

令和元年6月19日現在

機関番号：82407

研究種目：若手研究(B)

研究期間：2017～2018

課題番号：17K18372

研究課題名(和文)化石・考古学遺物用バイオマーカーとしての型コラーゲン配列情報の総合的解析

研究課題名(英文)The comprehensive analysis of type I collagen sequence as a biomarker for distinguishing fossils and archaeological samples.

研究代表者

大波 純一(Onami, Jun-ichi)

国立研究開発法人科学技術振興機構・バイオサイエンスデータベースセンター・研究員

研究者番号：10726623

交付決定額(研究期間全体)：(直接経費) 1,600,000円

研究成果の概要(和文)：コラーゲンファミリーは細胞外マトリクスを構成するタンパク質の一群であり、多くの生物が有している。その中でも型コラーゲンは特に体組織を構成する主要なコラーゲンタンパク質として知られている。

本研究ではこの型コラーゲンをコードする遺伝子配列を、分子生物学的手法と情報学的手法の両方を使ってゲノム内から集中的に探索した。この結果、複数の種から新しく型コラーゲンの相同領域を同定することができた。またこれらのコラーゲン配列をデータベースにまとめ公開した。

研究成果の学術的意義や社会的意義

本研究結果は、重要だと予想されていたながらその解析の困難さのためにあまり着手されていないコラーゲンタンパク質を、集中的に解析しその情報を公開したことに意義を持つ。これらの公開されたコラーゲン配列情報は生物の生体だけでなく、今後増えていくと考えられる化石由来のタンパク質や動物加工製品の迅速な同定・分類に役立つだけでなく、国際的な生物動態や環境保全のための情報として幅広く利用されることが見込まれる。

研究成果の概要(英文)：The collagen family forms a part of the extracellular matrix protein group. A lot of organisms have this protein family. The collagen type I family is known as the major fibrillar collagen constitutes body tissue.

In this project, undiscovered coding regions of collagen type I sequence were searched from animal genomes. With the method of molecular biological technique and bioinformatic approach, some collagen type I sequences were newly identified. A database website was developed and these sequences have been published from the server.

研究分野：ゲノム生物学

キーワード：コラーゲン リファレンスデータベース バイオマーカー

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

- (1) コラーゲンファミリーは細胞外マトリクスを構成するタンパク質の一群であり、多くの生物が有する。その中でもⅠ型コラーゲンは特に体組織を構成する主要なコラーゲンタンパク質として知られている。さらにタンパク質の中でも動物体内で大量に合成されて繊維状に蓄積し遺物から検出し易いこと、系統特異的な変異が検出でき、DNA よりも古い年代の試料から検出できることから、化石標本や考古学資料など多様な条件のサンプルから検出することのできる「優秀なバイオマーカー」として普及しつつある(表1)。このことからコラーゲンの配列情報は、今後さらに幅広く利用されていくと考えられていた。

表1. 古代バイオマーカーとしてのコラーゲンと DNA の特性

	十分な検証が行われた試料の最古の年代	より検証が必要な試料の最古の年代	保存環境の影響 (比較値)	DB 登録数 (比較値)
古代タンパク質 (コラーゲン)	150 万年前	8000 万年前	小	少
古代 DNA (*A : 核、*B : 葉緑体)	70 万年前 (*A)	2000 万年前 (*B)	大	多

- (2) 近年、化石残存タンパク質を利用した系統推定方法が確立され、これまで古代 DNA が抽出できず分子生物学的な解析が困難であった熱帯地域の絶滅哺乳類の分類が、化石に残存するⅠ型コラーゲンのアミノ酸構造の比較を元におこなわれた。また、ウシやシカのコラーゲンは古くから膠(にかわ)として利用され、しばしば日本の考古学においては発掘遺物の周辺情報解析の指標として利用されている。この一方でいくつかの研究では、コラーゲン配列を元にした系統樹が、哺乳類の獣類レベルや急速な分化が起きたと考えられる系統関係を正確に反映していない。また、2015年8月に申請者の研究により新しくシカの一様のコラーゲン配列が同定されたものの、このアミノ酸配列を指標とする偶蹄類内部の科間の系統関係は、現在正しいと考えられている系統樹と違ったものとなった。つまり、コラーゲン配列によって正しい解像度が得られる系統関係の範囲は「属」や「種」内レベルであり、コラーゲンを正確な系統同定のためのバイオマーカーとして利用するには、あらかじめ幅広い「属」や「種」の比較のためのリファレンス配列を網羅して比較対象として保持する必要がある。

- (3) しかし現在、データベースに登録されているコラーゲンのアミノ酸配列は登録種数が少なく、米国の遺伝子データベースである NCBI Gene においても、動物の分類群に対し「Collagen」で検索すると 28000 件ヒットする(2016/10/20 時点)が、殆どがヒトやマウス、ウシなど限られたモデル生物や家畜化動物のコラーゲン情報であり、シカ属や奇蹄類など広く知られた高等動物であってもヒットしない。この理由としてコラーゲンの遺伝子が他の一般的な遺伝子に比べ長く(14kb~)、介在するエキソンとイントロンの数が多く(エキソン数 50 以上~)、各エキソン・イントロンのモジュールが小さく(50bp 程度~)、随所に反復が見られることから、他の遺伝子に比べ配列解析や自動的なアノテーションが難しい為と考えられる。

2. 研究の目的

そこで本研究ではこのⅠ型コラーゲンの鎖の配列情報を、各種生物サンプルを使った実験とデータベース情報からのスクリーニングで収集し、分子生物学者だけでなく化石研究者や考古学研究者が簡単にアクセスし手元のサンプルと比較できる「バイオマーカーリファレンスデータベース」として公開することとした。

3. 研究の方法

本研究ではコラーゲン遺伝子の構造特異性のためにアノテーションがされにくい問題に対応するため、生物情報学的手法と、生体サンプルからの配列解読により、高等動物の中の幅広い種におけるコラーゲン配列情報を収集し、利便性の高いデータベースとして公開する。

4. 研究成果

(1) 保存領域の検討

高等動物の既知のコラーゲンコード遺伝子の配列比較を実施し、保存されているエキソン上の領域を選択してオリゴヌクレオチドプライマーをハイブリダイゼーションのために調製した。調製した 4 種類のプライマー(プライマー名: 5'末端~3'末端方向の配列、COLF1:

GATGATGTAAGTAGCCCTGG、COLR1 : GAGTTTACAGGAAGCAGACA、COLF2 : GGTCTAGACATGTTTCAGCTT、COLR2 : GGGGAGTTCAACTACTCTG) を組合せて実験に使用することとした。コラーゲンの内部における逆位や反復領域が存在しない部分に、プライマーがダイマーやヘアピン構造を取らないよう dot plot 図等を利用し慎重にプライマーを設計した(図1)。

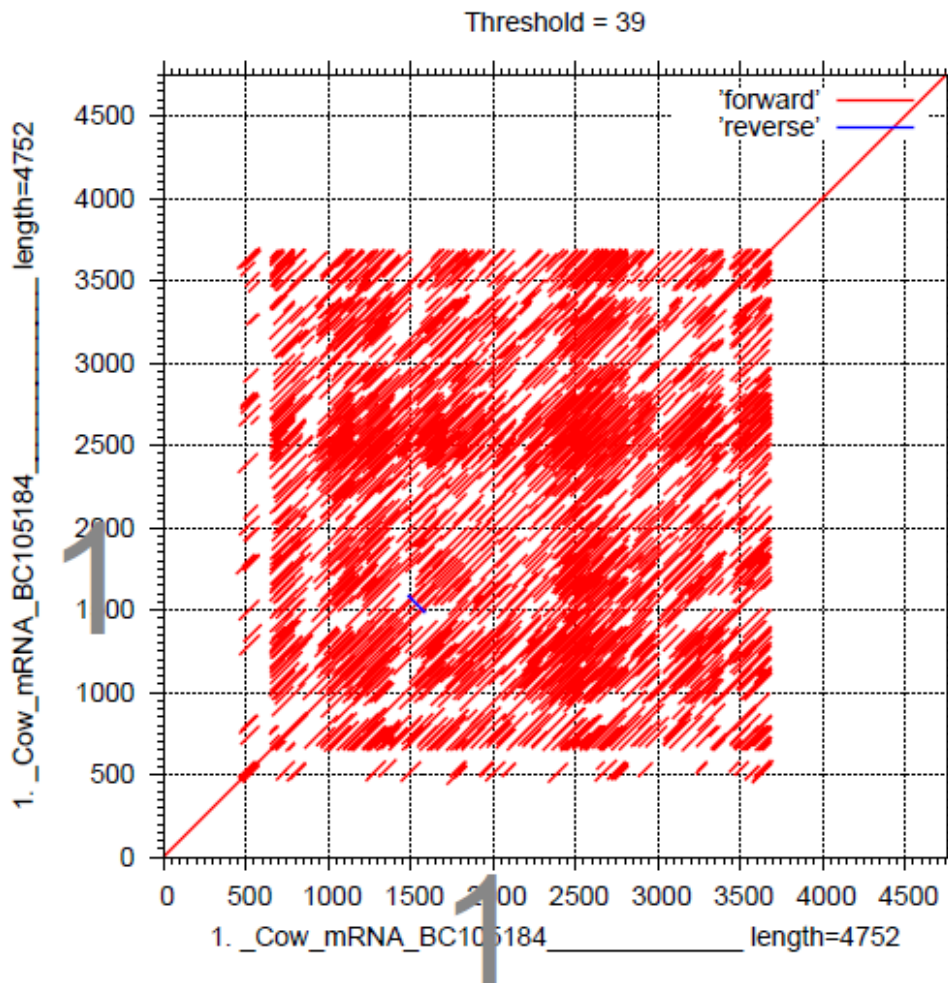


図1 ウシの 型コラーゲンの塩基配列同士を比較した dot plot 図

(2) コラーゲン配列の分子生物学的手法を利用した単離

国内で飼育・採取されたウシ、ニホンジカ、イノシシのゲノムをテンプレートとして、作成したプライマーで PCR を様々な条件で実施したが、十分な増幅産物を得ることができなかった。ゲノムサンプルの状態や抽出量も十分であったが、コラーゲン特有の配列構造のために、予想されていたものの、特異的なアニーリングは困難であった。

(3) コラーゲン配列の情報学的手法を利用した単離

比較的最近全ゲノム情報が公開された 5 種の高等動物、*Antilocapra americana*、*Moschus moschiferus*、*Tragulus javanicus*、*Odocoileus hemionus*、*Rangifer tarandus* のデータを取得し、既知のコラーゲン配列をクエリとして相同領域を抽出した。これらのコラーゲン相同領域は先行研究では報告されることがない情報だった。

(4) データベース構築、データ解析

最終的に 5 種から新しく 型コラーゲンの相同領域を同定することができ、合計 24 種の高等動物のサンプルから情報を収集することができた。得られたコラーゲン配列同士で比較解析をおこなった所、図2のように系統樹の枝長が属内の種によっても大きく異なり、進化速度が系統によって大きく異なることが示された。また図2では明らかに種の系統樹とは異なった系統関係を示しており、新しい知見を得た。

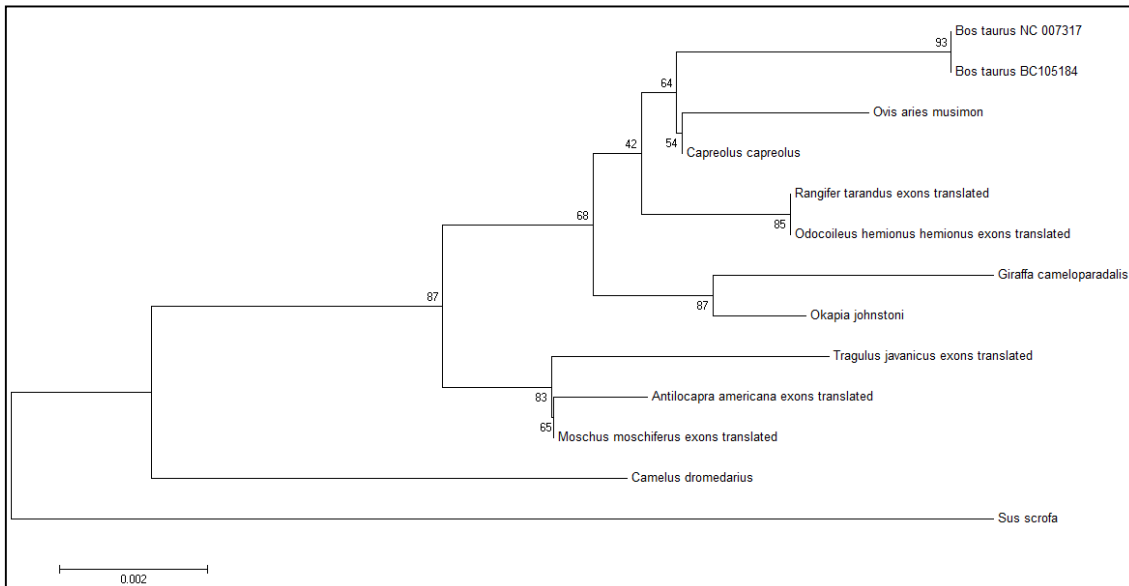


図2 近隣結合法で構築した偶蹄目内部の種Ⅰ型コラーゲンの塩基配列の系統樹

これらの配列をデータベースに集約し、バイオサイエンスデータベースセンターの基盤から「Collagen reference database」として公開をおこなった。この変異情報を利用して古代タンパク質や動物性遺物の種同定のためのリファレンスとされることを期待し、今後もデータベースの拡充を実施していく予定である。

5. 主な発表論文等

〔雑誌論文〕(計0件)

〔学会発表〕(計1件)

大波純一、村上勝彦、山崎朗子、高木利久, "高等動物におけるⅠ型コラーゲンの配列比較解析", 日本進化学会第21回大会, 2019年

〔図書〕(計0件)

〔産業財産権〕

出願状況(計0件)

取得状況(計0件)

〔その他〕

ホームページ等

Collagen reference database <https://togodb.biosciencedbc.jp/db/collagenseq>

6. 研究組織

(1)研究分担者

なし

(2)研究協力者

研究協力者 氏名: 山崎 朗子

ローマ字 氏名:(YAMAZAKI, akiko)

研究協力者 氏名: 村上 勝彦

ローマ字 氏名:(MURAKAMI, katsuhiko)

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。