



特邀稿件

Invited Review

吴恙, 周腾飞, 赖泽钿, 谢李华, 谢雨谷, 刘培文, 刘通, 杨文强, 金彬彬, 孔翎, 郭怡佳, 赵宜洁, 邓洁琳, 顾金保, 陈晓光. 蚊虫组学研究进展 [J]. 环境昆虫学报, 2020, 42 (4): 789–797.

蚊虫组学研究进展

吴恙, 周腾飞, 赖泽钿, 谢李华, 谢雨谷, 刘培文, 刘通,
杨文强, 金彬彬, 孔翎, 郭怡佳, 赵宜洁, 邓洁琳, 顾金保, 陈晓光*

(南方医科大学公共卫生学院病原生物学系暨广东省热带病防治研究重点实验室, 广州 510515)

摘要: 近年来, 随着高通量测序技术和生物信息学分析技术的成熟和不断革新, 蚊虫基因组学、转录组学、小 RNA 组学也获得了快速发展。迄今为止, 已有包括白纹伊蚊、埃及伊蚊、冈比亚按蚊在内约 22 种媒介蚊虫的基因组被解析报道。不同蚊种的基因组大小差异很大, 且与基因组中的重复序列的多少呈正相关; 蚊基因组的解析和比较基因组的分析有助于探索蚊基因组的结构和功能; 转录组的研究为蚊虫嗅觉、性别决定、胚胎发育等相关基因的研究提供了有效手段; 小 RNA 组的研究揭示了 miRNA 和 piRNA 在蚊媒抗病毒免疫通路中具有重要作用。综上所述, 蚊虫组学研究为防治媒介蚊虫和蚊媒传染病提供了理论基础和数据支撑。

关键词: 蚊; 媒介; 基因组学; 转录组学; 小 RNA 组学; 蚊媒传染病

中图分类号: Q968.1; S433

文献标识码: A

文章编号: 1674–0858 (2020) 04–0789–09

The research progress in the omics of vector mosquitoes

WU Yang, ZHOU Teng-Fei, LAI Ze-Tian, XIE Li-Hua, XIE Yu-Gu, LIU Pei-Wen, LIU Tong, YANG Wen-Qiang, JIN Bin-Bin, KONG Ling, GUO Yi-Jia, ZHAO Yi-Jie, DENG Jie-Lin, GU Jin-Bao, CHEN Xiao-Guang* (Key Laboratory of Prevention and Control for Emerging Infectious Diseases of Guangdong Higher Institutes, Department of Pathogen Biology, School of Public Health, Southern Medical University, Guangzhou 510515, China)

Abstract: Recently, thanks to the novel high-throughput sequencing techniques and bioinformatical methods, the studies on the genomics, transcriptomics, and small RNAomics of mosquitoes developed rapidly. Thus far, the genomes of 22 mosquitoes have been released, including *Aedes albopictus*, *Aedes aegypti*, *Anopheles gambiae* and other important vectors. The genome sizes vary among the different species, which is positively correlated with the percentage of the repeat regions in the genome. The genome studies reveal the function and structure of mosquito genome, and the transcriptomes provide a powerful tool to study on the olfactory, sex determination, and embryonic development of mosquitoes, and the small RNAomics indicate the important roles of miRNA and piRNA in the immunity pathway against virus infection. In conclusion, the studies on mosquito omics provide the theoretical basis and data support for the prevention and control of vector mosquitoes and mosquito-borne diseases.

Key words: Mosquito; vector; genomics; transcriptomics; small RNAomics; mosquito-borne disease

基金项目: 国家自然科学基金 (31830087); 美国 NIH R01 项目 (AI 136850); 广州市与发达国家和港澳台地区合作项目 (201803040006)

作者简介: 吴恙, 男, 1991 年生, 博士, 主要研究方向为生物信息学分析技术和媒介蚊虫的性别决定机制, E-mail: yangwu91@smu.edu.cn

* 通讯作者 Author for correspondence, 陈晓光, 男, 博士, 教授, 主要研究方向为媒介蚊虫及其传染病的防治, E-mail: xgchen@smu.edu.cn

收稿日期 Received: 2020–02–17; 接受日期 Accepted: 2020–03–12

蚊是一类重要的医学昆虫,能吸血传病,是多种疾病如登革热、寨卡病毒病、疟疾、流行性乙型脑炎等的传播媒介,对人类的日常生活和生命健康危害极大。长期以来,由于蚊基因组中高度重复序列的存在,使得蚊虫基因特别是非编码序列的分子克隆和功能分析非常困难。近年来,随着高通量测序和生物信息学技术的成熟和不断革新,蚊虫基因组学、转录组学、小 RNA 组学等也取得了快速发展。本文对蚊组学相关领域的研究进展综述如下。

1 基因组学研究

蚊基因组的解析和比较基因组分析有助于探索蚊基因组的结构和功能。大多数蚊染色体数是 $2n=6$,只有按蚊亚科夏蚊属的 *Chagasia bathana* 染色体数 $2n=8$ (Rao and Rai, 1987)。在常见的媒介蚊虫中,按蚊属 *Anopheles* 拥有 X 和 Y 染色体,这对性染色体的核型是不同的,被称为异态性染色体 (heteromorphic sex chromosomes) (Neafsey *et al.*, 2015); 伊蚊属 *Aedes*、库蚊属 *Culex*、阿蚊属 *Armigeres* 拥有一对核型相同的性染色体,被称为同态性染色体 (homomorphic sex chromosomes) (McKee and Handel, 1993)。蚊染色体有大量的移位和颠换,其结构和组成是蚊生物学性状差异的基础 (Breland, 1961; Rai, 1999; Touns and Hahn, 2010)。近年来,随着第二代测序、第三代测序、HiC 染色体构象捕获、Bionano 光学图谱等生物技术的快速发展和日益成熟,目前已公开发表了 22 种蚊的基因组,其中包括 19 种按蚊属、2 种伊蚊属和 1 种库蚊属。

1.1 蚊基因组组成

蚊基因组具有明显的多样性,表现在相对高比例的单拷贝、中等重复和高度重复的序列,导致蚊种间基因组大小有大约 8 倍的变异。按蚊约为 0.24 ~ 0.29 pg,巨蚊属 *Toxorhynchites* 和煞蚊属 *Sabethes* 基因组为中等大小 (0.62 pg),而库蚊属 0.54 ~ 1.02 pg,脉毛蚊属 *Culiseta* 为 0.92 ~ 1.25 pg,骚扰阿蚊 *Armigeres subalbatus* 和 *Heamagogus equinus* 分别为 1.24 pg 和 1.12 pg; 作为全世界广泛分布的伊蚊,其核 DNA 量变化超过 3 倍; 波利尼西亚的两种伊蚊基因组最小,仅 0.59 pg; *Ochlerotatus zoosophus* 则具有最大的基因组 (1.9 pg) (Besansky and Collins, 1992; Severson and Behura,

2012; Bonizzoni *et al.*, 2013)。一般来说,在进化关系中,基因组大小随着蚊科进化而增加。

在按蚊亚科,大约 60% ~ 80% 的基因组是单拷贝序列 (Neafsey *et al.*, 2015),而库蚊亚科大多数是中等和高度的重复序列 (Nene *et al.*, 2007; Arensburger *et al.*, 2010; Chen *et al.*, 2015; Matthews *et al.*, 2018)。蚊基因组组成的基本形式是短周期性散点式 (short-period interspersion),单拷贝序列长度为 1 000 ~ 2 000 bp,交替出现短的 (200 ~ 600 bp) 和中等长度 (1 000 ~ 4 000 bp) 的重复序列,这在库蚊亚科中是普遍现象 (Nene *et al.*, 2007; Arensburger *et al.*, 2010; Chen *et al.*, 2015; Matthews *et al.*, 2018); 另一种形式是长周期性散点式 (long-period interspersion),长的 ($\geq 5\ 600$ bp) 和极长的 ($\geq 13\ 000$ bp) 重复交替不间断地出现,这在按蚊亚科中是普遍现象 (Neafsey *et al.*, 2015)。

按蚊亚科和库蚊亚科基因组组成的差别对于应用多线染色体物理作图以及应用核糖体分子标志进行分类有重要的影响。在已经研究多线染色体的蚊属中,仅按蚊属具有染色中心 (chromocenter),而伊蚊属、库蚊属、曼蚊属、巨蚊属、直脚蚊属 *Orthopodomyia* 及怀蚊属 *Wyeomyia* 缺少明显的染色中心 (Munstermann *et al.*, 1985; Coluzzi *et al.*, 2002; Campos *et al.*, 2003a; Campos *et al.*, 2003b)。按蚊亚科容易制备高质量的多线染色体,被用于虫种鉴定和物理作图 (Coluzzi *et al.*, 2002)。库蚊亚科的多线染色体不易展开,可能与其基因组大量的重复序列易发生错配有关,可用分裂中期的染色体进行研究 (Campos *et al.*, 2003a)。此外,在按蚊中,用转录间隔区 (internal transcribed spacer, ITS) 进行蚊种鉴定能获得较为稳定的结果 (Paskewitz *et al.*, 1994),而在库蚊和伊蚊,由于存在较多的重复序列,用该标志进行虫种鉴定,不易获得重复稳定的结果 (Porter and Collins, 1991)。

1.2 蚊基因组测序

1.2.1 蚊基因组

冈比按蚊 *An. gambiae* PEST (Pink Eye STandard) 基因组于 2002 年首次公布 (Holt *et al.*, 2002),是 9 株野生株系经人工杂交产生的实验室品系,具有特征性的粉色眼,但该株系已于 2005 年遗失 (Sharakhova *et al.*, 2007)。通过构建该株系的细菌人工染色体 (bacterial artificial

chromosome, BAC) 并进行测序, 以及近年来众多研究者对其基因组不断完善组装, 目前冈比按蚊 PEST 株基因组版本号为 AgamP4, 基因组大小约为 273 Mb, Scaffold N50 为 49.4 Mb, 包括 8 个 scaffolds, 分别为 X 染色体、2 号染色体 L 臂、2 号染色体 R 臂、3 号染色体 L 臂、3 号染色体 R 臂、Y 染色体 (未确定序列具体位置)、线粒体基因组和未知位置的序列 (Holt *et al.*, 2002; Sharakhova *et al.*, 2007)。

埃及伊蚊 *Ae. aegypti* LVP_AGWG (Liverpool Aedes Genome Working Group) 株雄性基因组于 2017 年公布 (Matthews *et al.*, 2018), 主要结合使用了 PacBio 第三代测序、HiC 染色体构象捕获和 Bionano 光学图谱等生物技术, 基因组版本号为 AaegL5, 大小约为 1.278 Gb, Scaffold N50 为 409.8 Mb, 包括 2310 个 scaffolds, 其中最长的 3 个 scaffolds 分别为 1 号、2 号和 3 号染色体。该版本的基因组组装首次报道了埃及伊蚊 1 号染色体中的雄性决定区域 (M-locus), 这段区域富含重

复序列而少有编码基因, 功能和结构类似于 Y 染色体, 与埃及伊蚊雄性性别决定密切相关。

白纹伊蚊 *Ae. albopictus* 又名亚洲虎蚊, 是一种最常见的入侵物种和媒介蚊虫之一, 并在全球有蔓延扩散的风险和趋势, 是寨卡病毒 (Zika virus)、登革病毒 (Dengue virus)、基孔肯雅病毒 (Chikungunya virus) 等病原的传播媒介 (Bonizzoni *et al.*, 2013)。白纹伊蚊佛山株 (Foshan) 是我国学者于 1981 年开始近交培育的实验室品系, 其基因组巨大, 变异度高于 0.5%, 且重复序列非常多。为了攻克这些基因组研究中的技术难题, 我国学者通过纯化基因组背景, 结合全基因组扩增 (whole genome amplification, WGA), 并构建了不同插入长度的测序文库进行了大量测序, 于 2015 年首次公布了白纹伊蚊佛山株基因组, 版本号为 AaloF1, 基因组大小约为 1.923 Gb, 包括 154 782 个 scaffolds (Chen *et al.*, 2015)。

已发布的几种常见媒介蚊虫基因组基本数据归纳于表 1 中。

表 1 常见媒介蚊虫基因组的基本数据
Table 1 The basic information on the genome of major vector mosquitoes

蚊种 Species of mosquitoes	基因注释版本号 Version of genome annotation	基因组 大小 (Mb) Genome size	基因数 Number of genes	编码蛋白数 Number of proteins	Scaffold N50 (Mb)	主要参考文献 References
埃及伊蚊 <i>Ae. aegypti</i>	AaegL5. 2	1279	19 763	14 677	409. 8	Nene <i>et al.</i> , 2007; Matthews <i>et al.</i> , 2018
白纹伊蚊 <i>Ae. albopictus</i>	AaloF1. 2	1923	18 294	17 535	0. 1955	Chen <i>et al.</i> , 2015
致倦库蚊 <i>Cx. quinquefasciatus</i>	CpipJ2. 4	579. 1	19 793	18 965	0. 4868	Arensburger <i>et al.</i> , 2010
冈比按蚊 <i>An. gambiae</i>	AgamP4. 12	273. 1	13 796	13 057	49. 36	Holt <i>et al.</i> , 2002; Neafsey <i>et al.</i> , 2015
中华按蚊 <i>An. sinensis</i>	AsinS2. 5	375. 8	13 204	12 903	0. 5791	Zhou <i>et al.</i> , 2014a
斯氏按蚊 <i>An. stephensi</i>	AsteI2. 3	221. 3	12 189	11 789	1. 591	Jiang <i>et al.</i> , 2014
纯净按蚊 <i>An. merus</i>	AmerM2. 9	288. 0	13 605	13 176	1. 490	Neafsey <i>et al.</i> , 2015
达氏按蚊 <i>An. darlingi</i>	AdarC3. 8	136. 9	11 002	10 553	0. 1151	Marinotti <i>et al.</i> , 2013

不同蚊种的基因组大小差异巨大,这是由于重复序列是蚊虫基因组的主要构成,而且其所占比例与基因组大小呈正相关,比如冈比亚蚊基因组中重复序列比例约为 14% (Neafsey *et al.*, 2015), 致倦库蚊 *Cx. quinquefasciatus* 重复序列比例约为 29% (Arensburger *et al.*, 2010), 埃及伊蚊重复序列比例约为 65% (Matthews *et al.*, 2018), 白纹伊蚊基因组中重复序列比例约为 68% (Chen *et al.*, 2015)。这些基因组中大量的重复序列主要包括多基因家族 (multigene family)、微卫星 (microsatellite)、小卫星 (minisatellite)、核糖体 DNA (ribosomal DNA, rDNA) 和转座子 (transposable element, TE)。

(1) 微卫星: 是指基因组中由短的重复单元 (一般为 1~6 个碱基) 组成的 DNA 串联重复序列, 位于着丝粒和端粒区, 反映位点特异性, 具有高度多态性, 在按蚊亚科中是很好的遗传标志 (Field *et al.*, 1999), 但在库蚊亚科中的应用有限 (Bahnck and Fonseca, 2006)。

(2) 核糖体 DNA: rDNA 由外转录间隔区 (external transcribed spacer, ETS)、18S RNA 基因、内转录间隔区 1 (internal transcribed spacer, ITS1)、5.8S RNA 基因、ITS2、28S RNA 基因, 以及基因间非转录间隔区 (intergenic nontranscribed spacer, IGS) 组成。ITS 和 IGS 具有明显的种间和种内多态性, 是重要的分类鉴定标志 (Walton *et al.*, 1999)。大多数库蚊属 rDNA 位于 1 号染色体; 伊蚊、吸蚊属位于 2 号染色体 (Timoshevskiy *et al.*, 2012); 阿蚊属和竹生杓蚊 *Tripteroides bambusa* 位于 3 号染色体 (Kumar and RAI, 1990); 大多数按蚊属 rDNA 位于 X 染色体, 四环按蚊 *An. quadriannulatus*、米拉按蚊 *An. melas*、纯净按蚊 *An. merus*、四斑按蚊 *An. quadrimaculata* 位于 X 和 Y 染色体 (Collins and Paskewitz, 1996); 而三列骚扰蚊 *Ochlerotatus triseriatus* rDNA 位于 1 号和 3 号染色体上 (Graham *et al.*, 2004)。

(3) 转座子 (TE): 是中等重复 DNA 序列, 具有能够在基因组中移动并自身复制的功能。共有两种类型: 一类是反转座子 (retrotransposons), 通过 RNA 介导的反向转录而实现转座; 另一类直接从 DNA 到 DNA 实现转座。在多种蚊基因组中已经鉴定了大量的转座子, 相关具体内容可以搜索并查询 TEfam 数据库网站 (<https://tefam.biochem.vt.edu>) (Diao *et al.*, 2011)。

1.2.2 比较基因组

通过比较蚊与果蝇或不同蚊种之间基因组组成结构, 有助于发现物种特有的基因、研究蚊虫性别决定机制、鉴定杀虫剂抗性突变位点、分析不同蚊虫传播疾病的能力 (即媒介能量) 及了解蚊发育生物学和系统进化关系。

冈比亚蚊、致倦库蚊、埃及伊蚊和白纹伊蚊分别属于两类不同蚊亚科, 它们所携带传播的病原种类和媒介能量均有显著差异, 生活习性也不尽相同。比较基因组学分析将有助于了解蚊虫在这些生物学特性方面的区别、理解蚊媒病原体感染机制, 这对寻找阻断疾病传播的途径大有裨益。例如, 不同蚊种对血液的不同偏好, 在选择宿主时的行为差异, 传播特定病原体时的个体能力差别, 有些蚊在清晨、傍晚吸食人血, 有些蚊多在夜间活动等。比较埃及伊蚊与冈比亚蚊的基因组有诸多相似之处, 但在基因组整体规模、基因密度及基因家族的构成等方面有所差别。其中, 冈比亚蚊的气味结合蛋白、细胞色素 P450 以及表皮相关的基因数量和种类多于埃及伊蚊, 从基因组水平上显示了这两种蚊的生物学性状差别 (Severson *et al.*, 2004; Manoharan *et al.*, 2013)。

致倦库蚊基因组中约有 18 965 个编码蛋白基因, 比埃及伊蚊多 29%, 比冈比亚蚊多 45%。与两种蚊相比, 致倦库蚊的基因家族数量明显较多, 包括与嗅觉和味觉受体、唾液腺和免疫系统功能等有关的基因。此外, 致倦库蚊与免疫反应相关的基因大约有 500 个, 与伊蚊相似, 但明显少于冈比亚蚊和黑腹果蝇 (Arensburger *et al.*, 2010)。致倦库蚊是多种脑炎病毒及淋巴丝虫的媒介, 其复杂的基因结构有可能提高了其向人类和鸟类传播病毒的能力; 也有些基因可能与对不利或外来有害物的适应性有关, 因为库蚊的孳生地常常是污染严重的环境 (Reddy *et al.*, 2012)。

通过提取和比对不同蚊的上千个单拷贝同源基因并拟合模型、构建系统进化关系, 发现埃及伊蚊与白纹伊蚊大约在 7 140 万年前分化, 伊蚊属与库蚊属大约在 1.79 亿年前分化, 按蚊亚科与库蚊亚科大约在 2.18 亿年前分化, 蚊科的共同祖先与同属于双翅目的果蝇大约在 2.61 亿年前分化 (Chen *et al.*, 2015), 见图 1。

通过比较雌、雄基因组 Illumina 测序数据, 有学者发明了染色体商 (chromosome quotient, CQ) 的方法, 筛选并鉴定了埃及伊蚊的性别决定基因

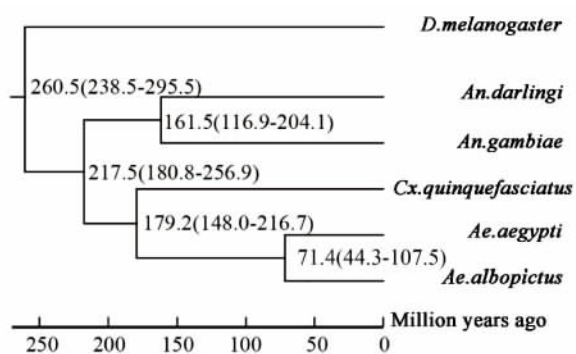


图1 主要的媒介蚊虫与黑腹果蝇的系统进化图

Fig. 1 The phylogeny tree of major vector mosquitoes and *Drosophila melanogaster*

Nix (Hall *et al.*, 2015) 和斯氏按蚊 *An. stephensi* Y 染色体基因 *Guy1* (Criscione *et al.*, 2013; Criscione *et al.*, 2016)。国内有学者使用类似的方法在白纹伊蚊中鉴定出雄性特异性基因 *Nix* (埃及伊蚊 *Nix* 基因的直系同源基因) (Liu *et al.*, 2020)。研究表明, 如果从雄性个体中敲除雄性特异性基因, 会导致雄性个体出现雌性化的性状 (Hall *et al.*, 2015; Liu *et al.*, 2020); 如果将这些雄性决定相关的基因转入雌性个体, 会导致雌性个体表现雄性化性状, 或出现特异性、稳定性的死亡 (Aryan *et al.*, 2019; Qi *et al.*, 2019)。这类研究为防控蚊媒疾病提供了新的思路 (Adelman and Tu, 2016)。

近年来, 随着化学杀虫剂的大量使用, 蚊虫对各类化学杀虫剂的抗性也被广泛报道。通过比较对不同种类化学杀虫剂产生抗性的蚊虫基因组和敏感的蚊虫基因组发现, 按蚊属和库蚊属基因组中编码乙酰胆碱酯酶 (acetylcholinesterase, AChE) 的基因第 119 位密码子由甘氨酸突变为丝氨酸 (G119S), 与蚊虫对有机磷酸酯类 (organophosphate) 和氨基甲酸酯类 (carbamate) 杀虫剂产生抗性相关 (Djogbénou *et al.*, 2010; Camerino *et al.*, 2015; Tmimi *et al.*, 2018); 伊蚊属基因组中编码电压门钠离子通道蛋白 (voltage-gated sodium channel, VGSC) 的基因第 1534 位密码子由苯丙氨酸突变为半胱氨酸 (F1534C), 与蚊虫对拟除虫菊酯类 (pyrethroid) 杀虫剂产生抗性相关 (Chen *et al.*, 2016; Xu *et al.*, 2016)。如果同时存在第 1016 位密码子由缬氨酸突变为异亮氨酸 (V1016I) 或甘氨酸 (V1016G), 则与蚊虫对杀虫剂 DDT 产生抗性相关 (Alvarez *et al.*, 2015; Sombié *et al.*, 2019)。

2 蚊虫转录组学研究

转录组学 (transcriptomics) 是一种生理状态下细胞所能转录出来的所有 mRNA 的总和, 是研究特定生理状态下机体表型和功能的重要手段。传统上用于转录组数据获得和分析的方法主要有基于杂交技术的芯片技术, 包括 cDNA 芯片和寡聚核苷酸芯片, 但目前使用最普遍的是转录组测序技术 (RNA-seq)。基于 Illumina 高通量测序平台的转录组测序技术能够在单核苷酸水平对任意物种的整体转录活性进行检测, 在分析转录本的结构和表达水平的同时, 还能发现未知转录本和低丰度转录本, 精确地识别可变剪切位点以及 cSNP (编码序列单核苷酸多态性), 提供最全面的转录组信息。

斯氏按蚊雄性个体具有 X 和 Y 染色体, 雌性个体具有一对 X 染色体, 因此雄性的 X 染色体基因只有一份拷贝, 雌性的 X 染色体基因有两份拷贝, 但雌性 X 染色体基因和雄性 X 染色体基因的表达水平是相同的, 这种现象被称为剂量补偿效应 (dosage compensation) (Jiang *et al.*, 2015)。斯氏按蚊 X 染色体上有上千个编码基因, 利用高通量的转录组测序的方法可以高效、快速、精确地计算每个编码基因的相对表达水平, 有研究发现在亲本斯氏按蚊常染色体中转入并表达 Y 染色体基因 *Guy1* 会导致雌性子代全部死亡, 通过比较雄性子代和雌性子代转录组数据, 证明了 Y 染色体基因 *Guy1* 是直接启动剂量补偿效应的信号, 异常表达 *Guy1* 基因使雌性 X 染色体基因表达水平被错误上调是导致雌性子代全部死亡的主要原因 (Qi *et al.*, 2019)。

另外, 转录组学分析还常用于蚊虫不同发育阶段、不同器官和不同性别的差异表达基因研究, 例如: 有研究通过比较和分析白纹伊蚊在卵 (egg)、幼虫 (larva)、蛹 (pupa)、雄性成蚊和吸血前后雌性成蚊、感染登革病毒前后等不同生理状态的转录组数据, 筛选得到了与胚胎发育相关基因、与雌性成蚊吸血相关基因、蚊虫感染登革病毒后免疫相关基因等 (Poelchau *et al.*, 2013; Grigoraki *et al.*, 2015; Esquivel *et al.*, 2016); 还有研究通过比较埃及伊蚊的头 (head)、触角 (antenna)、触须 (palp)、吻突 (proboscis)、喙 (rostrum)、腿 (leg)、腹节 (abdomere) 和卵巢

(ovary) 等器官组织的转录组数据,可以得到多种基因的空间表达谱,包括编码气味分子受体 (odorant receptor)、亲离子型受体 (ionotropic receptor, IR) 和味觉受体 (gustatory receptor, GR) 的基因,这些基因被认为与蚊虫搜寻宿主密切相关 (Price *et al.*, 2011; Alfonso-Parra *et al.*, 2016; Matthews *et al.*, 2016)。

转录组学的技术手段是深入研究蚊虫的分子生物学相关领域的强大工具,提供了精确的数字化信号、高效的检测通量和广泛的检测范围,适用于综合测量和计算蚊虫基因相对表达水平,有助于研究者了解蚊虫在不同生理状态下的整体转录活动。

3 蚊虫小 RNA 组学研究

小 RNA (small RNAs) 主要指长度在 18 ~ 30 nt 的一类非编码 RNA (ncRNA)。在真核生物中,具有基因表达调控功能的小 RNA 主要有微小 RNA (microRNA, miRNA)、内源小干扰 RNA (endo-siRNAs) 和 piwi 干扰 RNA (piRNA)。miRNA 和 endo-siRNA 长度主要集中在 20 ~ 24 nt, piRNA 长度集中在 26 ~ 31 nt。miRNA 在动植物和微生物中都普遍存在,据估计一个物种中约 1/3 的基因会受到 miRNA 的调控,大量的实验也表明 miRNA 参与了诸多生命过程的调控,例如细胞周期、细胞分化、组织器官的发生、营养代谢、信号途径以及对外界生物的、非生物环境的反应。piRNA 目前主要在动物的生殖系干细胞、果蝇的卵巢体细胞中被发现 (Klattenhoff and Theurkauf, 2008; Lau *et al.*, 2009), 其主要功能是参与转座子的沉默。以往用于寻找小 RNA 的方法主要有实验克隆法和计算机预测法。实验克隆法可以直接用于鉴定新的小 RNA,是初期发掘小 RNA 的常用方法,不足之处是实验周期较长,对低表达的小 RNA 的发现能力十分有限;计算机预测法多是针对某一已知的小 RNA 特征设计算法,从全基因组或 EST 数据库中快速发掘大量潜在的小 RNA,一定程度上弥补了克隆法的缺点,然而,预测的小 RNA 最终还需要实验证明,而且计算机预测法对新类型小 RNA 的发掘能力十分有限。随着第二代高通量测序技术的问世,小 RNA 测序 (small RNA-Seq) 技术开始逐渐取代原始的小 RNA 发掘法,该技术具有速度快、成本低、覆盖度深等多

方面的优点,对鉴定与发现生命体内的小分子 RNA 及其功能与机理研究具有重要作用。

目前为止,仅冈比亚蚊、斯氏按蚊、埃及伊蚊、致倦库蚊和白纹伊蚊有 miRNA 的鉴定报道 (Winter *et al.*, 2007; Mead and Tu, 2008; Li *et al.*, 2009; Thirugnanasambantham *et al.*, 2013; Biryukova *et al.*, 2014; Liu *et al.*, 2015)。miRNA 功能分析表明,miRNA 对蚊虫的卵巢发育和吸血后的血液消化具有调节作用 (Bryant *et al.*, 2010; Lucas *et al.*, 2015)。另外,病毒感染可以对宿主细胞 miRNA 的表达水平产生巨大影响,这可能与宿主抗病毒机制及病毒入侵后改变细胞内环境有关,雌蚊中 miRNA 的表达模式会随着病原体的感染而发生变化 (Hussain *et al.*, 2013; Zhou *et al.*, 2014b)。国外有学者对登革病毒 (DENV) 编码的 miRNA 或病毒小 RNA (vsRNA) 的进行了功能研究,他们发现 6 个 vsRNA 能通过作用于病毒基因组 RNA 茎环结构中的 5' 和 3' UTR 区,显著增加病毒复制 (Hussain and Asgari, 2014)。中肠屏障是蚊虫防止病原体入侵的一道重要屏障,有研究发现 miR-1174 仅在伊蚊和按蚊的中肠中表达,且雌蚊吸血后其表达量明显上调;而当 miR-1174 表达下调后,蚊虫吸血率明显降低,寿命明显缩短 (Liu *et al.*, 2014)。

国内有研究对白纹伊蚊不同发育时期 (卵、幼虫、蛹、雄蚊、雌蚊、吸血后雌蚊) 的小 RNA 进行了深度测序分析。结果在白纹伊蚊中筛选出 119 条已知的 miRNA 基因,确定了 15 条新的 miRNA 基因,其中 11 条是白纹伊蚊特异的,并且许多 miRNA 仅在特定的发育时期表达:经过实验验证,miR-286、miR-2492 和 miR-1891 分别对白纹伊蚊的卵、幼虫和成虫期特异高效表达,敲低或敲除这些 miRNA 会对蚊虫的生长发育造成显著影响 (Gu *et al.*, 2013; Liu *et al.*, 2015)。这些研究为新型生物杀虫剂的研发提供了靶标。另外还有研究对感染登革病毒前后白纹伊蚊的细胞和成虫的小 RNA 进行了深度测序分析,结果在感染登革病毒的白纹伊蚊中找到了 10 条表达上调的 miRNA 和 11 条表达下调的 miRNA (Skalsky *et al.*, 2010)。通过对这些差异表达 miRNA 的功能分析发现,miR-252 通过与 E 蛋白 3' UTR 区域的结合,对登革病毒的复制起到抑制作用 (Yan *et al.*, 2014); 而 miR-281 则通过与 E 蛋白 5' UTR 区域的结合,对登革病毒的复制具有促进作用。这些

研究为抗登革病毒药物的设计和研发提供了线索和方向 (Zhou *et al.*, 2014b)。

piRNA 来源于转座子、基因间隔区和一些编码蛋白质基因的 3' UTR 区, 对维持基因的完整性和稳定性有一定作用, 最近还有研究证明 piRNA 在抗病毒免疫中也有较大作用: 对蚊虫细胞感染虫媒病毒可以引发 piRNA 通路, 而敲除 piRNA 基因会使病毒滴度增加 (Lucas *et al.*, 2013; Schnettler *et al.*, 2013)。多个 24 ~ 30 nt 与 piwi 相互作用的 RNA 基因组簇可以比对到转座子和蛋白质编码基因的 3' UTR 区, 很多转座子和一些内源性基因的 3' UTR 区会产生大量具有 piRNA 样特征、长度为 29 nt 的小 RNA 峰 (Castellano *et al.*, 2015)。另有研究通过对比缺失 dicer-2 基因的蚊细胞系和野生型蚊细胞系发现, 病毒产生的 piRNA 样小 RNA 可以在病毒产生 siRNA 的过程中调节病毒感染的发生, 这可能是一种蚊虫抗病毒感染的途径 (Morazzani *et al.*, 2012)。

参考文献 (References)

- Adelman ZN, Tu Z. Control of mosquito-borne infectious diseases: Sex and gene drive [J]. *Trends in Parasitology*, 2016, 32 (3): 219 – 229.
- Alfonso-Parra C, Ahmed-Braimah YH, Degner EC, *et al.* Mating-induced transcriptome changes in the reproductive tract of female *Aedes aegypti* [J]. *PLoS Neglected Tropical Diseases*, 2016, 10 (2): e0004451
- Alvarez LC, Ponce G, Saavedra-Rodriguez K, *et al.* Frequency of V1016I and F1534C mutations in the voltage-gated sodium channel gene in *Aedes aegypti* in Venezuela [J]. *Pest Management Science*, 2015, 71 (6): 863 – 869.
- Arensburger P, Megy K, Waterhouse RM, *et al.* Sequencing of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics [J]. *Science*, 2010, 330 (6000): 86 – 88.
- Aryan A, Anderson M, Biedler JK, *et al.* Nix confers heritable sex-conversion in *Aedes aegypti* and myo-sex is needed for male flight [J]. *Biorxiv*, 2019: 595371.
- Bahnck CM, Fonseca DM. Rapid assay to identify the two genetic forms of *Culex* (*Culex*) *pipiens* L. (Diptera: Culicidae) and hybrid populations [J]. *The American Journal of Tropical Medicine and Hygiene*, 2006, 75 (2): 251 – 255.
- Besansky N, Collins F. The mosquito genome: Organization, evolution and manipulation [J]. *Parasitology Today*, 1992, 8 (6): 186 – 192.
- Biryukova I, Ye T, Levashina E. Transcriptome-wide analysis of microRNA expression in the malaria mosquito *Anopheles gambiae* [J]. *BMC Genomics*, 2014, 15 (1): 557.
- Bonizzoni M, Gasperi G, Chen X, *et al.* The invasive mosquito species *Aedes albopictus*: Current knowledge and future perspectives [J]. *Trends in Parasitology*, 2013, 29 (9): 460 – 468.
- Breland OP. Studies on the chromosomes of mosquitoes [J]. *Annals of the Entomological Society of America*, 1961, 54 (3): 360 – 375.
- Bryant B, Macdonald W, Raikhel AS. MicroRNA miR-275 is indispensable for blood digestion and egg development in the mosquito *Aedes aegypti* [J]. *Proceedings of the National Academy of Sciences*, 2010, 107 (52): 22391 – 22398.
- Camerino E, Wong DM, Tong F, *et al.* Difluoromethyl ketones: Potent inhibitors of wild type and carbamate-insensitive G119S mutant *Anopheles gambiae* acetylcholinesterase [J]. *Bioorganic & Medicinal Chemistry Letters*, 2015, 25 (20): 4405 – 4411.
- Campos J, Andrade CFS, Recco-Pimentel SM. Malpighian tubule polytene chromosomes of *Culex quinquefasciatus* (Diptera, Culicinae) [J]. *Memorias do Instituto Oswaldo Cruz*, 2003a, 98 (3): 383 – 386.
- Campos J, Andrade CFS, Recco-Pimentel SM. A technique for preparing polytene chromosomes from *Aedes aegypti* (Diptera, Culicinae) [J]. *Memorias do Instituto Oswaldo Cruz*, 2003b, 98 (3): 387 – 390.
- Castellano L, Rizzi E, Krell J, *et al.* The germline of the malaria mosquito produces abundant miRNAs, endo-siRNAs, piRNAs and 29-nt small RNAs [J]. *BMC Genomics*, 2015, 16 (1): 100.
- Chen H, Li K, Wang X, *et al.* First identification of kdr allele F1534S in VGSC gene and its association with resistance to pyrethroid insecticides in *Aedes albopictus* populations from Haikou City, Hainan Island, China [J]. *Infectious Diseases of Poverty*, 2016, 5 (1): 31.
- Chen X-G, Jiang X, Gu J, *et al.* Genome sequence of the Asian tiger mosquito, *Aedes albopictus*, reveals insights into its biology, genetics, and evolution [J]. *Proceedings of the National Academy of Sciences*, 2015, 112 (44): E5907 – E5915.
- Collins F, Paskewitz S. A review of the use of ribosomal DNA (rDNA) to differentiate among cryptic *Anopheles* species [J]. *Insect Molecular Biology*, 1996, 5 (1): 1 – 9.
- Coluzzi M, Sabatini A, Della Torre A, *et al.* A polytene chromosome analysis of the *Anopheles gambiae* species complex [J]. *Science*, 2002, 298 (5597): 1415 – 1418.
- Ciscione F, Qi Y, Saunders R, *et al.* A unique Y gene in the Asian malaria mosquito *Anopheles stephensi* encodes a small lysine-rich protein and is transcribed at the onset of embryonic development [J]. *Insect Molecular Biology*, 2013, 22 (4): 433 – 441.
- Ciscione F, Qi Y, Tu Z. GUY1 confers complete female lethality and is a strong candidate for a male-determining factor in *Anopheles stephensi* [J]. *Elife*, 2016, 5: e19281.
- Diao Y, Qi Y, Ma Y, *et al.* Next-generation sequencing reveals recent horizontal transfer of a DNA transposon between divergent mosquitoes [J]. *PLoS ONE*, 2011, 6 (2): e16743.
- Djogbénou L, Noel V, Agnew P. Costs of insensitive acetylcholinesterase insecticide resistance for the malaria vector *Anopheles gambiae* homozygous for the G119S mutation [J]. *Malaria Journal*, 2010, 9 (1): 12.

- Esquivel CJ, Cassone BJ, Piermarini PM. A de novo transcriptome of the Malpighian tubules in non-blood-fed and blood-fed Asian tiger mosquitoes *Aedes albopictus*: Insights into diuresis, detoxification, and blood meal processing [J]. *PeerJ*, 2016, 4: e1784.
- Field L, James A, Kamau L, *et al.* Analysis of genetic variability in *Anopheles arabiensis* and *Anopheles gambiae* using microsatellite loci [J]. *Insect Molecular Biology*, 1999, 8 (2): 287–297.
- Graham D, Holmes J, Black Iv W. Identification of quantitative trait loci affecting sex determination in the eastern treehole mosquito (*Ochlerotatus triseriatus*) [J]. *Journal of Heredity*, 2004, 95 (1): 35–45.
- Grigoraki L, Lagnel J, Kioulos I, *et al.* Transcriptome profiling and genetic study reveal amplified carboxylesterase genes implicated in temephos resistance, in the Asian tiger mosquito *Aedes albopictus* [J]. *PLoS Neglected Tropical Diseases*, 2015, 9 (5): e0003771.
- Gu J, Hu W, Wu J, *et al.* MiRNA genes of an invasive vector mosquito, *Aedes albopictus* [J]. *PLoS ONE*, 2013, 8 (7): e67638.
- Hall AB, Basu S, Jiang X, *et al.* A male-determining factor in the mosquito *Aedes aegypti* [J]. *Science*, 2015, 348 (6240): 1268–1270.
- Holt RA, Subramanian GM, Halpern A, *et al.* The genome sequence of the malaria mosquito *Anopheles gambiae* [J]. *Science*, 2002, 298 (5591): 129–149.
- Hussain M, Asgari S. MicroRNA-like viral small RNA from Dengue virus 2 autoregulates its replication in mosquito cells [J]. *Proceedings of the National Academy of Sciences*, 2014, 111 (7): 2746–2751.
- Hussain M, Walker T, O'Neill SL, *et al.* Blood meal induced microRNA regulates development and immune associated genes in the Dengue mosquito vector, *Aedes aegypti* [J]. *Insect Biochemistry and Molecular Biology*, 2013, 43 (2): 146–152.
- Jiang X, Biedler JK, Qi Y, *et al.* Complete dosage compensation in *Anopheles stephensi* and the evolution of sex-biased genes in mosquitoes [J]. *Genome Biology and Evolution*, 2015, 7 (7): 1914–1924.
- Jiang X, Peery A, Hall AB, *et al.* Genome analysis of a major urban malaria vector mosquito, *Anopheles stephensi* [J]. *Genome Biology*, 2014, 15 (9): 459.
- Klattenhoff C, Theurkauf W. Biogenesis and germline functions of piRNAs [J]. *Development*, 2008, 135 (1): 3–9.
- Kumar A, Rai KS. Chromosomal localization and copy number of 18S + 28S ribosomal RNA genes in evolutionarily diverse mosquitoes (Diptera, Culicidae) [J]. *Hereditas*, 1990, 113 (3): 277–289.
- Lau NC, Robine N, Martin R, *et al.* Abundant primary piRNAs, endo-siRNAs, and microRNAs in a *Drosophila* ovary cell line [J]. *Genome Research*, 2009, 19 (10): 1776–1785.
- Li S, Mead EA, Liang S, *et al.* Direct sequencing and expression analysis of a large number of miRNAs in *Aedes aegypti* and a multi-species survey of novel mosquito miRNAs [J]. *BMC Genomics*, 2009, 10 (1): 581.
- Liu P, Jin B, Li X, *et al.* Nix is a male-determining factor in the Asian tiger mosquito *Aedes albopictus* [J]. *Insect Biochemistry and Molecular Biology*, 2020, 118: 103311.
- Liu S, Lucas KJ, Roy S, *et al.* Mosquito-specific microRNA-4174 targets serine hydroxymethyltransferase to control key functions in the gut [J]. *Proceedings of the National Academy of Sciences*, 2014, 111 (40): 14460–14465.
- Liu Y, Zhou Y, Wu J, *et al.* The expression profile of *Aedes albopictus* miRNAs is altered by dengue virus serotype-2 infection [J]. *Cell & Bioscience*, 2015, 5 (1): 16.
- Lucas KJ, Myles KM, Raikhel AS. Small RNAs: A new frontier in mosquito biology [J]. *Trends in Parasitology*, 2013, 29 (6): 295–303.
- Lucas KJ, Roy S, Ha J, *et al.* MicroRNA-8 targets the Wingless signaling pathway in the female mosquito fat body to regulate reproductive processes [J]. *Proceedings of the National Academy of Sciences*, 2015, 112 (5): 1440–1445.
- Manoharan M, Ng Fuk Chong M, Vaitinadapoulé A, *et al.* Comparative genomics of odorant binding proteins in *Anopheles gambiae*, *Aedes aegypti*, and *Culex quinquefasciatus* [J]. *Genome Biology and Evolution*, 2013, 5 (1): 163–180.
- Marinotti O, Cerqueira GC, De Almeida LGP, *et al.* The genome of *Anopheles darlingi*, the main neotropical malaria vector [J]. *Nucleic Acids Research*, 2013, 41 (15): 7387–7400.
- Matthews BJ, Dudchenko O, Kingan SB, *et al.* Improved reference genome of *Aedes aegypti* informs arbovirus vector control [J]. *Nature*, 2018, 563 (7732): 501–507.
- Matthews BJ, McBride CS, Degennaro M, *et al.* The neurotranscriptome of the *Aedes aegypti* mosquito [J]. *BMC Genomics*, 2016, 17 (1): 32.
- Mckee BD, Handel MA. Sex chromosomes, recombination, and chromatin conformation [J]. *Chromosoma*, 1993, 102 (2): 71–80.
- Mead EA, Tu Z. Cloning, characterization, and expression of microRNAs from the Asian malaria mosquito, *Anopheles stephensi* [J]. *BMC Genomics*, 2008, 9 (1): 244.
- Morazzani EM, Wiley MR, Murreddu MG, *et al.* Production of virus-derived ping-pong-dependent piRNA-like small RNAs in the mosquito soma [J]. *PLoS Pathogens*, 2012, 8 (1): 1.
- Munstermann L, Marchi A, Sabatini A, *et al.* Polytene chromosomes of *Orthopodomyia pulcralpalpis* (Diptera, Culicidae) [J]. *Parassitologia*, 1985, 27 (3): 267–277.
- Neafsey DE, Waterhouse RM, Abai MR, *et al.* Highly evolvable malaria vectors: The genomes of 16 *Anopheles* mosquitoes [J]. *Science*, 2015, 347 (6217): 1258522.
- Nene V, Wortman JR, Lawson D, *et al.* Genome sequence of *Aedes aegypti*, a major arbovirus vector [J]. *Science*, 2007, 316 (5832): 1718–1723.
- Paskewitz SM, Wesson D, Collins F. The internal transcribed spacers of ribosomal DNA in five members of the *Anopheles gambiae* species complex [J]. *Insect Molecular Biology*, 1994, 2 (4): 247–257.
- Poelchau MF, Reynolds JA, Elisk CG, *et al.* RNA-Seq reveals early distinctions and late convergence of gene expression between diapause and quiescence in the Asian tiger mosquito, *Aedes albopictus* [J]. *Journal of Experimental Biology*, 2013, 216 (21): 1.

- 4082 – 4090.
- Porter CH, Collins FH. Species-diagnostic differences in a ribosomal DNA internal transcribed spacer from the sibling species *Anopheles freeborni* and *Anopheles hermsi* (Diptera: Culicidae) [J]. *The American Journal of Tropical Medicine and Hygiene*, 1991, 45 (2): 271 – 279.
- Price DP, Nagarajan V, Churbanov A, *et al.* The fat body transcriptomes of the yellow fever mosquito *Aedes aegypti*, pre-and post-blood meal [J]. *PLoS ONE*, 2011, 6 (7): e22573.
- Qi Y, Wu Y, Saunders R, *et al.* Gyl, a Y-linked embryonic signal, regulates dosage compensation in *Anopheles stephensi* by increasing X gene expression [J]. *Elife*, 2019, 8: e43570.
- Rai KS. Mosquito genomes: Structure, organization, and evolution [J]. *Advances in Genetics*, 1999, 41: 1 – 33.
- Rao PN, Rai K. Comparative karyotypes and chromosomal evolution in some genera of nematoceros (Diptera: Nematocera) families [J]. *Annals of the Entomological Society of America*, 1987, 80 (3): 321 – 332.
- Reddy BN, Labbé P, Corbel V. *Culex* genome is not just another genome for comparative genomics [J]. *Parasites & Vectors*, 2012, 5 (1): 63.
- Schnettler E, Donald CL, Human S, *et al.* Knockdown of piRNA pathway proteins results in enhanced Semliki Forest virus production in mosquito cells [J]. *The Journal of General Virology*, 2013, 94 (Pt 7): 1680.
- Severson D, DeBruyn B, Lovin D, *et al.* Comparative genome analysis of the yellow fever mosquito *Aedes aegypti* with *Drosophila melanogaster* and the malaria vector mosquito *Anopheles gambiae* [J]. *Journal of Heredity*, 2004, 95 (2): 103 – 113.
- Severson DW, Behura SK. Mosquito genomics: Progress and challenges [J]. *Annual Review of Entomology*, 2012, 57: 143 – 166.
- Sharakhova MV, Hammond MP, Lobo NF, *et al.* Update of the *Anopheles gambiae* PEST genome assembly [J]. *Genome Biology*, 2007, 8 (1): R5.
- Skalsky RL, Vanlandingham DL, Scholte F, *et al.* Identification of microRNAs expressed in two mosquito vectors, *Aedes albopictus* and *Culex quinquefasciatus* [J]. *BMC Genomics*, 2010, 11 (1): 119.
- Sombié A, Saiki E, Yaméogo F, *et al.* High frequencies of F1534C and V1016I kdr mutations and association with pyrethroid resistance in *Aedes aegypti* from Somgandé (Ouagadougou), Burkina Faso [J]. *Tropical Medicine and Health*, 2019, 47 (1): 2.
- Thirugnanasambantham K, Hairul-Islam VI, Saravanan S, *et al.* Computational approach for identification of *Anopheles gambiae* miRNA involved in modulation of host immune response [J]. *Applied Biochemistry and Biotechnology*, 2013, 170 (2): 281 – 291.
- Timoshevskiy VA, Sharma A, Sharakhov IV, *et al.* Fluorescent in situ hybridization on mitotic chromosomes of mosquitoes [J]. *Journal of Visualized Experiments*, 2012, (67): e4215.
- Tnimi F-Z, Faraj C, Bkhache M, *et al.* Insecticide resistance and target site mutations (G119S ace-1 and L1014F kdr) of *Culex pipiens* in Morocco [J]. *Parasites & Vectors*, 2018, 11 (1): 51.
- Toups MA, Hahn MW. Retrogenes reveal the direction of sex-chromosome evolution in mosquitoes [J]. *Genetics*, 2010, 186 (2): 763 – 766.
- Walton C, Sharpe R, Pritchard S, *et al.* Molecular identification of mosquito species [J]. *Biological Journal of the Linnean Society*, 1999, 68 (1–2): 241 – 256.
- Winter F, Edaye S, Hüttenhofer A, *et al.* *Anopheles gambiae* miRNAs as actors of defence reaction against Plasmodium invasion [J]. *Nucleic Acids Research*, 2007, 35 (20): 6953 – 6962.
- Xu J, Bonizzoni M, Zhong D, *et al.* Multi-country survey revealed prevalent and novel F1534S mutation in voltage-gated sodium channel (VGSC) gene in *Aedes albopictus* [J]. *PLoS Neglected Tropical Diseases*, 2016, 10 (5): e0004696.
- Yan H, Zhou Y, Liu Y, *et al.* MiR-252 of the Asian tiger mosquito *Aedes albopictus* regulates dengue virus replication by suppressing the expression of the dengue virus envelope protein [J]. *Journal of Medical Virology*, 2014, 86 (8): 1428 – 1436.
- Zhou D, Zhang D, Ding G, *et al.* Genome sequence of *Anopheles sinensis* provides insight into genetics basis of mosquito competence for malaria parasites [J]. *BMC Genomics*, 2014a, 15 (1): 42.
- Zhou Y, Liu Y, Yan H, *et al.* MiR-281, an abundant midgut-specific miRNA of the vector mosquito *Aedes albopictus* enhances dengue virus replication [J]. *Parasites & Vectors*, 2014b, 7 (1): 488.