

藻类叶绿体 DNA 和基因图谱

DNA and gene map of alga chloroplast genome

何培民

(上海水产大学渔业学院, 200090)

He Peimin

(Fisherise College, SFU, 200090)

张荣铄

(南京农业大学农学系, 210095)

Zhang Rongxian

(Agriculture Department, Nanjing Agricultural University, 210095)

关键词 藻类, 叶绿体, DNA 特性, 基因图谱

KEYWORDS alga, chloroplast, DNA, gene map

中图分类号 S917

叶绿体遗传学的研究最早可追溯到上世纪末, Strasburger 最早观察到藻类叶绿体能够分裂, 并在细胞分裂过程中完成了分裂的叶绿体分别进入子代细胞。同时其他研究者也认为, 叶绿体是已经存在的质体分裂形成的。随后, Baur 和 Correne 认为紫茉莉枝条三种不同颜色的花斑性状是由母本细胞质传递的, 与核基因没有直接关系, 从而产生了叶绿体自主性遗传概念^[1]。现已逐步知道植物细胞的核基因组和叶绿体基因组之间存在复杂相互作用关系, 叶绿体遗传应属于半自主系统 (Semiautonomous system)。在应用电镜首次证明衣藻和玉米叶绿体中具有 DNA 后, 有关叶绿体分子生物的研究迅速发展。仅仅十多年的时间, 叶绿体 DNA 的第 1 个物理图谱在玉米中建立^[2]。在藻类中, Rochaix^[3] 第一个完成了衣藻叶绿体物理图谱制作。随后, 藻类叶绿体分子生物学和转基因研究得到快速发展, 如利用叶绿体核编码的 18S rRNA 顺序同源性以确立藻类之间的亲源关系及建立系统发生树结构^[4]。迄今已发表了 50 多种植物叶绿体 DNA 物理图谱, 并已在高等植物叶绿体中定位了 120~160 个基因, 而在藻类定位了 255 个基因^[5], 可见藻类在叶绿体分子生物学研究中占有相当重要地位。本文就藻类叶绿体 DNA 特性及基因图谱等有关内容进行了综述。

1 藻类叶绿体 DNA

1.1 叶绿体 DNA 发现

Ris 和 Plant^[6] 应用电子显微镜首次证明了莱茵藻和玉米叶绿体中具有 DNA。他们发现莱茵藻, 玉米叶绿体基质中有 25Å 左右细纤维丝, 用 DNA 酶处理可使其消失。此后, 又从菠菜、甜菜、蚕豆、烟草、眼虫藻、莱茵藻和伞藻等叶绿体中分离到环状 DNA 分子。

1.2 叶绿体 DNA 的大小、形状

叶绿体 DNA 为双连环状分子。一般长约 40~45μm, 分子量为 90MDa, 高等植物每个叶绿体 DNA 含量为 0.8×10^{-14} ~ 6.9×10^{-14} g, 藻类为 0.01×10^{-14} ~ 1.2×10^{-14} g。维管植物叶绿体 DNA 大小一般为 120~160kb, 其中被子植物为 120kb 左右 (230 种植物), 裸子植物银杏 (*Jinkgo biloba*) 为 158kb, 蕨类植物为 140~150kb, 苔藓地钱 (*Marchantia nidua*) 为 120kb。而藻类 (主要为绿藻) 叶绿体 DNA 大小变化很大,

上海市教委曙光计划资助课题 (SG-98-032)。

第一作者简介: 何培民, 男, 1959 年 10 月生, 副研究员, 在职博士生。

收稿日期: 1998-11-10

如刺松藻 (*Codium fragile*) 最小为 86kb, 伞藻 (*Acetabularia*) 最大高达 2 000kb, 小球藻 (*Chlorella ellipsoidea*) 和衣藻 (*Chlamydomonas amoebii*) 则分别为 174kb 和 199kb。红藻为 175~190kb, 有色藻为 115~135kb。

叶绿体 DNA 分子以环状分子, 线型分子单体及其多聚体形式存在。高等植物叶绿体 DNA 的环状分子占 80%, 其中超螺旋结构占 15%~30%, 平均周长为 5 μ m; 而纤维裸藻有 34% 为环状, 平均周长为 40 μ m。但伞藻例外, 叶绿体 DNA 为线状, 不是环状分子。此外有 10% 环状二聚体 (如豌豆、菠菜、莴苣), 2.5% 的线状二聚体。200 多种被子植物叶绿体基因组均为单一类群的环状分子, 但间囊藻 (*Pr-laiella littoralis*) 存在 135kb 和 58kb 两种环状分子, *Sphacelaria* sp. 也为多环。叶绿体 DNA 常由两组分子组成, 仅其单拷贝区方向相反, 遗传内容相同。

1.3 叶绿体 DNA 其它特性

可用氯化铯梯度离心分离和鉴定植物的叶绿体 DNA。高等植物叶绿体 DNA 浮力密度为 1.696~1.698g/mL, 藻类为 1.685(眼虫藻)~1.706g/mL(伞藻); 高等植物叶绿体 DNA 的 dG + dC 含量在 36%~40%, 藻类为 24%(眼虫藻)28~47%(伞藻)。利用叶绿体 DNA 的浮力密度和核 DNA 的不同可以有效地将它们分离。

由于叶绿体 DNA 序列同源性高, 叶绿体 DNA 复性速率高于核 DNA。

所有研究过的叶绿体 DNA(包括藻类)都是裸露形式存在, 没有组蛋白存在^[7]。核 DNA 含有 10% 的 5-甲基胞嘧啶, 而在叶绿体 DNA 中未发现 5-甲基胞嘧啶, 这与兰藻、细菌 DNA 相似。该嘧啶存在与否可作为叶绿体 DNA 提纯指标。

每个叶绿体含有的 DNA 分子是多拷贝的, 一般每个叶绿体含有 10~50 个(份)拷贝, 最高可达 900 份拷贝。因此, 叶绿体中一个 DNA 拷贝发生突变, 会被其客观存在野生型拷贝所掩盖, 并要有相当长时间才可达到纯化状态。

用碱或 RNA 酶处理, 叶绿体 DNA 共价闭合环状分子会逐渐转变成开环结构^[8], 这是因为叶绿体 DNA 分子含有 12~18 个核苷酸, 与相邻的脱氧核糖核苷酸以共价键连接。叶绿体 DNA 中含有核糖核苷酸有何意义, 目前还不清楚。

叶绿体中有 DNA 复制机构。最初是在菠菜和衣藻中发现叶绿体内能合成 DNA。一般幼嫩细胞叶绿体的分裂速度和 DNA 复制速度均比细胞和核 DNA 复制要快。叶绿体 DNA 复制时, 10~20 个 DNA 聚成一族, 并且叶绿体 DNA 分子与包被的内膜相连。有时也与类囊体膜相连接, 以 DNA 膜复合体的形式进行分裂。叶绿体 DNA 复制过程与原核生物 DNA 复制模式相类似, 主要为滚环复制和堆叠复制 (caim) 两种。最近用免疫荧光显微技术证实某些藻类的叶绿体 DNA 是位于蛋白核 (pyrenoid) 中^[9, 10, 11]。

从理论上讲, 叶绿体 DNA 可编码 150~300 种多肽。利用离体叶绿体已证明叶绿体能合成全部叶绿体 rRNA, 几乎所有的 tRNA 和近百种蛋白质。

2 藻类叶绿体基因图谱

2.1 叶绿体基因图谱制作史

1976 年, Redbrook 和 Boyrad 建立了玉米叶绿体 DNA 第 1 个物理图谱。迄今大约发表了 50 多种植物叶绿体 DNA 物理图谱^[1]。

1978 年, Rochaix 首次在藻类建立了莱茵藻叶绿体 DNA 物理图谱^[3]; 1985 年, Lamber 等制作了红藻 (*Cyanochora paradoxa*) 叶绿体 DNA 物理图谱^[4]; Kuhse 和 Kowallik 制作了网地藻 (*Dictyota dichotoma*) 叶绿体 DNA 物理图谱^[12]; Bayen 等制作了角叉菜 (*Chondrus crispus*) 叶绿体 DNA 物理图谱^[13]; Loffelhaux de Gor [1988] 等建立了褐藻间囊藻 (*Pylaiella littoralis*) 叶绿体 DNA 物理和基因图谱^[14]; Shivji 制作了条斑紫菜 (*Porphyra yezoensis*) 叶绿体 DNA 基因图谱^[15]; 至今大约有 7 种红藻, 10 种色藻, 3 种绿藻等叶绿体制作了较有价值基因图谱, 并在进一步完善。目前已测定了红藻 (*Cyanophora paradoxa*)^[16]、紫红紫菜 (*Porphyra purpurea*)^[17, 18] 和纤细裸藻 (*Euglena gracilis*)^[19] 叶绿体 DNA 全部核苷酸序列, 并且分别定位了 150、255 和 108 个基因。以后, 又有红藻 (*Cyanidium caldarium*)、隐藻 (*Guillardia theta*)、有色藻 (*Odonzella sinensis*) 和

小球藻 (*Chlorella vulgaris*) 的叶绿体基因组测序被完成。其中 Reith 和 Munholland 建立的紫红紫菜 (*Porphyra purpurea*) 叶绿体基因图谱, 255 个基因被定位, 是目前植物中叶绿体基因定位最多的藻类^[5]。

在陆地植物中已测定了烟草 (*Nicotiana tabacum*)^[20], 水稻 (*Oryza sativa*)^[21] 和地钱 (*Marchantia polymorpha*)^[22] 叶绿体 DNA 全部核苷酸序列, 此后又完成了玉米 (*Zea mays*)、松树 (*Pinus thunbergii*)、月草 (*Oenothera elata*) 叶绿体基因组测序。表 1 显示了藻类和主要高等植物的质体基因的结构及内容。

表 1 质体基因组组织的基因内容

Tab.1 Plastid genome organization and gene content

种 名	大小 (kbp)	rRNA 重复序列数和方向 ^a	重复序列大小 (kbp)	已作图基因数	参考文献
红藻 (RHODOPHYTES)					
<i>Cyanophora paradoxa</i>	134	2i	12	150 ^c	[16]
<i>Goldieria sulphuraria</i>	?	2i	5	60	[17]
<i>Porphyra purpurea</i>	191	2d	4.8	255 ^c	[17, 18]
<i>Porphyra</i>	185	2i	6.6	25	[15, 23]
<i>Griffithsia pacifica</i>	178	1	-	24	[15]
<i>Chondrus crispus</i>	173	1	-	15	[13]
<i>Antithamnion</i> sp. ^b	180	1	-	55	[17]
有色藻 (CHROMOPHYTES)					
<i>Cryptomonas</i>	118	2i	6	70	[24]
<i>Pyrenomonas salina</i>	127	2i	5	20	[25]
<i>Casinianthus granii</i>	118	2i	10	35	[26]
<i>Cyclotella meneghiniana</i>	128	2i	17	25	[27]
<i>Odontella sinensis</i>	118	2i	8.7	96	[28]
<i>Ochromonas danica</i>	121	2i	7.5	19	[23]
<i>Heterosigma akashiwo</i>	154	2i	22	27	[23]
<i>Vaucheria hirsuta</i>	124	2i	7.7	34	[26]
<i>Dictyota dichotoma</i>	123	2i	4.7	30	[26]
<i>Pyruella litoralis</i>	133	2i	5	23	[14]
绿藻 (CHLOROPHYTES)					
<i>Chlamydomonas reinhardtii</i>	196	2i	20	75	[29]
<i>Chlamydomonas moewusii</i>	292	2i	41	74	[29]
<i>Euglena gracilis</i>	143	3.5d	5.9	108 ^c	[19]
陆生植物 (METAPHYTES)					
<i>Marchantia polymorpha</i>	121	2i	10	129 ^c	[22]
<i>Nicotiana tabacum</i>	156	2i	25	138 ^c	[20]
<i>Oryza sativa</i>	135	2i	21	133 ^c	[21]

注: ^a rRNA 重复的方向反向为 (d), ^b 其质体基因组还未完全作图, ^c 其质体基因组已完全测序。

2.2 叶绿体 DNA 物理图上的基因组织结构

一般把叶绿体基因组分为 4 个区: 反向重复区 A (inverted sequence region A, IRA), 反向重复区 B (inverted repeat sequence region B, IRB), 大单拷贝区 (Large single copy region, LSC) 小单拷贝区 (Small single copy region, SSC)。

大多数植物叶绿体的环状 DNA 的基因图谱中均有两个内容相同而方向相反的重复序列 IRA 和 IRB。但蚕豆、豌豆、一种苜蓿 (*Medicago truncatula*)、松树、纤细裸藻、*Griffithsia pacifica*、角叉菜 (*Chondrus crispus*)、对丝藻 (*Antithamnion* sp.) 叶绿体没有反向重复区, 而纤细裸藻含有 3 个同向串联方式排列的重复 DNA 片段。有趣的是, 在紫菜中, 紫红紫菜 (*P. purpurea*) 为同向重复, 而条斑紫菜 (*P. yezoensis*) 则为反向重复^[23, 17]。所有 rRNA 的基因均在 IR A 和 IR B 两个重复区内。陆生植物和所研究过的有色藻中, rRNA 操纵子 (16S-23S-4.5S-5S_{rrn}) 均以反向重复区的 5SrRNA 基因与小拷贝区相邻, 而红藻有 3 种邻接方式: 反向重复区以 5SrRNA 基因与小拷贝区相邻, 如 *Goldieria sulphuraria*; 反向重复区以 16SrRNA 基因与小拷贝区相邻, 如 *Cyanophora paradoxa* 和条斑紫菜 (*P. yezoensis*); 同向重复, 如紫红紫菜 (*P. purpurea*)。

IR A 和 IR B 两个重复区之间可分为大单拷贝区和小单拷贝区。

叶绿体 DNA 为双链,编码 RuBP 羧化酶(Rubisco)大亚基基因(*rbcL*)的链称为 A 链,其互补链则为 B 链。以烟草为例,A 链上含有 16 个 *trn* 基因;4 个 *rnm* 基因(即 16S, 23S, 4.5S 和 5S *rnm* 基因);19 个蛋白基因;12 个未知产物基因(*yef* 序列);2 个自主复制序列 *ars1* 和 *ars2*。B 链含有 20 个 *trn* 基因;4 个 *rnm* 基因;52 个蛋白基因;11 个 *yef* 序列。

将 IR A 和 LSC 之间的 BamHI 片段 8 上的 BamHI 切点定义为 0,核苷酸数目依次向 LSC 递增计数。

2.3 藻类叶绿体 DNA 上的基因内容

叶绿体 DNA 基因可分为三大类:①与叶绿体光合作用有关基因;②与叶绿体基因表达有关的基因;③其它有关生物合成的基因。

陆生植物地钱、烟草、水稻的叶绿体基因组已分别定位了 129、138、133 个基因,其中有关光合作用基因数和有关基因表达的基因数分别是 40(含 *ndh*)和 60 个左右。

红藻紫红紫菜(*P. purpurea*)叶绿体基因组现已定位了 259 个基因和阅读框架(DRFS),是目前所有植物叶绿体含有基因数最多的植物,大约是陆生植物的两倍。已确定了 194 个基因,其中有 150 个基因是陆地植物中没有发现的。大约还有 20% 的基因还没有确定。已知功能的基因包括 50 个光合蛋白基因,102 个为基因表达的有关基因。纤维裸藻叶绿体基因组定位了 108 个基因,其中有关光合作用基因数和有关基因表达的基因数分别是 27 和 61 个。

2.3.1 有关光合基因

高等植物叶绿体 30 个光合基因中有 5 个 *psa*, 12 个 *psb*, 5 个 *pet*, 6 个 *atp*, 11 个 *ndh*, 1 个 *rbc* 等基因。4 个 *rnm*, 12 个 *rps*, 7 个 *rpl*, 30~32 个 *trn*。

藻类紫红紫菜叶绿体 50 个光合基因中有 8 个 *psa*, 11 个 *psb*, 4 个 *pet*, 8 个, 2 个 *rbc* 基因和 10 个基因编码藻胆蛋白,几乎所有已知的类囊体膜复合体,以及移动电子携带者 Cyt c533 和铁氧还蛋白均在质体 DNA 中编码。仅一个 ATPase 蛋白基因(*atpC*)、一个 PA II 蛋白基因(*psbM*)、一个 Cyt b6/f 复合体蛋白基因(*petC*)、和编码氧释放复合体(*psb*●)的 2 个亚基和一个 9-12kDa 蛋白的基因缺失。

藻类紫红紫菜叶绿体 8 个编码 PSI 蛋白的 *psa* 中,有 4 个(*psaD*, *psaE*, *psaF*, *psaL*)也存在于硅藻(*Odontella sinensis*)叶绿体中。其中 *psaE*, *psaF*, *psaL* 为高等植物所没有。*psaF* 也存在于 *Cyanophora paradoxa*, *psaL* 也为有色藻(*Cryptomonas*)所具有。纤维裸藻特有的 *psaM*。

紫红紫菜编码 PSII 蛋白 *psb* 中,缺少 *psbI* 和 *psbM*,纤维裸藻也缺少 *psbM*,唯仅蓝细菌和地钱具有 *psbM*。其中 *psbA* 研究较多,它编码 PSII 的 D1 蛋白。*psbD* 编码 PSII 的 D2 蛋白。*psbE* 和 *psbF* 则是编码细胞色素 b599 的 α 、 β 亚基。

高等植物叶绿体编码细胞色素复合体基因为 *petA*、*petB*、*petD* 和 *petE*,紫红紫菜则还含特有的 *petF*、*petJ* 和 *petK*。高等植物的 *petF* 为核编码,也存在于 *Cyanophora paradoxa*^[30] 和 *Cryptomonas*^[24] 中。*petJ* 编码细胞色素 c553,在高等植物其功能为质体蓝素所代替。*petK* 编码细胞色素 c550 其功能目前还不清楚。纤维裸藻则只有 *petB* 和 *petG*。

紫红紫菜叶绿体所有的 ATP 酶蛋白(除 r 亚基[*atpC*]之外)均由叶绿体编码。而在绿藻中,*atpD* 和 *atpG* 基因为核编码,而在硅藻(*Odontella sinensis*)^[31]、*Cyanodinium caldarium*^[32] 和红藻(*Cyanophora paradoxa*)均由叶绿体编码。

绿藻和陆地植物叶绿体的光合作用关键酶 Rubisco 大亚基由叶绿体编码(*rbcL*),小亚基则由核编码(*rbcS*)。而红藻和有色藻的 Rubisco 大小亚基均由叶绿体编码。用菠菜 Rubisco 大亚基 *rbcL* 作为探针,从紫球藻(*Porphyridium aeruginum*)叶绿体 DNA 中筛选分离出紫球藻的 *rbcL*,在测序时发现其下游一段 130bp 的开放阅读框与高等植物和蓝藻中的 *rbcS* 同源,且为同一操纵子。*rbcL* 和 *rbcS* 是被同时转录的。隐藻(*Cryptomonas*)和一个硅藻细柱藻(*Cylindrotheca* sp.)也是如此^[33]。另外在裸藻 Rubisco 大亚基基因中发现有 9 个内含子。

陆地植物 NADH 脱氢酶(现称为 NADH——质体醌氧化还原酶)复合体有 11 个蛋白基因(*ndh*)定位

于叶绿体基因组上^[34],也未出现在 *P. purpurea*^[1]及有色藻^[28]、裸藻(*E. gracilis*)^[19]、2种衣藻^[29]的叶绿体基因组中,因而认为NADH脱氢酶复合体蛋白基因是存在于原始的叶绿体中,随着红藻和有色藻的进化而被转移到核基因组中或被丢失。

另外紫红紫菜还有10个藻胆蛋白体多肽基因(别藻蓝蛋白 *acp*、藻蓝蛋白 *cpc* 和藻红蛋白 *cpe*)^[5]。

2.3.2 转录和翻译基因

高等植物叶绿体核糖体基因包括23S,26S,5S,和4.5SrRNA基因,绿藻没有4.5SrRNA,但存在另外2个rRNA(3S和7SrRNA)。玉米、烟草、眼虫藻的16SrRNA基因(*rrn*)与大肠杆菌的同源程度分别为74%、74%、72%,说明它们是相当保守的序列^[1]。紫菜(*P. purpurea*)也只有3个*rrn*(23S,16S,5S)^[5]。

在烟草叶绿体基因组与定位了34个tRNA基因(*trn*),紫菜(*P. purpurea*)叶绿体基因组上定位了20个*trns*,其中*trnA*-GGC、*trnL*-GAG和*trnS*-CGA是在陆生植物中所没有发现的^[17]。叶绿体*trn*基因更类似于原核生物,不似于真核和线粒体的*trn*基因。但它们没有3'-CCA末端,这又类似于真核生物^[1]。

叶绿体核糖体含有60多种蛋白质,其中1/3由叶绿体DNA编码。烟草叶绿体基因组已定位25个核糖体蛋白基因。紫菜(*P. purpurea*)则含有21个核糖体蛋白基因,其中9个不存在于陆生植物叶绿体。

RNA聚合酶的4个亚基基因 *rpoA*、*rpoB*、*rpoC1*、和 *rpoC2* 均存在于陆生植物和藻类中,但纤细裸藻叶绿体缺 *rpoA*。

紫红紫菜叶绿体基因组还具有4个翻译因子基因,其中有翻译延长因子亚基基因 *tufA* 和 *tsf* 和2个起始因子 β 基因(*infB*、*infC*),而高等植物只有1个翻译起始因子基因 *infA* 和1个延长因子 *tufA*。

更有趣的是紫菜还含有2个tRNA合成酶基因(*syfB*、*syh*),1个DNA复制因子基因 *dnaB* 及2个涉及tRNA成熟的基因(*rapB*和*rne*),5个转录调节蛋白基因(*trnA*、*trnB*、*trnC*、*trnD*、*trnE*),其中*trnB*在江蓠(*G. sulphuraria*)及对丝藻(*Antithamnion* sp)质体基因组以及隐藻(*Cryptomonas*)、*R. vilacea*和红藻(*C. paradoxia*)等藻中得到证实^[35,24,36],而红藻(*C. paradoxia*)和硅藻(*O. sinensis*)质体基因组却含有*trnE*^[1,28]。

2.3.3 生物合成基因及其它各种杂基因

藻类叶绿体基因组与高等植物叶绿体基因组的差异更主要体现在其有关生物合成基因及杂蛋白基因的差异性。高等植物叶绿体编码有关生物合成基因较少,而藻类较多。

紫红紫菜叶绿体编码合成基因包括:氨基酸生物合成基因(*argB*、*carA*、*gltB*、*ihvH*、*tpa*、*trpG*),脂肪酸生物合成基因(*accA*、*accB*、*accD*、*acpA*、*fabH*),叶绿体生物合成基因(*chlB*、*chlI*、*chlL*、*chlN*),类胡萝卜素合成基因(*preA*)及藻胆素(*pbsA*)和硫胺素(*thiG*)等合成基因,这些基因均已在紫菜(*P. purpurea*)叶绿体基因组上被定位^[1]。现知道,*accD*编码乙酰辅酶A羧化酶-羧基转移酶^[37]。另外还有一些基因涉及氮同化蛋白基因(*glnB*、*gltB*、*hesB*),氧化还原反应<硫氧化还原蛋白(*trxA*)和铁氧还原蛋白-硫氧还原蛋白还原酶(*ftrB*)>、涉及通过类囊体膜蛋白运输(*secA*、*secY*)和糖酵解酶(*pgmA*)以及蛋白酶(*clpC*、*fisH*)、chaperonin(*dnak*、*groEL*)、和丙酮酸脱氢酶亚基(*odpA*、*odpB*)也在紫菜(*P. purpurea*)及其它红藻和有色藻中被检测到。此外,一个未知功能的ATP结合蛋白(*cfx*)、1个推论为血红素结合蛋白基因(*hbp*)也在紫菜叶绿体基因组定位。另外还有一些基因不在紫菜(*P. purpurea*)中出现,而在红藻(*C. paradoxia*)中检测出,这些基因为1个核糖体蛋白(*rpl2*),1个chaperonin(*groES*),1个组氨酸运输蛋白(*hisP*),1个细胞分裂蛋白(*fisW*),及涉及到NAD(*nadA*),血红素,组氨酸(*hisH*)和类胡萝卜素(*crtE*)生物合成的酶基因^[16]。

烟草叶绿体基因组大约有23个未知产物基因序列(*yf*序列)。纤细裸藻大约有5个*yf*序列,10个开放阅读框(ORFs)未知其功能。在紫菜中,36个ORFs类似于其它种类的ORFs,29个ORFs为紫红紫菜独特的。其中ORFs31、31a、62分别类似于陆生植物ORF35、31、62。ORFs174、265和283类似红藻(*Cyanophora paradoxia*)的ORFs,并定位于其相类似的基因上。ORF199类似于鼠的红白血细胞系的特异蛋白(MER5)、*Helicobacter pylori*抗原、*Entamoeba histolytica*表面抗原和 *Salmonella typhimurium* 烷基氢过氧化

物酶^[17,38]。OFR587 则是类似于一群真核蛋白:Pas1p、Sec18p、NSF、Cdc48p 和 VCP^[17]。

还有一些基因不出现在紫菜(*P. purpurea*)叶绿体基因组中,而出现在陆地植物中,这些基因包括 *mbpX*, *mbpY* 和几个保守开放阅读框(ORFs: *ycf1*、*ycf2*、*ycf14*、*ycf15*)。

表 2 叶绿体基因组的基因
Tab.2 Genes of chloroplast genomes

基因	产物	基因	产物	基因	产物
I 光合作用基因					
<i>psaA</i>	PSI P700 叶绿素 A 脱辅基蛋白 A1	<i>rpoC1</i>	RNA 聚合酶亚基 β'	<i>rp131</i> *	50S 核糖蛋白体 CL31
<i>psaB</i>	PSI P700 叶绿素 A 脱辅基蛋白 A2	<i>rpoC2</i>	RNA 聚合酶亚基 β''	<i>rp132</i>	50S 核糖蛋白体 CL32
<i>psaC</i>	PSI 铁硫中心(亚基Ⅶ蛋白)	<i>dnaB</i> *	DNA 复制旋转酶	<i>rp133</i>	50S 核糖蛋白体 CL33
<i>psaD</i> *	PSI 反应中心亚基Ⅱ	<i>infA</i>	起始因子 1	<i>rp134</i> *	50S 核糖蛋白体 CL34
<i>psaE</i> *	PSI 反应中心亚基Ⅳ	<i>infB</i>	起始因子 2	<i>rp135</i> *	50S 核糖蛋白体 CL35
<i>psaF</i> *	PSI 反应中心亚基Ⅲ	<i>infC</i> *	起始因子 3	<i>rp136</i>	50S 核糖蛋白体 CL36
<i>psaG</i>	PSI 8kDa 蛋白	<i>ropB</i> *	rRNA 和 tRNA 成熟	<i>rps1</i> *	30S 核糖蛋白体 CS1
<i>psaH</i>	PSI 10.4kDa 蛋白	<i>rne</i> *	RNA 酶 E	<i>rps2</i>	30S 核糖蛋白体 CS2
<i>psaI</i>	PSI 反应中心亚基Ⅶ	<i>syfB</i> *	tRNA 合成酶	<i>rps3</i>	30S 核糖蛋白体 CS3
<i>psaJ</i>	PSI 反应中心亚基Ⅸ	<i>syh</i> *	tRNA 合成酶	<i>rps4</i>	30S 核糖蛋白体 CS4
<i>psaK</i> *	PSI 反应中心亚基Ⅹ	<i>trxA</i> *	转录调节蛋白调节子	<i>rps5</i> *	30S 核糖蛋白体 CS5
<i>psaL</i> *	PSI 反应中心亚基Ⅺ	<i>trxB</i> *	转录调节蛋白	<i>rps6</i> *	30S 核糖蛋白体 CS6
<i>psaM</i>	PSI 反应中心亚基 M	<i>trxC</i> *	转录调节蛋白	<i>rps7</i>	30S 核糖蛋白体 CS7
<i>psbA</i>	PSII 32kDa D1 蛋白	<i>trxD</i> *	转录调节蛋白	<i>rps8</i>	30S 核糖蛋白体 CS8
<i>psbB</i>	PSII 47kDa 叶绿素脱辅基蛋白	<i>trxE</i> *	转录调节蛋白	<i>rps9</i>	30S 核糖蛋白体 CS9
<i>psbC</i>	PSII 44kDa 叶绿素脱辅基蛋白	<i>trfA</i> *	延长因子 Ts	<i>rps10</i> *	30S 核糖蛋白体 CS10
<i>psbD</i>	PSII 34kDa D2 蛋白	<i>tyfA</i>	延长因子 Tu	<i>rps11</i>	30S 核糖蛋白体 CS11
<i>psbE</i>	PSII 8kDa Cytb559α 蛋白	23 S <i>rrn</i>	23SrRNA	<i>rps12</i>	30S 核糖蛋白体 CS12
<i>psbF</i>	PSII 4kDa Cytb559β 蛋白	16 S <i>rrn</i>	16SrRNA	<i>rps13</i> *	30S 核糖蛋白体 CS13
<i>psbG</i>	(<i>ndhK</i>)	7 S <i>rrn</i> *	7SrRNA	<i>rps14</i>	30S 核糖蛋白体 CS14
<i>psbH</i>	PSII 10kDa 磷蛋白	5 S <i>rrn</i>	5SrRNA	<i>rps15</i>	30S 核糖蛋白体 CS15
<i>psbI</i>	PSII I 蛋白	4.5S <i>rrn</i>	4.5SrRNA	<i>rps16</i>	30S 核糖蛋白体 CS16
<i>psbJ</i>	PSII J 蛋白	3 S <i>rrn</i> *	3SrRNA	<i>rps17</i> *	30S 核糖蛋白体 CS17
<i>psbK</i>	PSII K 蛋白	<i>trnA</i> - GGC *	Ala - tRNA(GGC)	<i>rps18</i>	30S 核糖蛋白体 CS18
<i>psbL</i>	PSII L 蛋白	<i>trnA</i> - UGC	Ala - tRNA(UGC)	<i>rps19</i>	30S 核糖蛋白体 CS19
<i>psbM</i>	PSII M 蛋白	<i>trnC</i> - GCA	Cys - tRNA(GCA)	<i>rps20</i> *	30S 核糖蛋白体 CS20
<i>psbN</i>	PSII N 蛋白	<i>trnD</i> - GUC	Asp - tRNA(GUC)	III 生物合成基因	
<i>psbT</i>	PSII T 蛋白	<i>trnB</i> - UUC	Glu - tRNA(UUC)	<i>accA</i> *	乙酰辅酶 A 羧化酶羧基转移酶 α 亚基
<i>psbU</i> *	PSII U 蛋白	<i>trnF</i> - GAA	Phe - tRNA(GAA)	<i>accB</i> *	乙酰辅酶 A 羧化酶生物素载体蛋白
<i>psbV</i> *	CytC550	<i>trnG</i> - GCC	Gly - tRNA(GCC)	<i>accD</i>	乙酰辅酶 A 羧化酶羧基转移酶 β 亚基
<i>psbW</i> *	PSII W 蛋白	<i>trnG</i> - UCC	Gly - tRNA(UCC)	<i>accE</i> *	酰载体蛋白
<i>psbX</i> *	PSII X 蛋白	<i>trnH</i> - GUG	His - tRNA(GUG)	<i>argB</i> *	乙酰谷氨酸激酶
<i>psbY</i> *	PSII Y 蛋白	<i>trnI</i> - GAU	Ile - tRNA(GAU)	<i>carA</i> *	氮甲酰基磷酸合成酶小亚基
<i>petA</i>	Cyt f	<i>trnI</i> - GAU	Ile - tRNA(CAU)	<i>carA</i> *	CytC 生物起源蛋白
<i>petB</i>	Cyt b6	<i>trnK</i> - UUU	Lys - tRNA(UUU)	<i>chlB</i>	原叶绿素酸酯还原酶 <i>chlB</i> 亚基
<i>petD</i>	Cyt b6/f 复合体亚基Ⅳ	<i>trnL</i> - GAG *	Leu - tRNA(GAG)	<i>chlI</i> *	镁 - 螯合酶亚基
<i>petE</i>	Cyt b6/f 复合体亚基	<i>trnL</i> - UAG	Leu - tRNA(UAG)	<i>chlL</i> *	硫铁 ATP 结合蛋白
<i>petF</i> *	铁氧还蛋白	<i>trnL</i> - UAA	Leu - tRNA(UAA)	<i>chlN</i> *	原叶绿素酸酯还原酶 <i>chlN</i> 亚基
<i>petG</i>	Cyt b6/f 复合体亚基 V	<i>trnL</i> - CAA	Iea - tRNA(CAA)	<i>crsE</i> *	Geranylgeranyl 焦磷酸合成酶
<i>petJ</i> *	CytC553	<i>trnM</i> - CAU	Met - tRNA(CAU)	<i>cysA</i>	硫酸运输系统 ATP 结合蛋白
<i>petL</i>	Cyt b6/f 复合体 3.5kDa 亚基	<i>trnM</i> - CAU	Met - tRNA(CAU)	<i>cysT</i>	硫酸运输系统 透酶蛋白
<i>petM</i> *	Cyt b6/f 复合体 4kDa 亚基	<i>trnN</i> - GUU	Asn - tRNA(GUU)	<i>fabH</i> *	3 - 酮 - [酰载体蛋白]合成酶 III
<i>ndhA</i>	NADH - 质体醌氧化还原酶亚基 1	<i>trnP</i> - UGG	Pro - tRNA(UGG)	<i>fabB</i> *	硫氧还蛋白 - 铁氧还蛋白还原酶 β 亚基
<i>ndhB</i>	NADH - 质体醌氧化还原酶亚基 2	<i>trnP</i> - GGG	Pro - tRNA(GGG)	<i>ftsH</i> *	细胞分裂蛋白
<i>ndhC</i>	NADH - 质体醌氧化还原酶亚基 3	<i>trnQ</i> - UUG	Gln - tRNA(UUG)	<i>ftsW</i> *	细胞分裂蛋白
<i>ndhD</i>	NADH - 质体醌氧化还原酶亚基 4	<i>trnR</i> - ACG	Arg - tRNA(ACG)	<i>glnB</i> *	谷氨酰胺合成酶调节子
<i>ndhE</i>	NADH - 质体醌氧化还原酶亚基 4L	<i>trnR</i> - UCU	Arg - tRNA(UCU)	<i>gltB</i> *	谷氨酸合成酶
		<i>trnR</i> - CCG	Arg - tRNA(CCG)	<i>hemA</i> *	谷氨酰 - tRNA 还原酶

(续上表)

基因	产 物	基因	产 物	基因	产 物
<i>ndhF</i>	NADH-质体醌氧化还原酶亚基 5	<i>trnS - GGA *</i>	Ser-tRNA(GGA)	<i>hisS *</i>	组氨酰-tRNA 还原酶
<i>ndhG</i>	NADH-质体醌氧化还原酶亚基 6	<i>trnS - CGA</i>	Ser-tRNA(CGA)	<i>hlpA *</i>	DNA 结合蛋白
<i>ndhH</i>	NADH-质体醌氧化还原酶 49kDa 亚基	<i>trnS - GCU</i>	Ser-tRNA(GCU)	<i>ibaB *</i>	乙酰乳酸合成酶大亚基
<i>ndhI</i>	NADH-质体醌氧化还原酶 I 亚基	<i>trnT - GGU</i>	Thr-tRNA(GGU)	<i>ibaH *</i>	乙酰乳酸合成酶小亚基
<i>ndhJ</i>	NADH-质体醌氧化还原酶 J 亚基	<i>trnT - UGU</i>	Thr-tRNA(UGU)	<i>lipB *</i>	硫辛酸蛋白连接酶 B
<i>ndhK</i>	NADH-质体醌氧化还原酶 K 亚基	<i>trnV - GAC</i>	Val-tRNA(GAC)	<i>matK</i>	核内含子成熟酶
<i>fixc</i>	31kDa 蛋白	<i>trnW - CCA</i>	Trp-tRNA(CCA)	<i>minD *</i>	隔定位蛋白
<i>atpA</i>	ATP 酶 CF1 α 亚基	<i>trnY - GUA</i>	Try-tRNA(GUA)	<i>minE *</i>	细胞分裂拓扑特异因子
<i>atpB</i>	ATP 酶 CF1 β 亚基	<i>rp11 *</i>	50S 核糖蛋白体 CL1	<i>nodA *</i>	NAD 合成酶 A
<i>atpD *</i>	ATP 酶 CF1 δ 亚基	<i>rp12 *</i>	50S 核糖蛋白体 CL2	<i>odpA *</i>	丙酮酸脱氢酶 α 亚基
<i>atpE</i>	ATP 酶 CF1 ϵ 亚基	<i>rp13 *</i>	50S 核糖蛋白体 CL3	<i>odpB *</i>	丙酮酸脱氢酶 β 亚基
<i>atpF</i>	ATP 酶 CFO 亚基 I	<i>rp14 *</i>	50S 核糖蛋白体 CL4	<i>pbsA *</i>	血红素加氧酶
<i>atpG *</i>	ATP 酶 CFO 亚基 II	<i>rp15 *</i>	50S 核糖蛋白体 CL5	<i>pgmA *</i>	磷酸甘油酸成熟酶
<i>atpH</i>	ATP 酶 CFO 亚基 III	<i>rp16 *</i>	50S 核糖蛋白体 CL6	<i>pheT *</i>	苯丙氨酸 tRNA 合成酶 β 亚基
<i>atpI</i>	ATP 酶 CFO 亚基 IV	<i>rp19 *</i>	50S 核糖蛋白体 CL9	<i>preA *</i>	异戊烯转移酶
<i>apcA *</i>	别藻蓝蛋白 α 亚基	<i>rp111 *</i>	50S 核糖蛋白体 CL11	<i>thiG *</i>	硫氨基酸生物合成蛋白 G
<i>apcB *</i>	别藻蓝蛋白 β 亚基	<i>rp112 *</i>	50S 核糖蛋白体 CL12	<i>trpA *</i>	色氨酸合成酶 α 亚基
<i>apcD *</i>	别藻蓝蛋白 α - β 亚基	<i>rp113 *</i>	50S 核糖蛋白体 CL13	<i>trpG *</i>	邻氨基苯甲酸合成酶组分 II
<i>pcE *</i>	藻蓝蛋白锚蛋白	<i>rp114 *</i>	50S 核糖蛋白体 CL14	IV 其它基因	
<i>pcF *</i>	别藻蓝蛋白 B18	<i>rp116 *</i>	50S 核糖蛋白体 CL16	<i>gfaQ *</i>	<i>rbcS</i> 的 ORF 下游
<i>pcpA *</i>	藻蓝蛋白 α 亚基	<i>rp118 *</i>	50S 核糖蛋白体 CL18	<i>clpC *</i>	Clp 蛋白酶 A1P 结合蛋白
<i>pcpB *</i>	藻蓝蛋白 β 亚基	<i>rp119 *</i>	50S 核糖蛋白体 CL19	<i>clpP</i>	蛋白酶
<i>pcpC *</i>	藻胆体杆核连接多肽	<i>rp120 *</i>	50S 核糖蛋白体 CL20	<i>dnaK *</i>	热激蛋白 70 型蛋白
<i>pcpA *</i>	藻红蛋白 α 亚基	<i>rp121 *</i>	50S 核糖蛋白体 CL21	<i>graEL *</i>	60kDa chaperonin
<i>pcpB *</i>	藻红蛋白 β 亚基	<i>rp122 *</i>	50S 核糖蛋白体 CL22	<i>groES *</i>	10kDa chaperonin
<i>rbcL</i>	Rubisco 大亚基	<i>rp123 *</i>	50S 核糖蛋白体 CL23	<i>hbp</i>	(血红素结合蛋白)
<i>rbcS *</i>	Rubisco 小亚基	<i>rp124 *</i>	50S 核糖蛋白体 CL24	<i>secA *</i>	Preprotein 移位酶亚基
II 转录和翻译基因		<i>rp127 *</i>	50S 核糖蛋白体 CL27	<i>secY *</i>	Preprotein 移位酶亚基
<i>rpoA</i>	RNA 聚合酶亚基 α	<i>rp128 *</i>	50S 核糖蛋白体 CL28	<i>trxA *</i>	硫氧还蛋白
<i>rpoB</i>	RNA 聚合酶亚基 β	<i>rp129 *</i>	50S 核糖蛋白体 CL29		

注: * 为藻类所特有

参 考 文 献

- 刘良式. 植物分子遗传学. 北京:科学技术出版社, 1997. 1-56
- Bealbrook J R, Boyrad L. Endonuclease recognition sites mapped on *Zea mays* chloroplast DNA. Proc. Natl Acad Sci USA, 1976, 73:4309-4313
- Rochaix J D. Restriction endonuclease map of the chloroplast DNA of *Chlamydomonas reinhardtii*. J Mol Biol, 1978, 126:597-617
- 何培民. 海藻基因工程的研究进展与现状. 见:海藻生物技术, 王素娟等主编. 杭州:浙江科学技术出版社, 1994. 143-148
- Reith M. Molecular biology of Rhodophyte and Chlorophyte plastids. Annu Rev. Plant Physiol. Plant Mol Biol, 1995, 46:549-75
- Ris H, Plant W. Ultrastructure of DNA containing area in the chloroplast of *Chlamydomonas*. J Cell Biol, 1962, 13:383-391
- Kirk J T D, Thley-Bassett R A E. The plastid: their chemistry, structure, growth and inheritance. Amsterdam/New York/Oxford, 1978, 527-559
- Kolodner R, Tewari K K. The molecular size and conformation of chloroplast DNA from higher plants. Biochem Biophys Acta, 1975, 402:372-390
- Myyanura S, Hori T. Presence of DNA in the pyrenoid matrix of the siphonous green alga, *Caulerpa octocorae* Web. v. Bos. Plant Morphology, 1989, 1:19-22
- Myyanura S, Kori T. DNA is present in the pyrenoid core of the siphonous of the genus *Caulerpa* and yellow green algae of the genus *Pseudodictyosiphon*. Protoplasma, 1991, 161:192-196
- Myyanura S, Hori T, Ohta T, et al. Co-localization of chloroplast DNA and ribulose-1,5-bisphosphate carboxylase/oxygenase in the so-called pyrenoid of the siphonous green alga *Caulerpa lentillifera* (Caulerpales, Chlorophyta). Phycologia, 1996, 35(2):156-160
- Kuhnel M, Kowalik K V. The plastid of a brown alga, *Dictyota dichotoma*. Mol Genet, 1987, 14:155-62
- Boyer C, Sauerbrikke C C, Le Gall Y, et al. Physical mapping of the plastid genome from the rhodophyte *Chondria crispus*. J Phycol, 1991, 27:s11
- Loseaux-de Coer S L, Markowicz Y, Dalmon J, et al. Physical maps of the two circular plastid DNA molecules of the brown alga *Phyllocladus littoralis* (L.) Kjellm. Curr Genet, 1988, 14:155-162

- 15 Shivji M S. Organization of the chloroplast genome in the red alga *Porphyra japonica*. *Curr Genet*, 1991, 19:49 ~ 54
- 16 Löffhardt W, Bohner H J. Structure and function of the cyanelle genome. *Int Rev Cytol*, 1994, 151:29 ~ 65
- 17 Rzeith M, Munnholland J. A high-resolution gene map of the chloroplast genome of the red *Porphyra purpurea*. *Plant Cell*, 1993, 5:465 ~ 75
- 18 Reith M, Munnholland J. The complete nucleotide sequence of the chloroplast genome of *Porphyra purpurea*. Presented at 4th Int. Congr Plant Mol Biol Amsterdam, 1994, Abstract 118
- 19 Hallick R B, Hong L, Drager R, et al. Complete sequence of *Euglena gracilis* chloroplast DNA. *Nucleic Acids Res*, 1993, 21:3537 ~ 3544
- 20 Shiozaki K, Ohne M, Tanaka M, et al. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expansion. *EMBO J*, 1986, 5:2043 ~ 2049
- 21 Hiratsuka J, Shimada H, Whitler R, et al. The Complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol Gen Genet*, 1989, 217:185 ~ 94
- 22 Ohyama K, Fukuzawa H, Kohchi T, et al. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha*. *Nature*, 1986, 355:265 ~ 267
- 23 Shivji M S, Li N, Cattolico R A. Structure and organization of rhodophyte plastid genomes: implications for the ancestry of plastids. *Mol Gen Genet*, 1992, 232:65 ~ 73
- 24 Douglas S E. Eukaryote endosymbioses: insights from studies of a cryptomonad alga. *Biosystems*, 1992, 28:57 ~ 68
- 25 Maerz M, Wolters J, Hafmann CJ, et al. Plastid DNA from *Pyrenomonas salina* (Cryptophyceae): physical map genes, and evolutionary implications. *Curr Genet*, 1992, 21:73 ~ 81
- 26 Kowallik K V. Origin and evolution of plastids from Chlorophyll - a + c containing algae: suggested ancestral relationships to red green algal Plastids. In *Origins of plastids*, ed. R A Lewin. New York: Chapman & Hall, 1992, 223 ~ 263
- 27 Bourne C M, Plamer J D, Stoermer E F. Organization of the chloroplast genome of the freshwater centric diatom *Cyclotella meneghiniana*. *J Phycol*, 1992, 28:347 ~ 55
- 28 Freier U, Stöbe B, Kowallik KV. The diatom *Odontella sinensis*: a missing link in chloroplast genomes evolution. Presented at 4th Int Congr. Plant Mol Biol, Amsterdam, Abstr, 1994, 116
- 29 Boutreau E, Otis C, Turmel M. Conserved gene clusters in the highly rearranged chloroplast genomes of *Chlamydomonas moewusii* and *C. reinhardtii*. *Plant Mol Biol*, 1994, 24:585 ~ 602
- 30 Bryant D A, Schluchter W M, Steward V L. Ferredoxin and ribosomal protein S10 are encoded cyanelle genome of *Cyanophora paradoxa*. *Gene*, 1991, 98:169 ~ 175
- 31 Panic P G, Sturruann H V. Chloroplast ATPase genes in the diatom *Odontella sinensis* reflect cyanobacterial characters in structure and arrangement. *J Mol Biol*, 1992, 224:529 ~ 536
- 32 Kostrewa M, Zetsche K. Large ATP synthase operon of the red alga *Antithamion* sp. Resembles the corresponding operon in cyanobacteria. *J Mol Biol*, 1992, 227:961 ~ 970
- 33 吴相钰, 吴光耀, 赵进东. 叶绿体分子生物学. 见: 植物生理与分子生物, 余叔文和汤章成主编. 北京: 科学技术出版社, 1998. 155 ~ 170
- 34 Arizmenil J M, Runswick M J, Skehel J M, et al. NADH: ubiquinone oxidoreductase from bovine heart mitochondria. A fourth nuclear encoded subunit with a homologue encoded in chloroplast genomes. *FEBS Lett*, 1992, 301:237 ~ 242
- 35 Zhou Y H, Ragan M A. Cloning and characterization of the nuclear gene encoding chloroplast glyceraldehyde-3-phosphate dehydrogenase from the marine red alga *Gracilaria verrucosa*. *Curr Genet*, 1994, 23:483 ~ 89
- 36 Kessler U, Maid U, Zetsche K. An equivalent tobacterial ompR genes is encoded on the plastid genome of red algae. *Plant Mol Biol*, 1992, 18:777 ~ 780
- 37 Li S L, Rock C O, Cruman J E. The *dedB* (*usg*) open reading frame of *Escherichia coli* encodes a subunit acetyltransferase Acarboxylase. *J Bacteriol*, 1992, 174:5754 ~ 5757
- 38 Yamamoto T, Matsui Y, Natori S, et al. Cloning of a housekeeping-type gene (*MER5*) preferentially expressed in murine erythroleukemia cells. *Gene*, 1989, 80:337 ~ 343