

Malawimonadida に近縁な未記載真核微生物 SRT706 株と TS011B 株のミトコンドリアゲノム解析

松浦 眞子 (筑波大学生物学類)

指導教員：稲垣 祐司 (筑波大学計算科学研究センター)

【背景・目的】

ミトコンドリアは酸素呼吸を行う細胞小器官であり、現存真核生物系統の直近の共通祖先では既に確立していたと考えられている。ミトコンドリアは、真核生物進化のごく初期に細胞内共生したアルファプロテオバクテリアを起源とする。ほとんどの真核生物種のミトコンドリアでは、その進化的起源を反映し細菌型ゲノムを保持している。アルファプロテオバクテリアが宿主真核細胞内でオルガネラ化する過程と、それに続く真核生物進化中で、自由生活性生活に必要な機能を担う遺伝子は破棄され、オルガネラの機能とその維持に必要な遺伝子の大多数は宿主核へと転移した。その結果、現存真核生物のミトコンドリアゲノム (mtDNA) にコードされる遺伝子数は、アルファプロテオバクテリアのゲノムにコードされる遺伝子数と比べ極めて少なくなっている。これまでの研究により系統的に多様な真核生物種の mtDNA が解読された結果、系統間で mtDNA コードのタンパク質遺伝子レパートリーには、ある程度のばらつきがあることが分かっている。また少数の例外を除き、外来遺伝子が mtDNA に挿入されることは知られていない。従って、ある特定の真核生物が他の大多数の真核生物の mtDNA よりも大きなタンパク質遺伝子レパートリーをもつ場合、その真核生物は他の真核生物よりも初期に分岐したため mtDNA の縮退が進んでいないと解釈することができる。このロジックに従えば、系統的に多様な真核生物の mtDNA 遺伝子レパートリーを比較することで、真核生物の初期分岐系統について示唆を得ることも可能である。

所属研究室は白鳥峻志博士 (筑波大) と共同研究を進めており、白鳥博士が単離培養した 4 種の未記載真核微生物である TS010A 株、TS028 株、TS034 株および TS011B 株について RNA-seq データと DNA-seq データを既に取得している。さらに RNA-seq データから抽出した 328 種の核ゲノムコードタンパク質配列に基づく系統解析では、4 種の大まかな系統的位置が明らかとされた。その系統解析では TS010A 株と TS028 株は Metamonada クレードの基部から分岐する新奇系統、TS034 株は Breviatea の基部から分岐する新奇系統であると推測された。また TS011B 株は Malawimonadida と未記載真核微生物 SRT706 株との近縁性が強く支持された。このような背景のもと、本研究では上記 4 種を対象に mtDNA の解読とタンパク質遺伝子レパートリーの解明を試みた。

【材料と方法】

所属研究室の先行研究により、未記載真核微生物 4 種から、illumina NovaSeqX からの short reads として約 83-97 Mbp が取得されていた。本研究では mtDNA の解読を目指し、これらの short reads をアセンブルすることによりゲノム断片を復元した。まず、Fastp プログラム (<https://github.com/OpenGene/fastp.git>) を用いてアセンブルに適さないシーケンスクオリティが低い配列を取り除き、その後 unicycler プログラム (<https://github.com/rwick/Unicycler.git>) を用いてアセンブルを行った。アセンブルにより復元されたゲノム断片には、解析目的

の mtDNA 断片の他に、核ゲノム断片や培養液中に存在している細菌のゲノム断片が含まれるため、mtDNA 断片とそれ以外のゲノム断片を区別する必要がある。本研究では、tblastn を用いた相同性検索と、復元されたゲノム断片がどのように連結されているかを可視化するプログラムである Bandage (<https://github.com/rwick/Bandage>) を併用し、mtDNA 断片候補を抽出した。tblastn では、これまでに知られている最大のタンパク質遺伝子レパートリーをもつ Jakobida の mtDNA コードタンパク質配列 (67 種類) をクエリとし、復元したゲノム断片に対して相同性検索を実行した (bit score が 100 以下のヒットは破棄)。最終的に 50 Kbp 以下かつ GC 含量が 40% 以下であるものを mtDNA 断片候補とし、MFannot ウェブサイト (<https://megasun.bch.umontreal.ca/apps/mfannot/>) によりアノテーションを行った。

【結果と考察】

TS010A 株の DNA-seq データから、3 つの mtDNA 断片候補が抽出された。それぞれの断片長は、約 2.7、7.1、14.2 Kbp であった (GC 含量は 29-34%)。Bandage が作成した graph からは、この mtDNA は 7.1 Kbp 断片を 2 コピー含むと予測された。TS208 株のアセンブルデータからは mtDNA 断片として約 23.1、14.3、16.6 Kbp の 3 断片が発見され、いずれの断片も GC 含量は 31% 程度であった。TS208 株の mtDNA も、14.3 Kbp 断片を 2 コピー含む可能性が高い。TS034 株の DNA-seq データからは、約 30 Kbp 断片が 1 つだけ発見され、その GC 含量は 23.6% であった。TS011B 株からは、約 35.6 Kbp、GC 含量 23.6% の mtDNA 断片候補が発見された。

未記載真核微生物 4 種の mtDNA 断片候補を Mfannot でアノテーションした。TS028 株の 3 つの mtDNA 断片には、これまで報告された mtDNA コード遺伝子のレパートリーのほとんどが検出されたため、これらの断片により mtDNA が構成されている可能性が高い。TS010A 株と TS034 株の mtDNA 断片候補には、これまでに解読されたすべての mtDNA に発見されている遺伝子 (例えばチトクローム c オキシターゼのサブユニットやリボソーム RNA 遺伝子) の一部が同定されなかった。従って、本研究では見過ごした mtDNA 断片が存在すると思われる。同様な理由で、TS011B 株についても本解析で見過ごされた mtDNA 断片が存在すると思われる。特に TS010A 株、*Malawimonas jakobiformis*、SRT706 株間には明らかな近縁性があるため、3 種間での mtDNA コード遺伝子のレパートリーについて厳密な比較が可能である。*M. jakobiformis* の mtDNA は解読されており (<https://megasun.bch.umontreal.ca/ogmp/projects/mjako/gen.html>)、SRT706 株の mtDNA についても所属研究室の先行研究で完全解読されている (未発表)。TS011B 株 mtDNA の遺伝子レパートリーは完全に明らかとなっていないが、本発表では、TS011B 株、*M. jakobiformis* および SRT706 株の mtDNA の遺伝子レパートリーと比較し、議論を行う。