

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 8 日現在

機関番号：12601

研究種目：基盤研究(B) (一般)

研究期間：2013～2015

課題番号：25291104

研究課題名(和文) 古人骨DNA研究から古人骨ゲノム研究へ：次世代シーケンサによる新展開

研究課題名(英文) From DNA to genome: New development on human evolution study by next-generation sequencer

研究代表者

植田 信太郎 (Ueda, Shintaroh)

東京大学・理学(系)研究科(研究院)・教授

研究者番号：20143357

交付決定額(研究期間全体)：(直接経費) 11,500,000円

研究成果の概要(和文)：熱帯～温帯地域で出土した古人骨から得られるDNAの99%以上は土壌菌由来であり、次世代シーケンサで得られるヒト由来のDNA配列はごく僅かである。そこで、目的とするDNA領域を特異的に濃縮するターゲット・エンリッチメント、ならびに、シーケンス・ライブラリー作成での油滴乳濁液中DNA増幅により、様々な時代の古代遺跡から出土した古人骨のミトコンドリアゲノムの全塩基配列を、最小限のシーケンス・ランで高い信頼度をもって得ることに成功した。

研究成果の概要(英文)：Due to the difficulties in deep sequencing, high-throughput sequencing of ancient DNA has been limited to exceptionally well-preserved ancient materials. The primary factor is microbial attack popularly observed in the buried materials, and it causes drastic increase in relative ratio of microbial DNA in the extracted DNA. I developed a unified strategy in which emulsion PCR is coupled with target enrichment followed by next-generation sequencing. The method made it possible to obtain efficiently non-duplicated reads mapped to target sequences of interest, and this can achieve deep and reliable sequencing of ancient DNA from typical materials, even though poorly preserved. Then, I applied this method to ancient human remains from archaeological sites of various times, and succeeded in obtaining highly reliable and entire nucleotide sequences of mitochondrial genome.

研究分野：自然人類学、分子人類学

キーワード：人類進化 古人骨 ゲノム ミトコンドリア 次世代シーケンサ

1. 研究開始当初の背景

遺跡から出土したヒトや動植物の遺骸に残されたDNAを取り出し、それらの塩基配列を決定し、得られた遺伝情報をもちいて、文献記録に残されていない時代の有り様を明らかにしようとする挑戦が始まってから四半世紀になるうとしている。これら古DNA分析によって様々な新知見が明らかにされてきた。しかし、発掘され保存されている動植物遺骸の数に比べて、分析された試料数はあまりにも少ない。

最大のネックは、古DNA分析に必要とされる専門性の高さや時間を含めた労力である。この問題を一気に解決する方法として期待されるのが次世代シーケンサである。次世代シーケンサはDNA断片の集合体を一気に読み取るハイスループットな塩基配列決定のための機器で、この10年ほどで開発が進み、現生生物の全ゲノム配列決定や既知の生物ゲノムの再シーケンスに利用されている。そして、古DNA分析への応用が開始された。しかし、次世代シーケンサによる古DNA分析が成功したのは非常に稀な試料(毛髪や爪といったケラチン質の試料、永久凍土などの非常に良好な環境で保存されてきた骨試料)に限られている。熱帯～温帯地域からの出土古人骨から得られるDNAでは、ヒト由来のDNAは高々その0.1%ほどであり、ほぼ全ては骨へ侵入した土壌菌由来のDNAである。このため、次世代シーケンサで得られるヒト由来のDNAの塩基配列情報はごく僅かである。そこで、解決策としてターゲット・エンリッチメントと呼ばれる「ヒト由来のDNAを特異的に濃縮するための方法」がいくつか提唱されている。しかし、いずれの方法も極めて特殊な方法であるばかりでなく、得られるヒト由来のDNAの塩基配列情報に重複が多い。信頼のある古人骨DNA配列情報を構築するためには重複しない配列情報を多く得る必要があった。

2. 研究の目的

本研究では、研究代表者自身が開発した次世代シーケンサによる古代ゲノム分析法(“普通の研究室”で“普通の古人骨”から得られる“普通の古人骨DNA、すなわち、古人骨から抽出されるDNAの99.9%以上が土壌菌由来である「保存の良くないDNA」”をもちいても次世代シーケンサによる塩基配列決定が可能な実験手法プロトタイプ)を用い、これまで深く関わってきた中国・日本・メキシコの様々な時代の遺跡から出土した古人骨を分析する。その過程で、(1)数千年前の古代の人々の遺伝的多様性をミトコンドリアゲノム全塩基配列から明らかにする、(2)実験手法をさらに普遍的かつ廉価な手法へと改良する。これにより、古人骨DNA研究を古人骨ゲノム研究へ導くことを目的とする。

3. 研究の方法

塩基配列を決定したい“目的とするDNA”(本研究ではヒトのミトコンドリアゲノム全塩基配列)を特異的に濃縮するための方法であるターゲット・エンリッチメントで用いるベイトをバイオチン化RNAとし、固相ではなく液相でのハイブリダイゼーションによりフィッシングをおこなう、シーケンス・ライブラリー作成にあたって油滴乳濁液中でDNAの増幅(エマルジョンPCR)をおこなうことにより、ヒト・ミトコンドリアゲノムにマッピングされるシーケンス・リードを重複少なく得る。これにより、最小限のシーケンス・ランでシーケンス密度が高い、すなわち、高い信頼度をもった古人骨のミトコンドリアゲノムの全塩基配列の決定をおこなった。

4. 研究成果

本研究課題により、研究代表者が以前に開発した「抽出されたDNAの99.9%以上が土壌菌由来である“保存の良くないDNA”からでも、ミトコンドリアゲノム全塩基配列の決定が可能となる次世代シーケンサによる古代ゲノム分析プロトタイプ」を改良し、その手法を確立した。この手法を用い、中国・山東省の2000年前の遺跡から出土した古人骨の分析を行い、複数の古人骨で、それらのミトコンドリアゲノム全塩基配列を決定した(全長のカバレッジ率100%)。決定したミトコンドリアゲノム全塩基配列の平均デプス(同じ塩基サイトを独立に、すなわち、重複のない配列として読んだ回数)は50以上であり、極めて信頼度の高いミトコンドリアゲノム全塩基配列を決定することに成功している。次に、中国・山東省の2500年前の遺跡から出土した古人骨の分析を行い、複数の古人骨のミトコンドリアゲノム塩基配列を決定した。決定したミトコンドリアゲノム全塩基配列の平均デプスならびに全長のカバレッジ率は、山東省の2000年前の遺跡から出土した古人骨試料と比べて低かったが、複数個体に関してほぼ全長に近いミトコンドリア塩基配列情報を、高いデプスデータとして得ることに成功した。さらに、中国・河南省の3000年前の遺跡から出土した古人骨でも、それらのミトコンドリアゲノム全塩基配列を決定することに成功した。同様に、日本ならびにメキシコの様々な時代の遺跡から出土した古人骨のミトコンドリアゲノム塩基配列を決定した。地球上の多様な現代人類集団に属する数万以上の現代人に関して、そのミトコンドリアゲノムの全塩基配列に基づく精細なハプログループがヒトミトコンドリアDNAの多様性の指標として定義されている。これにより、ハプログループを基盤としたミトコンドリアゲノム全塩基配列の個々の塩基サイトの情報を重視した「ハプログループのサブグループの、さらに細かいサブタイプレベル」

での分析が可能となっている。そこで、本研究によって得られた古人骨のミトコンドリアゲノム全塩基配列の情報を用いて、それぞれのハプログループの、サブグループの、さらに細かいサブタイプに、さらに個体ごとの変異塩基情報を加えたミトコンドリア系統ネットワークを作成し、上記の各古人骨の遺伝的多様性情報を明らかにした。

ところで、古人骨DNA分析では、発掘者や実験者のDNAが混入すると、得られた結果への影響は甚大である。そのため、データの信頼性を検証するには、コンタミネーションの割合を推定する方法が必要である。そこで、コンタミネーションの影響について考察するためのシミュレーション・データを作成し、それを本研究によって得られた古人骨ミトコンドリアゲノム全塩基配列に適用した。検証の結果、約16000塩基中、最大で1塩基のコンタミネーションであると推定され、得られた配列データが高い信頼性をもつことが示された。

以上の解析に加え、次世代シーケンサから出力されるリードデータ(生データ)の分析に不可欠なバイオフィォマティックス処理を容易にする解析ツールの開発をおこなった。古人骨ゲノム解析では、長期に及ぶ時間経過の中で受けた様々な塩基修飾やコンタミネーションの割合の推定(データの信頼性の検証)など、現代人ゲノムの解析に加えて、古人骨ゲノムに固有の問題に対応した解析が不可欠である。そこで、現代人・古人骨共通の「ミトコンドリアゲノムにマッピングされた次世代シーケンサ出力生データに関するプロファイリングおよび解析」、そして、古人骨ゲノムに固有の「プロファイリングおよび解析」をおこなうためのツールを開発した。加えて、解析は煩雑なコマンド操作で実行するのではなく、マウス操作での実行を可能にするGUI解析ツールとして開発した。これにより、次世代シーケンサから出力される膨大なリードデータ(塩基配列情報)を視覚的に解析することを可能とした。

5. 主な発表論文等

〔雑誌論文〕(計 4件:すべて査読あり)

1. MtDNA diversity of the Zapotec in Mexico suggests a population decline long before the first contact with Europeans. Gojobori J., Mizuno F., Wang L., Onishi K., Granados J., Gomez-Trejo C., Acuna-Alonzo V., and Ueda S. *Journal of Human Genetics* 60(9), 557-559 (2015) doi:10.1038/jhg.2015.55 (査読あり)
2. Complete mitogenome analysis of indigenous populations in Mexico: Its relevance for the origin of Mesoamericans. Mizuno F., Gojobori J., Wang L., Ohnishi K., Sugiyama S., Granados J., Gomez-Trejo C., Acuna-Alonzo V., and Ueda S. *Journal of*

Human Genetics 59(7), 359-367 (2014) doi:10.1038/jhg.2014.35 (査読あり)

3. Mammalian-specific sequences in Pou3f2 contribute to maternal behavior. Nasu M., Yada S., Igarashi A., Suto D., Akiyama K., Ito M., Yoshida N., and Ueda S. *Genome Biology and Evolution* 6(5), 1145-1156 (2014) doi:10.1093/gbe/evu072 (査読あり)

4. Emulsion PCR-coupled target enrichment: an effective fishing method for high-throughput sequencing of poorly preserved ancient DNA. Kihana M., Mizuno F., Sawafuji R., Wang L., and Ueda S. *Gene* 528(2), 347-351 (2013) doi:10.1016/j.gene.2013.07.040 (査読あり)

〔学会発表〕(計15件)

1. 水野文月・熊谷真彦・黒崎久仁彦・王瀝・林美千子・杉山三郎・植田信太郎、低カバレジ NGS データから集団レベルの解析を可能にする配列構築のアプローチ、第69回日本人類学会大会、2015年10月10日~12日、産業技術総合研究所臨海副都心センター(東京都江東区)
2. 水野文月・石谷孔司・熊谷真彦・王瀝・黒崎久仁彦・植田信太郎、次世代シーケンサを用いた古代ゲノム - ミトゲノム解析の展望、第69回日本人類学会大会、2015年10月10日、産業技術総合研究所臨海副都心センター(東京都江東区)
3. 水野文月・王瀝・黒崎久仁彦・大橋順・中伊津美・植田信太郎、次世代シーケンサを用いた古代ゲノム - 核ゲノム解析の展望、第69回日本人類学会大会、2015年10月10日、産業技術総合研究所臨海副都心センター(東京都江東区)
4. 水野文月・王瀝・黒崎久仁彦・澤藤りかい・熊谷真彦・杉山三郎・植田信太郎、低カバレジ NGS データから集団レベルの解析を可能にする配列構築のアプローチ、NGS現場の会・第4回研究会、2015年7月1日~3日、つくば国際会議場(茨城県つくば市)
5. 澤藤りかい・吉田建朗・植田信太郎、ヒトの古代DNA データにおけるDNA混入の評価方法、NGS現場の会・第4回研究会、2015年7月1日~3日、つくば国際会議場(茨城県つくば市)
6. Ishiya K. and Ueda S., HuMAP: A comprehensive tool for assessing NGS data on human mitogenome, The 11th International Workshop on Advanced Genomics, 2015年5月20日~22日、National Center of Sciences (Tokyo, Japan)
7. 水野文月・澤藤りかい・木花牧雄・黒崎久仁彦・王瀝・植田信太郎、殷墟から出土した古人骨の次世代シーケンサによるゲノム解析、日本DNA多型学会 第23回学術集会、2014年11月27日、愛知県産業労働センター・ウインクあいち(愛知県名古屋)

8. 澤藤りかい・植田信太郎、動物遺存体における古代プロテオミクス解析の現状と展望、第68回日本人類学会大会、2014年11月3日、アクトシティ浜松・コンgresセンター（静岡県浜松市）

(2)研究分担者
なし

9. 石谷孔司・植田信太郎、古代ゲノムにおけるコンタミネーション推定法の開発とその応用、第68回日本人類学会大会、2014年11月1日、アクトシティ浜松・コンgresセンター（静岡県浜松市）

(3)連携研究者
なし

10. 澤藤りかい・吉田建朗・植田信太郎、古代DNA解析における現代DNAのコンタミネーション推定法、第68回日本人類学会大会、2014年11月1日、アクトシティ浜松・コンgresセンター（静岡県浜松市）

11. 水野文月・植田信太郎、次世代シーケンサによる古人骨ゲノム解析：コンセンサス配列の構築、第68回日本人類学会大会、2014年11月1日、アクトシティ浜松・コンgresセンター（静岡県浜松市）

12. 水野文月・澤藤りかい・木花牧雄・王瀝・植田信太郎、古代中国3000年前の殷墟から出土した古人骨ゲノム解析、第68回日本人類学会大会、2014年11月1日、アクトシティ浜松・コンgresセンター（静岡県浜松市）

13. 水野文月・王瀝・澤藤りかい・杉山三郎・植田信太郎、テオティワカン・月のピラミッドから出土した古人骨ミトコンドリアゲノム解析、第67回日本人類学会大会、2013年11月1日～2013年11月4日、国立科学博物館筑波研究施設（茨城県つくば市）

14. 澤藤りかい・吉田建朗・梅山翔平・植田信太郎、次世代シーケンサでの古代DNA解析におけるコンタミネーションの影響について、第67回日本人類学会大会、2013年11月1日～2013年11月4日、国立科学博物館筑波研究施設（茨城県つくば市）

15. 水野文月・澤藤りかい・木花牧雄・王瀝・植田信太郎、次世代シーケンサによる古人骨ゲノム解析：現在の問題点、第67回日本人類学会大会、2013年11月1日～2013年11月4日、国立科学博物館筑波研究施設（茨城県つくば市）

〔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

<http://www.biol.s.u-tokyo.ac.jp/users/shinkalab.html>

6. 研究組織

(1)研究代表者

植田 信太郎 (UEDA SHINTAROH)

東京大学・大学院理学系研究科・教授

研究者番号：20143357