

# 颈链血苔虫 (*Watersipora subtorquata*) 线粒体基因组的测定及其系统发育学意义

孙名安<sup>1,2</sup> 吴志刚<sup>1,2</sup> 申欣<sup>1,2</sup> 任建峰<sup>1,2</sup> 刘锡兴<sup>1,2</sup> 刘斌<sup>1,2</sup> 刘会莲<sup>1,2</sup>

(1. 中国科学院海洋研究所, 山东 青岛 266071; 2. 中国科学院研究生院, 北京 100039)

**摘要:** 苔藓动物为一类底栖、滤食性、营附着生活的小型水生群体动物, 关于其进化地位长期以来一直存在许多争议。本研究测定了颈链血苔虫的线粒体基因组。它是一个双链闭合环状分子, 全长 14,144 bp, 与其它后生动物相比相对较小。颈链血苔虫线粒体基因组存在 *atp8* 基因缺失现象, 所有的基因都编码在同一条链上。相比其它后生动物它具有独特的基因排列顺序, 而且与已发表的两个苔藓动物线粒体基因组相比基因排列顺序也显著不同。这说明苔藓动物线粒体基因组经历了大规模的基因重排过程。基因排列顺序比较分析的结果支持苔藓动物为原口动物的观点, 提示我们基因排列顺序的比较分析可能会成为苔藓动物门进化地位确定的良好途径。本研究为苔藓动物的分子系统发生研究提供了宝贵的分子数据, 对于其进化地位的确定有积极的推动作用。

**关键词:** 苔藓动物; 颈链血苔虫; 线粒体全基因组; 系统发生; 基因排列顺序

**中图分类号:** Q951+3

The complete mitochondrial genome of *Watersipora subtorquata* and its phylogenetic significance

SUN Ming-an<sup>1,2</sup> WU Zhi-gang<sup>1,2</sup> SHEN Xin<sup>1,2</sup> REN Jian-feng<sup>1,2</sup> LIU Xi-xing<sup>1,2</sup> LIU Bin<sup>1,2</sup>

LIU Hui-lian<sup>1,2</sup>

(<sup>1</sup> Key Laboratory of Experimental Marine Biology, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071 )

(<sup>2</sup> Graduate School of Chinese Academy of Sciences, Beijing 100039 )

**ABSTRACT** Bryozoans are small sessile aquatic colonial animals, and their phylogenetic position has

been controversial for long time. In this paper, we have determined the complete mitochondrial genome of *Watersipora subtorquata*. It is a circular molecule of 14,145 bp, relatively small compared with most other published metazoan mitochondrial genomes. One of the typical 37 genes (*atp8*) is missing from this genome. It is unusual that all genes are transcribed from the same strand. The gene order is unique compared with mtDNA of other metazoans, and drastic rearrangement was also found when compared with mtDNA of the other two bryozoans. It indicates that drastic rearrangement has happened in the mitochondrial genome of bryozoans. Comparison of gene order supports the protostome affinity of Bryozoa, and indicates that comparison of gene order may be especially useful for the establishment of the phylogenetic position of Bryozoa. This study provides important bryozoan molecular data which will facilitate the phylogenetic research of Bryozoa.

**KEY WORDS** Bryozoa *Watersipora subtorquata* Complete mitochondrial genome  
Phylogenetic Gene arrangement

苔藓动物（也称为外肛动物）为一类底栖、滤食性、营附着生活的小型水生群体动物。苔藓动物作为重要的海洋无脊椎动物，在一些底栖群落中常占据优势地位，同时苔藓动物在生物活性物质分离、抗肿瘤药物开发、生物污损以及成矿作用（刘锡兴等 2003；刘会莲 刘锡兴 2001；Blunt *et al.* 2007；Jones *et al.* 1990；Smith *et al.* 2006）等方面，和人类生产、生活密切相关。然而目前对于苔藓动物门的分子系统发生研究还相对较少，本门进化上的许多争议还没有得到解决（Halanych 2004）。到目前苔藓动物门在后生动物中的进化地位还没有得到确定，它与腕足动物门、帚虫动物门以及内肛动物门的进化关系也悬而未决，对于苔藓动物门内部的系统发生关系也尚未达成共识。更多的分子数据，尤其是基因组数据的获得，将有助于苔藓动物门进化地位的确定。

后生动物的基因组包括核基因组与线粒体基因组。动物的线粒体基因组一般都较小，通常在 15

---

基金项目：中国科学院百人计划（L97052417）及国家自然科学基金（30570212）资助。

作者简介：孙名安（1984-），男，硕士研究生，主要从事海洋生物基因组学研究，电话：(010)80481799，E-mail：minganthun@gmail.com；刘斌，通讯作者，电话：(010)80481799，E-mail: binliu\_cas@yahoo.com

kb~20kb 之间,除少数种类外都包括 37 个基因,即 2 个核糖体 RNA 基因,13 个蛋白质编码基因与 22 个转运 RNA 基因 (Boore 1999)。线粒体基因组相比单个基因信息量更为丰富,与核基因组相比具有大小相对较小、母系遗传、结构紧凑、基因相对保守等优点,同时它还具有基因排列顺序等基因组层次的特征。正因为如此,线粒体基因组的比较分析已经成为系统发生研究的常用手段 (Boore 1999; Boore and Brown 1998; Boore *et al.* 2005)。

本研究使用基于长 PCR 的基因组测序方法测定了颈链血苔虫 (唇口目,血苔虫科) 的线粒体基因组,并对其基本特征进行分析。通过基因排列顺序的比较分析,对苔藓动物门的进化地位进行了初步的探讨。

## 1 材料与方法

### 1.1 样品

*W. subtorquata* 采集自青岛,用蒸馏水清洗三次后 95%乙醇固定,4℃ 冰箱保存备用。

### 1.2 方法

#### 1.2.1 基因组 DNA 的提取和纯化

用刀片小心刮取颈链血苔虫样品于 Eppendorf 管中,去离子水浸泡半小时。剪碎后,加入 700 $\mu$ l 裂解液(40mM Tris-HCl,20 mM 醋酸钠,1 mM EDTA,1% SDS, pH 8.0)和 5 $\mu$ l 蛋白酶 K(20mg/mL),55℃ 消化约 2 小时。4000rpm 离心 5min 去除钙质沉淀,小心将上清液转入新的 Eppendorf 管中。等体积的酚:氯仿:异戊醇(25:24:1)抽提两次,等体积的氯仿:异戊醇(1:1)抽提一次。将上清液转入新的 Eppendorf 管,加入 70%体积比的异丙醇,充分混匀后 -20℃ 冰箱中放置 30min。12,000rpm 离心 15min 沉淀 DNA,70%乙醇洗涤 2 次后室温干燥。用 50 $\mu$ l TE (含 20 $\mu$ g/mL 的 RNase) 溶解 DNA,1%琼脂糖凝胶电泳检测质量。基因组 DNA 用 DNA 纯化试剂盒(天根生化科技有限公司)纯化后 -20℃ 保存备用。

#### 1.2.2 线粒体基因组的 PCR 扩增

先根据后生动物线粒体基因组通用引物 (表 1) 扩增 *cox1*, *cox2*, *16s* 基因的部分序列。PCR 为 25 $\mu$ l 反应体系, 包括 18.0  $\mu$ l 的灭菌去离子水, 2.5  $\mu$ l 10 $\times$ LA PCR bufferII (Mg<sup>2+</sup> plus, Takara), 0.5  $\mu$ l dNTP (各 10 mM), 正向引物与反向引物各 1 $\mu$ l, 1 $\mu$ l LA-Taq 聚合酶, 1 $\mu$ l 模板。反应程序为: 95 预变性 2 分钟; 94 变性 20 秒, 50 退火 50 秒, 72 延伸 2 分钟共 38 个循环; 最后 72 延伸 5 分钟。PCR 产物用 1%琼脂糖凝胶电泳进行检测, 纯化后用 ABI 3730 型测序仪测序。

根据已获得的线粒体基因组部分序列设计长 PCR 引物 (表 1), 将引物相互搭配用于整个线粒体基因组的扩增。PCR 为 25 $\mu$ l 反应体系, 包括 18.0  $\mu$ l 的灭菌去离子水, 2.5  $\mu$ l 10 $\times$ LA PCR bufferII (Mg<sup>2+</sup> plus, Takara), 0.5  $\mu$ l dNTP (各 10 mM), 正向引物与反向引物各 1 $\mu$ l, 1 $\mu$ l LA-Taq 聚合酶, 1 $\mu$ l 模板。反应程序为: 95 预变性 2 分钟; 94 变性 20 秒, 50 退火 50 秒, 65 延伸 10 分钟共 36 个循环; 最后 72 延伸 10 分钟。用 1%琼脂糖凝胶电泳检测 PCR 产物质量。最终获得了能够覆盖整个线粒体基因组的一组 PCR 产物, 其长度分别为: WScox1LF-WS16sLR 约 4 kb, WScox2LF-WScox1LR 约 4 kb, WS16sLF-WScox2LR 约 6 kb。

表 1 *W. subtorquata* 线粒体基因组测定中用到引物及其序列

Table 1 Primers used in amplifying and sequencing of *W. subtorquata* mitochondrial genome

	引物名称 Name	引物序列 Sequences	引物来源 References
通用引物 Universal primer	LCO	GGTCAACAAATCATAAAGATATTGG	Folmer <i>et al.</i> 1994
	HCO	TAAACTTCAGGGTGACCAAAAAATCA	
	cox2F	AAGCWAATWGGNCATCARTGRTATTG	Burger <i>et al.</i> 2007
	cox2R	CTCCRCATATTCNGARCATTGNCC	
	16SarF	CGCCTGTTTATCAAAAACAT	Palumbi 1996
	16SbrR	CCGGTCTGAACTCAGATCACGT	
长 PCR 引物 Long PCR primer	WScox1LF	TCCCATTATTGTTTGAGCCGTA	本研究中设计 This study
	WScox1LR	AGTTCATCCAGTCCCTGCCCTC	
	WScox2LF	TGCTGTTCTGGACGACTAAATC	
	WScox2LR	GATTTAGTCGTCAGGAACAGCA	
	WScox3LF	ATCAAGAACACACAAAGACACCC	
	WScox3LR	CTCCTACAGGTGGTCAAGTCGCC	
	WS16sLF	CTAATTGAAGGTAAGGATTGCGAC	
	WS16sLR	AGATTACGCTGTATCCCTAAGG	

### 1.2.3 线粒体全基因组序列的获得

将两个 4kb 的 PCR 产物用 DNA 纯化试剂盒纯化后直接测序，使用引物步移的方法得到它们的全长。而对于 6kb 的 PCR 产物，先用超声破碎仪随机打碎，用胶回收试剂盒（Qiagen 公司）回收 1~3 kb 的片段，用 DNA 聚合酶进行末端补平后插入到 pUC19 平端载体中。提取重组质粒后，使用 M13 通用引物测序。

使用 phredPhrap( Ewing and Green 1998a ;Ewing *et al.* 1998b )软件对测序的结果进行组装拼接，然后使用 CONSED 软件 ( Gordon *et al.* 1998 ) 进行人工校对。根据拼接的结果设计引物扩增并测定剩下的缺口，并补测低质量区，直到得到线粒体全基因组。

#### 1.2.4 基因注释

通过在线基因注释工具 DOGMA ( Wyman *et al.* 2004 )对蛋白质编码基因及两个核糖体 RNA 基因进行注释。Genome type 选择“Mitochondrial”，Percent identity cutoff for protein coding genes 设为 30，Percent identity cutoff for RNAs 设为 30，其它参数采用默认值。随后通过与其它后生动物对应的同源基因进行 BLAST 比对来确定每个基因的精确位置。

大多数转运 RNA 基因通过在线 *tRNA* 注释工具 tRNAscan-SE 1.21 ( Lowe and Eddy 1997 ) 来注释得到。Genetic Code 设为“invertebrate mitochondrial”，Source 设为“mito/chloroplast”，Cove score cutoff 设为 2，其它都取默认值。剩余的转运 RNA 基因通过人工寻找 *tRNA* 二级结构及反密码子来确定。

## 2 结果与分析

### 2.1 碱基组成分析

*W. subtorquata* 的线粒体基因组在 NCBI 的序列号为 EU365892。它为闭合双链环状分子，全长 14,144 bp，与已发表的其它后生动物线粒体基因组比较相对较小。在 *W. subtorquata* 的线粒体基因组中，其发现了 12 处基因间的重叠区，大小在 1 到 4 个碱基之间，共有 21 个碱基；发现了 16 个非编码区，共有 164 个碱基，其中 *cox1* 与 *tRNA<sup>Leu(CUN)</sup>* 间的非编码区具 100 个碱基且 AT 含量较高 (73%)，

推测为D-loop区。α-链的碱基组成为：A = 5,146 (36.4%), T = 4,842 (34.2%), C = 2,303 (16.3%), G = 1,853 (13.1%) , AT含量为 70.6%。

表 2 *W. subtorquata* 线粒体基因组基因组成

Table 2 Mitochondrial genome profiles of *W. subtorquata*

基因 Gene	编码链 Strand	位置 Position	大小 Size (bp)	起始密码子 Start codon	终止密码子 Stop codon	基因间区碱基 Intergenic nucleotides <sup>a</sup>
<i>cox1</i>	+	1-1548	1548	ATC	TAA	
nCR	+	1549-1648	100			0
<i>tRNA<sup>Leu(CUN)</sup></i>	+	1652-1708	57			3
<i>cox3</i>	+	1717-2502	786	ATA	TAG	8
<i>tRNA<sup>Ile</sup></i>	+	2501-2560	60			-2
<i>nad1</i>	+	2561-3483	923	ATG	TA <sup>b</sup>	0
<i>tRNA<sup>Gly</sup></i>	+	3483-3544	62			-1
<i>lrRNA</i>	+	3545-4676	1132			0
<i>tRNA<sup>Ser(AGN)</sup></i>	+	4679-4734	56			2
<i>tRNA<sup>Trp</sup></i>	+	4735-4801	67			0
<i>tRNA<sup>Asp</sup></i>	+	4801-4863	63			-1
<i>nad5</i>	+	4867-6546	1680	ATG	TAG	3
<i>tRNA<sup>Ala</sup></i>	+	6544-6598	55			-3
<i>tRNA<sup>Lys</sup></i>	+	6597-6659	63			-2
<i>tRNA<sup>Asn</sup></i>	+	6664-6730	67			4
<i>nad2</i>	+	6731-7693	963	ATG	TAA	0
<i>tRNA<sup>His</sup></i>	+	7694-7754	61			0
<i>cob</i>	+	7755-8849	1095	ATG	TAA	0
<i>tRNA<sup>Arg</sup></i>	+	8848-8914	67			-2
<i>srRNA</i>	+	8915-9684	770			0
<i>tRNA<sup>Ser(UCN)</sup></i>	+	9696-9748	53			11
<i>tRNA<sup>Leu(UUR)</sup></i>	+	9748-9803	56			-1
<i>cox2</i>	+	9803-10478	676	ATG	T <sup>b</sup>	-1
<i>tRNA<sup>Pro</sup></i>	+	10479-10541	63			0
<i>tRNA<sup>Glu</sup></i>	+	10541-10602	62			-1
<i>tRNA<sup>Phe</sup></i>	+	10602-10666	65			-1
<i>nad3</i>	+	10667-11018	352	ATG	T <sup>b</sup>	0
<i>tRNA<sup>Tyr</sup></i>	+	11019-11069	51			0
<i>tRNA<sup>Thr</sup></i>	+	11070-11133	64			0
<i>tRNA<sup>Cys</sup></i>	+	11134-11193	60			0
<i>nad4L</i>	+	11197-11490	294	ATG	TAA	3
<i>nad4</i>	+	11511-12824	1314	ATA	TAA	20
<i>atp6</i>	+	12830-13508	679	ATG	T <sup>b</sup>	5
<i>tRNA<sup>Gln</sup></i>	+	13505-13564	60			-4
<i>nad6</i>	+	13567-14028	462	ATA	TAA	2
<i>tRNA<sup>Met</sup></i>	+	14027-14089	63			-2
<i>tRNA<sup>Val</sup></i>	+	14093-14144	52			3

a 基因间间隔区的碱基数，负数代表相邻基因间的碱基重复数目

## b 通过多腺苷酸化来完成的终止子

### 1.3.2 蛋白质编码基因

后生动物的线粒体基因组通常都具有 13 个蛋白质编码基因。然而 *W. subtorquata* 的线粒体基因组缺失了 *atp8* 基因，只有 12 个蛋白质编码基因。*atp8* 基因从线粒体基因组中缺失的现象在其它后生动物中也曾发现过。所有的蛋白质编码基因都编码在  $\alpha$ -链上。其中有 8 个基因 (*cox2*, *atp6*, *cob*, *nad1*, *nad2*, *nad3*, *nad4L* 与 *nad5*) 使用了 ATG 起始密码子，*cox3*, *nad4* 与 *nad6* 使用了 ATA，*cox1* 使用了 ATC。8 个基因使用了完全终止密码子 (*cox1* 与 *nad5* 使用了 TAG; *cox3*, *cob*, *nad2*, *nad4*, *nad4L* 与 *nad6* 使用了 TAA)，4 个基因使用了不完全终止密码子 (*nad1* 使用了 TA，*cox2*, *atp6* 与 *nad3* 使用了 T)。

### 1.3.3 核糖体 RNA 与转运 RNA 基因

*W. subtorquata* 的两个核糖体 RNA 都编码在  $\alpha$ -链上。*lrRNA* 位于 *tRNA<sup>Gly</sup>* 与 *tRNA<sup>Ser(AGN)</sup>* 之间，具有 1132 个碱基，*srRNA* 位于 *tRNA<sup>Arg</sup>* 与 *tRNA<sup>Ser(UCN)</sup>* 之间，具有 770 个碱基。*W. subtorquata* 的线粒体基因组具有后生动物典型的 22 个 *tRNA* 基因，除 *tRNA<sup>Tyr</sup>* 的 DHU 臂缺失外，其它 21 个转运 RNA 都呈三叶草结构。

### 1.3.4 基因排列顺序

线粒体基因组基因排列顺序的比较分析用于推测物种间进化关系有很多优势。几乎所有的动物都具有同样的一系列基因，这些基因通过基因重排可以形成极多的状态，所以回复或趋同现象不太可能发生。虽然线粒体基因组进化速率较快，然而除极少数类群外，基因排列顺序大多都比较稳定，例如人和鲨鱼就具有同样的基因排列顺序。这使得通过基因排列顺序的比较分析来推测深层进化关系尤其有用 (Boore 1999; Boore *et al.* 2005)。目前线粒体基因组排列顺序比较分析已经成为研究后生动物系统发生的常用方法 (Boore *et al.* 1995; Boore *et al.* 2005)。原先已经发表的两条苔藓动物，即 *Flustrellidra hispida* (Waeschenbach *et al.* 2006) 与 *Bugula neritina* 的线粒体基因组都具有独特的



基因排列顺序。*W. subtorquata* 的线粒体基因组的基因排列顺序与其它后生动物也显著不同。与其它后生动物相比，基因排列顺序相同的最长的基因区段（不包括转运 RNA 基因）仅包括 4 个基因，即 *nad6-cox1-cox3-nad1*，存在于 *Heterodoxus macropus*（节肢动物门）中。长期以来对于苔藓动物是原口动物还是后口动物一直存在许多争议，基因组排列顺序的比较分析支持苔藓动物为原口动物的观点。另外，这三个苔藓动物的线粒体基因组的基因排列顺序也显著不同，可见苔藓动物在进化过程中经历了大规模的基因重排过程。

### 3 讨论

关于苔藓动物门的进化地位一直存在许多争议，目前关于苔藓动物的分子系统发生研究较少，分子数据相对缺乏。本研究测定了 *W. subtorquata* 的线粒体全基因组。相比其它后生动物的线粒体基因组，它具有一些显著特征：全长 14,144 bp，与其它后生动物相比相对较小；缺失了 *atp8* 基因；所有的基因都编码在同一条链上。*W. subtorquata* 的线粒体基因组相比其它后生动物它具有独特的基因排列顺序，而且与已发表的两个苔藓动物相比基因排列顺序也显著不同。这说明苔藓动物线粒体基因组经历了大规模的基因重排过程。基因排列顺序比较分析的结果支持苔藓动物为原口动物的观点。然而由于目前基因排列顺序比较分析的方法还不是很完善（Boore *et al.* 2005），另外苔藓动物门的线粒体基因组数据还较少，通过此方法尚不能确定苔藓动物门确切的进化地位。随着苔藓动物线粒体基因组数据的丰富，基因排列顺序的比较分析可能会成为苔藓动物门进化地位确定的良好途径。本研究为苔藓动物的分子系统发生研究提供了宝贵的分子数据，对于其进化地位的确定有积极的推动作用。

### 参考文献：

- 刘锡兴, 尹学明, 马江虎. 2001. 中国海洋污损苔虫生物学. 科学出版社  
刘会莲, 刘锡兴. 2003. 中国沿岸水域养殖贝类及其养殖笼网污损苔虫 7 新种. 海洋科学集刊, 45: 202~216  
Blunt, J.W., Copp, B.R., Hu, W.P. *et al.* 2007. Marine natural products. Nat Prod Rep, 24:31~86.  
Boore, J.L. 1999. Animal mitochondrial genomes. Nucleic Acids Res, 27: 1767~1780  
Boore, J.L., Collins, T.M., Stanton, D. *et al.* 1995. Deducing the pattern of arthropod phylogeny from mitochondrial DNA rearrangements. Nature, 376: 163~165  
Boore, J.L. and Brown, W.M. 1998. Big trees from little genomes: mitochondrial gene order as a phylogenetic tool. Curr Opin Genet Dev, 8: 668~674

- Boore, J.L., Macey, J.R. and Medina, M. 2005. Sequencing and comparing whole mitochondrial genomes of animals. *Methods Enzymol*, 395: 311~348
- Burger, G., Lavrov, D.V., Forget, L. and Lang, B.F. 2007. Sequencing complete mitochondrial and plastid genomes. *Nature Protocols*, 2: 603~614
- Ewing, B. and Green, P. 1998a. Basecalling of automated sequencer traces using phred. II. Error probabilities. *Genome Res*, 8: 186~194
- Ewing, B., Hillier, L., Wendl, M. and Green, P. 1998b. Basecalling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res*, 8: 175~185
- Folmer, O., Black, M., Hoeh, W., *et al.* 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Marine Biol Biotechnol*, 1994, 3: 294~299
- Gordon, D., Abajian, C. and Green, P. 1998. Consed: a graphical tool for sequence finishing. *Genome Res*, 8: 195~202
- Halanych, K.M. 2004. The new view of animal phylogeny. *Annu. Rev. Ecol. Evol. Syst.*, 35: 229~256
- Jones, R.J., Sharkis, S.J., Miller, C.B., *et al.* 1997. Bryostatin 1, a unique biologic response modifier: anti-leukemic activity in vitro. *Blood*, 75: 1319~1323
- Lowe, T.M. and Eddy, S.R. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res*, 25: 955~964
- Nickisch-Roseneck, M., Brown, W.M. and Boore, J.L. 2001. Sequence and structure of the mitochondrial genome of the tapeworm *Hymenolepis diminuta*: Gene arrangement indicates that plathyhelminths are derived eutrochozoans. *Mol. Biol. Evol.*, 18: 721~730
- Palumbi, S., Martin, A., Romano, S., *et al.* 1991. "The Simple Fool's Guide to PCR." Version 2.0, University of Hawaii, Department of Zoology and Kewalo Marine Laboratory
- Smith, A.M., Key, M.M. and Gordon, D.P. 2006. Skeletal mineralogy of bryozoans: Taxonomic and temporal patterns. *Earth-Science Reviews*, 78: 287~306
- Waeschenbach, A.M., Telford, J., Porter, J.S. and Littlewood, D.T.J. 2006. The complete mitochondrial genome of *Flustrellidra hispida* and the phylogenetic position of Bryozoa among the Metazoa. *Mol Phylogenet Evol*, 40: 195~207
- Wyman, S.K., Jansen, R.K. and Boore, J.L. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics*, 20: 3252~3255