

博士論文

ゲノム解読に基づく藻類進化の研究

田島 直幸

要旨

現在、地球上にはさまざまな種類の藻類が存在している。真核光合成生物が持つ葉緑体の起源は、その祖先の細胞内に共生したシアノバクテリアであるという細胞内共生説が広く受け入れられている。真核光合成生物は、シアノバクテリア由来の葉緑体を持つアーケプラスチダと、アーケプラスチダがさらに他の真核生物の細胞内に共生したことに由来する葉緑体を持つ二次共生藻類に分けられる。近年、次世代シーケンサーの発展に伴い、さまざまな生物のゲノムが解読されてきているが、ゲノムの解読された種が存在しない系統もまだ数多く存在する。広範囲な系統の生物のゲノムを解読することで、生物全体の進化を考察する新たな知見が得られると期待される。

アーケプラスチダは、紅色植物門、灰色植物門、緑色植物亜界に分けられ、二次共生藻類は紅色植物門由来の葉緑体を持つものと緑色植物亜界の由来の葉緑体を持つものに分けられる。紅色植物門および紅色植物門由来の二次共生藻類は、緑色植物亜界にも劣らない多様性があり、地球環境に与える影響も大きい。ゲノムに関する知見はまだ少ない。本博士論文では、紅色植物門および紅色植物門由来の二次共生藻類のゲノム解読に重点を置き、紅色植物門のチノリモ綱チノリモ *Porphyridium purpureum* の葉緑体ゲノム、紅色植物門由来の二次共生藻類であるパルマ藻 *Triparma laevis* の葉緑体ゲノムとミトコンドリアゲノム、シアノバクテリア *Limnothrix* のゲノムをそれぞれ解読し、コードされた遺伝子の種類やゲノムの構成を関連生物と比較した。最後に、これら 3 種の生物のゲノムと、今まで解読された藻類のゲノムを基に系統解析を行った。

紅色植物門は 2 種類の亜門、7 種類の綱に分類される。しかし、これまでに葉緑体ゲノムが解読されている紅色植物門の種類は限られており、紅色植物門の系統関係も完全に解明されてはいない。さらに、紅色植物門由来の二次共生藻類の起源も、どの系統であるか詳しく分かっていない。そこで、葉緑体ゲノムが未解読であるチノリモ綱に属し、純粋培養系が確立されているチノリモ (*P. purpureum* NIES 2140) の葉緑体ゲノムを解読した。チノリモの葉緑体ゲノムは全長 217,694 bp の環状 DNA であった。224 個のタンパク質遺伝子、29 個の tRNA 遺伝子、6 個の rRNA 遺伝子、1 個の ncRNA 遺伝子をコードすることが推定された。チノリモの rRNA オペロンは、rRNA 遺伝子はそれぞれ 2 コピーずつあり、逆向き反復配列の構造をとっていたが、16S rRNA 遺伝子と 23S rRNA 遺伝子の

間に異なる tRNA 遺伝子が存在することが分かった。多くの種の葉緑体ゲノムは、2 コピーの rRNA オペロンを互いに逆向きで持っているが、rRNA オペロン内の tRNA 遺伝子も含めて同じ構造をしており、チノリモ葉緑体の rRNA オペロン構造は独自のものであることが分かった。また、リボソームタンパク質遺伝子クラスターは、他の紅色植物門や紅色植物門由来の二次共生藻類と異なり、途中で途切れて離れた場所に存在するという構造をしており、チノリモの葉緑体ゲノムは、他の種と分岐した後で大規模なゲノムの再編成が起きたのではないかと考えられた。

パルマ藻は、数枚の珪酸の殻で覆われた、直径約 2-5 μm の単細胞藻類である。最近、珪酸の細胞壁を染色する技術により、パルマ藻を分離・単種培養する方法が確立された。リボソーム小サブユニット rDNA および *rbcL* の分子系統解析の結果から、パルマ藻は珪藻の姉妹群に位置することが示唆された。パルマ藻のゲノムを解読し、珪藻と比較することで、パルマ藻に加え、珪藻の進化についても知見が得られることが期待された。パルマ藻 *T. laevis* の葉緑体およびミトコンドリアゲノムを解読した結果、それぞれ全長 117,514 bp、39,580 bp の環状 DNA であった。葉緑体ゲノムは 141 個のタンパク質遺伝子、29 個の tRNA 遺伝子、6 個の rRNA 遺伝子、2 個の ncRNA 遺伝子をコードし、ミトコンドリアゲノムは 37 個のタンパク質遺伝子、25 個の tRNA 遺伝子、2 個の rRNA 遺伝子をコードすることが推定された。ミトコンドリアゲノムにコードされたタンパク質遺伝子の種類は、珪藻綱や褐藻綱等が属する不等毛植物門の間で、ほぼ共通していることが分かった。一方、葉緑体ゲノムにコードされたタンパク質遺伝子の種類は、いくつか異なっていた。不等毛植物門の系統関係から考えると、珪藻とパルマ藻の共通祖先種が、褐藻等の祖先種と分岐した頃に、いくつかの遺伝子が失われたと推定された。また、パルマ藻と珪藻の葉緑体ゲノムは、リボソームタンパク質遺伝子クラスターの位置が異なるものの、他の不等毛植物門の藻類と比べると遺伝子の並びが保存されており、パルマ藻と珪藻では、分岐した後も大規模な葉緑体ゲノムの再編成は起きていないと考えられた。

Limnothrix は *Pseudanabaena* に近縁の糸状性シアノバクテリアである。この種は、窒素欠乏時に窒素固定能に特化した細胞、ヘテロシストを形成しないことが知られている。今回解析に用いた ABRG5-3 株のゲノムを解読した結果、メインゲノムは全長約 4.8 Mbp の環状 DNA であった。4,318 個のタンパク質遺伝子、

52 個の tRNA 遺伝子、9 個の rRNA 遺伝子、4 個の ncRNA 遺伝子をコードすることが推定された。また、7 種のプラスミドを持つことが分かった。そのうち、窒素固定に関わる遺伝子を多く持つプラスミドがあることが分かった。

現在、ゲノムが解読されたシアノバクテリアの種はさらに増えてきている。本博士論文では、既に解読されているシアノバクテリアゲノムと *Limnothrix* のゲノムにコードされる遺伝子やゲノム構造を比較し、シアノバクテリアの進化と合わせて *Limnothrix* のゲノムについてまとめた。

これまでの結果および既に解読されている藻類のゲノムをもとに、葉緑体ゲノムに共通して保存されているタンパク質遺伝子のアミノ酸配列と 16S rRNA 遺伝子の塩基配列、ミトコンドリアゲノムに共通して保存されているタンパク質遺伝子のアミノ酸配列を用いて系統解析を行った。パルマ藻は、珪藻の姉妹群であることが全ての系統解析結果から支持された。チノリモ綱は、真正紅藻綱およびウシケノリ綱と、イデユコゴメ綱の間に位置することが葉緑体ゲノムに保存されているタンパク質遺伝子アミノ酸配列の系統解析結果から支持された。紅色植物門由来の二次共生藻類の葉緑体は、真正紅藻綱とウシケノリ綱の共通祖先よりさらに外側の、イデユコゴメ綱以外の紅色植物門が属する紅藻植物亜門の共通祖先に由来することが示唆された。*Limnothrix* の系統はシアノバクテリアの進化の根元の方に位置することが、タンパク質遺伝子アミノ酸配列の系統解析結果から支持された。この系統解析結果をもとにシアノバクテリアゲノムの構造を比較した結果、細胞内共生のもとになったと考えられるシアノバクテリアの系統とそれ以外の系統で、リボソームタンパク質遺伝子クラスターの構造が異なることが分かった。

目次

要旨	1
目次	4
略語	6
第1章 序論・・・藻類進化に関するこれまでの研究	9
1.1. 葉緑体の起源に関する論説	9
1.2. 葉緑体と真核光合成生物の分類	10
1.3. 紅色植物門の系統分類	11
1.4. チノリモ綱	13
1.5. クロムアルベオラータという呼称について	13
1.6. パルマ藻と珪藻およびボリド藻との関係	14
1.7. シアノバクテリアの分類	15
1.8. <i>Limnothrix</i> sp. ABRG5-3	16
1.9. ゲノム解読方法の今昔	17
1.10. 系統解析の方法	18
1.11. 葉緑体の起源となるシアノバクテリアについての議論	20
1.12. 本博士論文の目的と構成	20
図表	22
第2章 海洋性単細胞紅藻チノリモ <i>Porphyridium purpureum</i> の葉緑体ゲノム 解読	27
2.1. 序文	27
2.2. 材料と方法	28
2.3. 結果	33
2.4. 考察	38
図表	39
第3章 パルマ藻 <i>Triparma laevis</i> の葉緑体とミトコンドリアのゲノム解読	53
3.1. 序文	53
3.2. 材料と方法	54
3.3. 結果	57
3.4. 考察	60
図表	62
第4章 シアノバクテリア <i>Limnothrix</i> sp. ABRG5-3 のゲノム解読	73
4.1. 序文	73
4.2. 材料と方法	74
4.3. 結果	78
4.4. 考察	84
図表	86
第5章 藻類の系統解析	105
5.1. 序文	105
5.2. 材料と方法	106

5.3. 結果	108
5.4. 考察	111
図表	112
第6章 総括	119
6.1. 本博士論文の総論	119
6.2. 細胞内共生に成功したと考えられるシアノバクテリアに関する仮説	120
6.3. 葉緑体ゲノム構造から見る紅藻とクロムアルベオラータの進化的関係	122
図表	124
参考文献	126
謝辞	141
付録1	142
付録2	155

略語

adk : adenylate kinase

CRISPR-Cas : clustered regularly interspaced short palindromic repeats – CRISPR associated protein

CTAB : 臭化ヘキサデシルトリメチルアンモニウム
(Hexadecyltrimethylammonium Bromide)

EDTA : エチレンジアミン四酢酸 (ethylenediaminetetraacetic acid)

EST : expressed sequence tag

EtBr : 臭化エチジウム (ethidium bromide)

FBA : フルクトース-1,6-ビスリン酸アルドラーゼ (fructose-1,6-bisphosphate aldolase)

GAPDH : グリセルアルデヒド-3-リン酸脱水素酵素 (glyceraldehyde-3-phosphate dehydrogenase)

GTR+G+I : general time-reversible (Gibson et al. 2005), Gamma-distributed among site rate variation, Invariant sites

IR : 逆向き反復配列 (Inverted Repeat)

LG : Le & Gascuel 2008

LR-ELW : the Expected-Likelihood Weights applied to Local Rearrangements of tree topology

LSC : Large Single Copy

map : methionine aminopeptidase

MCMC : マルコフ連鎖モンテカルロ法 (Markov chain Monte Carlo methods)

MCMCMC : Metropolis 共役 MCMC (Metropolis-coupled MCMC)

ncRNA : non-coding RNA

PDMPO : 2-(4-pyridyl)-5-([4-(2-dimethylaminoethylaminocarbonyl)-methoxy]phenyl) oxazole

RDP : Ribosomal Database Project

SAR : ストラメノパイル(Stramenopiles)、アルベオラータ(Alveolates)、リザリア (Rhizaria)を1つの系統群にまとめたスーパーグループ

SBS : Sequence-by-Synthesis

SSC : Small Single Copy

TPR : tetratricopeptide repeat

Tris : トリス(ヒドロキシメチル)アミノメタン (Tris(hydroxymethyl)aminomethane)

WAG : Whelan & Goldman 2001

第 1 章

序論・・・藻類進化に関するこれまでの研究

1.1. 葉緑体の起源に関する論説

葉緑体は、真核光合成生物に特有のオルガネラで、酸素発生型の光合成を行い、独自のゲノムと遺伝子発現系をオルガネラ内に持つなどの特徴がある。

以前は、藻類や植物などの真核光合成生物は、原核生物から直接進化したものであるとする、直系分化説が支配的であった。その説では、葉緑体を含めたオルガネラは、原核的真核生物自身から分化し、内発自生的に出現したと考えられていた(Taylor 1976)。一方、Schimper (1883)は、顕微鏡観察の結果から、葉緑体が独自に分裂することで数を増やし、自然発生的に生まれるものではないことを示し、葉緑体が共生由来の組織であることを述べていた。Mereschkowsky (1905)は、シアノバクテリアが真核生物に共生したものが現在の葉緑体の起源であるという考えを述べた。その後、Margulis が葉緑体、ミトコンドリアおよび鞭毛の起源は、それぞれシアノバクテリア、好気性真正細菌およびスピロヘータが真核生物に共生した結果であるとする細胞内共生説を唱えた(Sagan 1967)。Margulis は、電子顕微鏡を用いて真核生物と原核生物の細胞を観察し、分裂様式や遺伝方式が両者で全く異なっており、その中間的生物が発見されなかったことを指摘した。他にもこの説を支持する研究として、ヌクレアーゼ処理と電子顕微鏡での観察による葉緑体内の核様体の発見(Ris & Plaut 1962, von Wettstein 1967)、2 種類の異なる沈降定数の DNA の存在(Chun et al. 1963, Bogorad 1967)およびそれらがそれぞれ核と葉緑体の DNA であること(Sagan et al. 1965)、葉緑体を失った細胞からは、葉緑体の DNA と思われるバンドの方が消失していること(Leff et al. 1963)、葉緑体が一度失われた細胞からは、葉緑体が再分化することは無かったこと(Edelman et al. 1964, 1965)や、葉緑体が細胞質遺伝すること(Lederberg 1952)が報告された。それ以外にも、葉緑体もシアノバクテリアも、カロテノイドを含んだ 2 枚の包膜に囲まれていること(Krinsky 1966)、葉緑体のリボソームは、細胞質のリボソームを阻害するシクロヘキシミドではなく、細菌のリボソームを阻害するストレプトマイシンによって阻害されること(Kirk &

Tilney-Bassett 1967)、細胞核の転写は影響を受けないが、原核生物の転写は阻害するリファンピシン(リファンピン)によって葉緑体の転写が阻害されること(Surzycki 1969)、葉緑体もシアノバクテリアもチラコイド膜で光の吸収や最初の光反応、電子伝達を行い、光合成は酸素発生を伴うこと等が報告され、葉緑体とシアノバクテリアの類似性が示された(Margulis 1970)。その後、シアノバクテリアや葉緑体のゲノム解読が進み、rRNA 遺伝子の配列(Urbach et al. 1992)や ATPアーゼのアミノ酸配列をもとにした系統解析の結果も、この説を裏付けた(Margulis 1993)。現在では、この細胞内共生説は鞭毛について以外は広く受け入れられている。

1.2. 葉緑体と真核光合成生物の分類

葉緑体には大きく分けて2つのタイプが存在する。1つは、シアノバクテリアが直接真核生物に細胞内共生してできたもの(一次共生)で、もう1つは、真核光合成生物がさらに別の真核生物に細胞内共生してできたもの(二次共生)である(Kim & Archibald 2009)。前者の葉緑体を持つ真核生物を一次植物と呼び、後者の葉緑体を持つ真核生物を二次植物と呼ぶ。主な一次植物は、灰色植物門、紅色植物門、緑色植物亜界の3グループに分かれるアーケプラスチダに分類される。アーケプラスチダの葉緑体の起源となるシアノバクテリアについては諸説ある(1.11.で詳しく述べる)。それ以外に、アーケプラスチダと異なる系統で、*Prochlorococcus/Synechococcus* の系統に所属するシアノバクテリアが起源であると考えられているクロマトフォアを持つ、有殻アメーバ *Paulinella chromatophora* がある (Marin et al. 2005)。二次植物は、紅色植物門由来の葉緑体を持つものと緑色植物亜界由来の葉緑体を持つものの2種類が存在する。二次植物では、ほとんどの場合、由来となる藻類の核やミトコンドリアは消失してしまっている。クリプト植物門とクロララクニオン植物門は、由来となる藻類の核をヌクレオモルフという、由来となる藻類の核が縮小したオルガネラとして保存している。(Greenwood 1974, Hibberd & Norris 1984)。渦鞭毛植物門は、紅藻由来の葉緑体を持つ種だけではなく、緑藻由来の葉緑体や、クリプト藻、ハプト藻、珪藻等、紅藻由来の二次共生藻類の三次共生が由来となる葉緑体を持つ種が存在することが知られている(Larsen 1992, Schnepf 1992) (図 1-1)。さまざまな系統が持つ葉緑体の特徴を表に示した(表 1-1)。

緑色植物亜界は、陸上植物を含むストレプト植物と緑藻植物門に分けられる。これらの分類群の葉緑体ゲノムは、既に数多く解読されている。灰色植物門で葉緑体ゲノムが解読された種は、*Cyanophora paradoxa* の1種のみである。紅色植物門は7綱に分けられ、そのうち3綱9種の葉緑体ゲノムが解読されている。解読された葉緑体ゲノムの種数については、全オルガネラゲノムが解読された種が登録されているサイト、GOBASE (O'Brien et al. 2009)を主に参照した。ただし、GOBASEの更新は2010年8月に終了したため、それ以降に全ゲノムが解読された種については、NCBIで検索した。3グループの葉緑体ゲノムにコードされたタンパク質遺伝子を比較すると、紅色植物門が最も多くのタンパク質遺伝子をコードしていることが分かる。また、緑色植物亜界は、NADH デヒドロゲナーゼ関連遺伝子群、紅色植物門と灰色植物門はフィコビリソーム関連遺伝子群、などグループ毎にそれぞれ他の系統には無い遺伝子を保存している(図1-2)。紅色植物門由来の二次植物は、緑色植物亜界由来の二次植物と比べ、より複雑な生物群となっている(1.5.で詳しく述べる)。陸上の多細胞真核光合成生物は、緑色植物亜界の陸上植物が主だが、海の多細胞真核光合成生物は、不等毛植物門に属する褐藻が主となっている。また、不等毛植物門に属する珪藻は、海洋全体での光合成量が地球全体の約40%ともいわれ(Nelson et al. 1995)、海洋の生態系を支えている。このような紅色植物門由来の二次植物についての知見を深めることは、緑色植物亜界にも劣らず重要であると考えられる。紅色植物門は、紅色植物門由来の二次植物とシアノバクテリアの間をつなげる生物群であり、こちらについて知見を深めることもまた、紅色植物門への理解だけでなく、紅色植物門由来の二次植物への理解に重要であると考えられるが、そのゲノムに関する知見はまだ少ない。

1.3. 紅色植物門の系統分類

紅色植物門は紅藻とも呼ばれる。紅藻は、生活環の全てを通じて鞭毛や中心小体を持たない(Gabrielson et al. 1990, Graham & Wilcox 2000)。紅藻は、かつて原始紅藻綱と真正紅藻綱の2種類の綱(または亜綱)に分けられており(Garbary & Gabrielson 1990)、原始紅藻綱はさらに6目に分けられていた。しかし、この分類に対して、分子系統解析の結果から、真正紅藻綱とウシケノリ目が単系統になる(Yoon et al. 2002)ことや、チノリモ目が3系統に分かれる(Oliveira &

Bhattacharya 2000)こと等の報告がされた。このことは、葉緑体の構造等、形態学から見ても矛盾していなかった(Gabrielson et al. 1985, 1990, Garbary & Gabrielson 1990)。

Saunders & Hommersand (2004)は、分子系統解析結果やゴルジ体と小胞体の位置関係等の超微細構造の違いから、紅藻の新しい系統分類を提唱した。この系統分類では、紅藻はイデユコゴメ門と紅色植物門の 2 門を持つ亜界として扱われた。紅色植物門は 4 綱に分けられたが、この分類でもロデラ綱は多系統になったままであった。

Yoon et al. (2006)は、紅藻の 25 属 48 種を用いて系統解析を行い、紅藻を 7 種類の綱に分けた(図 1-3)。この分類でもイデユコゴメ綱が一番先に分岐したが、イデユコゴメ綱も他の紅藻と同様に紅藻デンプンを細胞質に貯蔵する点(Barbier et al. 2005)や、葉緑体が 2 枚の包膜に囲まれ、クロロフィル *b* や *c* を持たない点等共通点が多いことから、Yoon 等は、紅藻を門で分ける必要はないと考え、イデユコゴメ綱をイデユコゴメ亜門、それ以外の綱を紅藻植物亜門という 2 種類の亜門に紅色植物門を分けた。紅藻を 7 種類の綱に分けたが、綱同士の系統関係は、イデユコゴメ綱が最初に分岐する点、真正紅藻綱とウシケノリ綱が単系統となる点、かつてのチノリモ目が 3 系統に分けられ、それらが単系統とはならない点以外は、系統樹作成の際に選んだ遺伝子の数、用いた配列がタンパク質遺伝子のアミノ酸配列か rRNA 遺伝子の塩基配列か、系統解析に用いた生物種数の違いによって異なる結果をもたらしていた。これまでに葉緑体ゲノムの全長配列が解読された紅藻も、イデユコゴメ綱、真正紅藻綱、ウシケノリ綱の 3 綱のみであり、紅藻の系統関係はその 3 綱以外、詳細は未解明のままである。また、オオイシソウ綱が単系統にならないこともある(Zuccarello et al. 2008)ため、今後この系統分類も書き換えられるかもしれない。

紅藻の進化の歴史を調べる上で、2 つの重要な化石がある。1 つはウシケノリ綱 *Bangia* によく似た糸状多細胞紅藻 *Bangiomorpha pubescens* (Butterfield 2000)である。この化石は、約 1200 Ma (1 Ma は 100 万年前)の地層から見つかっており、性分化の痕跡が含まれている。2 つ目は、599 Ma の地層から見つかったサンゴモ目(真正紅藻綱)のもので、典型的生殖構造を持っているものだった(Xiao et al. 1998, 2004)。Yoon et al. (2004)は、分子時計を利用して、紅藻の分岐は 1449-1513 Ma の間、イデユコゴメ綱の分岐は 1350-1416 Ma の間、真正紅藻綱の分岐は約

800 Ma に起きたと概算した。

1.4. チノリモ綱

チノリモ綱は、単細胞で枝分かれ状もしくは星状の葉緑体を1個持っている。ピレノイドは持つものと持たないものがある。紅藻の分類にはゴルジ体の位置関係も利用されており(Saunders & Hommersand 2004)、チノリモ綱はゴルジ体がミトコンドリアと小胞体に隣接している。低分子炭水化物としてフロリドシドを貯蔵する。増殖は細胞分裂によって行われる。チノリモ綱は、硫酸化したキシロース、ガラクトース、グルコース、グルクロン酸からなる多糖のゲルに包まれている(Patel et al. 2013)。海洋や汽水域、土壌中でも発見されている。本博士論文で葉緑体ゲノムを解読した *Porphyridium purpureum* NIES 2140 は、国立環境研究所保有株の星状の葉緑体を持つ海洋性単細胞紅藻である(図 1-4a)。

1.5. クロムアルベオラータという呼称について

紅色植物門由来の二次植物である不等毛植物門、クリプト植物門、ハプト植物門は、かつて一度の細胞内共生から生まれたとされ、単系統の分類群クロミスタ(Cavalier-Smith 1981)が提唱された。根拠として、これらの藻類が共通して4枚の葉緑体包膜に囲まれ、クロロフィル *c* を持っていること等があげられる。さらに、アルベオラータも含めたクロムアルベオラータ(Cavalier-Smith 1999)も単系統であると提唱された。根拠として、グリセルアルデヒド-3-リン酸脱水素酵素(GAPDH)やフルクトース-1,6-ビスリン酸アルドラーゼ(FBA)の系統解析の結果(Harper & Keeling 2003, Patron et al. 2004)と、葉緑体遺伝子の系統解析の結果(Yoon et al. 2002)があげられた。しかし、核にコードされたいくつかの遺伝子を結合した系統解析の結果から、クロムアルベオラータ全体は多系統群であることが指摘された(Burki et al. 2012)。また、ストラメノパイルとアルベオラータはリザリアと同じ共通祖先を持つという系統解析結果が報告され(Burki et al. 2007, 2008, Hackett et al. 2007)、SAR (Stramenopiles, Alveolates, Rhizaria)というスーパーグループが新しく提唱された。Cavalier-Smith は、二次共生が起きるより、一度獲得した葉緑体を失う方が起こりやすいという考えから、一度の紅藻の二次共生からクロムアルベオラータの共通祖先が生まれたと主張した。しかし、緑色植物亜界由来の二次植物であるユーグレノゾアとクロララクニオン植物門は、

それぞれユーグレノゾアはエクスカバータ、クロララクニオン植物門はリザリアといった異なる生物群に属している(Adl et al. 2005, Keeling et al. 2005)。さらに、葉緑体局在の PsbO タンパク質や葉緑体にコードされたタンパク質の結合配列の系統解析から、ユーグレノゾアとクロララクニオン植物門の葉緑体はそれぞれ別の起源ではないかという仮説が支持された(Rogers et al. 2007, Takahashi et al. 2007)。また、渦鞭毛藻類の三次共生は、種によって異なる起源の二次植物が細胞内共生したものだと言われる(Archibald 2009)。紅藻由来の二次共生でも、同じように別々の起源の藻類が複数回にわたって細胞内共生することがないとは言えない。本博士論文では簡便のため、紅藻由来の二次共生藻類を総称してクロムアルベオラータという用語を用いた。

1.6. パルマ藻と珪藻およびボリド藻との関係

パルマ藻は、数枚の様々な形をした珪酸の殻で覆われた、直径 2-5 μm の単細胞海洋性ピコ植物プランクトンである(Booth & Marchant 1987)。熱帯から寒帯の水域にかけて幅広く分布しており、特に寒帯や亜寒帯の水域に多く生息していることが知られている。パルマ藻は、光学顕微鏡を用いて野外の他のピコ植物プランクトンから分離するにはとても小さく、今まで単離培養されていなかった。初めは、ケイ素の同位体組成や細胞のサイズが似ていたため、珪酸の殻に覆われた襟鞭毛虫の休止のう胞だと指摘された(Booth et al. 1980, Silver et al. 1980)。Marchant & McEldowney (1986)が電子顕微鏡の観察結果から、成長する植物プランクトンの細胞であると結論を下した。Booth & Marchant (1987)は、黄金色藻綱パルマ目を新設した。パルマ藻は珪酸の殻の数や種類で分類され、2種類の科に分けられた。5枚の珪酸の殻に覆われた Pentalaminae は、1種類の属 *Pentalamina* 属に1種のみ分類される。8枚の珪酸の殻に覆われた Triparmaceae は、3種の *Triparma* 属と5種の *Tetraparma* 属に分けられた。単離培養できなかったために系統学的知見は限られていたが、近年、ケイ素に結合し、新しく作られる珪酸の細胞壁に取り込まれる蛍光色素 PDMPO (Shimizu et al. 2001)を利用することで、北太平洋の親潮海域から天然のパルマ藻が単離された(Ichinomiya et al. 2010)。本博士論文で解説したパルマ藻は、この時単離された *Triparma* 属に属する *Triparma laevis* である(図 1-4b)。また、色素の構成や 18S rRNA と *rbcL* の分子系統解析の結果から、パルマ藻はそれまで属していた黄金色藻綱ではなく、

ボリド藻綱に含まれていることが報告された(Ichinomiya et al. 2011)。

ボリド藻は、無殻の海産性ピコ植物プランクトンで、分子系統解析の結果から、珪藻の姉妹群であると見られている(Guillou et al. 1999, Lovejoy et al. 2006)。

珪藻は、珪酸の殻に覆われた植物プランクトンである。熱帯から極地にかけて、海水、淡水、土壌と、ありとあらゆる環境に生息している種が存在しており、種数は10万種を超るともいわれている。海洋全体での光合成量は地球全体の約40%ともいわれ(Nelson et al. 1995)、海洋の生態系を支えている。珪藻は、大きく分けて殻面の模様が放射相称な中心珪藻と左右相称な羽状珪藻に分類される(井上 2007)。中心珪藻の方が先に出現していることが、化石からも確認されている。珪藻の被殻は細胞分裂の際に、半分は娘細胞にそのまま再利用されており、そのままだと細胞は次第に小さくなってしまう。細胞サイズを増大させるために、珪藻は有性生殖を行う。中心珪藻は、卵と鞭毛を持つ精子による卵生殖を行う。珪藻において鞭毛が確認されるのは、中心珪藻の精子のみである。パルマ藻は鞭毛を持たず珪酸の殻を持ち、ボリド藻は珪酸の殻を持たず鞭毛を持つとされる(Guillou et al. 1999)。しかし、形態の違いに比べて分子系統解析をもとにした系統関係が近いこと、両者もしくは両者の共通祖先も、中心珪藻のように鞭毛を持たず珪酸の殻を持つ時期と珪酸の殻を持たず鞭毛を持つ時期を切り替える生活環を持っており、現在はそれぞれ片方の生活環のみ確認されているだけなのかもしれない、という考え方もある。このように、パルマ藻やボリド藻について調べることは、珪藻およびかつての珪藻の祖先の性質を調べる上でも役に立つことが考えられる。

1.7. シアノバクテリアの分類

シアノバクテリアは、酸素発生型の光合成を行う原核生物である。光合成により炭素源を供給するだけでなく、窒素源の固定(Zehr et al. 2008)やさまざまな独特の二次代謝物を生産すること(Welker & von Döhren 2006)が知られている。シアノバクテリアの系統は、細胞の形態に基づき5種類に分けられている(Rippka et al. 1979)。単細胞シアノバクテリアは、二細胞に分裂するもの(subsection I, Chroococcales)と内生胞子(baeocytes)によって増えるもの(subsection II, Pleurocapsales)に分けられる。多細胞シアノバクテリアは、細胞分化せずに全て栄養細胞で、それらが一列につながった糸状体(トリコーム(trichome))を形成する

もの(subsection III, Oscillatoriales)と窒素固定に特化した細胞(ヘテロシスト(heterocyst))や、種によっては休眠細胞(アキネート(akinete))と移動性細胞(ホルモゴニア(hormogonia))に細胞分化するものに分けられる。細胞分化するものはトリコームのみ形成するもの(subsection IV, Nostcales)と、時折分枝する能力があるもの(subsection V, Stigonematales)に分けられる。*Limnothrix* sp. ABRG5-3 のゲノムを解読し始めた頃は、まだ Pleurocapsales と Stigonematales には、ゲノムが解読された種がいなかった。一方で、*Prochlorococcus/Synechococcus* のグループは、ゲノム解読されたシアノバクテリア 72 種中 29 種におよび、グループによってゲノム解読された種の数に偏りが生じていた。Shih et al. (2013)は、ドラフトゲノムも含めて新しく 54 種のシアノバクテリアゲノムを解読した。Stigonematales はドラフトゲノム解読までだったが、Pleurocapsales は完全に解読された種も存在した。系統解析の結果により、Stigonematales の系統と Nostcales, Stigonematales の系統は単系統となり、Chroococcales と Pleurocapsales と Oscillatoriales は複数の系統に分かれていることが分かった。また、タンパク質遺伝子アミノ酸配列での系統解析の結果から、細胞内共生が起きた系統ではないシアノバクテリアの系統は、最初に分岐したとされる *Gloeobacter violaceus* (Nakamura et al. 2003)、イエローストーン国立公園から単離された 2 種類の *Synechococcus* (Bhaya et al. 2007)の系統に加え、*Pseudanabaena* の系統という合計で 3 系統あることが示唆された。

1.8. *Limnothrix* sp. ABRG5-3

Limnothrix sp. ABRG5-3 は、熱海にある植物研究園内の池から単離(Nishizawa et al. 2010)された Oscillatoriales の一種である(図 1-4c)。*Limnothrix* は、*Pseudanabaena* や *Arthronema* と近縁種であるが、それらの分類は形態の観察と rRNA 遺伝子の一部の配列のみを根拠としていた。ほぼ全域の 16S rRNA 遺伝子をもとにした系統解析の結果、3 グループは入り混じってしまい、命名し直す必要があるかもしれないと考えられた(Nishizawa et al. 2010)。*Limnothrix/Pseudanabaena* のグループは、ヨーロッパ(Acinas et al. 2009)だけではなくアジア(Wang et al. 2009)からも海水、淡水を含めて広く発見されているが、それに対してゲノムのデータは少なかった。

Limnothrix sp. ABRG5-3 のゲノムを解読している途中で、近縁種の

Pseudanabaena が、細胞内共生が起きた系統ではないシアノバクテリアの系統に属していることを知り (Shih et al. 2013)、*Limnothrix* sp. ABRG5-3 のゲノムを解読するだけでなく、*Limnothrix* sp. ABRG5-3 を含めた細胞内共生が起きた系統ではないシアノバクテリアの系統と、細胞内共生が起きたシアノバクテリアの系統のゲノム構成やコードされている遺伝子の種類を比較し、細胞内共生が起きた系統と起きなかった系統の違いを調べ、細胞内共生に成功した要因について推察することを目的とした。

1.9. ゲノム解読方法の今昔

今までは、ゲノムの解読をするためには、まず、染色体を制限酵素処理によって断片化し、その断片をλファージやコスミドに挿入したゲノムライブラリーを作成した。それぞれを超音波処理により断片化し、適度な長さのものを選りベクターに挿入し、大腸菌に形質転換した。確かに挿入されたことをコロニーPCRで確認し、それぞれの挿入DNAの塩基配列を解読した。塩基配列解読には、DNAポリメラーゼの合成を蛍光色素の付与等により識別可能にしたジデオキシヌクレオチドにより塩基特異的に停止させ、停止した塩基を読んでいくというサンガー法が用いられた。サンガー法は正確ではあるが、以上のような方法で全ゲノムを解読するのは手間も時間もかかった。

近年、次世代シーケンサーの発展により、数多くの生物のゲノムを短期間で解読することが可能になった。次世代シーケンサーを用いる場合、染色体を断片化し、アダプターを付けたものがゲノムライブラリーとなり、以降の操作は次世代シーケンサーの装置が行い、解読する。次世代シーケンサーは並列して数百万以上の解読を行うので、短時間に大量のリードが得られる。次世代シーケンサーでの解読方法は、機種によって異なる。

Roche Diagnostics 社の genome sequencer FLX instrument は、パイロシーケンス法を用いる。パイロシーケンス法は、DNAの伸長反応を光で観察し、配列を解読する方法である。具体的には、dNTPが鋳型DNAと結合した時に生じたピロリン酸が、スルホリラーゼによりアデノシン-5'-ホスホ硫酸と反応し、生成したATPとルシフェラーゼの反応によるルシフェリンの発光を、CCDカメラで検出する。連続した塩基の数は、検出した光の強度で推定する(例えば、TTAAACGGという配列の場合、Tの発光量が2、Aの発光量が3、Cの発光量

が 1、G の発光量が 2 と順に検出される)。genome sequencer FLX instrument は、dTTP、dATP、dCTP、dGTP の順に 1 種類ずつ 200 サイクル繰り返し反応させ、塩基配列を解読していく。

Illumina 社の次世代シーケンサーは、SBS (Sequence-by-Synthesis)法を用いる。3'末端がブロックされ、蛍光色素が結合した 4 種類の塩基を順に 1 塩基ずつ鋳型 DNA に対して反応させ、蛍光が確認できた塩基に配列を決定するという方法である。

解読法は、DNA 断片を片側からのみ読むシングルリード法と両側から読むペアエンド法の 2 種類がある。ペアエンド法は解読量が増えるだけでなく、DNA 断片の両端の配列データが得られるため、短い反復配列が連続している領域でも DNA 断片の長さが情報に加わることで、より正確な解読が可能になる。

ゲノム解読が以前より短時間かつ安価で行えるようになることで、実験室で培養している株のゲノムを解読し、登録株と配列を比較することも可能になった(Kanesaki et al. 2012, Tajima et al. 2011)。より多くの種の全ゲノムが解読され、登録されることで、特定の遺伝子だけではなく、ゲノム全体の構造を比較することや、多くのタンパク質遺伝子を用いた系統樹の作成も可能になるだろう。本博士論文では 3 種類の生物のゲノムについて、次世代シーケンサーを用いて解読した。

1.10. 系統解析の方法

塩基配列やアミノ酸配列から系統樹を推定する方法は、数多く提唱されている(Yang 2006)。本博士論文はその中から近隣結合法、最尤法、ベイズ法を利用した。

近隣結合法は、Saitou & Nei (1987)によって提唱された最小進化基準に基づくクラスタ解析アルゴリズムである。この方法は計算が速く、多くの種を用いた系統解析が可能であることから、広く用いられてきた。この方法では、まず星状の系統樹から始め、樹長の合計を最も短くできる 2 つの節を選び、その節を結合するという手続きを繰り返すことで系統樹を作成する。

最尤法(Felsenstein 1981)は、与えられたデータの、推定した系統樹に対する適合度を置換速度のモデルに基づいて計測し、最大の対数尤度値を示すものを選ぶ。本博士論文で最尤法の推定に用いた TREEFINDER (Jobb et al. 2004)は、それ

ぞれの分岐の信頼度を LR-ELW (the Expected-Likelihood Weights applied to Local Rearrangements of tree topology)によって求めている。この方法は、まず、系統樹のある分岐の近傍にある枝の位置を取り替える。そうすることで、2分木する系統樹の場合、注目した分岐毎にそれぞれ元の系統樹と異なる樹形が2種類ずつ出来上がる。枝の長さのみ最尤法によりもう一度計算し直してから、元の樹形と新しくできた樹形のどちらがより正しいかを ELW (Strimmer & Rambaut 2002)により計算する。信頼度はパーセントで表す。これを全ての分岐について調べる。この方法は、ブートストラップ解析より速く信頼度を求めることができるという利点がある。

ベイズ法(Rannala & Yang 1996)は、データと尤度モデルと事前分布が与えられた時に、その系統樹が真である確率が最大のものを選ぶ。ベイズ法は、Metropolis-Hastings アルゴリズムを利用したマルコフ連鎖モンテカルロ法(MCMC)が用いられる(Metropolis et al. 1953, Hastings 1970)。Metropolis-Hastings アルゴリズムは、ある試行(ベイズ法では世代と言われる)の結果から、事前確率、尤度モデル、次の世代へ変わる確率をもとに新しい樹形を求め、それを採択するかどうかを確率で決定する。採択されれば新しい樹形をもとに次の試行を行い、棄却されれば元の樹形で次の試行を行う。このアルゴリズムを繰り返し、定常状態に達した頃の樹形が採用される。MCMC では、とり得る樹形のピークが複数あった時、1つのピークから他のピークに移るのは困難とされる。本博士論文でベイズ法の推定に用いた MrBayes (Huelsenbeck & Ronquist 2001, Ronquist & Huelsenbeck 2003)では、この問題点を改良した Metropolis 共役 MCMC (Metropolis-coupled MCMC; MCMCMC) (Geyer 1991)が用いられている。MCMCMC は複数のマルコフ連鎖を同時に走らせる。1つは低温連鎖というMCMCと同じ計算を行い、サンプルをとる連鎖で、残りは高温連鎖という低温連鎖より攪乱が起こりやすい連鎖を行う。各世代のあとにランダムに選択された2つの連鎖の状態を混合することで、特定のピークのみにとられることが改善される。

近隣結合法には、特に複数の遺伝子を用いた系統解析の場合、長枝誘引の影響を受けてしまいがちになるという問題が知られている(Qiu et al. 2012)。そのため、本博士論文では、タンパク質遺伝子のアミノ酸配列を結合した系統解析の場合、近隣結合法は使わず、最尤法とベイズ法を使用した。rRNA 遺伝子の塩基

配列の系統解析には、伝統的に用いられている近隣結合法と最尤法を用いた。

1.11. 葉緑体の起源となるシアノバクテリアについての議論

アーケプラスチダの葉緑体はどのシアノバクテリアが細胞内共生したものかは現在でも議論がなされている。タンパク質遺伝子のアミノ酸配列を利用した系統解析では、シアノバクテリアの根元の方で細胞内共生が起きたと報告されている(Crisuolo & Gribaldo 2011, Shih et al. 2013)。コードされたタンパク質遺伝子の類似性を比較すると、Nostcales が最も近縁だとする結果(Deusch et al. 2008)や Stigonematales が最も近縁だとする結果(Dagan et al. 2013)が報告されている。rRNA 遺伝子を用いた系統解析では、窒素固定を行う単細胞シアノバクテリアが最も近縁であるという報告がされた(Falcon et al. 2010)。本博士論文では 16S rRNA 遺伝子の塩基配列とタンパク質遺伝子のアミノ酸配列についてそれぞれ系統解析を行った。

1.12. 本博士論文の目的と構成

本博士論文では、今までゲノム解読が報告されてこなかった系統のゲノムを解読し、上位分類群と比較することで、ゲノム解読をした種だけではなく、その上位分類群のゲノム構造の特徴等について知見を得ることを目的とした。本博士論文では紅藻系統の進化や細胞内共生に焦点を絞り、紅藻 *Porphyridium purpureum*、パルマ藻 *Triparma laevis* のオルガネラゲノムおよびシアノバクテリア *Limnothrix* sp. ABRG5-3 のゲノムを解読した。それぞれの生物についてゲノム構造やコードされた遺伝子について調べた。ゲノム構造の比較は、葉緑体ゲノムでも広く保存されている rRNA オペロンとリボソームタンパク質遺伝子クラスターについて重点的に行った。第 2 章では、*P. purpureum* の葉緑体ゲノムの特徴を調べるために、紅藻をはじめとした真核光合成生物の葉緑体ゲノムと比較した。第 3 章では、*T. laevis* と珪藻の類似点と相違点を調べ、それがさらに上位の分類群と共通しているか珪藻とパルマ藻だけの特徴なのかを調べるために、不等毛植物門のオルガネラゲノムと比較した。第 4 章では、シアノバクテリアゲノムの系統解析結果をもとに、*Limnothrix* sp. ABRG5-3 とシアノバクテリアのゲノムを系統単位で比較した。第 5 章では、今まで解読した生物を含めた系統

樹の作成を、葉緑体ゲノムは保存されたタンパク質遺伝子アミノ酸配列と 16S rRNA 遺伝子塩基配列について、ミトコンドリアゲノムは保存されたタンパク質遺伝子アミノ酸配列について行った。最後に、総括として、アーケプラスチダの葉緑体の由来となるシアノバクテリアが細胞内共生に成功した理由を探るため、細胞内共生のもとになったと考えられるシアノバクテリアの系統とそれ以外の系統の比較を行った。

表 1-1：シアノバクテリアと真核光合成生物の葉緑体の比較

分類群	シアノバクテリア	紅色植物門 (紅藻)	灰色植物門	緑色植物界	クリプト植物門	不等毛植物門	ハプト植物門	渦鞭毛植物門	クロララクニオン植物門	ユーグレノゾア	
真核・原核の別	原核生物	真核生物									
所属するグループ	—	アーケープラスチダ	不明	不明	不明	ストラメノパイル	不明	アルベオラータ	リザリア	エクスカバータ	
想定される共生と共生体	一次共生の共生体	一次共生・シアノバクテリア	二次共生・紅色植物門								
主要アンテナ色素系	フィコビリンはクロロフィル <i>b</i>	フィコビリン	クロロフィル <i>b</i> および β カロテン	クロロフィル <i>c</i>							クロロフィル <i>b</i> および β カロテン
				フィコピリン	フコキサンチン	19'ヘキサノイルフコキサンチン	ペリデイン				
葉緑体包膜	細胞膜と外膜	2枚	4枚								
ヌクレオモルフ	—	なし	あり								
チラコイド	一重	一重	多重 (グラナ)	二重							三重
	α -1,4 グルカン (ラン藻デンプン)	α -1,4 グルカン (紅藻デンプン)	α -1,4 グルカン (デンプン)	α -1,4 グルカン (デンプン)	あり	β -1,3 グルカン	β -1,3 グルカン	α -1,4 グルカン	β -1,3 グルカン?	β -1,3 グルカン、パラミロン	
光合成貯蔵物質	—	なし	多重 (グラナ)	あり	なし	なし	なし	なし	あり	なし	

井上 (2007)の表より改変

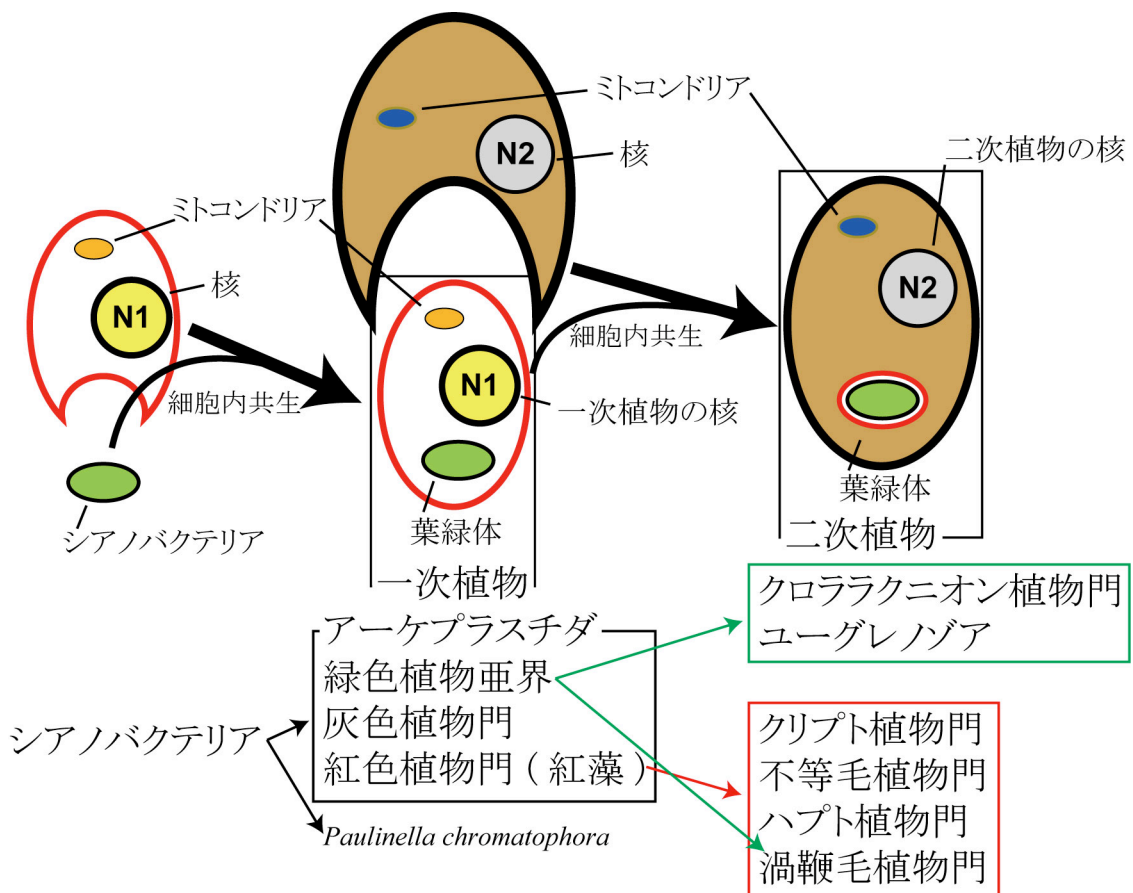


図 1-1 : 細胞内共生による葉緑体の獲得と真核光合成生物の分類

細胞内共生による葉緑体の獲得を模式図で示した。この模式図では一次植物が二次共生したものを二次植物として示したが、二次共生した藻類がさらに他の真核生物に細胞内共生(三次共生)したことに由来する葉緑体を持つ藻類も二次植物に含まれる。

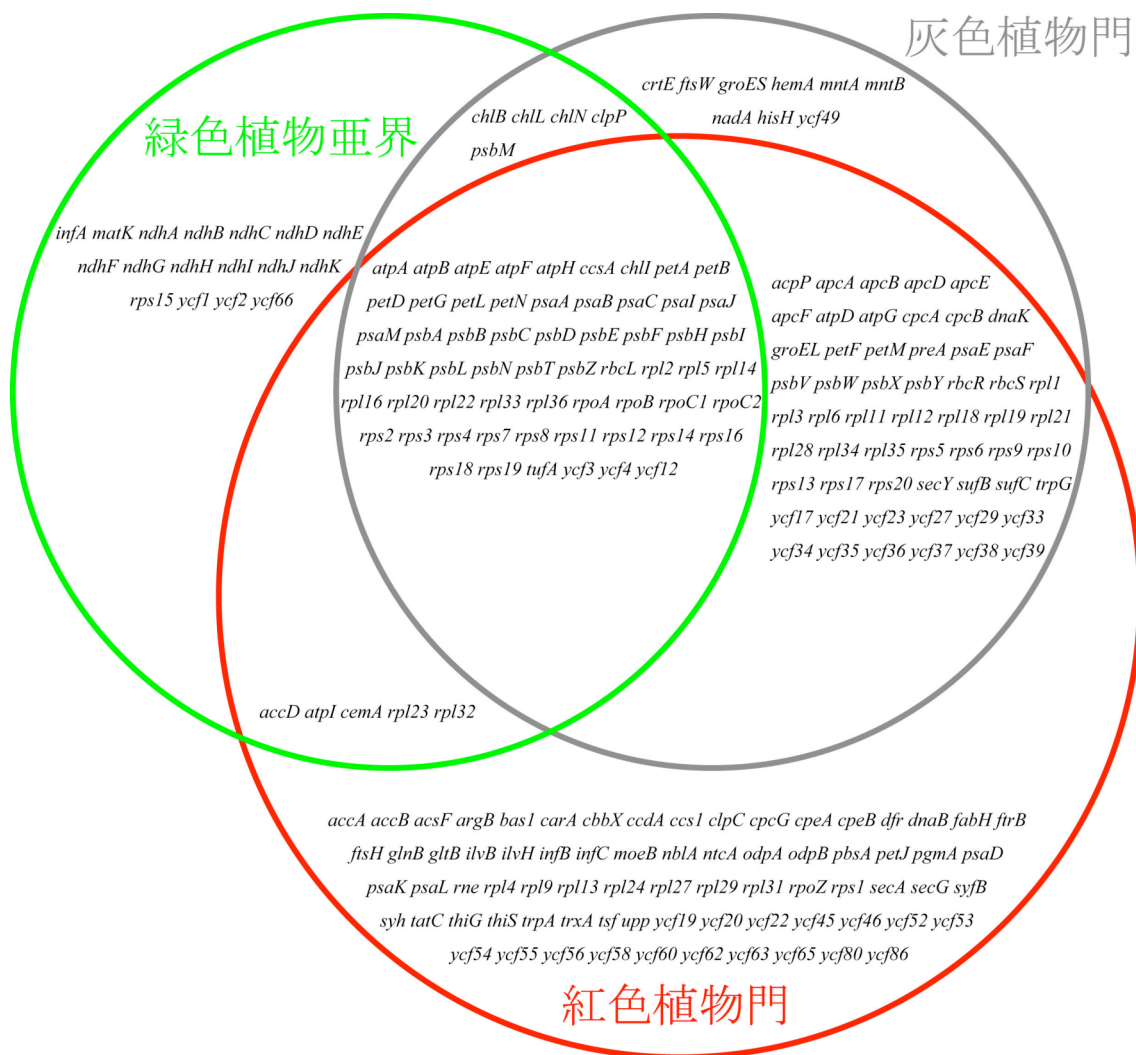


図 1-2 : アーケプラスチダの葉緑体ゲノムにコードされたタンパク質遺伝子の種類

5 種の紅色植物門(*Cyanidioschyzon merolae*, *Cyanidium caldarium*, *Gracilaria tenuistipitata* var. *liui*, *Porphyra purpurea*, *Porphyra yezoensis*)、1 種の灰色植物門 (*Cyanophora paradoxa*)、5 種の緑色植物亜界 (*Arabidopsis thaliana*, *Physcomitrella patens*, *Chaetosphaeridium globosum*, *Chlorella vulgaris* C-27, *Mesostigma viride*)の内それぞれ半数以上にコードされた hypothetical protein 以外のタンパク質遺伝子をベン図に表した。

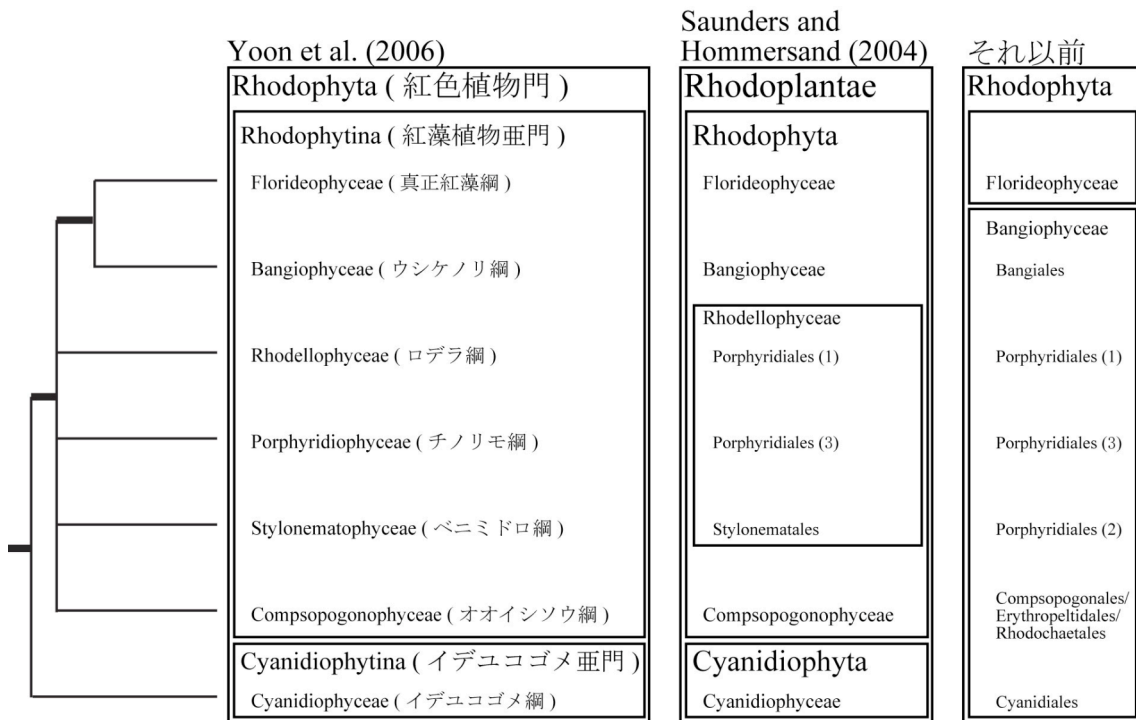


図 1-3 : 紅藻の系統分類の変遷

Yoon et al. (2006)より改変。系統関係が明らかになっている部分は太線で示した。

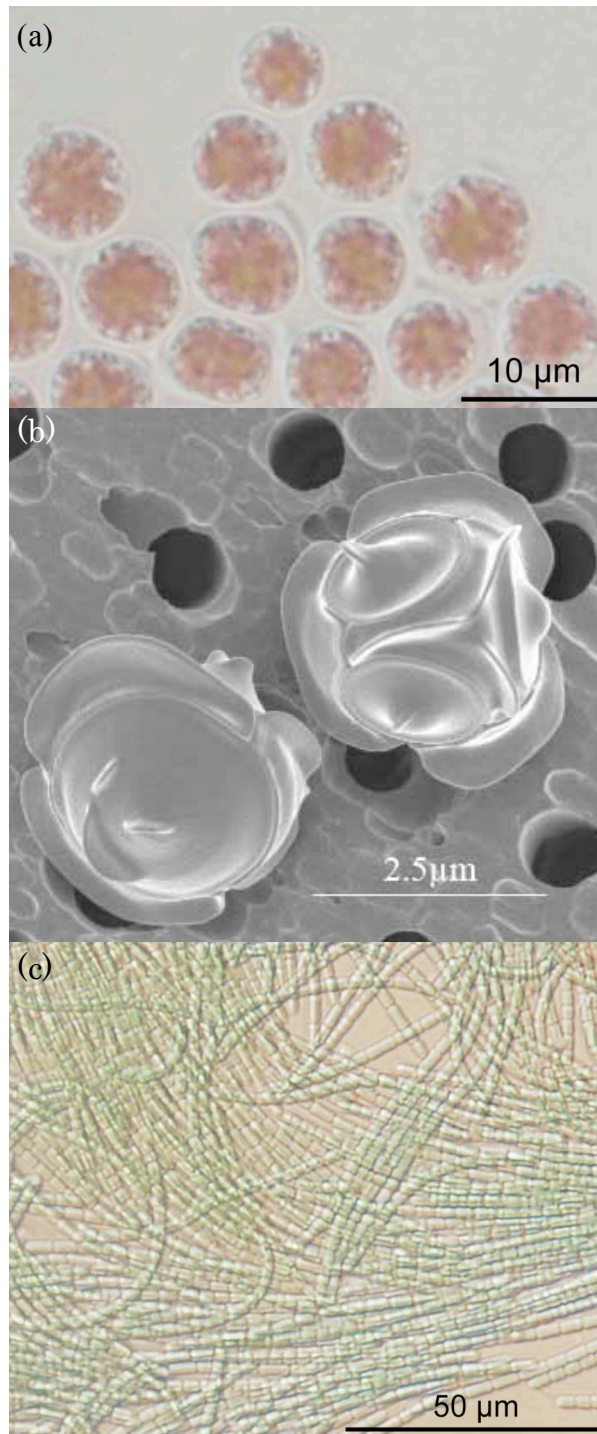


図 1-4 : ゲノムを解読した生物

(a) *Porphyridium purpureum* のノマルスキー微分干渉顕微鏡像

(b) *Triparma laevis* の走査電子顕微鏡像 (写真提供：桑田晃博士)

(c) *Limnothrix* sp. ABRG5-3 のノマルスキー微分干渉顕微鏡像

第2章

海洋性単細胞紅藻チノリモ *Porphyridium purpureum* の葉緑体ゲノム解読

2.1. 序文

紅藻は、2種類の亜門と7種類の綱に分けられる(Yoon et al. 2006)。イデユコゴメ亜門(Cyanidiophytina)はイデユコゴメ綱(Cyanidiophyceae)1綱のみに分けられ、紅藻植物亜門(Rhodophytina)は真正紅藻綱(Florideophyceae)、ウシケノリ綱(Bangiophyceae)、オオイシソウ綱(Compsopogonophyceae)、ベニミドロ綱(Stylonematophyceae)、ロデラ綱(Rhodellophyceae)、チノリモ綱(Porphyridiophyceae)の6綱に分けられる。これら7綱の内、葉緑体ゲノムの全長配列が解読されたものは3綱のみである;2種のイデユコゴメ綱(Glöckner et al. 2000, Ohta et al. 2003)、4種の真正紅藻綱(DePriest et al. 2013, Hagopian et al. 2004, Janouškovec et al. 2013)および3種のウシケノリ綱(Reith & Munholland 1995, Wang et al. 2013)。そのため、紅藻の系統関係は以上の3綱以外、未解明のままである。

最近、*P. purpureum* CCMP 1328 のドラフトゲノム配列が報告された(Bhattacharya et al. 2013)が、この種の葉緑体ゲノムの全塩基配列は報告されていない。

本博士論文ではCCMP 1328とは別株である*P. purpureum* NIES 2140の葉緑体ゲノムを解読した。この種は、葉緑体ゲノムが未解読であるチノリモ綱に属する海洋性単細胞紅藻である。この章では*P. purpureum*の葉緑体ゲノムの構造や、コードされた遺伝子の種類などについて述べた。

2.2. 材料と方法

2.2.1. 生物材料と培養

P. purpureum NIES 2140 は、国立環境研究所に保有されていた株を分けていただいたものである。改変 ESM 培地(通常の ESM 培地(Okaichi et al. 1982)に soil extract を入れず、海水の代わりに人工海水 Viesalt (MARINETECH 社)を 35 g/l 加えたもの)で培養を行った。前培養は、室温 22°C で、 $20 \mu\text{mol}/\text{m}^2 \cdot \text{s}$ の白色光を 12 時間明期 : 12 時間暗期の条件で当て、約一ヶ月間静置培養を行った。その後、DNA 精製を行うために、水温 30°C、横から $30 \mu\text{mol}/\text{m}^2 \cdot \text{s}$ の白色光が当たる水槽内に試料の入った試験管を置き、1%の CO₂ を含む空気を通気させ、約 6 日間培養を行った。

2.2.2. 試薬調整

- 1.0 M Tris-HCl: Tris 60.5 g を蒸留水に溶かし、HCl で pH を 8.0 に調整し、500 ml にした。作成後、オートクレーブで滅菌した。
- 0.5 M EDTA: 2NA(EDTA•2Na) 37.2 g を蒸留水に溶かし、NaOH で pH 8.0 に調整し、200 ml にした。作成後、オートクレーブで滅菌した。
- TE50/20: 5 ml の 1.0 M Tris-HCl (50 mM) と 4 ml の 0.5 M EDTA (20 mM) を混ぜ、蒸留水で 100 ml にした。作成後、オートクレーブで滅菌した。
- TE: 1 ml の 1.0 M Tris-HCl (10 mM) と 0.2 ml の 0.5 M EDTA (1 mM) を混ぜ、蒸留水で 100 ml にした。作成後、オートクレーブで滅菌した。

CTAB はヒートスターラーで温めながら混ぜた。作成後、冷凍保存した。

- 2×CTAB: 1.0 M Tris-HCl (pH 8.0) 10 ml (100 mM)、NaCl 8.18 g (1.4 M)、0.5 M EDTA 4 ml (20 mM)、CTAB 2.0 g を蒸留水に溶かし、100 ml にした。使用前直前に 2-メルカプトエタノールを 280 μl 加えた。
- 1×CTAB: 1.0 M Tris-HCl (pH 8.0) 5 ml (50 mM)、0.5 M EDTA 2 ml (10 mM)、CTAB 1.0 g を蒸留水に溶かし、100 ml にした。
- 10% CTAB: CTAB 10.0 g、NaCl 4.09 g (0.7 M) を蒸留水に溶かし、100 ml にした。
- クロロホルム液: クロロホルムとイソアミルアルコールを、体積比 24:1 で混合した。

- 20 mg/ml プロテイナーゼ K: 10 ml の蒸留水にプロテイナーゼ K 200 mg を溶かした。作成後、冷凍保存した。
- 20 mg/ml EtBr: 10 ml の蒸留水に EtBr 200 mg を溶かした。作成後、アルミホイルで遮光して、室温で保存した。
- 1 M CsCl/EtBr: CsCl 8.42 g、1.0 M Tris-HCl (pH 8.0) 2.5 ml (50 mM)、0.5 M EDTA 1 ml (10 mM)、20 mg/ml EtBr 0.5 ml を蒸留水に溶かし、50 ml にした。
- 7 M CsCl/EtBr: CsCl 58.9 g、N-ドデカノイルサルコシン酸ナトリウム 50 mg、20 mg/ml EtBr 0.5 ml を蒸留水に溶かし、50 ml にした。
- バランス用 CsCl/EtBr: TE 30 ml と CsCl 30 g と 20 mg/ml EtBr 0.5 ml を混合して CsCl を溶かした。
- 4 M NaCl 飽和 2-プロパノール: 4 M NaCl 水溶液と 2-プロパノールを等量加えてよく混ぜ、一晩静置した。上層の、水分が飽和した 2-プロパノール層を使用した。

2.2.3. DNA 精製

培養した細胞は、培養液ごと 50 ml チューブに注ぎ、30°C、3,000 g、10 分間の遠心を行った。上清の液は捨て、新たに培養液を注ぎ、同様の遠心を行った。この作業を培養液が無くなるまで繰り返し、細胞を回収した。途中、試験管にこびりついた細胞をブラシでこそげ落とし、それも回収した。沈殿を TE50/20 に懸濁し、軽く攪拌した。400 μ l ずつ計 12 本の 1.5 ml チューブに分けた。1 本あたり約 40 mg ずつガラスビーズ(150-212 microns (SIGMA))を加え、軽く混ぜ、30 分間強く攪拌した。20 mg/ml プロテイナーゼ K を 10 μ l ずつ加え、30 分間攪拌した。細胞が破碎されていることを顕微鏡観察で確認した後、さらに 15 分間攪拌した。液を 1 本の沈殿管に集め、液量を量り(3.6 g)、液量とほぼ同じ量の 2 \times CTAB (3.5 ml)を加え、50°C で 1 時間インキュベートした。液が濃かったため、TE50/20 と 2 \times CTAB を 1.5 ml ずつ加え、さらに 20 mg/ml プロテイナーゼ K を 50 μ l 加え、50°C で 1 時間インキュベートした。20 mg/ml プロテイナーゼ K を 110 μ l 加え、50°C で一晩インキュベートした。クロロホルム液を等量(10 ml)加え、30 分間よく混ぜてから 20°C、8,400 g、10 分間遠心した。上層を新しい沈殿管に移し、残った方をもう一度 20°C、8,400 g、5 分間遠心し、上層は先ほどの新しい沈殿管に移した。合わせた上層の液量の

約 1/10 の 10% CTAB (0.6 ml) とクロロホルム液を 7 ml 加え、30 分間よく混ぜた。20°C、8,400 g、10 分間遠心した。上層はシリコンコーティングしたガラスチューブに移し、等量の 1×CTAB (7 ml) を加えた。軽く混ぜてから、1 時間室温で放置した。20°C、3,000 g、20 分間遠心した。上清は取り除き、ガラスチューブをペーパータオルの上で逆立ちにし、10 分間沈殿を乾燥させた。1 M CsCl/EtBr を 3.5 ml 加え、遮光して 30 分間 50°C で温めながら沈殿を溶かした。7 M CsCl/EtBr を 4.5 ml 加え、遮光して 10 分間 50°C で温めた。超遠心用チューブに移し、20°C、352,000 g、38 時間超遠心した。DNA のバンド部分を新しい超遠心用チューブに移し、バランス用 CsCl/EtBr 溶液で少し薄め、20°C、315,000 g、20 時間超遠心した。DNA バンドを 1.5 ml チューブに回収し、4 M NaCl 飽和 2-プロパノールを等量加え、よく転倒攪拌した。上層の 2-プロパノール層を捨て、下層に新しく 4 M NaCl 飽和 2-プロパノールを等量加え、よく転倒攪拌した。同様のことを何度か繰り返し、最後に 100% 2-プロパノールで同様のことを行った。この作業により EtBr を除いた。事前に 1 mM EDTA で 40 分間程度煮沸し、オートクレーブで滅菌したセロハンチューブ (Spectra/Por® molecularporous membrane tubing (MWCO:12-14,000)) に回収した液を入れ、チューブの両端を Spectra/Por® CROSURES で挟み、液漏れが無いことを確認した後、TE で透析した。最初は 500 ml のビーカーに 100 ml の TE で 1 時間透析し、それ以降は 400 ml の TE で透析した。途中で液をかえながら、合計で 24 時間以上透析したものを回収した。濃度は 30.8 ng/μl だった。サンプルの濃度を DNA 配列解析に用いるために十分な濃度になるまで濃くするために、真空乾燥機で乾燥した。最終濃度は 92.6 ng/μl になった。

2.2.4. DNA 配列解析

得られたゲノム DNA のうち、5.5 μg を DNA 配列の解析に利用した。ゲノムの塩基配列解析には Roche Diagnostics 社の genome sequencer FLX instrument を利用した。得られたシーケンスデータは Newbler Matrics ver. 2.5.3 によってアセンブルされ、コンティグおよびスキップフォルドにまとめられた。得られたコンティグは 3,355 個(平均長=5,775 bp)だった。peak depth は 27.0 で、総ゲノムの長さは 19.4 Mbp だった。コンティグごとに blastn (Altschul et al. 1997) の解析を行い、それぞれ核ゲノムの塩基配列であるかオルガネラゲノム

の塩基配列であるかを振り分けた。葉緑体のゲノムに相当するコンティグは 13 個だった。アセンブル解析の結果から、コンティグ間やスキップフォールドのつながりも同時に推定された。推定された隣り合うコンティグの端からプライマーを作り、PCR 法でコンティグ間のギャップとなる部分の配列を増やした。増やした配列は、サンガー法により塩基配列を決定し、コンティグ間のギャップを埋めた。サンガー法による塩基配列解析は株式会社ファスマックに依頼した。

この葉緑体ゲノムには、反復配列様の領域が存在していることが分かった。反復配列様領域の外側から PCR でそれぞれの配列を増やし、両者の違いおよび反復配列様領域外とのつながりを確認した。解読した葉緑体ゲノムの全長配列は GenBank に登録した(accession number: AP012987)。

2.2.5. ゲノムアノテーションとデータ解析

タンパク質遺伝子を推定するために、まず MetaGeneAnnotator (Noguchi et al. 2008)を利用した。それから、*P. purpureum* の葉緑体ゲノムを参照配列として、Gclust (Sato 2009)を基に作られた CyanoClust (Sasaki & Sato 2010)のデータベースに登録されているタンパク質配列および 7 種の不等毛植物門(*Ectocarpus siliculosus*, *Fistulifera* sp. JPCC DA0580, *Fucus vesiculosus*, *Heterosigma akashiwo*, *Synedra acus*, *Thalassiosira oceanica*, *Vaucheria litorea*)のタンパク質配列を用いて、tblastn の解析を行った。また、tblastn の結果からタンパク質遺伝子のアノテーションをつけた。最後に、アミノ酸配列長が 30 を超え、ATG もしくは GTG を開始コドンとする領域をタンパク質遺伝子として推定した。タンパク質遺伝子の相同クラスタリング解析には CyanoClust に登録されているシアノバクテリアゲノムのタンパク質配列と葉緑体ゲノムのタンパク質配列および 7 種の不等毛植物門のタンパク質配列を利用した。

RNA 遺伝子の推定には *P. purpureum* の葉緑体ゲノムを参照配列として、紅藻と灰色植物門の葉緑体ゲノムにある RNA 遺伝子を用いた blastn の解析を行った。tRNA 遺伝子の検索には tRNAscan-SE (Lowe & Eddy 1997)を利用した。アンチコドンが CAU の tRNA (tRNA (CAU))は 3 種あるが、それらの同定方法は 2.3.4. で詳しく述べる。

イントロンの推定には RNAweasel (Lang et al. 2007)、近縁種の EST データ (Chan et al. 2011)、アラインメント解析を利用した。

葉緑体ゲノムの rRNA オペロン構造の比較には GOBASE データベース (O'Brien et al. 2009) に登録された種の葉緑体ゲノムおよび先程述べた 7 種の不等毛植物門、ポーリネラ (*Paulinella chromatophora*)、4 種の紅藻 (*Pyropia haitanensis*, *Calliarthron tuberculosum*, *Chondrus crispus*, *Grateloupia taiwanensis*)、2 種のハプト藻 (*Pavlova lutheri*, *Phaeocystis antarctica*)、1 種のクロララクニオン藻 (*Bigelowiella natans*) の葉緑体ゲノムを利用した。

3 種の tRNA(CAU) のアラインメント、rRNA オペロン構造の比較およびリボソームタンパク質遺伝子クラスターの比較解析には第 3 章で述べるパルマ藻の葉緑体ゲノムのデータも加えている。

DNA およびタンパク質配列の加工には SISEQ ver. 1.59 (Sato 2000) を利用した。tRNA(CAU) のアラインメント作成は Clustal X ver. 2.0.9 (Larkin et al. 2007) を利用し、その他のアラインメント作成は ClustalW/ClustalX ver. 1.8.3 (Thompson et al. 1994)、もしくは Muscle ver. 3.6 (Edgar 2004) を利用した。ゲノム配列の編集は Artemis ver. 13.0 (Rutherford et al. 2000) を利用した。

2.2.6. *P. purpureum* が持つイントロンの系統解析

P. purpureum のマチュラーゼを系統解析するために、相同タンパク質配列を Gclust のデータセット “CyanoClust4” から手に入れた (Cluster No. 5 および Cluster No. 4454)。その中で、アミノ酸配列長が 350 を超えるアミノ酸配列を実際の系統解析に利用した。アラインメント作成は、Muscle ver. 3.6 を利用し、LG モデルを用いた最尤法で系統解析を行った。最尤法の計算には TREEFINDER ver. March 2011 (Jobb et al. 2004) を利用した。それぞれの分岐の信頼度は LR-ELW によって求めた。

P. purpureum にある全てのグループ II イントロンを、GTR+G+I モデルを用いた最尤法で系統解析を行った。最尤法の計算には TREEFINDER ver. June 2008 (Jobb et al. 2004) を利用した。

2.3. 結果

2.3.1. *P. purpureum* 葉緑体ゲノムの構造

P. purpureum NIES 2140 の葉緑体ゲノムを Roche Diagnostics 社の genome sequencer FLX instrument を用いて解読した。葉緑体ゲノムは 13 個のコンティグにアセンブルされ、ギャップはサンガー法を用いて解読した。*P. purpureum* NIES 2140 の葉緑体ゲノムは全長 217,694 bp、GC 含量 30.3%の環状 DNA だった。これは他の紅藻の葉緑体ゲノムと比べると大きい値ではある(表 2-1)が、*P. purpureum* の葉緑体ゲノムが大量のイントロン(計 30,433 bp)を含むためであると考えられる。*P. purpureum* の葉緑体ゲノムは、1 個の ncRNA 遺伝子(*rnpB*)、6 個の rRNA 遺伝子(2 個の *rrs*、2 個の *rrl*、2 個の *rrf*)、29 個の tRNA 遺伝子および 224 個のタンパク質遺伝子をコードすることが分かった(表 2-2、図 2-1)。

一般に、葉緑体ゲノムは rRNA オペロンを含んだ逆向き反復配列(IR)領域を持っている。多くの藻類の葉緑体ゲノムは、rRNA オペロンに 3 個の rRNA 遺伝子(*rrs*, *rml*, *rrf*)と 2 個の tRNA 遺伝子($\text{tRNA}^{\text{Ile}}(\text{GAU})$, $\text{tRNA}^{\text{Ala}}(\text{UGC})$)を含んでいる(図 2-2)。 $\text{tRNA}^{\text{Ile}}(\text{GAU})$ 遺伝子と $\text{tRNA}^{\text{Ala}}(\text{UGC})$ 遺伝子は、*rrs* 遺伝子(16S rRNA 遺伝子をコードする)と *rml* 遺伝子(23S rRNA 遺伝子をコードする)の間に位置する。ウシケノリ綱の葉緑体ゲノムは、直列反復配列で 2 セットの rRNA オペロンを持ち、イデユコゴメ綱と真正紅藻綱の葉緑体ゲノムは単一の rRNA オペロンを持つ。

P. purpureum の葉緑体ゲノムの rRNA 遺伝子は逆向き反復配列の構造をしている。しかし、その内の一つ(IR1)は *rrs* と *rml* の間に $\text{tRNA}^{\text{Ala}}(\text{UGC})$ を持ち、もう片方(IR2)は $\text{tRNA}^{\text{Ile}}(\text{GAU})$ を持っている(図 2-1、図 2-2)。この 2 種類の rRNA オペロン構造は、反復配列様領域の外側から PCR でそれぞれの配列を増やし、確認した。EST データ(Chan et al. 2011)によると近縁種(CCMP 1328 株)も同様の rRNA オペロン構造を持つことが分かった。この非対称の rRNA オペロン構造は全ての解読された葉緑体ゲノムのうち、チノリモ綱に特有のものであることが推測された。

2.3.2. 系統間で保存されているタンパク質遺伝子の比較

P. purpureum の葉緑体ゲノムにコードされている全てのタンパク質遺伝子について、アミノ酸配列による相同クラスタリング解析を行った。*P. purpureum* を含めた 6 種の紅藻、13 種のクロムアルベオラータ、1 種の灰色植物門、40 種の緑色植物界の葉緑体ゲノムにコードされたタンパク質および 41 種のシアノバクテリアゲノムにコードされたタンパク質を解析に用いた(図 2-3)。ベン図内のタンパク質遺伝子は、それぞれの系統毎で半数以上の生物種に保存されているものを示している。イデユコゴメ亜門ではコードされていないが、紅藻植物界で保存されているタンパク質遺伝子を、紅色植物門の円より一回り大きい紅藻植物界の円と紅色植物門の円の間を示した。37 個のタンパク質遺伝子は他に相同性を持つものが検出されなかった。マチュラーゼ遺伝子とアノテーションされた 5 個の遺伝子(*mat1a*, *mat1b*, *mat1c*, *mat1d*, *mat1e*)は、クリプト藻 *Rhodomonas salina* といくつかのシアノバクテリアのリバーストランスクリプターゼとホモログだった。マチュラーゼタンパク質の系統解析については、2.3.3. でより詳しく述べる。これらの遺伝子と *ycf12*, *ycf80* 以外の全てのタンパク質遺伝子は、シアノバクテリアでも保存されていた。RNA ポリメラーゼ、リボソームタンパク質、シトクロム複合体、ATP 合成酵素、光化学系 I, II に関わる 57 個の遺伝子は、全ての光合成生物に保存されていることが分かった。アロフィコシアニン(*apcA*, *apcB*, *apcD*)、フィコシアニン(*cpcA*, *cpcB*)、フィコビリソーム(*apcE*, *apcF*)のサブユニットに関わる遺伝子は紅色植物門と灰色植物門で保存されており、クロムアルベオラータには保存されていなかった。リボヌクレアーゼ(*rne*, *ycf56*)やフィコエリスリン関連タンパク質(*cpeA*, *cpeB*)、tRNA 合成酵素(*syfB*, *syh*)の遺伝子は紅藻植物界に保存されていた。マンガン ABC トランスポーターの遺伝子(*mntA*, *mntB*)は灰色植物門と *P. purpureum* に保存されていた。

藻類の進化の過程で、葉緑体ゲノムにコードされたタンパク質遺伝子は、徐々に葉緑体ゲノムから失われていくという考え(Glöckner et al. 2000)をもとに、紅藻の葉緑体ゲノムにコードされたタンパク質遺伝子の種類を比べ、紅藻の進化の過程で葉緑体ゲノムから失われたと考えられるタンパク質遺伝子を調べた(図 2-4)。19 個の遺伝子はイデユコゴメ綱のみに保存されていた。

プロトクロロフィリドレダクターゼの遺伝子(*chlB*, *chlL*, *chlN*)はウシケノリ綱にのみ保存されていた。3-イソプロピルマレートデヒドラターゼサブユニットの遺伝子(*leuC*, *leuD*)は、*Gracilaria tenuistipitata* の葉緑体ゲノムのみが、遺伝子の水平伝播により手に入れたと考えられている(Janouškovec et al. 2013)。アセチル CoA カルボキシラーゼ(ACCase) (*accA*, *accB*, *accD*)、アセチルグルタミン酸キナーゼ(*argB*)、*ycf29* は *P. purpureum* 以外の紅藻に保存されていた。*P. purpureum* は、イネ科(Konishi et al. 1996)やクロムアルベオラータ(Livne & Sukenik 1990, Roessler & Ohlrogge 1993)のように、核にコードされる真核型 ACCase のホモログのみ核ゲノムに持ち、多くの植物や他の紅藻が葉緑体ゲノムに持つ原核型 ACCase を失っていることが分かった。

2.3.3. イントロンとマチュラーゼ遺伝子

RNAweasel (Lang et al. 2007)、近縁種の EST データ(Chan et al. 2011)との比較、アラインメント解析を利用し、イントロンの推定を行った(表 2-3)。224 個のタンパク質遺伝子の内、29 個の遺伝子が 1 つまたは複数のイントロンを含んでいることが分かった。*rpl28* に含まれるイントロンはシュードノット構造(グループ I イントロンでよく見られる二次構造(Cech 1988, Lang et al. 2007))を持つことから、グループ I イントロンであると推定した(図 2-5)。それ以外のイントロンはグループ II イントロンだった。*dnaK*, *infC*, *gltB*, *rpoC2* の遺伝子内にあるイントロンは、マチュラーゼタンパク質遺伝子を含んでいた。*P. purpureum* の葉緑体ゲノムには 5 個のマチュラーゼタンパク質遺伝子がある(図 2-6)。*mat1a* は *atpE* と *ORF44* の間に位置していた。全ての推定マチュラーゼタンパク質は RT ドメインと X ドメインを持っていた。*Mat1d* のみがマチュラーゼの移動性に関わる Zn ドメイン(ヌクレアーゼドメイン)を持っていた。*Mat1a* と *Mat1b* は YxDD モチーフ(5 ドメイン内にあるポリメラーゼ活性部位(Zimmerly et al. 2001))を失っていた。*dnaK* 内のイントロン(ここでは *int.b* (*dnaK*)と書く)は *mat1b* を含んでいるが、RNAweasel では検出されなかった。ほぼ全ての陸上植物の葉緑体ゲノムでは、*trnK* 遺伝子内にマチュラーゼタンパク質遺伝子(*matK*)が含まれている(Neuhaus & Link 1987)が、*P. purpureum* の葉緑体ゲノムにあるマチュラーゼタンパク質は *MatK* のホモログではなかった。

シアノバクテリア等に含まれるマチュラーゼタンパク質遺伝子の系統解析によると、*P. purpureum* には2種類のマチュラーゼタンパク質がコードされていることが分かった(図 2-7)。1 つは Mat1a, Mat1c, Mat1d, Mat1e が含まれていた。Mat1a は *Rhodomonas salina* の葉緑体ゲノムが持つマチュラーゼタンパク質に近縁だった。Mat1b はこれらとは違う系統のものとなった。イントロン配列の系統解析(図 2-8)によると、int (*mntA*), int.b (*rpoC2*), int.b (*atpI*)は移動性のマチュラーゼを持つイントロン(int (*gltB*))に近縁であった。他のイントロン配列はそれほど似てはいなかった。イントロン配列の合計塩基対数は 30,433 bp であり、葉緑体ゲノム全体の約 14%におよぶことが分かった。

2.3.4. アンチコドンが CAU である tRNA 遺伝子の同定

多くの場合、反復領域にある場合を除いて、1 種類のアンチコドンを示す tRNA 遺伝子は 1 種類である。しかし、アンチコドンが CAU である tRNA 遺伝子は tRNA^{Met}(CAU), tRNA^{fMet}(CAU), tRNA^{Ile}(CAU)の3種類が存在している。*P. purpureum* の葉緑体ゲノムにもこれら3種類の tRNA 遺伝子がコードされている(表 2-4)。

tRNA^{fMet}(CAU)は、開始メチオニンに対応した tRNA 遺伝子である。tRNA^{fMet}(CAU)の tRNA 遺伝子は、シアノバクテリアや葉緑体のゲノムでは、D アーム内の 11 番目と 25 番目の AU 塩基対と 12 番目と 24 番目の GC 塩基対およびアンチコドンアーム内の 2 対か 3 対の GC 塩基対が保存されている(Alkatib et al. 2012)。126,118-126,045 にコードされた tRNA(CAU)はそれらの特徴が保存された塩基配列だったので、tRNA^{fMet}(CAU)であると同定した。

tRNA^{Ile}(CAU)は、最初のアンチコドンであるシチジンが TilS (tRNA^{Ile}-リシジン合成酵素)によってリシジンに変換された tRNA 遺伝子である(Soma et al. 2003)。このタンパク質をコードする *ycf62* は、紅藻の葉緑体ゲノムに広く保存されている。tRNA^{Ile}(CAU)は、この変換によって認識するコドンが AUG から AUA へと変わり、メチオニンではなくイソロイシンに対応した tRNA 遺伝子となる。TilS は tRNA^{Ile}(CAU)遺伝子の 33 番目の C と 39 番目の A を認識する(Nakanishi et al. 2009)。しかし、残りの 2 つの tRNA(CAU)はどちらも 33 番目が C で 39 番目は A だった。ハウレンソウの葉緑体ゲノムにコードされた tRNA^{Ile}(CAU)遺伝子の塩基配列は実験的に確かめられている(Francis &

Dudock 1982)。ハウレンソウの配列を基にした tRNA(CAU)のアラインメント解析の結果から 114,922-114,837 にコードされた方が tRNA^{Ile}(CAU)で、19,649-19,577 にコードされた方が tRNA^{Met}(CAU)であることが分かった(図 2-9)。また、いくつかの他の紅藻や *C. paradoxa* の葉緑体ゲノムの tRNA(CAU)は、tRNA^{Met}(CAU)と tRNA^{Ile}(CAU)のアノテーションが逆になっていることが分かった。

2.3.5. 葉緑体ゲノムのリボソームタンパク質遺伝子クラスターの解析

P. purpureum 葉緑体ゲノムのリボソームタンパク質遺伝子クラスターの構造を *Synechocystis* sp. PCC 6803 と *Anabaena* sp. PCC 7120 のメインゲノム、アーケプラスチダやクロムアルベオラータ、ユーグレノゾア、クロララクニオン植物門の葉緑体ゲノムと比較した(図 2-10)。褐藻以外のクロムアルベオラータや紅藻は、*dnaK* が、他の遺伝子と比べると逆向きであるが *rpl3* の上流に位置していた。ハプト藻の葉緑体ゲノムは、リボソームタンパク質遺伝子クラスター内のいくつかの遺伝子(*rpl4*, *rpl29*, *rpl24*, *rpl18*, *rpl13*)が失われていた。珪藻の葉緑体ゲノムのみがコードしている *ycf88* は、*rps19* と *rpl22* の間に位置していた。

リボソームタンパク質遺伝子クラスター内の遺伝子の内容はこのような違いがあるが、遺伝子の並びは紅藻とクロムアルベオラータの間で強く保存されていた。*rps12-rps7* の部分が *rpl36-rpl31* の部分の下流に位置するのは、灰色植物門やシアノバクテリアでは見られない、紅藻特異的な転座の結果である(Ohta et al. 1997)と考えられている。

P. purpureum では、リボソームタンパク質遺伝子クラスターは 4 つに分けられる。その箇所は図 2-10 の矢尻で示す通り、*rps5* と *secY* の間、*secY* と *rpl36* の間、*rps7* と *tufA* の間である。*secY* はクラスターから離れた位置にあり、*ORF57c* と *rpoC2* の間に位置する。*secY*, *rps10*, *tufA* は他の遺伝子達とは逆方向であった。*C. paradoxa* も、*secY* と *rpl36* の間が途切れている(Stirewalt et al. 1995)が、シアノバクテリアは途切れていない。

2.4. 考察

P. purpureum の葉緑体ゲノムは、多くのイントロンを持っている、2つのrRNA オペロンが同一ではない、リボソームタンパク質遺伝子クラスターの構造が他の紅藻やクロムアルベオラータとは大きく異なる、という他の種では見られない独特な特徴を多く持つことが今回のゲノム解読で明らかになった。

いくつかのイントロンが紅藻の葉緑体ゲノムにも存在することは既に報告されている(Bernard et al. 1992, Janouškovec et al. 2013, Yoon et al. 2006)が、1種の葉緑体ゲノムに存在するイントロンは1個か2個のみである。真正紅藻綱は tRNA(CAU)内に逆転写酵素を含んだイントロンを持っている(Janouškovec et al. 2013)。ロデラ綱、チノリモ綱、ベニミドロ綱のいくつかの種は *psaA* 遺伝子内の決まった位置にイントロンを持っている(Yoon et al. 2006)。*P. purpureum* も *psaA* 遺伝子内にイントロンを持っているが、イントロンの挿入位置は他と異なっていた。

P. purpureum の葉緑体ゲノムにコードされた5個のマチュラーゼタンパク質遺伝子のアミノ酸配列を調べたところ、Mat1dのみがマチュラーゼの移動能力に関わるヌクレアーゼドメインを持っていた(図2-6)。マチュラーゼタンパク質の系統解析の結果(図2-7)から、Mat1a, Mat1c, Mat1d, Mat1eは同じ系統であることが示唆された。これらのことから、*P. purpureum* の葉緑体ゲノム上では、Mat1d型のマチュラーゼを含むイントロンが、いくつかのコピーを別の場所に挿入しながら移動し続けているのではないかと考えられた。Mat1a, Mat1c, Mat1eは、配列に変異が起こったことで移動能力を失い、その場にとどまることになったと考えた。また、マチュラーゼを含むイントロンが *P. purpureum* の葉緑体ゲノム上での移動を繰り返す過程で、ゲノムの再編成が起りやすくなり、現在の独特な特徴を多く持つ葉緑体ゲノムになったのではないかと考えた。

表2-1：現在までに報告されている紅藻の葉緑体ゲノム

種名	アクセッション 番号	タンパク質 遺伝子数	イントロン 数	総塩基対数 (bp)	GC 含量 (%)	Reference
<i>Porphyridium purpureum</i> (チノリモ綱)	AP012987	224	43	217,694	30.3	本博士論文
<i>Calliarthron tuberculosum</i> (真正紅藻綱)	KC153978	201	2	178,981	29.2	Janouškovec et al. 2013
<i>Chondrus crispus</i> (真正紅藻綱)	HF562234	204	1	180,086	28.7	Janouškovec et al. 2013
<i>Gracilaria tenuistipitata</i> var. <i>liui</i> (真正紅藻綱)	AY673996	204	1	183,885	29.2	Hagopian et al. 2004
<i>Grateloupia taiwanensis</i> (真正紅藻綱)	KC894740	234	0	191,270	30.6	DePriest et al. 2013
<i>Porphyra purpurea</i> (ウシケノリ綱)	U38804	209	0	191,028	33.0	Reith & Munholland 1995
<i>Porphyra yezoensis</i> (ウシケノリ綱)	NC007932	209	0	191,954	33.1	-
<i>Pyropia haitanensis</i> (ウシケノリ綱)	KC464603	211	0	195,597	33.8	Wang et al. 2013
<i>Cyanidioschyzon merolae</i> (イデユコゴメ綱)	NC004799	207	0	149,987	37.6	Ohta et al. 2003
<i>Cyanidium caldarium</i> (イデユコゴメ綱)	AF022186	199	0	164,923	32.7	Glöckner et al. 2000

表2-2 : *P. purpureum*葉緑体ゲノムにコードされたタンパク質遺伝子の機能分類

分類	遺伝子数	遺伝子名				
Transcription & replication						
Helicase	1	<i>dnaB</i>				
Ribonuclease	2	<i>rne</i>	<i>ycf56</i> (RNase_Z)			
RNA polymerase	5	<i>rpoA</i>	<i>rpoB</i>	<i>rpoC1</i>	<i>rpoC2</i>	<i>rpoZ</i>
Transcription	3	<i>ntcA</i>	<i>rbcR</i>	<i>ycf27</i>		
Maturase	5	<i>mat1a</i>	<i>mat1b</i>	<i>mat1c</i>	<i>mat1d</i>	<i>mat1e</i>
Chaperone	5	<i>clpC</i>	<i>dnaK</i>	<i>groEL</i>	<i>ycf39</i>	<i>ycf45</i>
Protease	1	<i>ftsH</i>				
Translation factors						
Translation	4	<i>infB</i>	<i>infC</i>	<i>tsf</i>	<i>tufA</i>	
Ribosomal protein	28	<i>rpl1</i>	<i>rpl2</i>	<i>rpl3</i>	<i>rpl4</i>	<i>rpl5</i> <i>rpl6</i>
		<i>rpl9</i>	<i>rpl11</i>	<i>rpl12</i>	<i>rpl13</i>	<i>rpl14</i> <i>rpl16</i>
		<i>rpl18</i>	<i>rpl19</i>	<i>rpl20</i>	<i>rpl21</i>	<i>rpl22</i> <i>rpl23</i>
		<i>rpl24</i>	<i>rpl27</i>	<i>rpl28</i>	<i>rpl29</i>	<i>rpl31</i> <i>rpl32</i>
		<i>rpl33</i>	<i>rpl34</i>	<i>rpl35</i>	<i>rpl36</i>	
	20	<i>rps1</i>	<i>rps2</i>	<i>rps3</i>	<i>rps4</i>	<i>rps5</i> <i>rps6</i>
		<i>rps7</i>	<i>rps8</i>	<i>rps9</i>	<i>rps10</i>	<i>rps11</i> <i>rps12</i>
		<i>rps13</i>	<i>rps14</i>	<i>rps16</i>	<i>rps17</i>	<i>rps18</i> <i>rps19</i>
		<i>rps20</i>	<i>ycf65</i> (PSRP-3)			
Photosynthesis						
Phycobilisome	11	<i>apcA</i>	<i>apcB</i>	<i>apcD</i>	<i>apcE</i>	<i>apcF</i> <i>cpcA</i>
		<i>cpcB</i>	<i>cpcG</i>	<i>cpeA</i>	<i>cpeB</i>	<i>pbsA</i>
Cytochrome complex	11	<i>ccdA</i>	<i>ccsA</i>	<i>ccs1</i>	<i>petA</i>	<i>petB</i> <i>petD</i>
		<i>petG</i>	<i>petJ</i>	<i>petL</i>	<i>petM</i>	<i>petN</i>
Carbon fixation	3	<i>cbbX</i>	<i>rbcL</i>	<i>rbcS</i>		
Redox system	4	<i>frbB</i>	<i>ORF107</i>	<i>petF</i>	<i>trxA</i>	
Chlorophyll biosynthesis	3	<i>acsF</i>	<i>chlI</i>	<i>ycf53</i>		
ATP synthase	8	<i>atpA</i>	<i>atpB</i>	<i>atpD</i>	<i>atpE</i>	<i>atpF</i> <i>atpG</i>
		<i>atpH</i>	<i>atpI</i>			
Photosystem						
Photosystem I	13	<i>psaA</i>	<i>psaB</i>	<i>psaC</i>	<i>psaD</i>	<i>psaE</i> <i>psaF</i>
		<i>psaI</i>	<i>psaJ</i>	<i>psaK</i>	<i>psaL</i>	<i>psaM</i> <i>ycf3</i>
		<i>ycf4</i>				
Photosystem II	19	<i>psbA</i>	<i>psbB</i>	<i>psbC</i>	<i>psbD</i>	<i>psbE</i> <i>psbF</i>
		<i>psbH</i>	<i>psbI</i>	<i>psbJ</i>	<i>psbK</i>	<i>psbL</i> <i>psbN</i>
		<i>psbT</i>	<i>psbV</i>	<i>psbW</i>	<i>psbX</i>	<i>psbY</i> <i>psbZ</i>
		<i>ycf12</i>				
Metabolism						
Pyruvate dehydrogenase	2	<i>odpA</i>	<i>odpB</i>			
Lipids	2	<i>acpP</i>	<i>fabH</i>			
Nucleotides	1	<i>carA</i>				
Amino acid synthesis	5	<i>glbB</i>	<i>ilvB</i>	<i>ilvH</i>	<i>trpA</i>	<i>trpG</i>
Cofactors	4	<i>moeB</i>	<i>preA</i>	<i>thiG</i>	<i>thiS</i>	
tRNA synthetase	3	<i>syfB</i>	<i>syh</i>	<i>ycf62</i>		
Phycobilisome degradation	1	<i>nblA</i>				
Transport						
Transport	6	<i>cemA</i>	<i>mntA</i>	<i>mntB</i>	<i>sufB</i>	<i>sufC</i> <i>ycf38</i>
Protein translocator	4	<i>secA</i>	<i>secG</i>	<i>secY</i>	<i>tatC</i>	
Unknown						
Conserved ORFs	13	<i>ycf17</i>	<i>ycf19</i>	<i>ycf20</i>	<i>ycf23</i>	<i>ycf33</i> <i>ycf34</i>
		<i>ycf35</i>	<i>ycf46</i>	<i>ycf52</i>	<i>ycf54</i>	<i>ycf55</i> <i>ycf60</i>
		<i>ycf80</i>				
Unique ORFs	37	<i>ORF30</i>	<i>ORF31</i>	<i>ORF32</i>	<i>ORF33</i>	<i>ORF34a</i> <i>ORF34b</i>
		<i>ORF34c</i>	<i>ORF34d</i>	<i>ORF35</i>	<i>ORF36</i>	<i>ORF38</i> <i>ORF39</i>
		<i>ORF40a</i>	<i>ORF40b</i>	<i>ORF43a</i>	<i>ORF43b</i>	<i>ORF44</i> <i>ORF45</i>
		<i>ORF48</i>	<i>ORF50</i>	<i>ORF52</i>	<i>ORF53</i>	<i>ORF54</i> <i>ORF55</i>
		<i>ORF57a</i>	<i>ORF57b</i>	<i>ORF57c</i>	<i>ORF61</i>	<i>ORF62</i> <i>ORF65a</i>
		<i>ORF65b</i>	<i>ORF84</i>	<i>ORF91</i>	<i>ORF93</i>	<i>ORF144</i> <i>ORF228</i>
		<i>ORF770</i>				
Total	224					

ORF と名付けた遺伝子の後ろの数字は、そのタンパク質遺伝子のアミノ酸残基数を示す。

表2-3 : *P. purpureum*葉緑体ゲノムのイントロン

Type	遺伝子名	イントロン名	判断基準	bp (maturase)	遺伝子名	イントロン名	判断基準	bp (maturase)
Group I	<i>rpl28</i>	int (<i>rpl28</i>)	a	195				
Group II	<i>yef46</i>	int.a (<i>yef46</i>)	a, c	485	<i>psbJ</i>	int (<i>psbJ</i>)	a, b	218
		int.b (<i>yef46</i>)	a, c	637	<i>yef55</i>	int (<i>yef55</i>)	a	200
		int.c (<i>yef46</i>)	a	204	<i>rpoA</i>	int (<i>rpoA</i>)	a, b	204
		int.d (<i>yef46</i>)	a	644	<i>gltB</i>	int (<i>gltB</i>)	a, b, c	2454 (<i>mat1d</i>)
		int.e (<i>yef46</i>)	a	435	<i>yef34</i>	int (<i>yef34</i>)	a, b	284
	<i>rpoZ</i>	int (<i>rpoZ</i>)	a, b, c	612	<i>psbN</i>	int.a (<i>psbN</i>)	a, b	225
	<i>psaA</i>	int (<i>psaA</i>)	a, b	257		int.b (<i>psbN</i>)	a, b, c	521
	<i>secG</i>	int (<i>secG</i>)	a, c	605	<i>preA</i>	int (<i>preA</i>)	a, b, c	643
	<i>yef62</i>	int.a (<i>yef62</i>)	a	215	<i>mntA</i>	int (<i>mntA</i>)	a, b, c	1411
		int.b (<i>yef62</i>)	a, c	1001	<i>yef23</i>	int (<i>yef23</i>)	a, b	771
		int.c (<i>yef62</i>)	a	519		<i>rpoC2</i>	int.a (<i>rpoC2</i>)	a, b, c
	<i>yef45</i>	int (<i>yef45</i>)	a	342	int.b (<i>rpoC2</i>)		a, b, c	1252
	<i>rps3</i>	int (<i>rps3</i>)	a, b	214	<i>rpoC1</i>	int (<i>rpoC1</i>)	a, b, c	660
	<i>rps19</i>	int.a (<i>rps19</i>)	a, b	197	<i>rpoB</i>	int.a (<i>rpoB</i>)	a, b, c	2194
		int.b (<i>rps19</i>)	a, b	207		int.b (<i>rpoB</i>)	a, b, c	811
	<i>rpl4</i>	int (<i>rpl4</i>)	a, b	241	<i>tsf</i>	int (<i>tsf</i>)	a, b	228
	<i>dnaK</i>	int.a (<i>dnaK</i>)	a, b, c	577	<i>atpI</i>	int.a (<i>atpI</i>)	a, b, c	607
		int.b (<i>dnaK</i>)	a, b	1779 (<i>mat1b</i>)		int.b (<i>atpI</i>)	a, b, c	1255
	<i>infC</i>	int.a (<i>infC</i>)	a, b	202	<i>atpG</i>	int (<i>atpG</i>)	a, b, c	530
		int.b (<i>infC</i>)	a	214	<i>atpF</i>	int (<i>atpF</i>)	a, b, c	650
int.c (<i>infC</i>)		a, b, c	2397 (<i>mat1c</i>)	<i>dnaB</i>	int (<i>dnaB</i>)	a, b, c	631	

判断基準 : a ; ホモログとのアラインメントを比較した結果から推定した。

b ; 近縁種のESTデータ(Chan et al. 2011)に基づいて推定した。

c ; RNAweaselプログラムで推定した。

表2-4 : *P. purpureum*葉緑体ゲノムのtRNA遺伝子の同定

First base	Second base								Third base
	U		C		A		G		
U	6427-6499	Phe	76836-76921	Ser	61289-61208	Tyr	159711-159640	Cys	U
		Leu							
	12139-12220						160459-160387	Trp	G
C	107977-107896	Leu	129623-129696	Pro	172308-172380	His	49830-49903	Arg	U
						172181-172110			Gln
A	208673-208600	Ile	61177-61106	Thr	148382-148453	Asn	19559-19469	Ser	U
			143603-143531			205509-205437			27424-27496
	*126118-126045, (I)114922-114837, 19649-19577	Met, Ile				Lys			Arg
G	27334-27405	Val	8249-8321	Ala	161600-161527	Asp	19361-19290	Gly	U
						68423-68495			Glu
								G	

数字はゲノム上でのtRNA遺伝子の領域を示す。アスタリスクはtRNA^{fMet}(CAU)を示し、(I)はtRNA^{Ile}(CAU)を示す。

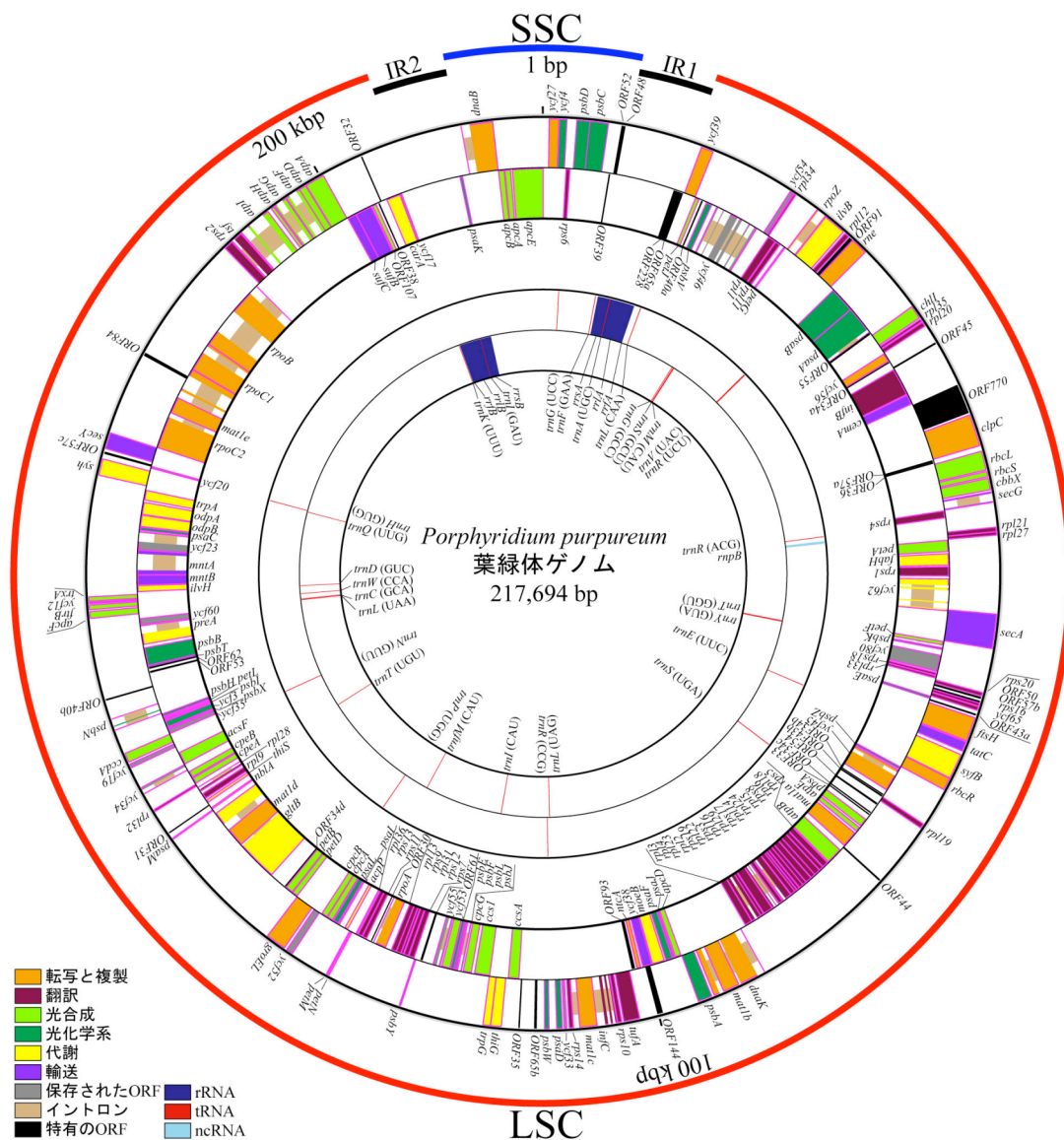


図 2-1 : *P. purpureum* 葉緑体ゲノムの遺伝子地図

内側の 2 つの輪が RNA 遺伝子を示し、外側の 2 つの輪がタンパク質遺伝子を示す。それぞれ外側の輪が時計回り方向にコードされた遺伝子を示し、内側の輪が反時計回り方向にコードされた遺伝子を示す。遺伝子は、機能毎に左下で示した色に塗り分けた。タンパク質遺伝子は枠で囲んでおり、枠が途切れ、帯でつないでいる領域はそこがイントロン領域であることを示す。

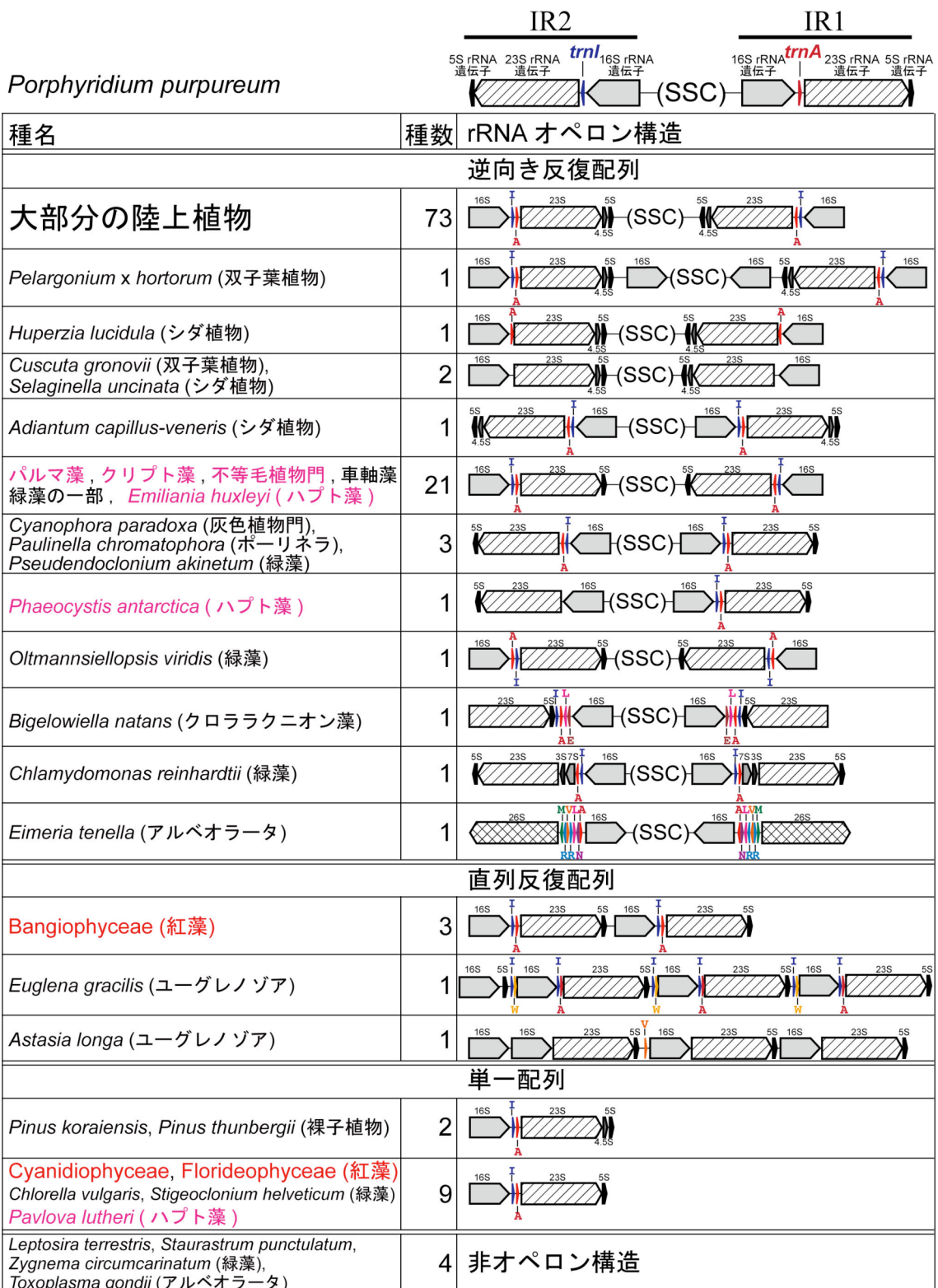


図 2-2 : 葉緑体ゲノムの rRNA オペロン構造

2 つの IR 領域に挟まれた部分は、SSC 領域を示す。tRNA 遺伝子は、それぞれ対応するアミノ酸の大文字一文字での略称で示した。

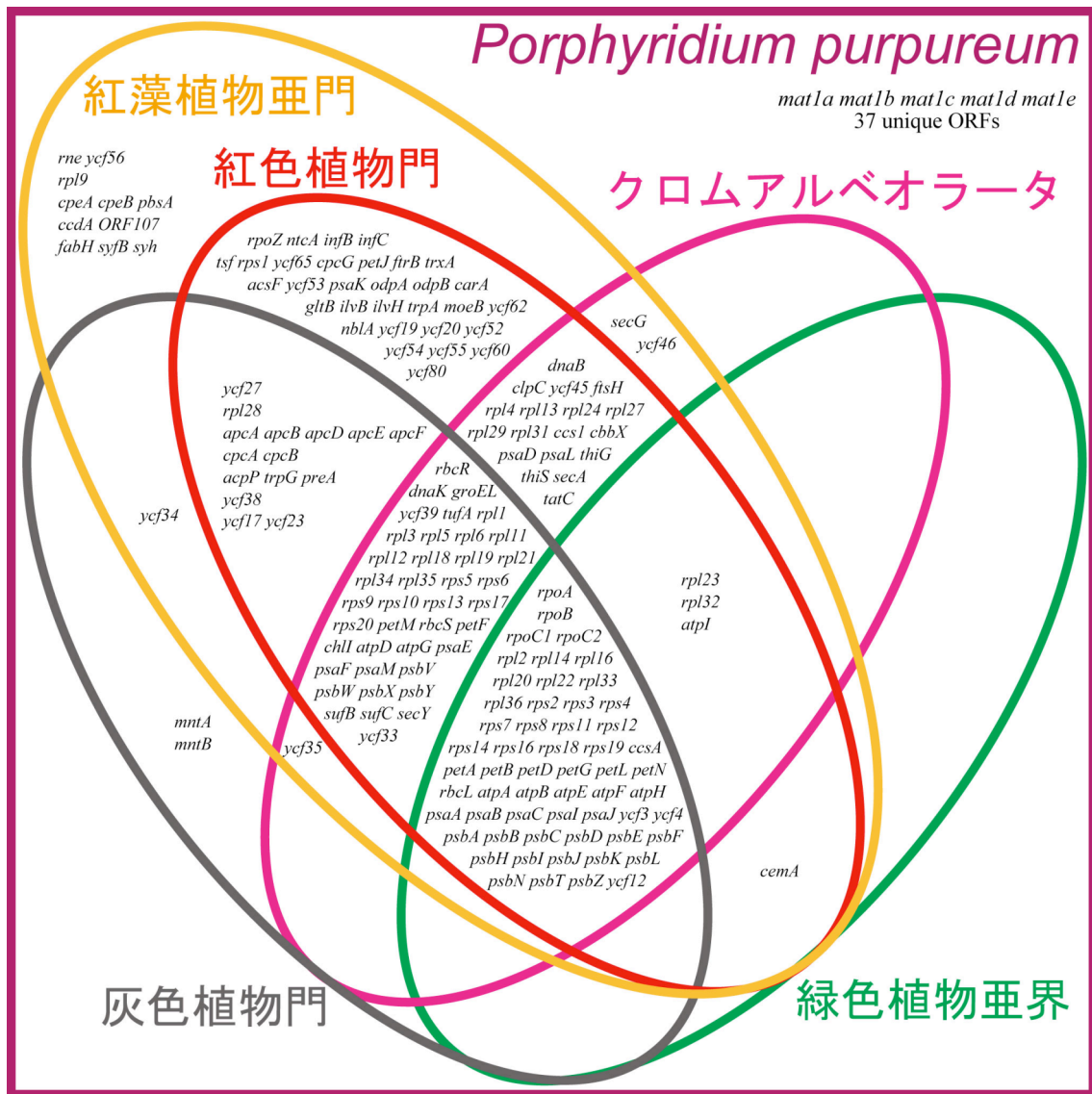


図 2-3 : 葉緑体ゲノムにコードされたタンパク質遺伝子のクラスタリング解析
 ベン図は *P. purpureum* 葉緑体ゲノムにコードされたタンパク質遺伝子の内、それぞれのグループが半数以上共通して保存していたタンパク質遺伝子を示す。

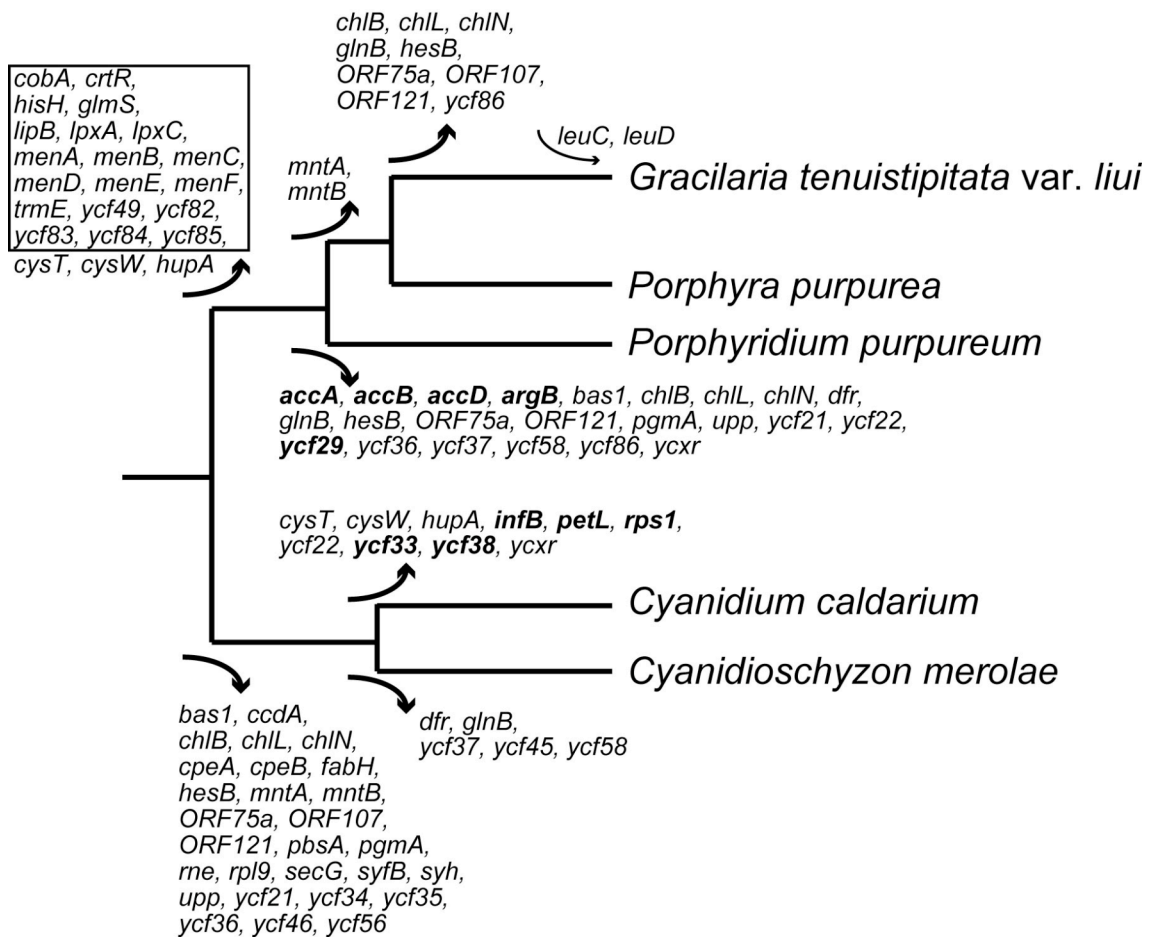


図 2-4 : 紅藻の進化の過程で失われたと考えられるタンパク質遺伝子
 それぞれの分岐の近くに書かれた遺伝子は、その系統が分岐した後に葉緑体ゲノムから失われたと考えられる遺伝子を示す。黒枠で囲った遺伝子達は、紅藻植物亜門でのみ共通して失われた遺伝子を示す。accA, accB, accD, argB, ycf29 は P. purpureum でのみ失われ、infB, petL, rps1, ycf33, ycf38 は C. caldarium でのみ失われた。leuC と leuD は G. tenuistipitata が遺伝子の水平移動により獲得したと考えられている。

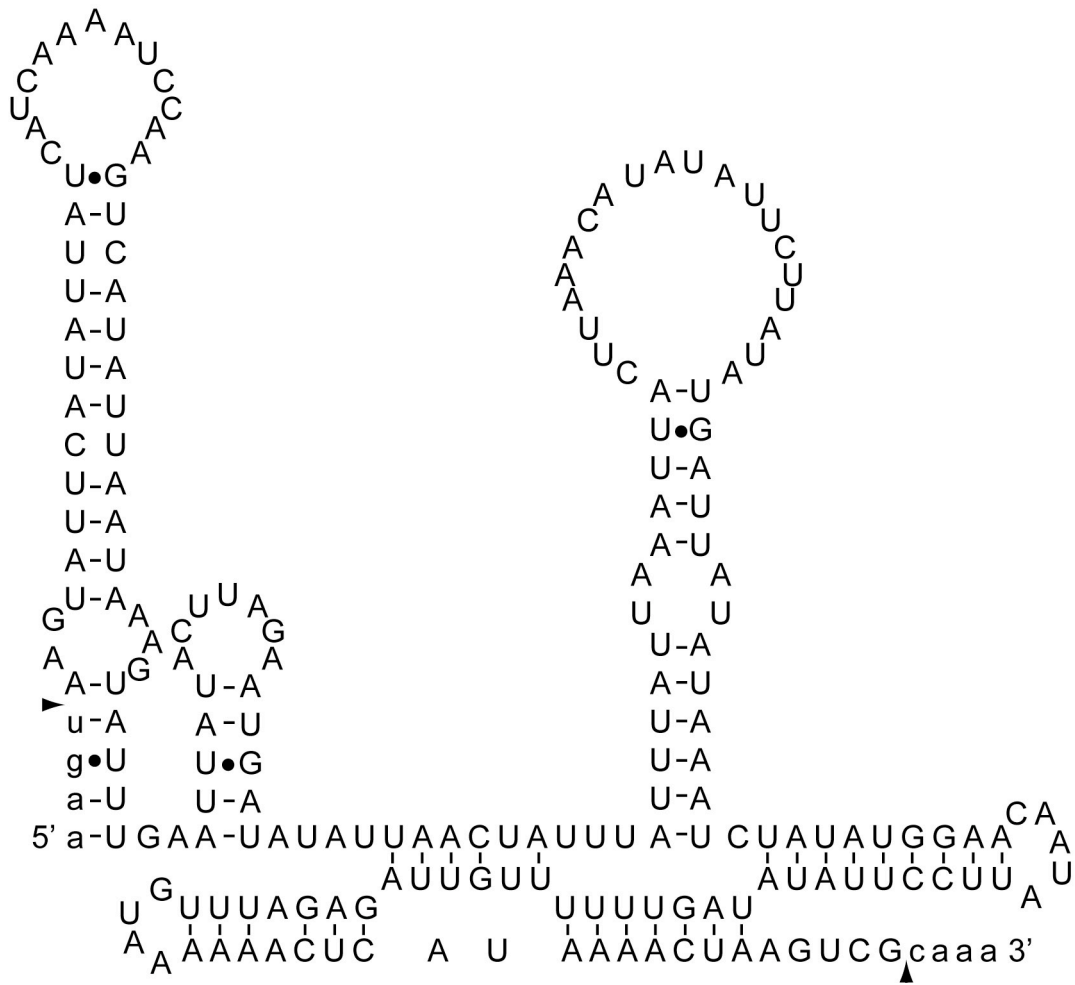


図 2-5 : *int(rpl28)* の推定二次構造

大文字はイントロン配列を示し、小文字はエキソン配列を示す。矢尻は切断部分を示す。黒丸は GU のウォッブルを示す。

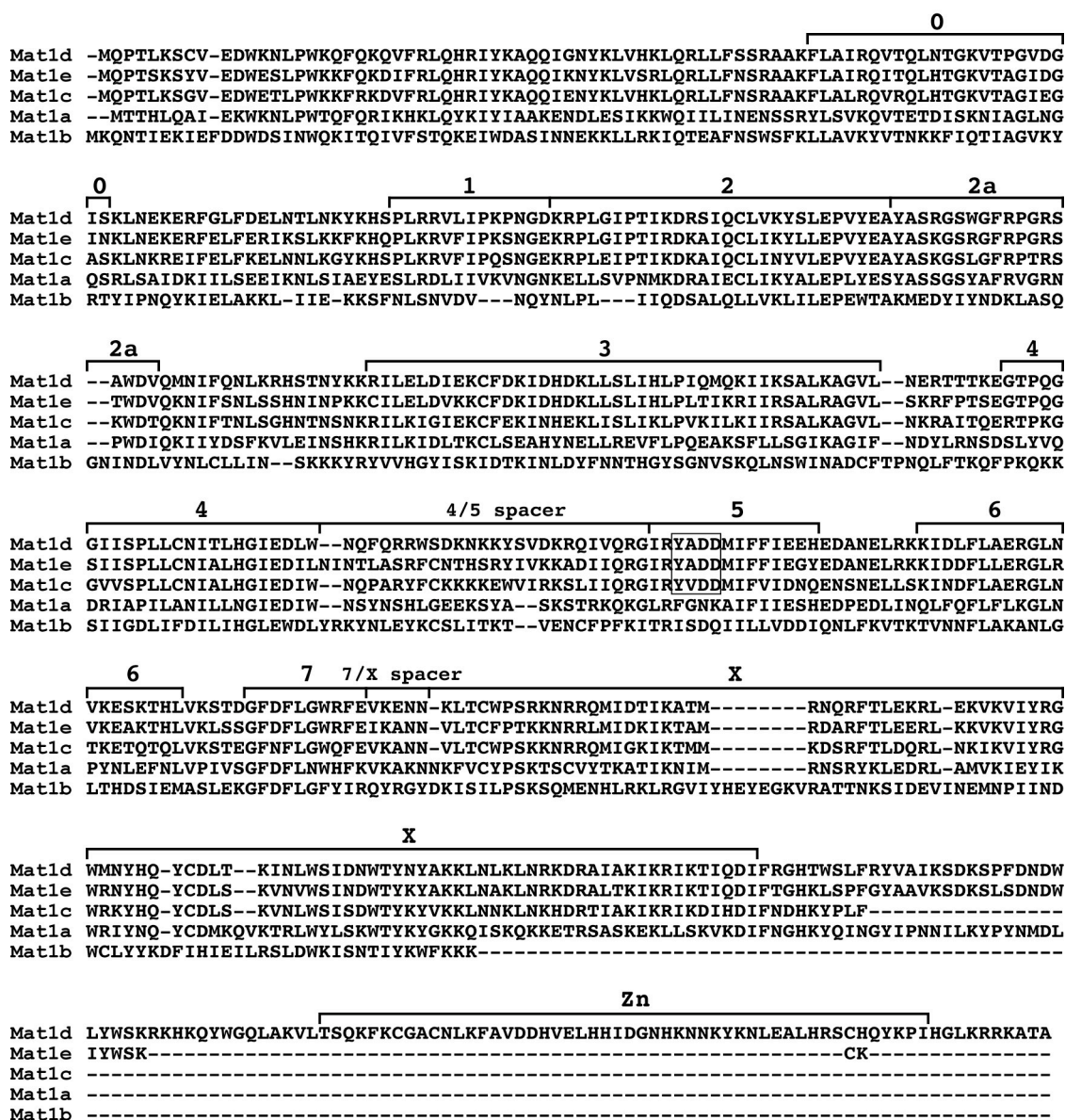


図 2-6: *P. purpureum* 葉緑体ゲノムにコードされた 5 個のマチュラーゼタンパク質のアラインメント

それぞれのドメイン領域(RT(0-7, 2a), X, Zn, 4/5 spacer, 7/X spacer)は、Zimmerly et al. (2001)の表記に従って示した。四角で囲った部分は、YxDDモチーフを示す。

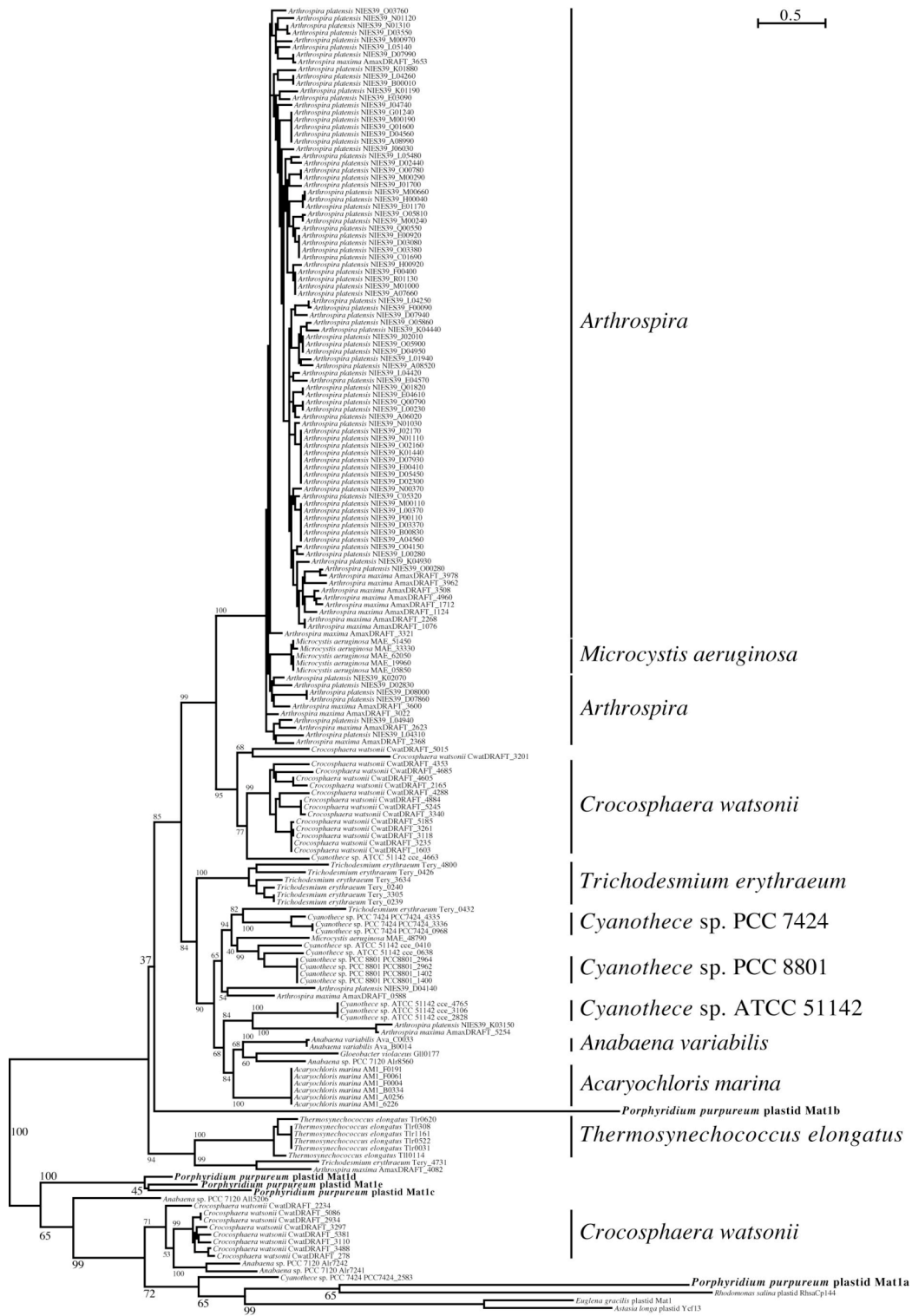


図 2-7: *P. purpureum* 葉緑体ゲノムにコードされたマチュラーゼタンパク質とそのホモログの系統樹
 右上のスケールバーは枝の長さを示す。それぞれの分岐の肩にある数字は最尤法での信頼度を示す。

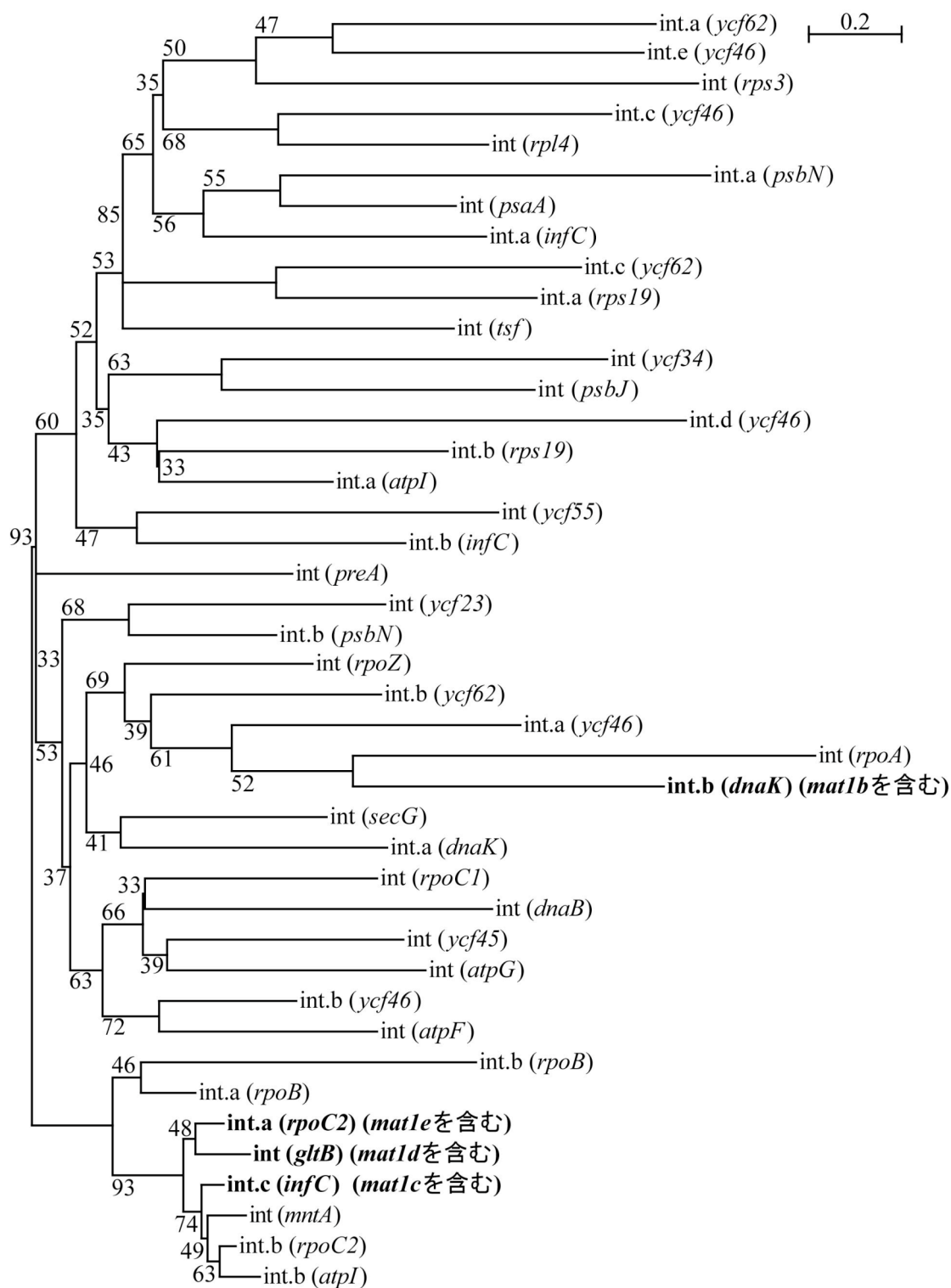


図 2-8 : グループ II イントロンの系統樹

右上のスケールバーは枝の長さを示す。それぞれの分岐の肩にある数字は最尤法での信頼度を示す。太字はマチュラーゼタンパク質遺伝子を領域内に含むイントロンを示す。

tRNA^{Met} (CAU)

PmPuc trnM
PPUC trnI
GTEc trnM (64208..64284)
Parmac trnM
CMEc trnI
CCAc tRNA-Met c(142204..142277)
CPAc trnI
ATHc trnM
Spinachc trnM
PPTc trnM
CREc trnM2
Ana trnM
Syn trnM
Lim trnM

tRNA^{fMet} (CAU)

PmPuc trnM
PPUC trnM
GTEc trnM (100175..100251)
Parmac trnM
CMEc trnM
CCAc tRNA-Met (156623..156696)
CPAc trnM
ATHc trnM
Spinachc trnM
PPTc trnM
CREc trnM3
Ana trnM
Syn trnM
Lim trnM2
Lim trnM

tRNA^{Ile} (CAU)

PmPuc trnI
PPUC trnM
Parmac trnI
CMEc trnM
CCAc tRNA-Met c(109401..109484)
CPAc trnM
ATHc trnI
Spinachc trnI
PPTc trnI
CREc trnM1
Ana trnI
Syn trnI
Lim trnI

図 2-9 : tRNA(CAU)遺伝子のアラインメント

C. caldariumとG. tenuistipitataは、複数のtRNA(CAU)遺伝子が同じ名前で登録されていたので、コード領域を後ろに書き加えた。コード領域の前にあるcは逆向きにコードされた領域であることを示す。それぞれのtRNA(CAU)遺伝子内で共通している塩基はコロンで、全てのtRNA(CAU)遺伝子に共通している塩基はアスタリスクで、それぞれのアラインメントの上部に示した。略称の後にあるcは葉緑体ゲノムにコードされているものであることを示す。

略称は、それぞれPmPU, Porphyridium purpureum; PPU, Porphyra purpurea; GTE, Gracilaria tenuistipitata; Parma, Triparma laevis; CME, Cyanidioschyzon merolae; CCA, Cyanidium caldarium; CPA, Cyanophora paradoxa; ATH, Arabidopsis thaliana; Spinach, Spinacia oleracea; PPT, Physcomitrella patens; CRE, Chlamydomonas reinhardtii; Ana, Anabaena sp. PCC 7120; Syn, Synechocystis sp. PCC 6803; Lim, Limnothrix sp. ABRG5-3を示す。

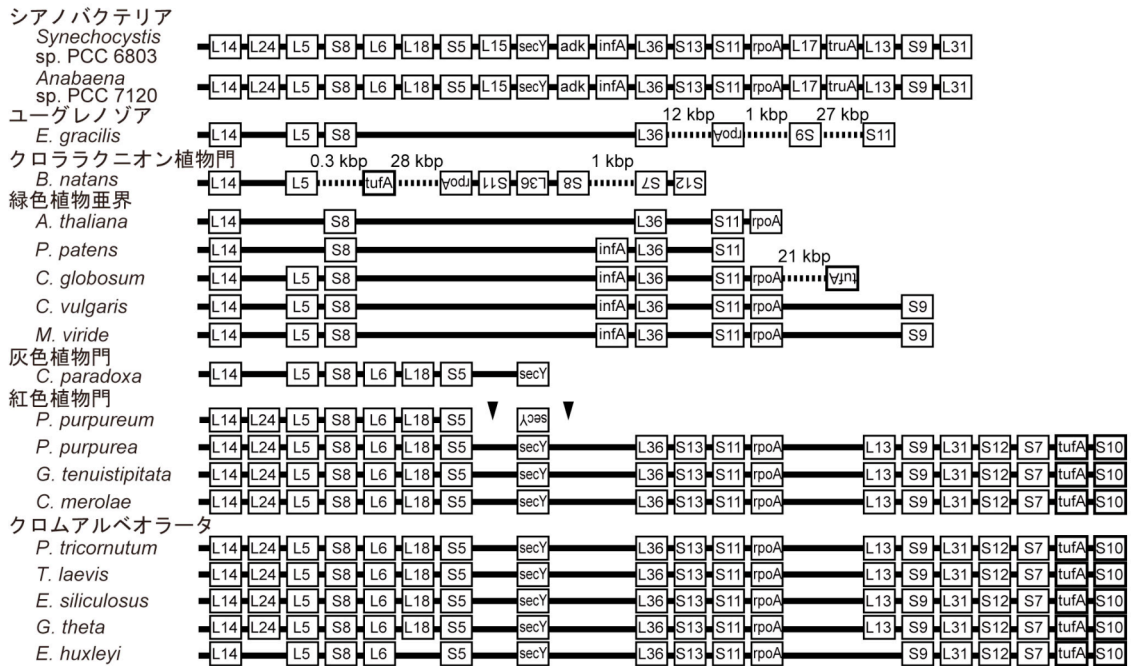
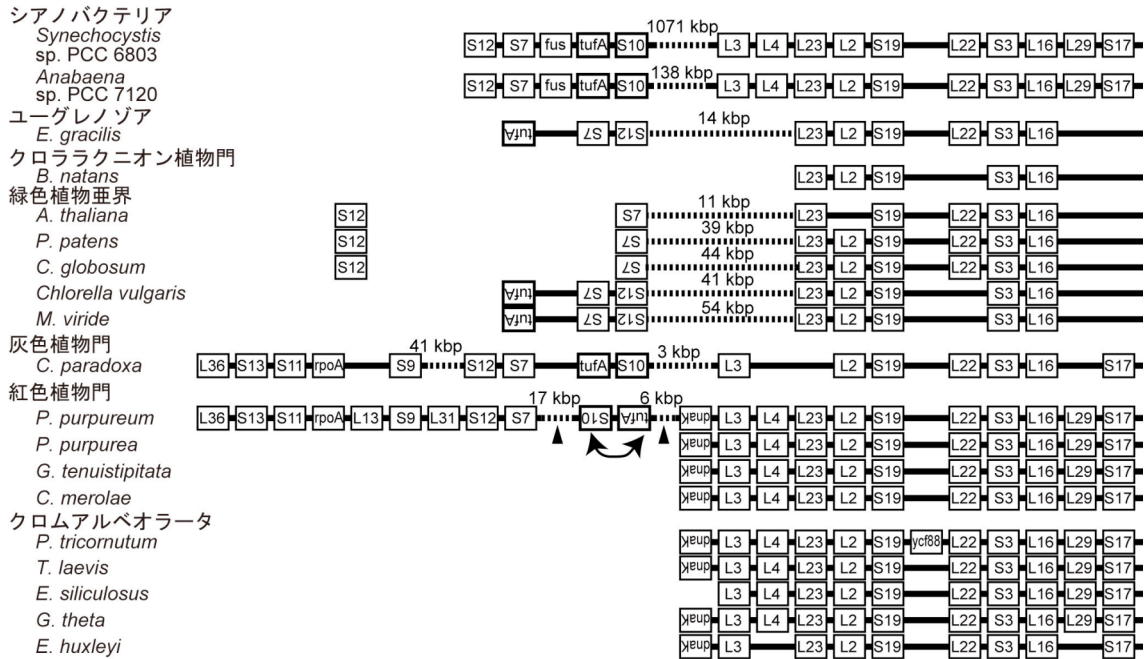


図2-10：リボソームタンパク質遺伝子クラスターの比較

連続している遺伝子は太線で繋げて示した。長いので2行に分けて示した。上下逆になっている遺伝子は逆方向にコードされた遺伝子を示す。矢尻は、*P. purpureum* で分断されている箇所を示す。*A. thaliana* と *P. patens* と *C. globosum* の *rps12* はトランススプライシングが起きている。

第3章

パルマ藻 *Triparma laevis* の葉緑体と

ミトコンドリアのゲノム解読

3.1. 序文

パルマ藻は数枚の様々な形をした珪酸の殻で覆われた、直径 2-5 μm の単細胞海洋性ピコ植物プランクトンである(Booth & Marchant 1987)。パルマ藻は、光学顕微鏡を用いて野外の他のピコ植物プランクトンから分離するにはとても小さく、単離培養は難しかったが、近年、新しく作られる珪酸の細胞壁に取り込まれる蛍光色素を利用することで、北太平洋の親潮海域から天然のパルマ藻 *Triparma laevis* が単離された(Ichinomiya et al. 2010)。また、色素の構成や 18S rRNA と *rbcL* の分子系統解析の結果から、パルマ藻は、それまで属していた黄金色藻綱ではなく、ボリド藻綱に含まれていると報告された(Ichinomiya et al. 2011)。ボリド藻は、無殻の海産性ピコ植物プランクトンで、分子系統解析の結果から珪藻の姉妹群であると見られている(Guillou et al. 1999, Lovejoy et al. 2006)。パルマ藻のゲノムを、より詳細に解読することで、パルマ藻、ボリド藻、珪藻の進化的関係への理解を深めることを目的としている。この章ではパルマ藻の葉緑体およびミトコンドリアゲノムを解読し、コードされた遺伝子の種類やゲノムの構造について、それぞれ珪藻およびその他の不等毛植物門と比較した。

3.2. 材料と方法

3.2.1. 生物材料と培養

100 μM のケイ酸と 100 μM の硝酸塩を加えた f/2 培地(Guilard and Ryther 1962)を培地とし、5°Cで、100 $\mu\text{mol}/\text{m}^2 \cdot \text{s}$ の光を 14 時間明期 : 10 時間暗期の条件で当て培養した(Ichinomiya et al. 2011)。培養し、メンブレン上に濾過して集めたものを水産総合研究センター東北区水産研究所の桑田晃博士から頂いた。頂いた株と同じ株は現在、国立環境研究所に登録されている(NIES-2565)。

3.2.2. 試薬調整

2.2.2. で表記済みの試薬は、2.2.2. と同じ方法で作成した。

- STE: 50 ml の TE50/20 に NaCl 0.29 g (0.1 M)を加えて溶かした。作成後、冷凍保存した。

3.2.3. DNA 精製

濾過したメンブレンを 50 ml チューブに入った 20 ml の STE で洗い、付着したパルマ藻を洗い落とした。洗い落とす作業は氷上で冷やしながら行った。洗い終わったら、5%量の N-ドデカノイルサルコシン酸ナトリウムと 20 mg/ml のプロテイナーゼ K を 100 μl 加え、50°Cで 1 時間揺らしながらインキュベートした。再び、20 mg/ml のプロテイナーゼ K を 100 μl 加え、50°Cで 16 時間インキュベートした。さらに、20 mg/ml のプロテイナーゼ K を 100 μl 加え、50°Cで 1 時間揺らしながら、その後は揺らさずに 30 分間インキュベートした。2 本の沈殿管に移し、フェノールを沈殿管の肩まで加え、10 分以上よく混ぜた。4°C、8,400 g、5 分間遠心した。上層部の水層を新しい沈殿管に、シリコンコーティングしたパスツールピペットで移した。下層部のフェノール層に STE を 5 ml ずつ加え、5 分間よく混ぜた後、4°C、8,400 g、5 分間遠心した。水層を既に移した沈殿管に移し、フェノールとクロロホルム液を 10 ml ずつ加えた。5 分間よく混ぜた後、4°C、8,400 g、5 分間遠心した。水層を新しい沈殿管に移し、クロロホルム液を 10 ml 加えた。5 分間よく混ぜた後、4°C、8,400 g、5 分間遠心した。同じ作業をもう一度行った。水層を 1 本の 50 ml

チューブに移し、水層と同じ重さの CsCl を加えた。50°C でしばらく温め、CsCl を溶かした。最終濃度 0.2 mg/ml になるよう 20 mg/ml EtBr を加え、遮光して 5 分間 50°C で温めた。超遠心用チューブに移して 25°C、315,000 g、18 時間超遠心した。紫外線で蛍光した DNA のバンド部分を新しい超遠心用チューブに移し、バランス用 CsCl/EtBr で少し薄め、25°C、315,000 g、18 時間超遠心した。DNA バンドを回収し、**2.2.3.**と同様に 2-プロパノールで EtBr を除き、TE で透析したものを回収した。濃度は 268.2 ng/μl だった。

3.2.4. DNA 配列解析

26.8 μg の DNA を **2.2.4.**と同様にゲノムの塩基配列解析を行い、サンガー法によりギャップを埋めた。解読したオルガネラゲノムの全長配列は GenBank に登録した。(葉緑体ゲノムの accession number: AP014625, ミトコンドリアゲノムの accession number: AP014626)

3.2.5. ゲノムアノテーションとデータ解析

葉緑体ゲノムのタンパク質遺伝子および tRNA、rRNA 遺伝子の推定は、**2.2.5.**と同様の方法で行った。ncRNA 遺伝子(*ffs*、*ssrA*)は、珪藻と共通している遺伝子の並びを利用して(*ffs* は tRNA^{Phe}(GAA)と *psbX* の間、*ssrA* は tRNA^{Arg}(CCG)と tRNA^{Ile}(CAU)の間)、*ffs* の場合は tRNA^{Phe}(GAA)と *psbX* の間の塩基配列と珪藻の *ffs* 遺伝子の塩基配列でアラインメントを作成し、相同性の高い部分を *ffs* 遺伝子の塩基配列として推定した。*ssrA* の場合も同様の推定方法を用いた。アラインメント作成は Clustal X ver. 2.0.9 を利用した。

ミトコンドリアゲノムは、まず *T. laevis* のドラフトゲノムに対して、3 種の紅藻、16 種の緑色植物亜界、1 種のハプト藻、1 種のクリプト藻、1 種の不等毛植物門、1 種のエクスカバータ、計 23 種の生物のミトコンドリアゲノムにコードされたタンパク質遺伝子のアミノ酸配列を用いた tblastn を行い、どのスキュアフォールドがミトコンドリアゲノムか推定した。得られたスキュアフォールドのギャップをサンガー法により埋めた後、MetaGeneAnnotator と不等毛植物門のミトコンドリアゲノムにコードされたタンパク質遺伝子のアミノ酸配列を用いた tblastn によって、アノテーションをつけた。RNA 遺伝子の推定には不等毛植物門のミトコンドリアゲノムにコードされた RNA

遺伝子の塩基配列を用いた blastn を行った。また、tRNA 遺伝子の推定のために tRNAscan-SE も利用した。DNA およびタンパク質配列の加工には SISEQ ver. 1.59 を利用した。ゲノム配列の編集は Artemis ver. 13.0 を利用した。

3.2.6. *T. laevis* の葉緑体ゲノムにコードされた 2 コピーの *psbA* の系統解析

T. laevis の葉緑体ゲノムにコードされた 2 コピーの *psbA* の系統解析をするために、Gclust のデータセット“CyanoClust4”に登録されている紅藻およびクロムアルベオラータと 2.2.5. で表記した 7 種の不等毛植物門および *P. purpureum* の相同タンパク質配列を利用した。アラインメント作成は、Muscle ver. 3.6 を利用し、LG モデルを用いた最尤法で系統解析を行った。最尤法の計算には TREEFINDER ver. March 2011 を利用した。また、Clustal X ver. 2.0.9 を用いて 2 コピーの *psbA* のアミノ酸配列の比較を行った。

3.3. 結果

3.3.1. *T. laevis* オルガネラゲノムの構造

T. laevis の葉緑体およびミトコンドリアゲノムを Roche Diagnostics 社の genome sequencer FLX instrument を用いて解読した。

葉緑体ゲノムは 21 個のコンティグにアセンブルされ、ギャップはサンガー法を用いて解読した。*T. laevis* の葉緑体ゲノムは、全長 117,514 bp、GC 含量 31.7%の環状 DNA だった(図 3-1a)。2 個の ncRNA 遺伝子(*ffs*, *ssrA*)、6 個の rRNA 遺伝子(2 個の *rrs*、2 個の *rpl*、2 個の *rpl*)、29 個の tRNA 遺伝子および 141 個のタンパク質遺伝子をコードすることが分かった(表 3-1、表 3-2)。また、系統関係は近いが、配列が異なる 2 コピーの *psbA* 遺伝子をコードしていた(図 3-2)。*T. laevis* の葉緑体ゲノムの rRNA オペロンは、2 つの逆向きで反復した構造だった(図 2-2)。

ミトコンドリアゲノムは 1 個のスキュップフォールドにアセンブルされ、ギャップはサンガー法を用いて解読した。*T. laevis* のミトコンドリアゲノムは、全長 39,580 bp、GC 含量 30.4%の環状 DNA だった(図 3-1b)。2 個の rRNA 遺伝子(*rrs*, *rpl*)、25 個の tRNA 遺伝子および 37 個のタンパク質遺伝子をコードすることが分かった(表 3-3a、表 3-4)。ミトコンドリアゲノムでは生物種によってはトリプトファンをコードすることもある TGA コドン(Fox 1979, Oudot-Le Secq & Green 2011)は、*T. laevis* のミトコンドリアゲノムでは終止コドンとしても使用されていなかった(表 3-3b)。

3.3.2. 系統間で保存されているタンパク質遺伝子の比較

2.3.2.と同じ考えをもとに、不等毛植物門の進化の過程で葉緑体ゲノムから失われたタンパク質遺伝子を調べた(図 3-3)。*T. laevis* と珪藻は、黄緑色藻、褐藻、ラフィド藻と分岐した後に、13 個のタンパク質遺伝子を失っていた。*bas1* は、紅藻(図 2-4)や珪藻の間では進化の過程で系統毎に何度も葉緑体ゲノムから失われており(Glöckner et al. 2000, Ruck et al. 2014)、*T. laevis* も失っていた。*T. laevis* と珪藻を比較すると、両者が分岐した後に珪藻は *chlB*, *chlL*, *chlN*, *rpoZ*, *ycf19* を失った。また、その他のいくつかの遺伝子は種によって失われるかもしくは核に移行していた(Lommer et al. 2010, Oudot-Le Secq et al.

2007, Ruck et al. 2014)。葉緑体ゲノムにはコードされていなかった *petJ* について、*T. laevis* のドラフトゲノムを検索した結果、核ゲノムの方に *petJ* のホモログが見つかり、*T. laevis* でも *petJ* が葉緑体ゲノムから核ゲノムに移行していることが考えられた。

一部の珪藻の葉緑体ゲノムにコードされている *serC* は、プラスミドゲノムから水平伝播した結果、獲得した遺伝子と考えられており、一部の珪藻およびラフィド藻の葉緑体ゲノムにコードされている *tyrC* は、プラスミドゲノムかバクテリアゲノムのどちらかから水平伝播した結果、獲得した遺伝子と考えられている(Ruck et al. 2014)。*ycf88*, *ycf89*, *ycf90* といった3種の機能未知タンパク質遺伝子は、珪藻以外にホモログは無く、珪藻に保存されたタンパク質遺伝子である(*Leptocylindrus danicus* のみ *ycf88* は持っていない)。これらの遺伝子は、それぞれの種や系統に分岐した後で獲得した遺伝子と考えた。

3.3.3. パルマ藻と珪藻の葉緑体ゲノム構造の比較

T. laevis と葉緑体ゲノムが解読されたその他の不等毛植物門の葉緑体ゲノムを比較した(図 3-4)。*T. laevis* と珪藻のゲノムサイズはほぼ同じであった。珪藻のなかにはゲノムサイズが他と比べて大きい種もあるが、それは遺伝子間領域が非常に長い箇所が多くあることや IR 領域が長いことが原因だと考えられる(Ruck et al. 2014)。遺伝子の並びは大まかな塊に分けて保存されていた。他の不等毛植物門では遺伝子の塊は保存されていなかった。

IR 領域の長さや遺伝子の並びに違いはあるが、SSC 側にコードされた遺伝子の種類は、その種が独自にコードしている遺伝子以外、全ての珪藻で共通していた。*T. laevis* と珪藻の最も大きな違いは、リボソームタンパク質遺伝子クラスターの塊(*dnaK*, *rpl3*, *rpl4*, *rpl23*, *rpl2*, *rps19*, *rpl22*, *rps3*, *rpl16*, *rpl29*, *rps17*, *rpl14*, *rpl24*, *rpl5*, *rps8*, *rpl6*, *rpl18*, *rps5*, *secY*, *rpl36*, *rps13*, *rps11*, *rpoA*, *rpl13*, *rps9*, *rpl31*, *rps12*, *rps7*, *tufA*, *rps10*)とそれに付随する2つの塊(*rps16*, *rps4*, tRNA^{His}(GUG), *thiS*, *syfB*, *psbW*, tRNA^{Gln}(UUG), tRNA^{Arg}(ACG), *groEL*)の塊と(*clpC*, tRNA^{Cys}(GCA), tRNA^{Leu}(UAA), *ilvH*)の塊)が、珪藻の葉緑体ゲノムでは SSC 側にコードされているのに対し、*T. laevis* では LSC 側にコードされていることだった。

3.3.4. ミトコンドリアゲノムの比較

不等毛植物門のミトコンドリアゲノムを比較した(表 3-5)。解読された不等毛植物門のミトコンドリアゲノムでは、ほぼ全ての遺伝子が保存されていた。また、スレオニンの tRNA 遺伝子が共通して失われていた。*T. laevis*、珪藻、ラフィド藻、シヌラ藻、黄金色藻は、全長の *nad11* をミトコンドリアゲノムにコードしているのに対して、褐藻は N 末端部分のみコードしていた (Oudot-Le Secq & Green 2011)。*T. laevis* と珪藻のミトコンドリアゲノムにコードされた遺伝子の種類の違いは、tRNA^{Gly}(UCC)が珪藻では失われていることのみだった。

3.4. 考察

Triparma laevis のオルガネラゲノムを解読し、不等毛植物門のオルガネラゲノムと比較した結果、コードされたタンパク質遺伝子の種類やゲノム構造が、他の不等毛植物門よりも珪藻に十分近いことが分かった。

T. laevis と珪藻の葉緑体ゲノムにコードされたタンパク質遺伝子を比較したところ、珪藻の進化の過程で系統毎に何度も葉緑体ゲノムから失われた *bas1* や *petJ*、そして珪藻以外にホモログが見つからず、珪藻が独自に獲得したかもしれない *ycf88*, *ycf89*, *ycf90* を除いて、珪藻から分岐した後に *T. laevis* から失われた遺伝子は無いことが分かった。*T. laevis* の葉緑体ゲノムに2コピーある *psbA* は、両者の系統関係も近いことから、*T. laevis* が珪藻から分岐した後で *psbA* に遺伝子重複が起き、片方はいくつかの塩基について変異が起きたのだろうと考えられた。

他の不等毛植物門と比べて、*T. laevis* と珪藻の葉緑体ゲノムの遺伝子の並びは、あまり入れ替わりが起きておらず、いくつかの大まかな塊毎に遺伝子の並びが保存されていた。このことから、*T. laevis* と珪藻の共通祖先が他の不等毛植物門から分岐してからは、大規模な葉緑体ゲノムの再編成は起きていないことが考えられた。*T. laevis* と珪藻の最も大きな違いは、リボソームタンパク質遺伝子クラスターの塊とそれに付随する2つの塊が、珪藻の葉緑体ゲノムではSSC側にコードされているのに対し、*T. laevis* ではLSC側にコードされていることだった。ボリド藻や他のパルマ藻の葉緑体ゲノムを解読することで、*T. laevis* が他の種から分岐した後でこれらの遺伝子群が移動したのか珪藻の系統とパルマ藻およびボリド藻の系統が分岐した後にどちらかでこれらの遺伝子群が移動したのかを確かめられると考えられる。

珪藻の葉緑体ゲノムは、遺伝子の並びが保存されており、*T. laevis* の葉緑体ゲノムもそれらと似た遺伝子の並びが保存されていた。一方で、珪藻のミトコンドリアゲノムは遺伝子の種類は保存していたが遺伝子の並びは保存していなかった。また、珪藻のミトコンドリアゲノムは大量の反復配列が一カ所にまとまって存在する領域やイントロン配列を持っているが(Oudot-Le Secq & Green 2011)、*T. laevis* には反復配列領域もイントロン配列も確認できなかった。珪藻の反復配列の構造は種毎に異なり、それぞれ独自に獲得した

領域であるとされている。ボリド藻や他のパルマ藻のミトコンドリアゲノムを解読することで、*T. laevis* が二次的にこのような構造を失ったのか珪藻になってからこのような構造を獲得したのかを確かめられると考えられる。

T. laevis の葉緑体ゲノムの IR 領域は、rRNA オペロンと機能未知の遺伝子 (*ORF50*, *ORF51*) の N 末端側の一部のみであり、珪藻と比べて短いことや、ミトコンドリアゲノムに反復配列領域もイントロン配列も見られないことから、珪藻のオルガネラゲノムよりも単純な構造をしていることが分かった。

珪藻では失われた遺伝子を保存していることや、珪藻よりも単純なゲノム構造をしていることから、*T. laevis* のオルガネラゲノムは、現在の珪藻よりも珪藻の共通祖先に近い構造を保存しているかもしれないと考えられた。今後、*T. laevis* の核ゲノムを解読することは、パルマ藻自身のゲノムの特徴について知るだけでなく、珪藻の共通祖先や進化について調べるうえでの足がかりとして有用であることが期待される。また、ボリド藻や他のパルマ藻のオルガネラゲノムを解読することで、実は *T. laevis* がたまたま単純な構造をしたオルガネラゲノムに進化しただけなのかどうかを確かめられると考えられる。

表 3-1 : *T. laevis* 葉緑体ゲノムにコードされたタンパク質遺伝子の機能分類

分類	遺伝子数	遺伝子名					
Transcription & replication							
Helicase	1	<i>dnaB</i>					
RNA polymerase	5	<i>rpoA</i>	<i>rpoB</i>	<i>rpoC1</i>	<i>rpoC2</i>	<i>rpoZ</i>	
Transcription	1	<i>rbcR</i>					
Chaperone	5	<i>clpC</i>	<i>dnaK</i>	<i>groEL</i>	<i>ycf39</i>	<i>ycf45</i>	
Protease	1	<i>ftsH</i>					
Translation							
Translation	2	<i>tsf</i>	<i>tufA</i>				
Ribosomal protein	26	<i>rpl1</i>	<i>rpl2</i>	<i>rpl3</i>	<i>rpl4</i>	<i>rpl5</i>	<i>rpl6</i>
		<i>rpl11</i>	<i>rpl12</i>	<i>rpl13</i>	<i>rpl14</i>	<i>rpl16</i>	<i>rpl18</i>
		<i>rpl19</i>	<i>rpl20</i>	<i>rpl21</i>	<i>rpl22</i>	<i>rpl23</i>	<i>rpl24</i>
		<i>rpl27</i>	<i>rpl29</i>	<i>rpl31</i>	<i>rpl32</i>	<i>rpl33</i>	<i>rpl34</i>
		<i>rpl35</i>	<i>rpl36</i>				
	18	<i>rps2</i>	<i>rps3</i>	<i>rps4</i>	<i>rps5</i>	<i>rps6</i>	<i>rps7</i>
		<i>rps8</i>	<i>rps9</i>	<i>rps10</i>	<i>rps11</i>	<i>rps12</i>	<i>rps13</i>
		<i>rps14</i>	<i>rps16</i>	<i>rps17</i>	<i>rps18</i>	<i>rps19</i>	<i>rps20</i>
Photosynthesis							
Cytochrome complex	9	<i>ccsA</i>	<i>ccs1</i>	<i>petA</i>	<i>petB</i>	<i>petD</i>	<i>petG</i>
		<i>petL</i>	<i>petM</i>	<i>petN</i>			
Carbon fixation	3	<i>cbbX</i>	<i>rbcL</i>	<i>rbcS</i>			
Redox system	1	<i>petF</i>					
Chlorophyll biosynthesis	4	<i>chlB</i>	<i>chlI</i>	<i>chlL</i>	<i>chlN</i>		
ATP synthase	8	<i>atpA</i>	<i>atpB</i>	<i>atpD</i>	<i>atpE</i>	<i>atpF</i>	<i>atpG</i>
		<i>atpH</i>	<i>atpI</i>				
Photosystem							
Photosystem I	12	<i>psaA</i>	<i>psaB</i>	<i>psaC</i>	<i>psaD</i>	<i>psaE</i>	<i>psaF</i>
		<i>psaI</i>	<i>psaJ</i>	<i>psaL</i>	<i>psaM</i>	<i>ycf3</i>	<i>ycf4</i>
Photosystem II	20	<i>psbA</i>	<i>psbA2</i>	<i>psbB</i>	<i>psbC</i>	<i>psbD</i>	<i>psbE</i>
		<i>psbF</i>	<i>psbH</i>	<i>psbI</i>	<i>psbJ</i>	<i>psbK</i>	<i>psbL</i>
		<i>psbN</i>	<i>psbT</i>	<i>psbV</i>	<i>psbW</i>	<i>psbX</i>	<i>psbY</i>
		<i>psbZ</i>	<i>ycf12</i>				
Metabolism							
Lipids	1	<i>acpP</i>					
Amino acid synthesis	2	<i>ilvB</i>	<i>ilvH</i>				
Cofactors	2	<i>thiG</i>	<i>thiS</i>				
tRNA synthetase	1	<i>syfB</i>					
Transport							
Transport	2	<i>sufB</i>	<i>sufC</i>				
Protein translocator	4	<i>secA</i>	<i>secG</i>	<i>secY</i>	<i>tatC</i>		
Unknown							
Conserved ORFs	6	<i>ycf19</i>	<i>ycf33</i>	<i>ycf35</i>	<i>ycf41</i>	<i>ycf46</i>	<i>ycf66</i>
Unique ORFs	7	<i>ORF211</i>	<i>ORF51</i>	<i>ORF4</i>	<i>ORF32</i>	<i>ORF79</i>	<i>ORF35</i>
		<i>ORF50</i>					
Total	141						

ORF と名付けた遺伝子の後ろの数字は、そのタンパク質遺伝子のアミノ酸残基数を示す。

表 3-3 : *T. laevis* ミトコンドリアゲノムにコードされたタンパク質遺伝子

a: タンパク質遺伝子の機能分類

分類	遺伝子数	遺伝子名				
Translation						
Ribosomal protein	5	<i>rpl2</i>	<i>rpl5</i>	<i>rpl6</i>	<i>rpl14</i>	<i>rpl16</i>
	11	<i>rps2</i>	<i>rps3</i>	<i>rps4</i>	<i>rps7</i>	<i>rps8</i>
		<i>rps11</i>	<i>rps12</i>	<i>rps13</i>	<i>rps14</i>	<i>rps19</i>
NADH dehydrogenase						
NADH dehydrogenase subunit	10	<i>nad1</i>	<i>nad2</i>	<i>nad3</i>	<i>nad4</i>	<i>nad4L</i>
		<i>nad6</i>	<i>nad7</i>	<i>nad9</i>	<i>nad11</i>	<i>nad5</i>
ATPase						
ATPase subunit	3	<i>atp6</i>	<i>atp8</i>	<i>atp9</i>		
Cytochrome						
Cytochrome b	1	<i>cob</i>				
Cytochrome oxidase subunit	3	<i>cox1</i>	<i>cox2</i>	<i>cox3</i>		
Transport						
Protein translocator	1	<i>tatC</i>				
Unknown						
Unique ORFs	3	<i>ORF61</i>	<i>ORF159</i>	<i>ORF277</i>		
Total	37					

ORF と名付けた遺伝子の後ろの数字は、そのタンパク質遺伝子のアミノ酸残基数を示す。

b: タンパク質遺伝子の開始コドンと終止コドン

		遺伝子数
開始コドン	ATG	35
	TTG	2
終止コドン	TAA	35
	TAG	2

表 3-4 : *T. laevis* ミトコンドリアゲノムの tRNA 遺伝子の同定

First base	Second base								Third base
	U		C		A		G		
U		Phe		Ser		Tyr		Cys	U
	26198-26126				26296-26214		8913-8841		C
	7740-7659	Leu	5169-5085					A	
						10353-10425	Trp	G	
C		Leu		Pro		His		Arg	U
					4073-4001				C
	7635-7554		35520-35593		22357-22286	Gln	6252-6325	A	
								G	
A		Ile		Thr		Asn		Ser	U
	11596-11523				22460-22388		9019-9106		C
					3974-3902		4151-4078	Arg	A
	8924-8997, (I)10589-10660, *36135-36209	Met, Ile			Lys		G		
G		Val		Ala		Asp		Gly	U
					33310-33383		5269-5198		C
	3396-3468		17067-16995		15480-15409	Glu	13195-13125	A	
							G		

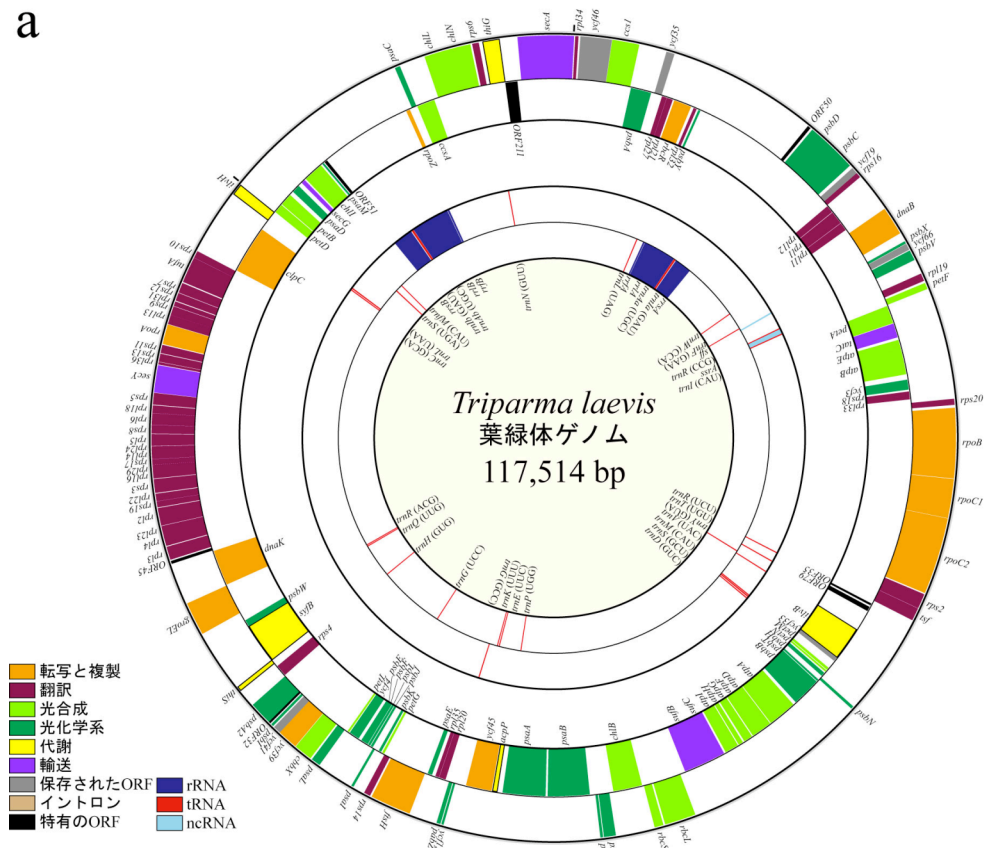
数字はゲノム上での tRNA 遺伝子の領域を示す。アスタリスクは tRNA^{fMet}(CAU)を示し、(I)は tRNA^{Ile}(CAU)を示す。

表 3-5：不等毛植物門のミトコンドリアゲノムの比較

	パルマ藻	珪藻	ラフィド藻	褐藻	シヌラ藻	黄金色藻
解読種数	1	3	2	6	1	1
<i>rrn5</i>	—	—	+	+	—	—
tRNA ^{Gly} (GCC)	+	+	+	+	—	—
tRNA ^{Gly} (UCC)	+	—	+	—	+	+
tRNA ^{Ile} (AAU)	—	—	—	—	—	+
tRNA ^{Leu} (CAA)	—	—	—	+	—	—
tRNA ^{Arg} (UCG)	+	+	+	—	—	—
<i>trnT</i>	—	—	—	—	—	—
<i>rpl31</i>	—	—	—	+	—	—
<i>rpl5</i>	+	+	+	+	—	—
<i>rps11</i>	+	+	+	+	—	+
<i>tatC</i>	+	+	+	+	+	—
<i>cox2</i> in-frame	—	—	—	+	—	—
insertion				(<i>D. dichotoma</i> 以外)		
<i>nad11</i>	全長	*全長	全長	N 末端	全長	全長

**Phaeodactylum tricornutum* では、N 末端部分と C 末端部分に分かれて、2 つのタンパク質遺伝子となっている。

a



b

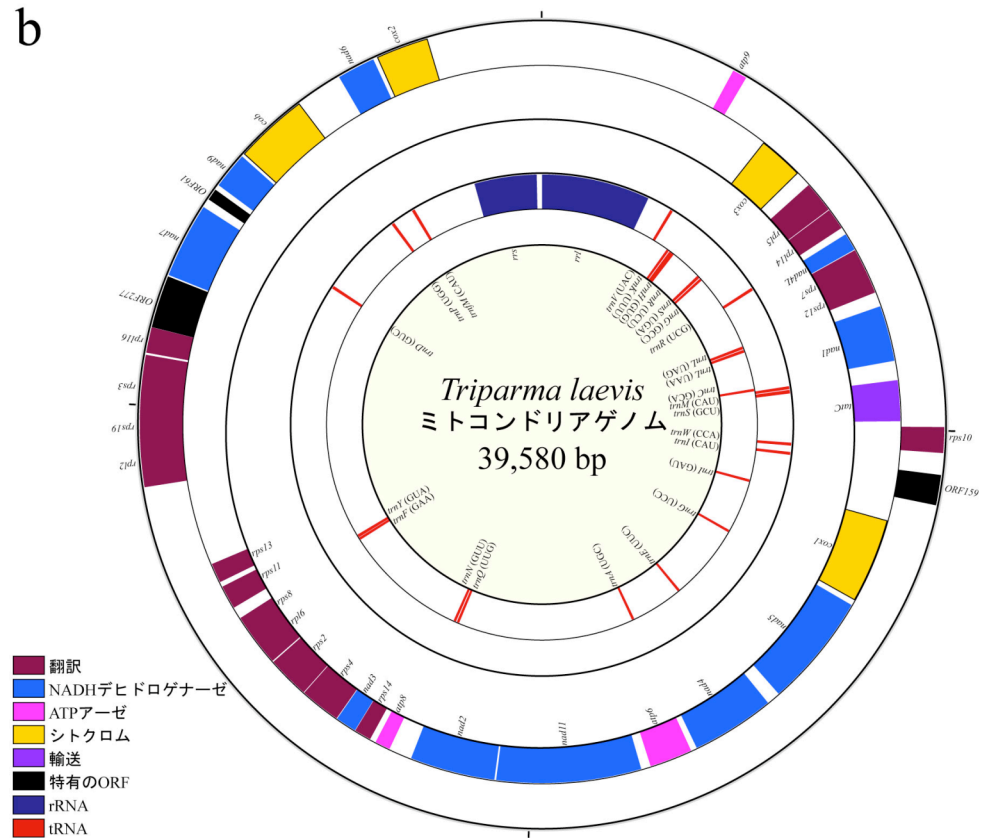


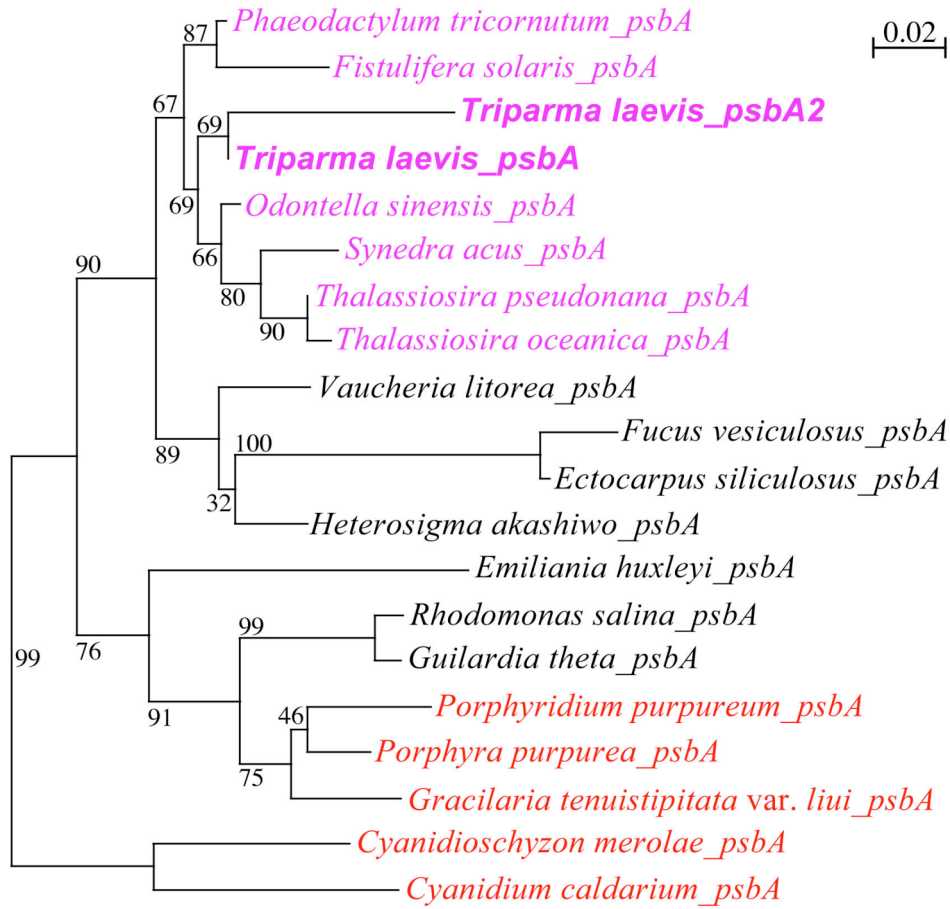
図 3-1 : *T. laevis* オルガネラゲノムの遺伝子地図(前ページ)

内側の 2 つの輪が RNA 遺伝子を示し、外側の 2 つの輪がタンパク質遺伝子を示す。それぞれ外側の輪が時計回り方向にコードされた遺伝子を示し、内側の輪が反時計回り方向にコードされた遺伝子を示す。遺伝子は、それぞれ機能毎に左下で示した色に塗り分けた。

a:葉緑体ゲノムの遺伝子地図

b:ミトコンドリアゲノムの遺伝子地図

a



b

```

psbA      MTATLERREGVSLWERFCAWITSTENRLYIGWFGCLMFPTLLTATSCYIIAFIAAPPVDI
psbA2    MTATLERNSSASLWERFALWITSTENRLYIGWFGCLMFPTLLSATAACYIIAFIAAPPVDI
*****. . . . *****. *****: ** : *****

psbA      DGIREPVAGSLLYGNNIISGAVIPSSNAIGIHFYPIWEAASVDEWLYNGGPYQLVVLHFL
psbA2    DGIREPVAGSLLYGNNIISGAVIPSSNAIGIHFYPIWEAASVDEWLYNGGPYQLVVLHFG
*****

psbA      LGVASYMGREWELSYRLGMRPWIFVAFSAPVAAASAVFLVYPIGQGSFSDGMPLGISGTF
psbA2    IGICAYMGREWELSYRLGMRPWIFVAFSAPVAAAYAVFIIYPIGQGSFSDGMPLGISGTF
: * : . : ***** ** : *****

psbA      NFMLVFQAEHNILMHPFHMAGVAGVFGGSLFSAMHGSLVTSSLIRETTENESTNYGYKFG
psbA2    NFMLVFQAEHNILMHPFHMAGVAGVFGGSLFSAMHGSLVTSSLIRETTENESTNYGYKFG
*****

psbA      QEEETYNIIVAAHGYFGRLIFQYASFNNRSLHFFLAAWPVVGIWLTSMGVSTMAFNLNGF
psbA2    QEEETYNIIVAAHGYFGRLIFQYASFNNRSLHFFLAAWPVIGIWIWLTSMGVSTMAFNLNGF
*****: ** : *****

psbA      NFNQSVVDSQGRVINTWADIINRADLGMEVMHERNAHNFPLDLASGEVLPVALTAPAING
psbA2    NFNQSVVDSQGRVINTWADIINRADLGMEVMHERNAHNFPLDLASGEVVPVALTAPAING
*****: *****

```

図 3-2 : *T. laevis* の葉緑体ゲノムにコードされた 2 コピーの *psbA* 遺伝子の系統解析とアラインメント(前ページ)

a:*psbA* 遺伝子の系統樹

右上のスケールバーは枝の長さを示す。それぞれの分岐の肩にある数字は最尤法での信頼度を示す。

b: 2 コピーの *psbA* 遺伝子のアラインメント

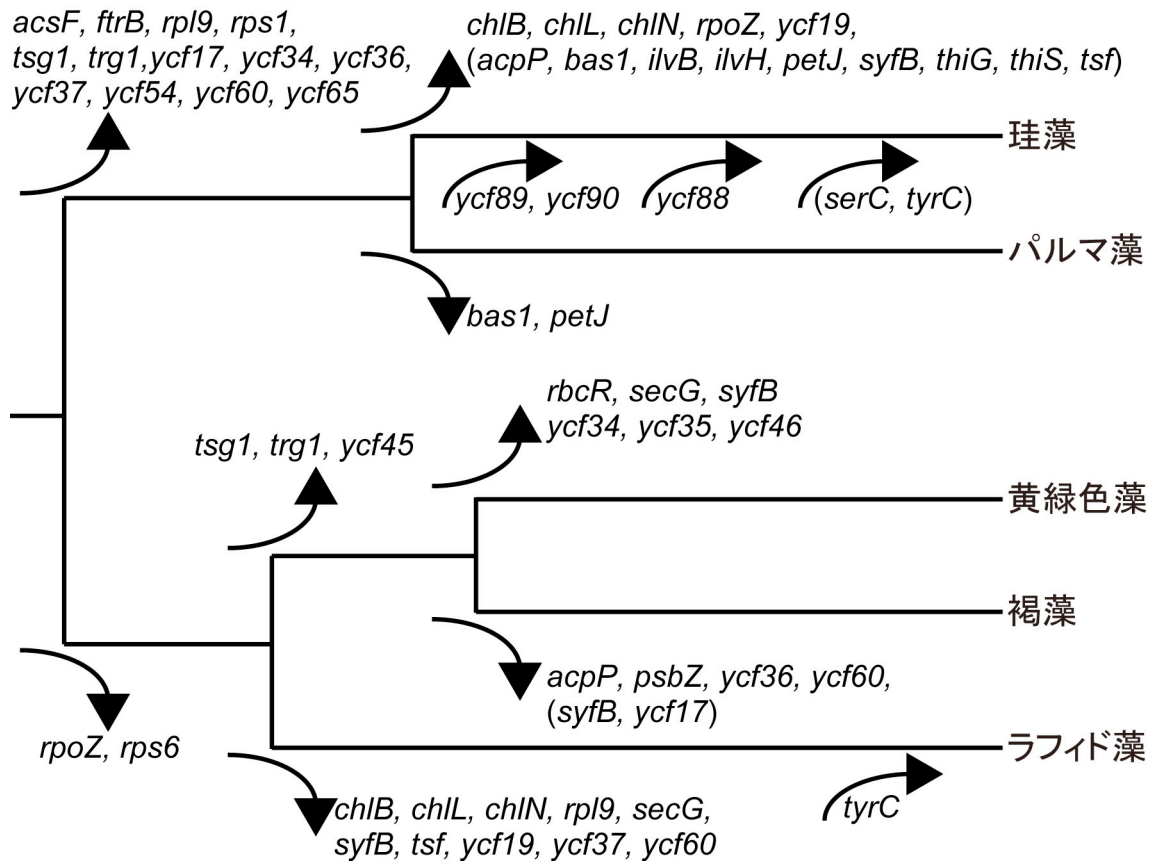


図 3-3 : 不等毛植物門の進化の過程で葉緑体ゲノムから失われたと考えられるタンパク質遺伝子と獲得したと考えられるタンパク質遺伝子
 それぞれの分岐の近くに書かれた遺伝子は、その系統が分岐した後に葉緑体ゲノムから失われたと考えられる遺伝子を示す。それぞれの枝の途中に書かれた遺伝子は、その系統が進化の過程で新しく葉緑体ゲノムに獲得したと考えられる遺伝子を示す。カッコで囲んだ遺伝子は、その系統の一部の種で失われた、もしくは獲得したと考えられる遺伝子を示す。

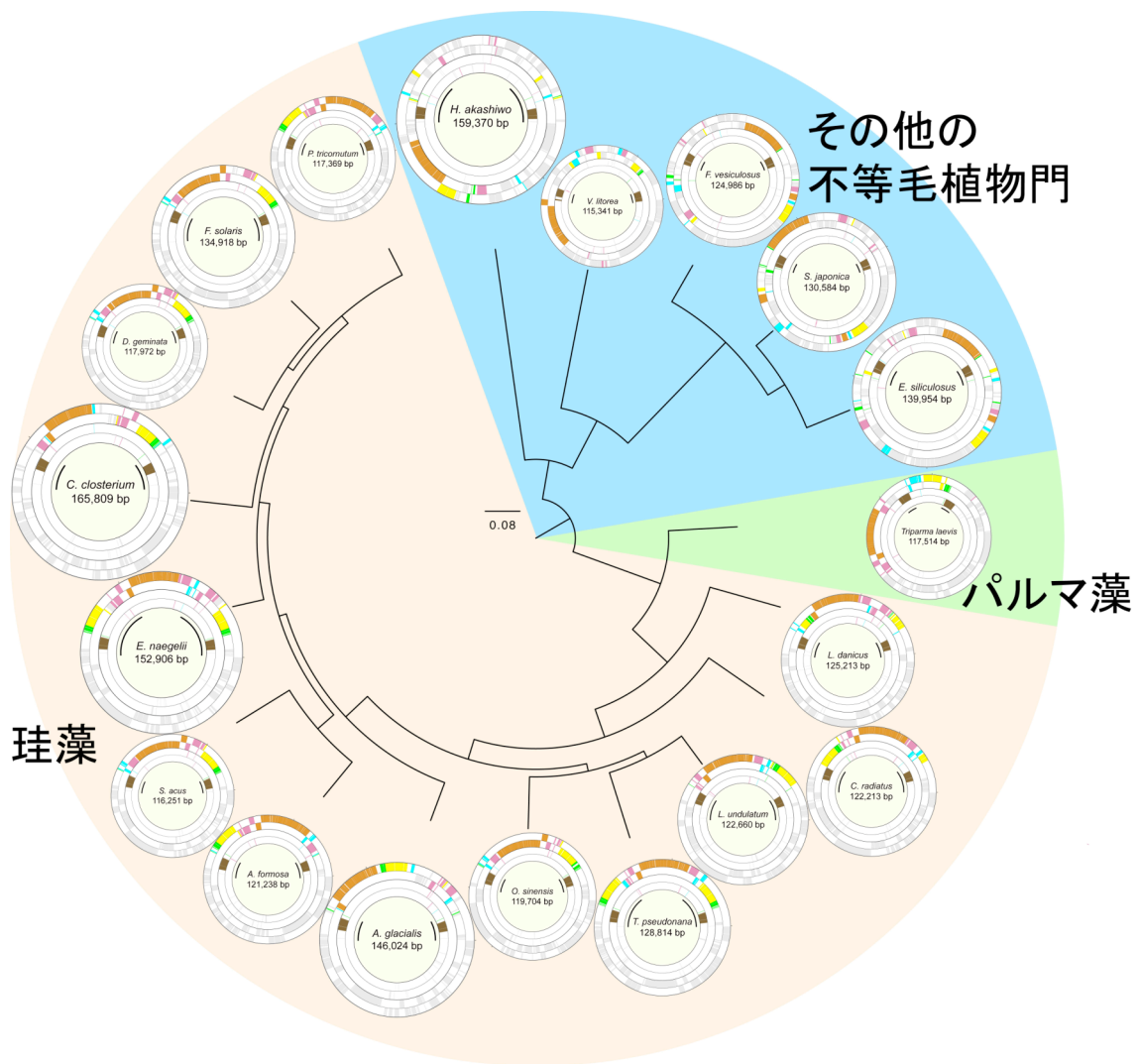


図 3-4 : 不等毛植物門の葉緑体ゲノム構造の比較

系統樹は、図 5-1b の一部を円状に表現し直したものである。スケールバーは、枝の長さを示す。遺伝子地図は葉緑体ゲノムのゲノムサイズに比例した大きさを示した。rRNA オペロンは茶色で示した。リボソームタンパク質遺伝子クラスターの塊は橙色で示した。珪藻の SSC 側にコードされたリボソームタンパク質遺伝子クラスター以外の遺伝子を、塊毎に 4 色に塗り分け、それ以外の不等毛植物門のホモログも同じ色で塗り分けた。IR 領域はそれぞれの円の最も内側に描かれた 2 本の黒い曲線で示した。

第4章

シアノバクテリア *Limnothrix* sp. ABRG5-3 の

ゲノム解読

4.1. 序文

Limnothrix は糸状性シアノバクテリアの一種である。今回解析に用いた ABRG5-3 株は熱海にある植物研究園内の池から単離された(Nishizawa et al. 2010)。この株の細胞は、長さ約 4 μm 、幅約 2 μm であり、端に節を持つ。また、運動性があり、大量に培養すると何本かまとまった集団となり、前後に動くなどの特徴がみられる。*Limnothrix* sp. ABRG5-3 の制限修飾酵素系は、*Anabaena* sp. PCC 7120 や *Synechocystis* sp. PCC 6803 と似た切断パターンを持っており、接合法での形質転換も確認されている(Nishizawa et al. 2010)。さらに、培養条件の変化により細胞が自己融解することも調べられている(Kitazaki et al. 2013)。

近年、次世代シーケンサーの発展により、多くのシアノバクテリアゲノムが解読され、ドラフトゲノムも含めて系統関係が調べられた(Shih et al. 2013)。*Limnothrix* の近縁種である *Pseudanabaena* は、タンパク質アミノ酸配列での系統解析の結果では、シアノバクテリアの進化の根元の方に位置していた。

この章では *Limnothrix* sp. ABRG5-3 のゲノムを解読し、構造やコードされた遺伝子の種類などについて、ゲノムが解読済みの他のシアノバクテリアと比較しながら述べた。

4.2. 材料と方法

4.2.1. 生物材料と培養

Limnothrix sp. ABRG5-3 は、茨城大学の朝山宗彦教授から頂いた。頂いた株は BG11 培地(Rippka et al. 1979)で培養した。水温 30°C、横から 30 $\mu\text{mol}/\text{m}^2 \cdot \text{s}$ の白色光が当たる水槽内に、試料の入った試験管を置き、1%の CO₂ を含む空気を通気し、培養を行った。

4.2.2. 試薬調整

2.2.2. で表記済みの試薬は、2.2.2. と同じ方法で作成した。

- SE buffer: NaCl 1.75 g (0.12 M)、0.5 M EDTA 25 ml (50 mM)を蒸留水に溶かし、250 ml にした。作成後、冷凍保存した。
- Lysis buffer: sucrose 2.5 g、1.0 M Tris-HCl (pH 8.0) 0.5 ml (50 mM)、0.5 M EDTA 2 ml (0.1 M)、lysozyme 20 mg を蒸留水に溶かし、10 ml にした。作成後、冷凍保存した。
- 10%サルコシル溶液: 10 g の N-ドデカノイルサルコシン酸ナトリウムを蒸留水に溶かし、100 ml にした。作成後、冷凍保存した。
- 5 M NaCl 水溶液: NaCl 2.9 g を蒸留水に溶かし、10 ml にした。作成後、冷凍保存した。

4.2.3. DNA 精製

集菌した細胞を 10 ml の SE buffer に懸濁し、8,400 g で 2 分間遠心した。沈殿に 20 ml の Lysis buffer を加えて懸濁し、37°C で 60 分間インキュベートした。1 ml の 10%サルコシル溶液と 200 μl の 20 mg/ml プロテインナーゼ K を加え、50°C で 2 時間インキュベートした。さらに 200 μl の 20 mg/ml プロテインナーゼ K を加え、50°C で一晩インキュベートした。350 μl の 5 M NaCl 水溶液を加えてからフェノールとクロロホルム液を 10 ml ずつ加え、よく混ぜた。4°C、8,400 g、10 分間遠心した。上清にクロロホルム液を 10 ml 加え、よく混ぜた。4°C、8,400 g、2 分間遠心した。以降は 3.2.3. と同様の手順で超遠心により DNA バンドを回収、2-プロパノールと透析により DNA を精製した。ただし、一度目の超遠心の時間は 6 時間とした。

4.2.4. DNA 配列解析

2.2.4.と同様にゲノムの塩基配列解析を行い、サンガー法によりギャップを埋めた。*Limnothrix* sp. ABRG5-3 の塩基配列解析については Illumina 社の Genome analyser II も利用した。また、サンガー法による塩基配列解析の一部はタカラバイオ社に依頼した。

4.2.5. ゲノムアノテーションとデータ解析

2.2.5.と同様に、まずタンパク質遺伝子を推定するために、MetaGeneAnnotator を利用した。MetaGeneAnnotator では見つからなかったが、開始コドンから十分な長さを持ち、タンパク質遺伝子をコードしていると思われる領域も推定した。逆に、推定した tRNA 遺伝子や rRNA 遺伝子と領域がかぶってしまったものは削除した。得られたタンパク質配列を、NCBI のタンパク質データベースに対して blastp を行い、アノテーションを付けた。アノテーションを付けるために、*Limnothrix* sp. ABRG5-3 および 66 種のシアノバクテリアゲノムと 16 種のバクテリアゲノムを用いた CyanoClust によるホモログ解析の結果も利用した。アミノ酸配列が短いため、ここまでの検索で見つけることはできなかったが、多くの他の種が持っているので、*Limnothrix* sp. ABRG5-3 も持っているかもしれないと思われたタンパク質遺伝子に対しては、該当するタンパク質遺伝子のホモログが存在する CyanoClust のクラスターにある全てのタンパク質遺伝子のアミノ酸配列を抜き出し、*Limnothrix* sp. ABRG5-3 の塩基配列に対して tblastn を行った。逆に、他種のシアノバクテリアゲノムに対しても、登録されていなくても存在するはずと思われたタンパク質遺伝子に対しては、同様の検索法で探した。そのため、表 4-4 で示したタンパク質遺伝子数は実際に登録されている数より多くなっている種も存在する。tRNA 遺伝子の推定には 2.2.5.と同様に tRNAscan-SE を利用し、その他の RNA 遺伝子は比較解析に利用した他のシアノバクテリアの該当配列を抜き出し、*Limnothrix* sp. ABRG5-3 の塩基配列に対して blastn を行うことで推定した。CRISPR-Cas の推定には、CRISPRFinder (Grissa et al. 2007)を利用した。DNA およびタンパク質配列の加工には SISEQ ver. 1.59 を利用した。ゲノム配列の編集は Artemis ver. 13.0 を利用した。

4.2.6. 窒素固定関連遺伝子群の系統解析

系統解析のために、21種のシアノバクテリアと2種の紅色光合成細菌に対して、共通してもつ窒素固定関連遺伝子群のうち *nifB*、*nifD*、*nifE*、*nifH*、*nifK*、*nifN*、*nifS*、*nifT*、*nifU*、*nifV*、*nifW*、*nifX*、*nifZ*、*d34*(プラスミド *d* にコードされ、登録上の1 bp から数えて34番目のタンパク質遺伝子) のホモログを用いてタンパク質配列のアラインメントを作成した。系統樹の計算には最尤法を用いて、モデルはLGモデルを選んだ。最尤法の計算には TREEFINDER ver. March 2011 を利用した。

4.2.7. シアノバクテリアの系統解析

系統解析のために、16S rRNA 遺伝子の塩基配列とタンパク質遺伝子アミノ酸配列のアラインメントを作成した。

タンパク質遺伝子アミノ酸配列のアラインメントは、まず、選んだ生物種に共通してコードされた個々のタンパク質のアミノ酸配列に対してアラインメントを作成し、最後に、個々の作成したアラインメントを一つに結合したものを利用した。66種のシアノバクテリアゲノムに共通してコードされた31個のタンパク質遺伝子で系統解析を行った。タンパク質配列の加工には SISEQ ver. 1.59 を利用した。アラインメント作成は ClustalW/ClustalX ver. 1.8.3、もしくは Muscle ver. 3.6 を利用した。系統樹の計算には最尤法とベイズ法を利用した。最尤法の計算には TREEFINDER ver. March 2011 を利用した。最適な進化モデルの選択は ProtTest ver. 2.4 (Abascal et al. 2005) を利用し、LGモデルを選んだ。ベイズ法ではLGモデルでの計算ができなかったため、WAGモデルで計算を行った。計算には MrBayes ver. 3.2.1 (Ronquist & Huelsenbeck 2003) を利用した。2連鎖で計算させ、100万世代のうち200世代ごとにサンプリングし、最初の2000サンプルは burn-in として除外した。

16S rRNA 遺伝子のアラインメントは、まず *Limnothrix* sp. ABRG5-3 以外のシアノバクテリアの16S rRNA 遺伝子塩基配列を Ribosomal Database Project (RDP) (Cole et al. 2014) から手に入れた。RDP には16S rRNA 遺伝子の塩基配列が16S rRNA 遺伝子の二次構造を考慮して作られたアラインメントの形で登録されている。手に入れたアラインメントから、共通しているギャップを取り除いた上で *Limnothrix* sp. ABRG5-3 の16S rRNA 遺伝子塩基配列を加え、

元のアラインメントに沿ったアラインメントを作成した。アラインメント作成は ClustalX ver. 2.0.9 を利用した。アラインメントは最尤法と近隣結合法により、系統樹の計算を行った。最尤法の計算には TREEFINDER ver. June 2008 を利用した。最適なモデルの選択も TREEFINDER で求め、GTR+G+I モデルで計算を行った。近隣結合法の計算には MEGA ver. 6.06 (Tamura et al. 2013) を利用した。モデルは Maximum Composite Likelihood で、レートは Uniform、パターンは Heterogeneous にした。ギャップの扱いに対しては、解析中必要があれば考慮しなくなる pairwise deletion を選んだ。

4.3. 結果

4.3.1. *Limnothrix* sp. ABRG5-3 のゲノム構造

Limnothrix sp. ABRG5-3 のゲノムを Roche Diagnostics 社の genome sequencer FLX instrument および Illumina 社の Genome analyser II を用いて解読した。

プラスミドゲノムを含めて合計で 223 個のコンティグにアセンブルされ、ギャップはサンガー法を用いて解読した。*Limnothrix* sp. ABRG5-3 のメインゲノムは、全長 4,796,642 bp、GC 含量 43.2%の環状 DNA だった。4 個の ncRNA 遺伝子(*rnpB*、*ssaA*、*ffs*、cyanobacterial functional RNA)、9 個の rRNA 遺伝子(3 個の *rrs*、3 個の *rri*、3 個の *rrf*)、52 個の tRNA 遺伝子および 4,318 個のタンパク質遺伝子をコードすることが分かった。7 種のプラスミドのゲノムも全て解読した(表 4-1、表 4-2)。

4.3.2. タンパク質遺伝子の機能分類

Limnothrix sp. ABRG5-3 のゲノムにコードされた全てのタンパク質遺伝子について、blastp の結果と、この株を含めたシアノバクテリア 67 種とバクテリア 16 種の計 83 種を用いた Gclust による相同クラスタリング解析の結果から、タンパク質遺伝子のアノテーションを付け、機能ごとに分類した(表 4-3)。全体の約 3 分の 1 は hypothetical protein であった。88 個の遺伝子はトランスポゼースと推定された。*Limnothrix* sp. ABRG5-3 は、いくつかのシアノバクテリアやバクテリアが細胞内に持ち、浮力の調整に用いる袋状でタンパク質性の構造体である gas vesicle (Walsby 1994)に関わるタンパク質遺伝子群をコードしていることが分かった。*Limnothrix* sp. ABRG5-3 の水平方向の運動について、Type IV pili 線毛による twitching motility に関わる遺伝子を多くコードしていることが分かった(Maruriello et al. 2010)。一方で、粘液を噴射しその推進力によって移動すると考えられている gliding motility に関わる遺伝子のホモログ(Cozy & Callahan 2014, Risser et al. 2014, Risser & Meeks 2013)は、ほとんど見つけることができなかった。

4.3.3. プラスミドゲノムにコードされた遺伝子

7 種のプラスミドゲノムにコードされた遺伝子は全てタンパク質遺伝子だ

った(表 4-1、表 4-3)。プラスミドや染色体の分裂に関わる *parA-parB* 遺伝子は、プラスミド g を除く全てのプラスミドにそれぞれコードされていた。プラスミド g には *parB* のホモログは無く、*parA* のホモログのみコードされていた。プラスミド a とプラスミド b のうち約 30 組の遺伝子は、ほとんどのアノテーションは hypothetical protein であるが互いにホモログであり、順番も保存されている領域があった。プラスミド a はヘリカーゼやトランスポゼーを多くコードしていた。また、機能未知であるが TPR ドメインを持つタンパク質も多くコードされていた。プラスミド b とプラスミド c は物質輸送に関わる遺伝子が多くコードされており、プラスミド b は銅や亜鉛、ニッケルやコバルト等重金属の排出に関わる遺伝子、プラスミド c はカドミウム排出に関わる遺伝子や ABC transporter をコードしていた。プラスミド d は窒素固定に関わるニトロゲナーゼ関連遺伝子を多くコードしていた(図 4-1)。これらの遺伝子の並びは、ヘテロシストは作らないが、窒素欠乏時に窒素固定を行う *Leptolyngbya boryana* dg5 の窒素固定遺伝子クラスターと似ていた(図 4-2) (Tsujiimoto et al. 2014)。さらに、窒素固定遺伝子群を調節する大元となる遺伝子 *cnfR* のホモログもコードしていることが分かった。*nifE* と *nifN* は一つのタンパク質遺伝子としてつながっており、*nifEN* と名付けた(図 4-3)。同様の遺伝子構造は *Trichodesmium erythraeum* IMS101 や *Chroococcidiopsis thermalis* PCC 7203 等でも見られる。窒素固定関連遺伝子を持つ種はヘテロシストを形成しないシアノバクテリアでも単細胞多細胞を問わずいくつか見られるが、いずれもメインゲノム上にコードしており、プラスミドゲノムにコードしている種は *Limnothrix* sp. ABRG5-3 のみだった。窒素固定関連遺伝子について系統解析を行った結果、本来近縁種である *Pseudanabaena* sp. PCC 6802 よりも *Cyanothece* sp. PCC 7425 と *Leptolyngbya boryana* dg5 に近いことが分かった(図 4-4)。プラスミド e は外来遺伝子からの防御系として知られている CRISPR-Cas 系をコードしていた(図 4-5)。検索の結果、CRISPR 領域はメインゲノム上にも 3 カ所存在していたが、Cas タンパク質群のホモログは見られなかった。プラスミド e の CRISPR 領域は、直列反復領域は 37 bp で、スペーサーは 74 個あると推定された。プラスミド f はプラスミド保持に関わる遺伝子(plasmid maintenance system antidote protein, plasmid maintenance system killer protein)や、アミノ酸アデニル化に関わる遺伝子をコードしていた。プ

ラスミド g は *parA* と 1 個の ATPase 以外は hypothetical protein だった。

4.3.4. シアノバクテリアの比較

Limnothrix sp. ABRG5-3 とゲノムが解読されたうち 65 種のシアノバクテリアを比較した(表 4-4)。メインゲノムの総塩基対数とタンパク質遺伝子数の間には正の相関関係が見られた(図 4-6)。多細胞シアノバクテリアは単細胞シアノバクテリアよりも塩基対数が長い傾向にあった。*Prochlorococcus/Synechococcus* のグループは他の単細胞シアノバクテリアよりも総塩基対数が短く、コードしているタンパク質遺伝子数も少ない傾向があった。総塩基対数約 6 Mbp までのシアノバクテリアの tRNA 遺伝子数は、約 45 個であり、それ以上長くなるとそれ以上の個数をコードする種が現れるようになった。ただし、一定でなくなった部分だけでも総塩基対数と tRNA 遺伝子数の間に正の相関関係は見られなかった。*Limnothrix* sp. ABRG5-3 は、tRNA 遺伝子数と総塩基対数の関係に当てはめて考えると、総塩基対数(約 4.8 Mbp)から推測されるより多くの tRNA 遺伝子(52 個)をコードしていることが分かった。総塩基対数と rRNA オペロン数の間にも強い正の相関関係は見られなかった。

4.3.5. シアノバクテリアの系統関係

Limnothrix sp. ABRG5-3 と 65 種のシアノバクテリアを用いて、16S rRNA 遺伝子の塩基配列と、共通してコードしているタンパク質遺伝子のうち 31 個のタンパク質遺伝子のアミノ酸配列で、それぞれ系統解析を行った(図 4-7)。

タンパク質遺伝子アミノ酸配列の系統解析結果では、*Limnothrix* sp. ABRG5-3 は *Synechococcus* sp. PCC 7502 と *Pseudanabaena* sp. PCC 7367 で 1 つの系統となっていた。系統樹の樹形はベイズ法でも 1 カ所を除いて完全に支持された。

16S rRNA 遺伝子の系統樹の樹形は、*Limnothrix* sp. ABRG5-3 の系統よりも *Geitlerinema* sp. PCC 7407 と *Prochlorococcus/Synechococcus* のグループが外側に位置することや、*Synechocystis* sp. PCC 6803 が *Cyanothece* よりも *Synechococcus* sp. PCC 7002 に近い位置に存在することなど、タンパク質遺伝子アミノ酸配列の系統樹と大きく異なっていた。さらに、最尤法の結果と近隣結合法の結果の間でも違いが数多く見られ、最尤法の信頼度や近隣結合法

のブートストラップ値が低い分岐も多かったため、16S rRNA 遺伝子の系統解析から系統関係を明らかにすることはできなかった。

より実際の系統関係を表現しているであろうと考えられるタンパク質遺伝子アミノ酸配列での系統解析結果を基にシアノバクテリアゲノムを4つのグループに分けた(表 4-4、図 4-7a)。*Prochlorococcus/Synechococcus* のグループは A2 として別グループにした。このグループ分けに従い、シアノバクテリアゲノムの比較を行った。

4.3.6. シアノバクテリアにコードされたタンパク質遺伝子のホモログ解析

アノテーション付けにも用いた Gclust による相同クラスタリング解析の結果から、タンパク質遺伝子のホモログ解析を行った。*Prochlorococcus/Synechococcus* のグループ(表 4-4 の A2)はコードしている遺伝子数が特に少なく、A グループと A2 グループをまとめてしまうと、そのグループが共通してコードしているタンパク質のクラスター数を少なく見積もることになる可能性が高いため、今回の解析からは除外した。A グループは全体の 9 割以上の種(36 種以上)が存在する相同クラスタを数え、その他のグループは全種が存在する相同クラスタを数えた。共通して存在するグループが無い区分(---)に、1 クラスタに 1 遺伝子のみのクラスタの数は含めなかった(表 4-5)。

911 個の相同クラスタは全てのシアノバクテリアグループが共通して存在していた。B グループのみ共通して存在する相同クラスタは 176 個で、逆に B グループ以外が共通して存在する相同クラスタは 50 個あった。

共通して存在するグループの有無によってクラスタを区分けし、それぞれの区分に含まれる *Limnothrix* sp. ABRG5-3 の遺伝子を機能により分類した。全てのシアノバクテリアグループが共通して存在する 911 個のクラスタに、*Limnothrix* sp. ABRG5-3 の遺伝子は 1489 個存在した。ID205 は全てのシアノバクテリアグループに共通して存在していた。ID105、ID106、ID107、ID201、ID202、ID204 に関わるタンパク質遺伝子のほとんどは全てのシアノバクテリアグループに共通して存在していた。Type IV pili 線毛による twitching motility に関わる遺伝子の半分以上も全てのシアノバクテリアグループに共通して存

在していた(ID304)。ID305 の全ての遺伝子は B グループには共通して存在していた。B, C, D グループのみに共通して存在する ID103 の遺伝子はほとんどがトランスポゼースだった。ID203 はあまり他のグループにも共通して存在はしていなかった。*Limnothrix* sp. ABRG5-3 の 589 個の遺伝子は 1 クラスターに 1 遺伝子のみのクラスター(unique)に属し、そのうち ID401 は 311 個だった。

4.3.7. シアノバクテリアゲノムの rRNA オペロン

2.3.1.と同様に、シアノバクテリアゲノムの rRNA オペロン構造を比較した(図 4-8)。多くは 16S rRNA 遺伝子、tRNA^{Ile}(GAU)遺伝子、tRNA^{Ala}(UGC)遺伝子、23S rRNA 遺伝子、5S rRNA 遺伝子の順でひとまとまりになっていた。*Synechocystis* sp. PCC 6803 のように、16S rRNA 遺伝子と 23S rRNA 遺伝子の間に tRNA^{Ile}(GAU)遺伝子のみコードしている種や、*Crinalium epipsammum* PCC 9333 のように 16S rRNA 遺伝子と 23S rRNA 遺伝子の間に tRNA 遺伝子が全く無い種も存在していた。*Oscillatoria acuminata* PCC 6304 や *Microcoleus* sp. PCC 7113 のように、16S rRNA 遺伝子と 23S rRNA 遺伝子の間に tRNA 遺伝子が無い rRNA オペロンと、ある rRNA オペロンが混じっている種も見られた。*Trichodesmium erythraeum* IMS101 のみ 5S rRNA 遺伝子が rRNA オペロンとして存在していなかった。また、*T. erythraeum* IMS101 などいくつかの種では、rRNA オペロンと離れた位置に 5S rRNA 遺伝子のみ存在している種もあったが、図には加えていない。*Limnothrix* sp. ABRG5-3 は 3 つの rRNA オペロンを持ち、1 つは他の 2 つと逆向きになっていた。*Synechococcus* sp. PCC 7502 と *Pseudanabaena* sp. PCC 7367 は互いに逆向きの 2 つの rRNA オペロンを持っていた。種毎に rRNA オペロンの構造、個数、向きやゲノム上の位置関係は異なり、特にヘテロシストを形成するシアノバクテリアの間では、違いは顕著だった。

4.3.8. シアノバクテリアのリボソームタンパク質遺伝子クラスター

2.3.5.と同様に、リボソームタンパク質遺伝子クラスターを比較した(図 4-9)。光合成をしないバクテリアのゲノムと比較するため、*Escherichia coli* str. K-12 と *Bacillus subtilis* subsp. *subtilis* str. 168 についても調べた。バクテリアゲノム

の *rpsE* と *rplO* の間にある *rpmD* はシアノバクテリアゲノムにコードされていなかった。*rpsJ* はシアノバクテリアゲノムでは *tufA* と隣接していたが、バクテリアゲノムでは *rplC* と隣接していた。*rpsD* と *rpsN* は *Gloeobacter violaceus* PCC 7421 と *E. coli* ではリボソームタンパク質遺伝子クラスター内に組み込まれていたが(*B. subtilis* は *rpsN* のみ)、他のシアノバクテリアゲノムでは遺伝子クラスターからは外れて種毎に異なる位置に存在していた。Cグループと *B. subtilis* では *adk* と *infA* の間にある *map* は、他のシアノバクテリアゲノムにもホモログは存在するが、リボソームタンパク質遺伝子クラスターからは外れており、種によっては2コピー以上コードしていた。Aグループでは *rplC-rpmE* は共通して1つの遺伝子クラスターとなっていた。例外として、*Microcoleus* sp. PCC 7113 は *infA* を2コピー連続でコードしており、*Acaryochloris marina* MBIC11017 は *infA* と *rpmJ* の間で分断されていた。*rpsL-rpsJ* の *rplC-rpmE* に対する位置関係や向きは種毎に異なっていた。Aグループでは *adk* と *rpmJ* の間にある *infA* がA2グループではリボソームタンパク質遺伝子クラスターから外れていた。Cグループは *A. marina* MBIC11017 と同様に *infA* と *rpmJ* の間で分断されていた。Bグループはそれに加えて *rpmJ* と *rpsM*、*rpsI* と *rpmE* の間が分断され、さらに *Synechococcus* sp. PCC 7502 と *Limnothrix* sp. ABRG5-3 では *tufA* と *rpsJ* の間も分断されていた。*G. violaceus* PCC 7421 は *rplW* と *rplB*、*rpsS* と *rplV*、*rplO* と *secY*、*adk* と *infA*、*infA* と *rpmJ*、*truA* と *rplM* の間で分断され他のシアノバクテリアゲノムと比べてバラバラになっていた。

4.4. 考察

Limnothrix sp. ABRG5-3 のゲノムを解読し、その構造やコードされたタンパク質遺伝子の機能等について調べ、他のシアノバクテリアゲノムと比較した。

コードしているタンパク質遺伝子から、*Limnothrix* sp. ABRG5-3 の水平移動は、Type IV pili 線毛の twitching motility によるものではないかと推定された。

Limnothrix sp. ABRG5-3 はそれぞれ異なる役割のプラスミドを 7 種持っていた。特にプラスミド d は、窒素固定に関わる遺伝子群をクラスターとしてまとめて保存していた。プラスミド d は、phage tail protein のアノテーションがつく多くのタンパク質遺伝子と、ヘテロシスト形成時には消失する光化学系 II (Wolk et al. 2004)に関わる *psbA* のホモログもコードしていた(図 4-1)。これらの遺伝子も窒素固定の働きに関わっているのかは分からない。今後、実際にこの株が窒素欠乏時に窒素固定が可能であるか調べた上で、特定のプラスミドを抜き出し別種のシアノバクテリアにそのまま移すことができるようになれば、あらゆるシアノバクテリアで窒素欠乏時に窒素固定が可能になるかもしれない。

シアノバクテリアの系統解析を行った結果、B グループは、タンパク質遺伝子アミノ酸配列、16S rRNA 遺伝子塩基配列どちらの結果でも単系統となることが分かった。16S rRNA 遺伝子塩基配列の場合、*A. marina* MBIC11017 や *Cyanotheca* sp. PCC 7425 とも単系統になった。16S rRNA 遺伝子塩基配列の系統解析結果では、*Prochlorococcus/Synechococcus* のグループはシアノバクテリアの進化の根元側に位置する傾向がある(Falcón et al. 2010, Shih et al. 2013, Urbach et al. 1992)。タンパク質遺伝子アミノ酸配列での解析結果の方が、解析方法の違いで樹形が乱されないため、より正しくシアノバクテリアの系統関係を表しているのではないかと考えた。

rRNA オペロンの構造を比較した結果、種毎に大幅に異なることが分かった。rRNA オペロンの数は、タンパク質遺伝子のように総塩基対数に比例しているわけではなく、総塩基対数が長くなるにつれて増えるものではないようである。特に、ヘテロシストを形成するシアノバクテリアの場合、近縁種のゲノムが解読済みであっても、それを基に rRNA オペロンの個数や向きを推定することはできないと考えられる。16S rRNA 遺伝子と 23S rRNA 遺伝子

の間の tRNA 遺伝子も、進化に伴い何度も 16S rRNA 遺伝子と 23S rRNA 遺伝子の中から抜けて別の場所にコードされるようになっているが、どのような仕組みが働いているかは分からない。

rRNA オペロンの構造に比べると、A グループでのリボソームタンパク質遺伝子クラスターの構造は保存されていた。*A. marina* MBIC11017 や *E. coli*、B, C, D グループを見ると、*infA* と *rpmJ* の間は比較的分断されやすく、A2 グループの *infA* が遺伝子クラスターから外れていたのも、A2 グループの共通祖先の進化の過程で、一度 *adk* と *infA* と *rpmJ* が分断され、その後 *infA* は他の場所に移り、*adk* と *rpmJ* が再びつながるようになったのではないかと推察した。

表 4-1 : *Limnothrix* sp. ABRG5-3 ゲノムの塩基対数とコードされた遺伝子の数

	塩基対数 (bp)	タンパク質	tRNA	rRNA	ncRNA
メインゲノム	4,796,642	4,318	52	9	4
プラスミド a	214,764	137	0	0	0
プラスミド b	152,084	126	0	0	0
プラスミド c	134,662	125	0	0	0
プラスミド d	125,503	109	0	0	0
プラスミド e	111,135	99	0	0	0
プラスミド f	110,608	53	0	0	0
プラスミド g	12,574	17	0	0	0

表 4-2 : *Limnothrix* sp. ABRG5-3 メインゲノムの tRNA 遺伝子の同定

First base	Second base						Third base				
	U		C		A			G			
U	3032664-3032591, 3590338-3590266	Phe	894039-893953	Ser	4063886-4063805	Tyr	2077255-2077185	Cys	U		
									C		
	2299134-2299216 1836107-1836190	Leu	1536928-1537012							A	
			1186871-1186957				3650766-3650694	Trp	G		
C	3936561-3936481	Leu	3420900-3420973	Pro	1420975-1420902	His	999028-999101	Arg	U		
									C		
	2688430-2688510, 3032456-3032384 2284798-2284717		3032342-3032271, 4330975-4331048		2629647-2629576	Gln	808955-809028		A		
			G								
A	847392-847465, 4563343-4563270, 4794737-4794664	Ile	732064-732135	Thr	1431137-1431066, 3032541-3032469	Asn	4556379-4556288	Ser	U		
									C		
	(I)2607838-2607767, *3032199-3032126, *3794699-3794772, 4170381-4170308	Met, Ile	3031987-3031915, 4033735-4033807		4251019-4250948	682956-683027	Lys		2116089-2116162, 3031255-3031184	2006981-2007054	A
											G
G	2085948-2085875	Val	1345641-1345569	Ala	4157587-4157660	Asp	3434401-3434472	Gly	U		
									C		
	1478640-1478568		847516-847588, 4563219-4563147, 4794613-4794541		1985210-1985282, 1985373-1985445	2893485-2893413	792653-792723		Glu	A	
										G	

数字はゲノム上での tRNA 遺伝子の領域を示す。アスタリスクは tRNA^{fMet}(CAU)を示し、(I)は tRNA^{Ile}(CAU)を示す。

表 4-3 : *Limnothrix* sp. ABRG5-3 にコードされたタンパク質遺伝子の機能分類

機能 ID	機能分類	メイン ゲノム	プラスミド							合計
			a	b	c	d	e	f	g	
101	ゲノム構造	55	11	4	1	1	0	3	0	75
102	複製	24	2	1	1	0	1	1	0	30
103	修復、遺伝子組み換え、修飾、ヌクレアーゼ	221	21	9	4	2	14	1	0	272
104	転写	128	2	5	6	4	5	1	0	151
105	リボソーム、翻訳	140	1	0	0	0	0	0	0	141
106	RNA 代謝	41	0	0	0	1	0	0	0	42
107	ソーティング	23	0	0	1	0	0	0	0	24
108	加工と分解	87	5	0	4	0	1	1	0	98
109	転写制御	92	2	4	1	1	2	2	0	104
110	シグナル伝達	359	8	7	5	7	6	2	0	394
111	ストレス反応とシャペロン	99	1	1	6	2	1	0	0	110
201	光合成	63	0	0	0	1	0	0	0	64
202	呼吸	29	0	0	0	0	1	0	0	30
203	ヒドロゲナーゼ	16	0	0	0	0	0	0	0	16
204	炭素同化	17	0	0	0	0	0	0	0	17
205	ATP 合成酵素	11	0	0	0	0	0	0	0	11
206	糖と中央代謝	131	3	1	1	0	1	2	0	139
207	脂質代謝	78	1	1	2	3	7	12	0	104
208	ポルフィリン、ヘム、シトクロム、色素代謝	106	0	0	0	1	0	0	0	107
209	リンとイオウ代謝	30	0	1	4	1	0	0	0	36
210	窒素とアミノ酸代謝	176	0	1	9	20	2	1	0	209
211	核酸代謝	53	1	1	2	2	0	1	0	60
212	補因子合成	129	0	3	0	3	1	0	0	136
213	その他の代謝	188	2	7	4	1	2	4	1	209
301	細胞外物質	125	2	3	4	2	0	5	0	141
302	細胞分裂	27	2	2	2	2	3	2	1	41
303	輸送と膜	304	0	14	27	6	19	1	0	371
304	運動	20	0	0	0	0	0	0	0	20
305	gas vesicle protein	9	0	0	0	0	0	0	0	9
400	未分類	114	15	1	4	17	1	3	0	155
401	hypothetical protein	1,423	58	60	37	32	32	11	15	1,668
合計		4,318	137	126	125	109	99	53	17	4,984

100 番台はゲノムそのものに関わる機能、200 番台は代謝関係の機能、300 番台は細胞の構造や運動に関わる機能を示す。

表 4-4：66 種のシアノバクテリアのゲノム情報

グループ	タイプ	総塩基対数 (bp)		ncRNA 遺伝子の個数	tRNA 遺伝子の個数	rRNA オペロン数	プラスミド数	アクセション番号
		メインゲノム	全ゲノム					
A	多細胞 (filamentous)	7,750,108	7,750,108	4,460	38	2	0	NC_008312
A	多細胞 (filamentous)	7,479,014	8,272,254	5,783	71	2	5	NC_019729
A	多細胞 (filamentous)	7,689,443	7,804,270	5,704	66	3	2	NC_019693
A	多細胞 (heterocystous)	6,767,834	6,960,392	5,393	41	4	3	NC_019751
A	多細胞 (heterocystous)	6,635,030	6,718,869	5,357	45	5	2	NC_019684
A	多細胞 (heterocystous)	6,413,771	7,211,789	5,365	48	4	6	NC_003272
A	多細胞 (heterocystous)	6,402,878	7,105,752	5,094	47	4	3	NC_007413, NC_014000
A	多細胞 (heterocystous)	6,329,823	6,329,823	5,243	78	4	0	NC_019676
A	多細胞 (heterocystous)	7,023,215	7,023,215	5,950	72	3	0	NC_019682
A	多細胞 (heterocystous)	8,234,322	9,059,191	6,090	88	4	5	NC_010628
A	多細胞 (heterocystous)	7,003,560	7,610,589	5,815	69	4	3	NC_019757
A	多細胞 (heterocystous)	6,395,836	7,063,285	5,325	61	4	6	NC_019771
A	多細胞 (heterocystous)	5,354,700	5,486,145	3,591	44	4	2	NC_014248
A	多細胞 (heterocystous)	5,149,229	5,305,675	4,354	44	5	3	NC_019427, NC_019439
A	多細胞 (heterocystous)	8,698,463	8,728,773	6,611	52	3	2	NC_019678
A	単細胞	5,431,448	5,882,710	4,685	41	2	4	NC_019745
A	単細胞 (baeocystous)	6,315,792	6,689,401	5,408	46	3	2	NC_019695
A	多細胞 (filamentous)	5,315,554	5,620,407	4,476	43	4	8	NC_019753
A	単細胞	6,284,095	6,761,765	5,498	54	3	2	NC_019697
A	多細胞 (filamentous)	7,470,429	7,966,510	6,029	70	3	8	NC_019738
A	単細胞	3,781,008	3,781,008	3,337	42	3	0	NC_019780
A	単細胞	4,179,170	4,179,170	3,710	46	3	0	NC_019779
A	単細胞	5,363,972	5,460,377	5,210	43	2	4	NC_010546, NC_010547
A	単細胞	4,669,813	4,803,347	4,320	42	2	4	NC_013161
A	単細胞	4,679,413	4,787,694	4,263	43	2	3	NC_011726
A	単細胞	3,571,103	3,944,652	3,170	42	2	4	NC_017277
A	単細胞 (baeocystous)	4,986,817	4,986,817	4,268	42	3	0	NC_019689
A	単細胞	6,091,620	7,841,948	5,424	47	3	6	NC_014501
A	単細胞	5,942,652	6,554,169	5,227	44	3	6	NC_011729
A	単細胞 (baeocystous)	5,041,209	5,544,990	4,328	43	2	5	NC_019748
A	単細胞	3,008,047	3,409,935	2,830	42	2	6	NC_010475
A	多細胞 (filamentous)	5,125,950	5,125,950	4,228	44	2	0	NC_019683
	<i>Trichodesmium erythraeum</i> IMS101							
	<i>Oscillatoria nigro-viridis</i> PCC 7112							
	<i>Oscillatoria acuminata</i> PCC 6304							
	<i>Calothrix</i> sp. PCC 6303							
	<i>Nostoc</i> sp. PCC 7524							
	<i>Anabaena</i> sp. PCC 7120							
	<i>Anabaena variabilis</i> ATCC 29413							
	<i>Nostoc</i> sp. PCC 7107							
	<i>Calothrix</i> sp. PCC 7507							
	<i>Nostoc punctiforme</i> PCC 73102							
	<i>Cylindrospermum stagnale</i> PCC 7417							
	<i>Anabaena cylindrica</i> PCC 7122							
	<i>Nostoc azollae</i> 0708							
	<i>Anabaena</i> sp. 90							
	<i>Rivularia</i> sp. PCC 7116							
	<i>Gloeocapsa</i> sp. PCC 7428							
	<i>Chroococcidiopsis thermalis</i> PCC 7203							
	<i>Crinatium epipsammum</i> PCC 9333							
	<i>Chamaesiphon minutus</i> PCC 6605							
	<i>Microcoleus</i> sp. PCC 7113							
	<i>Dactylococcopsis salina</i> PCC 8305							
	<i>Halothece</i> sp. PCC 7418							
	<i>Cyanothece</i> sp. ATCC 51142							
	<i>Cyanothece</i> sp. PCC 8802							
	<i>Cyanothece</i> sp. PCC 8801							
	<i>Synechocystis</i> sp. PCC 6803							
	<i>Pleurocapsa</i> sp. PCC 7327							
	<i>Cyanothece</i> sp. PCC 7822							
	<i>Cyanothece</i> sp. PCC 7424							
	<i>Stanieria cyanosphaera</i> PCC 7437							
	<i>Synechococcus</i> sp. PCC 7002							
	<i>Leptolyngbya</i> sp. PCC 7376							

A	<i>Cyanobacterium staniieri</i> PCC 7202	單細胞	3,163,381	2,837	2,837	3	43	3	0NC_019778
A	<i>Cyanobacterium aponinum</i> PCC 10605	單細胞	4,114,099	3,358	3,431	3	43	3	1NC_019776
A2	<i>Synechococcus</i> sp. RCC307	單細胞(pro/syn)	2,224,914	2,533	2,533	3	42	1	0NC_009482
A2	<i>Prochlorococcus marinus</i> str. NATL2A	單細胞(pro/syn)	1,842,899	2,162	2,162	3	38	1	0NC_007335
A2	<i>Prochlorococcus marinus</i> str. NATL1A	單細胞(pro/syn)	1,864,731	2,195	2,195	2	38	1	0NC_008819
A2	<i>Prochlorococcus marinus</i> subsp. <i>pastoris</i> str. CCMP1986	單細胞(pro/syn)	1,657,990	1,717	1,717	3	37	1	0NC_005072
A2	<i>Prochlorococcus marinus</i> str. MIT 9515	單細胞(pro/syn)	1,704,176	1,909	1,909	2	37	1	0NC_008817
A2	<i>Prochlorococcus marinus</i> str. AS9601	單細胞(pro/syn)	1,669,886	1,923	1,923	2	38	1	0NC_008816
A2	<i>Prochlorococcus marinus</i> str. MIT 9301	單細胞(pro/syn)	1,641,879	1,908	1,908	2	37	1	0NC_009091
A2	<i>Prochlorococcus marinus</i> str. MIT 9215	單細胞(pro/syn)	1,738,790	1,989	1,989	3	36	1	0NC_009840
A2	<i>Prochlorococcus marinus</i> str. MIT 9312	單細胞(pro/syn)	1,709,204	1,810	1,810	3	39	1	0NC_007577
A2	<i>Prochlorococcus marinus</i> str. MIT 9211	單細胞(pro/syn)	1,688,963	1,855	1,855	2	40	1	0NC_009976
A2	<i>Prochlorococcus marinus</i> subsp. <i>marinus</i> str. CCMP1375	單細胞(pro/syn)	1,751,080	1,882	1,882	3	40	1	0NC_005042
A2	<i>Prochlorococcus marinus</i> str. MIT 9313	單細胞(pro/syn)	2,410,873	2,269	2,269	3	44	2	0NC_005071
A2	<i>Prochlorococcus marinus</i> str. MIT 9303	單細胞(pro/syn)	2,682,675	2,999	2,999	3	43	2	0NC_008820
A2	<i>Prochlorococcus</i> sp. WH 8102	單細胞(pro/syn)	2,434,428	2,519	2,519	3	44	2	0NC_005070
A2	<i>Synechococcus</i> sp. CC9605	單細胞(pro/syn)	2,510,659	2,645	2,645	3	45	2	0NC_007516
A2	<i>Synechococcus</i> sp. CC9902	單細胞(pro/syn)	2,234,828	2,306	2,306	0	45	2	0NC_007513
A2	<i>Synechococcus</i> sp. WH 7803	單細胞(pro/syn)	2,366,980	2,533	2,533	3	44	2	0NC_009481
A2	<i>Synechococcus</i> sp. CC9311	單細胞(pro/syn)	2,606,748	2,892	2,892	2	44	2	0NC_008319
A2	<i>Cyanobium gracile</i> PCC 6307	單細胞(pro/syn)	3,342,364	3,280	3,280	4	44	2	0NC_019675
A2	<i>Synechococcus elongatus</i> PCC 7942	單細胞(pro/syn)	2,695,903	2,612	2,612	3	44	2	1NC_007604
A2	<i>Synechococcus elongatus</i> PCC 6301	單細胞(pro/syn)	2,696,255	2,524	2,524	4	45	2	0NC_006576
A	<i>Geitlerinema</i> sp. PCC 7407	多細胞(filamentous)	4,681,111	3,819	3,819	3	46	3	0NC_019703
A	<i>Thermosynechococcus elongatus</i> BP-1	單細胞	2,593,857	2,476	2,476	4	42	1	0NC_004113
A	<i>Synechococcus</i> sp. PCC 6312	單細胞	3,697,276	3,515	3,547	4	41	1	1NC_019680
A	<i>Cyanothece</i> sp. PCC 7425	單細胞	5,374,574	4,968	5,329	4	47	2	3NC_011884
A	<i>Acaryochloris marina</i> MBIC 11017	單細胞	6,503,724	6,257	8,386	4	69	2	9NC_009925
B	<i>Pseudanabaena</i> sp. PCC 7367	多細胞(filamentous)	4,557,046	3,563	3,856	3	46	2	1NC_019701
B	<i>Synechococcus</i> sp. PCC 7502	單細胞	3,510,253	3,249	3,319	2	43	2	2NC_019702
B	<i>Limnothrix</i> sp. ABRG5-3	多細胞(filamentous)	4,796,642	4,318	4,984	4	52	3	7本博士論文
C	<i>Synechococcus</i> sp. JA-2-3B'a(2-13)	單細胞	3,046,682	2,865	2,865	1	45	2	0NC_007776
C	<i>Synechococcus</i> sp. JA-3-3Ab	單細胞	2,932,766	2,760	2,760	2	47	2	0NC_007775
D	<i>Gloeobacter violaceus</i> PCC 7421	單細胞	4,659,019	4,430	4,430	4	45	1	0NC_005125

表 4-5：シアノバクテリアにコードされたタンパク質遺伝子のホモログ解析と機能分類

全体の クラス ター数	<i>Limnithrix</i> sp. ABRG5-3 を含む クラスター数	機能分類に含まれるタンパク質遺伝子数																			
		101	102	103	104	105	106	107	108	109	110	111	201	202	203	204	205				
ABCD	911	26	11	35	38	122	25	15	31	15	162	58	48	28	1	12	11				
ABC-	86	0	0	0	1	3	0	3	2	10	13	0	9	1	0	0	0				
AB-D	71	4	0	3	0	5	0	1	6	0	1	0	1	0	0	0	0				
AB--	32	0	0	1	0	0	0	0	2	4	4	0	3	0	0	0	0				
A-CD	50	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0				
A-C-	20	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
A--D	28	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0				
A---	32	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0				
-BCD	89	0	1	33	8	0	0	0	2	2	0	1	1	0	0	0	0				
-BC-	72	0	0	1	0	0	0	0	5	0	7	1	0	0	1	0	0				
-B-D	81	1	5	15	3	1	4	1	1	0	8	5	0	0	0	0	0				
-B--	176	2	0	14	4	2	0	1	8	7	27	2	1	0	0	0	0				
--CD	181	2	4	0	7	0	1	1	4	1	2	0	0	0	0	0	0				
--C-	518	66	0	7	1	0	1	0	1	2	4	3	0	0	0	0	0				
---D	1,064	245	5	0	47	37	1	5	0	4	10	14	15	1	0	0	2				
----	10,920	1,131	29	6	94	36	4	5	1	26	47	89	13	0	13	3	0				
unique		589	6	3	20	16	3	1	4	6	61	11	0	1	0	0	0				
合計		3,651	75	30	272	151	141	42	24	98	104	394	110	64	30	16	17	11			

シアノバクテリアのグループ分けは表 4-4 で示した。機能 ID は表 4-3 で示した。

共通して存在するグループの有無によってクラスターを区分けた。例えば、ABCD に属するクラスターは、A, B, C, D 全てのグループが共通して存在していることを示し、-B-- に属するクラスターは、B グループは共通して存在しているが、それ以外は共通して存在してはいないことを示す。

表 4-5 : シアノバクテリアにコードされたタンパク質遺伝子のホモログ解析と機能分類(続き)

	206	207	208	209	210	211	212	213	301	302	303	304	305	400	401	合計
ABCD	67	37	66	9	112	34	68	47	46	13	130	10	0	26	186	1,489
ABC-	1	3	7	1	4	1	2	1	1	0	19	0	0	1	25	108
AB-D	8	4	0	1	4	0	5	12	6	4	11	0	0	1	27	104
AB--	2	1	0	0	1	0	0	0	2	1	9	0	0	1	13	44
A-CD	1	0	0	0	0	0	1	1	0	0	1	0	0	1	7	15
A-C-	0	0	0	0	0	0	2	0	0	0	0	0	0	1	2	5
A--D	1	0	3	0	0	3	1	2	0	0	0	0	0	0	3	15
A---	0	0	0	2	0	0	1	0	2	0	1	0	0	0	9	18
-BCD	8	6	4	1	1	3	9	13	1	2	17	0	1	1	26	141
-BC-	1	1	8	2	7	1	4	4	0	1	17	1	7	1	23	93
-B-D	3	7	6	0	5	2	7	4	4	1	5	0	0	2	60	150
-B--	6	2	0	2	7	0	4	9	6	6	17	0	1	2	99	229
--CD	4	1	1	2	6	0	3	2	0	0	12	0	0	1	38	92
--C-	4	0	0	0	12	1	3	1	2	2	10	0	0	2	21	77
---D	7	8	2	6	6	3	3	19	12	5	29	0	0	12	122	375
----	23	26	9	9	36	9	18	71	41	6	65	5	0	60	696	1,440
unique	3	8	1	1	8	3	5	23	18	0	28	4	0	43	311	589
合計	139	104	107	36	209	60	136	209	141	41	371	20	9	155	1,668	4,984

機能分類に含まれるタンパク質遺伝子数

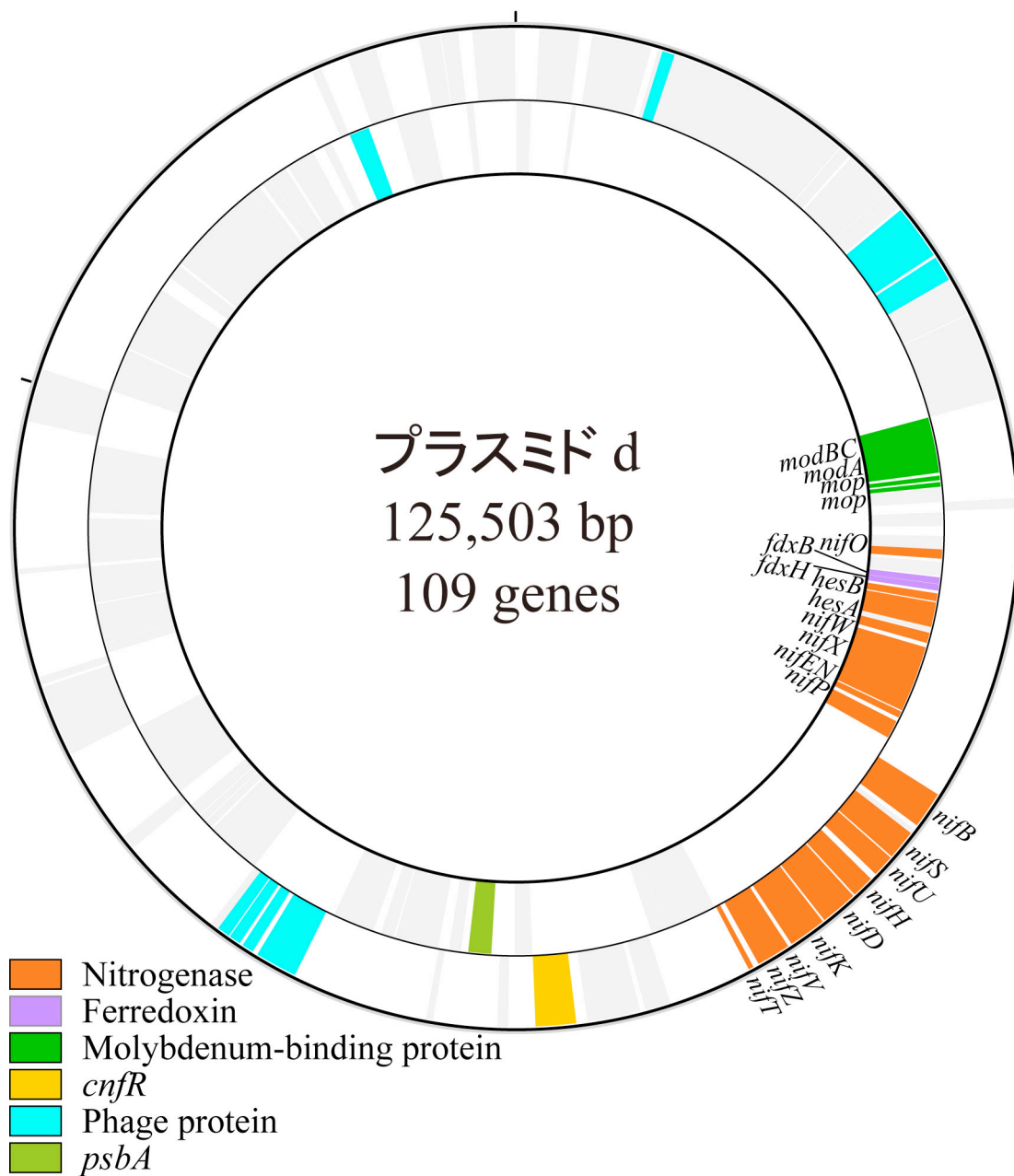


図 4-1 : プラスミド d ゲノムにコードされた窒素固定関連遺伝子群
 外側の輪が時計回り方向にコードされた遺伝子を示し、内側の輪が反時計回
 り方向にコードされた遺伝子を示す。遺伝子は、それぞれ機能毎に左下で示
 した色に塗り分けた。

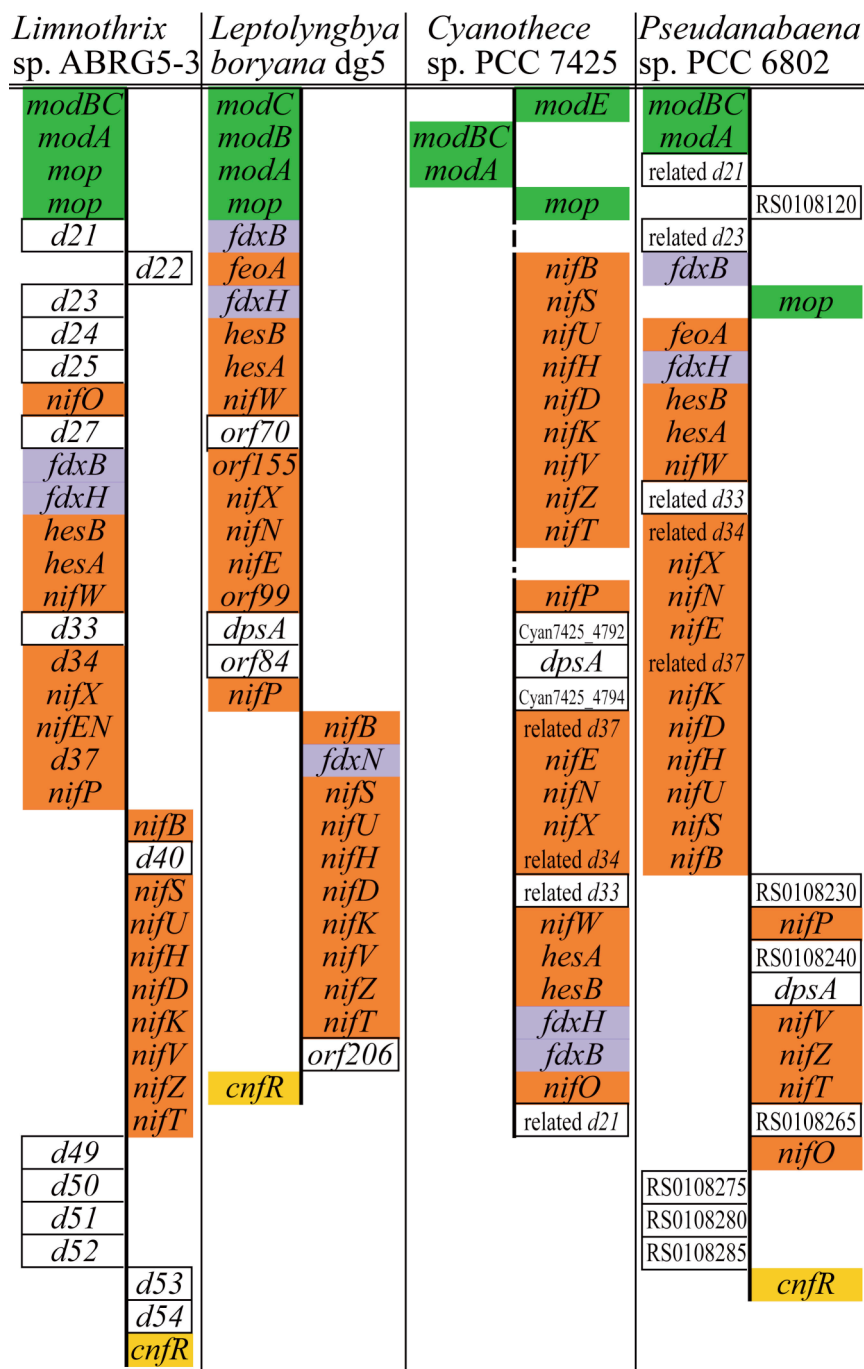


図 4-2 : 窒素固定遺伝子クラスターの比較

右側が正方向にコードされた遺伝子を示し、左側が逆方向にコードされた遺伝子を示す。遺伝子は、それぞれ機能毎に図 4-1 と同じ色に塗り分けた。

Cyanothece sp. PCC 7425 の中央の線で破線になっている部分は、その前後の遺伝子がつながっていないことを示した。

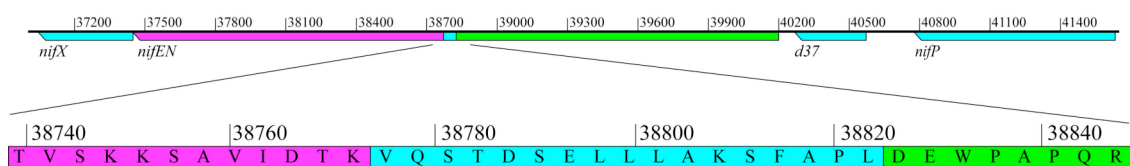


図 4-3 : *nifEN* の構造

上部は *nifEN* および近隣遺伝子の全体像を示し、下部は *nifE* 部分と *nifN* 部分の境目となる部分のアミノ酸配列を示した。N 末端側の緑色で塗った部分までが *nifE* と相同性が高く、C 末端側の紫色で塗った部分は *nifN* と相同性が高かった。系統解析にはそれぞれ塗り分けた部分のみ使用した。

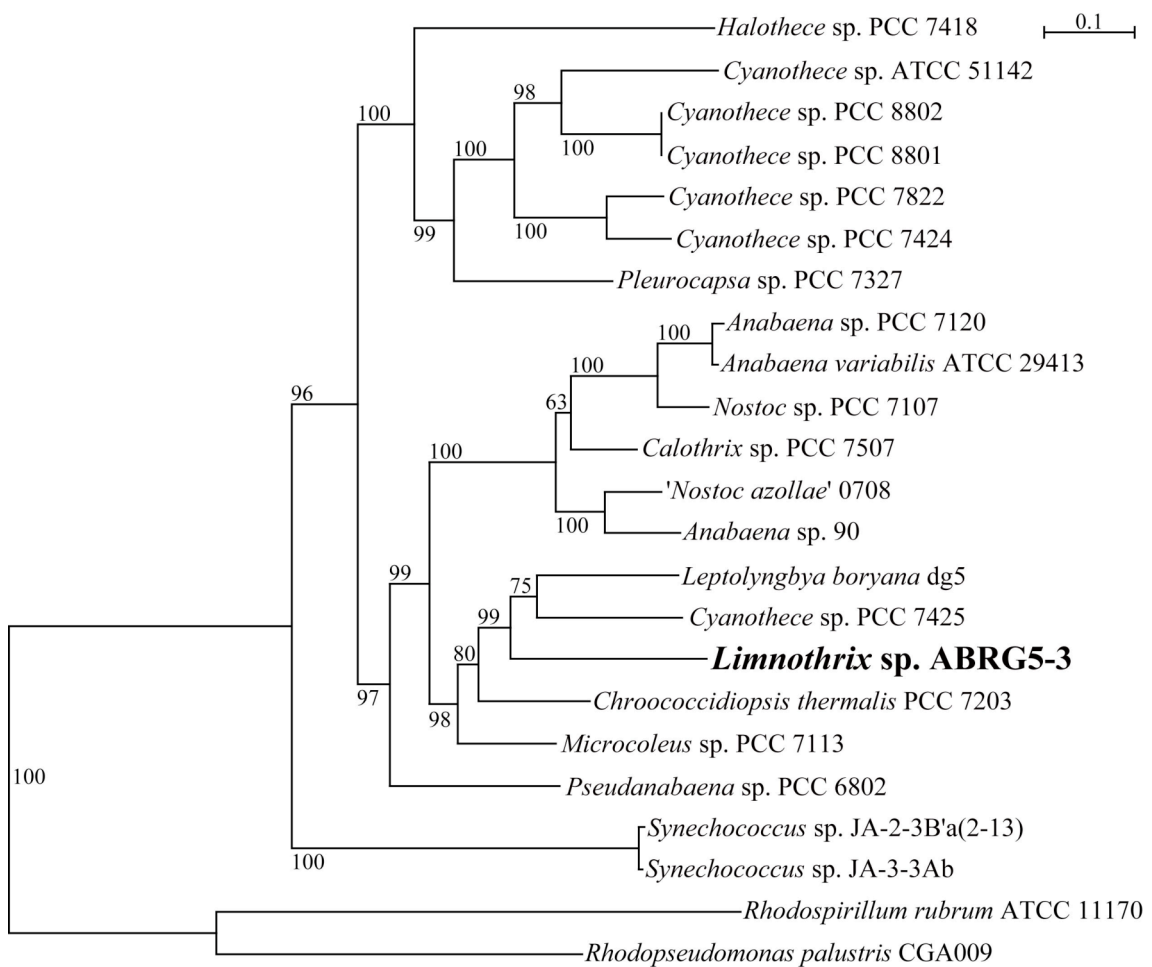


図 4-4 : 窒素固定関連遺伝子群の系統解析

窒素固定関連遺伝子群のうち 14 個のタンパク質遺伝子のアミノ酸配列を用いた。右上のスケールバーは枝の長さを示す。それぞれの分岐の肩にある数字は最尤法での信頼度を示す。

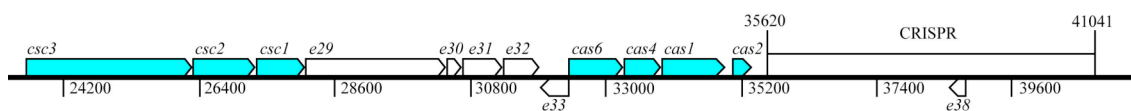


図 4-5 : プラスミド e ゲノムの CRISPR-Cas 系

上側が正方向にコードされた遺伝子を示し、下側が逆方向にコードされた遺伝子を示す。Cas タンパク質遺伝子は水色に塗って示した。

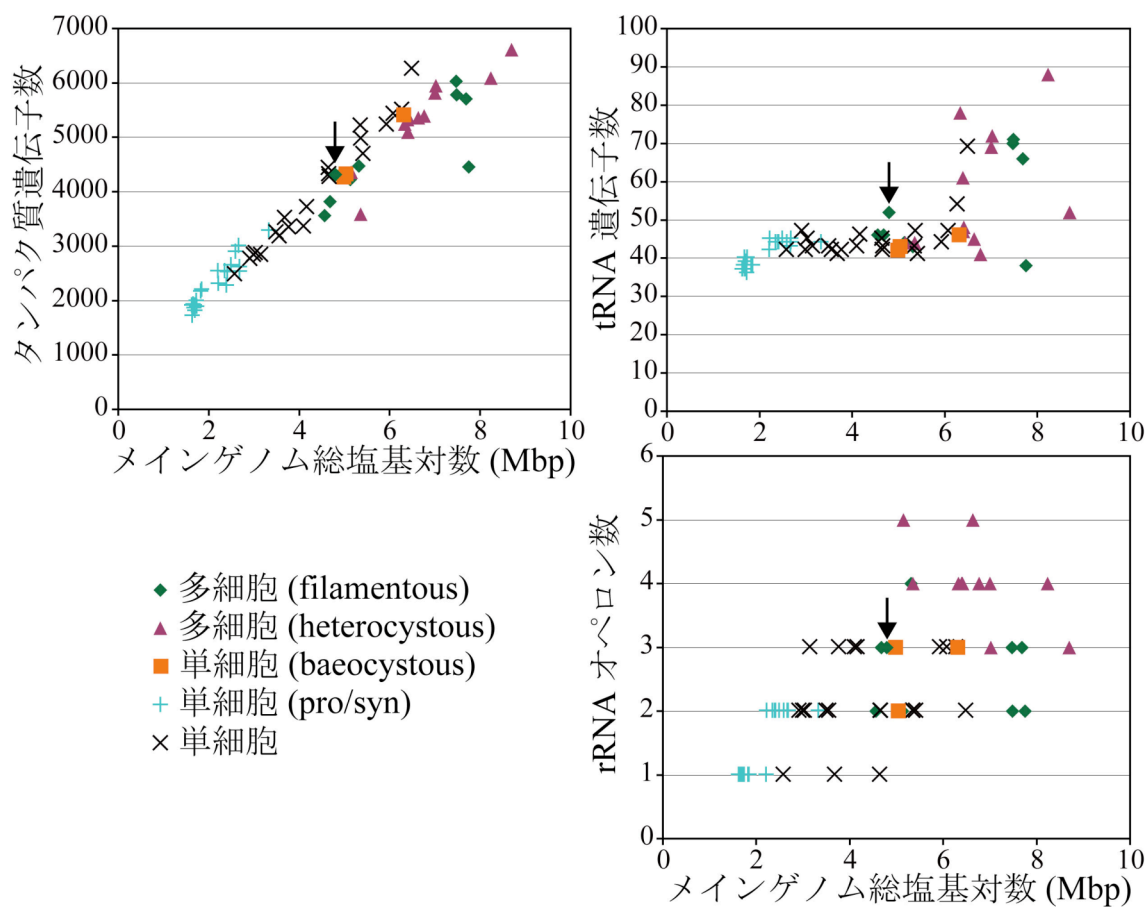


図 4-6 : シアノバクテリアの比較

(左上)メインゲノムの総塩基対数とタンパク質遺伝子数との関係

(右上)メインゲノムの総塩基対数と tRNA 遺伝子数との関係

(右下)メインゲノムの総塩基対数と rRNA オペロン数との関係

それぞれ *Limnothrix sp. ABRG5-3* のデータを矢印で指し示した。

図 4-7、図 4-8、図 4-9 の系統樹の枝は、この時の記号の色に従いそれぞれ塗り分けた。

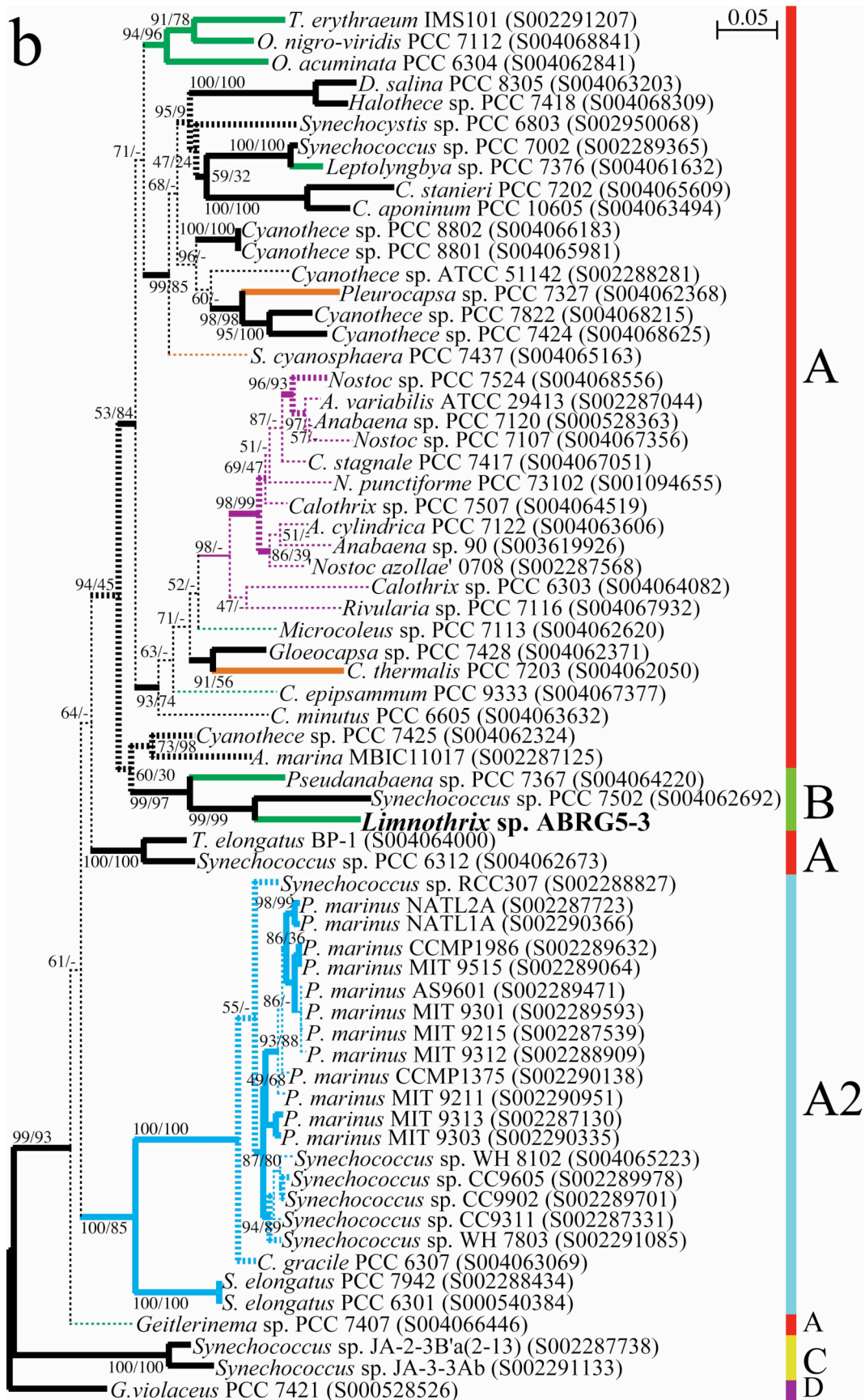


図 4-7 : シアノバクテリアの系統樹(前ページ)

表 4-4 で分けたシアノバクテリアのグループを、それぞれの系統樹の右側に太線で示した。

a: タンパク質遺伝子アミノ酸配列を用いた系統樹

シアノバクテリアに保存された 31 個のタンパク質遺伝子(*atpA*, *atpB*, *atpE*, *atpH*, *petA*, *petB*, *petD*, *psaA*, *psaB*, *psaC*, *psbB*, *psbC*, *psbE*, *rplB*, *rplN*, *rplP*, *rplT*, *rpoA*, *rpoB*, *rpoC1*, *rpsB*, *rpsC*, *rpsD*, *rpsG*, *rpsH*, *rpsK*, *rpsL*, *rpsN*, *rpsS*, *ycf3*, *ycf4*)のアミノ酸配列を用いた。右上のスケールバーは枝の長さを示す。それぞれの分岐の肩にある数字は最尤法での信頼度を示す。アスタリスクがついた枝以外はベイズ法では 100%支持した。アスタリスクがついた枝の事後確率は 73 だった。

b: 16S rRNA 遺伝子の塩基配列を用いた系統樹

右上のスケールバーは枝の長さを示す。それぞれの分岐の肩にある数字は最尤法での信頼度/近隣結合法でのブートストラップ値を示す。最尤法と近隣結合法で支持した枝の分かれ方が同じだった部分は太線で示した。タンパク質遺伝子アミノ酸配列を用いた系統樹と 16S rRNA 遺伝子の塩基配列を用いた系統樹で支持した枝の分かれ方が同じだった部分は実線で示し、違った部分は破線で示した。RDP の識別番号をそれぞれの種の後に表示した。

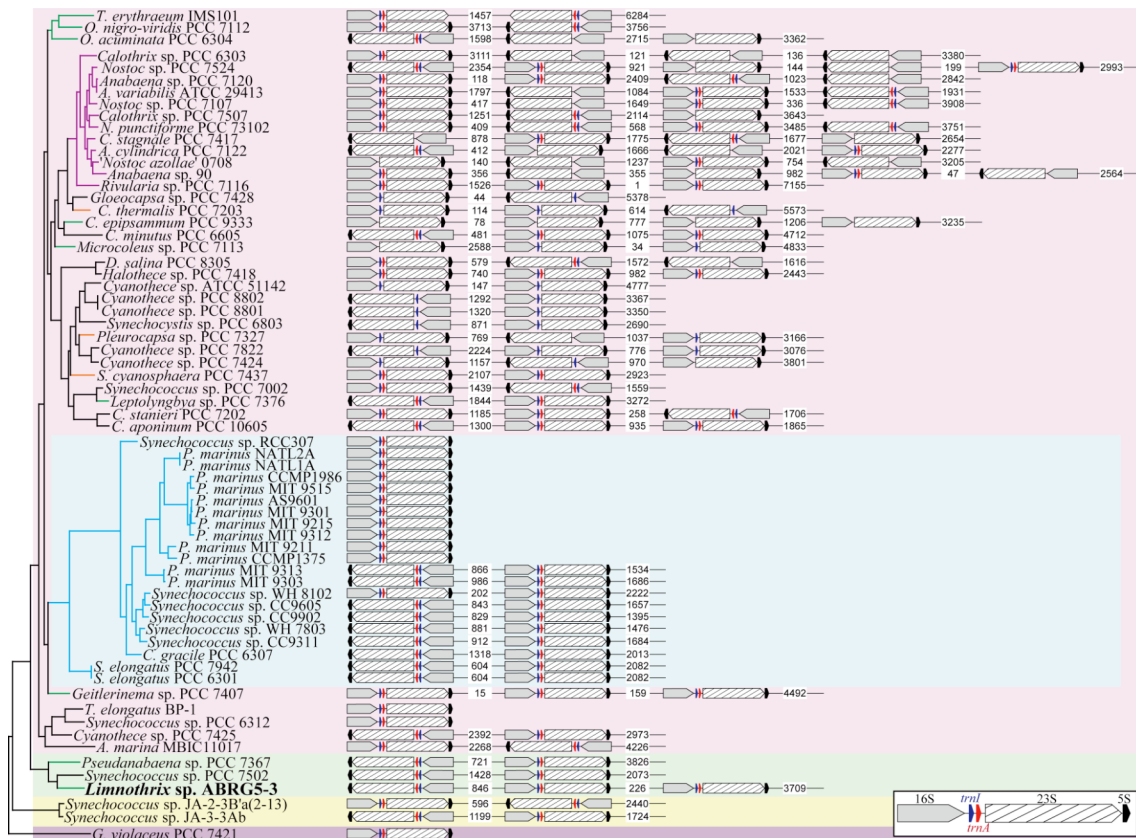


図 4-8 : シアノバクテリアゲノムの rRNA オペロン構造

系統樹は図 4-7a を利用した。rRNA オペロンは、右下の凡例に従って表した。rRNA オペロンの間にある数字はそれぞれ rRNA オペロンの間の塩基対数(単位は kbp)を示す。背景の色は、桃色は表 4-4 の A グループ、水色は A2 グループ、緑色は B グループ、黄色は C グループ、紫色は D グループをそれぞれ示す。

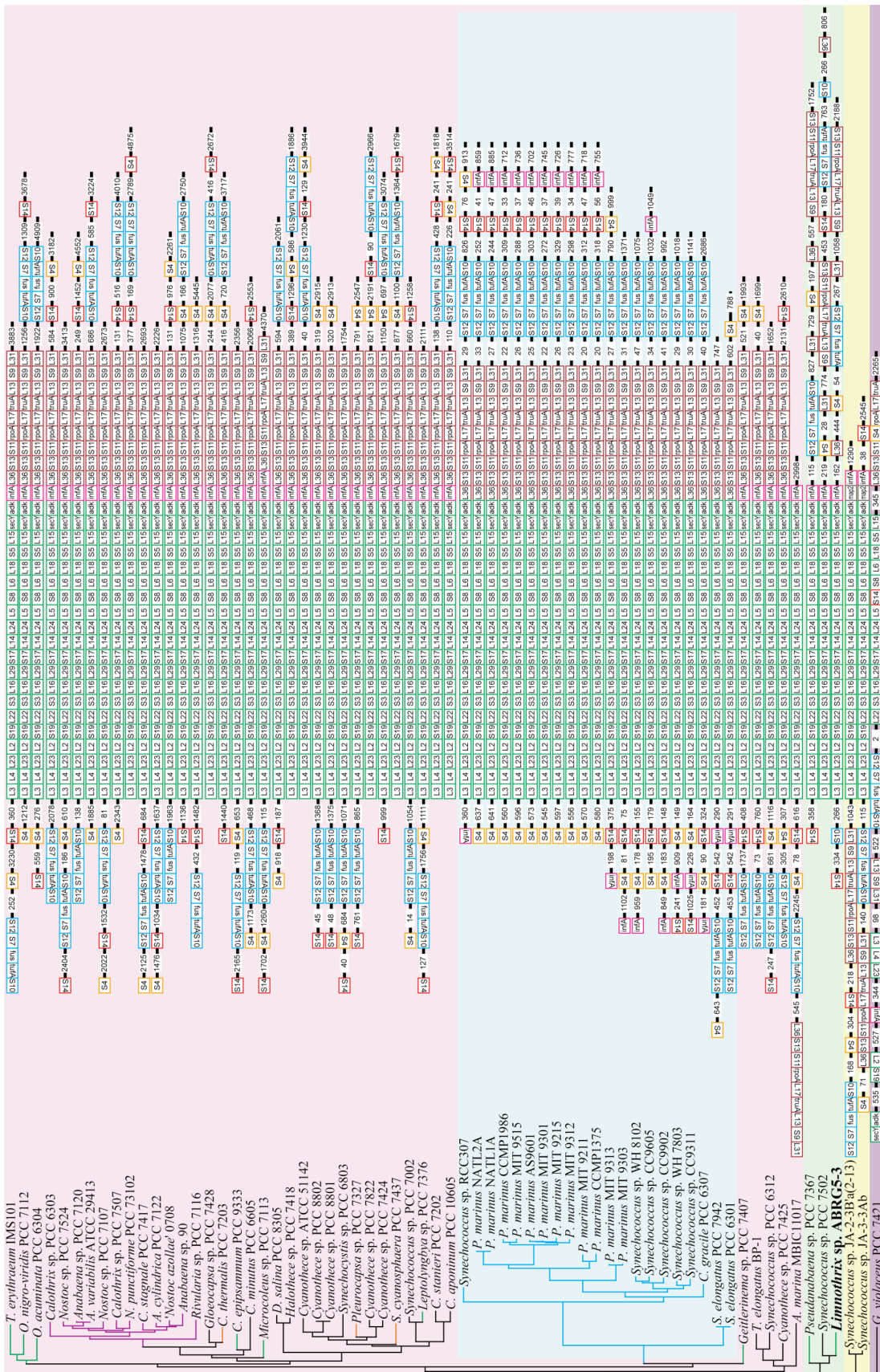


図 4-9 : シアノバクテリアゲノムのリボソームタンパク質遺伝子クラスター
(前ページ)

系統樹は図 4-7a を利用した。上下逆になっている遺伝子は逆方向にコードされた遺伝子を示す。遺伝子名の間にある数字はそれぞれリボソームタンパク質遺伝子クラスター以外の遺伝子を間に挟んだ部分のリボソームタンパク質遺伝子クラスター遺伝子間の塩基対数(単位は kbp)を示す。背景の色は、桃色は表 4-4 の A グループ、水色は A2 グループ、緑色は B グループ、黄色は C グループ、紫色は D グループをそれぞれ示す。

第5章

藻類の系統解析

5.1. 序文

第2章から第4章までで、紅色植物門の中でまだ解読されていなかった綱、チノリモ綱の葉緑体ゲノム、最近珪藻の姉妹群であることが分かったパルマ藻の葉緑体ゲノムやミトコンドリアゲノム、シアノバクテリアのゲノムを解読してきた。この章ではこれらの生物と、既に解読された真核光合成生物の葉緑体ゲノムとシアノバクテリアゲノムに共通してコードされるタンパク質遺伝子のアミノ酸配列を用いた系統樹と、16S rRNA 遺伝子の塩基配列を用いた系統樹を作成し、系統関係を調べた。また、パルマ藻のミトコンドリアゲノムについても、既に解読された真核光合成生物のミトコンドリアゲノムと、ミトコンドリアの起源といわれる α -プロテオバクテリアのゲノムに、共通してコードされるタンパク質遺伝子のアミノ酸配列を基にした系統樹を作成し、系統関係を調べた。

5.2. 材料と方法

5.2.1. 系統解析

系統解析のために、葉緑体の 16S rRNA 遺伝子の塩基配列とタンパク質遺伝子アミノ酸配列のアラインメントを作成した。

葉緑体ゲノムのタンパク質遺伝子アミノ酸配列での系統解析を行うために、2通りのタンパク質遺伝子アミノ酸配列のアラインメントを作成した。1つは5種の紅藻と24種のクロムアルベオラータの葉緑体ゲノム、および19種のシアノバクテリアゲノムに共通してコードされた64個のタンパク質遺伝子のアミノ酸配列のアラインメントを結合したもので、もう1つは6種の緑色植物亜界、1種の灰色植物門、5種の紅藻と24種のクロムアルベオラータの葉緑体ゲノム、および19種のシアノバクテリアゲノムに共通してコードされた33個のタンパク質遺伝子のアミノ酸配列のアラインメントを結合したものである。

ミトコンドリアゲノムの系統解析を行うために、4通りのタンパク質遺伝子アミノ酸配列のアラインメントを作成した。1つ目は13種の不等毛植物門のミトコンドリアゲノム、および5種の α -プロテオバクテリアゲノムに共通してコードされた29個のタンパク質遺伝子のアミノ酸配列のアラインメントを結合したものである。2つ目は6種の緑色植物亜界、2種の灰色植物門、3種の紅藻、1種のクリプト藻、1種のハプト藻と13種の不等毛植物門のミトコンドリアゲノム、および5種の α -プロテオバクテリアゲノムに共通してコードされた15個のタンパク質遺伝子のアミノ酸配列のアラインメントを結合したものとした。2つ目のデータセットにさらにもう1種ずつのクリプト藻とハプト藻のデータを加えたものとクリプト藻とハプト藻のデータを除いたものも作成した。ミトコンドリアの場合はアラインメント長が短いものを加えた方が、樹形が安定したので、アラインメント長が75に満たなかったタンパク質(Atp8、Atp9)も計算に加えた。

16S rRNA 遺伝子の系統解析は、6種の緑色植物亜界、1種の灰色植物門、5種の紅藻と24種のクロムアルベオラータの葉緑体ゲノム、および19種のシアノバクテリアゲノムに対して行った。これらの種の内、RDPに登録されている16S rRNA 遺伝子塩基配列はRDPから手に入れた。登録されていなかった

たものは NCBI に登録されたゲノムから 16S rRNA 遺伝子塩基配列部分を抜き出したものを利用した。RDP から手に入れたアラインメントから共通しているギャップを取り除いた上で、NCBI に登録されたゲノムから 16S rRNA 遺伝子塩基配列部分を抜き出したものを、ひとつずつ加えては元のアラインメントに沿ったアラインメントを作成するという作業を繰り返し、系統解析に用いたアラインメントを作成した。

アミノ酸配列の加工や系統樹の計算方法は **4.2.7.**のシアノバクテリアの系統解析と同じ方法である。

16S rRNA 遺伝子の系統樹の計算方法も **4.2.7.**のシアノバクテリアの系統解析と同じ方法、モデル等も同じ設定で行った。

5.3. 結果

5.3.1. 葉緑体ゲノムの系統解析

紅藻およびクロムアルベオラータの系統関係を調べるため、アーケプラスチダとクロムアルベオラータの葉緑体ゲノムとシアノバクテリアゲノムに共通してコードされた 33 個のタンパク質遺伝子のアミノ酸配列のアラインメントを結合し、系統解析を行った(図 5-1a)。さらに、より多くのタンパク質遺伝子を解析に用いるため、対象となる生物を減らし、紅藻およびクロムアルベオラータの葉緑体ゲノムとシアノバクテリアゲノムに共通してコードされた 64 個のタンパク質遺伝子のアミノ酸配列のアラインメントを結合し、系統解析を行った(図 5-1b)。過去に紅藻やクリプト藻で遺伝子の水平伝播が起きたと指摘されたタンパク質(RbcL、RbcS、Rpl36)は除外した(Delwiche & Palmer 1996、Rice & Palmer 2006)。また、Qiu et al. (2012)において、過去に複雑に重複、消失が起きた遺伝子は、系統解析に悪影響を与える可能性があることや、短い配列のアラインメントは系統関係を示すには不十分となっていることが指摘されていたため、共通してコードされていたが、過去に複雑に重複、消失が起きたと指摘されたタンパク質(PsbA、PsbD)とアラインメント長が 75 に満たなかったタンパク質(PsbF、PsbH、PsbJ、PsbK、PsbN、PsbT、Rpl31、Rpl33、Rps18)は除外した。これらの遺伝子を含めた場合と含めなかった場合で、樹形に違いが生じることも既に確認している(Tajima et al. 2014)。系統樹の計算には、最尤法とベイズ法を利用した。紅藻の系統関係について、全ての解析で同じ結果を支持し、これは Yoon et al. (2006)や Verbruggen et al. (2010)の結果と同じであった。また、パルマ藻が珪藻にもっとも近縁であることと、*Limnothrix* sp. ABRG5-3 が *Pseudanabaena* sp. PCC 7367 に最も近縁で、シアノバクテリアの進化の根元の方に位置することが完全に支持された。さらに、クロムアルベオラータの葉緑体の共通祖先が紅藻植物亜門の共通祖先であることと、葉緑体の共通祖先が第 4 章で示した A グループの共通祖先であることが強く支持された。ハプト藻とクリプト藻は、今回の解析では単系統にならなかった。*T. erythraeum* IMS101 の系統がヘテロシストを形成するシアノバクテリアと最も近縁になる点以外は、シアノバクテリアのみの系統樹(図 4-7 a)と同じ結果を示した。

アーケプラスチダとクロムアルベオラータの葉緑体およびシアノバクテリアの 16S rRNA 遺伝子による系統解析を行った(図 5-2)。系統樹の計算には、最尤法と近隣結合法を用いた。結果は、タンパク質遺伝子アミノ酸配列での系統解析結果と異なっていた。また、最尤法と近隣結合法の間でも結果は異なっていた。さらに、シアノバクテリアのみで行った 16S rRNA 遺伝子の系統解析結果(図 4-7b)とも異なっていた。最尤法では *Limnothrix* sp. ABRG5-3 と *Pseudanabaena* sp. PCC 7367 が姉妹群にならなかった。ハプト藻の系統が 2 つに分かれ、それぞれ別の紅藻から分岐する形になった。*Triparma laevis* が珪藻と姉妹群であることとアーケプラスチダの間では灰色植物門がまず分岐したことはタンパク質の系統解析結果と同じであった。最尤法の信頼度や近隣結合法のブートストラップ値が低い分岐も多かったため、16S rRNA 遺伝子の系統解析から系統関係を明らかにすることはできなかった。

5.3.2. ミトコンドリアゲノムの系統解析

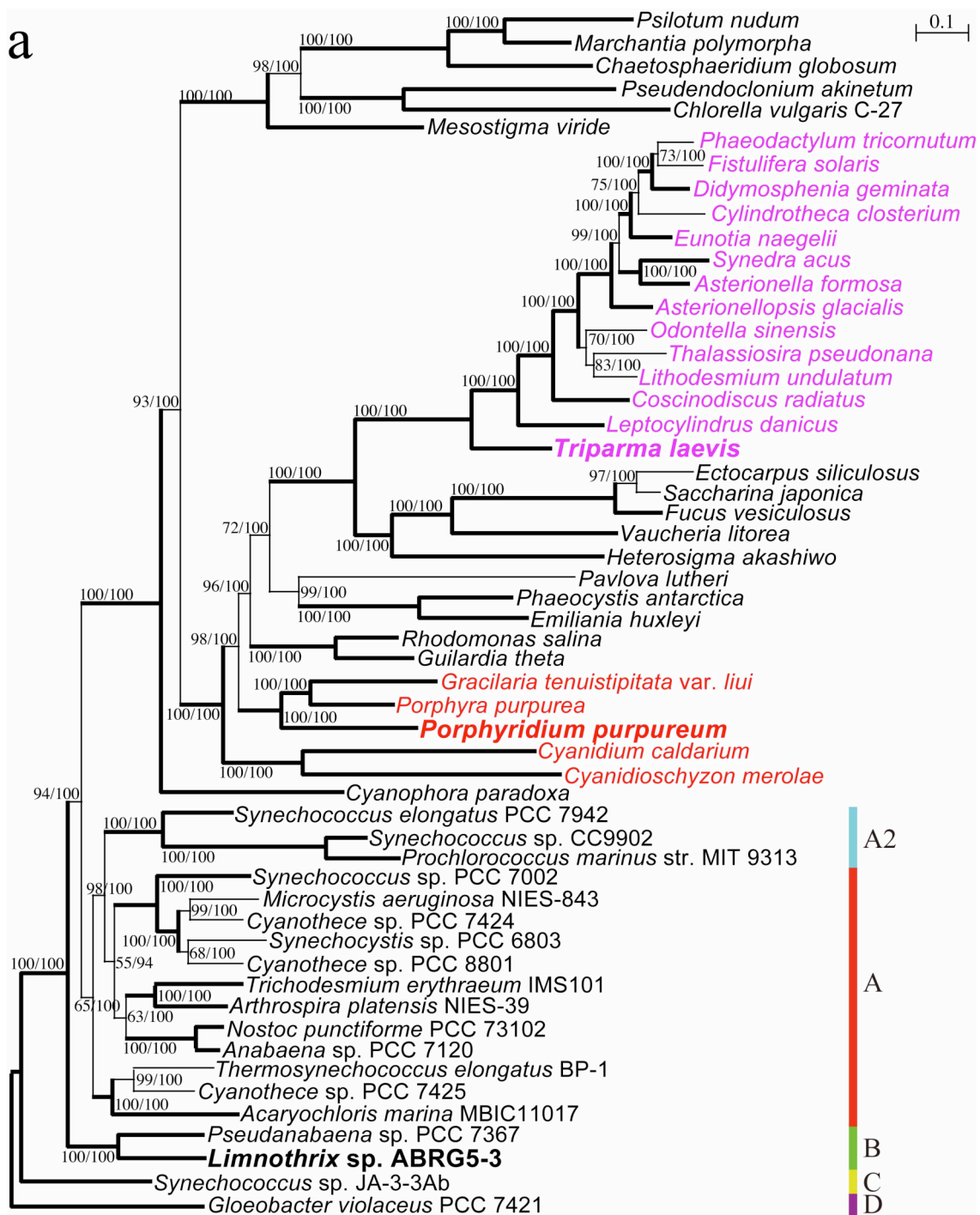
ミトコンドリアの系統関係を調べるため、不等毛植物門のミトコンドリアゲノムとミトコンドリアの起源といわれる α -プロテオバクテリアゲノムに共通してコードされた 29 個のタンパク質遺伝子のアミノ酸配列のアラインメントを結合し、系統解析を行った(図 5-3a)。さらに、より多くの生物との系統関係を調べるため、対象となる生物を増やし、アーケプラスチダとクロムアルベオラータのミトコンドリアゲノムと α -プロテオバクテリアのゲノムに共通してコードされた 15 個のタンパク質遺伝子のアミノ酸配列のアラインメントを結合し、系統解析を行った(図 5-3b(ii))。また、このデータセットに、さらにもう 1 種ずつクリプト藻とハプト藻のミトコンドリアゲノムのデータを加えたもの(図 5-3b(iii))と、クリプト藻とハプト藻を除いたもの(図 5-3b(i))についてもそれぞれ系統解析を行った。いずれの結果も、 α -プロテオバクテリアのリケッチア目を外群とした。不等毛植物門のミトコンドリアゲノムを用いた系統解析は、全てのデータセット、計算で同じ結果を示した。さらに、パルマ藻が珪藻に最も近縁であることは完全に支持された。しかし、それ以外の生物との系統関係は、データセットや計算方法によって異なっていた。クリプト藻とハプト藻が 1 種類ずつあるデータセットでは、クリプト藻とハプト藻と紅藻が単系統になっていた。また、灰色植物門と緑色植物門

界が単系統であり、不等毛植物門はこちらのクレードと1つになっていた。この結果は、クリプト藻とハプト藻を除いたデータセットの解析結果でも支持された。それに対して、クリプト藻とハプト藻が2種類ずつあるデータセットでは、緑色植物亜界、灰色植物門、紅藻、クリプト藻、ハプト藻が単系統になり、不等毛植物門はその外側に位置していた。また、最尤法の計算結果では、灰色植物門とクリプト藻、紅藻が単系統、ハプト藻と緑色植物亜界が単系統であるのに対して、ベイズ法の計算結果では、ハプト藻とクリプト藻、紅藻が単系統、灰色植物門と緑色植物亜界が単系統であるという異なったものとなった。クリプト藻とハプト藻を除いたデータセットの最尤法での計算結果のみ、*Nephroselmis olivacea* と *Mesostigma viride* の系統関係が入れ替わっていた。

5.4. 考察

これまで解読してきたゲノムと、過去に読まれてきたゲノムを合わせて、オルガネラゲノムの系統解析を行った。葉緑体ゲノムのタンパク質遺伝子アミノ酸配列の系統解析は、データセットの違いによって、一部の珪藻の系統関係に若干違いが見られた(*Fistulifera* sp.が *Phaeodactylum tricornutum* と単系統になるか、*Didymosphenia geminata* と単系統になるか)ぐらいで、紅藻、クリプト藻、ハプト藻、不等毛植物門の系統関係に違いは見られなかった。それに対して、ミトコンドリアゲノムの系統解析では、データセットの選び方によって、樹形が大きく変わることが分かった。さらに、以前行われたミトコンドリアゲノムの系統解析の結果とも異なっていた(Oudot-Le Secq et al. 2006)。ミトコンドリアゲノムでの系統解析の結果の違いは、葉緑体ゲノムと比べてミトコンドリアゲノムは、門毎に遺伝子の変異率が大きく異なること(Drouin et al. 2008、 Smith et al. 2012、 Smith et al. 2014)に影響を受けているのではないかと(Felsenstein 1978)と考えられた。

今回の結果では、細胞内共生した葉緑体の系統解析はクリプト藻、ハプト藻、不等毛植物門が単系統となったが、細胞内共生のもととなる細胞のものであると考えられるミトコンドリアの系統解析では、どのデータセットでも三者は単系統にならなかった。これらのことからクリプト藻、ハプト藻、不等毛植物門は、それぞれ別の種類の生物に紅藻植物亜門の共通祖先となる紅藻が細胞内共生したことから生まれたのではないかとこの可能性が考えられた。



b

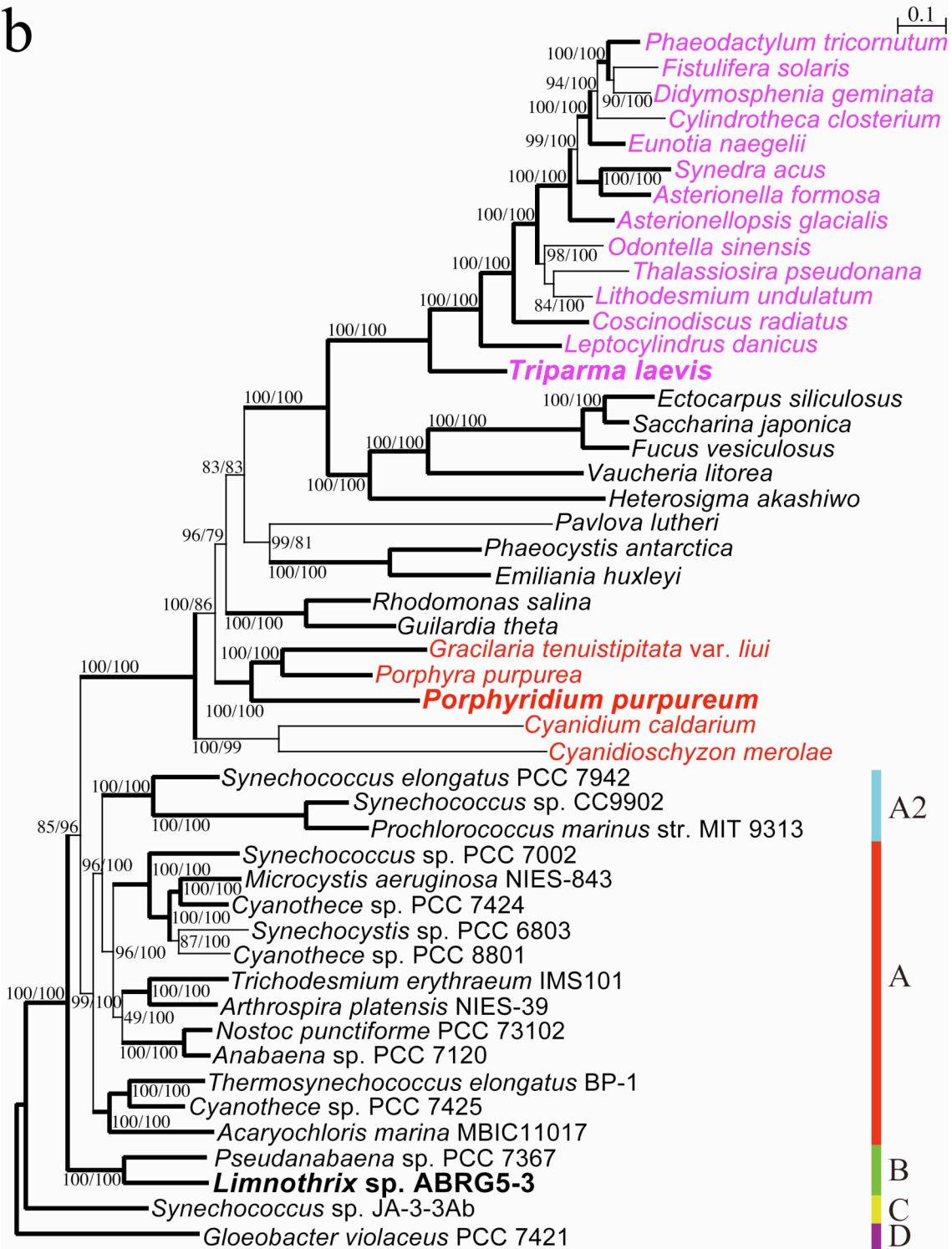


図 5-1 : 葉緑体とシアノバクテリアゲノムのタンパク質遺伝子アミノ酸配列を用いた系統樹(前ページ)

右上のスケールバーは枝の長さを示す。それぞれの分岐の肩にある数字は最尤法での信頼度/ベイズ法での事後確率を示す。太線はどちらの方法でも完全に支持された分岐を示す。表 4-4 で分けたシアノバクテリアのグループを、それぞれの系統樹の右側に太線で示した。

a:アーケプラスチダとクロムアルベオラータの葉緑体とシアノバクテリアの系統樹

アーケプラスチダとクロムアルベオラータの葉緑体ゲノムとシアノバクテリアゲノムに保存された 33 個のタンパク質遺伝子(*atpA*, *atpB*, *atpE*, *atpF*, *atpH*, *petA*, *petB*, *petD*, *psaA*, *psaB*, *psaC*, *psbB*, *psbC*, *psbE*, *rpl2*, *rpl14*, *rpl16*, *rpl20*, *rpoA*, *rpoB*, *rpoC1*, *rpoC2*, *rps2*, *rps3*, *rps4*, *rps7*, *rps8*, *rps11*, *rps12*, *rps14*, *rps19*, *ycf3*, *ycf4*)のアミノ酸配列を用いた。

b:紅藻およびクロムアルベオラータの葉緑体とシアノバクテリアの系統樹

紅藻およびクロムアルベオラータの葉緑体ゲノムとシアノバクテリアゲノムに保存された 64 個のタンパク質遺伝子(*atpA*, *atpB*, *atpD*, *atpE*, *atpF*, *atpG*, *atpH*, *atpI*, *ccsA*, *chlI*, *clpC*, *dnaK*, *groEL*, *petA*, *petB*, *petD*, *psaA*, *psaB*, *psaC*, *psaD*, *psaF*, *psaL*, *psbB*, *psbC*, *psbE*, *psbV*, *rpl2*, *rpl3*, *rpl5*, *rpl6*, *rpl14*, *rpl16*, *rpl19*, *rpl20*, *rpl21*, *rpl22*, *rpl23*, *rpl27*, *rpoA*, *rpoB*, *rpoC1*, *rpoC2*, *rps2*, *rps3*, *rps4*, *rps5*, *rps7*, *rps8*, *rps9*, *rps10*, *rps11*, *rps12*, *rps13*, *rps14*, *rps17*, *rps19*, *secA*, *secY*, *tatC*, *tufA*, *ycf3*, *ycf4*, *sufB*, *ccs1*)のアミノ酸配列を用いた。

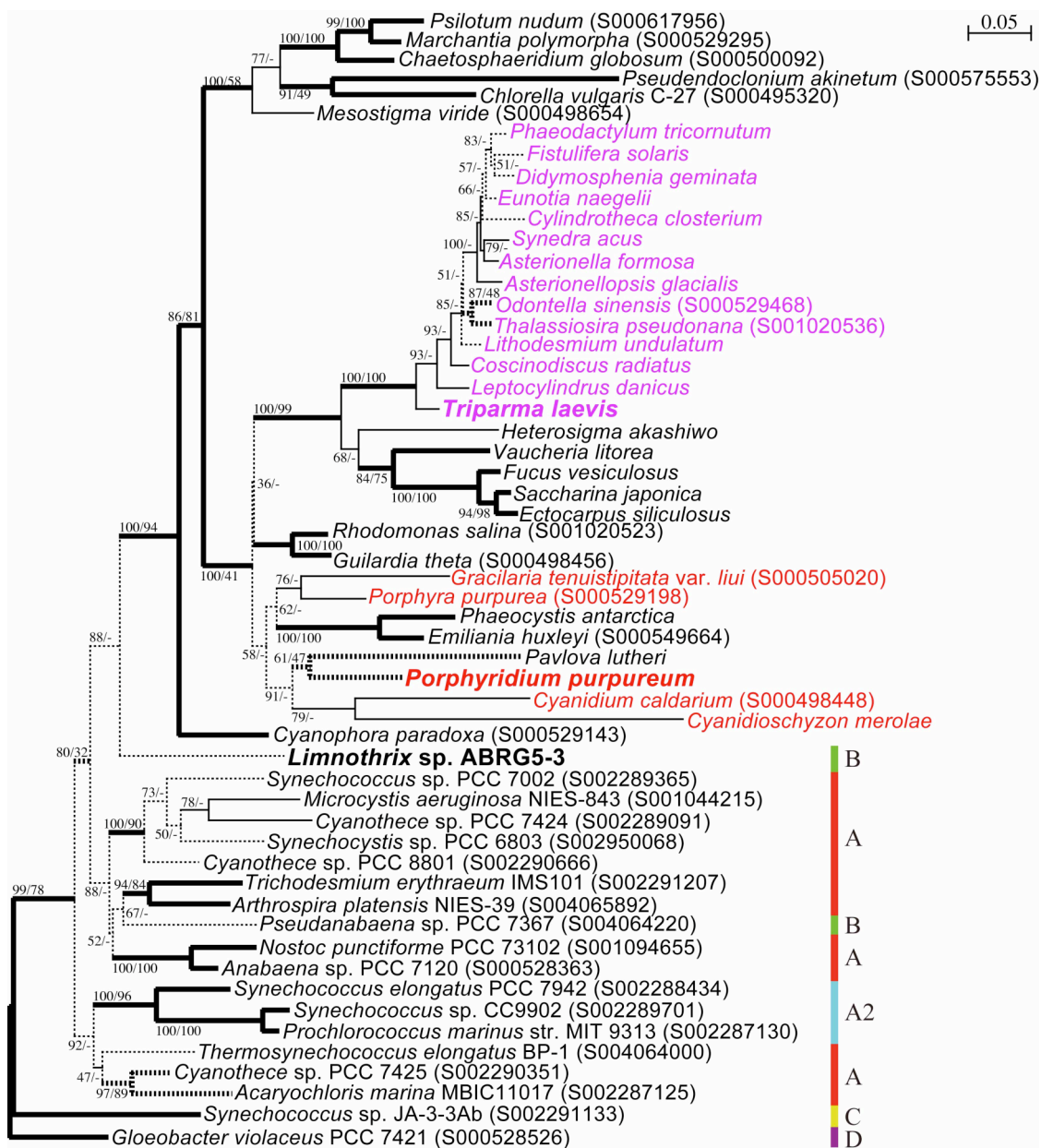


図 5-2 : 16S rDNA 遺伝子の塩基配列を用いた系統樹

右上のスケールバーは枝の長さを示す。それぞれの分岐の肩にある数字は最尤法での信頼度/近隣結合法でのブートストラップ値を示す。最尤法と近隣結合法で支持した枝の分かれ方が同じだった部分は太線で示した。タンパク質遺伝子アミノ酸配列を用いた系統樹と 16S rDNA 遺伝子の塩基配列を用いた系統樹で支持した枝の分かれ方が同じだった部分は実線で示し、違った部分は破線で示した。表 4-4 で分けたシアノバクテリアのグループを、それぞれの系統樹の右側に太線で示した。RDP から手に入れた配列については、RDP の識別番号をそれぞれの種の後に表記した。

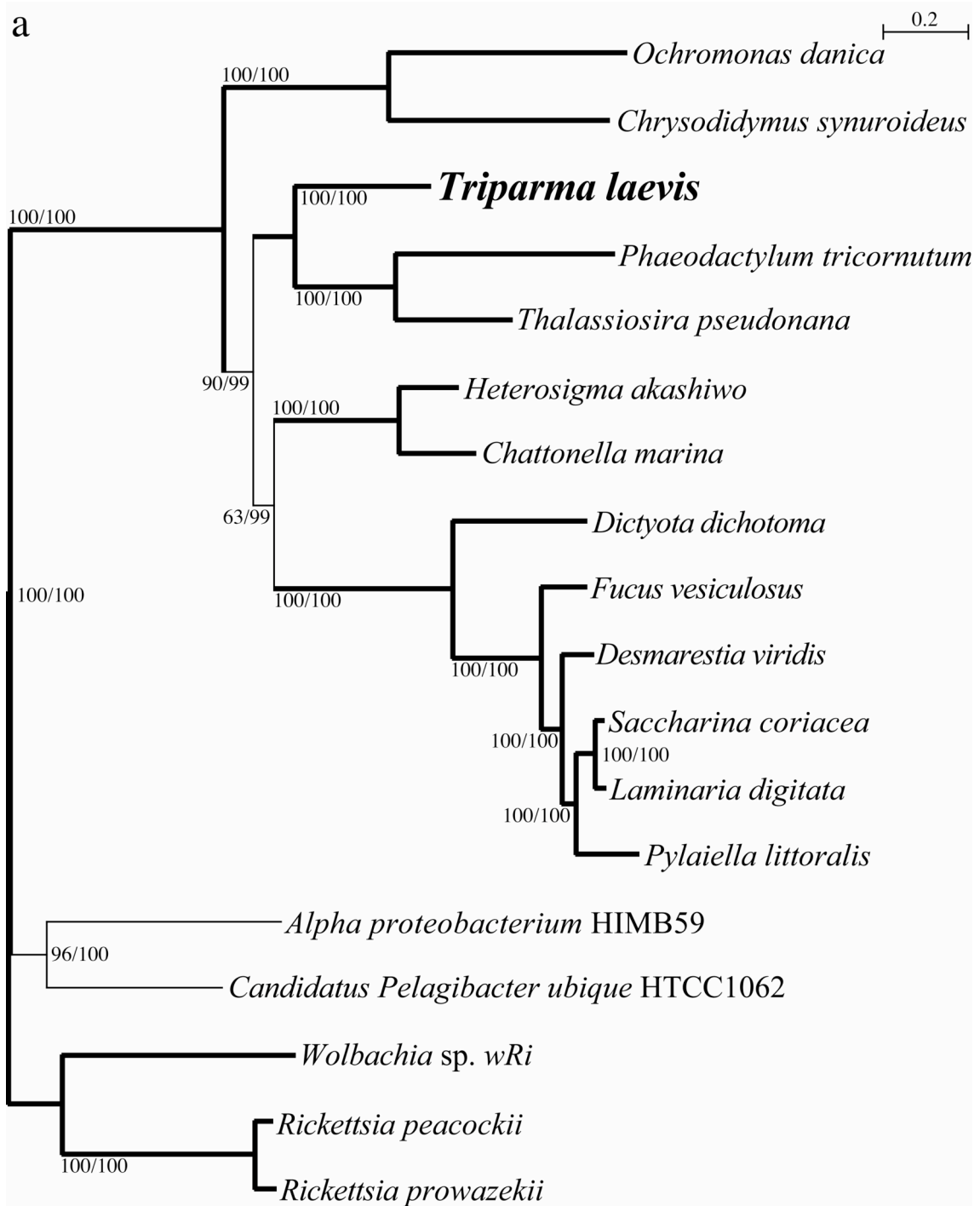


図 5-3 : ミトコンドリアゲノムのタンパク質遺伝子アミノ酸配列を用いた系統樹(前ページ)

右上のスケールバーは枝の長さを示す。それぞれの分岐の肩にある数字は最尤法での信頼度/ベイズ法での事後確率を示す。太線はどちらの方法でも完全に支持された分岐を示す。

a: 不等毛植物門のミトコンドリアゲノムの系統樹

不等毛植物門のミトコンドリアゲノムに保存された 29 個のタンパク質遺伝子(*atp6*, *atp9*, *cob*, *cox1*, *cox3*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *rpl16*, *rps3*, *rps12*, *atp8*, *nad7*, *nad9*, *rpl2*, *rpl6*, *rpl14*, *rps2*, *rps4*, *rps7*, *rps8*, *rps10*, *rps13*, *rps14*, *rps19*)のアミノ酸配列を用いた。

b: 一次共生藻類とクロムアルベオラータのミトコンドリアゲノムの系統樹

一次共生藻類とクロムアルベオラータのミトコンドリアゲノムに保存された 15 個のタンパク質遺伝子(*atp6*, *atp9*, *cob*, *cox1*, *cox3*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *rpl16*, *rps3*, *rps12*)のアミノ酸配列を用いた。

b(i): クリプト藻とハプト藻のデータは入れなかった系統樹

b(ii): クリプト藻とハプト藻のデータを 1 種ずつ含めた系統樹

b(iii): クリプト藻とハプト藻のデータを 2 種ずつ含めた系統樹

第6章

総括

6.1. 本博士論文の総論

本博士論文では、紅藻 *Porphyridium purpureum*、パルマ藻 *Triparma laevis* のオルガネラゲノムおよびシアノバクテリア *Limnothrix* sp. ABRG5-3 のゲノムを解読した。

第2章で解読した *P. purpureum* の葉緑体ゲノムは、他の紅藻の葉緑体ゲノムと比べて多くの独特な特徴を持っていることが分かった。*P. purpureum* の葉緑体ゲノムには、マチュラーゼタンパク質遺伝子が数個コードされていた。移動能力を保持しているものもあり、*P. purpureum* の葉緑体ゲノム内でマチュラーゼタンパク質遺伝子が転移増殖を繰り返した結果、*P. purpureum* の葉緑体ゲノムは多くのイントロンを持つようになったのではないかと考えられた。また、マチュラーゼタンパク質遺伝子が転移増殖を繰り返す過程でゲノムの再編成も起こりやすくなり、ゲノム構造の変異が蓄積した結果、現在の他の紅藻と比べて多くの独特な特徴を持った葉緑体ゲノムになっていったのではないかと考えた。

今回、*P. purpureum* のシーケンシングのアセンブル結果からはミトコンドリアゲノムに相当する配列を見つけ出すことができなかった。DNA を抽出する時点でミトコンドリアの DNA を失ってしまったのかもしれない。近縁種の EST データ(Chan et al. 2011)からもミトコンドリアゲノムに相当する配列を見つけることができなかったため、元々ミトコンドリア DNA の抽出が難しい生物なのかもしれない。

第3章で解読した *T. laevis* のオルガネラゲノムは、他の不等毛植物門と比べて珪藻のオルガネラゲノムと構造やコードされた遺伝子の種類等がよく似ていた。さらに、珪藻では失われた遺伝子を保存していることや珪藻よりも単純なゲノム構造をしていることから、*T. laevis* のオルガネラゲノムは、珪藻の共通祖先に近い構造を現在まで保存しているものなのかもしれないと考えられた。現在、*T. laevis* の核ゲノムについても解読中であるが、そのゲノ

ム情報はパルマ藻への知見を深めるだけでなく、珪藻の核ゲノムと比較解析し、お互いの特徴を明らかにする為に十分有用であると考えられる。

第4章では *Limnothrix* sp. ABRG5-3 のゲノムを解読した。*Limnothrix* sp. ABRG5-3 は特徴的な遺伝子群をコードしたプラスミドを多く持つことが分かった。特に窒素固定関連遺伝子群が実際に働くかは、調べる必要があると考えられる。*Limnothrix* sp. ABRG5-3 は、接合法での形質転換も確認されており(Nishizawa et al. 2010)、今後、遺伝子組み換えによる解析も可能だと考えられる。

第5章では、今まで解読した種を含めて、タンパク質遺伝子アミノ酸配列と 16S rRNA 遺伝子塩基配列で系統解析を行った。葉緑体ゲノムの系統解析結果(図 5-1)から、クロムアルベオラータの葉緑体はチノリモ綱とウシケノリ綱、真正紅藻綱の系統が起源であることが分かった。今後、残りの3綱の葉緑体ゲノムが解読されることで、紅藻の系統だけでなく、クロムアルベオラータの葉緑体が紅藻植物亜門の共通祖先が起源か紅藻植物亜門が現在の綱に分岐した後でどれか特定の綱の祖先が起源になったのかが明らかになるであろう。一方で、ミトコンドリアゲノムは門毎に遺伝子の変異率が大きく異なっており、データセットの選び方によって、樹形が大きく変わること(図 5-3b)などから、あまり広範囲な生物間の系統関係を調べるのには向いていないのではないかと考えられた。

6.2. 細胞内共生に成功したと考えられるシアノバクテリアに関する仮説

葉緑体ゲノムのタンパク質遺伝子アミノ酸配列の系統解析結果に基づいて細胞内共生のもとになったと考えられるシアノバクテリアの系統(第4章のAグループ)と、それ以外の系統(第4章のB, C, Dグループ)で違いがあるのか調べるため、シアノバクテリアゲノムを、コードしているタンパク質遺伝子の種類、rRNA オペロンの構造、リボソームタンパク質遺伝子クラスターの構造で比較した。シアノバクテリアは、進化の過程で過去に何度も多細胞化と単細胞化を繰り返しているようだった。単細胞のシアノバクテリアと多細胞のシアノバクテリアを比較しても特徴的なタンパク質遺伝子の有無は見つけることができず(Shih et al. 2013)、AグループとB, C, Dグループのコード

しているタンパク質遺伝子の種類の間にも、特徴的な違いを見つけることはできなかった。rRNA オペロンの構造も種毎に大きく異なるため、細胞内共生したシアノバクテリアが、当時どのような rRNA オペロン構造をしていたか推測することはできなかった。

一方、リボソームタンパク質遺伝子クラスターは、A グループと B, C, D グループの間で、明確な違いがあった(図 4-9)。A グループは、ほとんどの場合 *rplC-rpmE* がひとまとまりであったのに対して、B, C, D グループは、*infA* と *rpmJ* の間で分断され、種によってはさらに細かく切れていた。バクテリアは、*rpsJ-rplQ* までがプロモーターによって一つにまとまって制御されていることが明らかになっている(Suh et al. 1996, Li et al. 1997)。リボソームタンパク質遺伝子クラスターは、ゲノムにコードしているリボソームタンパク質遺伝子の半数以上を含み、オルガネラゲノムにも保存されている遺伝子クラスターである(Ohta et al. 1997)。また、アーケプラスチダと別の起源を持つとされる *Paulinella chromatophora* のリボソームタンパク質遺伝子クラスターも、クロマトフォアの起源と考えられている *Prochlorococcus/Synechococcus* の系統によく似た構造をしていた(図 4-9、図 6-1)。

このことから、ある仮説を考えた。バクテリアからシアノバクテリアへ分岐する時、*rpsJ-rplQ* のプロモーターが失われ、オペロン構造も一度壊れてしまった。その後、シアノバクテリアでは、*rpsJ* は *rpsL-tufA* のクラスターと一つになった (*Synechococcus* sp. PCC 7502 と *Limnothrix* sp. ABRG5-3 では離れてしまっている)。葉緑体ゲノムも、ほとんどは *rplC-rpmE* がひとまとまりとなっていた(図 2-10)。細胞内共生を成功させるためには、シアノバクテリアと共生先の真核生物の細胞分裂等のタイミングが合う、または共生先の真核生物による調節が、早期に完了する必要があると考えられる。もしも、シアノバクテリアの細胞分裂周期が真核生物よりも早かったら、真核生物の細胞内はやがてシアノバクテリアであふれてしまうだろうし、逆にシアノバクテリアの細胞分裂周期が真核生物よりも遅かったら、真核細胞は細胞内にシアノバクテリアを保持し続けることができなくなってしまうからである。リボソームタンパク質遺伝子クラスターがばらばらなシアノバクテリアも、必要な遺伝子が揃ってさえいれば独立して生存することは可能だと考えられる。しかし、リボソームタンパク質遺伝子群がクラスターとしてまとまっている

ことが、共生先の真核生物にリボソームタンパク質遺伝子群を制御されることを容易にし、真核生物に細胞内共生することがスムーズに行えるようになったのではないかと考えられた。つまり、葉緑体の共通祖先と思われるシアノバクテリアは、細胞内共生に成功した当时には既にリボソームタンパク質遺伝子群がクラスターとしてまとまっており、そのことが細胞内共生の成功につながったのではないかと考えた。*P. chromatophora* のクロマトフォアがかつて細胞内共生を成功させたのもそのためではないかと考えた(図 6-2)。最近、珪藻 *Rhopalodiaceae* に見られる細胞内構造体スフェロイドボディのゲノムが解読された(Nakayama et al. 2014)。この構造体は、シアノバクテリアの *Cyanothece* が起源であるとされている。解読された *Epithemia turgida* のスフェロイドボディゲノム(accession number: AP012549)のリボソームタンパク質遺伝子クラスターを確認したところ、*rplC-rpmE* と *rpsL-rpsJ* は離れて存在していた。*rplC-rpmE* と *rpsL-rpsJ* もつながっていないなければならないという訳ではないのかもしれない。今後、細胞周期とリボソームタンパク質遺伝子クラスターの発現の関係を調べることで、リボソームタンパク質遺伝子群が核に制御されているかどうか確かめることができるかもしれない。

6.3. 葉緑体ゲノム構造から見る紅藻とクロムアルベオラータの進化的関係

紅藻を起源とする二次共生藻、クロムアルベオラータの葉緑体ゲノムの多くは、逆向きに反復した同一の rRNA オペロン構造を持っていた。しかし、これと同じ rRNA オペロン構造は、報告済みの紅藻の葉緑体ゲノムの中では今のところ見つかっていない(図 2-2)。*P. purpureum* の葉緑体ゲノムは、rRNA 遺伝子達は逆向きで反復した配列だが、*rrs* と *rri* の遺伝子の間にある tRNA 遺伝子がそれぞれ違うものとなっている。

rRNA オペロン構造と葉緑体ゲノムの系統解析の結果から、葉緑体ゲノムのゲノム構造に関して、以下のような進化的シナリオを想定した。まず、紅藻葉緑体の共通祖先は2つの逆向きで反復した rRNA オペロンを持っていた。イデユコゴメ綱は紅藻植物亜門から分岐した後で、rRNA オペロンを1個失った。それから、紅藻植物亜門の共通祖先が起源となる二次共生が起きた。二次共生が起きた回数が一度か複数かは今回の解析では明らかにできなかつ

たが、この時期の紅藻を細胞内共生したものであると考えられる。二次共生藻の葉緑体ゲノムでは、2つの逆向きで反復した rRNA オペロンが保存されている。ウシケノリ綱では1つの rRNA オペロンが逆向きから直列に反転するという再編成が起こり、真正紅藻綱ではその内の1個が失われた。キノリモ綱では大規模なゲノムの再編成が起こり、非同一な2個の rRNA オペロンを持つようになった。これは、2つの rRNA オペロンからそれぞれ別の tRNA 遺伝子が失われたことで、生まれた構造なのかもしれない。

他の3綱(オオイシソウ綱、ロデラ綱、ベニミドロ綱)の rRNA オペロン構造はまだ明らかになっていない。また、不等毛植物門の葉緑体ゲノムに共通して保存されている *ycf66* をコードしている紅藻の葉緑体ゲノムも、まだ報告されていない。今後、紅藻の全ての綱の葉緑体ゲノムが解読した時に、系統解析だけでなくコードされた遺伝子の種類やゲノムの構造を比較することも、紅藻とクロムアルベオラータの進化的関係の理解に役立つと考えられる。

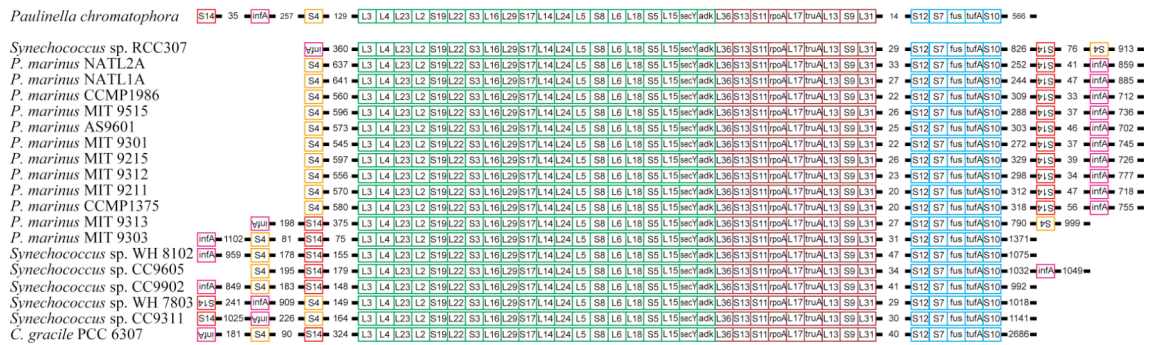


図 6-1: *Paulinella chromatophora* のクロマトフォアゲノムのリボソームタンパク質遺伝子クラスター

上下逆になっている遺伝子は逆方向にコードされた遺伝子を示す。遺伝子名の間にある数字はそれぞれリボソームタンパク質遺伝子クラスター以外の遺伝子を間に挟んだ部分のリボソームタンパク質遺伝子クラスター遺伝子間の塩基対数(単位は kbp)を示す。比較のため、下に *Prochlorococcus/Synechococcus* グループのリボソームタンパク質遺伝子クラスターを示した。

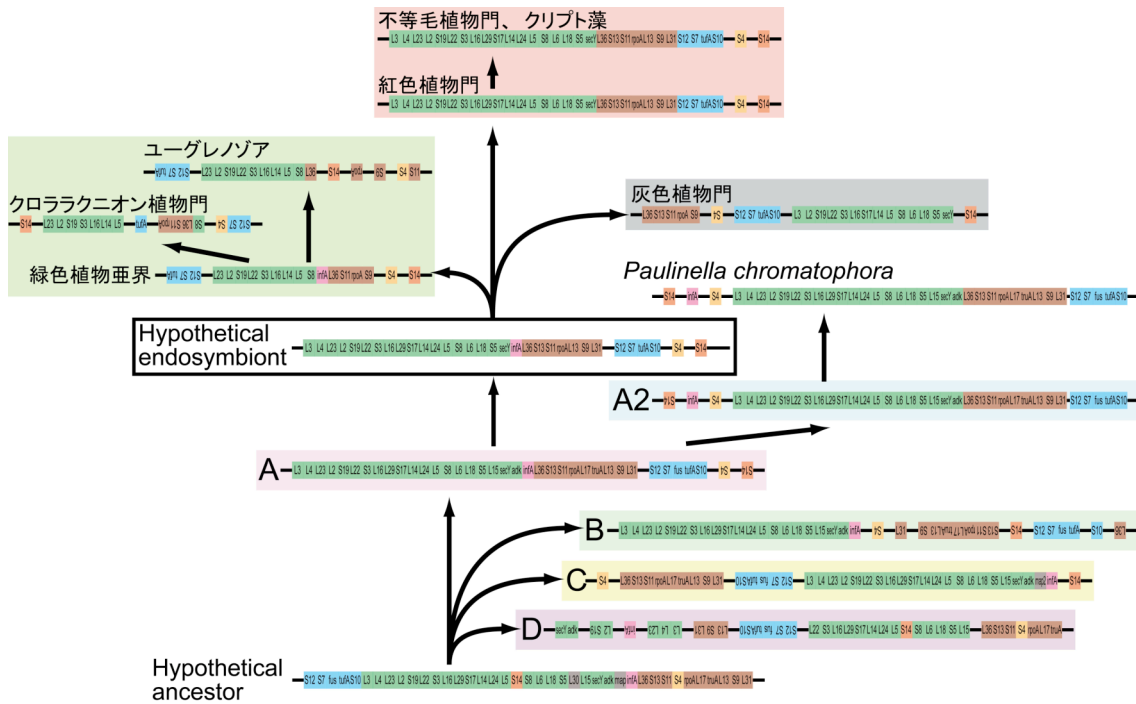


図 6-2 : リボソームタンパク質遺伝子クラスターから見るシアノバクテリアの進化と細胞内共生

シアノバクテリアの進化に伴うリボソームタンパク質遺伝子クラスターの変化を模式図で示した。シアノバクテリアとバクテリアのリボソームタンパク質遺伝子クラスターを基に考えた祖先的シアノバクテリアのリボソームタンパク質遺伝子クラスターの構造から、進化に伴い A, B, C, D グループのような構造へ、それぞれ変化した。アーケプラスチダの葉緑体も *Paulinella chromatophora* のクロマトフォアも、*rplC-rpmE* と *rpsL-rpsJ* がそれぞれひとまとまりになった A グループのシアノバクテリアが細胞内共生したことに由来していることを示す。ハプト藻は、この図では示していないが、クリプト藻や不等毛植物門と比べていくつかの遺伝子が失われている(図 2-10)。

参考文献

英文

- Abascal F, Zardoya R, Posada D (2005) ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* 21:2104-2105
- Acinas SG, Haverkamp TH, Huisman J, Stal LJ (2009) Phenotypic and genetic diversification of *Pseudanabaena* spp. (cyanobacteria). *ISME J* 3:31-46
- Adl SM, Simpson AGB, Farmer MA, Andersen RA, Anderson OR, Barta JR, Bowser SS, Brugerolle G, Fensome RA, Fredericq S, James TY, Karpov S, Kugrens P, Krug J, Lane CE, Lewis LA, Lodge J, Lynn DH, Mann DG, McCourt RM, Mendoza L, Moestrup O, Mozley-Standridge SE, Nerad TA, Shearer CA, Smirnov AV, Spiegel FW, Taylor MFJR (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J Eukaryot Microbiol* 52:399-451
- Alkatib S, Fleischmann TT, Scharff LB, Bock R (2012) Evolutionary constraints on the plastid tRNA set decoding methionine and isoleucine. *Nucleic Acids Res* 40:6713-6724
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389-3402
- Archibald JM (2009) The puzzle of plastid evolution. *Curr Biol* 19:R81-R88
- Barbier G, Oesterhelt C, Larson MD, Halgren RG, Wilkerson C, Garavito RM, Benning C, Weber APM (2005) Comparative genomics of two closely related unicellular thermo-acidophilic red algae, *Galdieria sulphuraria* and *Cyanidioschyzon merolae*, reveals the molecular basis of the metabolic flexibility of *Galdieria sulphuraria* and significant differences in carbohydrate metabolism of both algae. *Plant Physiol* 137:460-474
- Bernard C, Thomas JC, Mazel D, Mousseau A, Castets AM, Tandeau de Marsac N, Dubacq JP (1992) Characterization of the genes encoding phycoerythrin in the red

- alga *Rhodella violacea*: evidence for a splitting of the *rpeB* gene by an intron. Proc Natl Acad Sci USA 89:9564-9568
- Bhattacharya D, Price DC, Chan CX, Qiu H, Rose N, Ball S, Weber APM, Arias MC, Henrissat B, Coutinho PM, Krishnan A, Zäuner S, Morath S, Hilliou F, Egizi A, Perrineau MM, Yoon HS (2013) Genome of the red alga *Porphyridium purpureum*. Nat Commun 4:1941
- Bhaya D, Grossman AR, Steunou AS, Khuri N, Cohan FM, Hamamura N, Melendrez MC, Bateson MM, Ward DM, Heidelberg JF (2007) Population level functional diversity in a microbial community revealed by comparative genomic and metagenomic analyses. ISME J 1:703-713
- Bogorad L (1967) Chloroplast structure and development. In *Harvesting the Sun: Photosynthesis in Plant Life*. San Pietro FA, Greer T, Army J (ed) New York: Academic Press pp. 191-210
- Booth BC, Lewin J, Norris RE (1980) Siliceous nanoplankton. 1. Newly discovered cysts from the Gulf of Alaska. Mar Biol 58:205-209
- Booth BC, Marchant HJ (1987) Parmales, a new order of marine chrysophytes, with descriptions of three new genera and seven new species. J Phycol 23:245-260
- Burki F, Okamoto N, Pombert JF, Keeling PJ (2012) The evolutionary history of haptophytes and cryptophytes: phylogenomic evidence for separate origins. Proc Biol Sci 279:2246-2254
- Burki F, Shalchian-Tabrizi K, Minge M, Skjaeveland Å, Nikolaev SI, Jakobsen KS, Pawlowski J (2007) Phylogenomics reshuffles the eukaryotic supergroups. PLoS One 8:e790
- Burki F, Shalchian-Tabrizi K, Pawlowski J (2008) Phylogenomics reveals a new 'megagroup' including most photosynthetic eukaryotes. Biol Lett 4:366-369
- Butterfield NJ (2000) *Bangiomorpha pubescens* n. gen., n. sp.: implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. Paleobiology 26:386-404
- Cavalier-Smith T (1981) Eukaryote kingdoms: seven or nine? Biosystems 14:461-481

- Cavalier-Smith T (1999) Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. *J Eukaryot Microbiol* 46:347-366
- Cech TR (1988) Conserved sequences and structures of group I introns: building an active site for RNA catalysis -- a review. *Gene* 73:259-271
- Chan CX, Yang EC, Banerjee T, Yoon HS, Martone PT, Estevez JM, Bhattacharya D (2011) Red and green algal monophyly and extensive gene sharing found in a rich repertoire of red algal genes. *Curr Biol* 21:328-333
- Chun EHL, Vaughan MH Jr., Rich A (1963) The isolation and characterization of DNA associated with chloroplast preparations. *J Mol Biol* 7:130-141
- Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR, Tiedje JM (2014) Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucl Acids Res* 42:D633-D642
- Cozy LM, Callahan SM (2014) The *hmp* chemotaxis cluster regulates gliding in the filamentous cyanobacterium *Nostoc punctiforme*. *Mol Microbiol* 92:213-216
- Criscuolo A, Gribaldo S (2011) Large-scale phylogenomic analyses indicate a deep origin of primary plastids within cyanobacteria. *Mol Biol Evol* 28:3019-3022
- Dagan T, Roettger M, Stucken K, Landan G, Koch R, Major P, Gould SB, Goremykin VV, Rippka R, Tandeau de Marsac N, Gugger M, Lockhart PJ, Allen JF, Brune I, Maus I, Pühler A, Martin WF (2013) Genomes of Stigonematalean cyanobacteria (subsection V) and the evolution of oxygenic photosynthesis from prokaryotes to plastids. *PLoS One* 8:e68246
- Delwiche CF, Palmer JD (1996) Rampant horizontal transfer and duplication of rubisco genes in eubacteria and plastids. *Mol Biol Evol* 13:873-882
- DePriest MS, Bhattacharya D, López-Bautista JM (2013) The plastid genome of the red macroalga *Grateloupia taiwanensis* (Halymeniaceae). *PLoS One* 8:e68246
- Deusch O, Landan G, Roettger M, Gruenheit N, Kowallik KV, Allen JF, Martin W, Dagan T (2008) Genes of cyanobacterial origin in plant nuclear genomes point to a heterocyst-forming plastid ancestor. *Mol Biol Evol* 25:748-761

- Drouin G, Daoud H, Xia J (2008) Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. *Mol Phylogenet Evol* 49:827-831
- Edelman M, Cowan CA, Epstein HT, Schiff JA (1964) Studies of chloroplast development in euglena. VIII. Chloroplast associated DNA. *Proc Natl Acad Sci U S A* 52:1214-1219
- Edelman M, Schiff JA, Epstein HT (1965) Studies of chloroplast development in euglena. XII. Two types of satellite DNA. *J Mol Biol* 11:769-774
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792-1797
- Falcón LI, Magallón S, Castillo A (2010) Dating the cyanobacterial ancestor of the chloroplast. *ISME J* 4:777-783
- Felsenstein J (1978) Cases in which parsimony or compatibility methods will be positively misleading. *Syst Zool* 27:401-410
- Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* 17:368-376
- Fox TD (1979) Five TGA “stop” codons occur within the translated sequence of the yeast mitochondrial gene for cytochrome c oxidase subunit II. *Proc Natl Acad Sci U S A* 76:6534-6538
- Francis MA, Dudock BS (1982) Nucleotide sequence of a spinach chloroplast isoleucine tRNA. *J Biol Chem* 257:11195-11198
- Gabrielson PW, Garbary DJ, Scagel RF (1985) The nature of the ancestral red alga: inferences from a cladistics analysis. *Biosystems* 18:335-346
- Gabrielson PW, Garbary DJ, Sommerfeld MR, Townsend RA, Tyler PL (1990) Phylum rhodophyta. In Margulis L, Corliss JO, Melkonian M, Chapman DJ (Eds) *Handbook of protocista: the structure, cultivation habitats and life histories of the eukaryotic microorganisms and their descendants exclusive of animals, plants and fungi*. Jones and Bartlett Pub., Boston pp. 914
- Garbary DJ, Gabrielson PW (1990) Taxonomy and evolution. In Cole KM, Sheath RG (Eds) *Biology of the red algae*. Cambridge University Press, New York, pp. 477-498

- Geyer CJ (1991) Markov chain Monte Carlo maximum likelihood. In Computing Science and Statistics: Proc 23rd Symp Interface (ed. Keramidas EM) Interface Foundation, Fairfax Station, VA pp. 156-163
- Gibson A, Gowri-Shankar V, Higgs PG, Rattray M (2005) A comprehensive analysis of mammalian mitochondrial genome base composition and improved phylogenetic methods. *Mol Biol Evol* 22:251-264
- Glöckner G, Rosenthal A, Valentin K (2000) The structure and gene repertoire of an ancient red algal plastid genome. *J Mol Evol* 51:382-390
- Graham LD, Wilcox LW (2000) *Algae*. Prentice-Hall Inc. NJ, pp. 640
- Greenwood AD (1974) The Cryptophyta in relation to phylogeny and photosynthesis. In Sanders J, Goodchild D (eds) 8th International Congress of Electron microscopy. Australian Academy of Sciences, Canberra, pp. 566-567
- Grissa I, Vergnaud G, Pourcel C (2007) CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* 35:W52-W57
- Guillard RRL, Ryther JH (1962) Studies of marine planktonic diatoms I. *Cyclotella nana* Hustedt and *Detonula confervacea* (Cleve) Gran. *Can J Microbiol* 8:229-239
- Guillou L, Chrétiennot-Dinet M-J, Medlin LK, Claustre H, Loiseaux-de Göer S, Vaillot D (1999) *Bolidomonas*: a new genus with two species belonging to a new algal class, the Bolidophyceae (Heterokonta). *J Phycol* 35:368-381
- Hackett JD, Yoon HS, Li S, Reyes-Prieto A, Rümmele SE, Bhattacharya D (2007) Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of Rhizaria with Chromalveolates. *Mol Biol Evol* 24:1702-1713
- Hagopian JC, Reis M, Kitajima JP, Bhattacharya D, de Oliveria MC (2004) Comparative analysis of the complete plastid genome sequence of the red alga *Gracilaria tenuistipitata* var. *liui* provides insights into the evolution of Rhodoplasts and their relationship to other plastids. *J Mol Evol* 59:464-477
- Harper JT, Keeling PJ (2003) Nucleus-encoded, plastid-targeted glyceraldehyde-3-phosphate dehydrogenase (GAPDH) indicates a single origin for chromalveolate plastids. *Mol Biol Evol* 20:1730-1735

- Hastings WK (1970) Monte Carlo sampling methods using Markov chains and their application. *Biometrika* 57:97-109
- Hibberd DJ, Norris RE (1984) Cytology and ultrastructure of *Chlorarachnion reptans* (Chlorarachniophyta *divisio nova*, Chlorarachniophyceae *classis nova*). *J Phycol* 20:310-330
- Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754-755
- Ichinomiya M, Gomi Y, Nakamachi M, Ota T, Kobari T (2010) Temporal patterns in silica deposition among siliceous plankton during the spring bloom in the Oyashio region. *Deep-Sea Res II* 57:1665-1670
- Ichinomiya M, Yoshikawa S, Kamiya M, Ohki K, Takaichi S, Kuwata A (2011) Isolation and characterization of Palmales (Heterokonta/Heterokontophyta/Stramenopiles) from the Oyashio region, western North Pacific. *J Phycol* 47:144-151
- Janouškovec J, Liu SL, Martone PT, Carré W, Leblanc C, Collén J, Keeling PJ (2013) Evolution of red algal plastid genomes: ancient architectures, introns, horizontal gene transfer, and taxonomic utility of plastid markers. *PLoS One* 8:e59001
- Jobb G, von Haeseier A, Strimmer K (2004) TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. *BMC Evol Biol* 4:18
- Kanesaki Y, Shiwa Y, Tajima N, Suzuki M, Watanabe S, Sato N, Ikeuchi M, Yoshikawa H (2012) Identification of substrain-specific mutations by massively parallel whole-genome resequencing of *Synechocystis* sp. PCC 6803. *DNA Res* 19:67-79
- Keeling PJ, Burger G, Durnford DG, Lang BF, Lee RW, Pearlman RE, Roger AJ, Gray MW (2005) The tree of eukaryotes. *Trends Ecol Evol* 20:670-676
- Kim E, Archibald JM (2009) Diversity and evolution of plastids and their genomes. *The Chloroplast - Interactions with the Environment*. In: Aronsson H, Sandelius AS (ed) *Plant Cell Monographs*, vol. 13. Springer, Berlin, pp. 1-39
- Kirk JTO, Tilney-Bassett RAE (1967) *The plastids*. London and San Francisco: Freeman WH

- Kitazaki C, Numano S, Takanezawa A, Nishizawa T, Shirai M, Asayama M (2013) Characterization of lysis of the multicellular Cyanobacterium *Limnothrix/Pseudanabaena* sp. strain ABRG5-3. *Biosci Biotechnol Biochem* 77:2339-2347
- Konishi T, Shinohara K, Yamada K, Sasaki Y (1996) Acetyl-CoA carboxylase in higher plants: most plants other than gramineae have both the prokaryotic and the eukaryotic forms of this enzyme. *Plant Cell Physiol* 37:117-122
- Krinsky NI (1966) The role of carotenoid pigments as protective agents against photosensitized oxidations in chloroplasts. In *Biochemistry of Chloroplasts*, Vol. 1. Goodwin TW (ed) New York and London: Academic Press pp. 423-430
- Lang BF, Laforest MJ, Burger G (2007) Mitochondrial introns: a critical view. *Trends Genet* 23:119-125
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23:2947-2948
- Larsen J (1992) Endocytobiotic consortia with dinoflagellate hosts. In Reisser W (ed) *Algae and Symbioses*. Biopress, Bristol, pp. 427-442
- Le SQ, Gascuel O (2008) An improved general amino acid replacement matrix. *Mol Biol Evol* 25:1307-1320
- Lederberg J (1952) Cell genetics and hereditary symbiosis. *Physiol Rev* 32:403-430
- Leff J, Mandel M, Epstein HT, Schiff JA (1963) DNA satellite from cells of green and aplastidic algae. *Biochem Biophys Res Commun* 13:126-130
- Li X, Lindahl L, Sha Y, Zengel JM (1997) Analysis of the *Bacillus subtilis* S10 ribosomal protein gene cluster identifies two promoters that may be responsible for transcription of the entire 15-kilobase S10-*spc-alpha* cluster. *J Bacteriol* 179:7046-7054
- Livne A, Sukenik A (1990) Acetyl-coenzyme A carboxylase from the marine prymnesiophyte *Isochrysis galbana*. *Plant Cell Physiol* 31:851-858
- Lommer M, Roy AS, Schilhabel M, Schreiber S, Rosenstiel P, LaRoche J (2010) Recent transfer of an iron-regulated gene from the plastid to the nuclear genome in an oceanic diatom adapted to chronic iron limitation. *BMC Genomics* 11:718

- Lovejoy C, Massana R, Pedrós-Alió C (2006) Diversity and distribution of marine microbial eukaryotes in the Arctic Ocean and adjacent seas. *Appl Environ Microbiol* 72:3085-3095
- Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25:955-964
- Marchant HJ, McEldowney A (1986) Nanoplanktonic siliceous cysts from Antarctica are algae. *Mar Biol* 92:53-57
- Margulis L (1970) *Origin of Eukaryotic Cells*. NewHaven and London: Yale University Press
- Margulis L (1993) *Symbiosis in Cell Evolution* 2nd ed. *Microbial Communities in the Archean and Proteozoic eons*. Freeman WH
- Mauriello EM, Mignot T, Yang Z, Zusman DR (2010) Gliding motility revisited: how do the myxobacteria move without flagella? *Microbiol Mol Biol Rev* 74:229-249
- Marin B, Nowack EC, Melkonian M (2005) A plastid in the making: evidence for a second primary endosymbiosis. *Protist* 156:425-432
- Mereschkowsky C (1905) Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biologisches Centralblatt* 25:593-604
- Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equations of state calculations by fast computing machines. *J Chem Phys* 21:1087-1092
- Nakamura Y, Kaneko T, Sato S, Mimuro M, Miyashita H, Tsuchiya T, Sasamoto S, Watanabe A, Kawashima K, Kishida Y, Kiyokawa C, Kohara M, Matsumoto M, Matsuno A, Nakazaki N, Shimpo S, Takeuchi C, Yamada M, Tabata S (2003) Complete genome structure of *Gloeobacter violaceus* PCC 7421, a cyanobacterium that lacks thylakoids. *Genome Res* 13:137-145
- Nakanishi K, Bonnefond L, Kimura S, Suzuki T, Ishitani R, Nureki O (2009) Structural basis for translational fidelity ensured by transfer RNA lysidine synthetase. *Nature* 461:1144-1148
- Nakayama T, Kamikawa R, Tanifuji G, Kashiyama Y, Ohkouchi N, Archibald JM, Inagaki Y (2014) Complete genome of a nonphotosynthetic cyanobacterium in a diatom reveals recent adaptations to an intracellular lifestyle. *Proc Natl Acad Sci U S A* 111:11407-11412

- Nelson DM, Tréguer P, Brzezinski MA, Leynaert A, Quéguiner B (1995) Production and dissolution of biogenic silica in the ocean: revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Global Biogeochem Cycles* 9:359-372
- Neuhaus H, Link G (1987) The chloroplast tRNA^{Lys}(UUU) gene from mustard (*Sinapis alba*) contains a class II intron potentially coding for a maturase-related polypeptide. *Curr Genet* 11:251-257
- Nishizawa T, Hanami T, Hirano E, Miura T, Watanabe Y, Takanezawa A, Komatsuzaki M, Ohta H, Shirai M, Asayama M (2010) Isolation and molecular characterization of a multicellular cyanobacterium, *Limnothrix/Pseudanabaena* sp. strain ABRG5-3. *Biosci Biotechnol Biochem* 74:1827-1835
- Noguchi H, Taniguchi T, Itoh T (2008) MetaGeneAnnotator: detecting species-specific patterns of ribosomal binding site for precise gene prediction in anonymous prokaryotic and phage genomes. *DNA Res* 15:387-396
- O'Brien EA, Zhang Y, Wang E, Marie V, Badejoko W, Lang BF, Burger G (2009) GOBASE: an organelle genome database. *Nucleic Acids Res* 37:D946-D950
- Ohta N, Matsuzaki M, Misumi O, Miyagishima SY, Nozaki H, Tanaka K, Shin-I T, Kohara Y, Kuroiwa T (2003) Complete sequence and analysis of the plastid genome of the unicellular red alga *Cyanidioschyzon merolae*. *DNA Res* 10:67-77
- Ohta N, Sato N, Nozaki H, Kuroiwa T (1997) Analysis of the cluster of ribosomal protein genes in the plastid genome of a unicellular red alga *Cyanidioschyzon merolae*: translocation of the *str* cluster as an early event in the rhodophyte-chromophyte lineage of plastid evolution. *J Mol Evol* 45:688-695
- Okaichi T, Nishio S, Imatomi Y (1982) Collection and mass culture. *Jpn Fish Soc* (ed) *In Toxic phytoplankton - occurrence, mode of action, and toxins*. Koseisya-Koseikaku, Tokyo, pp. 22-34
- Oliveira MC, Bhattacharya D (2000) Phylogeny of the Bangiophycidae (Rhodophyta) and the secondary endosymbiotic origin of algal plastids. *Am J Bot* 87:482-492
- Oudot-Le Secq MP, Green BR (2011) Complex repeat structures and novel features in the mitochondrial genomes of the diatoms *Phaeodactylum tricorutum* and *Thalassiosira pseudonana*. *476:20-26*

- Oudot-Le Secq MP, Grimwood J, Shapiro H, Armbrust EV, Bowler C, Green BR (2007) Chloroplast genomes of the diatoms *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*: comparison with other plastid genomes of the red lineage. *Mol Genet Genomics* 277:427-439
- Oudot-Le Secq MP, Loiseaux-de Goër S, Stam WT, Olsen JL (2006) Complete mitochondrial genomes of the three brown algae (Heterokonta: Phaeophyceae) *Dictyota dichotoma*, *Fucus vesiculosus* and *Desmarestia viridis*. *Curr Genet* 49:47-58
- Patel AK, Laroche C, Marcati A, Ursu AV, Jubeau S, Marchal L, Petit E, Djelveh G, Michaud P (2013) Separation and fractionation of exopolysaccharides from *Porphyridium cruentum*. *Bioresour Technol* 145:345-350
- Patron NJ, Rogers MB, Keeling PJ (2004) Gene replacement of fructose-1,6-bisphosphate aldolase supports the hypothesis of a single photosynthetic ancestor of chromalveolates. *Eukaryot Cell* 3:1169-1175
- Qiu H, Yang EC, Bhattacharya D, Yoon HS (2012) Ancient gene paralogy may mislead inference of plastid phylogeny. *Mol Biol Evol* 29:3333-3343
- Rannala B, Yang Z (1996) Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. *J Mol Evol* 43:304-311
- Reith M, Munholland J (1995) Complete nucleotide sequence of the *Porphyra purpurea* chloroplast genome. *Plant Mol Biol Rep* 13:333-335
- Rice DW, Palmer JD (2006) An exceptional horizontal gene transfer in plastids: gene replacement by a distant bacterial paralog and evidence that haptophyte and cryptophyte plastids are sisters. *BMC Biol* 4:31
- Rippka R, Deruelles J, Waterbury JB, Herdman M, Stanier RY (1979) Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *J Gen Microbiol* 111:1-61
- Ris H, Plaut W (1962) Ultrastructure of DNA-containing areas in the chloroplast of *Chlamydomonas*. *J Cell Biol* 13:383-391
- Risser DD, Chew WG, Meeks JC (2014) Genetic characterization of the *hmp* locus, a chemotaxis-like gene cluster that regulates hormogonium development and motility in *Nostoc punctiforme*. *Mol Microbiol* 92:222-233

- Risser DD, Meeks JC (2013) Comparative transcriptomics with a motility-deficient mutant leads to identification of a novel polysaccharide secretion system in *Nostoc punctiforme*. *Mol Microbiol* 87:884-893
- Roessler PG, Ohlrogge JB (1993) Cloning and characterization of the gene that encodes acetyl-coenzyme A carboxylase in the alga *Cyclotella cryptica*. *J Biol Chem* 268:19254-19259
- Rogers MB, Gilson PR, Su V, McFadden GI, Keeling PJ (2007) The complete chloroplast genome of the chlorarachniophyte *Bigeloviella natans*: evidence for independent origins of chlorarachniophyte and euglenid secondary endosymbionts. *Mol Biol Evol* 24:54-62
- Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572-1574
- Ruck EC, Nakov T, Jansen RK, Theriot EC, Alverson AJ (2014) Serial gene losses and foreign DNA underlie size and sequence variation in the plastid genomes of diatoms. *Genome Biol Evol* 6:644-654
- Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B (2000) Artemis: sequence visualization and annotation. *Bioinformatics* 16:944-945
- Sagan L (1967) On the origin of mitosing cells. *J Theor Biol* 14:255-274
- Sagan L, Ben-Shaul Y, Schiff J, Epstein HT (1965) Autoradiographic localization of the DNA in the chloroplasts of *Euglena*. *J Plant Physiol* 40:1257-1259
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406-425
- Sasaki NV, Sato N (2010) CyanoClust: comparative genome resources of cyanobacteria and plastids. *Database* 2010:bap025
- Sato N (2009) Gclust: *trans*-kingdom classification of proteins using automatic individual threshold setting. *Bioinformatics* 25:599-605
- Sato N (2000) SISEQ: manipulation of multiple sequence and large database files for common platforms. *Bioinformatics* 16:180-181

- Saunders GW, Hommersand MH (2004) Assessing red algal supraordinal diversity and taxonomy in the context of contemporary systematic data. *Am J Bot* 91:1494-1507
- Schimper AFW (1883) Über die Entwicklung der Chlorophylkörner und Farbkörner. *Bot Zeit* 41:105-120, 121-136, 137-152, 153-162
- Shih PM, Wu D, Latifi A, Axen SD, Fewer DP, Talla E, Calteau A, Cai F, Tandeau de Marsac N, Rippka R, Herdman M, Sivonen K, Coursin T, Laurent T, Goodwin L, Nolan M, Davenport KW, Han CS, Rubin EM, Eisen JA, Woyke T, Gugger M, Kerfeld CA (2013) Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proc Natl Acad Sci U S A* 110:1053-1058
- Shimizu K, Del Amo Y, Brzezinski MA, Stucky GD, Morse DE (2001) A novel fluorescent silica tracer for biological silicification studies. *Chem Biol* 8:1051-1060
- Schnepf E (1992) From parasitism to symbiosis: the dinoflagellate example. In Reisser W (ed) *Algae and Symbiosis*. Biopress, Bristol, pp. 699-710
- Silver MW, Mitchell JG, Ringo DL (1980) Siliceous nanoplankton. II. Newly discovered cysts and abundant choanoflagellates from the Weddell Sea, Antarctica. *Mar Biol* 58:211-217
- Smith DR, Arrigo KR, Alderkamp AC, Allen AE (2014) Massive difference in synonymous substitution rates among mitochondrial, plastid, and nuclear genes of *Phaeocystis* algae. *Mol Phylogenet Evol* 71:36-40
- Smith DR, Hua J, Lee RW, Keeling PJ (2012) Relative rates of evolution among the three genetic compartments of the red alga *Porphyra* differ from those of green plants and do not correlate with genome architecture. *Mol Phylogenet Evol* 65:339-344
- Soma A, Ikeuchi Y, Kanemasa S, Kobayashi K, Ogasawara N, Ote T, Kato J, Watanabe K, Sekine Y, Suzuki T (2003) An RNA-modifying enzyme that governs both the codon and amino acid specificities of isoleucine tRNA. *Mol Cell* 12:689-698

- Stirewalt VL, Michalowski CB, Löffelhardt W, Bohnert HJ, Bryant DA (1995) Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*. *Plant Mol Biol Rep* 13:327-332
- Strimmer K, Rambaut A (2002) Inferring confidence sets of possibly misspecified gene trees. *Proc Biol Sci* 269:137-142
- Suh JW, Boylan SA, Oh SH, Price CW (1996) Genetic and transcriptional organization of the *Bacillus subtilis* *spc-alpha* region. *Gene* 169:17-23
- Surzycki SJ (1969) Genetic functions of the chloroplast of *Chlamydomonas reinhardtii*: Effect of rifampin on Chloroplast DNA-dependent RNA polymerase. *Proc Natl Acad Sci U S A* 63:1327-1334
- Tajima N, Sato S, Maruyama F, Kaneko T, Sasaki NV, Kurokawa K, Ohta H, Kanesaki Y, Yoshikawa H, Tabata S, Ikeuchi M, Sato N (2011) Genomic structure of the cyanobacterium *Synechocystis* sp. PCC 6803 strain GT-S. *DNA Res* 18:393-399
- Tajima N, Sato S, Maruyama F, Kurokawa K, Ohta H, Tabata S, Sekine K, Moriyama T, Sato N (2014) Analysis of the complete plastid genome of the unicellular red alga *Porphyridium purpureum*. *J Plant Res* 127:389-397
- Takahashi F, Okabe Y, Nakada T, Sekimoto H, Ito M, Kataoka H, Nozaki H (2007) Origins of the secondary plastids of Euglenophyta and Chlorarachniophyta as revealed by an analysis of the plastid-targeting, nuclear-encoded gene *psbO*. *J Phycol* 43:1302-1309
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30:2725-2729
- Taylor FJR (1976) Autogenous theories for the origin of eukaryotes. *Taxon* 25:377-390
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673-4680

- Tsujimoto R, Kamiya N, Fujita Y (2014) Transcriptional regulators ChlR and CnfR are essential for diazotrophic growth in nonheterocystous cyanobacteria. *111:6762-6767*
- Urbach E, Robertson DL, Chisholm SW (1992) Multiple evolutionary origins of prochlorophytes within the cyanobacterial radiation. *Nature 355:267-270*
- Verbruggen H, Maggs CA, Saunders GW, Le Gall L, Yoon HS, De Clerck O (2010) Data mining approach identifies research priorities and data requirements for resolving the red algal tree of life. *BMC Evol Biol 10:16*
- von Wettstein D (1967) Chloroplast structure and genetics. In *Harvesting the Sun: Photosynthesis in Plant Life*. San Pietro FA, Greer T, Army J (ed) New York: Academic Press pp.153-190
- Walsby AE (1994) Gas vesicles. *Microbiol Rev 58:94-144*
- Wang G, Murase J, Asakawa S, Kimura M (2009) Novel cyanophage photosynthetic gene *psbA* in the floodwater of a Japanese rice field. *FEMS Microbiol Ecol 70:79-86*
- Wang L, Mao Y, Kong F, Li G, Ma F, Zhang B, Sun P, Bi G, Zhang F, Xue H, Cao M (2013) Complete sequence and analysis of plastid genomes of two economically important red algae: *Pyropia haitanensis* and *Pyropia yezoensis*. *PLoS One 8:e65902*
- Welker M, von Döhren H (2006) Cyanobacterial peptides - nature's own combinatorial biosynthesis. *FEMS Microbiol Rev 30:530-563*
- Whelan S, Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol 18:691-699*
- Wolk CP, Ernst A, Elhai J (2004) Heterocyst metabolism and development. In Bryant D, editor. (ed), *The molecular biology of cyanobacteria*, vol 1 Springer, Dordrecht, Netherlands, pp. 769-823
- Xiao S, Knoll AH, Yuan X, Pueschel CM (2004) Phosphatized multicellular algae in the Neoproterozoic Doushantuo Formation, China, and the early evolution of florideophyte red algae. *Am J Bot 91:214-227*

- Xiao S, Zhang Y, Knoll AH (1998) Three-dimensional preservation of algae and animal embryos in a Neoproterozoic phosphorite. *Nature* 391:553-558
- Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D (2004) A molecular timeline for the origin of photosynthetic eukaryotes. *Mol Biol Evol* 21:809-818
- Yoon HS, Hackett JD, Pinto G, Bhattacharya D (2002) The single, ancient origin of chromist plastids. *Proc Natl Acad Sci U S A* 99:15507-15512
- Yoon HS, Müller KM, Sheath RG, Ott FD, Bhattacharya D (2006) Defining the major lineages of red algae (Rhodophyta). *J Phycol* 42:482-492
- Zehr JP, Bench SR, Carter BJ, Hewson I, Niazi F, Shi T, Tripp HJ, Affourtit JP (2008) Globally distributed uncultivated oceanic N₂-fixing cyanobacteria lack oxygenic photosystem II. *Science* 322:1110-1112
- Zimmerly S, Hausner G, Wu Xc (2001) Phylogenetic relationships among group II intron ORFs. *Nucleic Acids Res* 29:1238-1250
- Zuccarello GC, West JA, Kikuchi N (2008) Phylogenetic relationships within the Stylonematales (Stylonematophyceae, Rhodophyta): biogeographic patterns do not apply to *Stylonema alsidii*. *J Phycol* 44:384-393

和文

- 井上 勳 (2007) 藻類30億年の自然史 藻類から見る生物進化・地球・環境 第2版 東海大学出版会 pp. 354-355
- Yang Z (2006) *Computational Molecular Evolution*. Oxford University Press (邦訳)
- 藤 博幸, 加藤 和貴, 大安 裕美 (2009) 分子系統学への統計的アプローチ—計算分子進化学 共立出版

謝辞

本博士論文作成にあたって、日頃から多くの御指導、御鞭撻を賜りました東京大学大学院総合文化研究科教授佐藤直樹先生に厚くお礼を申し上げます。*Triparma laevis* を分けてくださり、*T. laevis* の研究に関して助言を頂いた水産総合研究センター東北区水産研究所生態系動態グループグループ長桑田晃先生にお礼を申し上げます。*Limnothrix* sp. ABRG5-3 を分けてくださった茨城大学農学部資源生物科学科教授朝山宗彦先生にお礼を申し上げます。次世代シーケンサーによるゲノムの塩基配列解読をしてくださった東北大学大学院生命科学科准教授佐藤修正先生、東京工業大学大学院生命理工学研究科教授太田啓之先生、かずさ DNA 研究所所長田畑哲之先生にお礼を申し上げます。また、解読した塩基配列をアセンブルしてくださった京都大学大学院医学研究科准教授丸山史人先生、東京工業大学地球生命研究所教授黒川顕先生にお礼を申し上げます。

本研究遂行にあたって、文部科学省グローバル COE プログラム「地球から地球たちへ」および科学技術振興機構 CREST から研究費の援助を頂きました。また、文部科学省グローバル COE プログラム「地球から地球たちへ」および文部科学省の「卓越した大学院拠点形成支援補助金」からリサーチアシスタントとして資金援助をして頂きました。

最後に、多くの助言をくださり、温かく見守ってくださった研究室の皆様をはじめとした多くの方々に、厚くお礼を申し上げます。

付録 1：解析に用いた生物(第 1 章、第 2 章)

種	第 1 章		第 2 章		タンパク質遺伝子の相同クラスタリング解析
	アクセッション番号	図 1-2	RNA 遺伝子の推定の比較	rRNA オペロン構造	
シアノバクテリア					
<i>Acarlyochloris marina</i> MBIC11017	NC_009925-9934				X
<i>Anabaena</i> sp. PCC 7120	AP003602-6, BA000019-20				X
<i>Anabaena variabilis</i> ATCC 29413	NC_007410-3, NC_014000				X
<i>Arthrospira maxima</i> CS-328	NZ_LBYK01000001-01000129				X
<i>Arthrospira platensis</i> NIES-39	AP011615				X
<i>Crocospaera waisonii</i> WH 8501	AAAY02000001-02000323				X
<i>Cyanobacterium</i> UCYN-A	NC_013771				X
<i>Cyanothece</i> sp. ATCC 51142	CP000806-811				X
<i>Cyanothece</i> sp. PCC 7424	CP001291-7				X
<i>Cyanothece</i> sp. PCC 7425	CP001344-7				X
<i>Cyanothece</i> sp. PCC 8801	CP001287-90				X
<i>Gloeobacter violaceus</i> PCC 7421	BA000045				X
<i>Microcystis aeruginosa</i> NIES-843	NC_010296				X
<i>Nostoc punctiforme</i> PCC 73102	CP001037-42				X
<i>Prochlorococcus marinus</i> str. AS9601	NC_008816				X
<i>Prochlorococcus marinus</i> str. MIT 9211	NC_009976				X
<i>Prochlorococcus marinus</i> str. MIT 9215	NC_009840				X
<i>Prochlorococcus marinus</i> str. MIT 9301	NC_009091				X
<i>Prochlorococcus marinus</i> str. MIT 9303	NC_008820				X
<i>Prochlorococcus marinus</i> str. MIT 9312	NC_007577				X
<i>Prochlorococcus marinus</i> str. MIT 9313	NC_005071				X
<i>Prochlorococcus marinus</i> str. MIT 9515	NC_008817				X
<i>Prochlorococcus marinus</i> str. NATL1A	NC_008819				X
<i>Prochlorococcus marinus</i> str. NATL2A	NC_007335				X
<i>Prochlorococcus marinus</i> subsp. <i>marinus</i> str. CCMP1375	NC_005042				X
<i>Prochlorococcus marinus</i> subsp. <i>pastoris</i> str. CCMP1986	NC_005072				X
<i>Synechococcus elongatus</i> PCC 6301	NC_006576				X
<i>Synechococcus elongatus</i> PCC 7942	NC_007595, NC_007604				X
<i>Synechococcus</i> sp. CC9311	NC_008319				X
<i>Synechococcus</i> sp. CC9605	NC_007516				X
<i>Synechococcus</i> sp. CC9902	NC_007513				X
<i>Synechococcus</i> sp. JA-2-3B ^{at} (2-13)	NC_007776				X
<i>Synechococcus</i> sp. JA-3-3Ab	NC_007775				X
<i>Synechococcus</i> sp. PCC 7002	NC_010474-80				X
<i>Synechococcus</i> sp. RCC307	NC_009482				X

<i>Synechococcus</i> sp. WH 5701	NZ_AAN001000001-01000135	X
<i>Synechococcus</i> sp. WH 8102	NC_005070	X
<i>Synechococcus</i> sp. WH 7803	NC_009481	X
<i>Synechocystis</i> sp. PCC 6803	NC_005229-32, NC_017277	X
<i>Thermosynechococcus elongatus</i> BP-1	BA000039	X
<i>Trichodesmium erythraeum</i> IMS101	NC_008312	X
綠色植物亜界		
<i>Acomis calamus</i> plastid	AJ879453	X
<i>Adiantum capillus-veneris</i> plastid	AY178864	X
<i>Aethionema cordifolium</i> plastid	AP009366	X
<i>Aethionema grandiflorum</i> plastid	AP009367	X
<i>Agrostis stolonifera</i> cultivar Penn A-4 plastid	EF115543	X
<i>Amborella trichopoda</i> plastid	AJ506156	X
<i>Angiopteris evecta</i> plastid	NC_008829, DQ821119	X
<i>Anthoceros formosae</i> plastid	AB086179	X
<i>Arabidopsis thaliana</i> plastid	AP000423	X
<i>Arabidopsis thaliana</i> plastid	AP009369	X
<i>Arabidopsis thaliana</i> plastid	NC_004561, AJ316582	X
<i>Atropa belladonna</i> plastid	AP009370	X
<i>Barbarea verna</i> plastid	EF380351	X
<i>Buxus microphylla</i> plastid	AJ428413	X
<i>Cayceanthus fertilis</i> var. <i>ferax</i> plastid	AP009371	X
<i>Capsella bursa-pastoris</i> plastid	NC_004115, AF494278	X
<i>Chaetosphaeridium globosum</i> plastid	DQ229107	X
<i>Chara vulgaris</i> plastid	NC_005353, BK000554	X
<i>Chlamydomonas reinhardtii</i> plastid	EF380352	X
<i>Chloranthus spicatus</i> plastid	AB001684	X
<i>Chlorella vulgaris</i> C-27 plastid	NC_008822, DQ422812	X
<i>Chlorokybus amorphyticus</i> plastid	DQ864733	X
<i>Citrus sinensis</i> plastid	EF044213	X
<i>Coffea arabica</i> plastid	AP009372	X
<i>Crucihimalaya wallichii</i> plastid	AJ970307	X
<i>Cucumis sativus</i> plastid	AM711639	X
<i>Cuscuta gronovii</i> plastid	AM711640	X
<i>Cuscuta reflexa</i> plastid	AP009339	X
<i>Cycas taitungensis</i> plastid	DQ898156	X
<i>Daucus carota</i> plastid	EF380353	X
<i>Dioscorea elephantipes</i> plastid	AP009373	X
<i>Draba nemorosa</i> plastid	DQ887676	X
<i>Drimys granadensis</i> plastid	M81884	X
<i>Epifagus virginiana</i> plastid	AY780259	X
<i>Eucalyptus globulus</i> subsp. <i>globulus</i> plastid	DQ317523	X
<i>Glycine max</i> cultivar PI 437654 plastid		X

<i>Gossypium barbadense</i> plastid	AP009123	X	
<i>Gossypium hirsutum</i> cultivar coker 310 FR plastid	DQ345959	X	
<i>Helianthus annuus</i> cultivar line HA383 plastid	DQ383815	X	
<i>Hordeum vulgare</i> subsp. <i>vulgare</i> cultivar Morex plastid	EF115541	X	
<i>Hyperzia lucidula</i> plastid	AY660566	X	X
<i>Illicium oligandrum</i> plastid	EF380354	X	
<i>Ipomoea purpurea</i> plastid	EU118126	X	
<i>Jasminum nudiflorum</i> plastid	DQ673255	X	
<i>Lactuca sativa</i> cultivar Salinas plastid	DQ383816	X	
<i>Lepidium virginicum</i> plastid	AP009374	X	
<i>Leptosira terrestris</i> plastid	NC_009681, EF506945	X	
<i>Liriodendron tulipifera</i> plastid	DQ899947	X	
<i>Lobularia maritima</i> plastid	AP009375	X	
<i>Lotus japonicus</i> plastid	NC_002694, AP002983	X	
<i>Lycopersicon esculentum</i> cultivar LA3023 plastid	DQ347959	X	
<i>Marchantia polymorpha</i> plastid	X04465	X	
<i>Mesostigma viride</i> plastid	AF166114	X	X
<i>Morus indica</i> plastid	DQ226511	X	
<i>Nandina domestica</i> plastid	DO923117	X	
<i>Nasturtium officinale</i> plastid	AP009376	X	
<i>Nephtrolepis olivacea</i> plastid	NC_000927, AF137379	X	
<i>Nicotiana sylvestris</i> plastid	AB237912	X	
<i>Nicotiana tabacum</i> plastid	NC_001879, Z00044	X	
<i>Nicotiana tomentosiformis</i> plastid	AB240139	X	
<i>Nuphar advena</i> plastid	DQ354691	X	
<i>Nymphaea alba</i> plastid	AJ627251	X	
<i>Oenothera elata</i> subsp. <i>hookeri</i> plastid	AJ271079	X	
<i>Olimarabidopsis pumila</i> plastid	AP009368	X	
<i>Oltmannsiellopsis viridis</i> plastid	NC_008099, DQ201132	X	
<i>Oryza nivara</i> plastid	AP006728	X	
<i>Oryza sativa Japonica Group</i> plastid	NC_001320, XI5901	X	
<i>Ostreococcus tauri</i> plastid	CR954199	X	
<i>Panax ginseng</i> plastid	AY582139	X	
<i>Pelargonium x hortorum</i> plastid	DQ897681	X	
<i>Phaseolus vulgaris</i> cultivar Negro Jamaica plastid	DQ886273	X	
<i>Physcomitrella patens</i> subsp. <i>patens</i> plastid	AP005672	X	
<i>Pinus thunbergii</i> plastid	D17510	X	X
<i>Pinus koraiensis</i> plastid	NC_004677, AY228468	X	
<i>Piper cenocladum</i> plastid	DQ887677	X	
<i>Platanus occidentalis</i> plastid	DQ923116	X	
<i>Populus alba</i> plastid	AP008956	X	
<i>Populus trichocarpa</i> plastid	EF489041	X	

<i>Phaeocystis antarctica</i> plastid	NC_016703	X	
<i>Phaeodactylum tricornutum</i> plastid	EF067920	X	X
<i>Rhodomonas salina</i> strain CCMP1319 plastid	NC_009573, EF508371	X	X
<i>Synedra acus</i> plastid	NC_016731	X	X
<i>Thalassiosira oceanica</i> CCMP1005 plastid	NC_014808	X	X
<i>Thalassiosira pseudonana</i> plastid	EF067921	X	X
<i>Toxoplasma gondii</i> plastid	U87145	X	
<i>Triparma laevis</i> plastid	AP014625	X	
<i>Vaucheria litorea</i> plastid	NC_011600	X	X
クロララクニオン植物門			
<i>Bigelowiella natans</i> plastid	NC_008408	X	
ユーグレラゾア			
<i>Astasia longa</i> plastid	AJ294725	X	
<i>Euglena gracilis</i> plastid	X70810	X	
ボウリネラ			
<i>Paulinella chromatophora</i> plastid	NC_011087	X	

付録 1 : 解析に用いた生物(続き)(第 3 章)

種	第 3 章			
	アクセッション番号	葉緑体ゲノムの RNA 遺伝子の推定	葉緑体ゲノムの比較	ミトコンドリアゲノムの比較
緑色植物亜界				
<i>Arabidopsis thaliana</i> mitochondrion	Y08501			X
<i>Chaetosphaeridium globosum</i> mitochondrion	AF494279, NC_004118			X
<i>Chara vulgaris</i> mitochondrion	NC_005255			X
<i>Chlamydomonas eugametos</i> mitochondrion	NC_001872			X
<i>Chlamydomonas reinhardtii</i> mitochondrion	U03843			X
<i>Marchantia polymorpha</i> mitochondrion	M68929			X
<i>Mesostigma viride</i> mitochondrion	AF353999, NC_008240			X
<i>Nephroselmis olivacea</i> mitochondrion	AF110138, NC_008239			X
<i>Nicotiana tabacum</i> mitochondrion	BA000042			X
<i>Oryza sativa Japonica Group</i> mitochondrion	BA000029			X
<i>Pedinomonas minor</i> mitochondrion	AF116775			X
<i>Physcomitrella patens</i> subsp. <i>patens</i> mitochondrion	AB251495, NC_007945			X
<i>Prototheca wickerhamii</i> mitochondrion	NC_001613			X
<i>Pseudoclonium akinetum</i> mitochondrion	NC_005926			X
<i>Scenedesmus obliquus</i> mitochondrion	AF204057			X
<i>Zea mays</i> mitochondrion	AY506529			X
紅色植物門				
<i>Chondrus crispus</i> mitochondrion	Z47547, NC_001677			X
<i>Cyanidioschyzon merolae</i> mitochondrion	D89861, NC_000887			X
<i>Porphyra purpurea</i> mitochondrion	AF114794, NC_002007			X
クロムアルベオラータ				
<i>Asterionella formosa</i> plastid	KC509519		X	
<i>Asterionellopsis glacialis</i> plastid	KC509520		X	
<i>Chattonella marina</i> mitochondrion	AB546636			X
<i>Chrysochloridium synuroideus</i> mitochondrion	AF222718			X
<i>Coscinodiscus radiatus</i> plastid	KC509521		X	
<i>Cylindrotheca closterium</i> plastid	KC509522		X	
<i>Desmaretia viridis</i> mitochondrion	AY500367			X
<i>Dicyota dichotoma</i> mitochondrion	AY500368			X
<i>Didymosphenia geminata</i> plastid	KC509523		X	
<i>Ectocarpus siliculosus</i> plastid	FP102343		X	
<i>Emiliania huxleyi</i> strain CCMP 373 mitochondrion	AY342361, NC_005332			X
<i>Eunotia naegelii</i> plastid	KF733443		X	

付録 1：解析に用いた生物(続き)(第 4 章)

種	アクセッション番号	第 4 章				リボソームタンパク質遺伝子クランプスターの比較
		CyanoClust による ホモログ解析	シアノバクテリアゲノムの比較、系統解析	窒素固定関連遺伝子群の解析		
バクテリア						
<i>Agrobacterium fabrum</i> str. C58	NC_003062-5	X				
<i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168	NC_000964	X				X
<i>Burkholderia mallei</i> ATCC 23344	NC_006348-9	X				
<i>Chlorobium tepidum</i> TLS	NC_002932	X				
<i>Chloroflexus aurantiacus</i> J-10-fl	NC_010175	X				
<i>Deinococcus radiodurans</i> R1	NC_000958-9, NC_001263-4	X				X
<i>Escherichia coli</i> str. K-12 substr. MG1655	NC_000913	X				
<i>Helicobacterium modesticaldum</i> Ice1	NC_010337	X				
<i>Mycobacterium tuberculosis</i> H37Rv	NC_000962	X				
<i>Pseudomonas syringae</i> pv. <i>Tomato</i> str. DC3000	NC_004578, NC_004632-3	X			X	
<i>Ralstonia solanacearum</i> GMI1000	NC_003295-6	X			X	
<i>Rhodospseudomonas palustris</i> CGA009	NC_005296-7	X			X	
<i>Rhodospirillum rubrum</i> ATCC 11170	NC_007641, NC_007643	X				
<i>Staphylococcus aureus</i> subsp. <i>aureus</i> MW2	NC_003923	X				
<i>Streptomyces coelicolor</i> A3(2)	NC_003888, NC_003903-4	X				
<i>Thermus thermophilus</i> HB27	NC_005835, NC_005838	X				
シアノバクテリア						
<i>Acryochloris marina</i> MBIC11017	NC_009925-9934	X	X			X
<i>Anabaena cylindrica</i> PCC 7122	NC_019771-5, NC_020056, NC_020157	X	X			X
<i>Anabaena</i> sp. 90	NC_019427-9, NC_019439-40	X	X	X	X	X
<i>Anabaena</i> sp. PCC 7120	NC_003240-1, NC_003267, NC_003270,	X	X			X
<i>Anabaena variabilis</i> ATCC 29413	NC_003272-3, NC_003276	X	X			X
<i>Calothrix</i> sp. PCC 6303	NC_007410-3, NC_014000	X	X			X
<i>Calothrix</i> sp. PCC 7507	NC_019727-8, NC_019751-2	X	X			X
<i>Chamaesiphon minutus</i> PCC 6605	NC_019682	X	X			X
<i>Chroococcidiopsis thermalis</i> PCC 7203	NC_019697-8, NC_020053	X	X			X
<i>Crinalium epipsammum</i> PCC 9333	NC_019695-6, NC_019699	X	X		X	X
<i>Cyanobacterium aponinum</i> PCC 10605	NC_019733-7, NC_019753-6	X	X			X
<i>Cyanobacterium gracile</i> PCC 6307	NC_019776-7	X	X			X
<i>Cyanobacterium stanieri</i> PCC 7202	NC_019675	X	X			X
<i>Cyanobacterium</i> UCYN-A	NC_019778	X	X			X
<i>Cyanothece</i> sp. ATCC 51142	NC_013771	X	X			X
<i>Cyanothece</i> sp. PCC 7424	NC_010539, NC_010541-3, NC_010546-7	X	X		X	X
	NC_011729-30, NC_011732-4, NC_011737-8	X	X		X	X

<i>Cyanothece</i> sp. PCC 7425	NC_011880, NC_011882, NC_011884-5	X	X	X	X
<i>Cyanothece</i> sp. PCC 7822	NC_014501-4, NC_014533-5	X	X	X	X
<i>Cyanothece</i> sp. PCC 8801	NC_011721, NC_011723, NC_011726-7	X	X	X	X
<i>Cyanothece</i> sp. PCC 8802	NC_013160-1, NC_013163, NC_013167-8	X	X	X	X
<i>Cylindrospermum stagnale</i> PCC 7417	NC_019744, NC_019757-8, NC_020050	X	X	X	X
<i>Dacylococcopsis salina</i> PCC 8305	NC_019780	X	X	X	X
<i>Geitlerinema</i> sp. PCC 7407	NC_019703	X	X	X	X
<i>Gloeobacter violaceus</i> PCC 7421	NC_005125	X	X	X	X
<i>Gloeocapsa</i> sp. PCC 7428	NC_019745-7, NC_019759, NC_020051	X	X	X	X
<i>Halothece</i> sp. PCC 7418	NC_019779	X	X	X	X
<i>Leptolyngbya boryana</i> dg5	AB808482	X	X	X	X
<i>Leptolyngbya</i> sp. PCC 7376	NC_019683	X	X	X	X
<i>Limnothrix</i> sp. ABRG5-3		X	X	X	X
<i>Microcoleus</i> sp. PCC 7113	NC_019738-43, NC_019760-2	X	X	X	X
<i>Nostoc azollae</i> ' 0708	NC_014248-50	X	X	X	X
<i>Nostoc punctiforme</i> PCC 73102	NC_010628-33	X	X	X	X
<i>Nostoc</i> sp. PCC 7107	NC_019676	X	X	X	X
<i>Nostoc</i> sp. PCC 7524	NC_019677, NC_019684-5	X	X	X	X
<i>Oscillatoria acuminata</i> PCC 6304	NC_019693-4, NC_019700	X	X	X	X
<i>Oscillatoria nigro-viridis</i> PCC 7112	NC_019729-32, NC_019763-4	X	X	X	X
<i>Pleurocapsa</i> sp. PCC 7327	NC_019689	X	X	X	X
<i>Prochlorococcus marinus</i> str. AS9601	NC_008816	X	X	X	X
<i>Prochlorococcus marinus</i> str. MIT 9211	NC_009976	X	X	X	X
<i>Prochlorococcus marinus</i> str. MIT 9215	NC_009840	X	X	X	X
<i>Prochlorococcus marinus</i> str. MIT 9301	NC_009091	X	X	X	X
<i>Prochlorococcus marinus</i> str. MIT 9303	NC_008820	X	X	X	X
<i>Prochlorococcus marinus</i> str. MIT 9312	NC_007577	X	X	X	X
<i>Prochlorococcus marinus</i> str. MIT 9313	NC_005071	X	X	X	X
<i>Prochlorococcus marinus</i> str. MIT 9515	NC_008817	X	X	X	X
<i>Prochlorococcus marinus</i> str. NATL1A	NC_007335	X	X	X	X
<i>Prochlorococcus marinus</i> str. NATL2A	NC_005042	X	X	X	X
<i>Prochlorococcus marinus</i> subsp. <i>marinus</i> str. CCMP1375	NC_005072	X	X	X	X
<i>Prochlorococcus marinus</i> subsp. <i>pastoris</i> str. CCMP1986	NZ_KB235914	X	X	X	X
<i>Pseudanabaena</i> sp. PCC 6802	NC_019690, NC_019701	X	X	X	X
<i>Pseudanabaena</i> sp. PCC 7367	NC_019678-9, NC_019686	X	X	X	X
<i>Rivularia</i> sp. PCC 7116	NC_019748-50, NC_019765-6, NC_020052	X	X	X	X
<i>Stantieria cyanosphaera</i> PCC 7437	NC_006576	X	X	X	X
<i>Synechococcus elongatus</i> PCC 6301	NC_007595, NC_007604	X	X	X	X
<i>Synechococcus elongatus</i> PCC 7942	NC_008319	X	X	X	X
<i>Synechococcus</i> sp. CC9311	NC_007516	X	X	X	X
<i>Synechococcus</i> sp. CC9605	NC_007513	X	X	X	X
<i>Synechococcus</i> sp. CC9902		X	X	X	X

<i>Synechococcus</i> sp. JA-2-3B'a(2-13)	NC_007776	X	X	X	X
<i>Synechococcus</i> sp. JA-3-3Ab	NC_007775	X	X	X	X
<i>Synechococcus</i> sp. PCC 6312	NC_019680-1	X	X	X	X
<i>Synechococcus</i> sp. PCC 7002	NC_010474-80	X	X	X	X
<i>Synechococcus</i> sp. PCC 7502	NC_019691-2, NC_019702	X	X	X	X
<i>Synechococcus</i> sp. RCC307	NC_009482	X	X	X	X
<i>Synechococcus</i> sp. WH 8102	NC_005070	X	X	X	X
<i>Synechococcus</i> sp. WH 7803	NC_009481	X	X	X	X
<i>Synechocystis</i> sp. PCC 6803	NC_005229-32, NC_017277	X	X	X	X
<i>Thermosynechococcus elongatus</i> BP-1	NC_004113	X	X	X	X
<i>Trichodesmium erythraeum</i> IMS101	NC_008312	X	X	X	X

付録 1 : 解析に用いた生物(続き)(第 5 章、第 6 章)

種	第 5 章		第 6 章	
	アクセッション番号	葉緑体ゲノムの系統解析	ミトコンドリアゲノムの系統解析	リボソームタンパク質遺伝子のクワスターの比較
バクテリア				
<i>Alpha proteobacterium</i> HIMB59	NC_018644		X	
<i>Candidatus Pelagibacter ubique</i> HTCC1062	NC_007205		X	
<i>Rickettsia peacockii</i>	NC_012730		X	
<i>Rickettsia prowazekii</i>	NC_000963		X	
<i>Wolbachia</i> sp. wRi	NC_012416		X	
シアノバクテリア				
<i>Acyrochloris marina</i> MBIC11017	NC_009925-9934	X		
<i>Anabaena</i> sp. PCC 7120	AP003602-6, BA000019-20	X		
<i>Arthrospira platensis</i> NIES-39	AP011615	X		
<i>Cyanobacterium gracile</i> PCC 6307	NC_019675		X	
<i>Cyanothece</i> sp. PCC 7424	CP001291-7	X		
<i>Cyanothece</i> sp. PCC 7425	CP001344-7	X		
<i>Cyanothece</i> sp. PCC 8801	CP001287-90	X		
<i>Gloeobacter violaceus</i> PCC 7421	BA000045	X		
<i>Limnothrix</i> sp. ABRG5-3		X		
<i>Microcystis aeruginosa</i> NIES-843	NC_010296	X		
<i>Nostoc punctiforme</i> PCC 73102	CP001037-42	X		
<i>Prochlorococcus marinus</i> str. AS9601	NC_008816		X	
<i>Prochlorococcus marinus</i> str. MIT 9211	NC_009976		X	
<i>Prochlorococcus marinus</i> str. MIT 9215	NC_009840		X	
<i>Prochlorococcus marinus</i> str. MIT 9301	NC_009091		X	
<i>Prochlorococcus marinus</i> str. MIT 9303	NC_008820		X	
<i>Prochlorococcus marinus</i> str. MIT 9312	NC_007577		X	
<i>Prochlorococcus marinus</i> str. MIT 9313	NC_005071		X	
<i>Prochlorococcus marinus</i> str. MIT 9515	NC_008817		X	
<i>Prochlorococcus marinus</i> str. NATL1A	NC_007335		X	
<i>Prochlorococcus marinus</i> str. NATL2A	NC_005042		X	
<i>Prochlorococcus marinus</i> subsp. <i>marinus</i> str. CCMMP1375	NC_005072		X	
<i>Prochlorococcus marinus</i> subsp. <i>pastoris</i> str. CCMPI986	NC_019690, NC_019701	X		
<i>Pseudanabaena</i> sp. PCC 7367	NC_007595, NC_007604	X		
<i>Synechococcus elongatus</i> PCC 7942	NC_008319		X	
<i>Synechococcus</i> sp. CC9311	NC_007516		X	
<i>Synechococcus</i> sp. CC9605			X	
<i>Synechococcus</i> sp. CC9902	NC_007513	X		

付録 2 : リボソームタンパク質および遺伝子の名称と本博士論文で用いた略称

タンパク質名	図 2-10、図 4-9 および 図 6-1, 2 で用いた略称	原核生物の 遺伝子名	真核生物の 遺伝子名
Rpl2	L2	<i>rplB</i>	<i>rpl2</i>
Rpl3	L3	<i>rplC</i>	<i>rpl3</i>
Rpl4	L4	<i>rplD</i>	<i>rpl4</i>
Rpl5	L5	<i>rplE</i>	<i>rpl5</i>
Rpl6	L6	<i>rplF</i>	<i>rpl6</i>
Rpl13	L13	<i>rplM</i>	<i>rpl13</i>
Rpl14	L14	<i>rplN</i>	<i>rpl14</i>
Rpl15	L15	<i>rplO</i>	<i>rpl15</i>
Rpl16	L16	<i>rplP</i>	<i>rpl16</i>
Rpl17	L17	<i>rplQ</i>	<i>rpl17</i>
Rpl18	L18	<i>rplR</i>	<i>rpl18</i>
Rpl22	L22	<i>rplV</i>	<i>rpl22</i>
Rpl23	L23	<i>rplW</i>	<i>rpl23</i>
Rpl24	L24	<i>rplX</i>	<i>rpl24</i>
Rpl29	L29	<i>rpmC</i>	<i>rpl29</i>
Rpl30	L30	<i>rpmD</i>	<i>rpl30</i>
Rpl31	L31	<i>rpmE</i>	<i>rpl31</i>
Rpl36	L36	<i>rpmJ</i>	<i>rpl36</i>
Rps3	S3	<i>rpsC</i>	<i>rps3</i>
Rps4	S4	<i>rpsD</i>	<i>rps4</i>
Rps5	S5	<i>rpsE</i>	<i>rps5</i>
Rps7	S7	<i>rpsG</i>	<i>rps7</i>
Rps8	S8	<i>rpsH</i>	<i>rps8</i>
Rps9	S9	<i>rpsI</i>	<i>rps9</i>
Rps10	S10	<i>rpsJ</i>	<i>rps10</i>
Rps11	S11	<i>rpsK</i>	<i>rps11</i>
Rps12	S12	<i>rpsL</i>	<i>rps12</i>
Rps13	S13	<i>rpsM</i>	<i>rps13</i>
Rps14	S14	<i>rpsN</i>	<i>rps14</i>
Rps17	S17	<i>rpsQ</i>	<i>rps17</i>
Rps19	S19	<i>rpsS</i>	<i>rps19</i>