

機関番号：17102

研究種目：挑戦的萌芽研究

研究期間：2013～2013

課題番号：25660124

研究課題名(和文) Long-PCRによる針葉樹葉緑体ゲノム構造の解明とゲノム完全解読システムの構築

研究課題名(英文) A simple sequencing system of complete chloroplast genome using long PCR and NGS in coniferous species

研究代表者

白石 進 (Shiraishi, Susumu)

九州大学・(連合)農学研究科(研究院)・教授

研究者番号：70226314

交付決定額(研究期間全体)：(直接経費) 3,000,000円、(間接経費) 900,000円

研究成果の概要(和文)：針葉樹(Coniferophyta)の葉緑体ゲノムの完全解読のための簡便迅速な分析系の構築を次の2つの方法で行った。1つは、針葉樹葉緑体ゲノムの中で塩基配列が高度に保持されている領域を探索し、それらの配列構造の規則性からLong PCRと次世代シーケンサ(NGS)で解読する方法である。他の方法は、簡単な細胞分画を行い、主に葉緑体DNAが占める画分のDNAを次世代シーケンサで判読し、バイオインフォマティクスを駆使してゲノムを解読する方法である。これら2つの方法により、針葉樹の多くの樹種で、葉緑体ゲノムの完全解読が可能となった。

研究成果の概要(英文)：Two simple complete-sequencing systems of chloroplast genome were constructed for Coniferophyta species. One was the sequencing system using a next generation sequencer (NGS), and long PCR s designed with the arrangement information of the highly conserved regions among genera. The other was the system using the next generation sequencing of the genome DNA extracted from crude chloroplast fraction of the leaf and bioinformatics. It could be possible to decode the chloroplast genome completely in a lot of coniferous species by using either of these two sequencing systems.

研究分野：農学

科研費の分科・細目：森林圏科学・森林科学

キーワード：葉緑体ゲノム ゲノム解析 針葉樹 マツ科 ヒノキ科 分子進化

1. 研究開始当初の背景

1986年に、タバコとゼニゴケで初めて葉緑体ゲノムの全塩基配列が決定され、現在までに多くの植物でゲノムの完全解読が行われている。針葉樹(球果植物門, Coniferophyta)においても、葉緑体ゲノムの完全解読が進められてきたが、現在までに報告されたものは、マツ科のマツ(*Pinus*)属、トウヒ(*Picea*)属、カラマツ(*Larix*)属、ヒマラヤスギ(*Cedrus*)属、トガサワラ(*Pseudotsuga*)属、*Cathaya*属、ユサン(*Keteleeria*)属と、ヒノキ科のスギ(*Cryptomeria*)属、タイワンスギ(*Taiwania*)属、コウヨウザン(*Cunninghamia*)属、シヨウナンボク(*Calocedrus*)属、ナンヨウスギ科のナンヨウスギ(*Araucaria*)属、イヌマキ科のナギ(*Nageia*)属とマキ(*Podocarpus*)属、イヌガヤ科のイヌガヤ(*Cephalotaxus*)属のわずか15属である。

葉緑体DNAは針葉樹では父親のみから子に遺伝するため、父系解析や父親鑑定に関する重要な遺伝情報を提供してくれるが、多数の葉緑体DNAマーカーの開発とその高度利用はほとんど行われていない。近年、スギで葉緑体ゲノムが完全解読され、この配列情報を活用し、葉緑体ゲノムを包括的にタイピング(global chloroplast genome typing, gCGT)することで、詳細な遺伝現象が解明できることが示された。今後、さらに多くの針葉樹でgCGTによる遺伝育種研究を進めるためには、葉緑体ゲノムの簡便な完全解読法の開発が不可欠である。また、葉緑体ゲノムには植物に特異的な機能である光合成に関連する多数の遺伝子が存在する。このことから、遺伝学、育種学および生理学において葉緑体ゲノム情報が必要不可欠となっている。しかし、針葉樹において完全解読された葉緑体ゲノムは極めて少ない。

2. 研究の目的

針葉樹の葉緑体ゲノム構造は、樹種間で大きく違っており、これが針葉樹での完全解読の障害となっている。今日、次世代シーケンサが登場し、膨大な塩基配列の解読が極めて容易となった。そこで、本研究では、次の2つの方法を用いて針葉樹葉緑体ゲノムを簡便・迅速に完全解読することに挑戦した。

葉緑体ゲノム全体をlong-PCRのみで増幅し、次世代シーケンサで塩基配列を判読、ゲノムの全塩基配列を決定する方法(以下、long-PCR法)

針葉よりクルードな葉緑体DNA分画を単離、これをシーケンサで判読した後、混入している核ゲノムとミトコンドリアゲノムの配列をバイオインフォマティクスを駆使して取り除き、葉緑体ゲノムを解読する方法(以下、葉緑体DNA分画法)

3. 研究の方法

3.1 Long-PCR法

葉緑体ゲノム構造解析のためのソフトウェアの開発: 針葉樹の中で系統的に最も遠縁にあるヨーロッパトウヒ(*Picea abies*)とタイワンスギ(*Taiwania cryptomerioides*)の2種間で塩基配列が高度に保持されている領域を探索するための解析プログラムをExcel VBAを用いて作成した。

Long PCR: 遺伝子配列が高度に保存されている3領域(ブロック)に8対、連結領域に8対の計16対のプライマーをOLIGOを用いて設計した。Long PCRによる増幅産物はアガロースゲルで電気泳動し、増幅したプライマー対からゲノム構造を判定した。

3.2 葉緑体DNA分画法

葉緑体DNAリッチ画分の単離: ヒノキ芽生え(約1g)に単離バッファー(0.33M sorbitol, 20mM Tricine, 5mM EGTA, 5mM EDTA, 10mM NaHCO₃, pH 8.4)を加え、Polytron(KINEMATICA)によって破碎し、遠心(4,800 rpm, 4, 5 min)後、上清を取り除き、沈殿をwash buffer(0.3M sorbitol, 20mM HEPES, 5mM MgCl₂, 2.5mM EDTA)で2回洗浄し、膜画分を得た。得られた膜画分と上清からDNeasy Plant Mini Kit(QIAGEN)を用いてDNAを抽出した。核DNA、葉緑体DNA、ミトコンドリアDNA上の遺伝子の塩基配列情報を基に設計したマーカーを用いて、各画分から抽出されたDNAをPCRし、その産物量によって核DNAとミトコンドリアDNAのコンタミを評価した。

次世代シーケンサによるゲノム解読: 上記で得られた膜画分DNAを、TruSeq DNA Sample Prep Kit(Illumina)で処理し、次世代シーケンサ(Illumina MiSeq)を用いて葉緑体ゲノムの塩基配列を解読した。また、針葉樹において葉緑体ゲノムが解読されているスギ(*Cryptomeria japonica*)、タイワンスギ(*Taiwania cryptomeroides*)、イヌガヤ(*Cephalotaxus wilsoniana*)、トウヒ(*Picea abies*)をリファレンスとしてシーケンサデータをゲノムマッピングした。

4. 研究成果

4.1 Long-PCR法

ヨーロッパトウヒとタイワンスギの2種間で塩基配列が高度に保持されている領域を探索した結果、ホモロジー(95%以上)の高い199個の領域が明らかとなった。さらに、各遺伝子で1領域を選び、43個の高ホモロジー領域を得た。これらの領域の配列順序を調べた結果、科間では大きな構造変異が存在するものの、マツ科(Pinaceae)内、ヒノキ科(Cupressaceae)内、マキ科(Podocarpaceae)内では非常に類似していることが明らかとなった。これら3科で針葉樹の大半を占めることから、針葉樹の多くの

種での効率的な葉緑体ゲノム解読の可能性が示された。

次に、上記の3科におけるゲノム解読システムについて検討した。高度に保持されたDNA領域上に、3科それぞれに特異的なプライマー群を設計し、long-PCRで増幅、PCR産物を酵素で切断し、次世代シーケンサで塩基配列を決定する新規のゲノム解読法を構築した。

マツ科は11属からなり、このうち既に7属で葉緑体ゲノムが解読されている。しかし、マツ属を除き、ほとんどの属では1樹種が解読されているに過ぎない。まず、マツ科の完全解読システムの確立を行った。マツ科の葉緑体ゲノム構造は比較的良好に保持されており、3ブロック(図1のL, M, S)間の組換えにより、ゲノム構造の多様性が作られていた。そこで、Lブロックを増幅するための6対のプライマー、MとSそれぞれのための2対のプライマー、L-M間、L-S間、M-S間を増幅する8対のプライマーを設計し、long-PCRシステムを作成した。

図1にマツ科の11属(太字:未解読の属)のブロックの配列型を、図2にカラマツとテッケンユサンでのlong-PCRの増幅結果を示した。L(レーン3~8)、M(レーン2)、S(レーン1)の3ブロックでは全ての樹種で増幅産物が得られた。また、ブロック間(レーン9~16)のPCRでは、3つのブロック間プライマー対で増幅産物が得られた。これら11個のPCR産物を次世代シーケンサで解読することにより全ゲノムの解読が可能となった。

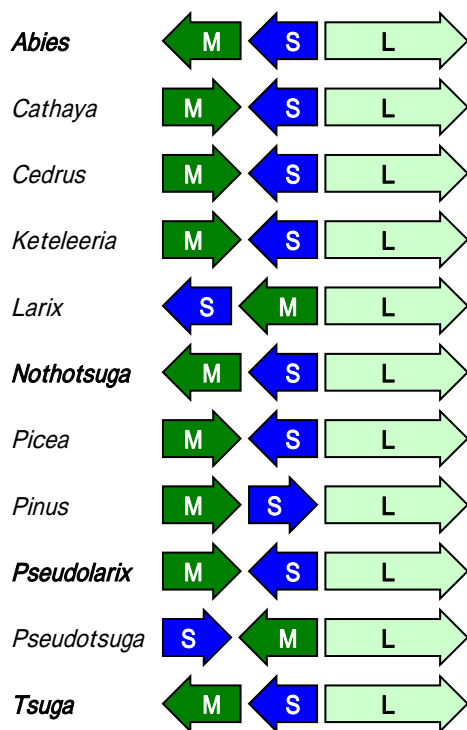


図1 マツ科葉緑体ゲノム増幅用 long-PCR の増幅結果から判定されたマツ科樹木 11 属 15 種のゲノム構成



図2 マツ科葉緑体ゲノム増幅用 long-PCR システムでの増幅結果

lane 1: S; 2: M; 3-8: L; 9: L-M1; 10: L-S1; 11: L-S2; 12: L-M2; 13: L-S3; 14: M-S1; 15: L-S4; 16: M-S2

4.2 葉緑体 DNA 分画法

これまで針葉樹の葉緑体ゲノム解読では、葉緑体の単離が大きな障害となってきた。植物の細胞には、核ゲノム、葉緑体ゲノム、ミトコンドリアゲノムの3種類のゲノムが存在する。葉緑体ゲノムを決定するためには、他のゲノムのコンタミがない、葉緑体だけを単離する必要がある。しかしながら、葉緑体の単離には膨大なサンプル量を必要とし、植物種によっては葉緑体の単離が困難であるため、葉緑体ゲノムの解読は限られた植物種にとどまっている。また、一般的なサンガー法による塩基配列の解読は、一度に 350~800 bp の解読が限界であり、おおよそ 130 kbp ある葉緑体ゲノムの完全配列を決定するためには、膨大な量のシーケンス解析が必要であった。また、葉緑体を単離する際にコンタミした微量な他ゲノムも解読されるため、コンタミ配列を排除するために、繰り返し解析を行う必要があり、葉緑体ゲノムの解読には多大な労力がかかっていた。

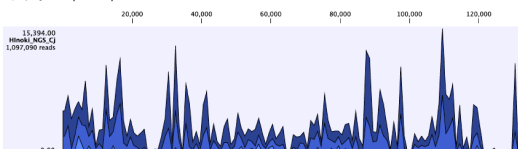
既往の知見により、シロイヌナズナ (*Arabidopsis thaliana*) の葉緑体 DNA の多くは、葉緑体の膜に付着していることが報告されており、被子植物でも同様に葉緑体 DNA が葉緑体膜に付着している。葉緑体膜と細胞膜の大部分は、破碎後、短時間の遠心分離により簡単に分画できる。短時間の遠心では、核ゲノムは膜に結合しないため上清に残る。一方、ミトコンドリアはサイズが小さいために沈殿しない。そのため、葉全体から抽出された膜画分は葉緑体 DNA リッチな画分であると考えられる。

核 DNA、葉緑体 DNA、ミトコンドリア DNA 上に設計したマーカーを用いて、それぞれの画分から抽出された DNA を PCR することでコンタミの程度を評価した結果、膜画分から抽出された DNA は、葉緑体 DNA マーカーである *rbcL* 遺伝子と *PsbA* 遺伝子の増幅が確認された。一方、ミトコ

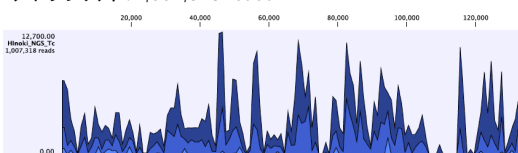
ンドリア DNA マーカーである *cox2* 遺伝子と *atp9* 遺伝子はわずかに増幅、核 DNA マーカーである *rbcS* 遺伝子と ITS 領域の増幅産物は検出できなかった。また、上清画分から抽出した DNA は、全てのマーカーで良好な増幅が確認された。これらの結果から、膜画分を得ることにより葉緑体ゲノムリッチな DNA を得ることができると明らかとなった。

膜画分の DNA を次世代シーケンサを用いて塩基配列を解読した結果、平均 190 bp の断片が約 700 万 read 得られた。得られた read から葉緑体ゲノムの特異的な配列のみを選択するために、カバー率が低い配列を取り除いた。残った配列で再度アセンブルを行った結果、293bp~20kbp の scaffold (/ contig) が得られ、合計の塩基長は 123,125 bp であった。相同性検索の結果、他の植物種の葉緑体ゲノムに局在する遺伝子は全て含んでいることが明らかとなった。また、針葉樹において葉緑体ゲノムが解読されているスギ、タイワンスギ、イヌガヤ、トウヒをリファレンスにしてゲノムマッピングを行った結果、得られた配列は他の葉緑体ゲノムにおいて高いカバー率を有することが分かった(図 3)。また、系統分類上近い種ほどカバー率が高く、ヒノキと系統分類上近いとされるスギ葉緑体ゲノムの約 97% をカバーしていた。一方、今回得られた断片の中にはスギにある *ycf1* 遺伝子(約 7 kbp) と相同性の高い断片は得られなかった。図 3 において他の植物種

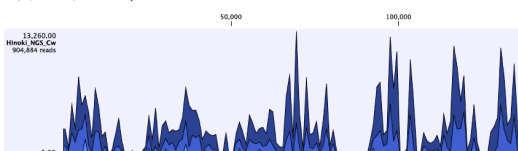
スギ: 1,097,090 reads



タイワンスギ: 1,007,318 reads



イヌガヤ: 904,884 reads



トウヒ: 581,506 reads

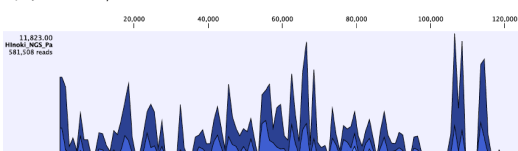


図 3 4 種のリファレンスシーケンスに対するゲノムマッピング

横軸はリファレンスゲノムの配列、縦軸は read 数を示す。

のリファレンスゲノムと一致しない領域のほとんどは *ycf1* 遺伝子によるものであり、ヒノキの葉緑体ゲノムには *ycf1* 遺伝子が存在しない可能性が示唆された。また、ミトコンドリアゲノムの配列と思われる read は非常に少なく、ゲノム解読の支障にはならなかった。

4.3 葉緑体ゲノムの簡便解読法の利用

次世代シーケンサの登場によりゲノム解読は一変し、大量の塩基配列の決定が極めて迅速・簡便に行えるようになってきている。針葉樹の葉緑体ゲノムでも、次世代シーケンサによる解読が行われているが、完全解読までには至っていない。これは、針葉樹においては、ゲノム構造変化が激しく、樹種間でゲノム構造に大きな違いがあることに起因しているとされてきた。しかし、本研究において long-PCR 法による新規解読法を構築する過程で、科レベルでは、ゲノム構造が比較的良好に保存されていることが明らかになった。

本研究で可能となった 2 つのゲノム解読法により、これまで遅れていた針葉樹(世界的に重要な森林資源樹種)の葉緑体ゲノムの完全解読が促進されるであろう。long-PCR 法は、既に数種の解読情報が蓄積されている科において、新規樹種での解読に威力を発揮することができる。一方、葉緑体 DNA 分画法は、まったく近縁種の情報が無い場合に有効である。いずれの方法も、少額の予算でゲノムプロジェクトを推進できることから、これら 2 つの分析システムを使い分けることによって、効率的な完全解読が達成できると考える。

育種学では交配家系の適切な管理が求められる。これまで、核 DNA マーカーのみが利用されてきたが、ゲノム情報を活用した多数の葉緑体 DNA マーカーの開発が容易となり、葉緑体ゲノムの包括的なタイピング (gCGT) 等の新しい研究法への発展が期待できる。特に父親を特定できる葉緑体 DNA マーカーは、育種を効率的に進める上での強力な手法として有用である。また、葉緑体ゲノムの詳細な情報を用いた集団遺伝学の新展開も期待される。

一方、生態学において分子生態学が創成され、DNA マーカーによる親子鑑定を用いた個体群の更新動態研究(花粉流動、種子散布、繁殖戦略等)が広く行われている。しかし、両親より遺伝する核 DNA マーカーによる両親鑑定(どちらが母親で、どちらが父親かは不明)にとどまっているのが現状である。片親のみ(針葉樹では父親)から子に伝わる葉緑体 DNA マーカーを利用すると、母親と父親を特定できるので、より詳細な生態系の解明が可能となる。

葉緑体ゲノムには光合成に関係する遺伝子が数多く存在し、これらの遺伝子が光合成の基本骨格となっている。近年、ゲノム解読

の技術進歩に伴い多くの植物種の葉緑体ゲノムが解読され、機能が未知の遺伝子 (open reading frame, ORF) が数多く存在することが報告されている。針葉樹の葉緑体ゲノムにも、他の植物種にはない ORF が確認されている。多くの針葉樹で完全なゲノム解読が進み、針葉樹の遺伝子群を詳細に解析できるようになることで、針葉樹に特異的な光合成メカニズムと光合成システムの分子進化の解明が期待される。

5. 主な発表論文等

(研究代表者, 研究分担者及び連携研究者には下線)

[学会発表](計 1 件)

柳原尚貴, 後藤栄治, 田村美帆, 渡辺敦史, ヒノキ葉緑体ゲノム解読に向けた効率的手法の開発, 森林遺伝育種学会, 2013.10, 東京 (東京大学)

6. 研究組織

(1) 研究代表者

白石 進 (Shiraisi, Susumu)
九州大学・大学院農学研究院・教授
研究者番号: 70226314

(2) 研究分担者

後藤栄治 (Gotoh, Eiji)
九州大学・大学院農学研究院・助教
研究者番号: 90614256