

Wolbachia 中一个新的可转移遗传因子的分离和鉴定*

周伟国 甘波谊 赵新燕 赵寿元 李昌本**

复旦大学生命科学学院遗传学研究所, 遗传工程国家重点实验室, 上海 200433

摘要 *Wolbachia* 是一类寄生于许多节肢动物细胞质内的类立克次氏体细菌. 该细菌与广泛存在于节肢动物群体内的生殖、发育异常现象相关. 为了研究这些现象的分子机制, 利用 RDA (representational difference analysis) 方法分析了不同表型 *Wolbachia* 的基因组差异. 发现在强烈诱导胞质不相容性的 *Wolbachia* 品系 wRi (存在于果蝇 *Drosophila simulans* Riverside DSR 品系) 的基因组中, 存在一个潜在的可转移遗传因子, 将其命名为 WISE (*Wolbachia* insertion sequence element).

关键词 *Wolbachia* 胞质不相容 可转移遗传因子

胞质不相容性 (cytoplasmic incompatibility, CI) 是存在于多种昆虫群体内的一种胚胎发育异常现象, 这种现象是由呈细胞质遗传方式的类立克次氏体细菌 *Wolbachia* 引发的. 当感染了 *Wolbachia* 雄性个体与非感染的雌性个体交配时, 通常不产生或者产生极少数目的子代. 而如果雌性个体感染了 *Wolbachia*, 与任何雄性个体交配均能产生正常数目的子代. 这样, 感染了 *Wolbachia* 的宿主相对于非感染的同种昆虫具有生殖上的优势, 从而能在一个群体中快速传播^[1]. 对于这一现象的群体生物学和细胞学研究已经进行得相当深入, 一般认为, 这一现象的产生是由于 *Wolbachia* 结合了宿主的某种生殖细胞特异染色体蛋白质并在随后精子发生过程中将这种蛋白质携至精细胞外, 从而影响了受精后的第 1 次有丝分裂过程^[2]. 不同品系的 *Wolbachia* 能产生不同的 CI 表型. 为了研究 CI 的分子机制, 我们采用了 RDA (representational difference analysis) 方法^[3] 分析不同表型 *Wolbachia* 之间的差异. 结果从 wRi *Wolbachia* 品系分离到一个潜在的可转移遗传因子并将其命名为 WISE (*Wolbachia* insertion sequence element).

1 材料与方法

1.1 果蝇品系

Drosophila simulans Riverside (DSR), *Drosophila simulans* DSW/Mau (Mau), *Drosophila simulans* Noumea (R3A), *Drosophila melanogaster* CantonS (DM), *Drosophila melanogaster*^{YW67c23} (YW), 均携带共生菌 *Wolbachia*, 其命名依次为 wRi, wMa, wNo, wMelCS, wMel. 其中 wRi, wNo, wMel 能诱导 CI 表

2000-08-23 收稿, 2000-09-14 收修改稿

* 美国 McKnight 基金和中国博士后基金资助项目

** 联系人

型,而 wMa, wMelCS 则不能^[4]. 5 种果蝇经四环素处理后体内的 *Wolbachia* 均被清除,其简称为 DSRT, MauT, R3AT, DMT, YWT.

```

1      TCGGTCTCAA TTTACCATG CGGTGACCCG GGAGATCTGA ATTCAAGCCT TTTACTTGT
61     TCCGGACAAC CTCTCATACC AATTTCTATT ACATAACAAA CAAATAGACA ATAGGAGTGA
121    GAAGGTTATA CTAAGGCAGT ATAATGATGG AGTTTGGAA TGGCATTGGT ATCAGGTTGC
181    AATTTTAGCT TACATTGTGC CAATCTTCT TCAATTGCTG CTCGCATAAA TTCTGCCTGT
241    CTTTTAGTTG TATTAATAA GACTTGACCT GACGCTTAA AAAAGTAAGT TAAAAATAA
      **
301    GGAGTGTTAG ATATGAGTAC AATACAGACA AAAATATTAA AACCAAAGCT AGAGAACTAG
361    CAAAACAACCT AGGAAATGTG TCTCAAGCAT GCAAAGTGAT GGGATACTCA AGAGATACAT
      M G Y S R D T
F421   TTTATCGCTT TAAGGAGTTA TATGAAATG GAGGAGAGGA AGCATTACAT GAAATAAGTA
      Y R F K E L Y E N G G E E A L H E I S K
481    AGAAAAAACC GCTATTAGCA AACAGAGTTT CCGACGATAT AGAAAGAGCA GTGATTGGTA
      K K P L L A N R V S D D I E R A V I G I
541    TAGCAACAGA ATTTCCAGCA TATGGGCAAG AAAGAGCTGC AAATGAACTG AGAAAAAGAG
      A T E F P A Y G Q E R A A N E L R K R G
601    GTATAATAAT TTCTCAGGGA GGAGTAAGGT CTGTATGGCT AAGAAATGAC CTGAAACTC
      I I I S Q G G V R S V W L R N D L E T L
661    TCAAAAAGAG ACTTAAAGCA CTAGAGACAA AAGTAGCTCA AGACGGAATC ATTTTAACTG
      K K R L K A L E T K V A Q D G I I L T E
721    AAGAACAAC TGCAGCTTA GAAAAGTGA AAGAACAAAG GGAAGCTCAT GGTGAAATG
      E Q L A A L E K V K E Q R E A H G E I E
781    AGACGCAACA TCCAGGTTAT TTGGGTCTC AAGATACCTA TTATGTGGGC AATATCAAAG
      T Q H P G Y L G S Q D T Y Y V G N I K G
841    GTATAGGGCG CATTATCAG CAACTTTTG TTGATACTTA TTCCAGAGTT GCAATGGTTA
      I G R I Y Q Q T F V D T Y S R V A M V K
901    AACTTTACAC GGACAGAACA GCTATTACAG CTGCAGATCT TCTCAACGAT AGAGTTATTC
      L Y T D R T A I T A A D L L N D R V I P
961    CATTTTTTGA TGAGCAGAAA ATTCATTAT TACGCATTTT AACTGATAGG GGTACGGAGT
      F F D E Q K I P L L R I L T D R G T E Y
1021   ACTGTGGCAA GCCCGAAAAT CACGCTTATC AGCTTTATTT GGCATAGAA AACATCGATC
      C G K P E N H A Y Q L Y L G I E N I D H
1081   ATTCTAGAAC TAAAGCCAAC TCTCCACAGA CTAATGGCAT ATGTGAAAGG TTTCATAGAA
      S R T K A N S P Q T N G I C E R F H R T
1141   CCATGCAAGA TGAGTGTTAC AATATCATCT TTAGAAAGAA AATCTACAGT TCTTTGGAAG
      M Q D E C Y N I I F R K K I Y S S L E D
1201   ATCTACAGAT TGATGTTGAT CATTGGTTGC GTTCTTATAA TGAGACAAGA CCTCACTCGG
      L Q I D V D H W L R S Y N E T R P H S G
1261   GTAAATATG CTATGGCAAA ACGCCTATGC AAACTTTCT TGATAGCAAA TATATTGCTT
      K Y C Y G K T P M Q T F L D S K Y I A F
1281   TTCAGAAAAA TATTAGTAGC ATTACACAAG AGCCTGATAT TAGTTTTGAC TACCTCAATT
      Q K N I S S I T Q E P D I S F D Y L N S
1341   CTTCGTGTCG TTAATCTTG TCTGTGAGAT CAAATCTTGT CTTTACATT TTAACATTG
      S V S
1441   TAAATTACC AAGGATTATG GAGAAGATAT GCTGTTTTGA TTTGTGTGAA AATTATTGTT
1501   GACTTATGTG CAATTAGAGA ACCTAAGTCA CTTAGCTAT AATTCAGAAT TGAGAGCTTT
1561   TAATAACTCT CTAACCTCCA CTTAGTCTT GGTGCAAAT TAAGTGATA

```

图 1 WISE 的结构

开放阅读框(ORF)以下划线标出,反向重复序列以箭头标出,两个正向重复的碱基下面以星号(*)指示

1.2 WISE 的分离

RDA 分析参见文献[3], RDA PCR 产物直接克隆连接至 pGEM-T 载体(Promega). 挑选含有插入片段的转化子进行序列测定. 所得顺序经 BLASTX 分析[5], 并选定与已知转座酶具有较高同源性的片段作进一步研究. 通过连接介导多聚酶链式反应(ligation-mediated PCR)获得该片段上下游序列[6].

1.3 Southern 印迹分析

取 5 μg 果蝇总 DNA 经 *Bgl* II 完全酶解后进行 Southern 印迹分析, 以转座酶基因的特异性引物 45F(5' CCA TGG GAT ACT CAA GAG AT 3')和 RDAR1 (5' AAT GAT CGA TGT TCT CTA TGC 3')扩增片段作为杂交探针. 同一张膜放射自显影后洗脱探针, 与 PCR 扩增的 *Wolbachia wsp* 基因片段[7]再次进行杂交.

2 结果与讨论

以 RDA 分析获得的 DNA 片段为起始, 经过 LM-PCR 扩增上下游序列, 我们获得了一段长约 1.6kb 的序列, 为了排除 *Taq* 酶扩增引入的差错, 以确定基因的确切序列, 我们测定了 3 个独立克隆的序列, 选定 3 个序列的一致顺序作为最终的正确顺序. 该序列包含了一个长 993bp 的完整开放阅读框, 对其进行的 BLASTX 分析表明, 这一段序列编码的多肽序列与一些已知的转座酶有较高的同源性. GCG 分析在开放阅读框的上下游识别出一对不完全匹配的反向重复序列 IRS (inverted repeat sequence), 31 个碱基中, 有 21 个能够配对. 在 IRS 的两侧, 则有两个碱基的顺向重复(图 1). 这些发现都符合可转移遗传因子的特征[8], 因此, 这可能是一个潜在的可转移遗传因子, 我们将其命名为 WISE.

Southern 印迹实验表明在 wRi 品系 *Wolbachia* 的基因组中, 至少含有 5 个 WISE 的拷贝(图 2(a)). 这一结果进一步证实了 WISE 极有可能是一个可转移的遗传因子, 在其他品系中未能检测到明显的阳性信号. 以 *wsp* 为探针的杂交结果证实了电泳上样总 DNA 中含有的 *Wolbachia* DNA 已足够进行 Southern 检测, 因此 WISE 阴性结果确是因为在基因组中没有这种因子(图 2(b)). 仅在 wMa 品系中出现很弱的信号, 这也许表明 wMa 品系的基因组中存在其他与 WISE 有一定同源性的可转移遗传因子. 不含 *Wolbachia* 的果蝇总 DNA 中没有阳性信号, 说明 WISE 的确来自 *Wolbachia*, 而并非是果蝇 DNA 的污染.

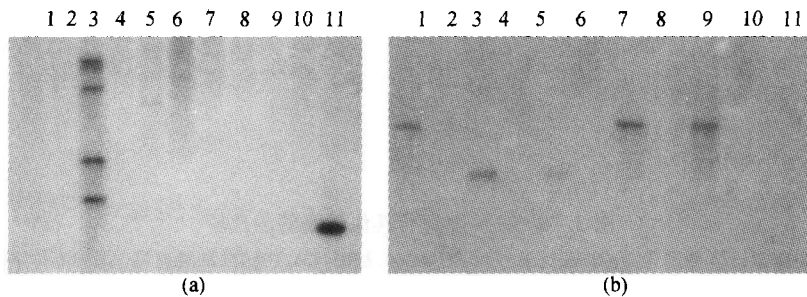


图 2 Southern 印迹法检测 WISE 在 *Wolbachia* 基因组中的存在及拷贝数
(a) WISE 转座酶探针杂交结果; (b) *wsp* 基因杂交. 两者系对同一张膜杂交的结果. 1~11 分别为 DM, DMT, DSR, DSRT, Mau, MauT, R3A, R3AT, YW, YWT, WISE 转座酶探针

WISE 的转座酶与数种转座酶表现出程度不等的同源性,几种同源性较高的转座酶之间的比较见图 3. 其中,对应于 WISE 转座酶 127 ~ 140, 193 ~ 209, 235 ~ 254 位氨基酸残基的 3 个区域保守性较强. 估计这些区段对转座酶的活性较为重要.

```

1  MGYSRDTFYRFKELYE---NGGEEALHE-----ISKKKPLLANRVSD--DIERAVIGIA
2  MRSYAGSDIVER-LMRRWCLCQIQARSWGWKDSVMPAPMKTQAVDGFPSGVFDKHALV-GR-
3  MDREIQQR-----LQ--W-VK-----MYEECG---DAGLVCRRCG
4  MKYIINNENLPD-LQAK--LKNFLSKNYKNKYKRIKQKIFSYINLCNDYYNGNFLKDLI

1  TE-FPA-----YGQ-----ERAAN---ELRKRGIIISQGGVRSVW
2  GRRLTVEDRVAIEAGCRVGDSARAIAQKINRHHSVVAREITRNGWKIVDEDEGTEQ---LR
3  -----ISRPTLRKWAKR----YKQCGIAGLESQSRRPHS---SPDTKLTDELRALILTMR
4  KKYFKNKFSTFYWANKILIAYKSNDFAELLKSTIPNNINYQYSNDVRQNICDLYFKYC

1  LRNDLE--TLK---KR--LKALETKVAQDGII-----LT-EEQ--LAALEKVKEQRE
2  YNAHNAAVSTAGRMVRPKLRKLDESPTLRGVVVDCLARRWSPGRISAWLEH--AFSDDES
3  DKRNLGARRLQTELIR--LHKIHLSTATLHKV-----LS-EAS--VKPIVTYRRKKD
4  NKHAGGVLSLFYNLKKG-IHGEELKNKAPKNLKTFF--RWLK-KDERWLKIKNKIKEIKK

1  AHGE-IETQHPGYLGSQDTYYVGNI-----K--GIGRIYQOTFVDTY
2  MRISH-EAIYSALYI-QKGSLRAELEEVMKTKDVLIRGGSTRKR----RARNAGVLTGR
3  -FQRYERPIPGDRV--QMDTCKIAPG-----IYQYTAIDDC
4  QH-SRYEVKEIG-LL-QMDAKYFVP-----SKFSVDKKYFVYDFIDEK

1  SRVAMVKLYTDRTAITAADLLNDRVIPFFDEQKIPLLRILTDRGTEYC--GK-----
2  PWIKGAEITHRSPEADDRAIPGHWEGDLVIGKGGKSALITLVERTSRYTLLGHLPDEHSS
3  SRYRVLRCYSRRTAANTVDFI----DCVVEEMPFPIQRIQTDRGREFF--AE-----
4  TRLALGVYDKLSTDNAIDAVKKAISDFKNIFGIIITRI-RTDNGSEFI--NNYRNNQKI

1  -PENHAYQLYLGIENIDHSRTKANSPQTN-GICERFHRTMQDE-CYNIIFRKKIYSSLED
2  HTVVATLQDMVKDLNTEQLKTITWDQGAEMAVTAQVQIKDGCQVFFCEPHSPWQRPTNEN
3  -KVQQLMIY---GIKFRPNKPGSPHLN-GKVERSQKTDKSEF-YPTIDVS---VGLQE
4  SVKKTNFTQFLTDKNILHQTTPVRSPQSN-GKIERFHQNYTKLFVFEE---KILNAVS-

1  LQIDVDHWLRSYNETRPHSGKYCYGKTPMQTFLDSKYI-----AFQKNISSITQEP--D
2  TNGEIRR--RFY-----KKGTDFATVTPEHVAWVQNELNETPRQILGGATPR-----E
3  LDLLLAEWQHYYNWERPHSSLNGLTPIDRITEISDQTPLSEEVSQNYQIKKERFQEQNYK
4  LQNKLNDYYYFYNFERVHKSL-----NFQ

1  ISFDYLN--SSVS
2  ILNELFKRGASTA
3  LDLQLRK-LKPSL
4  TPFNFLNSLSLIK

```

图 3 WISE 转座酶与其他转座酶的比较

1 为 WISE 转座酶, 2, 3, 4 分别为 GenBank 登录号为 AF024666, AF034434, U28974 的转座酶. 图中黑体字部分为保守性较强的氨基酸残基

自从 RDA 方法在 1993 年建立起来, 在许多方面得到了广泛的应用. 利用这一方法, 研究者们分离到了与肿瘤相关的基因片段或标记. *Wolbachia* 由于不能在体外培养, RDA 方法尤其显得有应用价值. 一般而言, RDA 方法较适用于获得仅存在于目的基因组中的 DNA, 或者差异

较大的 RFLP 标记. 我们的研究成功地分离到了一系列 DNA 片段, 其中与已知转座酶有较高同源性的一个片段尤其引起了我们的兴趣, 这是因为在前期的研究中, 我们曾以 *Wolbachia* 染色体基因为基础建立了分子系统发生, 发现 *Wolbachia* 品系诱导 CI 的能力与其在系统发生的位置无关^[4], 也就是说, 相近品系的 *Wolbachia* 并不表现同样的 CI 诱导表型, 而同样表达 CI 的品系则可能亲缘关系较远. 因此, 有理由认为, *Wolbachia* 诱导 CI 的基因(簇)可能位于某种可转移的遗传因子上. 所以我们推测 WISE 可能与 CI 现象有某种联系. 然而, 就我们目前的结果来看, WISE 与 CI 现象并没有直接联系. 比如, wRi, wNo, wMel 同样能诱导 CI, 但只有 wRi 的基因组内有 WISE 存在的证据. 但我们并不能排除这样一种可能性, 即 WISE 的一个拷贝也许位于某种质粒或者噬菌体因子上, 除 WISE 外, 该质粒或噬菌体携带了与 CI 相关的基因(簇).

之前, Masui 等曾报道了从 wTai *Wolbachia* 品系中分离到一种插入序列^[9]. 可转移遗传因子可能较普遍地存在于多种 *Wolbachia* 的基因组中. 可转移遗传因子在 *Wolbachia* 的理论研究或实践应用中都有较重要的意义, 可作为一种诱变因子产生各种 *Wolbachia* 突变型从而有助于对 *Wolbachia* 的遗传分析. 此外, 我们用可转移遗传因子转化 *Wolbachia*, 可使其携带针对某种病虫害的抗性基因, 从而降低其宿主群体传播疾病的能力.

参 考 文 献

- 1 Werren J H. Biology of *Wolbachia*. *Annu Rev Entomol*, 1997, 42: 587
- 2 Callaini G, et al. *Wolbachia*-induced delay of paternal chromatin condensation does not prevent maternal chromosomes from entering anaphase in incompatible crosses of *Drosophila simulans*. *J Cell Sci*, 1997, 110: 271
- 3 Lisitsyn N, et al. Cloning the differences between two complex genomes. *Science*, 1993, 259: 946
- 4 Zhou W, et al. Phylogeny and PCR-based classification of *Wolbachia* strains using wsp gene sequences. *Proc R Soc Lond B Biol Sci*, 1998, 265: 509
- 5 Altschul S F, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nuc Acid Res*, 1997, 25: 3389
- 6 Mueller R, et al. Ligation Mediated PCR for Genomic Sequencing and Footprinting. In: Ausubel F M, et al. eds. *Current Protocols in Molecular Biology*. Vol 2. New York: John Wiley & Sons Inc, 1992, 15. 5. 1 ~ 15. 5. 20
- 7 Braig H R, et al. Cloning and characterization of a gene encoding the major surface protein of the bacterial endosymbiont *Wolbachia pipientis*. *J Bacteriol*, 1998, 180: 2373
- 8 Mahillon J, et al. Insertion sequences. *Microbiology and Molecular Biology Reviews*, 1998, 62: 725
- 9 Masui S, et al. The first detection of the insertion sequence ISW1 in the intracellular reproductive parasite *Wolbachia*. *Plasmid*, 1999, 42: 13