模式生物。中国科学院动物研究所的康乐院士团队 2014 年完成了飞蝗的基因组 *de novo* 测序和组装工作 (Wang *et al.*, 2014b)。飞蝗的基因组大小达到 6.5 Gb,

是人类基因组的两倍多,也是迄今为止测序完成的最大的动物基因组。在克服大基因组的天然难度下,研究团队公布了 scaffold N50 达到 320 kb 的基因组草图,并预测出 17 307 个基因,转录组数据显示 93.8% 的基因都有表达,基本实现了基因组的全覆盖。研究人员发现转座元件占飞蝗基因组的 60%,是导致飞蝗比其它昆虫物种基因组都大的主要原因;同时发现飞蝗发生显著扩张的基因家族主要参与解毒、化学感受、染色体活动和营养代谢等通路。此外,转录组和表观组学的研究表明,脂肪酸代谢的显著扩张可能有助于蝗虫形成更高效的能量供应体系,在飞蝗的长距离飞行中实现集约型的能量消耗。蝗虫基因组的破译将大大助推害虫迁飞、昆虫型变以及生态基因组学的发展,并树立科学家攻克任何困难物种基因组的信心。

1.4 褐飞虱基因组研究

褐飞虱 *Nilaparvata lugens* 专门以水稻韧皮汁液为食,并能传播多种病毒病;同时具有两型转换和迁飞的能力,对生态系统适应性强。因此,褐飞虱对我国粮食安全构成很大威胁。浙江大学张传溪教授团队通过解析褐飞虱和其体内互作的两种共生菌的基因组,在揭示褐飞虱专一食性形成的分子机制以及褐飞虱-共生真菌 YLS-共生细菌三者的依存关系方面取得了系列进展 (Bao *et al.*, 2014; Xue *et al.*, 2014)。研究者通过结合鸟枪法 Fosmid to Fosmid 测序和 Illumina 高通量测序的方法, *de novo* 组装了 1.14 Gb 的褐飞虱基因组参考序列,注释了 27 571 个编码基因。比较基因组分析显示褐飞虱化学感应受体基因家族的基因数目呈现收缩现象,部分解毒酶基因家族以及消化相关基因也存在丢失现象,这一现象可能与褐飞虱取食的专一性相关 (Xi *et al.*, 2014)。对褐飞虱共生真菌 YLS 和内共生细菌的基因组进行了组装与注释,发现褐飞虱缺少的 10 种必需氨基酸合成途径在 YLS 中均能找到;YLS 还可协助褐飞虱宿主本身形成完整的氮素循环和胆固醇合成途径;另外,共生细菌则可以弥补 YLS 和褐飞虱在维生素生物合成途径上的

缺陷 (Huang *et al.*, 2015)。在褐飞虱多型性的分子调控方面,该团队取得了突破性进展,发现胰岛素受体是控制稻飞虱长翅型和短翅型的开关 (Xu *et al.*, 2015)。这些工作不仅阐明了褐飞虱对水稻专食性的分子机制,同时也为通过切断褐飞虱-共生真菌-共生细菌营养途径防治稻飞虱提供了新的视野。

1.5 柑橘凤蝶基因组研究

蝴蝶与人类生活关系密切,并具有特殊的进化地位和生物模型。柑橘凤蝶 *Papilio xuthus* 属鳞翅目凤蝶科,主要分布于我国南部、东南亚、东亚等地,是研究贝氏拟态、性二型的重要模型。在前期已有相关物种基因组公布的不利背景下,中国科学院昆明动物研究所王文团队于 2015 年在著名期刊《自然-通讯》上发表了柑橘凤蝶的基因组研究成果 (Li *et al.*, 2015)。研究团队成功克服了基因组高杂合度的难题,构建了高质量的柑橘凤蝶参考基因组。柑橘凤蝶基因组共 244 Mb,其中 scaffold N50 达 3.4 Mb,CEGMA 基因覆盖率达 99.60%。依据该基因组预测了 15 322 个编码基因。值得一提的是,研究人员成功地将 CRISPR/Cas9 基因编辑技术引入到柑橘凤蝶的研究中,对蝴蝶基因实现高效编辑,使柑橘凤蝶有望成为研究宏观形态进化遗传基础和新基因产生与进化的新模式生物,同时也为非模式物种基因组文章的发表提供了新的范式。

1.6 烟粉虱基因组研究

烟粉虱 *Bemisia tabaci* 是一类遗传多样性丰富,分布广泛的物种复合体,通过取食植物韧皮部,在宿主间传播病毒,进而危害农作物及观赏性植物。由于烟粉虱生殖力强、食性广泛、抗药性显著,该物种已成为极具破坏性的世界性农业害虫。近期国内外的研究团队先后对 *Bemisia* Middle East-Asia Minor 1 (MEAM1/B) 及 Mediterranean (MED/Q) 隐种进行了全基因组测序。康奈尔大学主导的合作团队首先报道了大小为 615 MB 的高质量 MEAM1/B 基因组参考序

列, 注释了 15 644 个编码基因, 并发现基因组中与抗药性、解毒相关的基因家族发生显著的扩张 (Chen *et al.*, 2016)。系统进化分析显示 MEAM1/B 基因组序列与已报道的半翅目昆虫基因组高度分化。随后, 中国农业科学院张友军团队报道了大小为 658 Mb 的 MED/Q 全基因组序列, 并对两个不同的烟粉虱隐种基因组特征进行比较 (Xie *et al.*, 2017)。MED/Q 基因组预测编码 20 786 个基因, 显著多于 MEAM1/B 的编码基因数量。此外, MEAM1/B 及 MED/Q 基因组中均含有大量的转座子元件, 分别占其基因组大小的 44% (269 Mb) 和 40% (265 Mb)。与 MEAM1/B 不同的是, MED/Q 与半翅目昆虫豌豆蚜 *Acyrtosiphon pisum* 处于同一个进化分支, 与含有长红猎蝽 *Rhodnius prolixus* 及褐飞虱 *Nilaparvata lugens* 的分支构成姐妹群关系。与其他两种韧皮部取食的昆虫 *A. pisum* 及 *N. lugens* 相似, MED/Q 基因组中同样包含大量物种特异性基因, 暗示种间特异性基因的数目变异可能与半翅目昆虫的取食特性有关。烟粉虱不同隐种基因组的测定, 为深入研究其宿主选择、抗药性、入侵、传毒等生物学过程的机理奠定了理论基础, 同时为烟粉虱的生物防控提供了重要的遗传信息。

以上有代表性地总结了近年来我国科学家贡献的重要昆虫物种的研究, 除此之外还包括中华按蚊等物种受限于篇幅未展开提及。总之, 基因组研究近年来在我国昆虫学发展中呈暴发式发展格局, 为国内外同行贡献了多种重要昆虫物种的参考基因组信息, 大大提高了我国昆虫学研究的国际地位。以华大基因为首的测序公司在我国基因组研究中发挥了不可磨灭的作用。以上重点提及的几种代表性物种的基因组, 均是国内优势团队联合华大基因共同完成的。除以上介绍, 近年来华大基因还参与了多种蚂蚁、蜜蜂的基因组项目 (Oxley *et al.*, 2014; Kapheim *et al.*, 2015)。在解决了分析瓶颈的前提下, 昆虫基因组项目近年来在我国呈现井喷, 包括斜纹夜蛾、二化螟、棉蚜等多种农业害虫的基因组业已完成或在投稿中。

2 昆虫演化和比较基因学研究

昆虫演化生物学通过研究六足总纲 (Hexapoda) 与外群以及昆虫纲 (Insecta) 内部支系间的系统演化关系, 阐明昆虫的起源, 明确重要性状的演化极性, 为昆虫学研究的各个分支方向提供重要的理论基础。近年来, 高通量测序技术及生物信息分析方法的快速进步推动了昆虫系统演化研究新一轮的突破。包括高阶昆虫系统关系在内的演化研究在近 5 年中取得了重大的突破。我国昆虫学者在运用线粒体基因组进行系统演化研究方面做出了大量的工作, 还首次作为核心成员发起并推动了利用转录组核心同源基因构建六足总纲演化关系的国际合作项目, 成为近年来该领域的标志性工作。另一方面, 昆虫全基因组 *de novo* 序列及基因组重测序被应用于重构昆虫近缘物种间的演化历史以及重要经济物种的种群或品种间的驯化历史。

由于昆虫在长期的演化过程中多次受到大规模灭绝事件的影响, 并在灭绝事件后的短时间内出现物种暴发, 这些因素造成了现生昆虫谱系关系的拓扑结构中含有大量典型的短中间节点, 而末端支系经常受“长枝吸引”效应的影响。这些特殊的拓扑结构使得昆虫系统演化历史的构建非常困难。针对一系列的重要中间节点, 有限的基因位点很难正确还原这些短枝的拓扑关系。早期昆虫分子系统学大量采用核糖体 RNA 序列, 特别是核糖体小亚基 18S 基因片段 (Kjer *et al.*, 2016)。至 2010 年前后, 更多的蛋白编码核基因及线粒体基因被应用于昆虫系统演化的研究中, 而基于高通量测序技术的实验和分析方法也迅速进入昆虫进化研究领域。Meusemann 等 (2010) 使用 RNA EST (Expressed sequence tags) 数据构建了节肢动物系统发育树, 标志着昆虫系统发育组学研究 (Phylogenomics) 时代的开始。昆虫基因组 *de novo* 测序大量展开, 使用基因组中的核心同源基因构建昆虫演化关系成为基因组常规分析中的重要组成部分。例如, Niehuis 等 (2012) 使用基因组数据解决了长期争议的捻翅目昆虫的系统关系的问题, 明确了捻翅目与鞘翅

目 Coleoptera 的姊妹群关系。同年, 两种核基因富集技术的发表, anchored hybrid enrichment (AHE) (Lemmon *et al.*, 2012) 及 ultraconserved elements (UCEs) (Faircloth *et al.*, 2012), 为分子系统学研究者提供了从保存标本中获取全基因组层面基因集的方案。

昆虫系统演化研究在 2014 年出现了重大突破。由任职深圳华大基因研究院的周欣团队与美国罗格斯大学 Karl Kjer 及德国波恩 Alexander Koenig 博物馆的 Bernhard Misof 团队于 2011 年共同发起并推动了“千种昆虫转录组演化”项目 (1 000 Insect Transcriptome Evolution, 1KITE)。该项目旨在通过针对涵盖所有现生昆虫主要支系的 1 000 种代表性物种进行全转录组测序组装, 通过寻找核心同源基因构建昆虫纲的系统演化树, 并结合形态学及化石证据等估算昆虫各类群的进化分歧时间。该项目集合了来自 10 余个国家近 20 家单位的超过 100 名昆虫学者, 依托高通量测序技术和生物信息学分析方法, 使用有史以来最大规模的基因集构建了迄今为止最稳定的昆虫系统发育关系。该项目的首个研究成果于 2014 年 11 月作为封面文章发表在 Science 杂志上 (Misof *et al.*, 2014)。该研究在海量大数据的基础上进行了系统性的分析, 排除了所有已知的数据缺陷 (如局部数据缺失等) 或分析缺陷 (如核酸或氨基酸序列在替换速率上可能存在的歧异性等) 可能造成的人为错误, 这项研究基本确立了六足总纲 (包括昆虫纲) 的系统关系, 解决了若干重要分支节点的关系, 如: 六足总纲的最近外群为桡足类的甲壳纲动物、狭义昆虫纲的外群为双尾纲 Diplura、确立了复新翅类 Polyneoptera、完全变态昆虫 Holometabola 的单系性等。古翅类 Palaeoptera (蜻蜓目+蜉蝣目) 的单系性虽然得到支持, 但 Misof 等 (2014) 在 quartet mapping 的分析中相关数据并不支持该结论, 因此新翅类 Neoptera 的起源仍然留有疑问。该研究根据化石证据, 结合分子系统树, 推断六足总纲的起源时间约为 4.78 亿年前, 昆虫和植物共同塑造了最早的陆生生态系统。研究同时表明昆虫的飞行能力起源于约 4 亿年前, 远远

早于其它动物, 与此同时, 陆生植物也开始大范围地繁衍成森林。该项工作得出的昆虫系统发育关系被广泛引用于 2015—2016 年昆虫学的相关研究中, 包括昆虫各类群的系统关系研究及功能基因的进化研究等, 同时该工作提出的使用转录组构建进化关系的方法也被应用于其它生物类群, 如植物 (Wickett *et al.*, 2014)、鱼类 (Sun *et al.*, 2016) 等。

3 昆虫功能基因组学研究进展

功能基因组学 (Functional genomics) 又被称为后基因组学 (Postgenomics), 它利用结构基因组所提供的信息和产物, 应用新的实验手段, 通过在基因组或系统水平上全面分析基因的功能, 使得生物学研究从对单一基因或蛋白质的研究转向多个基因或蛋白质同时进行系统的研究。这是在厘清基因组静态的碱基序列之后转入对基因组动态的生物学功能学研究。研究内容包括基因功能发现、基因表达分析及突变检测。5 年来, 我国昆虫功能基因组学发展也非常迅速, 在转录组、蛋白质组、代谢组、表观基因组和基因编辑等各个方面都取得了一系列成果。

基因表达谱和转录组分析已经成为功能基因组研究中最常用的技术手段, 为理解生物学现象提供了海量的信息和启发。近 5 年 Pubmed 数据库中, 检索出我国昆虫学领域近 400 多篇论文。研究对象主要集中在家蚕、蜜蜂等资源昆虫, 以及蚊虫和多种农业害虫。在家蚕的研究中, 通过转录组比较分析, 揭示了其耐温、可变剪接调控、免疫、发育等方面的机制 (Shao *et al.*, 2012; Ou *et al.*, 2014; Wang *et al.*, 2014a)。在蜜蜂研究中, 也有大量转录组分析的报道。例如, 蜂王和工蜂幼虫转录组比较 (Chen *et al.*, 2012)、内分泌腺体的转录组分析 (Liu *et al.*, 2014)、嗅觉识别机制 (Wang *et al.*, 2013c)。在蚊子方面的研究, 转录组分析涉及到抗药性机制、免疫等 (Zhu *et al.*, 2014; Wang *et al.*, 2015a)。在重要农业害虫方面, 转录组分析也取得了一系列成果。例如, 揭示了飞蝗转座子、神经、嗅觉受体、两型转变等方面的基因表达模式和机制 (Jiang

et al., 2012; Wang *et al.*, 2013b, 2015b; Hou *et al.*, 2015)。在飞蝗转录组分析中, 揭示了其多个生物学特性如取食、生殖、传毒等的分子调控机制 (Ji *et al.*, 2013; Zhai *et al.*, 2013; Zhao *et al.*, 2016)。

蛋白质组学 (Proteomics) 是一门大规模、高通量、系统化的研究某一类型细胞、组织或体液中的所有蛋白质组成及其功能的新兴学科。蛋白质组学分析是对蛋白质翻译和修饰水平等研究的一种补充, 是全面了解基因组表达的一种必不可少的手段。近 5 年我国学者在该领域发表了大量论文。中国农科院蜜蜂研究所李建科团队在蜜蜂蛋白质组学方面做出系列工作, 阐释了蜜蜂嗅觉识别、发育、免疫、级型分化和蜂王浆成分等 (Fang *et al.*, 2012; Zhang *et al.*, 2012; Gala *et al.*, 2013)。西南大学夏庆友等团队在家蚕蛋白质组学方面也发表了系列论文, 涉及家蚕的消化、免疫、发育等生物学特征的分子调控机制 (Qin *et al.*, 2012; Dong *et al.*, 2013)。另外, 在重要农业害虫中也有相关报道。例如, 通过比较飞蝗两型脑部的蛋白质谱, 揭示了 COP9 蛋白在型变中的重要调控作用 (Tong *et al.*, 2015)。

代谢表型是生物表型的一种直接体现, 某一时空节点下生物体液中代谢物的定性与定量信息能反映生命体当下的生理状态。因此, 代谢组学在复杂生命科学问题研究中呈现出广泛的应用价值。Wu 等 (2012) 应用 HPLC/MS 和 GC/MS 联用的分析手段, 分析了群散两型之间差异代谢物的种类和数量, 发现多种脂类分子在飞蝗两型间差异非常明显, 其中参与脂类氧化的乙酰肉碱在调控行为发生和转变的关键作用。该研究率先将飞蝗表型与代谢组相联系, 极大地加深对维持飞蝗两型转变机制的认识。利用代谢组手段, 发现调控滞育上游的脑激素关键基因是 PTTH, PTTH 通过影响下游的蜕皮激素, 抑制脂肪体的代谢活性, 脂肪体通过减少中间代谢物抑制脑的三羧酸循环, 控制葡萄糖分解的第一个限速酶——己糖激酶基因, 调节糖酵解和三羧酸循环能量产生的多或少, 从而影响脑的代谢活性, 决定个体的发育或滞育, 这表明昆虫的脂肪体与大

脑之间的互作调控了昆虫的滞育, 而三羧酸循环是发育停滞的关键点 (Xu *et al.*, 2012)。

表观遗传现象是指基因表达发生改变但不涉及 DNA 序列的变化, 能够在代与代之间传递。表观基因组学是从基因组水平来研究表观遗传现象的规律和机制, 尤其是在 DNA 和组蛋白修饰、非编码 RNA 等方面, 是当前研究的重点。有关昆虫表观基因组学的工作, 多在果蝇和社会性昆虫如蜜蜂、蚂蚁中开展, 并取得一系列重要进展。

小的非编码 RNA 广泛参与生长发育、细胞凋亡、肿瘤、抗病毒等一系列重要的生命过程, 具有重要的调控功能。我国近年来通过 miRNA 表达谱方面的研究, 取得很多成果, 涉及发育、免疫、可塑性、抗药性、昆植互作等各个领域。发现 miRNA let-7 与生物钟基因的互作, 以及在发育变态中的重要调控作用 (Chen *et al.*, 2014)。Liu 等 (2016) 报道了白纹伊蚊中肠 microRNAs 和转录组的综合分析, 发现登革热 2 型病毒感染后白纹伊蚊的免疫相关基因具有明显的表达差异。Yang 等 (2014, 2016) 发现 miR-133 能够调控飞蝗多巴胺通路上的两个关键基因——*Henna* 和 *Pale*, 进而调控飞蝗两型转变, 并且发现 MiR-71 和 MiR-263 能够分别调控表皮几丁质代谢关键基因几丁质合成酶和几丁质酶, 进而调控飞蝗蜕皮发育。He 等 (2016) 发现 MiR-276 能够促进飞蝗群居型卵发育的一致性, 能够介导母体的跨代遗传效应。

DNA 甲基化和组蛋白修饰是表观遗传学重要的调控机制。近年来昆虫 DNA 甲基化和组蛋白修饰的研究, 主要集中在社会性昆虫中。我国在这两方面仅有少量报道。例如, Wang 等 (2014b) 通过 RRBS 测序, 发现飞蝗散居型和群居型之间在 DNA 甲基化位点上具有明显差异。Guo 等 (2016) 通过系统鉴定飞蝗组蛋白乙酰化及甲基化酶类基因和表达谱分析, 发现飞蝗两型间多个组蛋白修饰酶基因表达发生明显差异。这些工作揭示了飞蝗型变的表观遗传调控基础。此外, 研究发现棉铃虫表观遗传基因 polycomb repressive complex 2 (PRC2) 能够应答光周期的变化, 进而调节 PTTH 基因的组蛋白甲

基化修饰, 导致 PTTH 基因表达的上调或下调, 进而诱导滞育或发育 (Lu *et al.*, 2013)。

基因编辑在基因功能分析方面发挥着重要的作用。当前基因编辑技术包括锌指核酸酶 (ZFNs, zinc-finger nucleases) 技术, 转录激活因子样效应物核酸酶 (TALENs) 技术和成簇的规律间隔短回文重复序列及其相关序列 (CRISPR/Cas) 技术 (Gaj *et al.*, 2013)。前两种技术通过蛋白与 DNA 的相互作用识别特异的基因序列, 而 CRISPR/Cas 系统是由短的向导 RNA 通过碱基配对的方式结合并识别靶标 DNA, 更快速简单, 成本更低, 用于基因组编辑更有优势。近 5 年, 我国在昆虫基因组编辑方面也取得的重要进展。例如, 家蚕中通过显微注射胚胎方法, 已获得多个基因的突变体, 并且能够稳定遗传 (Wang *et al.*, 2013a)。采用转入质粒 DNA, 利用 CRISPR/Cas 系统在家蚕核细胞对 BmKu70 基因进行定向敲除 (Ma *et al.*, 2014)。此外, 在飞蝗 (Li *et al.*, 2016)、小菜蛾 (Huang *et al.*, 2016) 中也已成功应用 CRISPR/Cas9 技术实现靶定基因突变并研究相应的表型变化。这些工作为未来研究我国重要经济昆虫基因功能、代谢通路等奠定了重要基础。

4 昆虫生物信息学与数据库

生物信息学在中国的起步可追寻到 20 世纪 80 年代末, 比国外的研究晚了近 30 年 (Wei and Yu, 2008)。由于我国的生物信息学起步晚, 以及生物信息学研究更加注重在医学和一些模式生物的研究, 目前我国昆虫学生物信息学相对还是薄弱, 从事昆虫生物信息学研究的人才和队伍也比较少的。依据研究方向, 生物信息学可分为 3 个主要部分: (1) 研发有效利用和管理数据的新工具, 构建新平台, 例如构建各种各样的生物信息学数据库; (2) 新算法的开发, 例如各类基因组测序数据的拼接和比对算法等; (3) 生物数据挖掘。

4.1 家蚕组学数据库

西南大学夏庆友团队开发了 SilkDB ([http://](http://silkworm.genomics.org.cn/)

silkworm.genomics.org.cn/) 家蚕数据库, 提供包括家蚕基因的功能注释、基因产物、ESTs 以及表达芯片数据等信息, 同时该网站还提供 BLAST、SilkMap、Wego、BmArray、Clustalw、SMS 等数据分析工具 (Duan *et al.*, 2010)。重庆大学的张泽团队主要构建了家蚕的两个生物信息学数据库。一个是家蚕转座子数据库 BmTEdb (<http://gene.cqu.edu.cn/BmTEdb/>), 目前收录了从基因组中预测得到的共计 1 308 个转座子, 对这些转座子进行了详细的归类 (Zhou *et al.*, 2016)。另一个是家蚕的非编码 RNA 数据库 BmncRNAdb (<http://gene.cqu.edu.cn/BmncRNAdb/index.php>), 从家蚕基因组中共鉴定出 6 281 个 lncRNA 以及 1986 个 microRNA, 数据库提供对家蚕非编码 RNA 的浏览查看、序列检索、序列比对等功能, 还在线提供一个 lncRNA 靶标预测软件 LncTar 和两个 miRNA 靶标预测软件 miRnada、PITA 在线服务 (Xu *et al.*, 2013)。

4.2 蝗虫组学资源数据库

中国科学院动物所康乐团队在 2006 年就开发了蝗虫的 EST 数据库 LocustDB (<http://locustdb.big.ac.cn/>), 目前共收录了 45 474 条 EST 序列, 提供序列的在线查询、比对、下载等服务, 为国内外蝗虫研究提供了重要的数据平台 (Ma *et al.*, 2006)。

4.3 二化螟基因组资源数据库

浙江大学李飞团队构建了世界上首个水稻害虫二化螟基因组数据库 ChiloDB (<http://ento.njau.edu.cn/ChiloDB>)。数据库收录了二化螟的基因组数据, CDS 和蛋白序列 10 221 条, 非编码 RNA miRNA 262 个和 8 万多条 piRNA, 为二化螟研究领域提供了开放的数据平台 (Yin *et al.*, 2014)。

4.4 小菜蛾基因组资源数据库

福建农林大学尤民生团队构建了世界上首个小菜蛾基因组数据库 DBM-DB (<http://www.iae.fafu.edu.cn/DBM/>), 数据库目前收录了小菜

蛾的最新的基因组数据, 包含 1 819 条基因组 scaffold 序列, 基因组 CDS 和蛋白序列 18 071 条, 数据库提供基因的序列比对、查询、可视化及下载服务 (Tang *et al.*, 2014)。

4.5 帝王蝶基因组资源数据库

中国科学院上海生命科学研究院詹帅团队构建了 MonarchBase (<http://monarchbase.umassmed.edu/>), 数据库收录了帝王蝶基因组测序的数据, 数据包含基因组序列、16 866 条基因的序列和蛋白序列, 以及基因家族分析数据、直系同源基因分析、miRNA 等数据, 网站提供这些序列的比对、查询、下载等服务 (Zhan and Reppert, 2013)。

4.6 昆虫基因组及转录组数据库

浙江大学李飞团队构建了昆虫信号通路数据库 iPathDB (<http://ento.njau.edu.cn/ipath/>), 数据库共包含了 52 种昆虫, 12 074 个不同的昆虫信号通路, 98 813 个基因的注释, 414 895 条序列 (Zhang *et al.*, 2014)。该团队还构建了世界上首个昆虫基因组与转录组综合数据库 InsectBase (<http://www.insect-genome.com>), 总数据量已超过 120 GB, 共收集了 155 种昆虫基因组, 116 个转录组数据, 237 个物种的 EST 序列, 69 个物种的 7 544 条 miRNA 序列, 2 个物种的 83 262 条 piRNA 序列, 构建了 78 个物种的 22 536 个信号通路, 116 个昆虫的 UTR 序列和 CDS 序列。针对 61 个有 OGS 注释的昆虫, 开展了数据挖掘。对研究较多的 36 个基因家族开展了系统分析, 运用 OrthoMCL 直系同源算法发现了 7 个物种中的直系同源基因, 共找到 1:1:1 直系同源基因 973 个。功能上提供序列查询、序列比对、基因组可视化、信号通路和注释、进化分析和进化树构建, 所有基因数据均可下载。从 PubMed 中下载了 94 758 昆虫研究相关文献, 通过数据挖掘, 建立了昆虫学领域的关系网络平台 iFacebook (Yin *et al.*, 2016)。此外, 中国农业大学周欣团队围绕转录组数据分析和挖掘, 构建了 1KITE 数据库 (<http://db.cngb.org/1kite/home>), 目前数据库收录了 1 200 多种昆虫的样品、测序信息、

基因注释等数据, 为昆虫转录组数据分析提供了重要的数据平台。

国内昆虫学家在生物信息学的算法开发方面也取的重要的进展。中国科学院动物所康乐院士团队以蝗虫为研究材料, 开发了基于 K-mer 策略的 piRNA 预测流程, 为后续开发和完善 piRNA 预测软件提供了基础 (Zhang *et al.*, 2011)。浙江大学李飞团队开发了一个昆虫基因信号通路构建算法 Insect Pathway Construction (iPathCons), 利用 KEGG 中的信号通路数据为模板, 以同源搜索的方法构建了 15 个基因组序列已知昆虫的基因信号通路。收集和建立了 35 个昆虫的信号通路。以此为模板, 开发了 iPathCons 软件, 可以用于昆虫基因组、转录组等数据的信号通路构建。另外, 该团队还利用 piRNA 与转座子互作形成的二级结构的三联体序列及结构特征, 利用支持向量机训练分类器, 开发 piRNA 的识别算法 Piano, 对二化螟的 piRNA 进行了识别, 发现了 82 639 条 piRNA。

5 发展趋势和展望

国际上昆虫基因组学的研究风起云涌。除了 130 多种昆虫全基因组序列的解析, 上千种昆虫转录组等组学数据在 NCBI 公共数据库中也被释放。还有 “i5K”、“1KITE” 等国际昆虫基因组测序计划, 以及蜂类基因组测序计划、500 种社会昆虫测序计划等不断被提出, 使得昆虫学研究发生了前所未有的深刻变化。这不仅为机理研究提供了序列基础, 也使我们能从更全局的角度认识昆虫生命活动规律。特别是在基因组和转录组基础上, 通过基因功能的研究和挖掘, 应用 CRISPR-Cas9 介导的基因组定点编辑技术和基因 RNAi 技术, 将使害虫的防治和资源昆虫的利用出现革命性的变化。

我国昆虫基因组学的研究, 紧紧跟上了时代的发展, 依赖后发优势, 在多个类群的研究中甚至取得了国际领先地位。未来基因组学研究势必快速渗透到昆虫学各个学科, 将出现一系列交叉学科。我们应加大国内外同行的交流和合作, 共

同促进这一学科的发展,未来应注重生物信息学和大数据技术等,培养交叉学科人才和相关的队伍建设,从而使昆虫基因组学在我国科学研究和国民经济发展中充分发挥应有的作用。

参考文献 (References)

- Bao YY, Qin X, Yu B, Chen LB, Wang ZC, Zhang CX, 2014. Genomic insights into the serine protease gene family and expression profile analysis in the planthopper, *Nilaparvata lugens*. *BMC Genomics*, 15(1): 507.
- Chen WB, Hasegawa DK, Kaur N, Klot A, Pinheiro PV, Luan JB, Stensmyr MC, Zheng Y, Liu WL, Sun HH, Xu YM, Luo Y, Kruse A, Yang XW, Kontsedalov S, Lebedev G, Fisher TW, Nelson DR, Hunter WB, Brown JK, Jander G, Cilia M, Douglas AE, Ghanim M, Simmons AM, Wintermantel WM, Ling KS, Fei ZJ, 2016. The draft genome of whitefly *Bemisia tabaci* MEAM1, a global crop pest, provides novel insights into virus transmission, host adaptation, and insecticide resistance. *BMC Biology*, 14(1): 110.
- Chen WF, Liu ZX, Li TJ, Zhang RF, Xue YB, Zhong Y, Bai WW, Zhou DS, Zhao ZW, 2014. Regulation of *Drosophila* circadian rhythms by miRNA let-7 is mediated by a regulatory cycle. *Nature Communications*, 5: 5549.
- Chen X, Hu Y, Zheng H, Cao L, Niu D, Yu D, Sun Y, Hu S, Hu F, 2012. Transcriptome comparison between honey bee queen- and worker-destined larvae. *Insect Biochem. Mol. Biol.*, 42(9): 665–673.
- Cook JM, Rasplus JY, 2003. Mutualists with attitude: coevolving fig wasps and figs. *Trends in Ecology & Evolution*, 18(5): 241–248.
- Dong Z, Zhao P, Wang C, Zhang Y, Chen J, Wang X, Lin Y, Xia Q, 2013. Comparative proteomics reveal diverse functions and dynamic changes of *Bombyx mori* silk proteins spun from different development stages. *J. Proteome Res.*, 12(11): 5213–5222.
- Duan J, Li RQ, Cheng DJ, Fan W, Zha XF, Cheng TC, Wu YQ, Wang J, Mita K, Xiang ZH, Xia QY, 2010. SilkDB v2.0: a platform for silkworm (*Bombyx mori*) genome biology. *Nucleic Acids Research*, 38(suppl.1): D453–D456.
- Faircloth BC, McCormack JE, Crawford NG, Harvey MG, Brumfield RT, Glenn TC, 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology*, 61(5): 717–726.
- Fang Y, Song F, Zhang L, Aleku DW, Han B, Feng M, Li J, 2012. Differential antennal proteome comparison of adult honeybee drone, worker and queen (*Apis mellifera* L.). *J. Proteomics*, 75(3): 756–773.
- Gaj T, Gersbach CA, Barbas CF, 2013. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in Biotechnology*, 31(7): 397–405.
- Gala A, Fang Y, Wolter DJ, Zhang L, Han B, Feng M, Li J, 2013. Changes of proteome and phosphoproteome trigger embryo-larva transition of honeybee worker (*Apis mellifera ligustica*). *J. Proteomics*, 78: 428–446.
- Guo SY, Jiang F, Yang PC, Liu Q, Wang XH, Kang L, 2016. Characteristics and expression patterns of histone-modifying enzyme systems in the migratory locust. *Insect Biochem. Mol. Biol.*, 76: 18–28.
- Guo W, Wang XH, Ma ZY, Xue LA, Han JY, Yu D, Kang L, 2011. CSP and takeout genes modulate the switch between attraction and repulsion during behavioral phase change in the migratory locust. *PLoS Genet*, 7(2): e1001291.
- He J, Chen Q, Wei Y, Jiang F, Yang M, Hao S, Guo X, Chen D, Kang L, 2016. MicroRNA-276 promotes egg-hatching synchrony by up-regulating brm in locusts. *Proc. Natl. Acad. Sci. U S A*, 113(3): 584–589.
- Heckel DG, Gahan LJ, Liu YB, Tabashnik BE, 1999. Genetic mapping of resistance to *Bacillus thuringiensis* toxins in diamondback moth using biphasic linkage analysis. *Proc. Natl. Acad. Sci. U S A*, 96(15): 8373–8377.
- Hou L, Jiang F, Yang P, Wang X, Kang L, 2015. Molecular characterization and expression profiles of neuropeptide precursors in the migratory locust. *Insect Biochem. Mol. Biol.*, 63: 63–71.
- Huang HJ, Liu CW, Cai YF, Zhang MZ, Bao YY, Zhang CX, 2015. A salivary sheath protein essential for the interaction of the brown planthopper with rice plants. *Insect Biochem. Mol. Biol.*, 66: 77–87.
- Huang Y, Chen Y, Zeng B, Wang Y, James AA, Gurr GM, Yang G, Lin X, Huang Y, You M, 2016. CRISPR/Cas9 mediated knockout of the abdominal-A homeotic gene in the global pest, diamondback moth (*Plutella xylostella*). *Insect Biochem. Mol. Biol.*, 75: 98–106.
- Ji R, Yu H, Fu Q, Chen H, Ye W, Li S, Lou Y, 2013. Comparative transcriptome analysis of salivary glands of two populations of rice brown planthopper, *Nilaparvata lugens*, that differ in virulence. *PLoS ONE*, 8(11): e79612.
- Jiang F, Yang M, Guo W, Wang X, Kang L, 2012. Large-scale transcriptome analysis of retroelements in the migratory locust, *Locusta migratoria*. *PLoS ONE*, 7(7): e40532.
- Kapheim KM, Pan HL, Li C, Salzberg SL, Puiu D, Magoc T, Robertson HM, Hudson ME, Venkat A, Fischman BJ, Hernandez A, Yandell M, Ence D, Holt C, Yocum GD, Kemp WP, Bosch J, Waterhouse RM, Zdobnov EM, Stolle E, Kraus FB, Helbing S, Moritz RFA, Glastad KM, Hunt BG, Goodisman MAD, Hauser F, Grimmelikhuijzen CJP, Pinheiro DG, Nunes FMF, Soares MPM,

- Tanaka ED, Simoes ZLP, Hartfelder K, Evans JD, Barribeau SM, Johnson RM, Massey JH, Southey BR, Hasselmann M, Hamacher D, Biewer M, Kent CF, Zayed A, Blatti C, Sinha S, Johnston JS, Hanrahan SJ, Kocher SD, Wang J, Robinson GE, Zhang GJ, 2015. Genomic signatures of evolutionary transitions from solitary to group living. *Science*, 348(6239): 1139–1143.
- Kjer KM, Simon C, Yavorskaya M, Beutel RG, 2016. Progress, pitfalls and parallel universes: a history of insect phylogenetics. *Journal of the Royal Society Interface*, 13(121): 20160363.
- Lemmon AR, Emme SA, Lemmon EM, 2012. Anchored hybrid enrichment for massively high-throughput phylogenomics. *Systematic Biology*, 61(5): 727–744.
- Li XY, Fan DD, Zhang W, Liu GC, Zhang L, Zhao L, Fang XD, Chen L, Dong Y, Chen Y, Ding Y, Zhao RP, Feng MJ, Zhu YB, Feng Y, Jiang XT, Zhu DY, Xiang H, Feng XK, Li SC, Wang J, Zhang GJ, Kronforst MR, Wang W, 2015. Outbred genome sequencing and CRISPR/Cas9 gene editing in butterflies. *Nature Communications*, 9: 8212.
- Li Y, Zhang J, Chen D, Yang P, Jiang F, Wang X, Kang L, 2016. CRISPR/Cas9 in locusts: Successful establishment of an olfactory deficiency line by targeting the mutagenesis of an odorant receptor co-receptor (Orco). *Insect Biochem. Mol. Biol.*, 79: 27–35.
- Liu H, Wang ZL, Tian LQ, Qin QH, Wu XB, Yan WY, Zeng ZJ, 2014. Transcriptome differences in the hypopharyngeal gland between western honeybees (*Apis mellifera*) and eastern honeybees (*Apis cerana*). *BMC Genomics*, 15: 744.
- Liu YX, Li FX, Liu ZZ, Jia ZR, Zhou YH, Zhang H, Yan H, Zhou XQ, Chen XG, 2016. Integrated analysis of miRNAs and transcriptomes in *Aedes albopictus* midgut reveals the differential expression profiles of immune-related genes during dengue virus serotype-2 infection. *Insect Sci.*, 23(3): 377–385.
- Lu YX, Denlinger DL, Xu WH, 2013. Polycomb Repressive complex 2 (PRC2) protein ESC regulates insect developmental timing by mediating H3K27me3 and activating prothoracicotropic hormone gene expression. *J. Biol. Chem.*, 288(32): 23554–23564.
- Ma S, Chang J, Wang X, Liu Y, Zhang J, Lu W, Gao J, Shi R, Zhao P, Xia Q, 2014. CRISPR/Cas9 mediated multiplex genome editing and heritable mutagenesis of BmKu70 in *Bombyx mori*. *Scientific Report*, 4: 4489.
- Ma ZY, Yu J, Kang L, 2006. LocustDB: a relational database for the transcriptome and biology of the migratory locust (*Locusta migratoria*). *BMC Genomics*, 7: 11.
- Meusemann K, von Reumont BM, Simon S, Roeding F, Strauss S, Kuck P, Ebersberger I, Walz M, Pass G, Breuers S, Achter V, von Haeseler A, Burmester T, Hadrys H, Wagele JW, Misof B, 2010. A phylogenomic approach to resolve the arthropod tree of life. *Molecular Biology and Evolution*, 27(11): 2451–2464.
- Misof B, Liu SL, Meusemann K, Peters RS, Donath A, Mayer C, Frandsen PB, Ware J, Flouri T, Beutel RG, Niehuis O, Petersen M, Izquierdo-Carrasco F, Wappler T, Rust J, Aberer AJ, Aspöck U, Aspöck H, Bartel D, Blanke A, Berger S, Böhm A, Buckley TR, Calcott B, Chen JQ, Friedrich F, Fukui M, Fujita M, Greve C, Grobe P, Gu SC, Huang Y, Jermiin LS, Kawahara AY, Krogmann L, Kubiak M, Lanfear R, Letsch H, Li YY, Li ZY, Li JG, Lu HR, Machida R, Mashimo Y, Kapli P, McKenna DD, Meng GL, Nakagaki Y, Navarrete-Heredia JL, Ott M, Ou YX, Pass G, Podsiadlowski L, Pohl H, Reumont BM, Schütte K, Sekiya K, Shimizu S, Slipinski A, Stamatakis A, Song WH, Su X, Szucsich NU, Tan MH, Tan XM, Tang M, Tang JB, Timelthaler G, Tomizuka S, Trautwein M, Tong XL, Uchifune T, Walz MG, Wiegmann BM, Wilbrandt J, Wipfler B, Wong TKF, Wu Q, Wu GX, Xie YL, Yang SZ, Yang Q, Yeates DK, Yoshizawa K, Zhang Q, Zhang R, Zhang WW, Zhang YH, Zhao J, Zhou CR, Zhou LL, Ziesmann T, Zou SJ, Li YR, Xu X, Zhang Y, Yang HM, Wang J, Wang J, Kjer KM, Zhou X, 2014. Phylogenomics resolves the timing and pattern of insect evolution. *Science*, 346(6210): 763–767.
- Niehuis O, Hartig G, Grath S, Pohl H, Lehmann J, Tafer H, Donath A, Krauss V, Eisenhardt C, Hertel J, Petersen M, Mayer C, Meusemann K, Peter RS, Stadler PF, Beutel RG, Borberg-Bauer E, McKenna DD, Misof B, 2012. Genomic and morphological evidence converge to resolve the enigma of *Strepsiptera*. *Current Biology*, 22(14): 1309–1313.
- Ou J, Deng HM, Zheng SC, Huang LH, Feng QL, Liu L, 2014. Transcriptomic analysis of developmental features of *Bombyx mori* wing disc during metamorphosis. *BMC Genomics*, 15: 820.
- Oxley PR, Ji L, Fetter-Pruneda I, McKenzie SK, Li C, Hu HF, Zhang GJ, Kronauer DJC, 2014. The genome of the clonal raider ant *Cerapachys biroi*. *Current Biology*, 24(4): 451–458.
- Pener MP, Simpson SJ, 2009. Locust phase polyphenism: An update. *Advances in Insect Physiology*, 36: 1–272.
- Qin L, Xia H, Shi H, Zhou Y, Chen L, Yao Q, Liu X, Feng F, Yuan Y, Chen K, 2012. Comparative proteomic analysis reveals that caspase-1 and serine protease may be involved in silkworm resistance to *Bombyx mori* nuclear polyhedrosis virus. *J. Proteomics*, 75(12): 3630–3638.
- Shao W, Zhao QY, Wang XY, Xu XY, Tang Q, Li M, Li X, Xu YZ, 2012. Alternative splicing and trans-splicing events revealed by analysis of the *Bombyx mori* transcriptome. *RNA*, 18(7): 1395–1407.
- Sun Y, Huang Y, Li XF, Baldwin CC, Zhou ZC, Yan ZX, Crandall KA, Zhang Y, Zhao XM, Wang M, Wong A, Fang C, Zhang XH,

- Huang H, Lopez JV, Kilfoyle K, Zhang Y, Orti G, Venkatesh B, Shi Q. 2016. Fish-T1K (Transcriptomes of 1,000 Fishes) Project: large-scale transcriptome data for fish evolution studies. *Gigascience*, 5: 18.
- Tang WQ, Yu LY, He WY, Yang G, Ke FS, Baxter SW, You SJ, Douglas CJ, You MS. 2014. DBM-DB: the diamondback moth genome database. *Database-the Journal of Biological Databases and Curation*, 2014: bat087
- Tong XW, Chen B, Huang LH, Feng QL, Kang L. 2015. Proteomic analysis reveals that COP9 signalosome complex subunit 7A (CSN7A) is essential for the phase transition of migratory locust. *Scientific Reports*, 5: 12542.
- Wang H, Fang Y, Wang L, Zhu W, Ji H, Wang H, Xu S, Sima Y. 2014a. Transcriptome analysis of the *Bombyx mori* fat body after constant high temperature treatment shows differences between the sexes. *Mol. Biol. Rep.*, 41(9): 6039–6049.
- Wang XH, Fang XD, Yang PC, Jiang XT, Jiang F, Zhao DJ, Li BL, Cui F, Wei JN, Ma CA, Wang YD, He J, Luo Y, Wang ZF, Guo XJ, Guo W, Wang XS, Zhang Y, Yang ML, Hao SG, Chen B, Ma ZY, Yu D, Xiong ZQ, Zhu YB, Fan DD, Han LJ, Wang B, Chen YX, Wang JW, Yang L, Zhao W, Feng Y, Chen GX, Lian JM, Li QY, Huang ZY, Yao XM, Lv N, Zhang GJ, Li YR, Wang J, Wang J, Zhu BL, Kang L. 2014b. The locust genome provides insight into swarm formation and long-distance flight. *Nature Communications*, 5: 1–9.
- Wang Y, Li Z, Xu J, Zeng B, Ling L, You L, Chen Y, Huang Y, Tan A. 2013a. The CRISPR/Cas system mediates efficient genome engineering in *Bombyx mori*. *Cell Res.*, 23(12): 1414–1416.
- Wang YD, Yang PC, Cui F, Kang L. 2013b. Altered immunity in crowded locust reduced fungal (*Metarhizium anisopliae*) Pathogenesis. *PLoS Pathog*, 9(1): e1003102.
- Wang YH, Hu Y, Xing LS, Jiang H, Hu SN, Raikhel AS, Zou Z. 2015a. A critical role for CLSP2 in the modulation of antifungal immune response in mosquitoes. *PLoS Pathog*, 11(6): e1004931.
- Wang Z, Yang P, Chen D, Jiang F, Li Y, Wang X, Kang L. 2015b. Identification and functional analysis of olfactory receptor family reveal unusual characteristics of the olfactory system in the migratory locust. *Cell Mol. Life Sci.*, 72(22): 4429–4443.
- Wang ZL, Wang H, Qin QH, Zeng ZJ. 2013c. Gene expression analysis following olfactory learning in *Apis mellifera*. *Mol. Biol. Rep.*, 40(2): 1631–1639.
- Wei LP, Yu J. 2008. Bioinformatics in China: A personal perspective. *Plos Computational Biology*, 4(4): e1000020.
- Wickett NJ, Mirarab S, Nguyen N, Warnow T, Carpenter E, Matasci N, Ayyampalayam S, Barker MS, Burleigh JG, Gitzendanner MA, Ruhfel BR, Wafula E, Der JP, Graham SW, Mathews S, Melkonian M, Soltis DE, Soltis PS, Miles NW, Rothfels CJ, Pokorný L, Shaw AJ, DeGironimo L, Stevenson DW, Surek B, Villarreal JC, Roure B, Philippe H, dePamphilis CW, Chen T, Deyholos MK, Baucom RS, Kutchan TM, Augustin MM, Wang J, Zhang Y, Tian ZJ, Yan ZX, Wu XL, Sun X, Wong GKS, Leebens-Mack J. 2014. Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proc. Natl. Acad. Sci. U S A*, 111(45): E4859–E4868.
- Wu R, Wu Z, Wang X, Yang P, Yu D, Zhao C, Xu G, Kang L. 2012. Metabolomic analysis reveals that carnitines are key regulatory metabolites in phase transition of the locusts. *Proc. Natl. Acad. Sci. U S A*, 109(9): 3259–3263.
- Xi Y, Pan PL, Ye YX, Yu B, Zhang CX. 2014. Chitin deacetylase family genes in the brown planthopper, *Nilaparvata lugens* (Hemiptera: Delphacidae). *Insect Mol. Biol.*, 23(6): 695–705.
- Xiao JH, Yue Z, Jia LY, Yang XH, Niu LH, Wang Z, Zhang P, Sun BF, He SM, Li Z, Xiong TL, Xin W, Gu HF, Wang B, Werren JH, Murphy RW, Wheeler D, Niu LM, Ma GC, Tang T, Bian SN, Wang NX, Yang CY, Wang N, Fu YG, Li WZ, Yi SV, Yang XY, Zhou Q, Lu CX, Xu CY, He LJ, Yu LL, Chen M, Zheng Y, Wang SW, Zhao S, Li YH, Yu YY, Qian XJ, Cai Y, Bian LL, Zhang S, Wang JY, Yin Y, Xiao H, Wang GH, Yu H, Wu WS, Cook JM, Wang J, Huang DW. 2013. Obligate mutualism within a host drives the extreme specialization of a fig wasp genome. *Genome Biology*, 14(12): R141.
- Xie W, Chen CH, Yang ZZ, Guo LT, Yang X, Wang D, Chen M, Huang JQ, Wen YA, Zeng Y, Liu YT, Xia JX, Tian LX, Cui HY, Wu QJ, Wang SL, Xu BY, Li XC, Tan XQ, Ghanim M, Qiu BL, Pan HP, Chu D, Delatte H, Maruthi MN, Ge F, Zhou XP, Wang XW, Wan FH, Du YZ, Luo C, Yan FM, Preisser EL, Jiao XG, Coates BS, Zhao JY, Gao Q, Xia JQ, Yin Y, Liu Y, Brown JK, Zhou XG, Zhang YJ. 2017. Genome sequencing of the sweetpotato whitefly *Bemisia tabaci* MED/Q. *Gigascience*, 6(5): 1–7.
- Xu HE, Zhang HH, Xia T, Han MJ, Shen YH, Zhang Z. 2013. BmTEdb: a collective database of transposable elements in the silkworm genome. *Database-the Journal of Biological Databases and Curation*, 2013: bat055.
- Xu HJ, Xue J, Lu B, Zhang XC, Zhuo JC, He SF, Ma XF, Jiang YQ, Fan HW, Xu JY, Ye YX, Pan PL, Li Q, Bao YY, Nijhout HF, Zhang CX. 2015. Two insulin receptors determine alternative wing morphs in planthoppers. *Nature*, 519(7544): 464.
- Xu WH, Lu YX, Denlinger DL. 2012. Cross-talk between the fat body and brain regulates insect developmental arrest. *Proc. Natl. Acad. Sci. U S A*, 109(36): 14687–14692.
- Xue J, Zhou X, Zhang CX, Yu LL, Fan HW, Wang Z, Xu HJ, Xi Y, Zhu ZR, Zhou WW, Pan PL, Li BL, Colbourne JK, Noda H,

- Suetsugu Y, Kobayashi T, Zheng Y, Liu S, Zhang R, Liu Y, Luo YD, Fang DM, Chen Y, Zhan DL, Lv XD, Cai Y, Wang ZB, Huang HJ, Cheng RL, Zhang XC, Lou YH, Yu B, Zhuo JC, Ye YX, Zhang WQ, Shen ZC, Yang HM, Wang J, Wang J, Bao YY, Cheng JA, 2014. Genomes of the rice pest brown planthopper and its endosymbionts reveal complex complementary contributions for host adaptation. *Genome Biology*, 15(12): 521.
- Yang M, Wang Y, Jiang F, Song T, Wang H, Liu Q, Zhang J, Zhang J, Kang L, 2016. miR-71 and miR-263 jointly regulate target genes chitin synthase and chitinase to control locust molting. *PLoS Genet*, 12(8): e1006257.
- Yang M, Wei Y, Jiang F, Wang Y, Guo X, He J, Kang L, 2014. MicroRNA-133 inhibits behavioral aggregation by controlling dopamine synthesis in locusts. *PLoS Genet*, 10(2): e1004206.
- Yin CL, Liu Y, Liu JD, Xiao HM, Huang SQ, Lin YJ, Han ZJ, Li F, 2014. ChiloDB: a genomic and transcriptome database for an important rice insect pest *Chilo suppressalis*. *Database-the Journal of Biological Databases and Curation*, 2014: bat065.
- Yin CL, Shen GY, Guo DH, Wang SP, Ma XZ, Xiao HM, Liu JD, Zhang Z, Liu Y, Zhang YQ, Yu KX, Huang SQ, Li F, 2016. InsectBase: a resource for insect genomes and transcriptomes. *Nucleic Acids Research*, 44(D1): D801–D807.
- You MS, Yue Z, He WY, Yang XH, Yang G, Xie M, Zhan DL, Baxter SW, Vasseur L, Gurr GM, Douglas CJ, Bai JL, Wang P, Cui K, Huang SG, Li XC, Zhou Q, Wu ZY, Chen QL, Liu CH, Wang B, Li XJ, Xu XF, Lu CX, Hu M, Davey JW, Smith SM, Chen MS, Xia XF, Tang WQ, Ke FS, Zheng DD, Hu YL, Song FQ, You YC, Ma XL, Peng L, Zheng YK, Liang Y, Chen YQ, Yu LY, Zhang YN, Liu YY, Li GQ, Fang L, Li JX, Zhou X, Luo YD, Gou CY, Wang JY, Wang J, Yang HM, Wang J, 2013. A heterozygous moth genome provides insights into herbivory and detoxification. *Nature Genetics*, 45(2): 220–225.
- Zhai Y, Zhang J, Sun Z, Dong X, He Y, Kang K, Liu Z, Zhang W, 2013. Proteomic and transcriptomic analyses of fecundity in the brown planthopper *Nilaparvata lugens* (Stal). *J. Proteome Res.*, 12(11): 5199–5212.
- Zhan S, Reppert SM, 2013. MonarchBase: the monarch butterfly genome database. *Nucleic Acids Research*, 41(D1): D758–D763.
- Zhang L, Fang Y, Li R, Feng M, Han B, Zhou T, Li J, 2012. Towards posttranslational modification proteome of royal jelly. *J. Proteomics*, 75(17): 5327–5341.
- Zhang Y, Wang XH, Kang L, 2011. A k-mer scheme to predict piRNAs and characterize locust piRNAs. *Bioinformatics*, 27(6): 771–776.
- Zhang Z, Yin CL, Liu Y, Jie WC, Lei WJ, Li F, 2014. iPathCons and iPathDB: an improved insect pathway construction tool and the database. *Database-the Journal of Biological Databases and Curation*, 2014: bau105.
- Zhao W, Lu L, Yang P, Cui N, Kang L, Cui F, 2016. Organ-specific transcriptome response of the small brown planthopper toward rice stripe virus. *Insect Biochem. Mol. Biol.*, 70: 60–72.
- Zhou QZ, Zhang BD, Yu QY, Zhang Z, 2016. BmncRNAdb: a comprehensive database of non-coding RNAs in the silkworm, *Bombyx mori*. *BMC Bioinformatics*, 17: 370.
- Zhu G, Zhong D, Cao J, Zhou H, Li J, Liu Y, Bai L, Xu S, Wang MH, Zhou G, Chang X, Gao Q, Yan G, 2014. Transcriptome profiling of pyrethroid resistant and susceptible mosquitoes in the malaria vector, *Anopheles sinensis*. *BMC Genomics*, 15: 448.