植物葉緑体ゲノムの遺伝子構成

京都大学理学部生物物理学教室 梅 園 和 彦

Phylogenetic conservation of a liverwort *Marchantia polymorpha* and tobacco *Nicotiana tabacum* chloroplast genomes has been elucidated by a comparative study of the complete nucleotide sequences. The genome encodes about 120 species of genes including those for proteins involved in photosynthesis and protein biogenesis and a complete set of rRNAs and tRNAs.

chloroplast genes/liverwort/tobacco/transcription/translation/intron

植物細胞は光合成や他の多くの生化学反応を行なう細胞内小器官（オルガネラ）として葉緑体を含んでおり、その内部に局在する環状二本鎖 DNA 分子には、葉緑体の維持や機能発現に必要な種々の遺伝子がコードされている。昨年わが国において、下等植物のゼニゴケ（121,024 塩基対）¹⁾と高等植物のタバコ（155,844塩基対）²⁾の両者について葉緑体の全塩基配列が解明され、多くの新しい知見が得られた。本稿では、これらの塩基配列データを比較しながら、植物葉緑体ゲノムの遺伝子構成について要約し、転写・翻訳機構などの特徴を議論する。なお、遺伝子地図の詳細についてはそれぞれの論文に譲り、紙面の関係上ここでは省略する。

葉緑体ゲノムの構成

環状の葉緑体 DNA 分子は多くの植物種から単離されており、その大きさは種によって12万塩基対から20万塩基対以上に及んでいる。また一般的に、葉緑体 DNA には1-7万塩基対から成る一対の逆位反復配列（Inverted Repeat, IR）が存在し、この領域によって残りの部分は大小ふたつの単一配列領域（Large Single Copy, LSC ; Small Single Copy, SSC）に区分されている³⁾。ゼニゴケ、タバコの葉緑体 DNA はいずれもこの構造を保持しており、ゼニゴケが最小クラス（IR の長さは10,058塩基対）、またタバコは高等植物の平均的な大きさといえる（IR, 25,339塩基対）。

塩基配列から推定された葉緑体 DNA 上の構造遺伝子はゼニゴケで128、タバコで154と見積られてい

る^{1),2)}。ただし、IR 領域内で重複している遺伝子を差し引いて種類で考えれば、この数は120ないし130程度となる。遺伝子産物の生体機能については、光合成や蛋白質合成への関与など、既に同定または推測がなされているものが約90種であり、残りの約30種は、31アミノ酸から2,136アミノ酸の大きさをもつ未同定蛋白質の遺伝子（ORF）である。しかしながら、未同定蛋白質遺伝子もその半数以上は、ゼニゴケ・タバコ間での対応関係が明らかであることから、塩基配列上の単なる“読み枠”にはとどまるものではなく、なんらかの機能をもった遺伝子である可能性が高い。

系統的に遠く離れた2種の葉緑体ゲノム間で遺伝子の配列順序を比較すると、LSC 領域で約3万塩基対ほどが逆向きに位置していること以外は、非常に良く保存されており、全ての植物葉緑体が共通の祖先に由来することを示唆している。さらに、葉緑体と大腸菌を比べても、rRNA やリボソーム蛋白質、ATPase サブユニットなどの遺伝子群では、同じ順序で対応する遺伝子が並んでおり、葉緑体の原核生物起源説を強く支持している。

光合成反応に関与する遺伝子

これまでに葉緑体ゲノム上に同定された遺伝子の多くは、光化学系Ⅰ（*psa* 遺伝子）、光化学系Ⅱ（*psb* 遺伝子）、電子伝達系（*pet* 遺伝子）、プロトン輸送性 ATPase（*atp* 遺伝子）など、光合成反応に関与するチラコイド膜蛋白質をコードしている⁴⁾（表1）。ストロマ蛋白質としては、Ribulose biphosphate carboxylase の大サブユニットの遺伝子（*rbcl*）が詳しく解析されている⁴⁾。これらの遺伝子発現については、光による転写制御の問題があり、光信号を受容するフィトクロムから核と葉緑体それぞれの転写系まで、どのよ

Gene organization of plant chloroplast genomes

Kazuhiko Umesono

Department of Biophysics, Faculty of Science, Kyoto University.

表1 A) 光合成に關与する遺伝子

ストロマの酵素	<i>rbcL</i> , リブロース 2リン酸カルボキシラーゼ (大サブユニット)
光化学系 I	<i>psaA</i> , P700クロロフィル <i>a</i> アポ蛋白質 A1 <i>psaB</i> , P700クロロフィル <i>a</i> アポ蛋白質 A2
光化学系 II	<i>psbA</i> , 32kd, D1, Q β 蛋白質 <i>psbB</i> , 51kd P680 クロロフィル <i>a</i> アポ蛋白質 <i>psbC</i> , 44kd P680 クロロフィル <i>a</i> アポ蛋白質 <i>psbD</i> , D2蛋白質 <i>psbE</i> , チトクローム <i>b</i> 559 サブユニット <i>psbF</i> , チトクローム <i>b</i> 559 サブユニット (タバコ, ORF39) <i>psbG</i> , G 蛋白質 (タバコ, ORF284または <i>bhpB</i>) <i>psbH</i> (ORF73), 10kd 燐蛋白質 (タバコ, <i>psbF</i>)
ATP アーゼ複合体	<i>atpA</i> , F1サブユニット α <i>atpB</i> , F1サブユニット β <i>atpE</i> , F1サブユニット ϵ <i>atpF</i> *, F0サブユニット I <i>atpH</i> , F0サブユニット III <i>atpI</i> , F0サブユニット VI または <i>a</i>
チトクローム <i>b6/f</i> 複合体	<i>petA</i> , チトクローム <i>f</i> <i>petB</i> *, チトクローム <i>b6</i> <i>petD</i> *, サブユニット 4

B) 新たに塩基配列から推定された遺伝子

<i>ndh</i> 遺伝子	<i>ndh1</i> *, <i>ndhA</i> *, ヒトミトコンドリア URF1 (ND1) に類似 <i>ndh2</i> *, <i>ndhB</i> *, ヒトミトコンドリア URF2 (ND2) に類似 <i>ndh3</i> , <i>ndhC</i> , ヒトミトコンドリア URF3 (ND3) に類似 <i>ndh4</i> , <i>ndhD</i> , ヒトミトコンドリア URF4 (ND4) に類似 <i>ndh4L</i> , <i>ndhE</i> , ヒトミトコンドリア URF4L (ND4L) に類似 <i>ndh5</i> , <i>ndhF</i> , ヒトミトコンドリア URF5 (ND5) に類似 <i>URF6</i> , ヒトミトコンドリア URF6 (ND6) に類似 (<i>ndh1-5</i> , <i>URF6</i> はゼニコケ, <i>ndhA-F</i> はタバコの名称)
その他	<i>frxA</i> , 原核生物の4Fe-4Sフェレドキシンに類似 <i>frxB</i> , 原核生物の4Fe-4Sフェレドキシンに類似 <i>frxC</i> , 原核生物の <i>nifH</i> 蛋白質に類似 <i>mbpX</i> , 大腸菌の <i>MalK</i> , サルモネラの <i>HisP</i> 蛋白質に類似 (<i>frx</i> , <i>mbp</i> 遺伝子はゼニコケのみ)

うな経路で情報が伝達・制御されているかが今後の課題である⁵⁾。また今回新たに、電子伝達に關与すると予想される一群の葉緑体遺伝子が見いだされており、ミトコンドリア呼吸鎖の NADH dehydrogenase のサブユニット遺伝子と類似する葉緑体遺伝子群 (*ndh* 遺伝子), さらにゼニコケでは、原核生物型の4Fe-4Sフェレドキシンや、窒素固定の際のニトロゲナーゼ (*nifH* 蛋白質) に高い相同性を示す鉄-硫黄蛋白質の遺伝子群 (*frx* 遺伝子) が報告されている (表1)。ゼニコケでのみ検出された *mbpX* 遺伝子は、その産物が能動輸送系の構成因子 (大腸菌の *MalK* 蛋白質など) に見られる特異的なアミノ酸配列 (特に ATP 結合領域)⁶⁾

を含んでおり、葉緑体内あるいは葉緑体-細胞質間における物質輸送の点からも興味を持たれる。ここで述べた遺伝子のうち, *petB*, *petD*, *atpF*, *ndh1* (*ndhA*), *ndh2* (*ndhB*) はイントロンを含んでいる。

遺伝情報の発現に關与する遺伝子

以上のような蛋白質遺伝子をもとより、葉緑体ゲノム上のすべての遺伝情報は、オルガネラ内部での転写・翻訳過程を経て発現している。葉緑体内には、固有の DNA ポリメラーゼ, RNA ポリメラーゼ, tRNA, リボソームが見いだされ、核-細胞質系とは独立した遺伝情報システムを形成している。

C) 遺伝子発現に関与する遺伝子

rRNA	16S, 23S, 5S, 4.5S (IR上で重複)
tRNA	ゼニコケ, 31種, 36遺伝子 (表 2) タバコ, 30種, 37遺伝子 (表 2)
RNA ポリメラーゼ	<i>rpoA</i> , <i>rpoB</i> , <i>rpoC1</i> *, <i>rpoC2</i> (タバコでは, <i>rpoA</i> , <i>rpoB</i> , <i>rpoC</i> *)
リボソーム蛋白質	30S: <i>rps2</i> , <i>rps4</i> , <i>rps7</i> , <i>rps8</i> , <i>rps11</i> , <i>rps12</i> *, <i>rps14</i> , <i>rps15</i> , <i>rps16</i> *(タバコのみ), <i>rps18</i> , <i>rps19</i> 50S: <i>rpl2</i> *, <i>rpl14</i> , <i>rpl21</i> (ゼニコケのみ), <i>rpl22</i> , <i>rpl23</i> , <i>rpl33</i> , <i>secX</i>
その他	<i>infA</i> , 大腸菌 IF-1に類似 <i>ssb</i> (ORF273), 大腸菌の単鎖 DNA 結合蛋白質に類似 (タバコのみ)

表 1 葉緑体 DNA にコードされる遺伝子¹⁾²⁾

*印はイントロンの存在を示す。表では省略してあるが、これらに加えて約30種の未同定蛋白質遺伝子があり、ゼニコケではその3種にもイントロが仮定されている (ORF135, ORF167, ORF203)。タバコの *rps7*, *rps12* (3'領域, 本文参照), *rpl2*, *rpl23*, *ndhB* 遺伝子は, rRNA 遺伝子同様, IR に在するので2コピーになっている。ゼニコケの IR は有意な蛋白質遺伝子を含んでいなかった。タバコの *infA* 遺伝子は開始コドンに欠いていることから、偽遺伝子である可能も指摘されている²⁾。

codon anticodon	codon anticodon	codon anticodon	codon anticodon
UUU Phe UUC Phe GAA UUA Leu TAA* UUG Leu CAA+	UCU Ser UCC Ser GGA UCA Ser TGA UCG Ser	UAU Tyr UAC Tyr GTA UAA ter UAG ter	UGU Cys UGC Cys GCA UGA ter UGG Trp CCA
CUU Leu CUC Leu CUA Leu TAG CUG Leu	CCU Pro CCC Pro CCA Pro TGG CCG Pro	CAU His CAC His GTG CAA Gln TTG CAG Gln	CGU Arg ACG # CGC Arg CGA Arg CGG Arg (CCG)
AUU Ile AUC Ile GAT* # AUA Ile CAT + AUG Met CAT fMet CAT	ACU Thr ACC Thr GGA ACA Thr TGT ACG Thr	AAU Asn AAC Asn GTT # AAA Lys TTT* AAG Lys	AGU Ser AGC Ser GCT AGA Arg TCT AGG Arg
GUU Val GUC val GAC # GUA Val TAC* GUG Val	GCU Ala GCC Ala GCA Ala TGC* # GCG Ala	GAU Asp GAC Asp GTC GAA Glu TTC GAG Glu	GGU Gly GGC Gly GCC CGA Gly TCC* GGG Gly

表 2 葉緑体 DNA にコードされる tRNA 遺伝子¹⁾²⁾

tRNA 遺伝子はアンチコドン位の DNA 塩基配列 (5'→3') で示した。CCG のアンチコドンをもつ tRNA 遺伝子がゼニコケで余分にあることを除けば、ゼニコケ、タバコ葉緑体ゲノムとも同種の tRNA 遺伝子をコードしている。アンチコドン右側の*印はイントロ、の存在を、#, +印は IR 領域に重複している tRNA 遺伝子 (#はゼニコケ、タバコの両方で、+はタバコで2コピー存在する)を示す。アンチコドン位が CAT の3種の異なる tRNA 遺伝子が存在するが、産物はそれぞれ tRNA-Ile, tRNA-Met, tRNA-fMet 分子であり、イソロイシンの AUA コドンに認識するものはアンチコドンの C に修飾を受けている。

これらの因子のうち、RNA 成分 (rRNA, tRNA) についてはその全ての遺伝子が葉緑体 DNA によってコードされ、4種の rRNA, 31種 (ゼニコケ) および 30種 (タバコ) の tRNA 遺伝子が同定されている。こ

のうち、6種の tRNA 遺伝子はイントロンを含み、4種の rRNA 遺伝子と tRNA 遺伝子の一部は、IR 領域に重複して存在する (表 1, 2)。翻訳の際、葉緑体で用いられる遺伝暗号は、変則性はなく普遍的なもの

であるが、tRNA 分子種が限られているため、ミトコンドリアと同様の拡張 wobbling^{71,81)}によってコドン認識を行っていると予想される(表2)。tRNA-Glu (UUC) 遺伝子の産物は、通常の転移 RNA としてだけでなく、クロロフィル合成にも必須であることが示されている⁹⁾。tRNA-Arg (CCG) 遺伝子は現在のところ、ゼニゴケにのみ見いだされているものである。

葉緑体リボソームは、30S, 50S サブユニットからなる70Sの大きさを持ち、その構成蛋白質も生化学また免疫学的な解析から、大腸菌のものと高い相同性を保持していることが指摘されてきた¹⁰⁾。葉緑体ゲノムにコードされるリボソーム蛋白質は、一次構造が既知である大腸菌蛋白質との構造的類似性によって同定され、ゼニゴケ、タバコともに19種が報告されている(*rps*, *rpl* 遺伝子)。*rpl21* (ゼニゴケのみ)と *rps16* (タバコのみ) 遺伝子以外は、*rps12*, *rpl2*, *rpl16* 遺伝子がイントロンを含むことなど、両ゲノムでその構成は一致している(表1)。また *secX* 遺伝子の産物も、大腸菌の相同蛋白質がリボソーム粒子中(50S)に見いだされる¹¹⁾ことから、事実上リボソーム蛋白質と判断される。この数(20種)は葉緑体リボソームを構成する蛋白質のはほぼ1/3にあたり、残りの2/3は細胞質で合成された後、葉緑体内へ移行して30S, 50S のサブユニット形成に参画すると考えられる。このほか、大腸菌の開始因子(IF1)に対応する *infA* 遺伝子も見いだされている。

RNA ポリメラーゼもリボソーム蛋白質と同様の手法によって葉緑体遺伝子が仮定されており、大腸菌 RNA ポリメラーゼの4種のサブユニット(α , β , β' , σ)のうち、 σ 以外の各サブユニットに対応する遺伝子が示されている。大腸菌にならって、 α サブユニット遺伝子は *rpoA*, β サブユニット遺伝子は *rpoB* と呼ばれているが、ゼニゴケ β' サブユニットは異なる2遺伝子によってコードされており、5'側から、*rpoC1* (イントロン1個を含み、 β' のN端側をコード)–73塩基対–*rpoC2* (イントロンはなく、C端側に対応)という順に並んでいる。一方、タバコでは *rpoC* という単一の遺伝子が想定され、多くのイントロンで分断されている可能性が指摘されている。RNA ポリメラーゼの σ サブユニットは転写プロモーターの特異的な識別に決定的な役割を持つことから、未同定蛋白質遺伝子のなかに σ 遺伝子があるか、細胞質からの供給が行なわれていると予想される。また、ホウレンソウを用いた生化学的な解析からは、全てのサブユニットが核DNAによってコードされる葉緑体

RNA ポリメラーゼの存在も報告されており¹²⁾、転写レベルでの遺伝子発現制御の問題と絡んで、今後の検討が期待される。

以上のように葉緑体の転写・翻訳装置は、その構成因子が少なくとも部分的には葉緑体ゲノム自身によってコードされているが、複製装置として既知のDNAポリメラーゼや大腸菌 *dna*, *pol* 遺伝子に相当するものを検出することはできなかった(大腸菌 *ssb* 蛋白質と部分的に類似するものはタバコで報告されている)。また、葉緑体DNA複製開始起点(*ori*)も、ゲノム上の明確な配列として特定されていない。

イントロン

葉緑体遺伝子のイントロンは、酵母など菌類ミトコンドリア遺伝子のイントロンと共通の構造的特徴を持ち、“Group I”, “Group II” という2群¹³⁾に分類できる。これらのグループに属するイントロンの中には、*in vitro* で蛋白質因子を全く要求せずに、いわゆる“self splicing”を行なうものが知られているが¹⁴⁾、葉緑体イントロンではこのような観察はされていない。

塩基配列からの予測を含めると、ゼニゴケで18種、タバコで16種の遺伝子にイントロンの存在が指摘されている(表1)。その長さは315から2,526塩基にわたり、最も長いtRNA-Lys(UUU)遺伝子のイントロンでは、内部に蛋白質をコードする領域がある(ゼニゴケでは370、タバコでは509アミノ酸)。遺伝子内での存在部位も含めて、イントロンの分布はゼニゴケ・タバコ間で非常に良く保存されており、ことにリボソーム蛋白質S12の遺伝子(*rps12*)では、第1エクソンと第2エクソンがゲノム上で数万塩基対も離れ、しかもそれぞれDNAの逆鎖でコードされていることが示されている^{15),16),17)}。これらふたつのエクソンが正しく連結されたRAN分子も既に同定されており^{18),19)}、独立に転写された前駆体間で *trans* にスプライシングの起こることが予想されている。

転写・翻訳単位とスプライシング

葉緑体の遺伝情報発現は、それを担うRNAポリメラーゼ、tRNA, rRNA, リボソーム蛋白質などが大腸菌のものと高い相同性を持っているように、システムとして非常に原核生物的である。転写プロモーター、ターミネーターの構造や、開始コドン上流のSD配列も保存されている。原核生物で転写調節の主要な単位となっているのがオペロンであるが、葉緑体でもこれに対応して、rRNAやtRNA遺伝子は勿論、複数の蛋

白質遺伝子が共転写を受けることが知られている。葉緑体の *atpB-atpE* 遺伝子のように、遺伝子の3'末端と5'末端がオーバーラップしている例からも明らかであろう。

原核生物においては、転写と翻訳が共役しているため転写途上のRNAもmRNAとしてポリソームへと取り込まれる。しかしながら葉緑体では、いくつかの蛋白質遺伝子がイントロンで分断されていること、イントロンを含む前駆体RNA分子が認められること（未発表）などから、転写と翻訳の場が区分されている可能性が高い。真核細胞の核遺伝子では、核で転写された前駆体RNAがスプライシングを受け、成熟mRNAとなって細胞質へ移行することでこの問題を回避している。スプライシングに限らず、多くの葉緑体蛋白質遺伝子について、ポリシストロニックなRNAがさらに短いRNA分子へとプロセッシングを受けることも示唆されているが^{(20), (21)}、このような方向の研究はまだ始まったばかりであり、データは十分ではない。

葉緑体はとかく光合成との関連においてのみ議論されることが多いが、全ゲノムの構造が明らかにされたことを含めて、今後さらに基本的な遺伝子発現に関する研究が進めば、核と葉緑体の相互作用や、光誘導の分子機構など、光合成器官としての全体像が浮かびあがってくるものと思われる。

本稿を執筆するにあたって、適切な御批判を頂きました京都大学理学部小関治男教授に感謝します。

文 献

- 1) Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H., Ozeki, H. (1986) *Nature*, **322**, 572-574.
- 2) Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B. Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H., Sugiura, M. (1986) *EMBO J.*, **5**, 2043-2049.
- 3) Palmer, J. D. (1985) *Ann. Rev. Genet.*, **19**, 325-354.
- 4) Crouse, E. J., Schmitt, J. H., Bohnert, H. J. (1985) *Plant Mol. Biol. Reporter*, **3**, 43-89.
- 5) Sasaki, Y., Sakihama, T., Kamikubo, T., Shinozaki, K. (1983) *Eur. J. Biochem.*, **133**, 617-620.
- 6) Higgins, C. F., Hiles, I. D., Salmond, G. P. C., Gill, D. R., Downie, J. A., Evans, I. J., Holland, I. B., Gray, L., Buckel, S. D., Bell, A. W., Hermodson, M. A. (1986) *Nature*, **323**, 448-450.
- 7) Barrell, B. G., Anderson, S., Bankier, A. T., de Bruijn, M. H. L., Chen, E., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R., Young, I. G. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 3164-3166.
- 8) Bonitz, S. G., Berlani, R., Coruzzi, G., Li, M., Macino, G., Nobrega, F. G., Nobrega, M. P., Thalenfeld, B. E., Tzagoloff, A. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 3167-3170.
- 9) Schön, A., Krupp, G., Gough, S., Berry-Lowe, S., Kannan-gara, C. G., Söll, D. (1986) *Nature*, **322**, 281-284.
- 10) Schmidt, R. J., Hosler, J. P., Gillham, N. W., Boynton, J. E. (1985) *Molecular Biology of the Photosynthetic Apparatus* (Steinback, K. E., Bonitz, S., Arntzen, C. J., Bogorad, L., eds.) 417-427, Cold Spring Harbor Laboratory, New York.
- 11) Wada, A. (1986) *J. Biochem.*, **100**, 1595-1605.
- 12) Lerbs, S., Brautigam, E., Parthier, B. (1985) *EMBO J.*, **4**, 1661-1666.
- 13) Michel, F., Dujon, B. (1983) *EMBO J.*, **2**, 33-38.
- 14) Cech, T. R. (1986) *Cell*, **44**, 207-210.
- 15) Fukuzawa, H., Kohchi, T., Shirai, H., Ohyama, K., Umesono, K., Inokuchi, H., Ozeki, H. (1986) *FEBS Lett.*, **198**, 11-15.
- 16) Fromm, H., Edelman, M., Koller, B., Goloubinoff, P., Galun, E. (1986) *Nucl. Acids Res.*, **14**, 883-898.
- 17) Torazawa, K., Hayashida, N., Obokata, J., Shinozaki, K., Sugiura, M. (1986) *Nucl. Acids Res.*, **14**, 3143.
- 18) Zaita, N., Torazawa, K., Shinozaki, K., Sugiura, M. (1987) *FEBS Lett.*, **210**, 153-156.
- 19) Hildebrandt, M., Hallick, R. B., Passavant, C. W., Bourquel, D. P. (1987) *Cell*, in press.
- 20) Rodermer, S. R., Bogorad, L. (1985) *J. Cell Biol.*, **100**, 463-476.
- 21) Bird, C. R., Koller, B., Auffret, A. D., Huttly, A. K., Howe, C. J., Dyer, T. A., Gray, J. C. (1985) *EMBO J.*, **4**, 1381-1388.

うめその かずひこ

昭和61年京都大学理学部生物物理学科、博士課程終了。理学博士。現在、日本学術振興会特別研究員。葉緑体に限らず、ひろく遺伝子の構造と発現様式に興味があります。塩基配列を眺めてああでもない、こうでもない長時間ボーッとしているので、周りから呆れられています。

☎075-751-2111 内線4200