

氏 名 塩 入 千 春

学位（専攻分野） 博士(理学)

学 位 記 番 号 総研大甲第709号

学位授与の日付 平成15年3月24日

学位授与の要件 先導科学研究科 生命体科学専攻

学位規則第4条第1項該当

学 位 論 文 題 目 ゲノム配列の数量的特徴づけ

論 文 審 査 委 員 主 査 教授 長谷川 政美  
教授 池村 淑道  
教授 高畑 尚之  
教授 堀内 嵩

## 論文内容の要旨

本学位論文は、ゲノム配列の大量データに対する数量的な特徴づけについて述べる。現在、利用可能なミトコンドリアゲノム 359 種、原核生物の染色体 105 配列、シロイヌナズナ、線虫、酵母など 6 種の真核生物の完全ゲノムのほか、ヒト、マウス、ショウジョウバエのドラフト配列を用いて解析を行った。解析方法は、突然変異の鎖非対称性による塩基組成の偏り (Skew) のパターンを配列全体において調べ、更に DNA の高次構造や機能と関係する二連塩基の頻度傾向を調べた。鎖非対称性は DNA の複製システムと強い相関があり、複製が一方向に進む動物のミトコンドリアでは単調なパターンを持つが、複製が二方向性である細菌では複製の開始点・終始点で Skew のパターンが反転し複製開始点の予測も容易に可能である。特に、大腸菌ではグアニン(G)とシトシン(C)との間で非常にきれいな Skew のパターンがあり、それは突然のホット・スポットである Chi 配列と呼ばれる 5'-GCTGGTGG-3'配列がもたらしているようである。真核生物においては、原核生物同様複製の方向は二方向性であるが、一本の染色体上に複製単位が複数あるため、Skew のパターンは複雑である。しかしながら、複製システムと関係しているのは明らかであり、リーディング鎖とラギング鎖が交互に入れ替わる様子が伺える。このような一塩基レベルの鎖非対称性にかかわらず、それぞれのゲノム配列全体に一貫した種固有的な二連塩基の頻度傾向がある。原核生物から真核生物まであらゆる種において TpA は少なく、ほぼ普遍的な傾向である。ヒトや哺乳類の核ゲノムでは、CpG 二連塩基が非常に低く抑えられており、期待値に比べて 20~40%ほどである。これは CpG の C がメチル化を受けやすいことに起因しており、メチル化を受けた C はデアミ化によってチミン (T) に変化しやすく、C->T 転移 (トランジション) が生じて、TpG/CpA になる。そのため、TpG/CpA の増加もみられるが、CpG は極端に抑えられているため、TpG/CpA の増加分だけで CpG の減少分を補填できない。つまり、メチル化-デアミ化による突然変異のほかに CpG を抑制する何らかの機構があると考えられ、最も有力な説は DNA の高次構造と関係するスタッキング・エネルギーが原因とする説である。スタッキング・エネルギーとは原子間あるいは分子間の相互作用による結合エネルギーであるため、それぞれの組み合わせで異なる。そして、このエネルギーによる違いが DNA の高次構造の可塑性を決めるが、TpA の組み合わせは最も可塑性が高く CpG は最も硬質であるうえポリ (CG) は左巻き二重らせんの Z 型 DNA 構造をとる。そのため、TpA や CpG が避けられているとされている。そして、その他の二重塩基も含めて 16 とおりの組み合わせ全てにおいて、ゲノムに一貫した頻度傾向があると説明できる。そして、このような二連塩基の頻度傾向は、アミノ酸をコードしているコード領域でも見られる傾向である。遺伝暗号は  $4 \times 4 \times 4 = 64$  通りのコドンからなるが、20 通りのアミノ酸に対応するためコドンには縮退がある。同じアミノ酸をコードするコドンを同義コドンの使用頻度は一律ではなくそれぞれに偏りがあり、下等な生物では tRNA の存在量とも強い相関があるが、ヒトの遺伝子では CpG や TpA を含むコドンの使用頻度が抑えられており、ゲノム全体における二連塩基の頻度傾向の影響を強く受けており、その他の真核生物においても同様に、コドンの使用頻度は二連塩基の頻度傾向の影響を強く受けており、二連塩基の組み合わせは基本的に非常に強い性質であるといえる。

生命の誕生や初期変化について考察を行うと、ゲノムも最初は DNA ではなく RNA で、熱に対する安定性から DNA 配列が一般的になったという説が有力であるが、本文でもふれるように DNA の高次構造は大変可塑的であり、B-型 DNA の二重らせんの他にも A-型や Z-型などの分子形態を持つことが、生命活動を行ううえで有利であったように思われる。そして、その構造は配列の並びによ

て決まるため、本研究の結果から明らかになるように、配列の最も基本的な一塩基レベルや二連塩基レベル、鎖非対称性やゲノム一貫性ともいえる様々なパターンが生物学的な機能と関係して観察されることは、配列と構造と機能が密接に関連している証拠である。

生命活動における有利性といっても種によって様々であり、雄大な長い年月を要する進化過程において、その時その時の環境や体制そして偶然性によって、普遍的な配列頻度傾向が保存されながらも、現在のような種固有のパターンを持つように至ったのであろう。今後の詳細な解析によって、それらのパターンを生み出す機構が一つ一つ明らかになっていくのかもしれない。そして、それらを全て総合することによって、生命とは何か？あるいは、どのようにして誕生し、どのような進化を遂げていくのか？という問題の理解が深められていくことと思う。

## 論文審査結果の要旨

本学位論文では、ゲノム配列の大量データに対する数量的な特徴づけについて述べられている。現在利用可能な 359 生物種のミトコンドリアゲノム、105 種類の原核生物ゲノム、シロイヌナズナ、線虫、酵母など 6 種の真核生物ゲノム、さらに、ヒト、マウス、ショウジョウバエのドラフト配列などのデータについて大規模な解析が行われ、一塩基と二連塩基組成頻度に関するゲノムレベルの普遍性や特異性がまとめられている。論文のポイントは、(1) DNA の複製エラーに鎖依存性がなければ、ワトソン・クリックの規則によって各鎖における A と T 及び G と C の頻度は等しくなること、及び (2) 二連塩基組成と DNA の公示構造との間に、エネルギー的な側面で相関関係があることに基づいている。

突然変異の鎖非対称性による塩基配列の偏り (skew) のパターンが配列全体において調べられ、さらに二連塩基頻度も調べられた。すべてのゲノムで塩基組成に偏りがあることが強く示唆された。こうした研究はすでに行われているが、これほど大規模に各種のゲノムを網羅し普遍性と特異性を考察したのは本論文がはじめてである。また二連塩基組成の偏りとして CpG や TpA が知られているが、低頻度 CpG は古細菌やある種の真核生物に特異な現象であるのに対して、低頻度 TpA は全ての生物で普遍的であることが示されている。本論文は、この観察事実を DNA の高次構造の融通性と関連してエネルギー的な側面から議論している。

鎖非対称性は DNA 複製システムと関係するものであり、一方向に複製が進む動物のミトコンドリアでは単調なパターンをもつが、複製が二方向性である原核生物では動物複製開始点・終始点で skew が反転する。真核生物の核でも、原核生物と同じく複製が二方向性であるが、一本の染色体上に複製単位がたくさんあるため、skew のパターンが複雑である。それにもかかわらず、それぞれのゲノム全体に一貫した種固有の二連塩基頻度の傾向があることが示された。配列の最も基本的な一塩基レベルや二連塩基レベルで、鎖非対称性やゲノム一貫性ともいえるさまざまなパターンが生物学的な機能と関連していることが示され、ゲノム配列と構造・機能との関連性が考察されている。

本論文では、限られた生物種のゲノム解析ではなく、生物界全体で入手可能なあらゆるゲノムデータについて、広範な解析が行われており、興味深い知見が得られた。真核生物ゲノムにおける一塩基レベルの鎖非対称性にもかかわらず、それぞれのゲノム全体に一貫した種固有の二連塩基頻度があるという知見は、新しい発見であり、これに関連した今後の研究の更なる展開が期待される。

以上の評価より、本論文の内容は博士(理学)に十分に値するものであると判定した。なお、本学位申請論文の内容に関する 1 篇の原著論文(申請者は筆頭著者)が、国際学術誌である Mol.Biol.Evol.に既に掲載されている。