

果蝇近缘物种间性腺特异 ncRNA 表达图谱 存在显著差异

杨勇飞 李政 樊启昶 龙漫远* 张文霞*

(北京大学生命科学学院, 北京 100871; Department of Ecology and Evolution, The University of Chicago, Chicago, IL60637 USA.

* 联系人, E-mail: wxzhang@pku.edu.cn, mlong@uchicago.edu)

摘要 以果蝇为材料, 检测 ncRNA 在物种间表达的保守性. 选取在黑腹果蝇(*D. melanogaster*)研究中获得的 12 个性腺特异表达的 ncRNA 基因, 用 RT-PCR 技术分别对黑腹果蝇及其近缘物种 *D. simulans*, *D. yakuba*, *D. pseudoobscura* 和 *D. virilis* 的头部、精巢和卵巢进行了 ncRNA 转录信号的检测. 结果显示, 在果蝇中, ncRNA 的表达与否及表达强度(表达图谱)在种间有显著差异, 并体现出组织特异性的明显变动, 表明 ncRNA 的表达在近缘物种间的保守性很低.

关键词 ncRNA 性腺特异表达 ncRNA基因 果蝇近缘物种

RNA可分为信使RNA (mRNA)和非编码RNA (non-coding RNA, ncRNA)两类. mRNA通过翻译的蛋白质行使其功能, ncRNA则不能翻译成蛋白质. 过去人们一直认为, ncRNA只包括 tRNA, rRNA和剪接体RNA (spliceosomal RNA)几类, 并且只能通过辅助蛋白质的方式来行使功能. 这些观念的形成主要是由于缺乏发现和鉴定ncRNA的有效实验手段. 随着基因组测序、大规模cDNA测序和全基因组的寡核苷酸覆瓦式芯片技术(tiling microarray)的发展, 人们发现, 人类有大约 50%的基因组DNA可以转录, 但其中只有 2%编码蛋白质, 其余的 98%都属于ncRNA. 目前, ncRNA包括传统的tRNA, rRNA和新发现的细菌调节核糖核酸(bacterial riboregulators)、信号识别颗粒(signal recognition particle, SRP)、核小RNA (small nuclear RNAs, snRNAs)、核酶(ribozyme)、核仁小RNA (small nucleolar RNAs, snoRNAs)、microRNA (miRNA)和类mRNA的ncRNA (mRNA-like non-coding RNA)等 [1-4]. 他们在诸如DNA结构调控、RNA剪切、蛋白质转运、基因转录和翻译调控等许多生物学过程中发挥着重要的作用 [5-9], 并且发现由于ncRNA基因突变可能导致疾病发生 [10], 更进一步表明了ncRNA 基因在生命活动中的重要意义.

目前, 在不同物种中都发现, 许多ncRNA的表达具有组织特异性. 例如, 果蝇中的*pgc*特异地表达在种质细胞中, 并对生殖细胞命运的决定起着重要的

作用 [11]; 果蝇*bereft*特异地表达在外周神经系统中 [12], 其突变体表现为感觉器官异常, 复眼小眼面间的小刚毛发育畸形; *EVF-1*在哺乳动物神经前体细胞中表达, 其表达与Shh和Dlx家族成员的表达紧密相关, 并随着前脑的发育其表达量逐渐降低, 从而调控脑的发育 [13]. Inagaki等人 [5]已在果蝇中鉴定了 27 个组织特异性表达的ncRNA, 分别表达在气管系统、中肠原始细胞、内脏中肠细胞、生殖细胞和中央及外周神经系统等不同组织中.

编码蛋白的RNA通常在DNA序列上高度保守. 比较而言, 一些重要的ncRNA基因的DNA序列高度保守 [14], 而其他许多ncRNA基因的DNA序列却显示出比较低的保守性. 引人注目的是, 越来越多的研究结果表明, 这些序列保守性低的ncRNA同样有着重要的作用 [15], 并且其功能异常保守. 一个重要的例子就是果蝇中的ncRNA基因*roX1*和*roX2*与哺乳动物中的ncRNA基因*Xist*, 他们的转录产物都结合于X染色体的特定区域, 并通过转录抑制作用产生对X染色体的剂量补偿效应 [16-18]. 有意思的是, 虽然*roX*和*Xist*具有类似的功能, 但是它们DNA序列间的比对却显示出高度不保守性, 甚至在按蚊中都找不到与果蝇*roX*相似的序列, 而在河豚中也没有发现与哺乳动物相似的*Xist*序列. 人们推测, 犹如不同物种的tRNA, 其核苷酸序列存在着巨大的差异, 但都行使转运氨基酸的重要功能一样, 这些ncRNA也主要是

依靠其二级结构和三级结构行使功能,就是说,其功能的行使不是依赖核苷酸序列,而是依赖分子二级结构和三级结构^[6].除此之外,人们还发现了物种特异的ncRNA,如锥虫(Trypanosome)线粒体中的指导RNA (guide RNAs, gRNAs)^[19].

以上现象的发现揭示了ncRNA基因在功能和进化上的重要性,而对其组织特异性和物种特异性现象的系统研究将有助于对这类基因的深入理解.其中,首要的工作是鉴别这些特异表达的ncRNA基因.最近,已有实验室报道了利用基因芯片技术研究果蝇胚胎期基因表达(包括ncRNA)的资料数据库^[20,21].在前期的工作中,我们利用RT-PCR技术获得了一批黑腹果蝇ncRNA的数据(结果将另文发表),结果显示,ncRNA的表达有明显的发育阶段和组织差异图谱.特别的是,精巢和卵巢有特异表达的ncRNA图谱.目前,在编码蛋白基因的水平上,数千个控制性腺发育的基因已在人及家鼠^[22,23]、黑腹果蝇^[8,24]和其近缘果蝇物种^[8,9]中被鉴别.研究性腺特异的ncRNA与这些蛋白基因的表达及功能间的关系将是一个重要的生物学课题.

在前期工作的基础上,本研究选取我们在黑腹果蝇中获得的 12 个生殖腺特异表达的 ncRNA 基因,进一步在果蝇近缘物种 *D. simulans*, *D. yakuba*, *D.*

pseudoobscura 和 *D. virilis* 中进行了表达图谱的检测,研究他们在果蝇近缘物种间的保守性,为今后深入的工作奠定基础.

1 材料与方法

() 序列比对与检测片段的选定. 本实验以黑腹果蝇中前期获得的组织特异表达的候选 ncRNA 片段(表 1 和 2)为研究对象,检测他们在 5 个近缘物种间的保守性. 利用 VISTA 软件(<http://genome.lblgov/vista/>),以 *D. melanogaster* 数据为参照对其近缘物种 *D. simulans*, *D. yakuba*, *D. pseudoobscura* 和 *D. virillis* 进行序列对比分析,其中 *D. melanogaster*, *D. simulans* 和 *D. yakuba* 间的亲缘关系更近.保守性计算滑动框的长度为 100 bp,保守片段的相似性百分数阈值为 70%(在多数比对中,只有 *D. simulans* 和 *D. yakuba* 能满足此阈值, *D. pseudoobscura* 和 *D. virillis* 则很少能达到此阈值.最后,我们选取的是那些在尽可能多的近缘物种中达到此阈值的 *D. melanogaster* 保守序列,并设计相应的引物),以此选定近缘物种间的保守区段,并设计对应的 PCR 扩增引物(表 3).

() cDNA 文库制备与 ncRNA 检测系统. 用常规方法分别制备整体果蝇与头部、卵巢和精巢的 cDNA 文库.主要操作包括有: TRIzol (Invitrogen)提取总 RNA,用 DNase (Promega)消化及 M-MLV

表 1 ncRNA 片段在黑腹果蝇基因组中的位置及其序列

ncRNA	基因组中位置	产物大小 小/bp	基因组的 DNA 序列(5'→3') ^{a)}
X31	X: 3696478~3696587	110	CCCAGACGAAGCAGGTCAATTTTTTTTTTCGTGTTACTTGCTCTTTTTCGCCAAGTGTGAGTTGTGT GTTTTGGCTTGGCTAACGTGATTGAAATCGAGACAGCCGGGACA
II 10	II: 859307~859416	110	CCCCAAGTCGCACCTTTACTGCAAAATGCAAAAACCTTTAGACAGGGAATCAATTGTTGTACCTTGT TTTATTTTCGCCAAGACACGCAACAACAATGGAACCGTCAAAAGG
II 25	II: 2307157~2307265	109	GGGTGGGGAAAACCTTATTTGCTTCGAAATTACGTGTCGCCAAAGTTGGCCGATGCGACATTATAC AATTTATATGCACGTATTGCGCTCTATCCAAGATGAAATGTCGA
II 50	II: 4838493~4838606	114	CGTAAGCCCGCCACAATCGCAATTGCAAAATATCGAATGTAAATAATGCAAAATAAGACCGTGGA CTAGTGATTAATTAGTGTGTGACTTCGTTTGTGACTGCGATAACGGGG
II 56	II: 5338314~5338427	114	GTTGCGATTCTGTTCTCGGCCATAATGTCAGATGACTAAAGCTGCTGTTATCGGAGTGTGCGA ACACACAGACATAATGGCCAATCGGTTGCCGGGTCAATAAGATCGAGC
II 62	II: 5971109~5971235	127	ACCTATTGGTATGCGGTGACCACATGGTTTATATGGGTAACCTTTAATGCCATTTGACTCAAGTTTGAG TGATTGAATGGCAACATTCCACTATTCCACCCCCCGTCAAAAATTGGCACTCCA
II 69	II: 6496957~6497083	127	GATATGCCAATAAAGGTAGCATGTTACAAGATAAACCCCTTCTTGCAAACTCTGTCTGCTTTTGTATCG TCTTCGTGGAAGCACAGTACCTTCTGTCAGATCTTCAAACCTCAATCGTTCAGACATT
II 73	II: 6804450~6804570	121	CATATTTTTGCGCTGTGTAGGCAGCACTATCGCCATGGCATGGGGATCGTATTGATAGGCTTATCAC CGACTGCTTGGAAACATTCCACTATTCCACCCCCCGTCAAAAATTGGCACTCCA
II 86	II: 8030379~8030492	114	TGTTTCTTTATCAGCGGGCTGATGTCGCCTAATTTGAAATCTATTTCTTTTGTTTTCGGAAATG ATCTTTTATAATGATGGACTTAGAACCTTTTAAATCCCAATCGCTGT
II 108	II: 10032554~10032684	131	TTTGAAAAAAGCTCGGACCGTTTCTCTCAAACCTGCGCATGCGCAGCCATCTAACTGCAGTTCCAAT GGAGCAGAGCGAAAGCAGACAAGGTGATGGAGACTTCTGGGCGTTAAACTCATTGCCCGAGC
II 117	II: 10426804~10426913	110	CGTTAGAGATGGTCAATTTTCATTAATAAAATCATGTGGCGGCACCTATCGATATCGATAGGCAG TGGCCATGTTATCAAACTTATTCTTATTCATCAGCCACGTCATT
II 218	II: 22121709~22121818	110	TAGTTGCTTGCATTTTATCATTTGAAATTGTTAACGTTACGATCAATTGGGCTGTGTGAATAGG TCGCATAATTGTTTTGCCATTAATTATTTTAAGAGGCTGTCGAC

a) 片段中两端的黑体字母分别为扩增的正向引物和反向引物序列的互补序列

表 2 ncRNA 片段在黑腹果蝇基因组中的表达特异性及其相邻基因

ncRNA	组织特异性	5'侧翼基因	距 5'侧翼基因的距离/bp	3'侧翼基因	距 3'侧翼基因的距离/bp
X31	卵巢、精巢	CG32789	871	CG32779	3
II 10	头部、精巢	CG13949	3071	CG13950	905
II 25	头部、卵巢	CG31949	6440	CG16995	762
II 50	头部、精巢	CG15626	497	CG12194	697
II 56	精巢	CG14021	261	CG12512	798
II 62	精巢	Ucp4B	125	chic	1665
II 69	精巢	CG31638	-29	CG9550	-30
II 73	精巢	CG11236	78	CG17375	49
II 86	精巢	CG7380	193	Cka	231
II 108	卵巢	CG13127	3255	CG4778	10
II 117	精巢	SmB	13	Kde1R	-46
II 218	精巢	CG6675	11260	CG3651	18

(Promega)反转录. 选取果蝇稳定表达的 ncRNA 7SL 和实验鉴定的非转录序列 X8c (X 染色体上选取的第 8 区间的第 3 对引物扩增区域)(7SL 正向引物: 5'-ACGCTTCTGTGAGGTCTGATT-3'; 7SL 反向引物: 5'-GAAGCTTGGCTGCTGGTTT-3'. X8c 正向引物: 5'-TTCACCTGTCCGACGTTGTT-3'; X8c 反向引物: 5'-CGTTTAGATCTGCGCAATT-3'), 对照果蝇基因组 DNA, 对制备的 cDNA 文库进行质量鉴定, 并同时以 7SL 和 X8c 的扩增条带作为本实验的正、负对照(图 1).

() 数据分析. 用 BandScan (version 4.50) 软件对 RT-PCR 获得的所有检测条带(图 2)进行灰度值扫描, 以 7SL 电泳条带为基准对各阳性检测条带进行归一化(normalization)处理. 设分子量 Marker 中 100 ng 条带为标准值 100, 计算各物种 7SL 条带的归一化因子值, 以此对检测数据进行归一化处理, 获得不同检测片段的相对表达量(表 4).

2 结果

2.1 检测到的 ncRNA 片段

我们在黑腹果蝇 X 染色体的 191 个基因间区和第 2 染色体左臂的 213 个间区中, 鉴定出 12 个性别特异性表达的 ncRNA 片段, 代表着 12 个非编码 RNA 基因. 表 1 和表 2 列出了这些 ncRNA 基因片段在基因组中的位置和 DNA 序列. 为描述这些基因的确切位置, 我们给出了他们分别距 5'侧翼基因和 3'侧翼基因的距离.

每个 ncRNA 基因在黑腹果蝇的表达图谱表现出特定的组织特异性(表 2), 在各组织中的相对表达量见表 4. ncRNA X31 在卵巢和精巢中都表达(图 2), 而卵巢的表达量显著地高于精巢(表 4); ncRNA 10 则

在头部和精巢中有相当的表达量, ncRNA 56 只在精巢中表达. 一般而言, 大部分 ncRNA 基因在精巢中表达, 少数在卵巢中表达.

2.2 ncRNA 表达图谱的种间差异

我们进而分别对 *D. simulans*, *D. yakuba*, *D. pseudoobscura* 和 *D. virilis* 的头部、精巢和卵巢的 cDNA 文库进行转录信号检测. 图 2 显示生殖腺特异表达的 ncRNA X31 在不同物种中的表达图谱, 结果显示出在不同物种间、甚至是在亲缘关系更近的物种间其表达存在明显差异性. 例如, ncRNA 10 在黑腹果蝇的头部和精巢中表达, 而其近缘种 *D. simulans* 则在卵巢和精巢中表达. 所有 12 个生殖腺特异表达的 ncRNA 片段在不同物种间都存在明显差异性(表 4), 但其差异程度不同: ncRNA 25 在 *D. melanogaster* 表达在卵巢和头部, 在其他 4 个近缘物种中则只表达在卵巢; ncRNA 108 在 *D. melanogaster* 和 *D. yakuba* 仅在卵巢表达, 而在 *D. simulans* 和 *D. pseudoobscura* 也在精巢表达, 但在 *D. virilis* 的 3 个组织都不表达. 其余在 *D. melanogaster* 精巢特异表达的 ncRNA 片段(10 也在头部表达), 没有一个在其余近缘物种中显示一致的表达图谱: 只有 62 和 73 在 *D. simulans* 与 *D. melanogaster* 表达一致, 在 *D. pseudoobscura* 根本就无法进行 RT-PCR 检测, 因为没有相对应的基因组序列; 73 在 *D. yakuba* 的 3 个组织都没有表达, 但 *D. virilis* 的卵巢和头部也都检测到了 ncRNA 的信号; 62 在 *D. yakuba* 的卵巢和头部也检测到了表达信号, 在 *D. virilis* 则与 *D. melanogaster* 相同. 极端的例子是 218, 它只在 *D. melanogaster* 的精巢表达, 其他物种

表3 果蝇近缘物种 12 个 ncRNA 扩增引物序列^{a)}

引物	<i>D. simulans</i> (5'→3')	<i>D. melanogaster</i> (5'→3')	<i>D. yakuba</i> (5'→3')	<i>D. pseudoobscura</i> (5'→3')	<i>D. virilis</i> (5'→3')
X31 正向	CCCAGACGAAGCAGGTCAAT	CCCAGACGAAGCAGGTCAAT	CAAGACAAAACAGGTCAAT	GACAGTGCCATCCAGTCC	TCCTGCTCCTCTGCGGC
X31 反向	TGTCCCGGCTGTCTCGATTT	TGTCCCGGCTGTCTCGATTT	ATGTCCTGGCGGCTGTCTCG	CAGTTCACGCGACGCCAA	ATTTGCTCAAAAAAATCGAC
10 正向	CCCCAAGTCGCACTTTAC	CCCCAAGTCGCACTTTAC	CCCCAAGTCGCACTTTAC	GGCCGGAGTCGCACTTTAC	TTGGGATGGGCGGTGTTCTG
10 反向	CCTTTTGACGGTTCCATT	CCTTTTGACGGTTCCATT	CCTTTTGACGGTTCCATT	CTTCTTCCATTGTTTCCATT	GCCATTGTTGTTGCGTGTC
25 正向	GGGTGGGGAAAACTTATT	GGGTGGGGAAAACTTATT	CGCCGGCAGCAAACTTATT	GGGTGGGGAAAACTTATT	CAAGGCAGAAAACTTATTT
25 反向	TCGACATTTTCATCTTGGA	TCGACATTTTCATCTTGGA	TCGACATTTTCATCTTGGA	TCGACATTTTCATCTTGGA	TCGACATTTTCATCTTGGA
50 正向	CGTAAGCCCGCCACAATC	CGTAAGCCCGCCACAATC	CGTAAGCCCGCCACAATC	AGCTAGTAATATTCGATG	CATGTGGCGCGCGCGCTT
50 反向	CCCCGTTATCGCAGTCAA	CCCCGTTATCGCAGTCAA	CCCCGTTATCGCAGTCAA	AAATGAAATTGACTTATCG	GTTTACATACAAGGCACAC
56 正向	GTTGCGATTCTGTCTC	GTTGCGATTCTGTCTC	GTTGCGATTCTGTCTC	GTTCTCGCTCTCGTTCTG	N/A
56 反向	GCTCGATCTTTATGACCC	GCTCGATCTTTATGACCC	GCTCGATCTTTATGACCC	TCGAGACTTTTATGACCC	N/A
62 正向	ACCTATTGGTATGCGGTGAC	ACCTATTGGTATGCGGTGAC	ACCTATTGGTATGCGGTGAC	N/A	GAATAACTATTGGCACATG
62 反向	GTTTTGGCGGGTTGTCTA	GTTTTGGCGGGTTGTCTA	GTTTTGGCGGGTTGTCTA	N/A	CTTTTCGCTTCCTCGCCATG
69 正向	GATATGCCAATAAAGGTAGC	GATATGCCAATAAAGGTAGC	GATATGCCAATAAAGGTAGC	CTACAACTGATATTCCAAT	GCTGTCTATCAAATTTGGT
69 反向	AATGTCTGAACGATTGAAGT	AATGTCTGAACGATTGAAGT	CAAAATCCGAATGGTTGAAG	GACTCGAAGTGCTCTATC	TTCTCAAGTCTCTTGAATA
73 正向	CATATTTTGCCTGTGTAGG	CATATTTTGCCTGTGTAGG	GATTTTGTCTCTGTGTAGG	N/A	CAAGCACGTTCTGCACATTG
73 反向	TGGAGTGCCAATTTTGACG	TGGAGTGCCAATTTTGACG	TGGTGGAGCACCAGTTTTTG	N/A	CGGAAGTTTTTCAAAAGCTT
86 正向	TGGTTCTTAATCAGCGGGCG	TGTTTCTTATCAGCGGGCT	TGATTTTTATTTTGGGCG	N/A	N/A
86 反向	ACAGCGATTGGGATTTAAAA	ACAGCGATTGGGATTTAAAA	GCTTCACTCTAATAATAAC	N/A	N/A
108 正向	TTTGAAAAAAGCTCGGAC	TTTGAAAAAAGCTCGGAC	CAGAGTGTAGAGCTTTCGT	GACTCAAAAGAGTGACTT	GCGCTAAAAGCTTCGAAAAG
108 反向	CGATCGGATCGTCTCGTC	GCTCGGGCAATGAGTTTT	GCTCGGGCAATGAGTTTT	GTGGACTTGGAGCTGGAA	AACTGTAAAAAAGTTTATG
117 正向	CGTTAGAGATGGTCGAATTT	CGTTAGAGATGGTCGAATTT	TGTTTCCTTAATGACCCTG	CTCGTTTTGTGCTGTACG	GTGCTCAAACGTCTCGAAT
117 反向	AATGACGTGGCTGATGAATA	AATGACGTGGCTGATGAATA	CACTAGTCTGTGCTGGAG	GTCGCAGACTTGAAGTTC	AACTTTTGGTGTTGTGCC
218 正向	TAGTTGCCTTGCAATTTATC	TAGTTGCCTTGCAATTTATC	CTTTAAATTAGTTCCTTTG	N/A	TGCCATTGGAGATAGCATG
218 反向	GTCGACAGCCTCTTAAAATA	GTCGACAGCCTCTTAAAATA	GAACGGTAAAAATGCCGAC	N/A	TGAATATACTTTCAAAAATA

a)“N/A”表示在该物种没有相对应的基因组序列

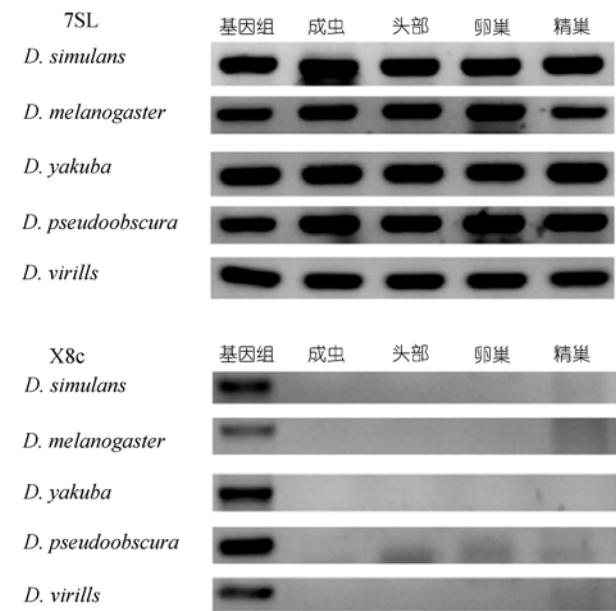


图 1 RT-PCR 检测 7SL 和 X8c 在不同物种头部、卵巢和精巢中的表达

和组织都没有检测到任何信号。86 在 *D. simulans* 除精巢表达外也在卵巢表达, 其他物种也没有任何信号。56 除在 *D. virilis* 没有对应的基因组序列外, 其他 3 个物种都同时在精巢和卵巢表达。这些表达图谱显示出显著的物种间差异, 这些差异与这些基因序列的保守性形成对比。

2.3 ncRNA 基因的基因组位置

我们观察到 ncRNA 在基因组的位置分布不符合传统理论的预期。在 12 个 ncRNA 基因中, 仅有 1 个位于 X 染色体, 另外 11 个全位于常染色体。在 10 个精巢表达的基因中, 9 个都在常染色体上, 1 个在 X 染

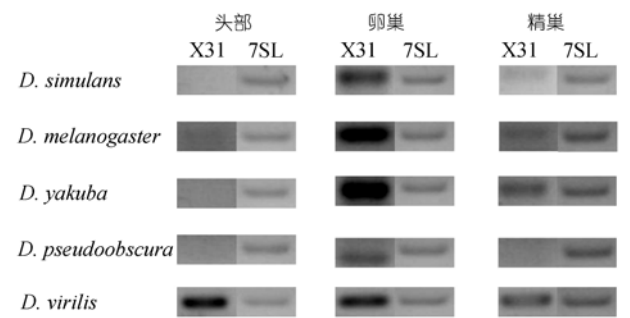


图 2 X31 在不同物种的头部、卵巢和精巢组织中的表达 7SL 为评估其相对表达量的内参(两次重复实验得到类似结果)

色体上(ncRNA X31)(表 2 和 4)。

3 讨论

根据前期实验在黑腹果蝇中检测片段的选定原则(挑选同源性高的区域), 我们原本期望在黑腹果蝇中所获得的精巢和卵巢特异表达的 ncRNA 阳性片段会在果蝇近缘物种中的表达显示出较高的保守性, 因为这些 ncRNA 基因的鉴定是基于种间 DNA 序列保守性并带来功能上的保守性而被选择的。然而, 实验结果却揭示了与预期不一致的事实, 即大部分 ncRNA 基因的表达与否和表达强度在近缘物种间表现出明显的差异。这种差异不仅表现为, 某些在黑腹果蝇中的阳性克隆(如 218)在其他物种中丢失, 而且出现器官特异表达图谱整体性的改变(如 73 在 *D. yakuba* 的 3 个器官中均未检测到, 而在 *D. virilis* 中均表达)。根据进化生物学研究, 这些近缘的果蝇物种的分歧时间很短, 特别是 *D. melanogaster* 和 *D. simulans*, 分歧时间仅仅是 2 百万~3 百万年, 黑腹果蝇与 *D. yakuba* 和 *D. pseudoobscura* 也仅分离 7 百万年

表 4 ncRNA 片段在果蝇近缘物种的头部、卵巢、精巢组织中的表达图谱(以归一化值表示)^{a)}

组织	<i>D. simulans</i>			<i>D. melanogaster</i>			<i>D. yakuba</i>			<i>D. pseudoobscura</i>			<i>D. virilis</i>		
	头部	卵巢	精巢	头部	卵巢	精巢	头部	卵巢	精巢	头部	卵巢	精巢	头部	卵巢	精巢
X31	0.00	204.88	0.00	0.00	558.08	49.18	0.00	747.72	67.71	0.00	66.99	0.00	178.97	213.17	93.00
II10	0.00	192.59	406.25	122.27	0.00	98.75	371.84	0.00	632.17	158.17	202.94	333.77	0.00	0.00	0.00
II25	0.00	38.82	0.00	74.69	244.48	0.00	0.00	44.45	0.00	0.00	93.23	0.00	0.00	121.75	0.00
II50	30.99	151.05	0.00	49.76	0.00	104.91	137.34	0.00	258.92	41.71	108.05	88.68	66.39	92.62	0.00
II56	0.00	67.93	66.07	0.00	0.00	88.20	0.00	177.97	292.14	0.00	236.56	257.95	N/A	N/A	N/A
II62	0.00	0.00	433.28	0.00	0.00	593.20	57.86	36.75	381.92	N/A	N/A	N/A	0.00	0.00	478.74
II69	0.00	98.62	458.20	0.00	0.00	82.09	56.78	0.00	0.00	139.50	150.77	161.36	49.22	0.00	75.05
II73	0.00	0.00	35.17	0.00	0.00	450.19	0.00	0.00	0.00	N/A	N/A	N/A	71.61	160.84	344.59
II86	0.00	30.73	44.30	0.00	0.00	54.71	0.00	0.00	0.00	N/A	N/A	N/A	N/A	N/A	N/A
II108	0.00	63.80	101.41	0.00	67.43	0.00	0.00	21.65	0.00	0.00	149.57	33.75	0.00	0.00	0.00
II117	0.00	94.83	118.40	0.00	0.00	123.44	0.00	0.00	0.00	28.44	108.44	33.97	34.98	576.17	498.43
II218	0.00	0.00	0.00	0.00	0.00	59.30	0.00	0.00	0.00	N/A	N/A	N/A	0.00	0.00	0.00

a) “N/A”表示在该物种没有相对应的基因组序列

到两千多万年^[25]。在这样的时间尺度, ncRNA表达图谱的显著改变, 强烈地暗示不同物种间基因组中相关的顺式和反式的调控元件发生了极快速的演化。

依据传统的进化遗传学理论, X染色体应该是性别控制基因发生的“温床”^[26], 并预期雄性特异基因应该富集在这一染色体上。但是, 控制雄性发育的编码蛋白基因在黑腹果蝇基因组中的分布并不支持这一理论^[8,24], 并且还发现由反转座产生的编码雄性调控蛋白的基因也存在逃离X染色体的演化过程^[24,27,28]。在我们检测的 12 个ncRNA基因中, 仅有 1 个位于X染色体, 另外 11 个全位于常染色体。虽然这一观察由于样本较小, 尚不能给出基因在基因组中分布的有效统计结论, 但是它与控制雄性发育的编码蛋白基因的数据是一致的, 明确显示出与传统理论相悖。

我们鉴别的 12 个ncRNA基因, 具有序列保守性和组织与发育特异性表达特性, 表明其功能上的重要性。这些基因分别在精巢和卵巢中表达, 表明他们有可能参与性别功能的执行和表达。已有的研究证明, 在分子水平上ncRNA基因是以对其他基因表达进行调控而发挥生物学功能的^[8,9,29,30]。那么, 前面提到的数千个在果蝇精巢和卵巢中表达的编码蛋白基因是否受这些ncRNA基因的调控? 虽然本研究并没有触及这个问题, 但是这些数据可能提供了为开展这一研究的基础平台, 即探索这些ncRNA基因与编码蛋白的性别控制基因的互作关系, 如开展RNA-蛋白互作实验和其他生物化学与遗传学实验^[31~34]。

致谢 感谢刘敏和林亚娟等在前期第 2 染色体左臂筛选 ncRNA 和 *D. melanogaster* 组织特异性表达的检测方面做的大量工作。

参 考 文 献

- 1 Eddy S R. Non-coding RNA genes and the modern RNA world. *Nat Rev Genet*, 2001, 2: 919—929 [DOI]
- 2 Griffiths-Jones S, Bateman A, Marshall M, et al. Rfam: An RNA family database. *Nucleic Acids Res*, 2003, 31: 439—441 [DOI]
- 3 Szymanski M, Erdmann V A, Barciszewski J. Noncoding regulatory RNAs database. *Nucleic Acids Res*, 2003, 31: 429—431 [DOI]
- 4 Pang K C, Stephen S, Engstrom P G, et al. RNAdb: A comprehensive mammalian noncoding RNA database. *Nucleic Acids Res*, 2005, 33(Database Issue): D125—D130 [DOI]
- 5 Inagaki S, Numata K, Kondo T. Identification and expression analysis of putative mRNA-like non-coding RNA in *Drosophila*. *Genes to Cells*, 2005, 10: 1163—1173 [DOI]
- 6 Mattick J S, Makunin I V. Non-coding RNA. *Human Mol Genet*, 2006, 15: R17—R29 [DOI]
- 7 齐力旺, Li X M, 张守攻, 等. 非编码蛋白 RNA 的遗传调控. *中国科学 C 辑: 生命科学*, 2006, 36: 193—208
- 8 Parisi M, Nuttall R, Naiman D, et al. Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. *Science*, 2003, 299: 697—700 [DOI]
- 9 Ranz J M, Castillo-Davis C I, Meiklejohn C D, et al. Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. *Science*, 2003, 300: 1742—1745 [DOI]
- 10 Ravasi T, Suzuki H, Pang K C, et al. Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. *Genome Res*, 2006, 16(1): 11—19 [DOI]
- 11 Martinho R G, Kunwar P S, Casanova J. A noncoding RNA is required for the repression of RNAPII-dependent transcription in primordial germ cells. *Curr Biol*, 2004, 14: 159—165 [DOI]
- 12 Hardiman K E, Brewster R, Khan S M, et al. The bereft gene, a potential target of the neural selector gene cut, contributes to bristle morphogenesis. *Genetics*, 2002, 161: 231—247
- 13 Kohtz J D, Baker D P, Corte G, et al. Regionalization within the mammalian telencephalon is mediated by changes. *Development*, 1998, 125: 5079—5089
- 14 Pollard K S, Salama S R, Lambert N, et al. An RNA gene expressed during cortical development evolved rapidly in humans. *Nature*, 2006, 443(7108): 167—172 [DOI]
- 15 Pang K C, Frith M C, Mattick J S. Rapid evolution of noncoding RNAs: Lack of conservation does not mean lack of function. *Trends Genet*, 2006, 22(1): 1—5 [DOI]
- 16 Lee J T, Strauss W M, Dausman J A, et al. A 450 kb transgene displays properties of the mammalian X-inactivation center. *Cell*, 1996, 86: 83—94 [DOI]
- 17 Herzing L B, Romer J T, Horn J M, et al. Xist has properties of the X-chromosome inactivation centre. *Nature*, 1997, 386: 272—275 [DOI]
- 18 Meller V H, Rattner B P. The roX genes encode redundant male-specific lethal transcripts required for targeting of the MSL complex. *EMBO J*, 2002, 21: 1084—1091 [DOI]
- 19 Madison-Antenucci S, Grams J, Hajduk S L. Editing machines: The complexities of trypanosome RNA editing. *Cell*, 2002, 108: 435—438 [DOI]
- 20 Manak J R, Dike S, Sementchenko V, et al. Biological function of unannotated transcription during the early development of *Drosophila melanogaster*. *Nat Genet*, 2006, 38(10): 1151—1158 [DOI]
- 21 Oliver B. Tiling DNA microarrays for fly genome cartography. *Nat Genet*, 2006, 38(10): 1101—1102 [DOI]
- 22 Khil P P, Smirnova N A, Romanienko P J, et al. The mouse X chromosome is enriched for sex-bias genes not subject to selection by meiotic sex chromosome inactivation. *Nat Genet*, 2004, 36: 642—646 [DOI]
- 23 Emerson J J, Kaessmann H, Betran E, et al. Extensive gene traffic on the mammalian X chromosome. *Science*, 2004, 303: 537—540 [DOI]
- 24 Betran E, Thornton K, Long M. Retroposed new genes out of the X in *Drosophila*. *Genome Res*, 2002, 12(12): 1854—1859 [DOI]
- 25 Ashburner M. *Drosophila: A Laboratory Handbook*. Cold Spring Harbor: Cold Spring Harbor Press, 1989
- 26 Bainbridge D. The X in sex: How the X chromosome controls our lives. London: Harvard University Press, 2003
- 27 Long M, Betran E, Thornton K, et al. The origin of new genes: Glimpses from the young and old. *Nat Rev Genet*, 2003, 4(11): 865—875 [DOI]
- 28 Dai H, Yoshimatsu T F, Long M. Retrogene movement within- and between-chromosomes in the evolution of *Drosophila* genomes. *Gene*, 2006, 385: 96—102 [DOI]
- 29 贺华良, 周惠, 肖振东, 等. 果蝇 3 个新的小分子非编码 RNA 的鉴定. *科学通报*, 2006, 51(20): 2393—2398
- 30 Deng W, Zhu X, Skogerbo G, et al. Organization of the *Caenorhabditis elegans* small non-coding transcriptome: Genomic features, biogenesis, and expression. *Genome Res*, 2006, 16: 20—29 [DOI]
- 31 Hayman M L, Miller M M, Chandler D M, et al. The trypanosome homolog of human p32 interacts with RBP16 and stimulates its gRNA binding activity. *Nucleic Acids Res*, 2001, 29(24): 5216—25 [DOI]
- 32 Deng X, Meller V H. Non-coding RNA in fly dosage compensation. *Trends Biochem Sci*, 2006, 31(9): 526—532 [DOI]
- 33 Akhtar A, Zink D, Becker P B. Chromodomains are protein: RNA interaction modules. *Nature*, 2000, 407: 405—409 [DOI]
- 34 Rattner B P, Meller V H. *Drosophila* male-specific lethal 2 protein controls sex-specific expression of the roX genes. *Genetics*, 2004, 166: 1825—1832 [DOI]