



昆虫长链非编码 RNA 的研究进展*

肖花美^{1**} 李美珍² 李 飞^{2**}

(1. 宜春学院生命科学与资源环境学院, 宜春 336000; 2. 浙江大学昆虫科学研究所, 杭州 310058)

摘 要 长链非编码 RNA (lncRNA) 是一类长度大于 200 nt 的非编码 RNA 基因, 在生物体内具有重要的生物学功能, 近年来逐渐引起人们的重视。利用生物信息学预测流程, 可以从 RNA-seq 测序数据中发现大量的 lncRNA 基因。迄今为止, 已在果蝇、家蚕、褐飞虱等 11 种昆虫中发现了 47 049 个 lncRNA 基因。目前 lncRNA 的分类标准尚未统一, 有的根据其在基因组上的位置进行分类, 有的根据其长度大小进行分类, 也有的根据其功能进行分类。研究发现, 大量的 lncRNA 在果蝇形态转变的关键时期、小菜蛾抗药性品系、褐飞虱高繁殖力种群及埃及伊蚊感病毒细胞中高水平表达, 显示其在昆虫变态发育、抗药性形成、繁殖力和抗病毒中具有重要的功能。lncRNA 还参与调控昆虫剂量补偿效应、睡眠行为、性别决定等重要生命活动过程。本文简要综述了 lncRNA 的分类、预测、昆虫 lncRNA 的发现、昆虫 lncRNA 功能研究的最新进展。

关键词 昆虫, 长链非编码 RNA, 预测, 功能

Progress in research on long, non-coding, insect RNA

XIAO Hua-Mei^{1**} LI Mei-Zhen² LI Fei^{2**}

(1. College of Life Sciences and Resource Environment, Yichun University, Yichun 336000, China;

2. Institute of Insect Sciences, Zhejiang University, Hangzhou 310058, China)

Abstract Long non-coding RNA (lncRNA) is a class of non-coding RNA gene with a length > 200 nt. lncRNA has important biological functions *in vivo*, and has attracted increasing research interest in recent years. There are currently no universal classification criteria for lncRNA which is classified on the basis of length, genome location, or function. Bioinformatics prediction pipelines for lncRNA are well established and a huge number of lncRNA have been identified from the RNA-seq data. So far, 47 049 lncRNA genes have been identified from 11 insects, including *Drosophila melanogaster*, *Bombyx mori* and *Nilaparvata lugens*. Some lncRNAs have been found to be highly expressed during the critical period of morphological change in *D. melanogaster*, the insecticide-resistant strain of the diamondback moth *Plutella xylostella*, a high fecundity population of *Nilaparvata lugens*, and dengue virus serotype 2 infected cells of *Aedes aegypti*. This suggests that lncRNA may perform important roles in regulating metamorphosis, development, insecticide resistance, viral resistance, and reproduction. It has also been reported that lncRNAs are involved in the insect dosage compensation effect, sleep behavior and sex determination. Here, we briefly review progress in research on the classification, identification, discovery, and function of lncRNAs.

Key words insect, lncRNA, prediction, function analysis

2003 年, 人类基因组百科全书 (Encyclopedia of DNA Elements, ENCODE) 计划启动, 致力

于研究人类基因组中的所有功能组件。研究发现, 超过 80% 的基因组具有生理活性, 93% 的基

*资助项目 Supported projects: 国家重点研发计划 “2016YFC1200602”; 宜春学院科研启动基金

**通讯作者 Corresponding author, E-mail: xiaohuamei625@163.com; lifei18@zju.edu.cn

收稿日期 Received: 2017-01-03, 接受日期 Accepted: 2017-01-09

因组被转录成 RNA, 打破了“垃圾 DNA”的理论 (International Human Genome Sequencing, 2004; Consortium *et al.*, 2007; Consortium, 2012; Pennisi, 2012)。超高比例的 RNA 序列不翻译成蛋白质, 被称为非编码 RNA (non-coding RNA, ncRNA), 研究发现其具有重要的生物学功能 (Huttenhofer *et al.*, 2005; Mattick and Makunin, 2006)。

非编码 RNA 种类繁多, 包括核糖体 RNA (Ribosomal RNA, rRNA)、转运 RNA (Transfer RNA, tRNA)、核小 RNA (Small nuclear RNA, snRNA)、核仁小 RNA (Small nucleolar RNA, snoRNA)、内源性小干扰 RNA (Endogenous small interfering RNA, siRNA)、与 PIWI 蛋白相作用的 RNA (PIWI-interacting RNA, piRNA)、微小 RNA (microRNA, miRNA)、长链非编码 RNA (Long non-coding RNA, lncRNA)、外显子-内含子环形 RNA (EIciRNA) 及许多功能未知的非编码 RNA (Surono *et al.*, 1999; Mattick and Makunin, 2006; Lauressergues *et al.*, 2015)。

1 lncRNA 基因的发现与分类

lncRNA 是一类长度大于 200 nt 的不编码蛋白的 RNA 基因, 于 20 世纪 90 年代初被发现, 研究人员发现 *H19* 和 *Xist* 两个 lncRNA 具有特异性调控作用, 参与表观遗传的调控 (Brannan *et al.*, 1990; Brockdorff *et al.*, 1992; Brown *et al.*, 1992)。通过对小鼠大规模测序的 60 770 条全长 cDNA 序列分析, 获得了 4 280 个 lncRNA 基因 (Okazaki *et al.*, 2002)。这些 lncRNA 基因没有长阅读框, 但加工过程类似于蛋白质编码基因, 由 RNA 聚合酶 转录, 发生剪接, 具有 5' 帽子结构和 3' PloyA 尾巴, 最后形成的基因结构类似于 mRNA (Guttman *et al.*, 2009)。由于 lncRNA 的低丰度及物种间的序列不保守, 一度被认为是由于 RNA 聚合酶的低准确性而导致的“转录噪音”, 不具有生物学功能 (Huttenhofer *et al.*, 2005)。

随着人类基因组 ENCODE 计划和模式生物 modENCODE 计划的进行, 越来越多物种的 lncRNA 被发现 (Ulitsky and Bartel, 2013), 表

明 lncRNA 并不是“转录噪音”, 其广泛存在于真核生物中, 作为功能元件在基因表达调控中起重要的作用, 可在表观遗传水平、转录水平和转录后水平调控基因的表达, 参与 X 染色体失活、基因组印迹、染色质修饰、转录激活、转录干扰、蛋白质折叠与变性、核内运输以及疾病的发生 (Hung and Chang, 2010; Rinn and Chang, 2012; Batista and Chang, 2013)。例如, 大量的小鼠 lncRNA 在胚胎干细胞分化期特异性表达, 并且在脑部呈现精确的亚细胞定位 (Dinger *et al.*, 2008; Mercer *et al.*, 2008)。

目前对 lncRNA 的分类还没有统一标准, Kung 等 (2013) 将 lncRNA 分为 5 类: 1) 单独转录, 不与蛋白质编码基因发生重叠, 这一类 lncRNA 称为基因间区 lncRNA (Long intervening non-coding RNA, lincRNA), 如 *H19* (Brannan *et al.*, 1990), *Xist* (Brockdorff *et al.*, 1992), *HORTAIR* (Rinn *et al.*, 2007), *MALAT1* (Ji *et al.*, 2003); 2) 来自于其他基因的内含子区, 如植物的 *COLD AIR*, 位于 *FLC* 基因的内含子区 (Heo and Sung, 2011); 3) 蛋白质编码基因的反义转录本, 与蛋白质编码基因发生不同程度的重叠 (完全不重叠的分散反义转录本; 完全重叠的反义转录本; 部分重叠的反义转录本), 形成正义-反义对 (Sense-antisense pairs), 如 *Xist/Tsix*, 共同调控 X 染色体的失活 (Lee *et al.*, 1999); 4) 来源于假基因, 如 *Xist* 被认为是由蛋白质编码基因 *LnX3* 假基因化, 并整合不同转座子来源的重复序列而形成 (Duret *et al.*, 2006; Elisaphenko *et al.*, 2008); 5) 由增强子 (Enhancer), 启动子 (Promoter) 转录而来 (图 1: A)。

依据 lncRNA 在基因组上的位置信息, 将 lncRNA 分为四类 (Ma *et al.*, 2013): 1) 正义 lncRNA (Sense lncRNA); 2) 基因间区 lncRNA (Intergenic lncRNA); 3) 内含子区 lncRNA (Intronic lncRNA); 4) 反义 lncRNA (Antisense lncRNA) (图 1: B)。

St Laurent 等 (2015) 综述了目前 lncRNA 的分类标准。依据转录本的长度将 lncRNA 分为 4 类: 1) lncRNA (长度大于 200 nt); 2) 基因

间区 lncRNA ; 3) 基因间区超长 lncRNA (长度在 50 kb 与 1 mb 之间); 4) 与启动子相关联的 lncRNA (Promoter-associated long RNA)。依据与蛋白质编码基因的相似程度可以分为类似 mRNA 的非编码 RNA (mRNA-like noncoding

RNA , mlncRNA) 和基因间区 lncRNA。依据其生物学功能分为 4 类 类似增强子功能的 lncRNA、miRNA 的初始转录本、piRNA 的初始转录本、竞争性内源 RNA (Competing endogenous RNA , ceRNA)。

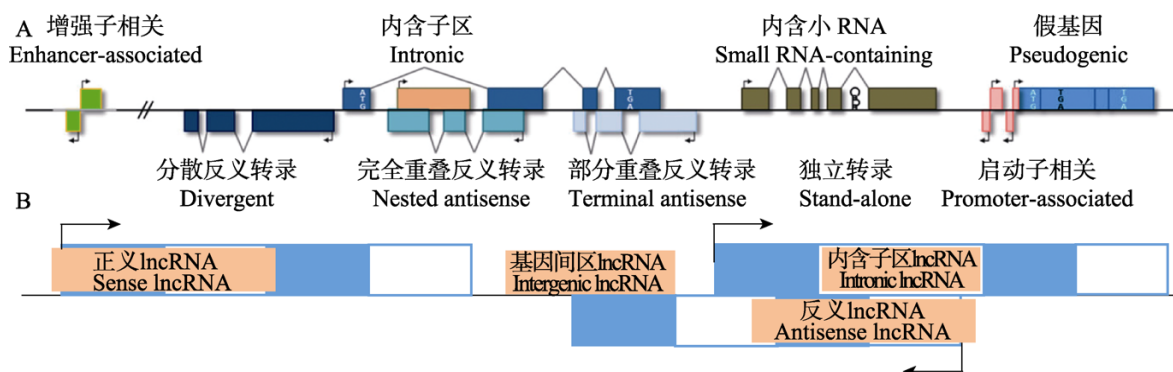


图 1 lncRNA 的分类情况 (改自 Kung *et al.*, 2013; Ma *et al.*, 2013)

Fig. 1 Classification of lncRNA (modified from Kung *et al.*, 2013; Ma *et al.*, 2013)

A. 基于 lncRNA 在蛋白质编码基因位于基因组位置的转录信息分类 ; B. 基于基因组位置信息的 lncRNA 分类。

A. Classification of lncRNA according to the genome loci of protein coding gene from where these RNAs are transcribed;

B. Classification of lncRNA based on the loci in the genome.

2 lncRNA 基因的预测方法

随着测序技术的进步,涌现出了不同类型的数据,如 EST、cDNA、Tiling array、ChIP-seq、RNA-seq、poly(A)位点图谱测序数据,针对这些数据,研究人员提出了不同的 lncRNA 发现方法。

(1) 基于 EST 数据和全长 cDNA 文库。生物体内存在独立转录的 lncRNA 分子,早期的发掘方法利用 lncRNA 这一分子特征,通过构建全长 cDNA 文库,结合 EST 数据及基因组注释信息,发掘潜在的 lncRNA (Okazaki *et al.*, 2002)。

(2) 基于染色质修饰特征。Khalil 等 (2009)、Nam 和 Bartel (2012) 利用启动子区存在 H3K4me3 修饰的特性,从小鼠、人、线虫的 ChIP-seq 测序数据中,进行全基因组范围内搜索“K4-K36”修饰信号,发现了大量的 lncRNA 基因。

(3) 基于 RNA-seq 测序数据。近年测序技术发展迅速,涌现出大量的高通量测序数据。高通量 RNA-seq 测序不仅可以检测到低丰度的转录本,也可以发现新的转录本。同时,基于转录组技术衍生出 poly(A)位点图谱测序和链特异性

RNA 测序 (Strand-specific RNA-seq), 这些测序技术都广泛地应用于 lncRNA 发现。传统转录组测序方法无法确定转录本的转录方向,而这些测序技术采用去除 rRNA 和链特异性文库构建的方法,能够保留较为完整的 lncRNA 和 mRNA 序列信息,确定 lncRNA 的序列方向,发现更多的反义 lncRNA,便于开展 lncRNA 的分类及序列结构分析,同时也可用于 lncRNA 与 mRNA 的关联分析,深入研究 lncRNA 调控网络。基于 RNA-seq 测序数据预测 lncRNA 的方法流程如图 2 所示,依赖基因组进行拼接形成转录本,转录本经长度的过滤,编码能力的预测过滤等步骤,最后获得 lncRNA 基因。

3 昆虫 lncRNA 基因的发现

早期的 lncRNA 发现主要集中于人、小鼠、线虫、斑马鱼等模式物种中,由于 lncRNA 的重要功能、高通量测序技术的广泛应用以及生物信息学的快速发展,昆虫 lncRNA 的发现受到了越来越多研究人员的重视。依据不同的测序数据,在果蝇中发现了大量的 lncRNA 基因 (表 1),

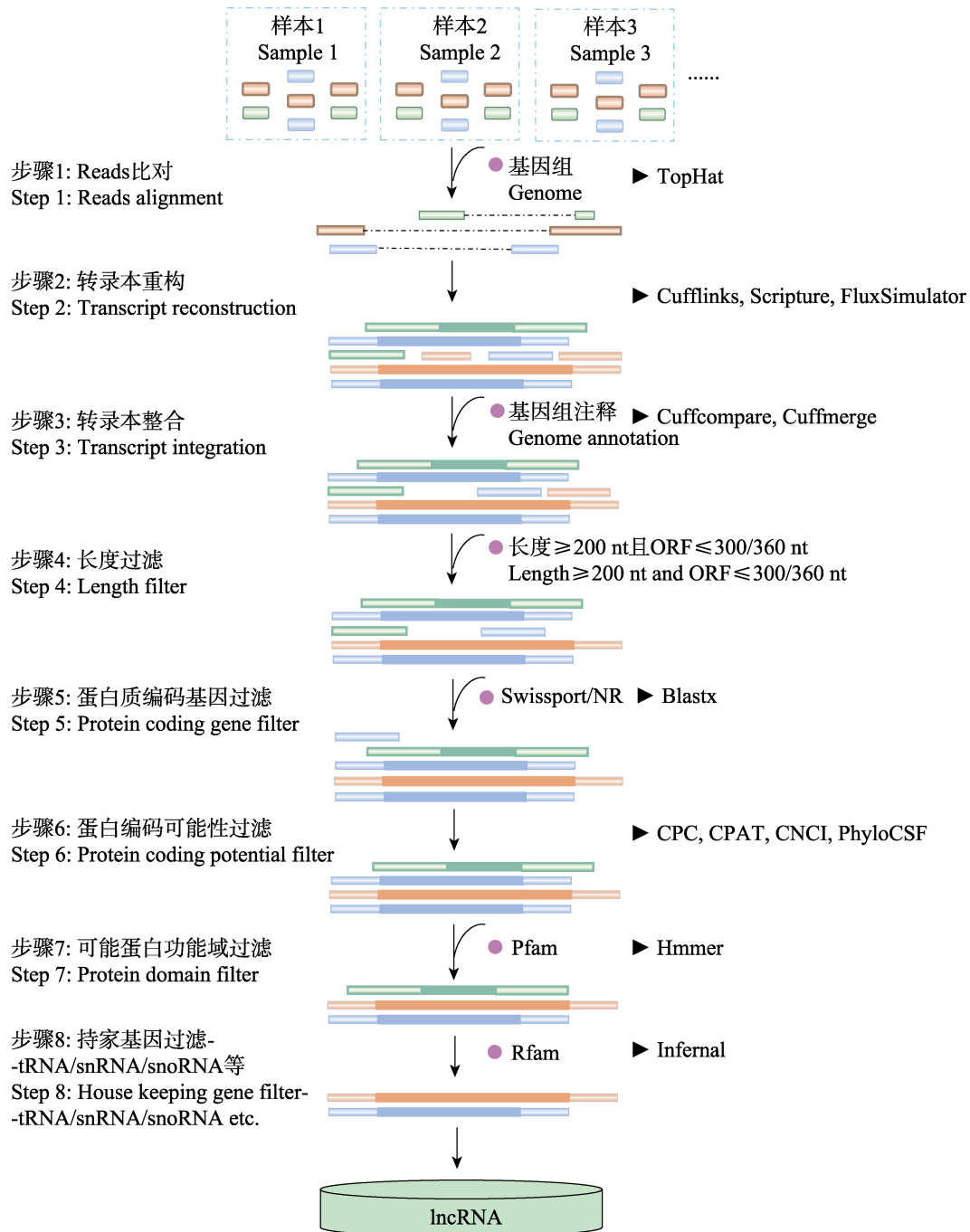


图2 长链非编码 RNA 预测流程
Fig. 2 lncRNA prediction pipeline

Young 等 (2012) 利用 30 个果蝇不同发育时间点的 RNA-seq 数据, 发现了 1 119 个 lncRNA 基因, 进一步分析发现这些 lncRNA 可能对果蝇的神经系统、精巢的发育起重要作用。Chen 等 (2016) 利用现有数据库中 RNA-seq 和 ChIP-seq 数据, 发现了 3 816 个 lncRNA 基因, 其中 462 个为新的 lncRNA, 同时更正了目前 Flybase 及

NONCODE 数据库中收集的果蝇 lncRNA 数据集。Jayakodi 等 (2015) 利用中华蜜蜂 *Apis cerana* 和意大利蜜蜂 *Apis mellifera* 的 RNA-seq 数据, 分别预测了 2 470 和 1 514 个 lncRNA 基因。Xiao 等 (2015) 利用 12 个 RNA-seq 数据, 在褐飞虱中发现了 1 882 个 lncRNA 基因。

目前, 已在果蝇、冈比亚按蚊 *Anopheles*

表 1 昆虫 lncRNA 的发现
Table 1 The identified lncRNAs in different insects

物种 Species	材料 Material	编码可能性分析 Coding potential analysis	lncRNA 总数 Number of lncRNA	参考文献 References
	cDNA	依据 Ka/Ks, QRNA, ORF 长度, 保守性及其他特征手动校正	17	Tupy <i>et al.</i> , 2005
黑腹果蝇 <i>D. melanogaster</i>	Intronscan 提供的供体和受体位点信息 Donor and acceptor sites identified by intronscan	N/A	129	Hiller <i>et al.</i> , 2009
	RNA-seq	CPC, PhyloCSF	1 119	Young <i>et al.</i> , 2012
	单链 polyA+ RNA-seq	PhyloCSF	1 875	Brown <i>et al.</i> , 2014
	RNA-seq, ChIP-seq	CPAT	3 816	Chen <i>et al.</i> , 2016
	RNA-seq (modENCODE)	CPAT	1 077	Chen <i>et al.</i> , 2016
拟暗果蝇 <i>D. pseudoobscura</i>	RNA-seq	CPAT	1 589	Nyberg and Machado, 2016
沙蝇 <i>P. perniciosus</i>	RNA-seq	CPAT, CPC	981	Petrella <i>et al.</i> , 2015
冈比亚按蚊 <i>A. gambiae</i>	RNA-seq	CPAT	9 863	Padron <i>et al.</i> , 2014
	RNA-seq	PhyloCSF	2 949	Jenkins <i>et al.</i> , 2015
埃及伊蚊 <i>A. aegypti</i>	RNA-seq	CPAT	3 482	Etebari <i>et al.</i> , 2016
中华蜜蜂 <i>A. cerana</i>	RNA-seq	CPC	2 470	Jayakodi <i>et al.</i> , 2015
意大利蜜蜂 <i>A. mellifera</i>	RNA-seq	CPC	1 541	Jayakodi <i>et al.</i> , 2015
多胚跳小蜂 <i>C. floridanum</i>	cDNA	Blast+	3	Inoue <i>et al.</i> , 2014
褐飞虱 <i>N. lugens</i>	RNA-seq	CPC	1 882	Xiao <i>et al.</i> , 2015
家蚕 <i>B. mori</i>	RNA-seq	CPC, CNCI	4 856	Zhou <i>et al.</i> , 2016
	RNA-seq	CPC, CPAT, CNCI	5 556	Wu <i>et al.</i> , 2016
小菜蛾 <i>P. xylostella</i>	RNA-seq	CPAT, CPC	3 844	Etebari <i>et al.</i> , 2015

gambiae、意大利蜜蜂、褐飞虱等共 11 种昆虫中发现了 47 049 个 lncRNA 基因 (表 1)。绝大多数昆虫物种的 lncRNA 基因都来源于 RNA-seq 测序数据。例如, Wu 等 (2016) 利用家蚕 *Bombyx mori* 21 个链特异性与非链特异性 RNA-seq 数据, 在家蚕中发现了 11 810 条 lncRNA 转录本, 其中 474 条为内含子区 lncRNA, 6 250 条为基因间区 lncRNA, 5 086 条为反义 lncRNA。Zhou 等 (2016) 利用家蚕 45 个链特异性与非链特异性 RNA-seq 数据, 构建了严格的 lncRNA 发现流程,

结合 Wu 等 (2016) 发现的 11 810 条 lncRNA 转录本, 进行了质量过滤, 在家蚕中共发现 6 821 个高质量的 lncRNA 基因, 并构建了家蚕非编码 RNA 数据库 BmncRNAdb, 这也是昆虫首个长链非编码 RNA 数据库。昆虫中 lncRNA 的发现, 有利于深入研究 lncRNA 在昆虫生长发育过程中的功能。

4 昆虫 lncRNA 基因的功能研究

昆虫 lncRNA 基因的功能研究相对较少, 表

2 总结了现有的昆虫 lncRNA 的功能研究。研究表明, lncRNA 参与调控果蝇的热击反应、X 染色体剂量补偿效应、通过调控 *BX-C* 基因表达进而调控生长发育、睡眠行为、求偶行为、运动行为、交配行为、胚胎发育中的背向闭合、刚毛的形成、性别决定等多个生物学过程(Smith *et al.* , 2001 ; Hardiman *et al.* , 2002 ; Deng and Meller ,

2006 ; Soshnev *et al.* , 2008 ; Chen *et al.* , 2011 ; Lakhotia , 2011 ; Soshnev *et al.* , 2011 ; Gummalla *et al.* , 2012 ; Lakhotia *et al.* , 2012 ; Li *et al.* , 2012 ; Pease *et al.* , 2013 ; Mulvey *et al.* , 2014 ; Rios-Barrera *et al.* , 2015)。在意大利蜜蜂中, *Nb-1* 被发现调控工蜂行为多态性(Tadano *et al.* , 2009) ; *Ks-1*、*AncR-1*、*Kakusei* 等 3 个 lncRNA 基因于脑

表 2 昆虫 lncRNA 基因的功能
Table 2 The function analysis of insect lncRNA

基因 Gene	物种 Species	功能 Function	参考文献 References
<i>hsr-ω</i>	黑腹果蝇 <i>D. melanogaster</i>	应对热激反应 In response to heat shock stress	Lakhotia, 2011; Lakhotia <i>et al.</i> , 2012
<i>rox1/2</i>	黑腹果蝇 <i>D. melanogaster</i>	参与剂量补偿效应 Involve in dosage compensation	Smith <i>et al.</i> , 2001; Deng and Meller, 2006
<i>bithorax</i>	黑腹果蝇 <i>D. melanogaster</i>	调控生长发育 Regulate growth and development	Lipshitz <i>et al.</i> , 1987; Pease <i>et al.</i> , 2013
<i>yar</i>	黑腹果蝇 <i>D. melanogaster</i>	调控果蝇的睡眠行为 Regulate sleep behavior	Soshnev <i>et al.</i> , 2008, 2011
<i>sphinx</i>	黑腹果蝇 <i>D. melanogaster</i>	调控果蝇的求偶行为 Involve in the regulation of male courtship behavior	Chen <i>et al.</i> , 2011
<i>acal</i>	黑腹果蝇 <i>D. melanogaster</i>	在果蝇胚胎形成过程中负调控其背向闭合 Regulate dorsal closure	Rios-Barrera <i>et al.</i> , 2015
<i>SxlPe R1/R2</i>	黑腹果蝇 <i>D. melanogaster</i>	调控果蝇的性别决定 Facilitate sex determination decision	Mulvey <i>et al.</i> , 2014
<i>CRG</i>	黑腹果蝇 <i>D. melanogaster</i>	正向调控果蝇的运动行为和攀爬能力 Locomotor activity and climbing ability	Li <i>et al.</i> , 2012
<i>iab-8</i>	黑腹果蝇 <i>D. melanogaster</i>	调控果蝇的交配行为 Regulate the mating behavior	Gummalla <i>et al.</i> , 2012
<i>Bft</i>	黑腹果蝇 <i>D. melanogaster</i>	调控果蝇刚毛的形成 Contribute to bristle morphogenesis	Hardiman <i>et al.</i> , 2002
<i>Lnccov1/2</i>	意大利蜜蜂 <i>A. mellifera</i>	胚胎形成期调控工蜂卵巢巢管细胞自噬 Involve in the autophagic cell death of ovarioles during worker embryogenesis	Humann and Hartfelder, 2011
<i>Nb-1</i>	意大利蜜蜂 <i>A. mellifera</i>	调控工蜂行为多态性 Regulate behavioral polyethism of workers	Tadano <i>et al.</i> , 2009
<i>Ks-1</i>	意大利蜜蜂 <i>A. mellifera</i>	蜜蜂脑部 Kenyon 细胞蕈型体中特异性表达 Expressed restrictively in the mushroom body of Keyon cells in the honeybee brain	Sawata <i>et al.</i> , 2004
<i>AncR-1</i>	意大利蜜蜂 <i>A. mellifera</i>	倾向于脑部、性组织、分泌器官及核内表达 Preferentially expressed in the brain, sexual tissues and some secretory organs and accumulates in nuclei	Sawata <i>et al.</i> , 2002
<i>Kakusei</i>	意大利蜜蜂 <i>A. mellifera</i>	神经核内特异性表达, 特异性调控蜜蜂脑部 RNA 代谢 Localized exclusively in neural nuclei and play specific roles in RNA metabolism in the honeybee brain	Kiya <i>et al.</i> , 2008

部特异性表达, 调控神经发育 (Sawata *et al.*, 2002, 2004; Kiya *et al.*, 2008)。Humann 和 Hartfelder (2011) 发现 lncRNA 基因 *Lnccov1* 在工蜂和蜂后的胚胎发育阶段的卵巢中差异表达, 在工蜂胚胎形成期调控其卵巢管细胞的自噬死亡过程。

在赤拟谷盗 Hox 基因簇的反义链上, 发现了 2 个能表达的非编码 RNA 基因, 这 2 个非编码 RNA 可能是潜在的 lncRNA 基因 (Shippy *et al.*, 2008)。Fben-1 是家蚕中首个被发现的 lncRNA 基因, Taguchi 等 (2011) 发现 Fben-1 在家蚕雌成虫脑部特异性表达, 可能参与家蚕的性二态分化。Li 等 (2014) 发现 Bm-102/Bm-159 在家蚕丝腺中高表达, 这两个中等长度的 ncRNA 可能参与家蚕的表观遗传修饰, 在家蚕的丝腺发育中起重要作用。

4.1 lncRNA 基因与昆虫发育

昆虫的生活史各异, 但绝大多数昆虫都是从卵孵化而来, 其生长受限于非弹性的外骨骼, 经历一系列的蜕皮而发育成熟。促前胸腺激素 (Prothoracicotrophic hormone, PTTH) 前胸腺分泌的蜕皮激素 (Ecdysone, E) 以及咽侧体分泌的保幼激素 (Juvenile hormones, JH) 共同调控昆虫的变态发育 (Truman and Riddiford, 2002)。Chen 等 (2016) 分析果蝇从胚胎至成虫 27 个不同发育时间点的 RNA-seq 数据, 发现了 1 077 个 lncRNA 基因, 其中 21% 在胚胎的 22~24 h 高表达, 42% 在幼虫末期高表达, 这两个时期均为果蝇发育过程形态转变的关键时期, 说明 lncRNA 可能参与果蝇从胚胎发育至幼虫、从幼虫发育至蛹的变态发育过程, 在变态发育过程中起重要调控作用。研究表明, lncRNA 基因的表达受激素的调控, 果蝇变态发育过程中 lncRNA 的表达上调可能受其体内蜕皮激素的作用 (Riddiford *et al.*, 2000; Knoll *et al.*, 2015)。Liu 等 (2016) 利用 CRISPR 干扰技术 (CRISPR interference, CRISPRi) 靶向人类 7 个不同细胞系中 16 401 个 lncRNA 基因位点, 发现其中 499 个 lncRNA 基因对细胞发育至关重要, 其中 89% 的 lncRNA 只在一个细胞系

中出现, 呈现严格的功能定向性。lncRNA 在动物胚胎发育过程中功能保守, 部分 lncRNA 参与胚胎干细胞分化 (Dinger *et al.*, 2008)。

4.2 lncRNA 基因与害虫抗药性

由于杀虫剂的长期广泛使用, 害虫的抗药性日趋严重, 已成为害虫综合治理面临的重大问题。而杀虫剂种类的增加及高剂量的使用, 使害虫的抗药性更为复杂。Etebari 等 (2015) 对小菜蛾抗性品系与敏感品系共 7 个 RNA-seq 数据分析共发现 3 844 个 lncRNA 基因, 进一步差异表达分析发现 358 个 lncRNA 在毒死蜱抗性品系中高表达、280 个 lncRNA 在氟虫腈抗性品系中高表达、59 个 lncRNA 在 Bt 抗性品系中高表达, 其中 3 个 lncRNA 在这三类杀虫剂抗性品系中均高表达。表明这些 lncRNA 可能参与调控小菜蛾对这三类杀虫剂的抗药性。为进一步验证 lncRNA 在小菜蛾敏感品系与抗性品系中的差异表达情况, 随机抽取了 9 个 lncRNA 基因进行 RT-PCR 验证, 发现 5 个 lncRNA 基因可能参与调控小菜蛾对溴氰菊酯的抗药性, 表明 lncRNA 可能在昆虫解毒过程与杀虫剂抗性形成过程发挥重要作用。与表观遗传效应相似, 接触杀虫剂可能诱导 lncRNA 基因表达上调但不改变 DNA 序列, 可能是由于杀虫剂诱导相关 lncRNA 调控通路进而引起表观遗传效应 (Collotta *et al.*, 2013; Joh *et al.*, 2014)。

4.3 lncRNA 基因与抗病毒反应

Etebari 等 (2016) 利用埃及伊蚊 35 个 RNA-seq 数据, 深度挖掘并分析了埃及伊蚊的 lncRNA, 发现埃及伊蚊 lncRNA 与邻近物种致倦库蚊 (*Culex quinquefasciatus*) 和冈比亚按蚊 lncRNA 进化保守性低, 感染型登革病毒的埃及伊蚊 Aa20 细胞其 lncRNA 表达丰度会上调, 进而抑制病毒的复制。沉默 lincRNA-1317 基因的表达能增强型登革病毒的复制; 在感染 *Wolbachia* 的埃及伊蚊中也观察到了 lincRNA-1317 表达量上升的情况, 这表明 lncRNA 可能参与了埃及伊蚊体内的抗病毒反应。lncRNA 参与免疫反应,

同时也发现病毒侵染能改变 lncRNA 的正常表达或者诱导 lncRNA 的表达 (Pan *et al.*, 2013; Fitzgerald and Caffrey, 2014)。埃及伊蚊体内 lncRNA 对病毒侵染的反应可能是通过控制免疫反应相关蛋白的表达而形成的 (Roberts *et al.*, 2014)。

4.4 lncRNA 基因与昆虫繁殖力

昆虫的繁殖力受环境条件和内部因子的双重调控, lncRNA 参与调控昆虫的繁殖力。Xiao 等 (2015) 分析了褐飞虱 12 个转录组数据, 发现 146 个 lncRNA 基因在高繁殖力种群中特异性表达, 同时发现 3 个 lncRNA 基因分别与 3 个繁殖力相关的蛋白质编码基因发生重叠, 其中 2 个 lncRNA 基因在高繁殖力种群中特异性表达, 这 2 个 lncRNA 基因和与其重叠的蛋白质编码基因独立转录; 1 个 lncRNA 基因在低繁殖力种群中特异性表达。RT-PCR 实验表明, 此 lncRNA 基因和与其重叠的蛋白质编码基因可能共用相同转录本。Wen 等 (2016) 在果蝇中发现了 128 个精巢特异性表达的 lncRNA 基因, 通过构建 CRISPR/Cas9 系统, 敲除了果蝇体内 105 个 lncRNA 基因, 发现 31% 的敲除都能导致果蝇雄性生殖力部分或完全缺失, 并呈现出可视的精子形成后期发育受损。这些结果表明 lncRNA 基因与昆虫的生殖能力密切相关, 调控昆虫繁殖力。

5 总结与展望

lncRNA 的研究已经成为生命科学领域的前沿热点。综上所述, lncRNA 在剂量补偿、基因组印迹、细胞周期、发育、配子形成等多个重要的生命活动中发挥重要的调控功能。随着高通量测序技术的发展, 越来越多的昆虫基因组和转录组数据提交至 SRA 数据库, 目前已有 156 个昆虫基因组完成测序, 转录组数据更是庞大 (Yin *et al.*, 2016)。然而, 传统的数据分析主要集中于蛋白质编码基因的发现和研究, 大量的非编码 RNA 信息被忽略。随着 lncRNA 研究的进一步深入, 昆虫 lncRNA 的研究也有了一定的进展, 大量转录组数据被重新利用。最新发展的 CRISPR/

Cas9、CRISPR 干扰技术, 将有利于深入研究昆虫 lncRNA 基因的功能。

相信在不久的将来, 随着昆虫高通量测序数据的不断积累, 以及 lncRNA 基因的发现、预测及功能研究的深入与完善, 会有越来越多的昆虫 lncRNA 被发现。由于昆虫种类繁多, 将成为 lncRNA 基因进化分析和功能研究的极佳材料。同时, 对昆虫 lncRNA 基因功能的深入研究, 有助于更全面理解昆虫发育、免疫、抗药性等重要性状的分子机制, 为害虫防治和新型杀虫剂的研发提供理论基础, 为资源昆虫的利用和有害昆虫的防治开辟新天地。

参考文献 (References)

- Batista PJ, Chang HY, 2013. Long noncoding RNAs: cellular address codes in development and disease. *Cell*, 152(6): 1298–1307.
- Brannan CI, Dees EC, Ingram RS, Tilghman SM, 1990. The product of the *H19* gene may function as an RNA. *Mol. Cell Biol.*, 10(1): 28–36.
- Brockdorff N, Ashworth A, Kay GF, McCabe VM, Norris DP, Cooper PJ, Swift S, Rastan S, 1992. The product of the mouse *Xist* gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell*, 71(3): 515–526.
- Brown CJ, Hendrich BD, Rupert JL, Lafreniere RG, Xing Y, Lawrence J, Willard HF, 1992. The human *XIST* gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell*, 71(3): 527–542.
- Brown JB, Boley N, Eisman R, May GE, Stoiber MH, Duff MO, Booth BW, Wen J, Park S, Suzuki AM, Wan KH, Yu C, Zhang D, Carlson JW, Cherbas L, Eads BD, Miller D, Mockaitis K, Roberts J, Davis CA, Frise E, Hammonds AS, Olson S, Shenker S, Sturgill D, Samsonova AA, Weizmann R, Robinson G, Hernandez J, Andrews J, Bickel PJ, Carninci P, Cherbas P, Gingeras TR, Hoskins RA, Kaufman TC, Lai EC, Oliver B, Perrimon N, Graveley BR, Celniker SE, 2014. Diversity and dynamics of the *Drosophila* transcriptome. *Nature*, 512(7515): 393–399.
- Chen B, Zhang Y, Zhang X, Jia S, Chen S, Kang L, 2016. Genome-wide identification and developmental expression profiling of long noncoding RNAs during *Drosophila* metamorphosis. *Sci. Rep.*, 6: 23330.

- Chen MJ, Chen LK, Lai YS, Lin YY, Wu DC, Tung YA, Liu KY, Shih HT, Chen YJ, Lin YL, Ma LT, Huang JL, Wu PC, Hong MY, Chu FH, Wu JT, Li WH, Chen CY, 2016. Integrating RNA-seq and ChIP-seq data to characterize long non-coding RNAs in *Drosophila melanogaster*. *BMC Genomics*, 17(1): 220.
- Chen Y, Dai H, Chen S, Zhang L, Long M, 2011. Highly tissue specific expression of *Sphinx* supports its male courtship related role in *Drosophila melanogaster*. *PLoS ONE*, 6(4): e18853.
- Collotta M, Bertazzi PA, Bollati V, 2013. Epigenetics and pesticides. *Toxicology*, 307(1416): 35–41.
- Consortium EP, 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489(7414): 57–74.
- Consortium EP, Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET, Thurman RE, Kuehn MS, Taylor CM, Neph S, Koch CM, Asthana S, Malhotra A, Adzhubei I, Greenbaum JA, Andrews RM, Flicek P, Boyle PJ, Cao H, Carter NP, Clelland GK, Davis S, Day N, Dhami P, Dillon SC, Dorschner MO, Fiegler H, Giresi PG, Goldy J, Hawrylycz M, Haydock A, Humbert R, James KD, Johnson BE, Johnson EM, Frum TT, Rosenzweig ER, Karnani N, Lee K, Lefebvre GC, Navas PA, Neri F, Parker SC, Sabo PJ, Sandstrom R, Shafer A, Vetric D, Weaver M, Wilcox S, Yu M, Collins FS, Dekker J, Lieb JD, Tullius TD, Crawford GE, Sunyaev S, Noble WS, Dunham I, Denoeud F, Reymond A, Kapranov P, Rozowsky J, Zheng D, Castelo R, Frankish A, Harrow J, Ghosh S, Sandelin A, Hofacker IL, Baertsch R, Keefe D, Dike S, Cheng J, Hirsch HA, Sekinger EA, Lagarde J, Abril JF, Shahab A, Flamm C, Fried C, Hackermuller J, Hertel J, Lindemeyer M, Missal K, Tanzer A, Washietl S, Korbel J, Emanuelsson O, Pedersen JS, Holroyd N, Taylor R, Swarbreck D, Matthews N, Dickson MC, Thomas DJ, Weirauch MT, Gilbert J, Drenkow J, Bell I, Zhao X, Srinivasan KG, Sung WK, Ooi HS, Chiu KP, Foissac S, Alioto T, Brent M, Pachter L, Tress ML, Valencia A, Choo SW, Choo CY, Ucla C, Manzano C, Wyss C, Cheung E, Clark TG, Brown JB, Ganesh M, Patel S, Tammana H, Chrast J, Henrichsen CN, Kai C, Kawai J, Nagalakshmi U, Wu J, Lian Z, Lian J, Newburger P, Zhang X, Bickel P, Mattick JS, Carninci P, Hayashizaki Y, Weissman S, Hubbard T, Myers RM, Rogers J, Stadler PF, Lowe TM, Wei CL, Ruan Y, Struhl K, Gerstein M, Antonarakis SE, Fu Y, Green ED, Karaoz U, Siepel A, Taylor J, Liefer LA, Wetterstrand KA, Good PJ, Feingold EA, Guyer MS, Cooper GM, Asimenos G, Dewey CN, Hou M, Nikolaev S, Montoya-Burgos JI, Loytynoja A, Whelan S, Pardi F, Masingham T, Huang H, Zhang NR, Holmes I, Mullikin JC, Ureta-Vidal A, Paten B, Seringhaus M, Church D, Rosenbloom K, Kent WJ, Stone EA, Program NCS, Baylor College of Medicine Human Genome Sequencing C, Washington University Genome Sequencing C, Broad I, Children's Hospital Oakland Research I, Batzoglou S, Goldman N, Hardison RC, Haussler D, Miller W, Sidow A, Trinklein ND, Zhang ZD, Barrera L, Stuart R, King DC, Ameur A, Enroth S, Bieda MC, Kim J, Bhinge AA, Jiang N, Liu J, Yao F, Vega VB, Lee CW, Ng P, Shahab A, Yang A, Moqtaderi Z, Zhu Z, Xu X, Squazzo S, Oberley MJ, Inman D, Singer MA, Richmond TA, Munn KJ, Rada-Iglesias A, Wallerman O, Komorowski J, Fowler JC, Couttet P, Bruce AW, Dovey OM, Ellis PD, Langford CF, Nix DA, Euskirchen G, Hartman S, Urban AE, Kraus P, Van Calcar S, Heintzman N, Kim TH, Wang K, Qu C, Hon G, Luna R, Glass CK, Rosenfeld MG, Aldred SF, Cooper SJ, Halees A, Lin JM, Shulha HP, Zhang X, Xu M, Haidar JN, Yu Y, Ruan Y, Iyer VR, Green RD, Wadelius C, Farnham PJ, Ren B, Harte RA, Hinrichs AS, Trumbower H, Clawson H, Hillman-Jackson J, Zweig AS, Smith K, Thakapallayil A, Barber G, Kuhn RM, Karolchik D, Armengol L, Bird CP, de Bakker PI, Kern AD, Lopez-Bigas N, Martin JD, Stranger BE, Woodroffe A, Davydov E, Dimas A, Eyraes E, Hallgrimsdottir IB, Huppert J, Zody MC, Abecasis GR, Estivill X, Bouffard GG, Guan X, Hansen NF, Idol JR, Maduro VV, Maskeri B, McDowell JC, Park M, Thomas PJ, Young AC, Blakesley RW, Muzny DM, Sodergren E, Wheeler DA, Worley KC, Jiang H, Weinstock GM, Gibbs RA, Graves T, Fulton R, Mardis ER, Wilson RK, Clamp M, Cuff J, Gnerre S, Jaffe DB, Chang JL, Lindblad-Toh K, Lander ES, Koriabine M, Nefedov M, Osoegawa K, Yoshinaga Y, Zhu B, de Jong PJ, 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, 447(7146): 799–816.
- Deng X, Meller VH, 2006. roX RNAs are required for increased expression of X-linked genes in *Drosophila melanogaster* males. *Genetics*, 174(4): 1859–1866.
- Dinger ME, Amaral PP, Mercer TR, Pang KC, Bruce SJ, Gardiner BB, Askarian-Amiri ME, Ru K, Solda G, Simons C, Sunkin SM, Crowe ML, Grimmond SM, Perkins AC, Mattick JS, 2008. Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Res.*, 18(9): 1433–1445.
- Duret L, Chureau C, Samain S, Weissenbach J, Avner P, 2006. The *Xist* RNA gene evolved in eutherians by pseudogenization of a protein-coding gene. *Science*, 312(5780): 1653–1655.
- Elisaphenko EA, Kolesnikov NN, Shevchenko AI, Rogozin IB, Nesterova TB, Brockdorff N, Zakian SM, 2008. A dual origin of the *Xist* gene from a protein-coding gene and a set of transposable elements. *PLoS ONE*, 3(6): e2521.
- Etebari K, Asad S, Zhang G, Asgari S, 2016. Identification of *Aedes aegypti* long intergenic non-coding RNAs and their association with *Wolbachia* and dengue virus infection. *PLoS Negl. Trop.*

- Dis.*, 10(10): e0005069.
- Etebari K, Furlong MJ, Asgari S, 2015. Genome wide discovery of long intergenic non-coding RNAs in diamondback moth (*Plutella xylostella*) and their expression in insecticide resistant strains. *Sci. Rep.*, 5(11): 14642.
- Fitzgerald KA, Caffrey DR, 2014. Long noncoding RNAs in innate and adaptive immunity. *Curr. Opin. Immunol.*, 26(2): 140–146.
- Gummalla M, Maeda RK, Castro Alvarez JJ, Gyurkovics H, Singari S, Edwards KA, Karch F, Bender W, 2012. *abd-A* regulation by the *iab-8* noncoding RNA. *PLoS Genet.*, 8(5): e1002720.
- Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, Cabili MN, Jaenisch R, Mikkelsen TS, Jacks T, Hacohen N, Bernstein BE, Kellis M, Regev A, Rinn JL, Lander ES, 2009. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature*, 458(7235): 223–227.
- Hardiman KE, Brewster R, Khan SM, Deo M, Bodmer R, 2002. The *bereft* gene, a potential target of the neural selector gene cut, contributes to bristle morphogenesis. *Genetics*, 161(1): 231–247.
- Heo JB, Sung S, 2011. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science*, 331(6013): 76–79.
- Hiller M, Findeiss S, Lein S, Marz M, Nickel C, Rose D, Schulz C, Backofen R, Prohaska SJ, Reuter G, Stadler PF, 2009. Conserved introns reveal novel transcripts in *Drosophila melanogaster*. *Genome Res.*, 19(7): 1289–1300.
- Humann FC, Hartfelder K, 2011. Representational Difference Analysis (RDA) reveals differential expression of conserved as well as novel genes during caste-specific development of the honey bee (*Apis mellifera* L.) ovary. *Insect Biochem. Mol. Biol.*, 41(8): 602–612.
- Hung T, Chang HY, 2010. Long noncoding RNA in genome regulation: prospects and mechanisms. *RNA Biol.*, 7(5): 582–585.
- Huttenhofer A, Schattner P, Polacek N, 2005. Non-coding RNAs: hope or hype? *Trends Genet.*, 21(5): 289–297.
- Inoue H, Yoshimura J, Iwabuchi K, 2014. Gene expression of protein-coding and non-coding RNAs related to polyembryogenesis in the parasitic wasp, *Copidosoma floridanum*. *PLoS ONE*, 9(12): e114372.
- International Human Genome Sequencing C, 2004. Finishing the euchromatic sequence of the human genome. *Nature*, 431(7011): 931–945.
- Jayakodi M, Jung JW, Park D, Ahn YJ, Lee SC, Shin SY, Shin C, Yang TJ, Kwon HW, 2015. Genome-wide characterization of long intergenic non-coding RNAs (lincRNAs) provides new insight into viral diseases in honey bees *Apis cerana* and *Apis mellifera*. *BMC Genomics*, 16(1): 680.
- Jenkins AM, Waterhouse RM, Muskavitch MA, 2015. Long non-coding RNA discovery across the genus anopheles reveals conserved secondary structures within and beyond the *Gambiae* complex. *BMC Genomics*, 16(1): 337.
- Ji P, Diederichs S, Wang W, Boing S, Metzger R, Schneider PM, Tidow N, Brandt B, Buerger H, Bulk E, Thomas M, Berdel WE, Serve H, Muller-Tidow C, 2003. *MALAT-1*, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene*, 22(39): 8031–8041.
- Joh RI, Palmieri CM, Hill IT, Motamedi M, 2014. Regulation of histone methylation by noncoding RNAs. *Biochim. Biophys. Acta.*, 1839(12): 1385–1394.
- Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A, Regev A, Lander ES, Rinn JL, 2009. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl. Acad. Sci. U S A*, 106(28): 11667–11672.
- Kiya T, Kunieda T, Kubo T, 2008. Inducible- and constitutive-type transcript variants of *kakusei*, a novel non-coding immediate early gene, in the honeybee brain. *Insect Mol. Biol.*, 17(5): 531–536.
- Knoll M, Lodish HF, Sun L, 2015. Long non-coding RNAs as regulators of the endocrine system. *Nat. Rev. Endocrinol.*, 11(3): 151–160.
- Kung JT, Colognori D, Lee JT, 2013. Long noncoding RNAs: past, present, and future. *Genetics*, 193(3): 651–669.
- Lakhotia SC, 2011. Forty years of the 93D puff of *Drosophila melanogaster*. *J. Biosci.*, 36(3): 399–423.
- Lakhotia SC, Mallik M, Singh AK, Ray M, 2012. The large noncoding *hsr omega-n* transcripts are essential for thermotolerance and remobilization of hnRNPs, HP1 and RNA polymerase II during recovery from heat shock in *Drosophila*. *Chromosoma*, 121(1): 49–70.
- Lauressergues D, Couzigou JM, Clemente HS, Martinez Y, Dunand C, Becard G, Combier JP, 2015. Primary transcripts of microRNAs encode regulatory peptides. *Nature*, 520(7545): 90–93.
- Lee JT, Davidow LS, Warshawsky D, 1999. Tsix, a gene antisense to *Xist* at the X-inactivation centre. *Nat. Genet.*, 21(4): 400–404.
- Li DD, Liu ZC, Huang L, Jiang QL, Zhang K, Qiao HL, Jiao ZJ, Yao LG, Liu RY, Kan YC, 2014. The expression analysis of silk gland-enriched intermediate-size non-coding RNAs in silkworm *Bombyx mori*. *Insect Sci.*, 21(4): 429–438.
- Li M, Wen S, Guo X, Bai B, Gong Z, Liu X, Wang Y, Zhou Y, Chen X, Liu L, Chen R, 2012. The novel long non-coding RNA *CRG* regulates *Drosophila* locomotor behavior. *Nucleic Acids Res.*,

- 40(22): 11714–11727.
- Lipshitz HD, Peattie DA, Hogness DS, 1987. Novel transcripts from the *Ultrabithorax* domain of the bithorax complex. *Genes Dev.*, 1(3): 307–322.
- Liu SJ, Horlbeck MA, Cho SW, Birk HS, Malatesta M, He D, Attenello FJ, Villalta JE, Cho MY, Chen Y, Mandegar MA, Olvera MP, Gilbert LA, Conklin BR, Chang HY, Weissman JS, Lim DA, 2016. CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science*, 355(6320): aah7111.
- Ma L, Bajic VB, Zhang Z, 2013. On the classification of long non-coding RNAs. *RNA Biol.*, 10(6): 925–933.
- Mattick JS, Makunin IV, 2006. Non-coding RNA. *Hum. Mol. Genet.*, 15 (1): R17–29.
- Mercer TR, Dinger ME, Sunkin SM, Mehler MF, Mattick JS, 2008. Specific expression of long noncoding RNAs in the mouse brain. *Proc. Natl. Acad. Sci. U S A*, 105(2): 716–721.
- Mulvey BB, Olcese U, Cabrera JR, Horabin JJ, 2014. An interactive network of long non-coding RNAs facilitates the *Drosophila* sex determination decision. *Biochim. Biophys. Acta*, 1839(9): 773–784.
- Nam JW, Bartel DP, 2012. Long noncoding RNAs in *C. elegans*. *Genome Res.*, 22(12): 2529–2540.
- Nyberg KG, Machado CA, 2016. Comparative expression dynamics of intergenic long noncoding RNAs in the genus *Drosophila*. *Genome Biol. Evol.*, 8(6): 1839–1858.
- Okazaki Y, Furuno M, Kasukawa T, Adachi J, Bono H, Kondo S, Nikaido I, Osato N, Saito R, Suzuki H, Yamanaka I, Kiyosawa H, Yagi K, Tomaru Y, Hasegawa Y, Nogami A, Schonbach C, Gojobori T, Baldarelli R, Hill DP, Bult C, Hume DA, Quackenbush J, Schriml LM, Kanapin A, Matsuda H, Batalov S, Beisel KW, Blake JA, Bradt D, Brusica V, Chothia C, Corbani LE, Cousins S, Dalla E, Dragani TA, Fletcher CF, Forrest A, Frazer KS, Gaasterland T, Gariboldi M, Gissi C, Godzik A, Gough J, Grimmond S, Gustincich S, Hirokawa N, Jackson JJ, Jarvis ED, Kanai A, Kawaji H, Kawasawa Y, Kedzierski RM, King BL, Konagaya A, Kurochkin IV, Lee Y, Lenhard B, Lyons PA, Maglott DR, Maltais L, Marchionni L, McKenzie L, Miki H, Nagashima T, Numata K, Okido T, Pavan WJ, Pertea G, Pesole G, Petrovsky N, Pillai R, Pontius JU, Qi D, Ramachandran S, Ravasi T, Reed JC, Reed DJ, Reid J, Ring BZ, Ringwald M, Sandelin A, Schneider C, Semple CA, Setou M, Shimada K, Sultana R, Takenaka Y, Taylor MS, Teasdale RD, Tomita M, Verardo R, Wagner L, Wahlestedt C, Wang Y, Watanabe Y, Wells C, Wilming LG, Wynshaw-Boris A, Yanagisawa M, Yang I, Yang L, Yuan Z, Zavolan M, Zhu Y, Zimmer A, Carninci P, Hayatsu N, Hirozane-Kishikawa T, Konno H, Nakamura M, Sakazume N, Sato K, Shiraki T, Waki K, Kawai J, Aizawa K, Arakawa T, Fukuda S, Hara A, Hashizume W, Imotani K, Ishii Y, Itoh M, Kagawa I, Miyazaki A, Sakai K, Sasaki D, Shibata K, Shinagawa A, Yasunishi A, Yoshino M, Waterston R, Lander ES, Rogers J, Birney E, Hayashizaki Y, Consortium F, I RGERGP, Team II, 2002. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature*, 420(6915): 563–573.
- Padron A, Molina-Cruz A, Quinones M, Ribeiro JM, Ramphul U, Rodrigues J, Shen K, Haile A, Ramirez JL, Barillas-Mury C, 2014. In depth annotation of the *Anopheles gambiae* mosquito midgut transcriptome. *BMC Genomics*, 15(1): 636.
- Pan YF, Qin T, Feng L, Yu ZJ, 2013. Expression profile of altered long non-coding RNAs in patients with HBV-associated hepatocellular carcinoma. *J. Huazhong Univ. Sci. Technolog. Med. Sci.*, 33(1): 96–101.
- Pease B, Borges AC, Bender W, 2013. Noncoding RNAs of the *Ultrabithorax* domain of the *Drosophila bithorax* complex. *Genetics*, 195(4): 1253–1264.
- Pennisi E, 2012. Genomics. ENCODE project writes eulogy for junk DNA. *Science*, 337(6099): 1159, 1161.
- Petrella V, Aceto S, Musacchia F, Colonna V, Robinson M, Benes V, Cicotti G, Bongiorno G, Gradoni L, Volf P, Salvemini M, 2015. De novo assembly and sex-specific transcriptome profiling in the sand fly *Phlebotomus perniciosus* (Diptera, Phlebotominae), a major Old World vector of *Leishmania infantum*. *BMC Genomics*, 16(1): 847.
- Riddiford LM, Cherbas P, Truman JW, 2000. Ecdysone receptors and their biological actions. *Vitam. Horm.*, 60(1): 1–73.
- Rinn JL, Chang HY, 2012. Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.*, 81(1): 145–166.
- Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, Chang HY, 2007. Functional demarcation of active and silent chromatin domains in human *HOX* loci by noncoding RNAs. *Cell*, 129(7): 1311–1323.
- Rios-Barrera LD, Gutierrez-Perez I, Dominguez M, Riesgo-Escovar JR, 2015. *acal* is a long non-coding RNA in JNK signaling in epithelial shape changes during drosophila dorsal closure. *PLoS Genet.*, 11(2): e1004927.
- Roberts TC, Morris KV, Weinberg MS, 2014. Perspectives on the mechanism of transcriptional regulation by long non-coding RNAs. *Epigenetics*, 9(1): 13–20.
- Sawata M, Takeuchi H, Kubo T, 2004. Identification and analysis of the minimal promoter activity of a novel noncoding nuclear RNA gene, *AncR-1*, from the honeybee (*Apis mellifera* L.). *RNA*, 10(7):

- 1047–1058.
- Sawata M, Yoshino D, Takeuchi H, Kamikouchi A, Ohashi K, Kubo T, 2002. Identification and punctate nuclear localization of a novel noncoding RNA, *Ks-I*, from the honeybee brain. *RNA*, 8(6): 772–785.
- Shippy TD, Ronshaugen M, Cande J, He J, Beeman RW, Levine M, Brown SJ, Denell RE, 2008. Analysis of the *Tribolium homeotic* complex: insights into mechanisms constraining insect *Hox* clusters. *Dev. Genes Evol.*, 218(3): 127–139.
- Smith ER, Allis CD, Lucchesi JC, 2001. Linking global histone acetylation to the transcription enhancement of X-chromosomal genes in *Drosophila* males. *J. Biol. Chem.*, 276(34): 31483–31486.
- Soshnev AA, Ishimoto H, McAllister BF, Li X, Wehling MD, Kitamoto T, Geyer PK, 2011. A conserved long noncoding RNA affects sleep behavior in *Drosophila*. *Genetics*, 189(2): 455–468.
- Soshnev AA, Li X, Wehling MD, Geyer PK, 2008. Context differences reveal insulator and activator functions of a *Su* (Hw) binding region. *PLoS Genet.*, 4(8): e1000159.
- St Laurent G, Wahlestedt C, Kapranov P, 2015. The Landscape of long noncoding RNA classification. *Trends Genet.*, 31(5): 239–251.
- Surono A, Takeshima Y, Wibawa T, Ikezawa M, Nonaka I, Matsuo M, 1999. Circular dystrophin RNAs consisting of exons that were skipped by alternative splicing. *Hum. Mol. Genet.*, 8(3): 493–500.
- Tadano H, Yamazaki Y, Takeuchi H, Kubo T, 2009. Age- and division-of-labour-dependent differential expression of a novel non-coding RNA, *Nb-I*, in the brain of worker honeybees, *Apis mellifera* L. *Insect Mol. Biol.*, 18(6): 715–726.
- Taguchi S, Iwami M, Kiya T, 2011. Identification and characterization of a novel nuclear noncoding RNA, *Fben-I*, which is preferentially expressed in the higher brain center of the female silkworm moth, *Bombyx mori*. *Neurosci. Lett.*, 496(3): 176–180.
- Truman JW, Riddiford LM, 2002. Endocrine insights into the evolution of metamorphosis in insects. *Annu. Rev. Entomol.*, 47(1): 467–500.
- Tupy JL, Bailey AM, Dailey G, Evans-Holm M, Siebel CW, Misra S, Celniker SE, Rubin GM, 2005. Identification of putative noncoding polyadenylated transcripts in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U S A*, 102(15): 5495–5500.
- Ulitisky I, Bartel DP, 2013. lincRNAs: genomics, evolution, and mechanisms. *Cell*, 154(1): 26–46.
- Wen K, Yang L, Xiong T, Di C, Ma D, Wu M, Xue Z, Zhang X, Long L, Zhang W, Zhang J, Bi X, Dai J, Zhang Q, Lu ZJ, Gao G, 2016. Critical roles of long noncoding RNAs in *Drosophila* spermatogenesis. *Genome Res.*, 26(9): 1233–1244.
- Wu Y, Cheng T, Liu C, Liu D, Zhang Q, Long R, Zhao P, Xia Q, 2016. Systematic identification and characterization of long non-coding RNAs in the silkworm, *Bombyx mori*. *PLoS ONE*, 11(1): e0147147.
- Xiao H, Yuan Z, Guo D, Hou B, Yin C, Zhang W, Li F, 2015. Genome-wide identification of long noncoding RNA genes and their potential association with fecundity and virulence in rice brown planthopper, *Nilaparvata lugens*. *BMC Genomics*, 16(1): 749.
- Yin C, Shen G, Guo D, Wang S, Ma X, Xiao H, Liu J, Zhang Z, Liu Y, Zhang Y, Yu K, Huang S, Li F, 2016. InsectBase: a resource for insect genomes and transcriptomes. *Nucleic Acids Res.*, 44(D1): D801–D807.
- Young RS, Marques AC, Tibbit C, Haerty W, Bassett AR, Liu JL, Ponting CP, 2012. Identification and properties of 1,119 candidate lincRNA loci in the *Drosophila melanogaster* genome. *Genome Biol. Evol.*, 4(4): 427–442.
- Zhou QZ, Zhang B, Yu QY, Zhang Z, 2016. BmncRNAdb: a comprehensive database of non-coding RNAs in the silkworm, *Bombyx mori*. *BMC Bioinformatics*, 17(1): 370.