

## Press Release

文部科学記者会・科学記者会  
同時発表

2022年11月22日  
横浜市立大学

### 言語は遺伝子と共に進化するのか？ ダーウィン「種の起源」以来の論争にゲノム DNA データで挑む

横浜市立大学木原生物学研究所 清水健太郎客員教授（チューリッヒ大学教授兼任）およびキアラ・バルビエリ博士（チューリッヒ大学・マックスプランク進化人類学研究所）らの研究グループは、「言語進化は遺伝子進化と同調するのか、しないのか？」という 1859 年のダーウィン「種の起源」以来の論争について、同調することが多いものの、同調しない場合も 20%程度みられることを示しました。遺伝子情報は親から子にのみ伝わりますが、言語などの文化はより複雑な形で継承されます。言語進化と遺伝子進化の関連性を調べるために、295 言語を話す 4 千人以上のゲノム DNA 多型情報と言語情報を含んだ大規模データベースを作成し、言語情報とゲノム DNA に含まれる遺伝子情報の比較解析を行いました。その結果、過半数の言語は遺伝子の進化に伴って分岐進化していた一方、ハンガリー語を含む約 20%の言語で言語と遺伝子の進化が一致していないことが分かりました。このような“言語転換”は人類の歴史上例外的ではなく普遍的な現象であることが分かりました（図 1）。本研究は、ゲノム DNA 情報が生命科学分野に限らず人間の文化史の研究にも欠かせないことを示しています。

本研究成果は、アメリカ国際誌「米国科学アカデミー紀要」に掲載されました。（2022 年 11 月 21 日オンライン）

#### 研究成果のポイント

- 295 言語の話者 4,030 人の約 60 万 DNA 多型情報データベース GeLaTo を構築
- インド・ヨーロッパ語族などでは言語と遺伝子が同調して分岐したことが多かったが、言語転換も世界各地で見られた
- ゲノム DNA 情報解析は、ヒトの言語や文化の進化と多様性の解明にも寄与する



# Press Release

(図1) 9世紀末に少数の移住者によってハンガリー語がカルパチア盆地にもたらされ、地域住民の言語転換が起きたと考えられている。絵画「ハンガリー人のカルパチア盆地征服」(製作 1892-1894年、作者フェスティ・アールバード。en.wikipedia.org)

## 研究背景

人間の遺伝子は親から子に伝わります。これを垂直伝達と言います。一方、言語や文化は、垂直伝達に加えて水平伝達も見られます。日本語を話す親から生まれて育てられた子供は垂直伝達によって日本語を話すようになることが多いかも知れませんが、学校で習ったり外国に移住したりして、水平伝達によって別の言語を話すこともあるのです。

進化学の祖チャールズ・ダーウィンは、その代表的著書「種の起源」(1859年)の中で、「ヒトの遺伝的な系統関係が完全に分かれば、現在世界で話されている様々な言語の最もよい分類ができるだろう」という仮説を述べました。この大胆な仮説は、言語進化には垂直伝達が主要な役割を果たすことを前提としています。1980年代から遺伝学者によってこの説を支持するデータが出されてきました。我々も2021年に日本語を含む北東アジア地域で遺伝進化と言語進化が相関していることを示しました ([https://www.yokohama-cu.ac.jp/news/2021/202108shimizu\\_SA.html](https://www.yokohama-cu.ac.jp/news/2021/202108shimizu_SA.html))。

一方で、ダーウィンの仮説に合わない例も多く指摘されてきました。中世にごく少数のハンガリー語話者が東ヨーロッパに移住した後に、多くの周辺住民がハンガリー語を話すようになったことはよく知られた例です(図1)。このハンガリー語における言語転換、つまり言語集団がその使用言語を変えることは、水平伝達による大きな言語変化の一例です。このような例外の存在から、ダーウィンの仮説が正しいのか正しくないのか、長年論争が続きました。生物学で長く論争が続く場合、両者ともあるケースでは正しく一部では正しくないと最終的に結論付けられることも往々にして見られます。そのため、この論争に決着をつけるには、個々の例を議論するだけでなく、大規模データベースによる世界規模での解析が求められていました。

## 研究内容

本研究では、295言語を話す397人類集団の計4,030人のゲノムDNA多型情報\*1と言語情報を集めたデータベース GeLaTo を作成しました。集団間の遺伝的距離( $F_{ST}$ と呼ばれる統計量)を、ゲノム中の約60万箇所のDNA配列多型情報に基づいて計算しました。一方で、言語間の類縁関係は、「同じ語族\*2に属するかどうか」で判定しました。例えば、ハンガリー語はウラル語族に属しますが、ハンガリー語話者と遺伝的に最も近いのは、ウラル語族の他の言語話者ではなく、インド・ヨーロッパ語族\*3の話者でした。このようなケースを「飛び地」(enclave)と名付けました(図2)。こうして調べていくと、全言語集団の約20%もが飛び地のパターンを示しました。これは、言語進化と遺伝的進化が同調しないケースです。一方、ダーウィンの仮説と合致して遺伝的に近い集団が同じ語族の言語を話すケースはより多く見られました。これらの結果をまとめると、ダーウィンの仮説のように言語進化と

## Press Release

遺伝的進化は同調することが多いが、しない場合も頻繁にあることがわかりました。

さらに、それぞれの言語族が、遺伝的に近縁の集団だけで話されているのか遠縁の集団でも話されているのかについて、地理的な距離も考慮に入れて解析しました。その結果、インド・ヨーロッパ語族（英語、ドイツ語、ペルシャ語などを含む）やシナ・チベット語族（中国語やブルマ語を含む）などは、遺伝的に近い集団で話されている傾向がありました。一方、ウラル語族（ハンガリー語やフィンランド語）やトルコ語属（トルコ語やウイグル語）などは遺伝的に遠い集団でも話されていました。

最後に、言語と遺伝的關係それぞれについて分岐年代の推定を行いました。集団の分岐時間は、遺伝的距離  $F_{ST}$  から推定可能です。一方、言語の分岐年代は言語データの系統解析から推定されます。両者の分岐年代が対応していれば、移住などによって集団が分かれながら言語もそれぞれに進化していったという垂直伝達の歴史が支持されます。インド・ヨーロッパ語族の起源は 5,500-8,000 年前程度と推定されており、遺伝的な分岐年代と言語分岐年代に相関がみられました。しかし、トルコ語族などでは相関が弱く、言語転換が頻繁に起こったと考えられました。これまでの言語学的な研究ではインド・ヨーロッパ語族が最もよく研究されていたため、チャールズ・ダーウィンの仮説もインド・ヨーロッパ語族を念頭に置いていた可能性があります。しかし、インド・ヨーロッパ語族は世界の言語の一部であり、今回、網羅的に語族を解析することで、言語転換は例外的でなく世界中で起きていることを示すことができました。

### 今後の展開

本研究では、GeLaTo データベースのうち言語数の多い語族を対象として解析を行いました。日本語族（日本語と琉球語）のような言語数の少ない語族や、アイヌ語のような孤立言語の進化は今後の課題であり、さらに多くのデータを集積することが必要です。しかし、グローバル化などの影響で各地の伝統文化が失われたり言語が使われなくなる例が多発しており、世界に知られる 7,000 以上の言語のうち相当部分が、話者の減少による絶滅の危機に瀕しています。また、遺伝的データが得られているのは医学の発達した先進国に偏っています。人類の文化の多様性を理解し維持していくには、話者の少ない言語集団も含めて世界各地で文化を保護・継承し、研究をすすめていくことが欠かせません。

本研究では、遺伝学、言語学、データ科学など学際的な協力が重要な役割を果たしました。本研究でデータベース構築を主導したドイツ・マックスプランク進化人類学研究所\*4をはじめとして、ゲノム DNA データを用いたヒトの文化の多様性の研究が各地で急速に進められています。農耕・言語拡散仮説によれば、農耕の拡散とともに言語も拡散し、遺伝的分岐とも同調しつつ言語が分岐し語族が形成されたと考えられています。この重要な仮説の検証のためには、人類学、植物学、言語学、考古学、ゲノム科学などを統合したさらなる学際的研究が必要です。

# Press Release

## 研究費

本研究は、学術変革領域研究 (A) 「中国文明起源解明の新・考古学イニシアティブ」、チューリッヒ大学学内重点領域「進行中の進化」などの支援を受けて実施されました。

## 論文情報

タイトル : A global analysis of matches and mismatches between human genetic and linguistic histories

著者 : Chiara Barbieri, Damián E. Blasi, Epifanía Arango-Isaza, Alexandros G. Sotiropoulos, Harald Hammarström, Søren Wichmann, Simon J. Greenhill, Russell D. Gray, Robert Forkel, Balthasar Bickel, Kentaro K. Shimizu

掲載雑誌 : 米国科学アカデミー紀要 Proceedings of the National Academy of Sciences, U.S.A. (PNAS)

DOI : 10.1073/pnas.2122084119



## 参考

### 用語説明

\*1 DNA 多型 : DNA 配列の個体差。DNA は ATGC の 4 種類の塩基からなっており、ヒトゲノムは約 30 億塩基ある。このうちで、個体差がよく見られることが分かっている 60 万カ所について本研究ではデータを取得した。

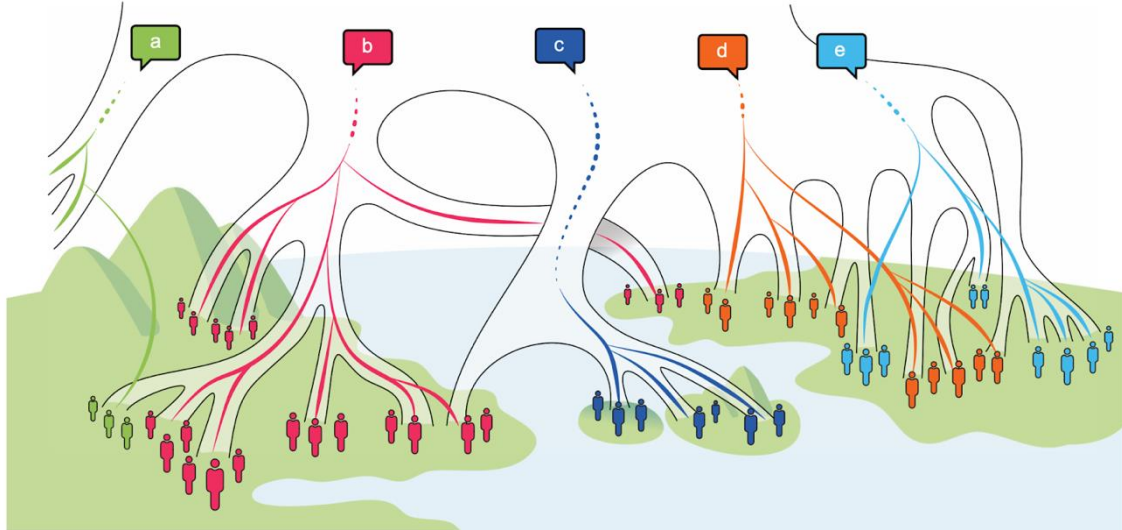
\*2 語族 : 共通の祖先の言語から分岐してきたと考えられる言語のグループ。Glottolog データベースによれば、世界の 7638 の言語は 245 の語族と、語族に分類されない 182 の孤立言語からなる。

\*3 インド・ヨーロッパ語族 : 英語・スペイン語などヨーロッパで話されている言語群とペルシア語・ヒンディー語などインドを含むアジアで話されている言語群は共通の祖先を持つ言語であると知られており、一つの語族としてインド・ヨーロッパ語族と呼ばれる。言語の歴史が最もよく研究されている語族である。

\*4 マックス・プランク進化人類学研究所 : 1997 年にドイツ・ライプツィヒに設立され、古代 DNA、言語学、心理学などの学際的研究を行っている。2022 年ノーベル医学・生理学賞を受賞したスヴァンテ・ペーボが所長を務める。

## Press Release

### 図、画像、表



(図 2) 遺伝子と言語が伴って進化する場合と、伴わずに進化する場合の模式図。遺伝的進化を太い黒線で、人類集団を人の形で示している。5つの仮想的な語族(a から e)を5つの色で示した。言語の歴史は集団の遺伝的歴史と一致することもあるが異なることもある。例えば左端では赤の語族(b)の言語を話していた集団が緑の語族(a)の言語への言語転換を経験して「飛び地」となった。Barbieri et al. PNAS より。