

八肋游仆虫大核基因组中组蛋白基因 *H4A* 和 *H4B* 的克隆及分析

杨旭霞 许 静 梁爱华 王 伟*

山西大学生物技术研究所 化学生物学与分子工程教育部重点实验室 太原 030006

摘要:组蛋白作为核小体的基本组分,是染色质的结构和功能必需的。组蛋白的变体和修饰共同参与染色质修饰及基因的表达调控。真核生物细胞中的 5 种组蛋白在进化中高度保守,然而纤毛虫的组蛋白 H4 与其他真核生物相比有较大的差异。本实验应用 PCR 技术从八肋游仆虫 (*Euplotes octocarinatus*) 中获得了 2 种组蛋白 H4 基因,分别为 *H4A* 和 *H4B*, GenBank 登录号为: JN715068 和 JN715069。序列分析表明, *H4A* 基因开放阅读框 324 bp, 预测编码 107 个氨基酸, 分子量为 11.6 ku, 等电点为 10.99。而 *H4B* 基因编码框 384 bp, 编码 127 个氨基酸, 分子量为 14.4 ku, 等电点为 9.93。Blast 结果显示, *H4A* 序列与其他生物中 H4 的一致性相对较高, 达 81% ~ 94%, 而 *H4B* 的一致性为 36% ~ 70%。 *H4A* 和 *H4B* 的一致性仅为 44.7%。实时荧光定量 PCR 表明, *H4A* 的转录本高于 *H4B*。结果提示: 在进化过程中八肋游仆虫可能进化出特殊的组蛋白 H4 基因, 不同的组蛋白 H4 可能发挥不同的功能。

关键词:八肋游仆虫; 克隆; 组蛋白 H4; 转录

中图分类号: Q953 **文献标识码:** A **文章编号:** 0250-3263(2012)03-73-08

Cloning and Analysis of *H4A* and *H4B* Genes from the Macronucleus of *Euplotes octocarinatus*

YANG Xu-Xia XU Jing LIANG Ai-Hua WANG Wei*

Key Laboratory of Chemical Biology and Molecular Engineering of Ministry of Education,
Institute of Biotechnology, Shanxi University, Taiyuan 030006, China

Abstract: Histones are basic components of nucleosome and essential to chromatin structure and function. Histone posttranslational modifications and sequence variants jointly participate in modification of chromatin and regulation of gene expression. Histones are highly conserved proteins in different organisms. However, a high degree of variation was found in ciliate. In the study, *H4A* and *H4B* were cloned from *Euplotes octocarinatus* by PCR, GenBank accession number: JN715068 and JN715069. Sequences analysis showed that opening reading frame of the *H4A* gene was 324 bp, which encoded a 107 amino acid polypeptide with a predicted molecular mass of 11.6 ku and isoelectric point of 10.99. The opening reading frame of the *H4B* gene was 384 bp, which encoded a 127 amino acid polypeptide with a predicted molecular mass of 14.4 ku and isoelectric point of 9.93. *H4A* gene shared a high identity of 81% - 94% with reported *H4* gene, while *H4B* shared an identity of 36% - 70% with *H4* gene from other eukaryotes. *H4A* shared 44.7% identity with *H4B*. Real-time PCR showed that

基金项目 国家自然科学基金项目 (No. 30770295, 31072000), 山西省自然科学基金项目 (No. 2008011063);

* 通讯作者, E-mail: gene@sxu.edu.cn;

第一作者介绍 杨旭霞, 女, 硕士研究生; 研究方向: 细胞分子生物学; E-mail: 200923002016@mail.sxu.cn.

收稿日期: 2011-10-24, **修回日期:** 2012-02-27

H4A transcript was higher than that of *H4B*. The results indicate that the *H4* gene has different variants in *E. octocarinatus*.

Key words: *Euplotes octocarinatus*; Cloning; Histone H4; Transcription

染色质作为真核生物遗传信息的载体,其基本结构与功能单位是核小体。组蛋白作为核小体的基本组分,是染色质的结构和功能所必需的。组蛋白的成分除了经典的 H2A、H2B、H3、H4 和 H1,还包括其他的组蛋白变体,组蛋白的变体和修饰共同参与染色质表观遗传调控^[1]。处于不同状态的染色质需要相应的组蛋白变体维持结构,以完成其生物学功能。对组蛋白变体的研究是近年来表观遗传学新的研究热点,也是对“组蛋白密码”新的诠释。真核生物的 5 种组蛋白中,组蛋白 H4 最为保守^[2-3],而纤毛虫的组蛋白 H4 与其他真核生物相比出现较大的差异,且不同纤毛虫之间组蛋白 H4 也表现出较大的差异^[4-5]。这种特殊性可能是由于纤毛虫特殊的细胞核结构和染色体的组织所决定,然而确切的机制仍不清楚。

八肋游仆虫(*Euplotes octocarinatus*)和其他大多数纤毛虫一样,在同一细胞体内既含有负责营养的大核,也含有负责有性生殖的小核。特别是游仆虫大核包含“gene-sized”微小染色体,每一条染色体中均无着丝粒结构^[6-8]。在细胞营养生长期,大核是细胞生长所必需的 RNA 的转录和合成的场所,它决定了细胞的表型,小核无转录活性;在有性生殖阶段,新的大核及小核由亲本小核通过减数分裂和有丝分裂形成的合子核发育而来,在此过程中组蛋白参与了新大核发育过程基因组的删除、断裂和扩增^[9]。游仆虫核二态性及染色体的特殊组成暗示组蛋白可能进化出不同的结构和修饰类型。

本研究从八肋游仆虫大核基因组中,克隆得到了 2 种序列存在较大差异的组蛋白 *H4* 基因变体,实时荧光定量 PCR 在 mRNA 水平对 2 种组蛋白 *H4* 进行了转录水平的比较,研究结果为进一步探讨纤毛虫中组蛋白 *H4* 的进化及分析组蛋白的新功能奠定了基础。

1 材料与方法

1.1 细胞株 八肋游仆虫、绿梭藻虫(*Chlorogonium longatum*)由本实验室培养;大肠杆菌(*Escherichia coli*) DH5 α 为本实验室保存。

1.2 试剂和工具酶 DNA 回收试剂盒(Gel Extraction Mini Kit)购自 BioMIGA 公司,质粒抽提试剂盒(Plasmid Mini Kit)购自 OMEGA 公司,Easy Taq DNA 聚合酶购于北京全式金生物技术有限公司,pMD18-T Vector 购于 TAKARA 公司,PCR 引物合成由华大基因公司完成,DNA 序列测定由北京奥科生物公司完成,SYBR Green Real-time PCR Master Mix 购自 TAKARA 公司。

1.3 八肋游仆虫细胞的培养 八肋游仆虫培养于 SMB 培养液中,以绿梭藻虫为食物,最适生长温度为 22 $^{\circ}\text{C}$ ^[10]。

1.4 八肋游仆虫大、小核 DNA 的提取 八肋游仆虫大小核 DNA 的提取见相关文献^[11-12]。

1.5 八肋游仆虫总 RNA 的提取及 cDNA 的合成 八肋游仆虫总 RNA 的提取方法见参考文献^[13]。紫外分光光度计测量其 A_{260} 与 A_{280} 值,估算 RNA 浓度与纯度。利用购自 TaKaRa 的 RNase-free DNase I 处理总 RNA,以除去总 RNA 中的污染 DNA。cDNA 的合成参照 TaKaRa 公司的 cDNA 合成试剂盒说明。

1.6 PCR 扩增和测序 八肋游仆虫中每个大核微小染色体两端都具有相同的端粒序列,据此设计端粒引物 Telo:5'-CCC CAA AAC CCC AAA ACC CCA AAA C-3'。根据组蛋白 H4 保守的氨基酸序列设计特异引物 H4A1F:5'-GGT AAA GGT GGA AAG GGA GGA AAG-3'; H4A1R:5'-CAC ATC AAT GGG AAG GAT AGT C-3'。以 H4A1F、H4A1R 为引物,游仆虫大核 DNA 为模板,进行 PCR 扩增,PCR 反应条件为:94 $^{\circ}\text{C}$ 4 min;94 $^{\circ}\text{C}$ 30 s,50 $^{\circ}\text{C}$ 30 s,68 $^{\circ}\text{C}$ 30 s,

共 35 个循环;68℃ 10 min。PCR 扩增产物经 1% 琼脂糖凝胶电泳回收纯化,将目的片段克隆于 PMD18-T 载体,测序分析。根据测序结果在该基因的上下游分别设计引物 H4AF: 5'-ATG GCC GGA AGA GGT AAA GGT G-3' 和 H4AR: 5'-ACC ACC GAA TCC GTA GAG GG-3'。以 H4AF 和 H4AR 为引物,分别以大核 DNA 和 cDNA 为模板进行 PCR 扩增,获得了 *H4A* 基因的完整编码框序列。

H4B 基因的克隆,设计引物 H4B1F: 5'-GCT ACA AAG TCT CAG GTA TCT CTG-3' 和 H4B1R: 5'-CCA CAA AGC GAG TAA AAA TAG-3', 以 H4B1F、H4B1R 为引物,游仆虫大核 DNA 为模板,进行 PCR 扩增,获得长度为 120 bp 的 *H4* 基因片段。根据这段序列,在该基因的上下游设计引物 H4B2F: 5'-CAA ACT CGC AAC TGA CAG CTA C-3'; H4B2R: 5'-GTC TTT CTC TTG GCA CAC TCG C-3'; 以 H4B2F、Telo 为引物,八肋游仆虫大核基因组为模板,PCR 扩增获得基因上游到下游端粒的序列;以 H4B2R、Telo 为引物进行 PCR 扩增,获得了该基因下游到上游端粒的序列。根据 2 次测序结果设计特异引物 H4BF: 5'-ATG GTT GCA ATG AAG AGA AAG TC-3'; H4BR: 5'-CTT CTT TTT CTT GCT GCC TTC-3', 以大核基因组、cDNA 为模板,进行 PCR 扩增,得到了 *H4B* 基因的编码区序列。大核 *H4* 基因的克隆过程见图 1。

1.7 序列分析及同源性比对 将获得的 2 种 *H4* 基因进行序列分析, BLAST 比对。运用 DNASTar 软件预测 2 种 *H4* 基因拟编码蛋白的氨基酸序列,分析拟编码蛋白的特征。采用 MEGA4 软件中的邻接法 (neighborjoining method, NJ 法) 进行组蛋白 H4 的聚类分析, bootstrap 值为重复 1 000 次。

1.8 实时荧光定量 PCR 检测基因的表达差异

根据测序得到的基因序列的特点设计 Real time PCR 引物, AF: 5'-CTA TCA GAA GAT TAG CAA GAA GAG GAG-3'; AR: 5'-TCT TGG CGT GTT CAG TGT ATG-3'。BF: 5'-GTG GAC AAA CTC GCA ACT GAC A-3'; BR: 5'-GTC ATA GAC ATA GCC ATA GAG GTT TC-3'。内参基因选择 β -tubulin rRNA, 引物为 β F: 5'-CTG GAA ACA ACT GGG CTA AGG-3'; β R: 5'-GAC GAC GGT ATC TGA GAC TTT TG-3'。反应体系为 20 μ l; SYBR Green real-time PCR Master Mix 10 μ l, 10 μ mol/L 的上下游引物各 0.8 μ l, ROX 0.4 μ l, 10 \times cDNA 模板 1 μ l, ddH₂O 7 μ l。反应在 Stepone Plus Real Time PCR System Cyclor (Applied Biosystem) 上进行, 反应程序为: 95℃ 15 min 预变性, 95℃ 15 s 变性, 50℃ 30 s 退火, 68℃ 35 s 延伸, 样品相对于对照 (未处理样品) 的表达通过 β -tubulin RNA 进行校正。每个样品做 3 个平行, 稀释成 5 个浓度梯度构建标准曲线; 数据使用 Real

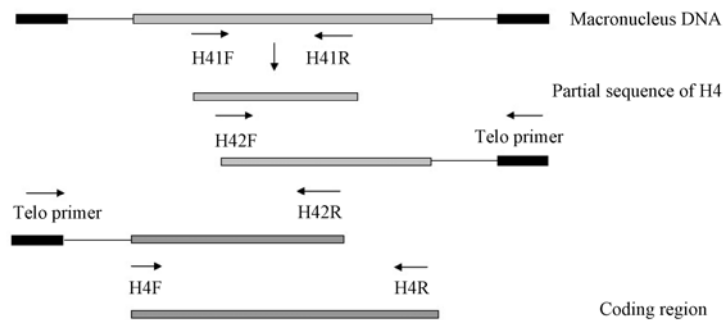


图 1 八肋游仆虫大核基因组中 *H4* 的扩增策略

Fig. 1 The PCR walking strategy for amplification of *H4* from *Euplotes* macronuclear DNA

“Macronucleus DNA”表示 *H4* 的大核 DNA 序列; “Partial sequence of *H4*”表示 *H4* 基因的部分序列; “coding region”表示 *H4* 基因的编码区序列; 端粒用黑色框表示, 编码区用灰色框表示, 非编码区用线条表示。Black boxes represent telomeric repeats, gray boxes represent coding regions, lines represent non coding regions.

Time PCR System Cycler 自带软件进行分析, Ct 值及扩增效率依照 Pfaffl 计算方法进行^[14]。

2 结果

2.1 *H4A* 和 *H4B* 基因序列的克隆及分析

以 H4A1F 和 H4A1R 为引物, 八肋游仆虫大核 DNA 为模板, PCR 扩增出一段 261 bp 的片段。测序结果比对发现: 该段序列预测编码的蛋白序列与其他生物的组蛋白 H4 具有较高的一致性。进而设计引物 H4AF 和 H4AR, 分别以大核 DNA、cDNA 为模板进行 PCR 扩增, 得到 324 bp 的片段(图 2A)。测序表明: *H4A* 基因开放阅读框为 324 bp, 编码 107 个氨基酸, 开放阅读框起始于 ATG, 终止于 TAA。

以 H4B1F 和 H4B1R 为引物, 扩增出 120 bp 的片段。测序得到的序列拟编码的蛋白序列同样与组蛋白 H4 具有较高的一致性。根据测序结果, 在该片段的 5' 端设计特异引物 H4B2F, 以 Telo 引物为下游引物, 大核 DNA 为模板进行 PCR 扩增, 结果得到了序列长度为 339 bp 的片段。再以该基因片段的 3' 端设计特异性引物 H42R, 以 Telo 引物为上游引物, PCR 获得序列长度为 312 bp 的片段。将 2 次测序结果拼

接, 设计特异引物 H4BF、H4BR。以 H4BF 和 H4BR 为引物, 分别以大核 DNA、cDNA 为模板进行 PCR 扩增, 均得到 384 bp 的序列(图 2B)。*H4B* 基因所在的大核染色体全长 633 bp, 两端含有非编码区和端粒序列。5' 端由 20 核苷酸 (C₄A₄C₄A₄C₄) 的端粒序列及另外 73 个核苷酸的非编码区组成。3' 端包括 132 个核苷酸的非编码区及后面的 24 个端粒核苷酸。开放阅读框为 384 bp。编码 127 个氨基酸。开放阅读框内含有一个 TGA (232~234) 密码子, 在游仆虫中编码半胱氨酸^[15]。开放阅读框起始于 ATG, 终止于 TAG。RT-PCR 表明, *H4A* 和 *H4B* 2 个基因均有转录活性, 且无内含子。GC 含量分析表明, *H4A* 和 *H4B* 的编码区的 GC 含量分别为 27.7% 和 24.9%, 与游仆虫基因组的 GC 含量相当。

2.2 *H4A* 和 *H4B* 基因拟编码的蛋白序列分析

应用 DNASTar 软件分析表明, *H4A* 基因拟编码的蛋白由 107 个氨基酸残基构成, 预测其分子量 11.6 ku, 等电点 10.99, 且在该蛋白的 N 端第 6、9、12、16、20 位均为赖氨酸, 这与组蛋白 H4 的甲基化及乙酰化的修饰位点相一致(图 3A)。对该蛋白序列进行 BLAST 分析表明, 该

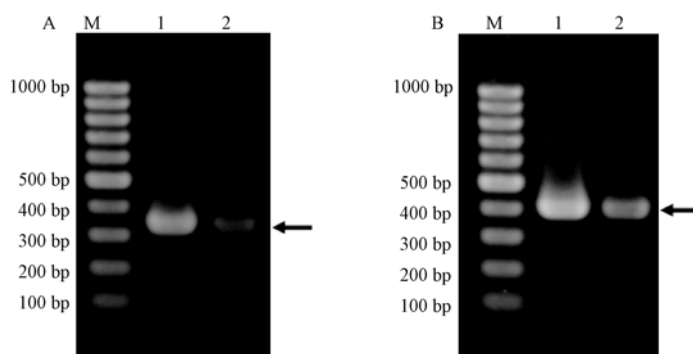


图 2 *H4A* 和 *H4B* 的 PCR 扩增

Fig. 2 The PCR products of *H4A* and *H4B*

A: M. 100 bp 分子量标记; 1. 以大核 DNA 为模板扩增到的 *H4A* 基因开放阅读框; 2. 以 cDNA 为模板扩增到的 *H4A* 基因开放阅读框。B: M. 100 bp 分子量标记; 1. 以大核 DNA 为模板扩增到的 *H4B* 基因开放阅读框;

2. 以 cDNA 为模板扩增到的 *H4B* 基因开放阅读框。

A: M. 100 bp DNA ladder; 1. The opening reading frame of the *H4A* gene from macronuclear DNA; 2. The opening reading frame of the *H4A* from cDNA. B: M. 100 bp DNA ladder; 1. The opening reading frame of the *H4B* gene from macronuclear DNA;

2. The opening reading frame of the *H4B* from cDNA.

蛋白与小腔游仆虫 (*Euplotes aediculatus*) 的组蛋白 H4 (AAM77588.1) 一致性最高, 高达 94%, 说明该基因编码蛋白属于一种八肋游仆虫的组蛋白 H4。

H4B 基因拟编码的蛋白由 127 个氨基酸残基构成, 预测其分子量 14.4 ku, 等电点 (pI) 为 9.93, 且在 N 端的第 5、7、12、16、20 位均为赖氨酸 (图 3B)。对该蛋白序列进行 BLAST 分析表明, 该蛋白与扇形游仆虫 (*E. vannus*) 的组蛋白 H4 (CAA06067.1) 一致性最高, 达到 70%, 说明该基因编码蛋白同样属于组蛋白 H4。

2.3 H4 蛋白的聚类分析 将 *H4A*、*H4B* 蛋白与其他游仆虫的组蛋白 H4 进行比对分析。结

果表明, 阔口游仆虫 (*E. eury stomus*) 的 2 种组蛋白 H4 (GenBank: CAA06072.1 和 CAA06071.1) 序列仅有 1 个氨基酸的差异, 而扇形游仆虫 2 种组蛋白 H4 (GenBank: CAH04403.1 和 CAA06067.1) 的一致性仅为 50%。八肋游仆虫的 *H4A* 与 *H4B* 的氨基酸序列一致性仅为 44.7%。说明在游仆虫细胞中存在 2 种不同的组蛋白 H4 变异体。将游仆虫与其他生物的 H4 蛋白进行聚类分析 (图 4), 结果表明, 八肋游仆虫的 *H4A* 蛋白与小腔游仆虫、厚体游仆虫 (*Moneuplotes crassus*) 的 H4 聚在一起, 尽管 *H4B* 蛋白与阔口游仆虫、扇形游仆虫的 H4 蛋白聚在一起, 但相对于 *H4A*, *H4B* 更为特殊。

```

A  ATGCCCGGAAGAGGTTAAAGGTGGAAAAGGAGGAAAGGGATACGGAAAGGTTGGAGCCAAG60
    M A G R G [K] G G [K] G G [K] G Y G [K] V G A [K]
    AGACACTAAGAAGGCTCTTAGAGAGACCATCTTGGAGTCACCAAGCCAGCTATCAGA 120
    R H T K K A L R E T I L G V T L P A I R
    AGATTAGCAAGAAGAGGAGGTGTTAAGAGAATCTCTTCACTCGTCTACGAAGAGACCAGA 180
    R L A R R G G V K R I S S L V Y E E T R
    GCTGTCCTCAAGGGATTCITA GAGA ACGTCATTAGAGACTCAGTAACATACTGAAACAC 240
    A V L K G F L E N V I R D S V T Y T E H
    GCCAAGAGAAAGACTGTACTGCCTAGACGT TGTITACGCCCTCAAGAGACAAGGAAAAG 300
    A K R K T V T A L D V V Y A L K R Q G K
    ACCCTACGGATTTCGGTGGTTAA 324
    T L Y G F G G

B  CCCCCAAACCCCAAAACCCCTGTTATAGCGGAAT -61

    GTGCTTGATAAATCAAAGAT TTCAGTAAATATAAATTATTAAAAAGTCGTTGAGTCATT -1

    ATGGTTGCAATGAAGAGAAAAGTCAAGAACCAAGAAGAGAGAGACAAATCAAAGGGTAAG 60
    M V A M [K] R [K] S R T K [K] R G D [K] S K G [K]
    GCTTGGATGACAACAGATACAAAGTCTCAGGTATCTCTGATGGTGGCCAT TAGAAGATTA 120
    A L D D N S Y K V S G I S D G A I R R L
    GCAAGAAGAGGAGGCATCAAGAGAATTGCAGAGGGAGTTTACTCAGAAATCAGAACTATT 180
    A R R G G I K R I G E G V Y S E I R T I
    TTACCCGCTTTGTGGCAAACTTGACCATGACAGCTACAACACTACTGTGAGTGAGCCAAG 240
    F T R F V D K L A T D S Y N Y C E C A K
    AGAAAGACTATTCCTCAAT TGATG TGATCT ATGCTTTGAAGAGAC AAGGAAGAAACCTC 300
    R K T I L P I D V I Y A L K R Q G R N L
    TATGGCTATGTCATGACAACGATGTCACGGACATTTAGGAGTCACTCATGATGAAGAA 360
    Y G Y V Y D N D V T G H L G V T H D E E
    GAAGGCAGCAAGAAAAGAAGTAG 384
    E G S K K K K
    ACTGTTTCAAGATT TTCAACATCACTTTCAAT TGAATTAAGTTAAGAGTTTCTTGA 444

    TACTCTTAAACAGCCITATTCTTTT TGCITTCITCCACTGATTGTCATTTTCAAT 504

    TAGCAGCTAATTAAGGGGTTTTGGGGTTTTGGGG 540
  
```

图 3 八肋游仆虫组蛋白 *H4A* 和 *H4B* 的核苷酸序列及拟编码的氨基酸序列

Fig. 3 Nucleotide sequence and deduced amino acid sequence of histone *H4A* and *H4B* from *Euplotes octocarinatus*

A: *H4A* 基因的开放阅读框及拟编码的氨基酸序列; B: *H4B* 基因序列及拟编码的氨基酸序列。单下划线所示为编码半胱氨酸的 TGA 密码子, 阴影所示为端粒序列。方框所示为保守的赖氨酸。

A: Nucleotide sequence and deduced amino acid sequence of histone *H4A*; B: Nucleotide sequence and deduced amino acid sequence of histone *H4B*. TGA codons that code for cysteines are underlined. The terminal C4A4 repeats are shaded. N-terminal conservative lysine of histone H4 are boxed.

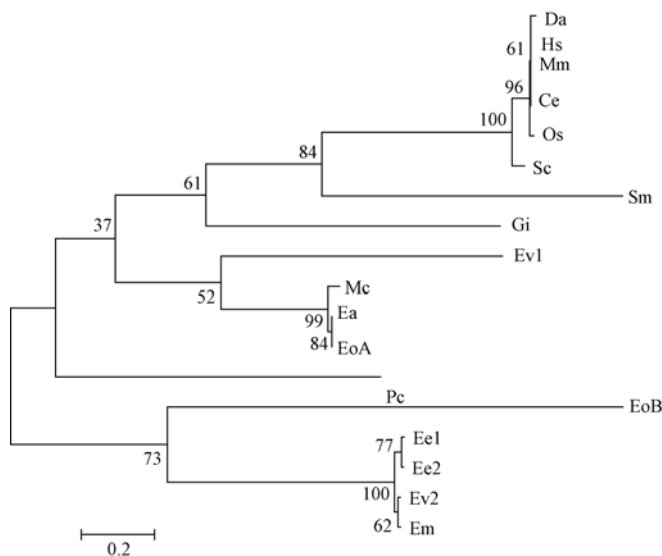


图 4 组蛋白 H4 的聚类分析

Fig. 4 Cluster analysis of histone H4

不同物种用前缀以示区别;Da. 果蝇; Hs. 人类; Mm. 小家鼠; Ce. 线虫; Os. 水稻; Sc. 酿酒酵母; Sm. 贻贝棘尾虫; Gi. 贾第虫; Ev1、Ev2. 扇形游仆虫的 2 种不同的组蛋白 H4; Mc. 厚体游仆虫; Ea. 小腔游仆虫; EoA、EoB. 八肋游仆虫的 2 种不同的组蛋白 H4; Pc. 草履虫; Ee1、Ee2. 阔口游仆虫的 2 种不同的组蛋白 H4; Em. 小游仆虫。标尺表示进化距离为每个位点发生 0.2 个碱基替换。分支处数字表示 bootstrap 支持率。

Da: *Drosophila americana* (BAF56194.1); Hs: *Homo sapiens* (NP_003539.1); Mm: *Mus musculus* (NP_291074.1); Ce: *Caenorhabditis elegans* (CAA33643.1); Os: *Oryza sativa Japonica Group* (AAM93740.1); Sc: *Saccharomyces cerevisiae* (EDN62788.1); Sm: *Stylonychia mytilus* (AAM77593.1); Gi: *Giardia intestinalis* (EET01662.1); Ev1: *Euplotes vannus* (CAH04403.1); Ev2: *E. vannus* (CAA06067.1); Mc: *Moneuplotes crassus* (AAB39722.1); Ea: *E. aediculatus* (AAM77588.1); EoA: *E. octocarinatus H4A*; EoB: *E. octocarinatus H4B*; Pc: *Paramecium caudatum* (BAB64430.1); Ee1: *E. eurystomus* (CAA06072.1); Ee2: *E. eurystomus* (CAA06071.1); Em: *E. minuta* (CAA06068.1). Scale bar indicates a genetic distance of 0.2 nt substitutions per site. Numbers in the nodes of cladogram indicate bootstrap values.

2.4 H4A 与 H4B 在八肋游仆虫中的表达差异

为了在转录水平比较 H4A 与 H4B, 首先验证实时定量 PCR 引物的实用性(无引物二聚体、无表 1 八肋游仆虫 H4A 与 H4B 的实时荧光 qRT-PCR 数据

Table 1 The real time qRT-PCR data of H4A and H4B from *Euplotes octocarinatus*

	目的基因 Target gene		
	β -微管蛋白 β -tubulin	H4A	H4B
循环次数 Threshold cycle	17.204	15.707	17.799
扩增效率 Efficiency	0.980	0.920	0.960
斜率 Slope	-3.206	-3.563	-3.248
方差 Δ^2	0.993	0.992	0.986

非特异性扩增), 建立实时定量标准曲线。H4A 和 H4B 的相对表达分析数据见表 1。H4A 与 H4B 在八肋游仆虫中的表达情况存在差异(图 5)。

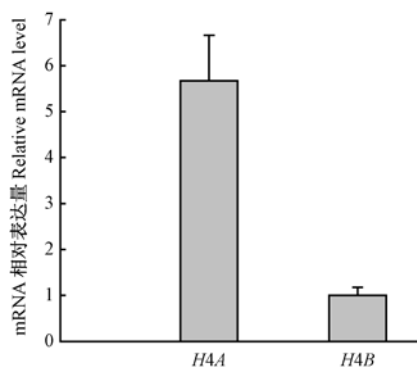


图 5 八肋游仆虫 H4A 与 H4B 的相对表达量
Fig. 5 The relative mRNA expression analysis of H4A and H4B from *Euplotes octocarinatus*

3 讨 论

本研究从八肋游仆虫大核基因组中克隆到 2 种新的编码组蛋白 *H4* 的基因,命名为 *H4A* 和 *H4B*。序列分析发现在 *H4A* 和 *H4B* 中均不存在内含子,但是 2 个基因使用不同的终止密码子,分别使用 TAA 和 TAG 终止翻译。*H4A* 和 *H4B* 编码区的 GC 含量分别为 24.9% 和 27.7%,与游仆虫基因组的 GC 含量相当。*H4B* 基因的上游和下游调控序列全部获得,与端粒部分一起组成微小染色体结构。但是 *H4A* 基因仅获得编码区序列,目前仍未能获得调控序列。将这 2 个基因拟编码蛋白的氨基酸序列进行 BLAST 搜索比对,发现 *H4A* 与人 (*Homo sapiens*)、小鼠 (*Mus musculus*)、斑马鱼 (*Danio rerio*)、贾第虫 (*Giardia intestinalis*)、贻贝棘尾虫 (*Stylonychia mytilus*)、小腔游仆虫等生物中组蛋白 *H4* 的氨基酸序列一致性相对较高,高达 81%~94%,而 *H4B* 与它们的一致性较低,仅为 36%~70%。这 2 种组蛋白 *H4* 之间的一致性为 44.7%。由于 *H4B* 与这些游仆虫的组蛋白 *H4* 之间的相似性较差,因此我们认为,该基因是八肋游仆虫进化出的特殊的组蛋白 *H4* 变异体。目前所知,组蛋白 *H4* 变异体仅在纤毛虫中有报道^[5]。2 种不同的组蛋白 *H4* 序列首次是在四膜虫中发现的^[16],在嗜热四膜虫 (*Tetrahymena thermophila*) 中 2 种 *H4* 基因只有 14 个核苷酸的差异,但编码氨基酸序列一致^[17]。同样,在浮萍棘尾虫 (*Stylonychia lemnae*) 中也有 2 种不同序列的 *H4* 被发现,它们有 23 个核苷酸的差异,但编码氨基酸序列一致^[18]。然而游仆虫中 2 种不同组蛋白的差异不仅是核苷酸序列上的不同,还发生在蛋白水平上。目前已经发现阔口游仆虫有 2 种不同的组蛋白 *H4* 序列,它们有 2 个核苷酸的差异且编码序列有 1 个氨基酸的差异。扇形游仆虫也有 2 种不同的组蛋白 *H4* 序列,且其氨基酸序列的一致性为 50%,这与本文报道的 *H4A* 和 *H4B* 44.7% 一致性是接近的。表明游仆虫中组蛋白 *H4* 的序列在进化过程中出现了更为特

殊的变异体。

为什么游仆虫组蛋白 *H4* 不像其他真核生物那么保守。一种可能的解释是游仆虫组蛋白 *H4* 的差异性与细胞中存在 2 个具有明显功能差异的基因组——转录沉默小核和转录活跃的大核有关^[19-20]。不同的组蛋白可能分别参与大核和小核基因组的组装。如在四膜虫中组蛋白 *H3* 的 3 种变异体分别定位在不同的核: major *H3s* 在大核、小核都有定位; minor *H3.3s* 只定位在大核; Cen $H3s$ 只定位在小核^[21-22]。如果纤毛虫核的二态性促使其组蛋白 *H4* 进化速度变快,那么我们期望氨基酸替换率的提高也应该在纤毛虫的其他蛋白中发现。事实上,蛋白的快速进化已经在数种纤毛虫的基因中被报道,包括延伸因子 1 α , 热休克蛋白 70, 肌动蛋白以及真核释放因子^[23-25], 且趋异进化的 α -微管蛋白的旁系同源物也在大核基因组中发现^[26]。多数情况下蛋白进化速率的提高伴随着趋异进化的旁系同源物的出现,表明通过基因复制产生的适应进化可能与纤毛虫基因组的二态性相互影响,从而解释了为何纤毛虫的组蛋白 *H4* 变异较大。在整个游仆虫中,具有两类不同的组蛋白 *H4* 基因,而在其他类群中不具有这种现象,可能的推测是这 2 种组蛋白 *H4* 变异体可能与游仆虫特殊的大核结构有关。是否具有显著差异的 2 种组蛋白 *H4* 在不同的游仆虫中普遍存在,目前仍不清楚。不同种类的游仆虫细胞的组蛋白序列克隆或全基因组序列的测序分析能够提供新的数据。

参 考 文 献

- [1] Sarma K, Reinberg D. Histone variants meet their match. *Nat Rev Mol Cell Biol*, 2005, 6(2): 139-149.
- [2] Jin J J, Cai Y, Li B, et al. In and out: histone variant exchange in chromatin. *Trends Biochem Sci*, 2005, 30(12): 680-687.
- [3] Piontkivska H, Rooney A P, Nei M. Purifying selection and birth-and-death evolution in the histone *H4* gene family. *Mol Biol Evol*, 2002, 19(5): 689-697.
- [4] Bernhard D, Schlegel M. Evolution of histone *H4* and *H3* genes in different ciliate lineages. *J Mol Evol*, 1998, 46(3): 344-354.

- [5] Katz L A, Bornstein J G, Lasek-Nesselquist E, et al. Dramatic diversity of ciliate histone H4 genes revealed by comparisons of patterns of substitutions and paralog divergences among eukaryotes. *Mol Biol Evol*, 2004, 21 (3): 555 – 562.
- [6] Jahn C L, Klobutcher L A. Genome remodeling in ciliated protozoa. *Annu Rev Microbiol*, 2002, 56: 489 – 520.
- [7] Duharcourt S, Yao M C. Role of histone deacetylation in developmentally programmed DNA rearrangements in *Tetrahymena thermophila*. *Eukaryotic Cell*, 2002, 1(2): 293 – 303.
- [8] 李江姣, 聂宇, 许静, 等. 八肋游仆虫 *Rab* 家族基因克隆和多样性分析. *动物学杂志*, 2010, 45(6): 1 – 8.
- [9] Jahn C L, Ling Z, Tebeau C M, et al. An unusual histone H3 specific for early macronuclear development in *Euplotes crassus*. *Proc Natl Acad Sci USA*, 1997, 94(4): 1332 – 1337.
- [10] Schulze D H, Freiburg M, Heckmann K. The isolation of gametes 3 and 4 of *Euplotes octocarinatus*. *Eur J Biochem*, 1987, 168(1): 89 – 94.
- [11] Abend M, Frombeck S, Baaske C, et al. Isolation of micronuclei: high yield by sucrose gradient versus maximum purity by cell sorting. *Mutat Res*, 1996, 360 (1): 23 – 28.
- [12] Tan M, Liang A H, Brünen-Nieweler C, et al. Programmed translational frameshifting is likely required for expressions of genes encoding putative nuclear protein kinases of the ciliate *Euplotes octocarinatus*. *J Eukaryot Microbiol*, 2001, 48(5): 575 – 582.
- [13] Voss T. Extraction and purification of *Plasmodium* total RNA. *Methods Mol Med*, 2002, 72: 151 – 157.
- [14] Pfaffl M W, Horgan G W, Dempfle L. Relative expression software tool (REST[©]) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res*, 2002, 30(9): e36.
- [15] Meyer F, Schmidt H J, Plümper E, et al. UGA is translated as cysteine in pheromone 3 of *Euplotes octocarinatus*. *Proc Natl Acad Sci USA*, 1991, 88(9): 3758 – 3761.
- [16] Hayashi H, Nomoto M, Iwai K. *Tetrahymena* histone H4. Complete amino acid sequences of two variants. *J Biochem*, 1984, 96(5): 1449 – 1456.
- [17] Horowitz S, Bowen J K, Bannon G A, et al. Unusual features of transcribed and translated regions of the histone H4 gene family of *Tetrahymena thermophila*. *Nucleic Acids Res*, 1987, 15(1): 141 – 160.
- [18] Wefes I, Lipps H J. The two macronuclear histone H4 genes of the hypotrichous ciliate *Stylonychia lemnae*. *DNA Seq*, 1990, 1(1): 25 – 32.
- [19] 许静, 王伟, 柴宝峰, 等. 八肋游仆虫一种富含三核苷酸重复序列的新基因 GARP 的克隆与序列分析. *遗传*, 2007, 29(1): 87 – 91.
- [20] Prescott D M. The DNA of ciliated protozoa. *Microbiol Mol Biol Rev*, 1994, 58(2): 223 – 267.
- [21] Cui B W, Liu Y F, Gorovsky M A. Deposition and function of histone H3 variants in *Tetrahymena thermophila*. *Molecular and Cellular Biology*, 2006, 26 (20): 7719 – 7730.
- [22] Cui B W, Gorovsky M A. Centromeric histone H3 is essential for vegetative cell division and for DNA elimination during conjugation in *Tetrahymena thermophila*. *Molecular and Cellular Biology*, 2006, 26 (12): 4499 – 4510.
- [23] Budin K, Philippe H. New insights into the phylogeny of eukaryotes based on ciliate Hsp70 sequences. *Mol Biol Evol*, 1998, 15(8): 943 – 956.
- [24] Moreira D, Le Guyader H, Philippe H. Unusually high evolutionary rate of the elongation factor 1 alpha genes from the Ciliophora and its impact on the phylogeny of eukaryotes. *Mol Biol Evol*, 1999, 16(2): 234 – 245.
- [25] Moreira D, Kervestin S, Jean-Jean O, et al. Evolution of eukaryotic translation elongation and termination factors: variations of evolutionary rate and genetic code deviations. *Mol Biol Evol*, 2002, 19(2): 189 – 200.
- [26] Israel R L, Kosakovsky P S L, Muse S V, et al. Evolution of duplicated alpha-tubulin genes in ciliates. *Evolution*, 2002, 56(6): 1110 – 1122.