

Thermoproteota 門古細菌における PCNA 進化

上西 慧莉紗（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】

DNA 複製はセントラルドグマの一部を構成し、その分子機構は生命にとって必要不可欠である。その重要性ゆえに、DNA 複製機構に関わるタンパク質の多くは、生物界全般で共通している。広範な生物系統間で高度に保存された DNA 複製機構に関わるタンパク質の 1 つが、環状構造を持つ DNA クランプである。環状の DNA クランプは、その穴に複製中の DNA を通すことで DNA 分子を固定する。さらに DNA と結合した DNA クランプは、複数の DNA 複製関連タンパク質と結合することで DNA 複製の効率を上昇させる。DNA クランプは生物の系統により、環状構造を構成するサブユニットの数と種類が異なることが分かっている。

細菌類では単一の PCNA タンパク質の 2 量体が DNA クランプを形成するのに対し、真核生物では単一の PCNA タンパク質が 3 量体の環状構造を形成する。多くの古細菌（例えばユーリ古細菌）は真核型 DNA クランプをもち、ホモ 3 量体が環状構造を形成する。一方、Thermoproteota 門古細菌の一部（例えば *Sulfurisphaera tokodaii*）ではゲノム中に 3 種類の PCNA 遺伝子が同定されている。さらに、これら 3 種類の PCNA 遺伝子産物がヘテロ 3 量体 DNA クランプを形成することが実験的に明らかとなっている（Lu ら BBRC 2008 376 369-374）。しかし、古細菌（特に Thermoproteota 門）でヘテロ 3 量体 DNA クランプが、いつ確立したのか、どの程度広い系統に分布しているのかなどは明らかにされていない。

本研究は Thermoproteota 門古細菌におけるヘテロ 3 量体 DNA クランプの起源と進化を考察するため、GTDB データベース（<https://gtdb.ecogenomic.org>）で公開されている古細菌ゲノムおよびメタゲノムデータ中の PCNA 遺伝子を探索し、ゲノム中に複数の PCNA 遺伝子をもつ Thermoproteota 門古細菌を探索した。また複数タンパク質配列に基づく最尤法系統解析によりゲノム間の系統関係を推測し、Thermoproteota 門古細菌においていつヘテロ 3 量体 DNA クランプが確立したのかを推測した。

【手法】

古細菌 PCNA 配列の取得と系統解析 GTDB データベースからメタゲノムを含む古細菌ゲノムデータをダウンロードし、データベースを作成した。あらかじめ入手した PCNA アミノ酸配列（113 配列）をクエリとし、作成したデータベースに対して BLASTP を用いた相同性解析を行った。クエリ配列に E value が $1e-4$ 以下でヒットした配列を PCNA 配列と見做し、後述する解析に用いた。本研究で取得した PCNA 配列とクエリ配列中の冗長性を CD-HIT プログラムを用いて冗長性を削除後、MAFFT プログラムによりアライメント座位の整列を行い、その後 trimAl プログラムによる系統解析に不適切だと考えられる座位の削除を行った。最終的に、619 配列（228 アミノ酸座位）から構成される PCNA アライメントを、IQ-TREE プログラムを用いた最尤法系統解析に供した。置換モデルは LG+F+G4 を選択した。最尤系統樹の二分岐に対する信頼性評価は UFBoot2 を用いた。

Thermoproteota 門古細菌の系統解析 Thermoproteota 門古細菌において、いつヘテロ 3 量体 DNA クランプが確立したのかを精度よく推測するには、Thermoproteota 門古細菌の系統関係を正確に把握する必要がある。そこで GTDB が多遺伝子系統解析用に選別した単一コピータンパク質データのうち 48 種類を連結することで「arc48」アライメントを作成した（7952 アミノ酸座位）。このアライメントは 361 種/ゲノムから構成され、IQ-TREE プログラムを用いた最尤法系統解析に供した（置換モデル、LG+F+I+G4；二分岐に対する信頼性評価、UFBoot2）。

【結果・考察】

これ以降 GTDB の古細菌分類体系を用いて議論する。本研究で行った PCNA 配列の探索の結果、Thermoproteota 門古細菌において Thermoprotei 綱および Thermoprotei_A 綱の一部の種で複数の PCNA 遺伝子が検出された。arc48 系統解析では上記 2 綱は姉妹群となったが、後述するように PCNA タンパク質の多様化は 2 綱間で独立に起こったと考えられる。それ以外の古細菌ゲノムからは PCNA 遺伝子は 1 種類のみ検出された。

Thermoprotei_A 綱に含まれる *S. tokodaii* など多様な古細菌ゲノムから 3 種類の PCNA 配列が検出された。3 種類の PCNA 配列はそれぞれ独立したクレードを形成し、それぞれの単系統性は UFBoot2 により強く支持された。Lu ら（BBRC 2008 376 369-374）の同定した *S. tokodaii* の PCNA 配列一すなわち SsoPCNA1、PCNA2 および PCNA3—は、それぞれ異なるクレードに包含されるため、本研究では「PCNA-1」、「PCNA-2」、「PCNA-3」クレードと呼称する。PCNA 系統解析では、上記 3 クレード間に強い近縁性は検出できなかった。従って、3 種類の PCNA タンパク質がどう形成されたかは不明であるが、この綱の共通祖先の DNA クランプは既にヘテロ 3 量体であったと推測できる。

Thermoprotei 綱古細菌では特定の PCNA タンパク質（PCNA-X と呼称）が普遍的に分布していた。arc48 系統解析で、Thermoprotei 綱クレード中で Thermocladaceae 科と Thermoproteaceae 科は姉妹群になるが、両科に含まれる一部の種において PCNA-X に加え、系統的に独立した PCNA 配列（PCNA-Y と呼称）も検出された。従って Thermocladaceae 科と Thermoproteaceae 科の共通祖先は PCNA-X に加え、PCNA-Y を追加獲得したと考えられる。またその後、一部の種では 2 次的に PCNA-X あるいは PCNA-Y を欠失した可能性がある。発表では、2 種類の PCNA タンパク質がどのような 3 量体を形成するかについても議論したい。

本研究の結果を総合すると、Thermoproteota 門古細菌の進化中で複数の PCNA タンパク質を保持する系統は独立に複数回確立したと考えられる。